

This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (<https://dspace.lboro.ac.uk/>) under the following Creative Commons Licence conditions.



For the full text of this licence, please go to:  
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

**University Library**

Author/Filing Title ..... RAMBL .....

Class Mark ..... T .....

Please note that fines are charged on ALL  
overdue items.

0402941098







# **Factors Affecting Spatial Awareness in Non-Stereo Visual Representations of Virtual, Real and Digital Image Environments**

by

**Dayang Rohaya Awang Rambli**

**A Doctoral Thesis**

**Submitted in partial fulfilment**

**of the requirements for the award of**

**Doctor of Philosophy**


**Department of Computer Science**

**Loughborough University, UK**

**December 2004**

© by Dayang Rohaya Awang Rambli 2004

**Supervisor: Professor Roy S. Kalawsky  
Director of Research: Dr. Chris J. Hinde**

	Loughborough University
Date	SEP'05
Class	T
40294109	

## ABSTRACT

The increasing number of applications employing virtual environment (VE) technologies as a tool, particularly those that use VE as surrogates, makes it important to examine the ability of VE to provide realistic simulations to users. Accurate space and distance perceptions have been suggested as essential preconditions for the reliable use of VE technologies in various applications. However, space and distance perception in the VE has been reported by some investigators as being perceived differently from the real world. Thus, the overall aim of this thesis is to improve our understanding of factors affecting spatial awareness in the VE. The general approach is based on a strategy of conducting empirical investigations comparing tasks performed in the VE to similar tasks performed in the real world. This research has examined the effect of display related factors on users' spatial task performance in the context of static, dynamic and interactive presentations. Three sets of experiments in these respective contexts were conducted to explore the influence of image type, display size, viewing distance, physiological cues, interface device and travel modes on distance estimate and spatial memory tasks. For distance perception, results revealed that the effect of image type depends on the context of presentations, the type of asymmetrical distances and image resolution. The effect of display size in static and dynamic presentations is consistent with the results of previous investigations. However, results from evaluations conducted by the author have indicated that other factors such as viewing distance and physiological cues were also accountable. In interactive presentations, results indicated that display size had different effects on different users whereby familiarity with display size may influence user's performance. Similarly, it was shown that a commonly used interface device is more useful and beneficial for user's spatial memory performance in the VE than the less familiar ones. In terms of travel mode, the natural method of movement available in the real world may not necessarily be better than the unnatural movement which is possible in the VE. The results of investigations reported in this thesis contribute towards knowledge and understanding on factors affecting spatial awareness in the real and VE. In particular, they highlight the influence of these factors in space and distance perception in different contexts of VE presentations which will serve as important scientifically based guidelines for designers and users of VE applications.

## ACKNOWLEDGEMENTS

First and foremost, gratitude goes to Almighty God for the mercy and strength given to me throughout my entire study and completion of work.

I would like to express my very deepest gratitude and sincere appreciation to my supervisor, Professor Roy S. Kalawsky for his expert guidance, support and encouragement throughout my research work. His ideas, suggestions and comments form a valuable part of this thesis. Special thanks to my Director of Research, Dr Chris J. Hinde for his continuous encouragement and motivation.

A big thank you goes to the staff and students of the Department of Computer Science for their support in my experiments. I would like also to thank Mr Patrick Hooligan, Mr Gurbinder Samra, Mr Andre Schappo, Mr Richard Mee and Mr Kip Sahnsi for their technical support, Dr Waad Yousif for teaching and giving me valuable advice and guidance on statistics, and Mr Nick Smith for his advice on statistics and using SPSS software package. Thanks also go to Mrs Judith Poulton for her kind administrative services.

My acknowledgement and appreciation also goes to my sponsor, Universiti Teknologi Petronas of Malaysia for providing me the financial assistance to pursue my study further. Without their support my period of research study would not have been possible.

Very special thanks to my mother for her unreserved love, support, encouragement and prayers. Thanks also to my sisters - Aisyah, Ani, Rima and brothers - Iderus and Edris, in-laws, and other family members for their support and motivation. Thanks are also due to all my friends, especially Suzi, Harlina & family, Halimah, for their encouragements and motivations.

My warmest gratitude to my husband, Mohd Yunus, for his love, patience, understanding, encouragement, as well as the sharing of housework and babysitting, while both of us were busy studying. Lastly, thank you to my children Faiz, Nisa and Syarafana. Thank you for asking "Mama tak pergi kerja hari ni?", thank you for the encouraging remark "Mama, work really hard okay" and thank you for understanding that I could not spend time with them when I was working.

## DEDICATIONS

*This work is especially dedicated to*

*My mother, Hajah Dyg Kijam Minan,*

*In memory of my late father, Awang Rambli,*

*and*

*My husband, Mohd Yunus and my children, Faiz, Nisa and Syarafana*

# TABLE OF CONTENTS

CERTIFICATE OF ORIGINALITY.....	ii
ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	iv
DEDICATIONS.....	v
TABLE OF CONTENTS.....	vi
LIST OF FIGURES.....	xi
LIST OF TABLES.....	xv
GLOSSARY OF TERMS.....	xvii
GLOSSARY OF ABBREVIATIONS.....	xx
 <b>PART I --THESIS BACKGROUND.....</b>	<b>1</b>
CHAPTER 1 INTRODUCTION.....	2
1 BACKGROUND.....	2
1.1 PERCEPTION AND SPATIAL AWARENESS IN VE.....	4
1.2 RESEARCH PROBLEM, AIM AND QUESTIONS.....	7
1.3 RESEARCH SCOPE.....	9
1.4 METHODOLOGY.....	9
1.4 PUBLICATIONS.....	10
1.5 THESIS OVERVIEW.....	10
 CHAPTER 2 PERCEPTUAL ISSUES OF VE.....	15
2 OVERVIEW.....	15
2.1 THE HUMAN VISUAL SYSTEMS (HVS).....	16
2.2 PERCEPTION OF SPACE.....	17
2.3 PICTORIAL DEPTH CUES.....	18
2.3.1 Static cues to depth.....	19
2.3.2 Cues from motion.....	22
2.4 PHYSIOLOGICAL DEPTH CUES.....	24
2.5 BINOCULAR DEPTH PERCEPTION.....	26
2.5.1 Stereo vision - illustration.....	26
2.5.2 The importance of stereo cues.....	28
2.6 COMBINATION AND RELATIVE IMPORTANCE OF DEPTH CUES.....	28
2.7 OTHER FACTORS THAT AFFECT DEPTH PERCEPTION.....	31
2.7.1 Size constancy.....	31
2.7.2 Prior knowledge.....	33

2.7.3	Cognitive Dissonance .....	33
2.7.4	Effects of sensory conflicts .....	34
2.8	PICTORIAL PERCEPTION .....	36
2.8.1	Geometrical theory on picture perception .....	39
2.9	PERCEPTION IN VE .....	41
2.9.1	Spatial awareness .....	42
2.9.2	Distance perception studies in VE .....	44
2.9.3	Spatial representation studies in VE .....	48
2.9.4	Factors affecting spatial perception in VE .....	51
2.10	CONCLUDING REMARKS .....	59
CHAPTER 3 TECHNOLOGICAL ISSUES IN VE .....		61
3	OVERVIEW .....	61
3.1	FUNDAMENTAL CONCEPT AND ISSUES IN VE MODELLING AND RENDERING .....	63
3.1.1	VE models .....	64
3.1.2	VE Image realism .....	68
3.1.3	Viewing/simulation of the VE model .....	75
3.1.4	Image Realism versus System Performance .....	77
3.2	TYPES OF VE .....	79
3.2.1	Non-immersive systems .....	82
3.2.2	Semi-immersive systems .....	82
3.2.3	Fully Immersive Systems .....	85
3.2.4	Comparison among the types of VR systems .....	88
3.3	CHAPTER SUMMARY .....	90
CHAPTER 4 BASIS FOR EXPERIMENTAL APPROACH TO UNDERSTANDING SPATIAL WARENESS .....		92
4	OVERVIEW .....	92
4.1	BASIS FOR EXPERIMENTAL WORKS .....	93
4.1.1	Task performance measures .....	95
4.1.2	Selection of factors to be examined and scope of investigation .....	100
4.1.3	VE Images modelling .....	117
4.1.4	Summary of experimental basis/approach .....	121
4.2	EXPERIMENTAL METHODOLOGY .....	123
4.2.1	Data collection .....	125
4.2.2	Data preparation and data analysis .....	138
4.2.3	Summary of all experiments in the thesis .....	142
4.3	CHAPTER SUMMARY .....	147



**PART II -- EXPERIMENTAL APPROACH, RESULTS AND ANALYSIS....148**

**CHAPTER 5 EXPERIMENT 1: DISTANCE PERCEPTION IN  
STATIC IMAGES.....149**

5 OVERVIEW ..... 149

5.1 EXPERIMENT 1A: EFFECT OF IMAGE TYPE AND DISPLAY TYPE..... 150

5.1.1 Rationale ..... 150

5.1.2 Experimental aim and hypotheses..... 151

5.1.3 Methodology ..... 152

5.1.4 Results..... 156

5.1.5 Analysis..... 158

5.1.6 Discussion ..... 160

5.1.7 Conclusion ..... 169

5.2 EXPERIMENT 1B: EFFECT OF DISPLAY SIZE ..... 170

5.2.1 Methodology ..... 171

5.2.2 Results..... 174

5.2.3 Analysis..... 179

5.2.4 Discussion ..... 179

5.2.5 Conclusion ..... 181

5.3 OVERALL CONCLUSIONS..... 182

**CHAPTER 6 EXPERIMENT 2: DISTANCE PERCEPTION IN  
DYNAMIC MAGES.....183**

6 OVERVIEW ..... 183

6.1 EXPERIMENT 2A: EFFECT OF IMAGE TYPE AND DISPLAY SIZE ..... 184

6.1.1 Rationale ..... 184

6.1.2 Experiment aims and hypotheses ..... 185

6.1.3 Methodology ..... 186

6.1.4 Results..... 192

6.1.5 Analysis..... 202

6.1.6 Discussion ..... 203

6.1.7 Conclusion ..... 207

6.2 EXPERIMENT 2B: EFFECT OF IMAGE TYPE AND DISPLAY SIZE..... 208

6.2.1 Rationale ..... 208

6.2.2 Experiment aim and hypothesis ..... 211

6.2.3 Methodology ..... 211

6.2.4 Results..... 213

6.2.5 Analysis..... 221

6.2.6 Experiment 2B Discussion..... 233

6.2.7	Experiment 2B Conclusion .....	226
6.3	OVERALL CONCLUSIONS.....	227

## CHAPTER 7 EXPERIMENT 3: DISTANCE PERCEPTION AND SPATIAL MEMORY TASK IN INTERACTIVE IMAGES.....228

7	OVERVIEW .....	228
7.1	EXPERIMENT 3A: EFFECT OF VIEWING DISTANCE, PHYSIOLOGICAL CUES, INTERFACE DEVICE AND TRAVEL MODES.....	229
7.1.1	Experimental Aims & Hypotheses.....	229
7.1.2	Methodology .....	230
7.1.3	Results.....	244
7.1.4	Discussion .....	262
7.1.5	Conclusions.....	275
7.2	EXPERIMENT 3B: EFFECT OF DISPLAY SIZE .....	277
7.2.1	Display questionnaire.....	281
7.2.2	Results.....	281
7.2.3	Discussion .....	288
7.2.4	Conclusion .....	290
7.3	OVERALL CONCLUSIONS.....	290

## PART III -- IMPLICATIONS DRAWN FROM THE LITERATURE AND EXPERIMENTS UNDERTAKEN.....292

### CHAPTER 8 OVERALL ANALYSIS OF RESULTS.....293

8	OVERVIEW .....	293
8.1	SUMMARY OF ALL EXPERIMENTS' RESULTS .....	294
8.2	OVERALL ANALYSIS BASED ON DISTANCE ESTIMATE TASK PERFORMANCE .....	297
8.2.1	Comparison based on image types (Real and VE).....	297
8.2.2	Comparison based on display size (Large and Small) .....	300
8.2.3	Comparison among asymmetrical distances .....	304
8.3	OVERALL ANALYSIS BASED ON SPATIAL MEMORY TASK .....	306
8.3.1	Comparison based on image types (Real and VE).....	306
8.3.2	Comparison based on display size (Large and Small) .....	308
8.3.3	Interface device.....	311
8.3.4	Travel mode .....	311
8.4	CONCLUDING REMARKS .....	311

### CHAPTER 9 FINAL CONCLUSIONS, RESEARCH CONTRIBUTION AND IMPLICATION FOR SPATIAL AWARENESS PERCEPTION IN VE.....313

9	OVERVIEW .....	313
9.1	FINAL CONCLUSIONS AND CONTRIBUTIONS.....	314
9.1.1	Distance estimates tasks.....	314
9.1.2	Spatial memory tasks in interactive images experiment .....	318
9.1.3	Effect of image modelling on real and VE comparisons .....	320
9.1.4	Key Research Questions Addressed.....	321
9.1.5	Scope of conclusions.....	323
9.1.6	Method contributions .....	324
9.2	IMPLICATION FOR SPATIAL AWARENESS PERCEPTION IN VE.....	325
9.2.1	Implications of results from distance estimate task performance .....	325
9.2.2	Implications of results from spatial memory task performance .....	329
9.3	Concluding remarks.....	332
	CHAPTER 10 RECOMMENDATIONS AND FUTURE WORKS.....	334
10	OVERVIEW .....	334
10.1	RECOMMENDATIONS AND FUTURE WORKS.....	334
10.2	CONCLUDING REMARKS .....	339
	REFERENCES.....	341
	APPENDIX A: EXPERIMENT 1 ON DISTANCE IN STATIC IMAGES: SUMMARIES, TEST MATERIALS & COLLECTED DATA.....	356
	APPENDIX B: EXPERIMENT 2 ON DISTANCE PERCEPTION IN DYNAMIC IMAGES: SUMMARIES, TEST MATERIALS & COLLECTED DATA .....	365
	APPENDIX C: EXPERIMENT 3 ON DISTANCE PERCEPTION AND SPATIAL MEMORY IN INTERACTIVE IMAGES: SUMMARIES, TEST MATERIALS & COLLECTED DATA.....	376
	APPENDIX D: PUBLICATIONS.....	405

# LIST OF FIGURES

Figure 1-1 Logical structure of thesis.....	11
Figure 2-1 Human FOV. Adapted from Kalawsky (1993).....	17
Figure 2-2 Static cues to depth .....	19
Figure 2-3 Woodcut by Albrecht Durer's 1525 illustrating perspective. ©Bettmann/CORBIS. Adapted from Sekular and Blake (2002), pp306.....	20
Figure 2-4 Height in the visual field.....	21
Figure 2-5 Motion parallax Adapted from Coren and Ward et al. (1999), pp264 .....	22
Figure 2-6 Optical flow pattern Adapted from Schiffman (1990), pp320.....	22
Figure 2-7 Accommodation of an image by adjusting the shape of the lens. ....	24
Figure 2-8 A version of horopter. Points of X on the horopter will fall in the corresponding retinal points of the two eyes yielding a single image. Other points outside the Panum' area will yield double image. Adapted from Schiffman (1990), pp356.....	26
Figure 2-9 Some basic geometry for stereoscopic viewing .....	27
Figure 2-10 Categorization of sources of depth information into three types of spaces based on distance from the observer. The more potent cues are reflected by the small values of the depth contrast. (Adapted from Cutting (1997)).....	29
Figure 2-11 Size constancy.....	31
Figure 2-12 Ponzo illusion .....	32
Figure 2-13 Muëller-Lyer illusion.....	32
Figure 2-14 Close viewing compresses geometrically specified depth .....	40
Figure 2-15 Close viewing leaves geometrically specified virtual frontal dimension unchanged (Adapted from Sedgwick 1991) .....	41
Figure 2-16 Close viewing distorts geometrically specified virtual shape.....	41
Figure 3-1 A polygonal 3-D elephant.....	65
Figure 3-2 Spheres of varying level realism based on the number of polygons used.....	65
Figure 3-3 Cartesian coordinates systems .....	66
Figure 3-4 A partial database representations for a room model in Multigen II Pro Software.....	68
Figure 3-5 Examples of texture mapping. Adapted from Wolfe (1997).....	74
Figure 3-6 Rendering pipeline (Adapted from Bethel (1999)) .....	76
Figure 3-7 Zeltzer unit cube model (AIP cube) (1992) .....	80
Figure 3-8 Examples of semi immersive systems: Immersadesk(left) located in SL 239 at Indiana University - Purdue University Indianapolis,USA; panorama systems (middle) developed by Panoramtech at <a href="http://www.panoramtech.com">www.panoramtech.com</a> ; ImmersiveWall (right) located at AVVRC, Loughborough University, UK.....	83
Figure 3-9 Shutter glasses.....	84
Figure 3-10 Examples of fully-immersive systems .....	85

Figure 4-1 Asymmetrical distances of vertical, horizontal and transverse .....	97
Figure 4-2 Relationship between FOV and viewing distance .....	101
Figure 4-3 The figures indicate comparison of display FOV with human FOV. The dark black lines indicate the left and right eye FOV. The box in the images represent the FOV of a 17" CRT monitor viewed at 50 cm (left image) and the FOV of typical HMD (right image). Adapted from Pfautz (2000).....	102
Figure 4-4 Experiment setup of previous investigations: Similar visual angle ( $x = y$ ) and.....	105
Figure 4-5 Experiment setup: Different FOV ( $x > y$ ) and similar viewing distance ( $d$ ).....	106
Figure 4-6 Factors and tasks performance investigated in this thesis.....	123
Figure 4-7 Experiment setup for all experiments .....	127
Figure 4-8 For Experiment 1A and 2: 2 x 2 factorial designs (2 factors: each factor has two levels; 4 experimental conditions). Factor 1 is image type (real and VE). Factor 2 is display type (desktop monitor and projected display).....	128
Figure 4-9 Experiment 1B: One-factor design with two levels (2 experimental conditions). Factor 1 is display factor (large and small).....	128
Figure 4-10 For Experiment 3: 2 x (2 x 2) mixed design (3 factors: each factor two levels; one between-subject and two within subject design; 8 experimental conditions for each between-subject factor level). Factor 1 is display factor (large and small). Factor 2 is interface device factor (mouse and trackball) and Factor 3 is travel mode factor (drive and fly). It should be noted that the fourth factor (image type) is not depicted here. This is because the real condition does not investigate factor 2 and 3. ....	137
Figure 5-1 Real picture .....	152
Figure 5-2 Virtual picture .....	153
Figure 5-3 Experiment 1A setup.....	155
Figure 5-4 Mean estimated distance for horizontal distance .....	157
Figure 5-5 Mean estimated distance for transverse distance .....	158
Figure 5-6 Percentage of estimate from actual for transverse and horizontal distance – comparing image types .....	159
Figure 5-7 Percentage of estimate from actual for transverse .....	160
Figure 5-8 Egocentric distance verses exocentric transverse distance .....	163
Figure 5-9 Exocentric distance (Horizontal distance) .....	165
Figure 5-10 Experiment 1A Setup.....	168
Figure 5-11 Real picture .....	172
Figure 5-12 Experiment 1B setup for small (a) and large display (b). X and Y represent the projected image size on the screen and $\alpha$ and $\beta$ represent the FOV for small and large display condition respectively. 'd' represent the fixed distance.....	173
Figure 5-13 Comparison of mean estimated distance between display size for vertical distance .....	174
Figure 5-14 Comparison of mean of estimated vertical distance to the actual distance between the large and small displays. Accompanying table indicates points on the graph.....	175
Figure 5-15 Comparison of mean estimated distance between display size for horizontal distance .....	176

Figure 5-16 Comparison of mean of estimated horizontal distance to the actual distance between the large and small displays. Accompanying table indicates points on the graph .	176
Figure 5-17 Comparison of mean estimated distance between display size for transverse distance .....	177
Figure 5-18 Comparison of mean of estimated transverse distance to actual distance between large and small display. Accompanying table indicates points on the graph.....	178
Figure 5-19 Comparison among asymmetrical distances .....	178
Figure 6-1 Real Environment .....	186
Figure 6-2 Virtual Environment .....	187
Figure 6-3 (a) original texture copy (b) repetition of the same texture pattern (c) copies opposite edge of the pattern to make it similar to the edges of the centre pattern.....	188
Figure 6-4 Images - Method of acquisition and display .....	189
Figure 6-5 Experiment 2A display setup. X is the viewing distance., $2\alpha$ is the display FOV and $2y$ is the image size. ....	190
Figure 6-6 Experiment 2A setup. Small display condition indicated by viewing distance $d_1$ , FOV $\alpha$ and image size X. Large display condition is indicated by viewing distance $d_2$ , FOV $\beta$ and image size Y. The eye level is set to be at the centre of projection.....	191
Figure 6-7 Vertical distance: Mean percentage of estimate for the real and VE on a large and small display .....	193
Figure 6-8 Mean percentage of estimated distances for all the five vertical distances among all experimental conditions .....	195
Figure 6-9 Horizontal distance: Mean percentage of estimate for the real and VE.....	<b>Error!</b>
<b>Bookmark not defined.</b>	
Figure 6-10 Mean percentage of estimated distances for all the five horizontal distances among all experimental conditions. ....	196
Figure 6-11 Transverse distance: Mean percentage of estimate for the real and VE .....	197
Figure 6-12 Mean percentage of estimated distances for all the five transverse distances among all experimental conditions .....	198
Figure 6-13 Comparison among distance types under the four experimental conditions.....	200
Figure 6-14 Experiment setup for Experiment 2B (a) and Experiment 2A (b). X and Y represent the physical image size for small and large display respectively. $\alpha$ and $\beta$ are the corresponding FOV for both display size .....	210
Figure 6-15 The original high resolution VE movie.....	212
Figure 6-16 Vertical distance: Mean percentage of estimate for the real and VE .....	213
Figure 6-17 Mean percentage of estimated distances for all the five vertical distances among all experimental conditions .....	214
Figure 6-18 Horizontal distance: Mean percentage of estimate for the real and VE.....	215
Figure 6-19 Mean percentage of estimated distances for all the five horizontal distances among all experimental conditions .....	216
Figure 6-20 Transverse distance: Mean percentage of estimate for real and VE on .....	217
Figure 6-21 Mean percentage of estimated distances for all the five transverse distances among all experimental conditions.....	218

Figure 6-22 Comparison among asymmetrical distances under the four experimental conditions .....	219
Figure 7-1 Layout of the experimental room.....	231
Figure 7-2 Picture of the Real Environment with objects for the .....	232
Figure 7-3 Different views of the 3D VE model .....	233
Figure 7-4 Practice environment for VE conditions .....	233
Figure 7-5 Experiment 3A setup. Viewing distance (d) was fixed for both the large and small displays but the FOV was varied ( $\beta > \alpha$ ).....	236
Figure 7-6 Mouse ( left image) & Trackball (right image).....	237
Figure 7-7 Participants' initial position in the Real and VE conditions for exploration.....	238
Figure 7-8 Summary of experimental variables for Experiment 3A .....	241
Figure 7-9 Real condition test trial .....	243
Figure 7-10 VE conditions: (a) Large display (test trial) (b) Small display (practice session).....	244
Figure 7-11 Comparison among real and VE conditions for height, width, and length .....	247
Figure 7-12 Comparison among real environment, VE/large and VE/small conditions in terms of the mean number of object correctly placed for all interface device/travel conditions... <b>Error!</b>	
<b>Bookmark not defined.</b>	
Figure 7-13 Comparison of practice time (left) and maptest time (right) among conditions .....	256
Figure 7-14 Median score for Q2i, Q2ii, Q2iii and Q2iv of interface device questionnaire ... <b>Error!</b>	
<b>Bookmark not defined.</b>	
Figure 7-15 Median scores for Q3i, Q3ii, Q3iii, Q3iv, Q3v of interface device questionnaire....	260
Figure 7-16 Distance estimates for each participants in the VE conditions (Case 1-10 is real world conditions data which was not shown).....	266
Figure 7-17 Experiment 3A setup. Left diagram represent small display setup and right diagram represent the large display setup .....	279
Figure 7-18 Experiment 3B setup. Retinal image size ( $x_1 = y_1$ ). FOV ( $\alpha = \beta$ ). Viewing distance ( $d_1 \neq d_2$ ). Physical image size ( $x \neq y$ ).....	279
Figure 7-19 Comparison of median score between large and small display for each question of display size questionnaire.....	282
Figure 7-20 Comparisons between large and small display in terms of asymmetrical distances .	283
Figure 7-21 Comparison among large and small display, mouse and trackball and travel modes	284
Figure 7-22 Median score for Q2i, 2ii, 2iii, 2iv of the Interface device questionnaire.....	286
Figure 7-23 Median score for Q3i, 3ii, 3iii, 3iv, 3v of the Interface device questionnaire.....	287

## LIST OF TABLES

Table 3-1 Comparison between different VR systems implementations in terms of qualitative performance (Adapted from Kalawsky (1996)) .....	88
Table 4-1 Summary of IV and DV variables examined in all experiments.....	126
Table 4-2 Methods of measuring distance perception: comparison of direct and indirect methods .....	136
Table 4-3 Statistical test used to analyze data in all experiments.....	141
Table 4-4 Summary of the three experiments.....	143
Table 5.1 Summary of experiment setup variables.....	154
Table 5-2 Comparisons among display types for horizontal and transverse distances.....	159
Table 6-1 Summary of experiment variables.....	191
Table 6-2 Results of comparisons among experimental conditions using Students' t-test.....	195
Table 6-3 Results of comparisons among experimental conditions using Students' t-test.....	197
Table 6-4 Results of comparisons among experimental conditions using Students' t-test.....	199
Table 6-5 Results of t-test values for comparison among distance types under the four experimental conditions .....	199
Table 6-6 Summary of result from ANCOVA analysis using sport variable as covariate for all distance type (vertical, horizontal and transverse).....	201
Table 6-7 Results of comparisons among experimental conditions using Students' t-test.....	215
Table 6-8 Results of comparisons among experimental conditions using the Students' t-test.....	217
Table 6-9 Results of comparisons among experimental conditions using Students' t-test.....	219
Table 6-10 Results of t-test values for comparison among distance types under the four experimental conditions .....	220
Table 6-11 Results of MANOVA analysis (Multivariate tests) using sport variables as covariate for all distance types (vertical, horizontal and transverse).....	221
Table 7-1 Summary of Real and VE participants' background.....	231
Table 7-2 Question 2 on Mode of travel in VE .....	239
Table 7-3 Question 3 on Interface device use in VE .....	239
Table 7-4 Summary of effect of sport variable and display on distance estimate .....	248
Table 7-5 Comparison of adjusted mean score between display size for all distance type .....	250
Table 7-6 Mean, standard deviation and standard error for all conditions among display types..	250
Table 7-7 Results of one-way ANOVA for comparison among groups (Real, VE/large and VE/small ) for all conditions .....	252
Table 7-8 Post-hoc comparisons among display types.....	252
Table 7-9 Comparisons among experimental conditions.....	253
Table 7-10 Summary of mean and standard error .....	255
Table 7-11 Summary of analysis on the effect of practice time and map-test time.....	255
Table 7-12 Summary of main effect and interaction for Q2(i) –(iv) .....	259



Table 7-13 Summary of main effect and interaction for Q3(i)-(v).....	261
Table 7-14 Summary of main effects and interaction for Q4 .....	261
Table 7-16 Comparison of means among experimental conditions.....	283
Table 7-17 Comparison of means among experimental conditions.....	285
Table 8-1 Summary of results for distance estimates for all experiments (1,2, and 3).....	294
Table 8-2 Summary of results for spatial memory task and post-test questionnaire for Experiment 3A and 3B.....	295
Table 8-3 Comparison of range of estimates among experiments in outdoor and indoor setting.....	306

## GLOSSARY OF TERMS

Accommodation	The process whereby the eye adjusts the lens shape to focus on near and far objects for a sharp image on the retina
Alpha level	The alpha ( $\alpha$ ) level is defined as the probability of what is called a Type I error in statistics. That is the probability of rejecting null hypotheses when in fact it was true. Also referred to as significant level.
Asymmetrical distance	Refers to vertical, horizontal and transverse
AVI	Audio Video Interleave. A common for audio/video data on the PC.
Binocular disparity	Refers to the difference between perceived images received from both eyes
Bi-ocular display	Displays that present the same image to both eyes
Constructive Solid Geometry	A computer-aided design strategy, based on fact that some objects such as sphere, cylinder, and ellipsoid can be described using mathematical equations.
Convergence	Inward movements of the two eyes to focus on near objects
Cyber sickness	Cyber sickness or simulator sickness occurs when the user is stationary but experience a compelling sense of self-motion though moving visual imagery. User often exhibits the following symptoms of cyber sickness: eyestrain, ataxia, fatigue, and drowsiness
Divergence	Outward movements of the two eyes to focus on far objects
Egocentric distance	Distance from self to an object
Exocentric distance	Distance between two objects or points within the same objects
Field of view	The angle subtended by the display device on the viewer's retina.
Frame rate	The rate at which new updated scene is rendered or prepared for drawing to the screen.
Geometric field of view	The visual angle subtended by the virtual scene
Horizontal distance	Refers to distance across the screen
Immersion	Refers to the extent of the peripheral imagery. It also refers to the extent in which the computer display are extensive, surrounding, inclusive, vivid and matching

Level of Details	A strategy which creates several levels of details (LOD) of objects in the database and present the appropriate LOD of these objects based on their distance from the viewer. A high LOD representation will be presented for near distance and a low LOD representation will be selected for far distance.
MIP mapping	A technique which uses a set of texture maps of different resolutions to represent different distances of object from the viewer. Similar to LOD concept.
Monocular cues	Cues that can be viewed using either eye alone. Examples are linear perspective, occlusion and texture gradient
Motion cues	Cues that is available when there is motion in either the viewer or the viewed scene or both. Examples are motion parallax and motion perspective
Physiological cues	Refers to accommodation and vergence cues
Polygon	Flat surfaces which have at least three edges or lines. Also known as faces
Presence	The sense of “being there” in the virtual environment
Proprioception	Refers to the awareness of the body. This awareness is derived from the information provided by the receptors in our muscles, tendons and joints. It gives information about the movement and positions of parts of our body
Real-time	Refers to the presenting and updating of images according to the observer’s current view
Refresh rate	Refresh rate refers to the frequency which the display hardware can draw the image on the display surface.
Scene graph	A scene graph is a collection of objects organized in a hierarchical tree-like form called directed-acyclic graph where objects are grouped according to location in the scene.
Simulator sickness	See cyber sickness
Situational awareness	The ability to know what is happening around us
Spatial awareness	Spatial awareness refers to the awareness of the 3D environment, which includes knowledge and understanding of objects’ spatial locations and relative distances within that environment.
Spatial resolution	The number, angular size and the spacing of the pixels
Stereopsis	Refers to the unique appearance of depth with solidity
Surface patch	Surface patch is based on mathematical techniques to create a

small smooth surface and these surfaces can be combined to forms larger complex smooth surfaces.

Texture mapping	A technique to improve image realism based on projecting photographic images (textures) onto polygon-based objects
Transverse distance	Refers to distance into the screen
Update rate	See frame rate
Vertical distance	Refers to height distance
Vertical-Horizontal illusions	Refers to a condition when a physical vertical extent is overestimated in length relative to a comparable physical horizontal extent
Vergence	Refers to the inward or outward movement of the two eyes in order to focus a sharp image on the retina.
Virtual environment	A computer-generated simulation of an environment typically designed to represent and provide experience of places or locations in a real, abstract or even a non-existent environment.
Visual acuity	The ability of a person or an animal to detect fine spatial pattern and resolve details

## GLOSSARY OF ABBREVIATIONS

3D	Three dimension
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
AVI	Audio Video Interleave
DV	Dependent variable
FOV	Field of view
ft	feet
GFOV	Geometric field of view
HMD	Head-mounted display
HVS	Human Visual Systems
IV	Independent variable
LOD	Level of Details
MANOVA	Multivariate analysis of variance
NURBS	Non-Uniform Rational Bezier-Spline
PC	Personal computer
Q	Question
SA	Situation awareness
SGI	Silicon Graphics Inc.
VE	Virtual environment
VHI	Vertical-Horizontal illusions
VOR	Vestibular Ocular Reflex
vs.	versus
cm	centimetre
m	metre
%	percent
°	degree

# PART I

## THESIS BACKGROUND

**Chapter 1 – Introduction**

**Chapter 2 – Perceptual Issues of VE**

**Chapter 3 – Technological Issues of VE**

**Chapter 4 – Basis for an Experimental Approach to  
Understanding Spatial Awareness**

# CHAPTER 1

---

## INTRODUCTION

### 1 BACKGROUND

Virtual environment (VE) has been attracting profound interest in various fields of applications. The past decade has seen the adoption of VE technologies in diverse areas such as training (e.g. vehicle simulation, fire fighting and flight training), prototyping (e.g. product design), medicine (e.g. psychiatric treatment and surgery training), tele-operations of robots, visualization of complex data sets, architecture (e.g. walkthrough and design), entertainment (e.g. virtual rides and virtual games), archaeology and education (Kalawsky 1993, Brooks 1999). There exist various definitions of VE in the literature. Collating definitions from several researchers (Kalawsky 1993, Barfield and Furness 1995, Ellis 1994) basically defines VE as an interactive immersive experience for the user in a computer-simulated world. In this definition, VE is referred to as a computer-generated environment rather than the technologies that are often associated with it. Interactive experience means the ability to manipulate objects in the VE. The term immersive denotes “the extent of the peripheral display imagery” (Kalawsky 2000). It includes the extent in which the computer displays are *extensive, surrounding, inclusive, vivid and matching* (Slater, Usoh et al. 1996). In more general terms, a VE is a computer-generated environment typically designed to represent and provide experience of places or locations in a real world, abstract or even a non-existent world. VE

that represents real world spaces or as surrogates to real world places is useful when the real environment is not safe, practical or too costly to be explored (Witmer, Bailey et al. 1996). For example, in training (such as flight training and fire fighters training) trainees can practise in a safe VE instead of training in actual places or situations which are rare, remote or dangerous (Waller, Hunt et al. 1998a, Sinai, Krebs et al. 1999). In prototyping, architecture and regional planning, VE allows designers, clients, and decision makers an early preview of the planned 3D space through simulated environments, thus enabling cost and time saving decisions to be made prior to the delivery of the actual physical structure or product (Henry and Furness 1993). In crime scene reconstruction, the preservation of the crime scene in computer generated VE allows investigators to “revisit” the crime scene for subsequent investigations even though it may no longer be available (Howard, Murta et al. 2000, Morley 2002). In addition, VEs can also be used to model complex environments which are inaccessible in the real world such as atomic structures and living cells or environments which do not exist in the physical form in the real world such as scientific and financial data (Witmer, Bailey et al. 1996).

To simulate a real environment may involve reproducing its aspect as accurately as possible in order to give the illusion of alternate reality. This includes emulating the spatial representation (dimensions of width, height, and depth) and the spatial awareness aspects that will be experienced by the user. This may be necessary particularly for VE applications that require a user to use spatial judgment tasks or to learn the spatial characteristics of the VE in order to apply them to a real world setting. Whilst a VE provides a user with access to information that would not be available at that particular space or time based on the human perception of visual information in three spatial dimension (which may be enhanced by other sensory stimuli), Wann and Mon-Williams (1996) further pointed out that, “there is no implicit assumption that it provides all categories of information or that it perfectly mimics a natural setting.” Thus, whilst it is critical for some applications such as flight simulator training to closely imitate the real world setting in all respects, for others (such as in architectural design, education and entertainment) it may not be necessary to do so (Riley and Kaber 1999).

It has been suggested that as VE enables a user to explore and interact within a 3D virtual space this requires human spatial perception for its effective use (Wann and Mon-Williams 1996). Therefore for applications that use VE as surrogates (such as in visualization and training), it is important to allow users to perceive the virtual 3D space and spatial relations in the VE similar to the real world. The increasing number of such applications has made it essential to examine the ability of VE technologies to provide a convincing simulation of the



real world places. The results of recent investigations into comparing perception in both environments, however, have been varying. Some studies have reported that it is possible to perceive the VE similar to the real world (Waller 1999, Yoon, Byun et al. 2000). Others have reported perception in VE as not being very accurate in terms of distance perception compared to the real world (Henry and Furness 1993, Lampton, McDonald et al. 1995, Witmer and Singer 1998, Witmer and Kline 1998, Waller, Hunt et al. 1998a, Sinai, Krebs et al. 1999, Loomis and Knapp 2003). These inconsistencies make it difficult to generalize findings from these studies. As the success of the applications that uses VE as surrogates to the real world places depends on VE technologies providing similar spatial perception and experience in both worlds, it is important to examine and understand factors that affect user's perception of the VE in order to inform the efficient and effective design of the VE. Various factors, particularly those relating to display systems (Waller 1999), have been suggested and investigated but the exact reasons for perceptual difference between real and VE are still unknown (Willemson and Gooch 2002). Thus, entails the need for further examining of factors influencing perception in the VE.

In this dissertation, the research studies presented aim to provide and enhance current knowledge and understanding towards similar perception of VE to the real world counterpart. To assess simulation fidelity of a VE, a commonly used strategy which compares task performance in the VE to similar task performance in the real world is employed (Witmer and Sadowski 1998, Mania 2001). It has been asserted that this method could provide knowledge on aspects of VE technologies that need to be improved (Witmer and Sadowski 1998). Kalawsky (2000) further stated that this comparison is useful particularly if the VE is to imitate the real world in some respects. Since spatial awareness is crucial for human performance efficiency (Mania 2001) and the utility of VE for any application for which they are being proposed is predicated upon the accuracy of spatial representation in the VE (Arthur, Hancock et al. 1997), examining factors affecting spatial awareness in VE is the focus of this thesis, whereby the spatial tasks performed in the VE are compared to similar tasks in the real counterparts.

## 1.1 PERCEPTION AND SPATIAL AWARENESS IN VE

In many instances, the presentation of 2-D images is adequate for some applications. However, expanding this presentation to a 3D format which can be explored interactively would be more useful as it provide more information to the viewer (Wann and Mon-William 1996). This is because the levels of details of a 3D structure presented in 2-D images are hidden unless the viewpoint can be changed interactively. This is due to the effect of

occlusion or interposition (see also Section 2.3.1.1 of Chapter 2). Consequently, presenting a 3D simulation of the real world and providing the flexibility to explore and view this environment from different perspectives interactively sets VE applications apart from their other traditional counterparts such as pictures, computer animations and movies. These features make it potentially an attractive tool for a wide range of areas of applications described earlier.

However, Wartenberg and Wiborg (2003) argued that the “accuracy of space perception and distance estimation in VE is an important precondition for the reliable use of virtual techniques in the design of products, workplaces, architecture and production systems.” The success of the applications that uses VE as surrogates to the real world places would depend on VE technologies offering similar spatial perception and experience in both worlds. Thus, one of the goals of the VE technologies is to create an environment that faithfully represents the real world environment where users must be allowed to perceive spatial relations in the VE in the equivalent way as they would in the real world. However, most available VEs are not modelled as exact replicas of the real world places whereby spatial properties and not all sensory cues are not available to the viewers. As such, several questions arise from the use of such VEs as surrogates to real world places:

- To what extent can experience gained in the VE be used to represent the real world? In other words, how similar is experience gained in the VE to the real world?
- If the experience is not similar, how can the user’s perception of a VE be made similar to the real world? What are factors affecting a user’s perception in the VE?

Similar questions have been the focus of several researchers (Henry and Furness 1993, Ruddles and Paynes et al. 1998, Waller and Hunt et al. 1998, Yoon, Byun et al. 2000). These questions serve to motivate the research works in this dissertation. As stated earlier, previous studies indicated that VE is often perceived differently from the real world and results from these studies are often varying and contradicting (Henry and Furness 1993, Lampton, Bliss et al. 1994, Witmer and Kline 1998, Witmer and Sadowski 1998, Waller 1999, Patrick, Cosgrove et al. 2000, Yoon, Byun et al. 2000, Willemson and Gooch 2002, Bideau, Kulpa et al. 2003, Youngblut and Huie 2003, Messing 2004, Thompson, Willemson et al. in press). Whilst there is continuing interest in the research community in this direction as shown by the number of related studies, current knowledge on factors necessary to provide similar perceptual and performance experience to the real world is still limited. Despite the popular interests in the use of VE for various applications such as training, visualization and entertainment, there is still a paucity of knowledge on factors influencing the user’s

perception and spatial knowledge using VE technology (Cutmore, Hine et al. 2000). This signifies the need for further research work. Since the usefulness of application that employ VE as surrogates to real world places relies upon similar perception of space in VE to the real world, this thesis examines factors affecting user's spatial perception in the VE in comparison to similar perception in the real environments.

The display systems types and properties have been suggested as one of the factors affecting distance underestimation in the VE (Egglestons, Jansen et al. 1996, Witmer and Kline 1998, Willemson and Gooch 2002). A VE experience depends on the visual display system's ability to simulate the human visual sensory channels. Although for spatial orientation tasks, some researchers (Bakker, Werkhoven et al. 1998) have shown that information from the vestibular channel is more important than visual cues, synonymous with the human perception system, the visual channel represents the most dominant sensory channel compared to other channels (for examples auditory, haptic and tactile) in a VE (Pfautz 2002). This highlights the need to enhance the capabilities of the visual display system in order to closely match the VE visual experience to the real one.

Various related aspects of the display or the computer systems have been the focus of intensively studied factors to explain perceptual differences between real and VE (Waller 1999). This includes variables such as display types, scene contrast; navigational interface and field of view (FOV). However, very few studies have examined the effects of display size on spatial awareness, indicating that more research is needed.

Spatial awareness refers to our awareness of the elements within a 3D environment which includes the knowledge and understanding of object locations and relative positions in that 3D space. Basically, it refers to the perception of the 3D space layout. In the real world, spatial awareness is critical to human performance efficiency, as such spatial tasks are often used in benchmarking processes (Mania 2001). The usefulness of applications which utilize VE technologies depends on the accuracy of how space is represented in the VE (Arthur, Hancock et al. 1997). Additionally, it has been suggested that accurate perception of space and distance estimates forms an important prerequisite for the reliable use of the VE technologies in such applications (Wartenberg and Wiborg 2003). For these reasons spatial tasks will be used in this thesis as a performance measure in the evaluation of perceptual experience in the VE compared to the real world environment.

Some initial studies (Patrick, Cosgrove et al. 2000, Czerwinski, Tan et al. 2002, Tan, Gergle et al. 2003) which investigated the effect of display size on task performance (such as spatial

orientation, spatial memory and navigation) reported better performance of large display participants over small display whilst others reported no difference between large and small display on distance estimation, reading and spatial memory tasks (Arthur 2000, Johnson and Stewart 1999, Tan, Gergle et al. 2003). These inconsistencies provide motivation for further investigation.

In order to investigate the effect of display size, in these previous studies, it was necessary to maintain a similar visual angle for both display sizes to remove the effect of FOV. In order to maintain a similar visual angle, the distance of the observer from the display needs to be varied for both display sizes. By employing this method, these studies failed to take the effect of viewing distance into consideration. It has been reported that an object viewed at a greater distance portrayed large distances compared to equivalent scene viewed at a shorter distance (Gooding, Miller et al. 1991). This implies that viewing distance may also contribute for the better performance of large display in these previous studies. Moreover, different viewing distances from the picture may also result in different physiological cues acting at these different distances. It has been asserted that the distance of accommodation may influence the perceived size and the distance of an image (Iavecchia, Iavecchia et al. 1988). Some empirical evidence has revealed that our eyes converge and accommodate at varying distances in the picture (Enright 1987a, Enright 1987b, cited in Coren, Wards et al. 1999). Although limited in the range of distance for their effectiveness (Sekuler and Blake 1994), these physiological cues may also contribute to the better performance of large display participants over small display participants in the previous investigations. As such the research presented in this thesis expands on the previous works by considering and examining the effects of these latter factors (viewing distance and physiological cues).

In the real world, the perceptual understanding of the 3D environment is derived from different sources of information (or depth cues). Many of these cues can be represented in the computer generated VE (Witmer and Kline 1998). An understanding of perception in the real world is essential to comprehend perception in the VE. The perception of space in the real world and in the VE is further discussed in the Chapter 2. A review of the perceptual issues in a VE and the related literature is also presented.

## 1.2 RESEARCH PROBLEM, AIM AND QUESTIONS

A VE provides users with a 3D visual experience of a computer-generated representation of real, abstract or non-existent places or locations. It permits them to interact and explore this virtual 3D space in real-time which requires the user's spatial perception for effective use

(Wann and Mon-William 1996). Additionally, to be effective most applications using VE technologies rely on the need for these technologies to enable users to form an accurate perception of the space in these virtual spaces (Arthur and Hancock 1997, Wartenberg and Wiborg 2003). This indicates that the usefulness of VEs in these applications depends on how it provides the spatial experience that closely matches those of the real environment.

Empirical evidence from recent studies examining factors necessary to yield similar perception to the real world reveals inconsistent results. Some researchers (Witmer and Kline 1998, Henry and Furness, Lampton, Bliss et al. 1994, Lampton, McDonald et al. 1995, Witmer and Sadowski 1998) indicated that the user's spatial performance in a VE differs from the real environment while others demonstrated that it is possible to perceive a VE as being similar to the real world within the given constraint (Waller 1999, Yoon, Byun et al. 2000). Moreover, some researchers reported an underestimation of distance perceived compared to the real world estimates (Witmer and Kline 1998, Henry and Furness, Lampton, Bliss et al. 1994, Lampton, McDonald et al. 1995), others reported overestimation of estimated distance compared to the actual distance (Waller 1999).

Various factors have been attributed for the perceptual difference in the real and VE but the exact reasons for these differences are still far from being resolved (Willemson and Gooch 2002). It has been suggested that comparing human task performance in the VE to a similar task performance in the real world can provide clues as to which aspect of the VE technologies require improvements (Witmer and Sadowski 1998). In addition, determining the circumstances under which perceptions are systematically distorted in VE represents a major step towards understanding the limit of VE (Waller 1999). Consequently, examining factors that influence users' spatial perception performance in VE could contribute towards a more effective and efficient design of the VE, where the task performed in the VE is similar to the task performed in the real world. Thus, the overall aim of this thesis is

“To examine spatial awareness in the real and VE by evaluating factors influencing spatial performance in these environments”

The four main research questions evaluated in this thesis are as follows:

- Is there a difference in spatial tasks (distance estimation and spatial memory task) performed in the real and VE?
- How does the display size (large and small) affect users' spatial task (distance estimation and spatial memory task) performance in the real and VE?

- How does the type of interface device (mouse and trackball) affect the users' spatial task performance (spatial memory task) in a VE?
- How does the type of travel mode (drive mode and fly mode) affect the user's spatial task performance (spatial memory task) in a VE?

The reasons for evaluation of these research questions in this thesis are further discussed in Chapter 4. The real and VE refer to comparisons among different forms of image presentations: static, dynamic and interactive forms. This involves comparisons between static real and static VE image, dynamic real and dynamic VE images and real physical environment and interactive VE (see section 4.1.3 of Chapter 4).

Knowledge gained from this research work will augment existing literature on spatial awareness in the VE and provide information and guidelines for designers and users of VE applications. Factors which may contribute towards cost effective use of VE and human performance efficiency in the VE will be highlighted. The output of the research will also explain the effects of sensory conflicts which exist in the VE compared with the real world.

### 1.3 RESEARCH SCOPE

The scope of investigation of the research presented in this thesis is limited to the followings:

- Non stereo image presentation
- Static, dynamic and interactive presentation of images
- Presentation of image on non-head-tracked, non-immersive and semi-immersive displays
- The VE models used are based on accurate geometric representation and photorealistic texture object only to create the realism effect.

In addition, initially, it was intended to include the examination of the impact of audio cues on spatial tasks performance. Due to time constraints, the research's main focus is on visual factors only. In the real world, the visual cues for spatial information are often redundantly supplemented by these later cues. Moreover, the purpose of the research was to examine users' perception of space and spatial relations in VE in comparison to similar perception in the real world. Therefore only visual cues were examined in this thesis. Further arguments for confining the scope of investigations to these scopes are discussed in Chapter 4.

### 1.4 METHODOLOGY

The general methodology employed in this thesis is a pragmatic combination of literature review, analysis and experimentation. The literature, which encompasses a wide-ranging area,

(including psychology, perception and human factors) provides for the understanding and predicting performance in the VE. In addition to the related literature in VE, research findings from the real world studies in the related areas can also be used as a baseline comparison for judging spatial task performance in the VE.

The research approach presented in this thesis extends the previous investigations by empirically examining factors affecting spatial awareness in VE in comparison to its real counterpart. Three sets of experiments which examine factors affecting a user's spatial awareness in three different forms of VE presentations (static, dynamic, and interactive images) have been undertaken. The research approach and methods are further described in Chapter 4.

## 1.5 PUBLICATIONS

Some of the works from this thesis have been published and presented at international conferences under the following titles (see Appendix D):

1. Awang Rambli, D. R. and R. S. Kalawsky (2003). Effects of varying display size on user's asymmetrical distance perception in the real and virtual environment. Virtual Concept 2003, Biarritz-France.
2. Awang Rambli, D. R. and R. S. Kalawsky. (2002). The Effect of Display and Image Type on Inter-Object Distance Estimation in Virtual and Real Environment. SIGCHI-NZ Symposium on Computer-Human Interaction, Hamilton, New Zealand.

## 1.6 THESIS OVERVIEW

The chapters in this thesis are presented in three major parts:

- I. Thesis Background
  - a. Chapter 1 Introduction
  - b. Chapter 2 Perceptual Issues in VE
  - c. Chapter 3 Technological Issues in VE
  - d. Chapter 4 Basis for the experimental approach for understanding spatial awareness
- II. Experimental approach, results and analysis
  - a. Chapter 5 Experiment on distance perception in static images
  - b. Chapter 6 Experiment on distance perception in dynamic images
  - c. Chapter 7 Experiment on distance and spatial memory task in interactive images
- III. Implications Drawn From The Literature And Experiments Undertaken
  - a. Chapter 8 Overall analysis of results
  - b. Chapter 9 Final Conclusions, Research Contributions Implications on spatial awareness in VE,
  - c. Chapter 10 Recommendation and Future works

The logical presentation of these chapters in this thesis is illustrated in Figure 1-1.

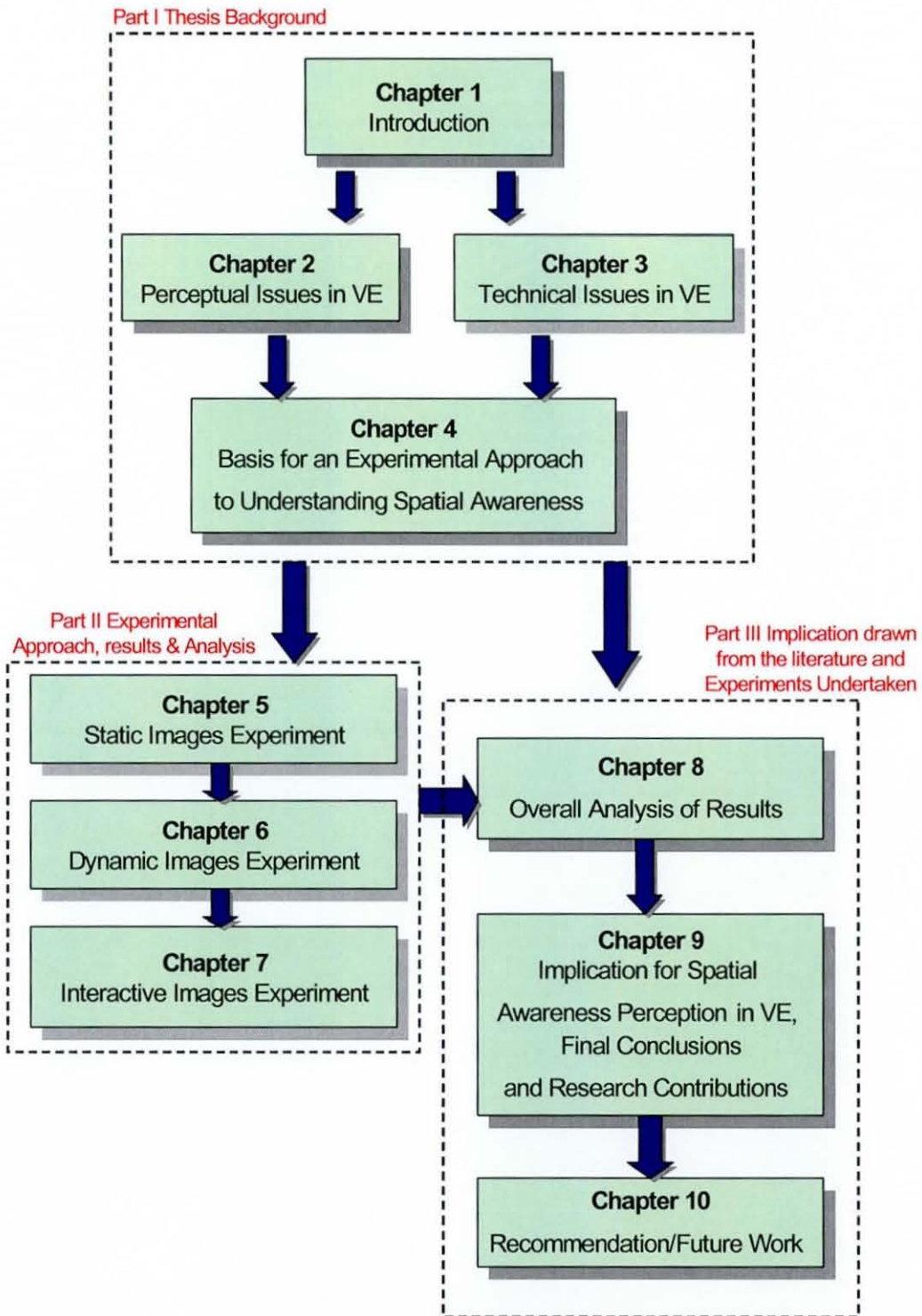


Figure 1-1 Logical structure of thesis



A brief overview of each of these chapters is presented in the following paragraphs.

### **Part I Thesis Background**

Chapter 1 provides a brief introduction to the background of the research explored in this thesis which includes the definition of VE, statements of the research problem, aims, questions, scope and methodology.

In Chapter 2, the perceptual issues relevant to the perception of 3D space in a VE are presented. A brief overview of the human visual system is introduced followed by a discussion on how we perceive space in the real world, the cues for perception and the related literature. The perception of space in pictures and the related literature is reviewed next and finally perception in the VE with emphasis on spatial awareness is presented. This covers a review of studies related to distance perception and spatial representation and related literature on factors affecting spatial awareness in VE.

The technological issues related to VE are discussed in Chapter 3. First, the fundamental concept and issues in the modelling and rendering of the VE are mentioned. This includes the techniques and software algorithms used to generate visual realism in real-time VE. Discussions on trade-offs between image realism and system performance follow next. The types of VE systems and the technological limitations and advantages of each system type are then presented. The VE systems are also compared in terms of qualitative performance.

Chapter 4 draws upon the issues highlighted in Chapter 2 and 3 and from the literature. It outlines the basis for the experimental works and methods for the research presented in this thesis. Discussions and arguments for the basis for the experimental works are first given. This includes highlighting the overall research aims, questions, scope and assumptions. The general research methodology employed and the arguments for the specific choice of research methods used to address the research questions in this thesis are provided next which involves presenting the methods employed for data collection and data analysis.

### **Part II Experimental Approach, Results and Analysis**

The experimental methods and the results of the three sets of experiments that were undertaken are described in the next three chapters. Chapter 5 outlines the experimental

methodology and the results of the first set of studies (Experiment 1A and 1B) which compare participants' spatial awareness in static images of real and VE. The first study examines the effect of image type and display type while the second study analyses the effect of display size, viewing distance and physiological cues. Discussion of results from both studies and conclusions drawn concludes the chapter.

In Chapter 6, the experimental methodology and the results of two studies (Experiment 2A and 2B) investigating user's spatial awareness in dynamic images are mentioned. The first study (Experiment 2A) investigates the effect of image type and display size while the second study (Experiment 2B) explores the effect of display size, viewing distance (that is physiological cues) and image resolution. Discussion of the results and conclusions drawn are presented at the end of the chapter.

Chapter 7 outlines the experimental methodology and the results of Experiment 3A and 3B which examine users' spatial awareness in interactive real and VE. The first study (Experiment 3A) investigates the effect of environment types (real and virtual), display size, viewing distance and physiological cues on distance estimate and spatial memory tasks. Additionally for spatial memory task, the effect of interface devices and travel modes was also examined. The second study (Experiment 3B) was undertaken and reported following the results of Experiment 3A. Finally, the results and conclusions from both studies are discussed.

### **Part III Implications drawn from the literature and experiment taken**

An overall analysis of the three experiments on static, dynamic and interactive images from Chapters 5, 6 and 7 is presented in Chapter 8. Prior to the presentation of the overall analysis of the results of Experiment 1, 2 and 3, a summary of the main findings from each experiment is given.

Chapter 9 recapitulates the overall conclusions and research contributions and provides a discussion on the implications of the experimental results on spatial awareness perception in a VE. The first section includes discussions on the major findings and contributions from the research and the impact of image modelling on the conclusions drawn. These results are considered with respects to the key research questions being proposed. The method contributions concerning the approach to investigate the display related factors examined in this thesis are also highlighted. The second section presented a discussion on the impact of these implications on VE-related applications.

Finally, Chapter 10 outlines several recommendations and directions for future works based on the research work conducted in this thesis. The recommendations include the proposition for new methods and improvement of the methods employed by the research in this thesis, which would clarify, enhance and provide support for some of the findings from the research presented in this thesis. Several potential areas highlighted for further research work include suggestions for the investigations of other aspects and factors that are related to and could affect a user's spatial awareness in VE.

# CHAPTER 2

---

## *PERCEPTUAL ISSUES OF VE*

### 2 OVERVIEW

In terms of perception there are several issues that directly influence the design of VEs: visual perception, auditory perception, haptic and kinaesthetic perception. Due to time constraints, it is not within the scope of this thesis to investigate all these issues. This thesis is concerned with evaluation of perceptual space and as such the scope is limited to the evaluation of the visual perception, though other related sensory experience (such as kinaesthetic perception) will also be discussed.

In the real world, we derive the perceptual understanding of the 3D space from different sources of information (Cutting and Vishton 1995). As such an understanding of perception in the real world forms the basis to comprehend perception in the VE.

In this chapter, we introduce the perceptual issues relevant to the perception of 3D space in a VE. Before presenting the perceptual issues in a VE, a brief overview of the human visual system is presented. This is followed by a discussion on how we perceive space in the real world, the cues for perception and the related literature. Next, the perception of space in pictures and the related literature is reviewed. Lastly, perception in a VE with emphasis on

spatial awareness, that is our ability to perceive objects within 3D environments, is presented. Studies related to distance perception and spatial representation, the two basic spatial tasks considered in this thesis, are also presented followed by a review of the related literature on factors affecting spatial awareness or perception in VE.

## 2.1 THE HUMAN VISUAL SYSTEMS (HVS)

The ability of a person or an animal to detect fine spatial patterns and resolve details is termed *visual acuity* (Bruce, Green et al.1997). At any instant the human eye samples a relatively large segment of the optic array (the peripheral field) with low acuity, and a much smaller segment (the central, or fovea) with high acuity. Thus, visual acuity is optimal for objects presented at the centre of the visual field compared to those at the peripheral view. However, smooth and saccadic eye movements shift the high-acuity segment about rapidly so that acute vision over a wide range of angle is achieved. Saccadic eye movements refer to the sudden, intermittent changes of the eye position to focus on an object. The visual acuity performance decreases with increases in distance of the viewed objects from the viewer. However, there is also a limit to focus on nearby objects where objects closer than this point are blurred and resolution is reduced. The HVS, which is sensitive to a broad range of ambient illumination, contains two types of photoreceptors (rods and cones). These photoreceptors vary significantly in sensitivities. Visual acuity increases with increase in luminance but contrast sensitivity decreases with luminance increase (May and Badcock 2002).

The human FOV spans an area of  $120^{\circ}$  vertically and  $150^{\circ}$  horizontally (Kalawsky 1993). This area could be further increased with eye movement and head movement, giving the maximum FOV for an individual using both eyes is approximately  $200^{\circ}$  (Barbour and Meyer 1992). The overlapped regions resulting from the two monocular FOV from both eyes is termed the binocular visual field. Stereoscopic vision occurs in this region and this is further discussed in section 2.5 of this chapter. Figure 2-1 illustrates the monocular visual fields and binocular visual fields.

Visual angle is usually used to indicate the dimension of objects (Kalawsky 1993). This angle is the visual angle subtended at the eye with respect to the viewed object. The value of this angle is inversely related to the distance of the object from the viewer; the farther the object is the smaller is the visual angle.

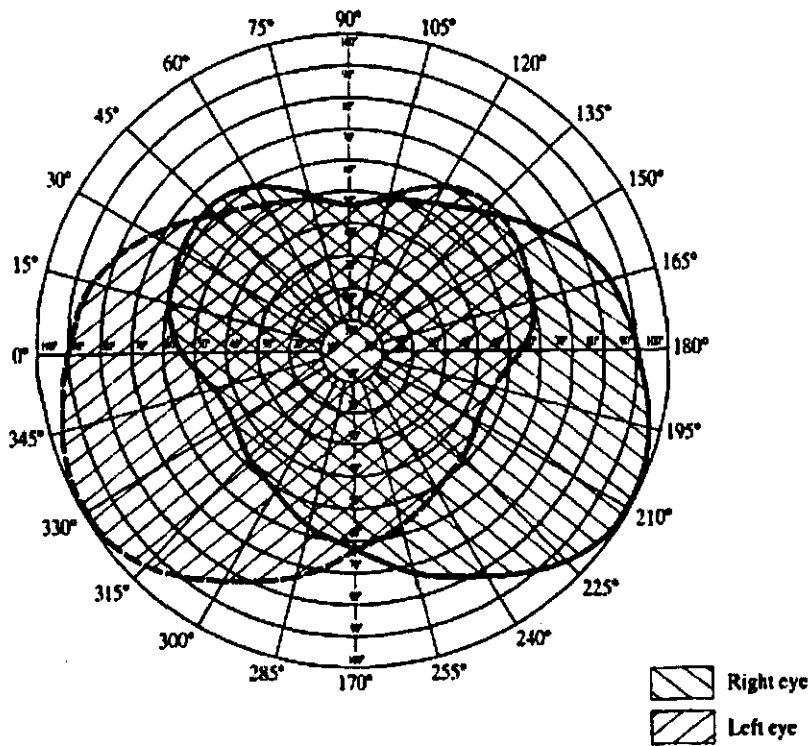


Figure 2-1 Human FOV. Adapted from Kalawsky (1993)

2.2 PERCEPTION OF SPACE

The ability to perceive 3D space is very important for our survival as it allows us to interact safely and effectively with the environment (Sekuler and Blake 1994). More specifically it guides our behaviour in the environment (Wade and Swanston 1999). In order to interact with an object within the 3D space we need to know where it is located and its shape. When driving a car we constantly judge the distance of our car from another car in front of us. Similarly, crossing a street or reaching for objects requires us to make similar judgments. In fact, most of our daily tasks depend on the accuracy of such judgements. Our ability to know where objects are located in space, that is how far objects are from us, is referred to as depth perception (Sekuler and Blake 1994).

There are two aspects of depth or space perception: the estimate of distance from self to objects and the estimate of distance between objects (Coren, Ward, et al.1999). The former is often referred to as egocentric distance perception (absolute distance) and the latter is referred to as exocentric distance perception (relative distance). Studies have shown our ability to make relative distance judgment is more accurate than on absolute distance (Sekuler and Blake 1994).

Based on the depth cue theory, a main theory of depth perception, our perceptual understanding of space in the natural environment is derived mainly through the use of various sources of information and the images created on the retina (Goldstein 1996). An alternative theory to visual perception is called the “Ecological” approach developed over a 35-year period by J.J.Gibson (Bruce, Greene et al.1997). The depth cue theory states that the visual system computes the distances of objects in the environment based on the information from the posture of the eyes and pattern of light projected onto the retinas by the environment (Wanger, Ferweda et al.1992). However, in the “Ecological” approach Gibson argues that the light reflected from surfaces and objects possesses structure which gives information about the spatial characteristics of the visual world; that is the information carried by the reflected light is responsible for the perception of the visual world.

In this thesis, the depth cue theory is the main theory used but we also acknowledge the importance of the “Ecological theory” where appropriate. Despite the differences on how information from the light results in visual perception, both theories agreed that there exist some sources of information about the 3D layout of the space. This information is sometimes referred to as cues to depth (Sekuler and Blake 1994). Such cues which can be categorized as pictorial cues, physiological cues and binocular cues (Gillam 1995; cited in Pfautz 2000) are presented in the next section.

## 2.3 PICTORIAL DEPTH CUES

Pictorial depth cues are those cues that are found in pictures to give the impression of three dimensionality. They are also called monocular depth cues because they can be viewed with either eye alone. Some of these cues have been used by artists since the Renaissance period to create an impression of 3D space on a 2-D display. Monocular depth cues are also richly available from our surrounding environment. They not only allow us to perceive the spatial layout of our three-dimensional world but also assist us to perform visually guided skilled tasks (Schiffman 2000). Some of these pictorial cues are available when the observers and the viewed scene are motionless and some are available when there is movement in the observers or the viewed scene or both (Schiffman 2000). The former is referred to as static cues to depth (Figure 2-2) and the latter is referred to as motion cues to depth.

### 2.3.1 Static cues to depth

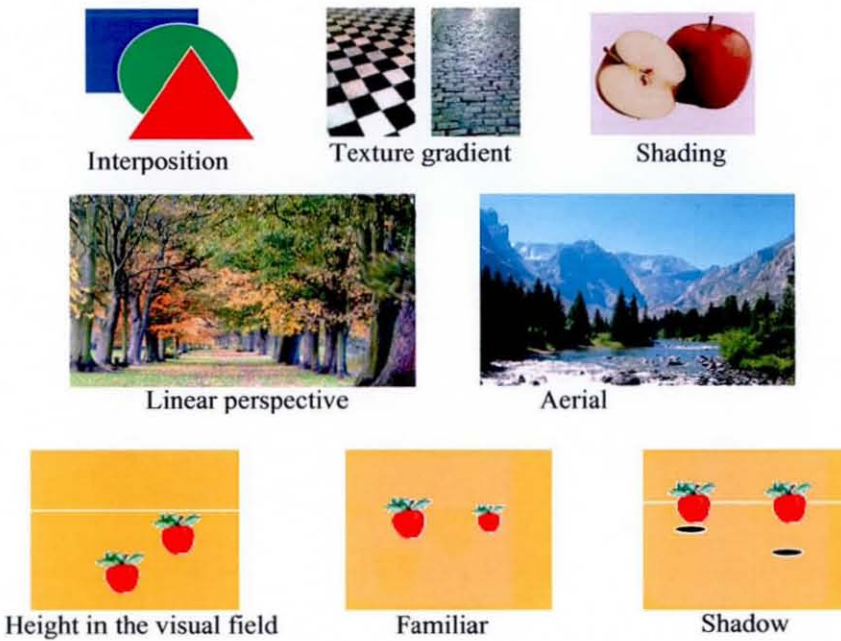


Figure 2-2 Static cues to depth

#### 2.3.1.1 Interposition

This cue refers to the hiding of part of a farther object by a nearer object. It is often called occlusion. This cue is an effective cue for determining relative depth between objects. It only indicates whether one object is farther or closer to the observer. No information on the actual distance of the objects is provided. The effectiveness of this cue does not decrease with increasing distance of the object from the observer (Cutting and Vishton 1995).

#### 2.3.1.2 Size

When two objects of the same size are located at different distances, we often judge the smaller one to represent the farther object. The size of the image on the retina depends on the distance of the objects from the observer. The farther the object from the observer, the smaller the retinal image size becomes. However, this cue depends on the familiarity with the object size; otherwise retinal image will provide no information about the object's distance. When too few cues are available, viewers may rely on the familiar size of objects to judge the object's distance (Schiffman 1994).

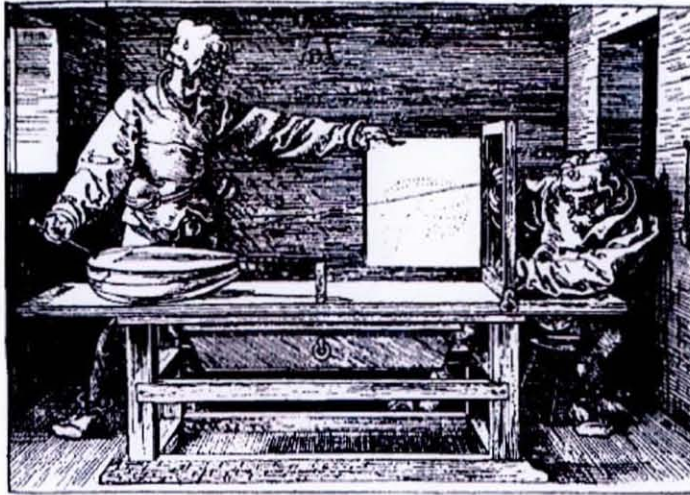
#### 2.3.1.3 Perspective

Perspective cues are based on the geometrical relationship. The size of the retinal image is inversely proportional to the distance of the object from the observer: that is, the farther the



object from the observer the smaller the image on the observer's retina and vice versa. Perspective cues are used by artists who want to realistically portray a 3D scene on a 2-D flat surface such as paintings and drawings. Examples of perspective cues are linear perspective, aerial perspective, shading, elevation and texture gradient.

- **Linear perspective** refers to apparent convergence of parallel lines as they recede toward the horizon. A good illustration is a railway track or a road, it seems to narrow at the farther distance when actually it is still the same size as the near one. This narrowing actually provides a sense of depth to the observer. This technique was successfully used by Albrecht Durer to portray a 3D scene on a 2D flat wood piece (Figure 2-3). It has been suggested that the depth cue implied by the linear perspective cue can be strong enough to contradict the depth information portrayed by retinal disparity (Steven and Brooks 1988; cited in Sekular and Blake 1994).



**Figure 2-3** Woodcut by Albrecht Durer's 1525 illustrating perspective. ©Bettmann/CORBIS.  
Adapted from Sekular and Blake (2002), pp306

- **Aerial perspective** or atmospheric perspective effects allow us to view closer objects as clearer compared to distant objects. This is because to view distant objects, we have to view through the air that contains small particles such as dust and moisture, thus making distant objects appearing to be dimmed and blurred. This cue provides an effective cue to relative distance. Its effectiveness increases with distances but at larger distance objects becomes less discernible (Cutting & Vishton 1995). Artists usually employ this technique by portraying distant object as blurred and less clear than nearer object.
- **Shading** refers to the viewing of a shaded two-dimensional image as three-dimensional due to the effect of lighting. The surface which faces the light source, will have the greatest illumination (that is brighter), and this illumination will decrease as the surface is further away from the light source. Thus, shading gives an object its solid look as well as depth information. The presence of an object's shadow has been shown to aid

participants' in their distance estimate performance (Wanger 1992). Other researchers have empirically shown that shadows are significant cues for certain performance tasks (Hu, Gooch et al. 2002, Hubona, Wheeler et al.1999). In their review of studies, Sekuler and Blake (2002) found that perceived depth varies depending on the position of the shadow relative to the object casting the shadow.

- **Elevation or height in the visual field.** Objects (B and C) that are located closer to the horizon are perceived as further compared to objects (A and D) located distant from the horizon (see Figure 2-4). Thus, B seems farther away than A because the base of B seems closer to the horizon. Similarly, C appears farther away than D because it is closer to the horizon. This cue can be used for the perception of relative distance and absolute distance. However, unlike occlusion its effectiveness decreases with the increasing distance of the object from the observer and at 2m it is nearly as effective as occlusion (Cutting and Vishton 1995)

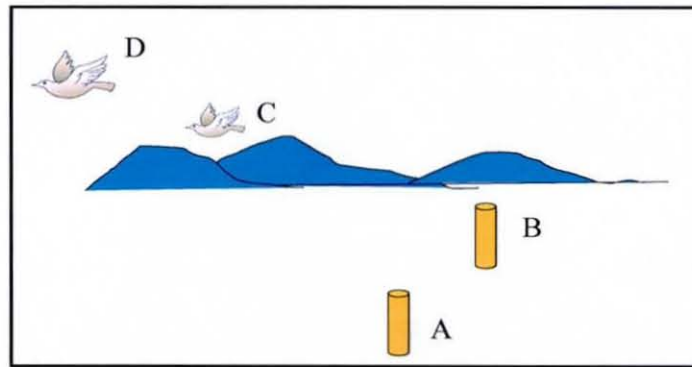


Figure 2-4 Height in the visual field

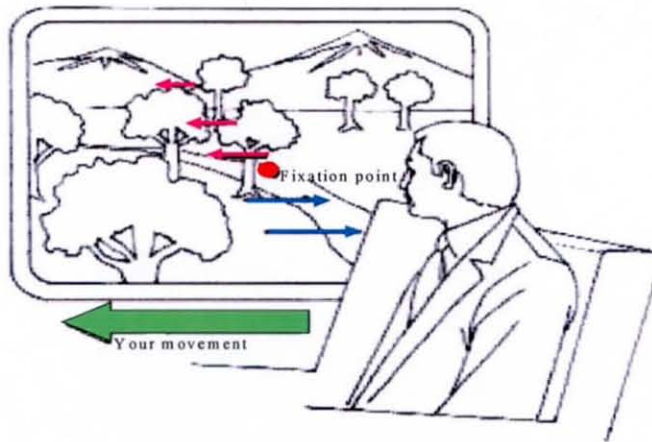
- **Texture gradient** refers to the changes of the size and the spacing of the elements comprising the texture of the surface as a function of distance. When the distance gets larger, the sizes of the elements appear to reduce in size and the spacing of the elements appears to be closer. According to Gibson (1950) (cited in Sekuler and Blake 1994), texture gradient provides precise and unambiguous information about distances and slant surfaces, including the size of the objects located on those surfaces. In the VE, some empirical evidence has suggested that texture is a weak cue to distance (Witmer and Kline 1998). Other researchers, however, found significant effect of texture on distance judgment using perceptual matching tasks (Sinai, Krebs et al.1999). They found that medium density texture yields very accurate results. Another study however revealed that a rich, fine resolution texture pattern yields the most accurate result (Kline and Witmer 1996). Differences in experimental methods may contribute to these differences in the results of these studies. James and Caird's (1995) study showed that participants tend to overestimate distance in a textured VE and underestimate distance to target in a polygonal VE. The shape of the texture may also determine perception of distance. An elongated,



regularly spaced element in a consistent orientation has been proposed as the best texture for determining depth (Carr and England 1993).

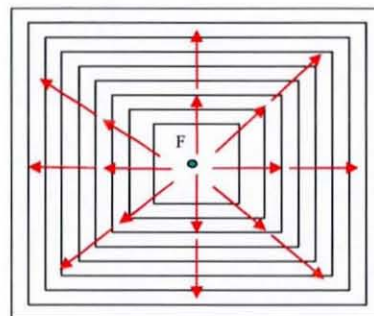
### 2.3.2 Cues from motion

The presence of motion in either the viewer or the objects in the environment allows for more cues for the perception of depth and distance. Motion parallax and motion perspective are examples of motion cues (Witmer and Kline 1998). Motion parallax refers to the apparent relative motion of objects in the visual field when there is movement either in the viewed scene or the observer. Movement is not restricted to moving the entire body; a simple head movement would produce the same effect. When an observer moves their head laterally, near objects seem to pass by quickly in the opposite direction of their movement and farther objects appear to pass by more slowly in the same direction as they are moving (Figure 2-5).



**Figure 2-5 Motion parallax Adapted from Coren and Ward et al. (1999), pp264**

This apparent difference in movement speed and direction of objects provides a very effective cue for perception of depth and distance (Schiffman 1990). It has been noted that relative depth judgment based on motion parallax are almost as accurate as binocular disparity (Graham 1965, cited in Sekuler and Blake 1994).



**Figure 2-6 Optical flow pattern Adapted from Schiffman (1990), pp320**

When an observer moves towards a surface (or away from it), a pattern of continuous changes called an optical flow pattern is created (Figure 2-6). The information, also known as motion

perspective, provides the viewer with a reliable source of relative velocity and direction of movement. It also provides information on the relative distance of objects from the moving observer. Movement of the observer through space additionally provides information about the topography and the layout of the environment (Bloomer 1990). As such, movement through space is necessary in order to form a mental representation of the space.

A study conducted by McCandless, Ellis et al. (1999) revealed a significant effect of motion parallax cues on a virtual object localization tasks. They reported that motion parallax induced by participant's head-movement is more influential than accommodation cues. A similar result was also obtained by Ferris (1972) who compared fixed head to head movement on a distance estimation task found that motion parallax can be useful for absolute distance estimation. Other researchers have found that it is possible to train participants to make accurate absolute distance estimates based on motion parallax cues (Dees 1966). Similarly, it has been indicated that motion parallax cues are notably salient for spatial tasks such as positional and rotational tasks (Morar 2002).

During motion the changes in the shapes and forms of objects and changes in spatial relationship among objects or between self and environment are perceived (Bloomer 1990). These changes register displacement of images on the retina. The rate of displacement that takes place will indicate whether we perceive motion or not. If displacement occurs too fast or too slow, we will not perceive any motion or movement. A plant growing and bullet trajectory are examples of the former and the latter case. Perception can be either real (there is actual movement) and apparent (appear to move but actually there is no movement). The latter is experienced in motion pictures, computer animations, or VE. Movies, television, video games and computer animations often employ motion cues to create a realistic sense of three dimensional spaces.

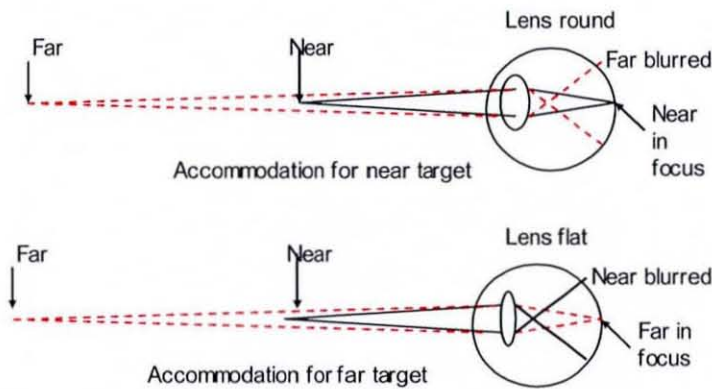
It was demonstrated that people can use optic flow to estimate distance provided scaling information is available (Redlick, Jenkin et al. 2001). Their findings suggest an impoverished VE (few details) might contribute to the overestimation of distance and we can rely upon optic flow for navigation when strong visual cues are available.

The absence or inaccurate simulation of motion cues such as optic flow pattern in a VE may lead users to perceive the motion as unnatural because the users are aware of the experience (Stanney and Mourant et al. 1998). The generation of realistic feelings of self-motion in the VE would contribute to the overall sense of presence in the virtual space (Hettinger 2002). As such an accurate depiction motion cues would improve realism in perception of VE.



## 2.4 PHYSIOLOGICAL DEPTH CUES

These cues are derived from the muscular responses and adjustment of the eyes in order to bring objects in view into clear focus on the retina. When we look at an object, our eyes will focus and converge so that the image projected on the retina is sharp. The amount of focus and convergence depends on how far the objects are from us. There are two types of physiological cues: accommodation and vergence cues. Accommodation is the process whereby the eye adjusts the lens shape to focus on near and far objects for a sharp image on the retina (Figure 2-7).



**Figure 2-7 Accommodation of an image by adjusting the shape of the lens.**

**Adapted from Coren, Ward et al.1999, pp55**

Relaxed accommodation occurs when the lens is flattened in order to clearly focus distant objects on the retina. Conversely, the lens thickens for nearer objects. The degree of contraction of eye muscle for the accommodation, first processed by the brain, gives us information or cues on how far a given object is (Schiffman 1990). Viewing a blurred image indicates that the object is not focused correctly. This however may be used as a cue for relative distance. Mon-Williams and Tresilian (2000) conducted a study on how much blur driven accommodation can provide information on target distance in the absence of any retinal cues to distance. The study results indicated that accommodation can act as a source of ordinal depth information in the absence of other cues but its role is questionable in full-cue condition. In a review of related literature, Howards and Rogers (2002) reported that earlier evidence suggested that people cannot judge distance based on accommodation but recent studies indicate that people can judge absolute distance up to a certain extent. In a recent study comparing actual versus virtual environment in a reaching task, Bingham and Bradley (2001) found that egocentric distance was overestimated. The authors suggested that in VE accommodation is beyond reach, thus when they reduce the focal distance in the VE using 2-diopter glasses, overestimation is reduced by half.

Another method the HVS uses to bring a viewed object into focus on the retina is through the movement of inward and outward movement of the two eyes. These eye movements are referred to as vergence eye movement. Convergence is when the eyes move inwards towards the nose to focus on near objects in front of us while divergence is when the eyes move outwards to focus on objects farther away. Similar to accommodation, information from the muscular contraction as a result of vergence movement can be used to determine distances of objects (Coren, Ward et al. 1999). In his studies on subject convergence response to monocularly viewed objects, Predebon (1994) found that convergence was influenced by the implied distance from the familiar size but not from the implied distance of suggested size. It is generally accepted that the judgment of distance is based to some extent upon the physiological process of accommodation and vergence (Swenson 1932). He further found that accommodation cues comprise only one-third of the effectiveness of convergence. However, distances investigated were limited to 25 and 30 cm.

Accommodation and vergence cues work in concert with one another, thus a change in one will result in a change in the other. Both cues are limited in their effectiveness as depth cues; as such they are useful for nearby objects (Sekuler and Blake 1994). For accommodation its effectiveness is up to 2m (Schiffman 1990) and for vergence cues it is useful for a distance of up to 6m, beyond this would reflect only small vergence changes (Howard and Rogers 2002). Early empirical evidence showed that accommodation is a determining factor in monocular vision while convergence is in binocular vision (Baird 1903).

Smith and Smith (1961) suggested that a monocularly viewed picture would permit perception of absolute distances that are independent of the accommodation and vergence cues and these cues could only carry information about optical distances of the photographs and not the portrayed distances. However, empirical evidence revealed that our eyes do converge and accommodate at various distances in responses to the pictorial depth cues found pictures, paintings and line drawings (Enright 1987a, Enright 1987b; cited in Coren and Ward et al.1999).

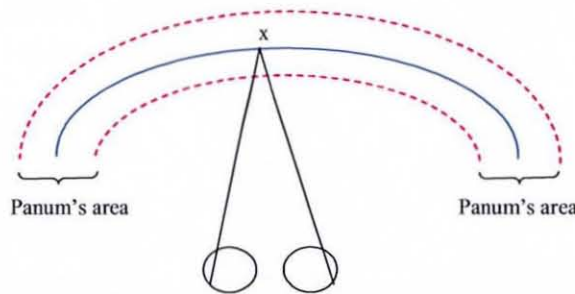
It has been suggested that the relationship between accommodation and vergence cues in stereo display might cause visual fatigue in the viewer (Takeda and Hashimoto et al.1999; Howard and Costello 1996). In stereo display the eyes accommodate at the plane of the display but may converge at difference distances. The conflict between these cues may result in visual fatigue (see Section 4.1.2.5 of Chapter 4).



## 2.5 BINOCULAR DEPTH PERCEPTION

Binocular depth cues are cues that are based on two eyes. The binocular cues of accommodation and vergence cues were discussed in prior section. In this section, the stereoscopic cues for binocular depth perception which are based on the fact that we have two horizontal separated eyes with overlapping views of the world are described. Each eye perceives a slightly different image of the world. This difference is referred to as binocular disparity or binocular parallax, which results in a unique appearance of depth with solidity called stereopsis. We are unaware of this difference because our brain combines information from both eyes yielding a single image through a process called fusion. Our visual system utilizes this information to accurately perceive depth between objects (Schiffman 1990).

We perceive a single image of the two different images from both eyes when the sets of spatial locations in space for a given a degree of convergence are projected to corresponding retinal points of the two eyes. The locus of all these spatial points in space is termed horopter; when the image is located in front or behind this horopter, double images are perceived (Schiffman 1990). However, under normal conditions we do not see double images because the visual system suppresses it. Panum's area is an additional region on either side of horopter. Any image located within this space will still be perceived as a single image (Figure 2-8).

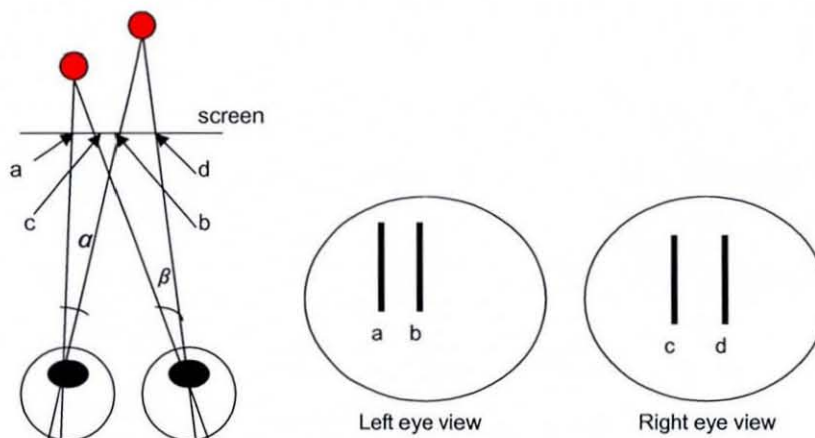


**Figure 2-8 A version of horopter. Points of X on the horopter will fall in the corresponding retinal points of the two eyes yielding a single image. Other points outside the Panum' area will yield double image. Adapted from Schiffman (1990), pp356.**

### 2.5.1 Stereo vision -illustration

In stereo viewing, when the eyes is fixated at vertical line a, a second line b appear closer to a in the right eye's than in the left eye's image (as shown in Figure 2-9). This discrepancy is resolves by the perceiving the lines as being perceived at different depths as shown. Retinal disparity refers to the difference between the angular separation of line **a** and **b** in the two eyes, that is disparity is equal to  $\alpha$  minus  $\beta$ . The closer the object is, the greater is the disparity of the images on the retinas.

For a discussion on stereo vision from a human perspective, readers are referred to Patterson (1992).



**Figure 2-9 Some basic geometry for stereoscopic viewing**

### 2.5.1.1 Cue conflicts

A conflict of cues occurs when the disparity information from stereo viewing causes an object to appear in front of the display. This is because information from the edge of the screen (occlusion cue) may appear to occlude the object, thus contradicting the disparity information provided by the stereo cues. Since occlusion cue is the stronger depth cue it dominates over stereo cues thus eliminating the illusion of depth (Ware 1995).

A cue conflict can also arise in a dynamic environment. Moving through an environment causes the disparities to change dynamically which in turn causes changes in the relative depths in the scene. However, since motion parallax cue has been shown to be a more important cue to 3D space perception compared to stereopsis (Arthur, Booth et al. 1993; Cutting 1986; Ware 1995), the influence of the cues from the changing disparities would be less effective. Thus, these evidences suggested that stereo cue is less effective when the other cues are more dominant (see section 2.5.2).

In the research described in this thesis, the image is presented to the user in a non-immersive and non-stereo mode presentation (see Chapter 4) as such the conflict of cues described in the prior paragraphs is not relevant. The images are presented to the users non-stereoscopically as opposed to stereoscopically. Similar to natural viewing, viewers used both eyes to view the images. The viewers however can still perceive depth in the images but at the same time they are aware of the flat screen. This occurs as a result of the perceptual conflict between the monocular and binocular cues whereby the monocular cues indicate depth but the binocular



cues indicate flatness. The amount retinal disparity is the same for all objects in the images; thus controlling for these cues in our study. Thus, the results of the experiments and the conclusions drawn from the research presented in this thesis are only valid for monoscopic vision only.

### 2.5.2 The importance of stereo cues

The stereo cues are effective only for objects less than 25 meters away and it is optimal for nearer objects. According to Kalawsky (2003), the effectiveness of stereoscopic cues is up to 9.2m for peripheral viewing and up to 500m in the fovea.

However, not everyone can use stereopsis for perceiving depth. It has been estimated that about 5-10 % of the human population were not able to perceive depth from this cue (Sekular and Blake 1994). Some people with the presence of stereo cues alone without the presence of monocular cues, found it difficult to perceive depth (Barbour and Meyer 1992). This indicates the importance of monocular cues for perception. Thus the proper rendering and emphasis of monocular cues in images might help overcome the absence of stereo cues.

Further discussions on the importance and the drawbacks from stereo presentations of image are presented in Chapter 4 (Section 4.1.2.5), which consequently argues for the use of non-stereo image presentation in the research presented in this thesis.

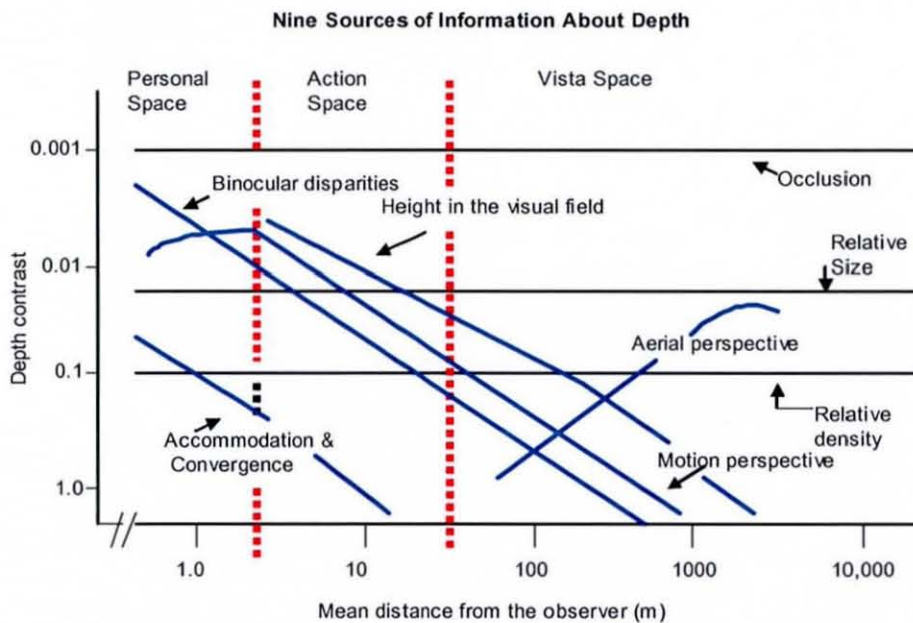
## 2.6 COMBINATION AND RELATIVE IMPORTANCE OF DEPTH CUES

Many of the depth cues described earlier are combined in a complex way by the HVS to give the impression of three dimensionality of the space. According to Pfautz (2000), generally, the more cues are presented the better is the sense of depth.

These cues vary in effectiveness depending on the distance of the viewer from the objects to be estimated (Cutting and Vishton 1995). Cutting (1997) groups these cues based on their relative utility into three regions of space: personal, action and vista space (Figure 2-9). Personal space is within 1.5m from the observer, action space is up to 30 m and vista space is beyond 30m.

With respect to effectiveness, Cutting and Vishton (1995) indicate that some cues' effectiveness is unaffected by distance, some cues' effectiveness decreases with distance and some cues effectiveness increases with distance. He ranks occlusion as the most effective across all viewing space followed by relative size. Occlusion, relative size and relative density

effectiveness is consistent across distances where occlusion will always dominate size and size will always dominate density. Some cues such as height in the visual field, motion perspective, the binocular disparities, accommodation and convergence cues decrease in effectiveness with increasing distance of the viewer from the viewed objects. Thus, these cues accuracy is dependent on the distance of the viewer from the viewed objects. For example, stereo cues are primarily used when threading a needle; to further increase the accuracy of stereo and physiological cues both the thread and needle are brought closer to the viewer (Pfautz 2000). Thus, when viewing near objects, the relative importance of cues such as linear perspective, relative brightness and size, height in the visual plane should diminish and the importance of physiological cues increases (Kline and Witmer 1996). For the viewing of objects at larger distances, the pictorial cues may be employed by the viewer. According to Cutting and Vishton (1995), aerial perspective is the only source of information that increases in effectiveness with distance.



**Figure 2-10 Categorization of sources of depth information into three types of spaces based on distance from the observer. The more potent cues are reflected by the small values of the depth contrast. (Adapted from Cutting (1997)).**

Most of the pictorial cues above can be accurately represented in the computer-generated images (Witmer and Kline 1998), giving these images a sense of depth and three-dimensionality. The use of more or redundant pictorial cues and depth information would yield a more realistic and compelling sense of 3D space. Kunnapas (1968) concluded from his study that increasing the number of cues increases the accuracy of distance judgment. However, utilizing a lot of pictorial cues would be computationally expensive. It would involve more processing overhead to calculate the shading, lighting and colour of complex



scenes. Because of this many real-time applications settle for low realism in images. This will inadvertently reduce computational complexities and rendering time but the lower level of realism would sometimes mean that the depth cues are less accurately represented. Thus, the choice of which cues to include becomes important design decisions so that only effective cues are included and less effective ones eliminated.

While it is not the aim of this thesis to provide a comparison among all these depths cues, understanding the influence of these cues on perception of space is important towards understanding the effect of other factors. Moreover, several researchers have investigated the relative effectiveness of several depth cues in computer generated images (Surdick, Davis et al.1997, Wanger, Ferweda et al.1992). Surdick and colleagues (1994) reported that perspective cues (linear perspective, foreshortening and texture gradient) were more effective than other depth cues such as relative brightness, relative size and relative height. Additionally, they conclude that the use of perspective cues in a simulated display to be more important than other depth cues because they are not only effective and accurate but they are easily perceived by participants and easily incorporated in less complex displays (bi-ocular display as opposed to binocular or stereo display).

In their investigation of perceived spatial relations in computer-generated images, Wanger and colleagues (1992) examined the influence of several pictorial cues on participants' accuracy in a position, orientation and size matching tests. The pictorial cues investigated included projection, shadow, object and ground texture, motion and elevation cues. They found that on positional accuracy, shadow had a dominant effect over other cues. Motion, object texture and ground texture however did not affect positional accuracy. On orienting tasks, perspective cues were shown to have a dominant effect over other cues. Motion effect is better than shadow but textures (object and ground) and elevation cues do have a significant effect. For scaling tasks, shadow is the most effective followed by motion cues, elevation, and perspective with texture cues being the least effective. Their results showed that the effectiveness of these cues is task dependent. Other researchers (Hubona, Wheeler et al.1999) indicated that the presence of shadows (object shadow) enhances positioning performance but not resizing performance. They further indicated that stereo cues are more effective at enhancing performance than shadows. Motion parallax cues have been shown to be more effective than accommodation cues (McCandless and Ellis et al.1999).

While the presence of these cues provides information for the perception of 3D space in the real and VE, various other factors may influence the accuracy of such perceptions. These factors are discussed next in Section 2.7.

## 2.7 OTHER FACTORS THAT AFFECT DEPTH PERCEPTION

The presence of cues in the real world and pictures or computer generated images, discussed in the previous section, allows for the perception of 3D space. However, there are several factors that may cause errors in perceptual judgment of distance and space. These factors which include size constancy, prior knowledge, cognitive dissonance and effects of sensory conflicts are discussed in the following subsections. The last factor is related to perception of 3D space in dynamic images such as video, computer animation and VE.

### 2.7.1 Size constancy

The size of object image casts on our retina varies with varying objects' distance from us. Figure 2-10 shows that the retinal image size is inversely related to the object's distance from the observer. This means that if an object ( $s_2$ ) is twice as far; the image size will be reduced by half. However, these changes usually are not realized by the observer under normal viewing conditions. When we look at a familiar object located at a distance, we find that its size tends to remain the same even if that object is twice as far away from us. This is called size constancy. Thus, in normal viewing conditions, the perceived size of an image does not depend entirely on the retinal image size. Perceived size can be independent of retinal image for a considerable amount of distance (Schiffman 1990).

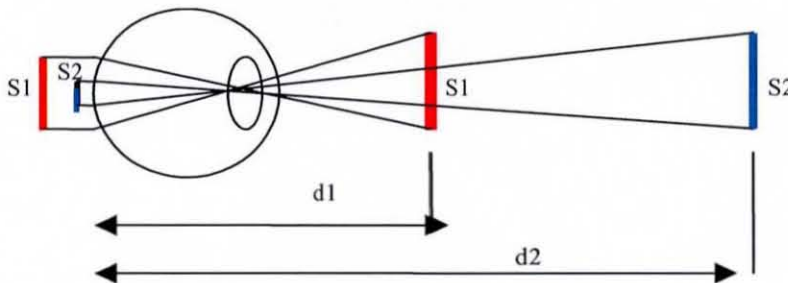


Figure 2-11 Size constancy

This size constancy phenomenon can also be explained in terms of visual angle. The visual angle (angular size) and the retinal image size is influenced by the distance of the object from the viewer: larger object distance will result in a smaller visual angle and smaller retinal image, while a nearer object will result in a larger visual angle and a larger retinal image size. However, visual angle (or the retinal image size) has been regarded as a weak cue because it easily overridden by other cues (Beall, Loomis et al.1995). This might explain the size-constancy experience in our perception of size.



It has been suggested that size-constancy may fail in photographs and drawings due to the perception of picture lacking depth and the fact that perception of the picture surface cannot be wholly eliminated may further reduce this effect (Boring 1964). However, we can get perfect size-constancy on television because size-constancy can be given perspective cues motives of distance on object location by showing the focus of near object and the blurring of far objects.

Bloomer (1990) asserted that “the context in which you see an object is most important influence on your perception of object and its size.....as long as the relationships among object is realistic we will not be conscious of it miniaturization”. As such image realism and the type of scene and the objects within it do play a role in how we perceive the objects.

Even though size constancy is useful for providing stable perception of the world, this is not always the case; under certain circumstances size-constancy might give us perceptual error and illusion (Coren, Ward et al.1999). This can be illustrated in the two commonly known illusions: Ponzo illusion and the Mueller-Lyer illusion. In Figure 2-11 (a) we assume that the farther line is similar in size to the nearer line even though they are drawn with different length. While in Figure 2-11 (b) we usually assume that the farther line is longer than the nearer line even though they are drawn with the same length. A possible explanation for this is that perspective cues (linear perspective) are strong enough to evoke size-constancy but not strong enough to apprehend distance (Gillam 1980 ; cited in Coren, Ward et al.1999).

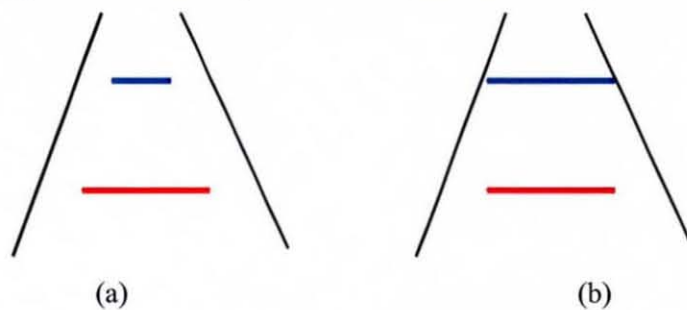


Figure 2-12 Ponzo illusion

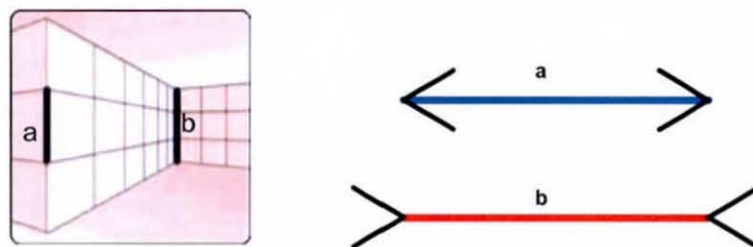


Figure 2-13 Mueller-Lyer illusion

Muëller-Lyer illusion is illustrated in Figure 2-12. In both cases, line b is overestimated compared to line a even though in both situations lines a and b are of equal length. This indicates that the context the object is in might influence our vertical distance and horizontal distance estimate of that object as illustrated by the Muëller-Lyer illusion and Ponzo illusion respectively. Thus the presence of strong perspective cues in a picture might influence our perception of distance. The context in which the object is located might also influence its height estimation as illustrated in the Muëller-Lyer illusion. Both conditions might cause error in perceptual judgment especially on distance perception.

### 2.7.2 Prior knowledge

The experience gained in an occupational setting or other setting may also influence individuals' perception and interpretation of various stimuli (Coren, Ward et al.1999). They suggested that specific experiences "produce a sensitization or predisposition to 'see' a situation in a certain way, especially when several alternative perceptual experiences exist". Thus a person's background such as jobs, skills, and other experiences would contribute to their ability to perceive depth or distance. For example, a user's sport skill ability might affect their distance estimation. Professional sport persons such as golfers, basketball players or tennis players rely on good distance estimation for their good performance in conjunction with proprioceptions. One might presume that their distance judgement would be fairly accurate due to their frequent training. However, in our daily experiences we also rely on distance estimations for everyday tasks (such as driving a car or playing games); it is reasonable to assume that our distance judgment would be less accurate compared to these professionals. Given this consideration, in the experimental work reported in this thesis information on participants' sport experiences was collected to examine whether this information influenced their depth perception.

Additionally, a person's prior experience or prior knowledge of an environment might influence their perception of it when compared to a person who has never seen the environment. Thus the environment location used as stimulus in the study must be carefully chosen as participants' familiarity with environment location might influence their depth perception. There exists empirical evidence to suggest that recognition of the scene is a critical step in perceiving depth based on pictorial information (Rock, Shallo et al.1978).

### 2.7.3 Cognitive Dissonance

The Cognitive Dissonance Theory developed by Leon Festinger (1957) is concerned with the relationship among cognitions. This theory asserts that people have the tendency to seek



consistency among their cognitions (i.e. belief, opinions, and worldview), when there is a contradiction among the cognitions they experience cognitive dissonance. When this occurs, they tend to eliminate this dissonance either by reducing the importance of the conflicting beliefs, acquiring new beliefs that change the balance or removing the conflicting attitude or behaviour. The removal of conflicting behaviour is usually the hardest to do. Thus, participants' previous experience, knowledge and expectations on different experimental conditions may cause cognitive dissonance.

### 2.7.4 Effects of sensory conflicts

In most VE, during simulation, the observers remain physically static while the environment passes by them. This experience is assumed to represent observer movement or navigation through the environment. This is especially true when an observer uses input devices such as a mouse, trackballs or joysticks. The represented movement presents a conflict between what observers see as opposed to what they feel: the visual system tells them that they are moving in a certain direction but since they are not moving, the vestibular cues indicate no movement. When this conflict of sensory information occurs the observer may exhibit symptoms of cyber sickness (Harris, Jenkin et al. 1999, Cobb, Nichols et al. in press).

Cybersickness or simulator sickness occurs when the user is stationary but experiences a compelling sense of self-motion through moving visual imagery (LaViola 2000). Symptoms of cybersickness include eyestrain, ataxia, fatigue, and drowsiness (Kennedy, Lanham et al. 1995). Under normal conditions, when the user is moving, both the visual and vestibular systems provide information of movement. In a VE, the visual systems only provide the user with visual information of movements. No vestibular information is provided to the user since the user is stationary. The conflict between sensory cues may cause the user to experience cybersickness. Thus, exposure to visual display which provides a compelling sense of motion but is not replicated in actual body movement might induce such an effect (Cobb, Nichols et al, in press).

#### 2.7.4.1 Proprioception

Generally, proprioception refers to the awareness of the body. This awareness is derived from the information provided by the receptors in our muscles, tendons and joints. Proprioception gives information about the movement and positions of parts of our body (Kalawsky 1993). When we move in the real world, proprioceptive cues can provide us with information on how far and how fast we walk and move (Reiser, Ahmead et al. 1990). Movement in the VE however is less natural and has limited proprioceptive information. Furthermore the interface

device is not directly related to the movement in the VE; this is true for relative interface devices such as the mouse, trackball or joystick. For example, movement of the mouse on the mouse pad yields relative movement in the VE. Thus, proprioceptive feedback is limited to those received from the muscles and joints of the wrist, arms and shoulder only. However, the visual information indicated movement in the VE but the proprioceptive feedback from other parts of the body indicates that the body is stationary. The inaccurate simulation of such cues might influence a user's performance in the VE.

It is possible to simulate natural movement (such as walking) using a head/body tracking system in the VE. However, this method requires a large space for the user to move around in the real world. This is especially true when dealing with a large VE where it is not feasible to provide a space as large as the space simulated by the VE. Moreover, tracking systems fail to function accurately over wide area. Some researchers (Allison, Harris et al. 2002) have tried to resolve the large space problem by developing the virtual reality tricycle (a stationary bicycle) which provides the non-visual cues (proprioceptive cues) in addition to the visual cues. Other researchers have tried to simulate more natural walking movement in the real world by the use of a treadmill (Witmer and Kline 1998). However, they found that a participant's performance on distance estimation is no better using a treadmill than using a joystick. In a different study, Grant and Magee (1998) examined the contribution of proprioception to navigation by providing participants with a walking interface and a joystick. Participants were asked to navigate a virtual building and their navigational abilities were tested on the actual building. Results showed that the transfer of spatial knowledge was significantly enhanced when using the walking interface which afforded proprioceptive cues. However, the walking interface participants were no better than joystick participants on an orientation task.

By simulating motion such as walking using a treadmill or the virtual reality tricycle, users would get both the vestibular and visual stimulation. But an incorrect alignment of the visual stimuli and the motion simulators lead to a conflict between the visual systems and the vestibular system, which could lead to users experiencing motion sickness (La Viola 2000). This might in turn affect users in performing the required tasks.

#### 2.7.4.2 The vestibular system

The vestibular system, located in the inner ear is the system concerned with orientating the body posture and balance (Schiffman 1990). It consists of two structures: the semicircular canals and the otolith organs. The systems sense and signal the movement of the head which results in the coordination of motor response, eye movement and posture (Draper 1996). The



vestibular system is responsive only to acceleration or deceleration of the body movement but not constant velocity.

#### 2.7.4.3 Vestibular-ocular reflex (VOR)

VOR is a “primitive eye movement reflex that stabilizes images on the retina during movement” (Draper 1996). Basically, it allows us to see clearly when we are in motion. Any movement of the head will be detected by the vestibular systems which send information on direction or rate of movement of the head to the oculomotor systems. The oculomotor systems then respond by moving the eye in an equal but opposite direction to keep the image stabilized on the retina. Inaccurate simulation of motion by the VE technologies would result in conflicting cues between the visual and vestibular information in where a stationary observer views a moving image, a conflict between sensory of information occurs. The visual information registers movement while the vestibular information registers no motion and this may lead to motion sickness in a vision-only display.

A common experience known as self-vection occurs when observers feel they are moving when in fact they are not. An illustration of this is when we watch a moving nearby train from the window of a stationary train, we feel that our train is moving. Vestibular information is absent and the visual stimulation is ambiguous, implying that either train can be moving. Because of our tendency to perceive a stable environment, we thus perceive that our train is moving (Schiffman 1990). From his review of several studies, Schiffman (1990) concluded that self-vection or visually induced illusion of motion appear equivalent to those produced from the actual motion. This means movement of the visual scene has the same effect on the individual nervous system as the stimulation of the vestibular system. In fact, navigation or movement in a VE can provide the stationary observer a compelling sense of movement (Harris, Jenkin et al. 2002) even though the vestibular feedback is not available. However, this conflict of sensory cues may lead to motion sickness and may reduce the users’ performance.

## 2.8 PICTORIAL PERCEPTION

Besides understanding of perception in the natural environment, designers of VE need to understand how people perceive photographs or pictures (Cutting 1997). In this section a discussion of perception of pictures and related work is presented. In addition, the geometrical theory of picture perception which describes the effect of viewing distance, position, height and angles is also presented.

The presence of all the static monocular cues mentioned in earlier sections in a picture enables pictorial perception. These cues allow us to perceive depth or three-dimensionality on a flat 2-D surface. However, viewing a picture of a scene is different from viewing a real world scene due to the character of the picture and the character of picture perception (Bengston, Stergios et al.1980). Pictures have dual reality: first, they are objects themselves and secondly, the marks on them represent other objects and space (Wade and Swanston 2001). Thus when viewing a picture of a scene, people are aware of its 2-D surface and at the same time people are aware that a 3D scene is being depicted (Yang and Dixon et al.1999).

Cutting (in press) explains the nature of perception of a picture in terms of depth cues. Despite the presence of static pictorial cues, he stated that when viewing a picture, "the status of accommodation and convergence cues, the absence of binocular disparities and motion cues tells us that we are not really looking into the distance." Our eyes tend to accommodate at the picture surface and may converge at different locations.

When we move our head in the real world, object positions and relationships change with our movement; however, in picture the objects and the relationships remain unchanged when the viewer moves. Thus, the status of the physiological cues and absence of motion cues tell us that we are looking at a picture not a real scene. Moreover, the picture frame and its surrounding context will also remind us of this viewing of a picture. We could eliminate this frame effect and its context by viewing a large picture at a close range such that the boundaries of the pictures are not visible. The result is an illusion of space (Bloomer 1990). Alternatively, viewing the picture through a rolled up paper tube will also eliminate this frame effect and enhance depths effect (Schiffman 1990).

Several studies have demonstrated such effects via monocular viewing of a picture where the participant's FOV is restricted to the image area on the picture (Smith 1958, Smith and Gruber 1958, Smith and Smith 1961, Hagen, Jones et al.1978). This impression of realism is flexible such that it is least when the photograph is close to the viewer and greatest when it is far from the viewer in these studies. Smith (1958) demonstrated that an increase in visual angle subtended by the photograph results in a decrease in apparent depth. However, Smith and Gruber (1958) study results showed that the perceived depth in the picture was consistently overestimated, the height and width of the horizontal remained constant even with varying depth distance. They attributed the constant values of the height and width of the corridor to the size-constancy illusion in the presence of a strong perspective cue, as in the real world viewing. These studies indicate that perception of depth in pictures is less accurate when compared to perception in the real world. Smith and Smith (1961) further demonstrated



that a picture could induce a realistic 3D space impression such that it can serve as distal stimuli for motoric responses. Their study results showed that participants tended to underestimate distance for far targets and overestimate near targets. Restriction of the FOV of the viewing apparatus was suggested to cause participants to increase perceived distance.

The less veridical estimates in pictorial perception have been attributed to the conflicting nature of the picture flat 2D surface with depth information it convey (Hagen, Jones et al. 1978). In addition to the conflicting nature of the picture perception, Hagen, Jones et al. (1978) demonstrated that the truncation of the visual field might also account for the compression of distance in picture. In their study, participants were asked to judge the distance and size of isosceles triangles viewed in four viewing conditions: unobstructed static monocular viewing, peephole view, view through a frame, and slides conditions. The study results showed that all the three conditions revealed smaller estimates compared to the unobstructed static monocular viewing condition. The truncation of the visual field in the three conditions causes a shift in the localization of the visual field which results in size and distance compression. Thus, the compression of distance in the earlier mentioned studies by Smith and colleagues could be attributed to the visual field truncation effect.

In the previous paragraphs, the studies examined perception in an indoor setting. Several researchers have also examined the perception of pictures of natural scenes or outdoor setting (Kraft, Patterson et al.1986, Hecht, Doorn et al.1999). Kraft, Patterson et al.(1986) presented participants with slides of natural terrain and asked participants to make direct distance estimates from self to targets in the scene. The pictures were captured using four different lens focal lengths: 48mm, 28mm, 24mm and 17mm. Longer focal length results in smaller viewing angle. Two types of terrain were used: cluttered and uncluttered terrain. Their results showed that distance estimation along the sagittal plane increases with increasing viewing angle, while distance along the lateral plane was not affected by the focal length. They concluded that wide angle (shorter focal length) results in more accurate estimate than small angle (longer focal lengths). With wide angle, the shorter focal length is associated with a decrease in truncation of the visual field, that is, the foreground is closer to the viewer in the wide angle condition. Furthermore as focal length decreases, the parallel lines and texture elements shrink more quickly along the depth plane; these changes in pictorial information would indicate an increase in distance. The study also indicates that estimates in the light wooded terrain tend to be greater than in the open terrain but these results may be due to the distances used in light wooded terrain being much shorter than those used in the open terrain condition.

Hecht, Doorn et al. (1999) compared distance and angle perception of real building corners and pictures of them. For angle estimates they found no significant difference between real and picture viewing of the angle subtended by building corners, especially for large distances. However, for perception of distance from self to the corners, they found that in the real world participants tend to overestimate near distance by 36% and far distance (5- 15m) was overestimated by 7.6%. In contrast, pictorial far distance is underestimated by 30.8% and near distance was overestimated by 71.1%. However, their study results show that the main effect of the condition (photograph versus real world) did not reach a significant difference.

Bengston, Stergios et al.(1980) examined the effect of viewing positions in the perception of distance and size in pictures. Participants were presented with five pictures of different perspectives depicting the layout of two dolls and were asked to estimate the distance between the dolls and the size of the dolls. Participants viewed the pictures at five different positions which corresponded to the five different perspectives. They found that viewing photographs from incorrectly large distances would result in an overestimation in pictorial distances. However, Cutting (1987) found that physical viewing distance from the computer screen has no effect on perceived distance.

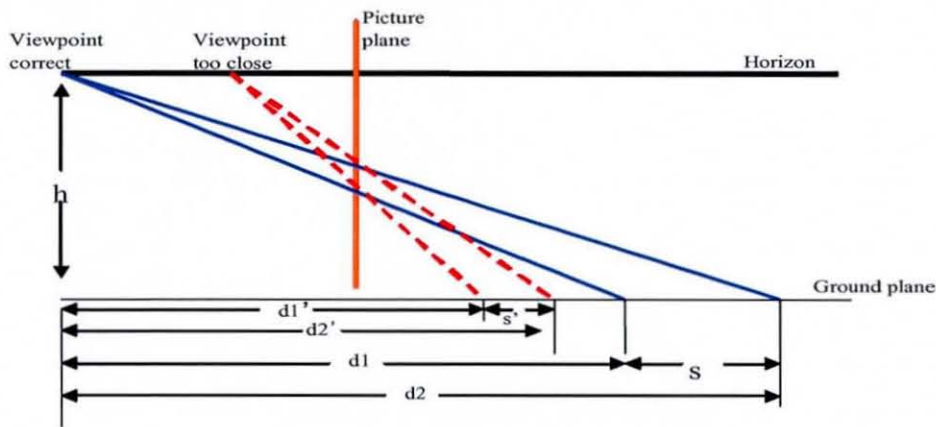
Studies reviewed in the prior paragraphs showed that distance perception in pictures is less accurate when compared to the real world. However, Cutting (1997) argued that there is actually nothing special about picture perception as compared to perception of natural scenes, except that in picture, as discussed earlier “cue conflicts” are present due to its dual aspects. In cinema, viewing at distances greater than 15 or 30 m would avoid these cue conflicts, thus producing effects that viewing the movie is similar to viewing a natural scene (except it is limited by the screen frame and choice of lenses and shooting distances). However, for most applications, this would not be practical due to space constraints. For VE systems, the presence of cue conflict might not allow generalization of perception in VE systems to perception in real world. Thus, potential cue conflicts must be removed in order to achieve the goal of mimicking everyday perception. Despite such cue conflicts, particularly in history of art, photography and cinema, the HVS has performed very well (Cutting 1997).

### 2.8.1 Geometrical theory on picture perception

In this section, the geometrical theory of viewing pictures or pictorial display is presented. One of the purposes of picture or pictorial display is to provide information about the 3D layout of an environment. The creation of the image involves the perspective projection of a three-dimensional scene onto a 2-D image plane (display). This is done by following the



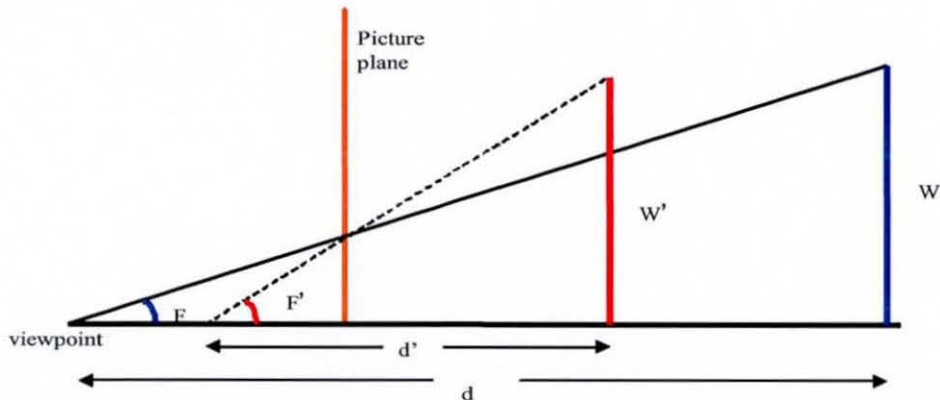
projection lines from a fixed point of view back into the scene and then determining the point of intersection between these lines and image plane (Barbour and Meyer 1992). The accuracy of such ability is discussed by Sedgwick (1991) using the perspective structure of the optic array to determine the geometrically specified sizes, distances and orientation of surfaces and edges in the pictures. Optics array here is referred to as a ‘structured array of light reflected to a point of observation by the surfaces of the environment’. He presented an analysis on the theoretical effect of viewing distance from the picture (far and close), viewing position (sides) and viewing heights on virtual space of a picture.



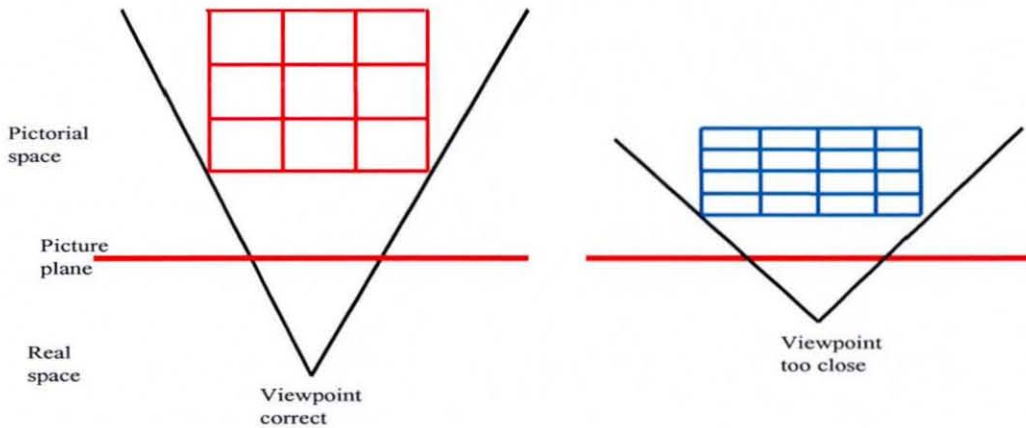
**Figure 2-14 Close viewing compresses geometrically specified depth**  
(Adapted from Sedgwick 1991)

When we approach a picture, the geometrically specified depths in the picture are compressed proportionately. Thus, the geometrically specified depths in the picture are compressed ( $s'$ ) when we are at a close distance from the picture and expand proportionately ( $s$ ) when we move away from it (Figure 2-13).

Close viewing has no effect on the geometrically specified frontal dimensions (Figure 2-14), but this distorts the geometrically specified virtual shape (Figure 2-15).



**Figure 2-15 Close viewing leaves geometrically specified virtual frontal dimension unchanged**  
(Adapted from Sedgwick 1991)



**Figure 2-16 Close viewing distorts geometrically specified virtual shape**  
(Adapted from Sedgwick 1991)

Moving laterally parallel to the picture causes a shearing of the virtual space. Viewing height has the same effect of moving laterally, except that the virtual space is sheared vertically in the former. To see the theoretical prediction, the picture needs to contain a strong linear perspective; a weak one may find it difficult to see these distortions. Most empirical investigations found such distortion but not at the predicted magnitude.

In his studies, Goldstein (1991) observed that the perception of the spatial layout remains constant with a changing viewing angle; that is the ability of the participant to reproduce spatial layout is not much affected by the change in viewing angle. He also found that changes in viewing distance have no effect on the observer's perception of spatial layout. On the contrary he found that perceived orientation was affected largely by changes in the viewing angle. He suggested participant's awareness of the picture plane as one possible cause for this perceived orientation.

## 2.9 PERCEPTION IN VE

The usefulness of applications that use VE to represent its real world counterpart depends on the VE technologies providing similar perception and experience in both worlds. As such users must be allowed to perceive spatial relations in the VE in the equivalent way as they would in the real environment. However, to date, VE technologies have not been able to allow the user similar perception and experience to the real world. Several studies have indicated that VE allows users to perceive the VE space differently from the real world (Henry and Furness 1993, Lampton, McDonald et al. 1995, Waller and Hunt et al. 1998, Witmer and Kline 1998, Sinai and Krebs et al. 1999, Loomis and Knapp 2003). The magnitude and

direction of spatial perception in these studies were often varying and contradicting due to the differences in factors investigated and the experimental methods employed. It is thus difficult to generalize findings from these studies. As such current knowledge on factors necessary to provide similar perceptual and performance experience to the real world is still limited (Witmer and Kline 1998). While significant advances have been made in the display and computing technology (Stanney and Zyda 2002), until recently little work has been done on evaluation of how users perceive such environments (Witmer and Kline 1998, Rushton and Wann 1993). As the utility and the effective and efficient design of an application using VE technologies depend on the user's ability to perceive VE similar to its real counterpart, it is thus essential to examine and understand factors that affect the user's perception in the VE.

The ability to perceive 3D space (spatial awareness) in the real world is crucial for our safe and effective interaction with the environment (Sekuler and Blake 1994). As such making the perception of the virtual space similar to the real environment would also be of prime importance because a VE also allows a user to experience and explore a 3D but computer-generated space. More research is thus required to add to the knowledge and understanding of how to allow user's perception and performance in VE to be similar to the real environment.

### 2.9.1 Spatial awareness

Basically, spatial awareness refers to our awareness of elements within an environment. It includes awareness of object locations or relative positions between objects in the space. Several researchers have included spatial awareness as one of the important components of a broader and complex concept of situation awareness (SA) (Venturino and Kunzo 1989, Fracker and Davis 1990; cited in Draper 1995, Endsley 2000). In its simplest term SA is to know what is happening around us (Endsley 2000). While there is no single definition of it, the following is a common and generally accepted definition:

“... perception of elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future” (Endsley, 1988, cited in Endsley (2000)).

According to this definition SA encompasses three levels: perception, comprehension and projection. Perception refers to an awareness of the elements' status, attributes and dynamics. Comprehension is the decision-maker's overall picture of the environment, including the relevance of objects and events. Projection is the ability to predict the future states of objects. It is very important for operators of complex systems to achieve and maintain SA as the cause

of many accidents in complex systems in the past has been attributed to its operator's lack of SA (Bass, Zenyuh et al. 1996). Despite this, however, achieving and maintaining SA remains a difficult task for humans. As SA is a very broad issue, the focus of this thesis is on spatial awareness as one important aspect of SA. Moreover, as emphasized by Endsley (2000), perception (that is awareness of elements' status, attributes and dynamics) forms the basis of SA where incorrect perception might drastically affect the next two levels of SA which are comprehension and projection.

In geography, spatial awareness takes a wider concept, it encompasses spatial movement in the environment, identifying and interpreting spatial patterns and understanding decision making that affects spatial arrangements, perceptions and understanding of the physical and social environment (Catling 2000). Crvarich (1995) and Draper (1995) however refer to spatial awareness as a person's perception and understanding of the 3D layout of an environment. Similarly, in this thesis, we define spatial awareness as the awareness of the 3D environment, which includes knowledge and understanding of objects' spatial locations and relative distances within that environment. Spatial awareness is a requirement for several tasks in the physical or virtual world (Draper 1995). Such tasks include object manipulation, navigation and way finding.

Several terms are often associated with spatial awareness. Information about the space or environment is referred to as spatial knowledge. Spatial knowledge may be learned through various sources: direct experience, maps, photos, drawings, video movies and videos, verbal written language and simulation (Witmer, Bailey et al. 1996). Spatial representation refers to human representation of space. Spatial representation, as defined by Siegel and White (1975), functions to facilitate location and movement within a large environment. Other terms have been used: cognitive maps, mental maps, survey maps, configuration maps and environmental maps. These terms including spatial knowledge and spatial representation are often used interchangeably by researchers. Basically these terms refer to knowledge of the space perceived; as such this thesis also makes no distinction between these terms. Siegel and White's (1975) model of spatial representations is made up of three important elements: landmarks, routes and configuration. The formulations of spatial representations begin with noticing and remembering landmarks. When people have developed an ordered sequence of landmarks, they have acquired the route knowledge. Configuration knowledge provides a person with survey-like knowledge. This knowledge is useful for way-finding and organising experience.



Following a review of several studies, Arthur, Hancock et al. (1997) concluded that “a central issue for the use of VEs both as an interface and training tool is how users mentally represent that virtual space”. They further asserted that the utility of VE for any application for which they are being intended is predicated upon the accuracy of this spatial representation formed in the VE. As such it is essential for a user to understand the space in which the tasks are to be performed. Thus, understanding how people form cognitive maps or spatial representation of VE is very important for effective VE design. Because the perception of distance forms the basis of our understanding of the physical structure or space perception (Golledge 1991; Coren, Ward et al. 1999), accordingly it is essential to understand factors affecting distance perception. This implies understanding and knowledge of distances and spatial representations are critical for the perception of space or our spatial awareness in the VE as well. The next two sections will present related literature on distance perception and spatial representation (or spatial memory studies). The last section will review studies that examine factors affecting distance perception in the VE.

### 2.9.2 Distance perception studies in VE

Studies concerning distance judgment in the real world are numerous. An overview of these studies is provided by Waller (1999). Generally, estimated distances were not veridical with respect to the actual physical space. Collating data from several studies, Wright (1995) reported that typical real world estimates are in the range of 87-91 percent of the actual distances. In VE, research interest on distance estimation in VE is just recent (Witmer and Kline 1998). Studies related to distance perception in VE were reviewed in the following paragraphs.

Caird and Hancock (1991) examined participant estimation of an object's location in a simulated scene of a traffic intersection as a function of their experience in it. The scene, which consists of coloured polygons, was projected on a 10ft diagonal screen, 8.4ft from the participant. Participants were asked to make relative and absolute distance judgments of nine objects presented to them. Estimates for participants with experience were shown to be more accurate than participants without experience.

Henry and Furness (1993) reported findings that people perceive real and virtual spaces differently. In their study they asked participants who had experienced a 15-minutes guided tour of a virtual and real museum to perform spatial dimension, orientation and evaluation tasks. Four viewing conditions were compared: stereoscopic head-tracked HMD, stereoscopic non head-tracked HMD, desktop monitor and real environment. Their study result showed

distances were underestimated more in the VE compared to the real environment. They also found that participants tended to underestimate distance more in the head-tracked HMD condition compared to the non-head-tracked HMD condition and monitor condition.

Lampton, Bliss et al. (1994) who compared participants' performance in real and VE in terms of object recognition, height estimation and egocentric distance judgement tasks also found that real world participants were more accurate compared to the VE results. In their study participants were required to recognize object (person), estimate height and judge distance to object as the object moved closer to the participants. The range of distance estimation was 2.5 – 40 ft. The VE was presented stereoscopically on a visual research flight helmet. Different groups of participants performed similar tasks in the real world setting. Results showed that participants tended to underestimate height in the VE but overestimate egocentric distance.

Lampton, McDonald et al. (1995) findings revealed that real world estimates were significantly more accurate than the VE conditions. In their study, the participant's distance estimations for static and moving images were compared under four viewing conditions: stereo head-tracked HMD (234 lines), stereo head-tracked Binocular Omni-Oriented Monitor (BOOM) display (1280x472), computer monitor (1024x1248) and real world settings. For moving image, the real condition performed significantly better than the three VE conditions and the BOOM display was significantly better than the desktop monitor. Participants tended to underestimate egocentric distance in the VE. Distance estimation was least accurate with monitor condition. For static distance estimate, HMD participants highly overestimated the distance and the BOOM display gave the lowest among all conditions.

Witmer and Sadowski (1998) showed that egocentric distance judgment in a VE average 85% of actual compared to 92% of actual for real environment. The authors compared distance judgment based on blind walking task in a real hallway (46m long) to a real hallway viewed binocularly using a head-tracked stereoscopic display (1280x1024 resolution) in monochrome mode. A manual treadmill calibrated to the user's walking speed was used to represent locomotion in the VE. They attributed the underestimation of distance in the VE to distance cues which was not perceived similarly to the real world. The narrow FOV might have degraded height in the visual field, linear perspective and relative size such that it compresses objects into a smaller visual frame as they recede into the distance, making distant objects appear closer than they would in the real world. Moreover, the binocular disparity cues may be erroneously represented in VE especially for shorter distances.

Witmer and Kline (1998) examined perceived egocentric distance and traversed distance judgement in VE. They found that participants greatly underestimated distances in VE. Participants' performance in the real condition was 72% of actual while performance in the virtual hallway was about 47% of actual distance. They attributed the difference in performance to the fewer cues present in the VE compared to more cues present in the real environment. Their results also indicated that estimates were more accurate for small cylinders compared to large cylinders; this led them to conclude that decreasing the size of the object might compensate partially for the underestimation of distance in VE. They also found that texture was not a reliable cue for distance estimate.

Sinai, Krebs et al. (1999) found that egocentric distance judgements were relatively more accurate when assessed using a perceptual matching task although participants tended to overestimate distance. However, far distances tended to be underestimated. Their study result showed distance tended to be overestimated by approximately 7%. The authors also found that texture significantly affected distance perception with the medium symmetrical brick pattern giving the highest user performance.

Eggleston, Janson et al. (1996) evaluated the effects of the VR system factors on users' size and distance judgments. Factors evaluated included mode of viewing (stereo vs. bi-ocular viewing), image resolution (1280 x 1024 vs. 640 x 480), field of view (60° x 60° vs. 60° x 100°) and scene contrast (single vs. multiple luminance). Participants were presented with two VE corridors (constructed from shaded polygons) on a monochrome HMD display. Participants' tasks were to adjust object size in one corridor to match the object in the other corridor. Results showed that impression of depth was greater in multiple luminance compared to single luminance conditions. Performance on higher resolution display was significantly better than on low resolution display but the difference between the field of view conditions was very small. Their study results also showed significant interactions between mode of viewing with field of view, image resolution with field of view and mode of viewing with image resolution. The authors concluded from their study that there was a difference between perception of 3D information in VE and real conditions however they asserted that "it is not clear what is missing in a VE and how the deficiency could be corrected".

Kline and Witmer (1996) examined the effects of system-related cues on user's estimation of distance within the personal space of 1-12 feet in a VE. The system-related cues investigated were texture type (rich emergent vs. poor non emergent), texture resolution (512 x 512 vs. 16 x 16) and FOV (140 x 90 vs. 60 x 38.5). Participants were asked to estimate the distance to the wall at the end of a virtual corridor presented on a monochrome high stereoscopic display

without head-tracking. Their results showed that the distances estimated were significantly affected by the FOV and texture type. More accurate estimates were found for wider FOV compared to narrow FOV. Generally, participants overestimated distance with narrow FOV but underestimated distance with wide FOV. Fine texture resolution provided more accurate estimates than coarse texture. Rich texture was more effective than poor texture at near distance but at larger distances the difference was very small. The absence of texture and insufficient perspective cues led participants to overestimate their distance. Overall results showed that with a wide FOV, rich and fine textures yielded the most accurate estimates.

Wright (1995) on the other hand investigated participant perception of forward, lateral, height and speed while viewing a computer-generated image of a terrain on a high resolution, wide angle head-tracked HMD. Viewpoints were adjusted using a joystick. Results showed large underestimates for forward distance (41% of actual), lateral distance (50% of actual), height (72% of actual) and speed (41% of actual).

Yoon, Byun et al. (2000) compared users' perception of psychological properties of a real and virtual room and their judgment of size of the rooms in terms of the width, height and length. A simple room which consists of one door, one window and one chair were used as stimuli. A counterbalanced design was employed. Free exploration of the room was allowed but no time allowance was reported. Prior to estimation, participants practiced estimation in three real rooms. A HMD was used to view the virtual room and navigation was controlled using a mouse device. Their results showed no significant difference between distance estimates in the real and virtual room. However both differ significantly from the actual distance. It was found that participants tended to make more errors in height estimation compared to width and length. Overall distance estimates were reported to be accurate even though users tended to overestimate distance.

In his studies of exocentric distance estimation in VE, Waller (1999) found that distances were generally overestimated. Participants were asked to freely explore a cube room and estimate distance between two red boxes placed at a random location in the room. The presence of a grid (perspective cue) had a significant effect on distance estimates. While the effect of display type (head-tracked HMD and desktop monitor) was less influential, it does approach significance for HMD. This small difference was attributed to small VE. His study revealed that estimations were more accurate for GFOV between  $50^{\circ}$  and  $80^{\circ}$ . On wide GFOV ( $100^{\circ}$ ) participants tended to overestimate while in low GFOV distance judgement tended to be inferential rather than perceptual. He concluded that distances in VE were not necessarily

perceived differently from real environment when participants were presented with sufficiently wide GFOV with feedback.

The above studies showed that distances estimated in the VE were less accurate than those found in the real world. While some researchers reported an overestimation in distance perception in the VE (Sinai, Krebs et al. 1999, Waller 1999, Caird and Hancock 1995), others found distance perception in the VE to be underestimated (Henry and Furness 1993, Witmer and Kline 1998, Witmer and Sadowski 1998, Wright 1995, Eggleston, Janson et al. 1996). Still other researchers (Waller 1999, Yoon, Byun et al. 2000) revealed that the differences between real and VE were too small to be of practical significance. The differences in the variables investigated such as egocentric distance versus exocentric distance, length of distances tested, methods of measuring distance (direct measures such as verbal measures and indirect measures walking task), display types (desktops, HMD, head-tracked and non-head-tracked) and other differences in experimental methods might have contributed to these differences. These differences make it difficult to generalize findings from these studies. As such, the exact reasons for these differences in distance perception in the VE are still unknown (Willemson and Gooch 2002), though various factors have been suggested to account for these differences.

### 2.9.3 Spatial representation studies in VE

Perception of spatial layout of complex scenes has not been widely researched (Caird & Hancock, 1991). As previously mentioned, understanding how people formed mental representation of the space or spatial memory of the VE is very important for its effective design. In this section, studies involving spatial representation/memory task in the VE are presented.

Arthur, Hancock et al. (1997) investigated user's perception of the real and VE by comparing participant's performance on their mapping accuracy and relative inter-object distance judgments. Participants were exposed to a spatial layout of nine objects on the floor under three viewing conditions: free binocular viewing of VE condition, free binocular real environment condition and static monocular view of the real world. Participants were allowed to freely navigate the VE and real environment for the first two conditions and view the environment from a small hole for the third condition. Participants were told to observe the spatial layout of objects. They were given as much time needed to explore or view the environment. For the map test, two of the objects were given to provide them with scale and orientation information. For the relative distance judgment, participants were asked to rank

order the distance between three possible pairs of any triad combination. Their results suggested that spatial representation resulting from interaction with small scale VE was comparable to real world experience. Thus, VE can be effectively used to simulate spatial relations. However, they found out that the monocular viewing condition yielded superior results compared to the real and VE. Explanations offered are: viewing is similar to map viewing that is from a single orientation, thus allowing focus on spatial layout of the objects, whereas real and VE conditions allow multiple orientations and so focus might be on the objects themselves and not on their locations. Additionally, strategies are less constrained for monocular viewing compared to real and VE conditions.

In their study, Johnson and Stewart (1999) assessed their participants' spatial knowledge acquisition in the VE using an object placement test where participants were required to place 34 objects in an outline of the heliport. Participant performance was compared in three viewing conditions: a wide FOV HMD, a narrow FOV HMD and a non-immersive rear-projection wide-screen. Participants were initially trained for two 30 minutes session. The HMD condition resolution (1024x1048) was higher than the wide-screen resolution (946 lines). Participants navigated the VE using a virtual carpet controlled by using 2 joysticks. Score results showed that all conditions were not significantly different from each other (76%, 78% and 83% for wide FOV HMD, narrow FOV HMD and large screen respectively. The implication of their study is that immersive visual displays are not necessarily more effective for spatial learning than other types of display such as the non-immersive large screen.

Patrick, Cosgrove et al. (2000) investigated spatial knowledge learned under three viewing conditions: HMD, large projection screen and desktop monitor. The display resolution and FOV were matched in all viewing conditions and no condition used stereo presentation. FOV of device was set to 60° x 46.5° and the resolution was 640 x 480. Participants were asked to produce a map of the layout of the VE after guided exploration of the VE. The interface device used was a steering wheel and navigation was restricted to driving mode only (no vertical movement). Their study results showed no significance difference between HMD and large screen or between HMD and monitor but there was a difference between monitor and large screen. The better performance in large screen condition over HMD and monitor was suggested by the authors due to the large display inducing more presence and that the images are big enough to appear real and thus resulted in better estimation.

Rossano and Moak's (1998) study involved exposing participants to two experimental conditions: a map of a campus and a computer model of it. The computer model was created with precise and realistic details and was presented on a 15" colour monitor. Participants

viewed a 20 minutes guided tour of the computer model and were asked to learn objects' locations and the layout of the campus. View was limited to ground level only. For the map participants, they were asked to study a map of the campus for two minutes. All participants were given 4 sessions with a week between sessions. Participants were given an orientation and configuration test. Their study revealed orientation specificity was eliminated by reducing the cognitive load of the participants through the use of actual test of orientation as opposed to simulated test of orientation. However, there was no significance difference between map group and computer view group for configuration or survey test.

In another similar study, Goerger, Darken et al. (1998) compared a map only exposure group to a map and VE exposure group on distance and direction tasks. For the VE, a high fidelity and accurate seven-storey building model was used and presented on three projection screen of 145° FOV. Participants were given a map of the VE and passively viewed the VE while giving the experimenter command on movements through the VE. The other group were just given a map of the building. Participants were later tested in the real building. Their study result showed that map only group participants performed better than map and VE groups participants on target placement tasks. They attributed this to the short exposure time of 30 minutes exploration of the VE. However, the passive experience of the viewer might also be a contributing factor.

The first study by Arthur, Hancock et al. (1997) indicated it is possible to perceive spatial relation in VE similar to the real world. Their study further revealed that viewing from a single orientation results in a better spatial representation compared to viewing in multiple orientations as in real and VE conditions. This is because users can focus more on objects' locations rather than the objects themselves. The second study by Johnson and Stewart (1999) compared spatial representation in displays of different FOV. They found that spatial knowledge formed was similar in the three conditions tested (wide FOV HMD, narrow FOV HMD and non-immersive rear-projection screen). However, no comparable real world conditions were carried out. Patrick, Cosgrove et al. (2000) also investigated spatial knowledge formed after a guided tour of a VE presented on three display types: desktop monitor, HMD and large projection screen. However, they found no significant difference between HMD and a large projection screen condition but there was a difference between these displays from the desktop monitor. No difference between large projection screen and HMD was explained due to the wide FOV where objects perceived are large enough to induce realism similar to the real world. However, again no comparable real condition was performed for comparison. No real conditions were included for these studies (Johnson and

Stewart 1999, Patrick, Cosgrove et al. 2000) because their main emphasis was on comparison among display types.

The next two studies compared performance using a map and a VE. Rossano and Moak (1998) compared user formation of the spatial layout of the campus after participants learned it through a map or by using a computer tour of it. No real world condition was conducted. They found no significant difference in performance in the spatial representation test between the map group and the computer group participants in the survey test. However, the next study by Goerger, Darken et al. (1998) indicated that map participants tended to perform better than VE + map participants. In a different study, the acquisition of spatial representations of an environment acquired was compared between map, real and VE (Richardson, Montello et al. 1999). Participants were asked to learn two floors of a complex building under these three conditions. Results showed that VE learners were the poorest on the learning of a complex building and they were more susceptible to disorientation after rotation compared to other conditions. However, using a simple single floor, all conditions revealed similar levels of participants' performance.

In the three later studies described above, the acquisition of spatial representation from VE was compared to those from map and real condition. The results from these studies however are not consistent. The study by Richardson, Montello et al. (1999) indicated that spatial representation acquired for a simple environment is similar for all conditions but for a complex environment, the VE condition yielded the least accuracy. But Rossano and Moak (1998) using a more complex campus environment revealed no significant difference between map and computer view participants on the spatial map test. Goerger and colleagues (1998) however indicated map participants performed better than VE participants who also used maps. Differences in experimental methods and different test methods may account for these inconsistent results. As such it is not possible to generalise findings from these studies.

## 2.9.4 Factors affecting spatial perception in VE

In the real environment, factors affecting distance perception have been intensively studied (see Cutting and Vishton (1995) for reviews). However, similar studies in VE have only received research attention in the recent years (Witmer and Kline 1998).

A VE experience depends on the VE system's ability to simulate the human sensory channels. Synonymous with the human perception system, the visual channel represents the most dominant sensory channel compared to other channels (for example auditory, haptic, tactile)



in VE. This highlights the need to enhance the capabilities of the visual display system in order to closely match the VE visual experience to the real environment. Unfortunately, the VE is often perceived differently from its real counterpart. In fact, the display system has been suggested as one of the probable causes of distance underestimation in VE. Roscoe (1984) suggested that the basic problem with all computer-animated, sensor-generated, and optically generated displays is that they produce systematic errors in size and distance judgments. He concluded that spatial information on a computer display requires modification for it to appear normal. A magnification of approximately 1.25 will cause objects to be perceived at their objective distances for most observers, though this may vary with different imaging systems and individuals. He found that pilots tend to overestimate distance for minification of 0.86 and underestimated distance for a magnification 2.00.

Related aspects of the display or the computer systems have been the focus of intensively studied factors to influence spatial perception in the VE (Waller 1999). System-related factors such as variation in display-types, FOV, image quality, image type, scene contrast, resolution, viewing modes, interface devices, modes of travel in VE, mismatch of cues have been suggested as potentially contributing to the misperception of distance in the VE. In this section studies on examining factors influencing spatial awareness in VE are reviewed.

#### 2.9.4.1 Display types

The visual display system forms an integral part of a VE system and many VEs are often characterized by the display they used (see Section 3.2 of Chapter 3 on types of VE systems). Few empirical evidences exist to provide an understanding of the strengths and weaknesses of the VE display devices (Bowman and Datey et al. 2002). However, several researchers have compared users' performances on display types used to view the VE (Willemson and Gooch 2002, Henry and Furness 1993, Lampton, McDonald et al. 1995, Waller 1999, Heineken and Shultze 2000, Riley and Kaber 1999, Johnson and Stewart 1999, Youngblut and Huie 2003, Patrick, Cosgrove et al. 2000). These studies are reviewed in the following paragraphs.

Willemson and Gooch (2002) provide empirical evidence suggesting that the display techniques cause distance underestimation in the VE. Their study compares participants' egocentric distance judgement task using a directed motor action in three conditions: real hallway, virtual image-based hallway on HMD and computer-generated image of the hallway on HMD. Participants viewed the images binocularly. The HMD resolution was 1280 x 1024. Their study revealed a significant difference in performance between real world conditions and virtual conditions. Distance judgement between image-based and computer-generated VE were not significant though the image-based participants performed slightly better. This led them to conclude that underestimation of distance in VE was due to the display factor and not

the images. They attributed the better estimate of VE participants in their study compared to past studies to the geometric complexity of their VE model and the high resolution of their display system. The authors suggested that understanding the causes and magnitude of spatial compression in VE requires still requires more investigation.

Henry and Furness (1993) compared users' performance on a virtual desktop monitor, head-tracked HMD and non head-tracked HMD conditions to real world conditions. Their study showed that estimation in the VE is less accurate compared to that in the real environment. They found that participants tended to underestimate distance more in the head-tracked HMD condition compared to the non-head-tracked HMD condition and monitor condition.

Lampton, McDonald et al. (1995) investigated participants' distance estimation under four viewing conditions: stereo head-tracked HMD, stereo head-tracked Binocular Omni-Oriented Monitor (BOOM) display, computer monitor and real world settings. The resolution of head-tracked HMD, the BOOM display and computer monitor was (234 lines), (1280x472) and (1024x1248) respectively. For static distance estimates, distances were highly overestimated by the HMD participants but the BOOM participants yielded the lowest estimates among conditions. Performance in real conditions was significantly better than all VE conditions for moving images. Estimates however were least accurate in monitor condition. In VE conditions, participants tended to underestimate egocentric distance.

In his investigations on distance perception, Waller (1999) compared distance estimations on a head-tracked HMD to those on a desktop monitor. He found that the effect of display type (head-tracked HMD and desktop monitor) was less influential but it did approach significance for HMD. This small difference was attributed to the small VE and the between-subject design which tended to yield less significant results.

Heineken and Shultze (2000) however showed that the distance estimation task using the bisection method was more accurate in head-tracked HMD condition than on a desktop monitor even though the FOV of both conditions were equated. The participant task was to bisect a route in a simple low resolution VE which had been explored earlier. Route lengths were 1.5 and 6.0m. More error was reported on a desktop monitor compared to a HMD.

In another study, Riley and Kaber (1999) examined the effect of display types (desktop monitor, HMD, large screen projection) and navigational aids on participants' navigation performance, presence, and workload during exploration of a virtual office using a telerobotic vehicle. Participants used a conventional mouse to control the movement of the vehicle.

Display types had a significant effect on presence with the monitor condition revealing the highest sense of presence. Navigation time was faster on desktop monitor compared to other display types. The difference in resolution might account for this unexpected result (monitor – 1280x1024; HMD – 640x480, large screen – 600x800). Additionally the familiarity of the participants with the desktop and unfamiliarity with the other two displays may have affected their presence rating. However, display types had no significant effect on workload as reported by the participants.

A study by Johnson and Stewart (1999) compared a wide FOV HMD, a narrow FOV HMD and a non-immersive rear-projection wide-screen on participant's spatial knowledge acquisition task. The resolution for HMD condition was 1024x1048 while the wide-screen resolution was 946 lines. Participants' spatial knowledge acquisition (landmark and configuration knowledge) was assessed using an object placement test where participants were required to place 34 objects in an outline of the heliport. Their study result revealed that all three conditions were equally able to develop spatial representation of a virtual heliport. Scores results do not differ greatly among conditions (76%, 78% and 83% for wide FOV HMD, narrow FOV HMD and large screen respectively). The implication from their study result shows that immersive visual display is not necessarily more effective for spatial learning than other types of display such as the non-immersive large screen.

Similarly, a recent study done by Youngblut and Huie (2003) showed the difference in users' performances between desktop and rear projected display. In their study, participants were asked to train on mission procedures in two virtual training sites (a warehouse and an office building). They were tested on this knowledge in a real world training transfer test. The results showed no significant difference in performance for both displays which led the authors to conclude that the immersiveness of the display did not influence participants' performance. Additionally, they found no significant difference in the sense of presence during the training session in both displays.

Patrick, Cosgrove et al. (2000) compared spatial knowledge learned through three display conditions: HMD, large projection screen and desktop monitor. All conditions were matched for resolution, FOV and non-stereo viewing. Participants were asked to perform a guided exploration of a VE (virtual amusement park) followed by a cognitive map test of the visited VE. Participants were given as much time as needed on both the exploration and map tests. Scoring was based on the distance errors on the map tests. Results showed the difference between HMD and large screen or between HMD and monitor on mean error scores was not significant. But mean error scores of large screen participants were significantly less than the

monitor conditions participants. This better performance in large screen conditions over the monitor conditions was suggested by the authors due to the large display inducing more presence and the images are big enough to appear real and thus give better estimations.

The first three studies described in the previous paragraphs showed that participant's performance in a VE differed significantly from a similar task performed in the real world (Willemson and Gooch 2002, Henry and Furness 1993, Lampton, McDonald et al. 1995). Some studies indicated there was a performance difference between HMD and desktop monitor (Heineken and Shultze 2000, Riley and Kaber 1999, Patrick, Cosgrove et al. 2000). Similarly some studies showed there was a significant difference between large projection displays and desktop monitors (Patrick and Cosgrove et al. 2000, Riley and Kaber 1999). With the exception of Riley and Kaber's (1999) study results, other studies indicate better performance on HMDs and large projection displays than on desktop monitors. It was shown that there was no significant difference between HMDs and large projected displays (Patrick, Cosgrove et al. 2000, Johnson and Stewart 1999). But Riley and Kaber (1999) indicated better performance on desktop monitors than on large screens. They attributed the better performance of the desktop monitor participants to the higher resolution and participants more familiarity with desktop monitors.

The focus of the aforementioned studies was on comparing spatial performance on various display types used to view the VE. Very few studies have directly examined the effects of display size on spatial performance especially on distance perception (Kline and Witmer 1996), indicating the need for more research. In this thesis, display size is one of the main factors examined. Rationales and reasons on the need to understand the effect of display size factor are further discussed in Chapter 4 (see Section 4.1.2.1).

It should be noted however there is a distinction between GFOV and FOV. The former refers to the visual angle subtended by the virtual scene while the latter is often referred to by most researchers as the angle subtended by the display device on the viewer's retina. Generally, a wide or narrow GFOV allows more or less of the virtual scene to be seen respectively without changing the viewing area on the screen. Generally a wide GFOV leads to scene compression and minimisation and this might cause perceptual error in distances, angle and shapes of objects (Lumsden 1980).

Studies have demonstrated that truncated FOV or narrow FOV may also result in misperception of distance (Hagen, Jones et al. 1978). In this thesis the GFOV is not manipulated but remains constant for all conditions. However, the physical FOV or FOV is

manipulated to examine the effect of varying physical display size. According to geometrical prediction the FOV size (and consequently the retinal image size) depends on the distance of the object from the viewer and also the size of the object. Though, this might be affected by size constancy where perception maintains size constancy under full cue conditions and perception follows retinal image size when the cue to perception is minimal (Eggleston and Jansen et al. 1996)

#### 2.9.4.2 Image quality, resolution and luminance

Besides display types, the influence of other display related factors such as image quality, resolution and luminance on spatial performance were also investigated (Jää-Aro and Kjelldahl 1997, Kline and Witmer 1996, Duh, Lin et al. 2002, Eggleston, Jansen et al.1996, Willemson and Gooch 2002, Loomis and Knapp 2003, Thompson, Willemson et al. in press). Jää-Aro and Kjelldahl (1997) examined the effect of image resolution on distance perception in stereo and non-stereo images presented on a HMD. Participants were presented abstract objects of different polyhedron shapes and void of shadows and textures and were asked to estimate the distance to them. The five levels of image resolutions compared were 832 x 560, 416 x 280, 208 x 140, 104 x 70 and 52 x 35. Their results showed that low resolution has a negative impact on distance judgment. Stereo images yielded the worst estimates at low resolution when compared to non-stereo image.

Kline and Witmer (1996) reported that distance estimation was significantly improved when higher resolution textures were used. Similarly, Duh, Lin et al.'s (2002) study revealed that scene content with high resolution appeared to influence simulator sickness and sense of presence. They reported that participants exhibited greater postural imbalance and more difficulty in maintaining upright posture with a fountain scene than with a simple radial pattern scene presented at high resolution especially with wide FOV. They explained that the fountain scene provided more 2-D (monocular) depth cues, more up-and down polarity cues and more meaningful information than the simple radial pattern scene. They concluded that higher image resolution together with wide FOV might offer more sense of realism than low resolution image.

Additionally, it was indicated that the presence of multiple luminances yielded more depth impression than single luminance (Eggleston, Jansen et al.1996). This is expected as visual acuity increases with increase in luminance (May and Badcock 2002) which will result in better depth perception. Loomis and Knapp (2003) suggested that the compression of distance in their study was due to limited rendering quality of their VE which lacked important cues

such as natural texture and highlights. To support this hypothesis, they provide evidence of their informal observation whereby viewing the real environment with a HMD appeared more realistic in terms of distance and scale. Though they argue more research is needed to determine other factors that underlie this difference.

However, Willemson and Gooch (2002) concluded from their study that image quality has little effect on distance perception in the VE. In their studies they compared perceived egocentric distances in three types of environment: real environment, stereoscopic photographic panorama, and virtual stereoscopic computer model. Their results indicated that while there was a significant difference between real and VE, the difference between photographic panorama VE and the computer model VE was not significantly different which led them to conclude that the underestimation of distance in VE was not caused by image rendering quality.

Investigations by Thompson and Willemson et al.(in press) arrived at similar a conclusion: the image rendering quality has little effect on the perception of egocentric distance judgment. They suggest that the possible explanation for the compression of virtual space in immersive VE is the low image quality used in previous studies which fails to generate familiar size effect. The authors based this on the assumption that the effectiveness of the familiar size cue depends partly on the realism of the images. The authors investigated this possibility by comparing real world condition to three types of images rendering viewed using a stereo high resolution HMD: photo-realistic 360° panoramic images, low textured mapped computer generated images and wireframe rendering. Participants were tested on distance judgment using triangulation walking tasks. Results showed that all distance judgements in the VE were significantly different from the real world judgment. Distance in the VE tends to be largely underestimated. Comparisons among image type showed no difference indicating that distance judgments were unaffected by the image quality. They concluded that photo-realistic improvements in computer generated images such as textures and illumination might not improve egocentric distance perception. The authors further suggest this similarity in performance might be due to the hallway scaling and geometry cues available in all conditions; that is, visual angle cues might dominate perception of distance. The large difference between real and VE distance estimation were attributed to the limitation of natural viewing in the VE. The low sense of presence and ergonomic factors associated with HMD was also implicated.



### 2.9.4.3 Other factors affecting spatial awareness in VE

Besides display and system related factors described in this chapter, other variables have also been attributed towards perceptual difference between real and VE. There is evidence to suggest that participant development of spatial knowledge increases as they become familiar with the environment (Ruddle, Paynes et al. 1998). In their study, the authors asked participants to repeatedly navigate complex virtual buildings presented on a non-immersive desktop monitor display. They found that participants' route findings, direction and relative distance estimate accuracy improves with experience in the VE. However, Allen, McDonald et al. (1997) found that experience improves landmark direction but it has no effect on distance estimation accuracy.

Several researchers have compared the effect of passive and active exploration of the VE on spatial knowledge acquisition. Peruch, Vercher et al. (1995) presented evidence that active exploration promotes wayfinding in the VE. However, Wilson and colleagues (Wilson, Foreman et al. 1997; Wilson 1999) reported no difference between active and passive exploration on pointing and map-drawing tasks. Studies by Brooks, Attree et al. (1999) however showed that active participants recalled spatial layout (room plan without objects locations) of the VE better than passive participants but there was no significant difference between both groups on their recall of correct locations of objects in the VE. Waller (1999) suggested that allowing participants free exploration of the VE helped them to improve their exocentric distance estimate. These studies indicated that the superior performance of active participants is task dependent.

The effect of the interface device was also examined by several researchers. VE navigation, one of the most prevalent user actions, can be implemented using a variety of input devices: mouse, trackball, joystick, position trackers, locomotive devices, eye tracking, haptic devices (see Baldi (1997)) for an overview of these devices). It has been suggested that the choice of device could affect participant's spatial performance (Ruddles and Jones 2001). Allen, McDonald et al. (1997) compared two types of interface device: treadmill and joystick for movement and visual control in the VE. Participants were asked to make distance and direction estimation in large scale VE. They found that head-tracked HMD/treadmill condition participants severely underestimate distance compared to the non-head tracked HMD/joystick participants. The implication of this result is that due to the increase in the task and cognitive demand on user a more immersive display may not necessarily improve participant's performance.

A similar result was obtained by Witmer and Kline (1998). The authors compared three methods of movement: joystick, treadmill and teleportation. Their result showed that participants' performance using a treadmill is comparable to those using a joystick. However, it was shown that the use of treadmill induced more sense of presence on the participants. The similar performance between both groups was also attributed to the treadmill participants paying more attention to control their movement and speed and less on attendance on distance cues during travel.

Physiological cues (accommodation and vergence), pictorial cues (linear perspective, occlusion, shading and shadows, aerial perspective, retinal and familiar size, texture gradients and heights in the plane), and motion factors (motion parallax, motion perspective, optic flow) have also been attributed as factors. These factors have been intensively studied in psychology research of visual perception in the real world and only recently in the VE. Witmer & Kline (1998) suggested that pictorial cues are adequately represented in VE but deficiencies in the VE display resolution or FOV may reduce their potential as distance or depth cues. Similarly, motion cues were also fairly represented in the VE but reduced display resolution and systems lag may reduce their usefulness. However, future technology is likely to produce higher resolution and more encompassing display which lead to more realistic representation of object motion and scene translation (May and Badcock (2002).

However, physiological cues were poorly represented in VE. Stereo VE display allows presentation to each eye slightly different virtual image to create a stereoscopic image of the virtual scene. In the real world, our eyes accommodate and converge at the same point to focus an image on the retina whereas in the VE eyes accommodate at the display plane and may converge at a different distance. This conflict however may cause problems to the viewer (see Section 1.3 (Chapter 1) and Section 4.1.2.5 (Chapter 4).

## 2.10 CONCLUDING REMARKS

This chapter introduced an overview of the human perceptual system which provides knowledge on visual performance. Terms such as FOV and visual angles which are closely related to the VE display systems were also introduced. Subsequently, the perception of space in the real world through the use of various types of cues or information was presented. This included a review of the depth cues such as pictorial depth cues, physiological depth cues, binocular cues. Other factors influencing depth perception were also discussed. In addition to understanding perception in the real world, knowledge on picture perception also informed

the design of acceptable and useful VEs. Thus, a review of perception in pictures which included the geometrical perception of pictures was presented next.

A discussion of perception in VE was presented which highlighted the requirement of applications that use VE technologies to represent the real world counterparts; that is to provide similar perception in both worlds. However, current VE technologies have not been able to provide similar perceptual experience to the real world as indicated by the results of some studies. Additionally inconsistent findings make generalization of results difficult.

The importance of spatial awareness in both the real and VE was highlighted, focusing on the distance perception and spatial representation as essential and basic tasks of perception of space. A review of studies on distance perception and spatial representation in VE was presented next. For distance perception, studies reviewed revealed inaccurate perception in the VE compared to the real world. For spatial representation studies, some researchers argue it is possible to perceive spatial representation in a VE similar to the real environment while others limit this to simple environments only. However, exact reasons for the difference in spatial perception in real and VE are still unknown, thus requiring further investigations.

Finally, factors affecting spatial perception in VE were presented in detail focusing on the related aspects of VE display systems. This included a review of studies that compared factors such as display types, image quality, resolution and luminance. The focus of the studies reviewed on display types was on the comparison of spatial performance on various display types used to view the VE. Few studies have examined directly the effects of display size on spatial performance especially on distance perception indicating the need for more research. Studies that examined other factors affecting spatial perception were also reviewed. This included participants' experience, passive versus active participants; interface devices for interactions, pictorial cues, and physiological cues. Interface devices were presented briefly in this chapter. Since the choice of interface device might have an impact on user performance, this factor will be discussed in more detail in Chapter 4 with regards to the selection of input device for interaction.

Whilst in this chapter the factors affecting perception in VE are reviewed, the technological issues regarding the creation of the VE model and its implementations are however presented in the next chapter (Chapter 3). The basis for the experimental approach undertaken to understand spatial awareness in VE is presented in Chapter 4.

# CHAPTER 3

---

## *TECHNOLOGICAL ISSUES OF VE*

### 3 OVERVIEW

One of the aims of a VE is to provide a synthetic experience indistinguishable from the real world by matching the capabilities of human sensory channels (Barfield et al.1995; Durlach and Mavor 1995; cited in Pfautz 2000). Thus, many of the design decisions of VE applications, including hardware and software design, are based on the capabilities and limitation of the user (Kessler 2002). Kessler (2002) further suggest that:-

“To be interactive, a VE software application must constantly present the current view of a computer-generated world and have the world quickly react to the user’s actions. To be convincing, the presentation must provide enough detail to make the object easily recognized and enough objects to give the user the sense of being in the world. To be useful, the environment must respond to the user. The user’s location in the world should change when a navigation action is performed. Objects that the user grab or nudge should move as expected. Manipulation of three-dimensional interface elements, such as floating buttons, tabs and sliders, should have the desired effect on the environment, perhaps by changing the appearance of an object.”

Thus, in order to provide a convincing simulation, a VE system must be able to accurately emulate its real world's counterpart in terms of image and behaviour presentation while maintaining an acceptable frame rate. The production of a realistic image or model requires a detailed geometric model of the scenes and an accurate simulation of the lighting effects similar to the real world. Additionally, for increased realism, objects in the VE must behave according to the physical law (Slater, Linakis et al. 1996). However, current technology is incapable of handling such amount of information and processing in order to generate VE with high degrees of realism (Kessler 2002). To present such an environment in real-time requires very powerful computer workstations such as the Silicon Graphics workstations or high-end personal computers. In order to maintain an acceptable frame rate, most often image and behaviour realism must sometimes be compromised (Bastos, Hoff et al. 1999). Fortunately most available applications (such as training, architectural walkthrough, and entertainment) do not require such a high level degree of realism but still, creating a VE that has some degree of realism and that is convincing enough to the user is a tough challenge by itself (Kessler 2002).

Image realism may influence user's evaluation of the sense of presence (Slater, Linakis et al. 1996) which may in turn influence participants' performance in the VE. Slater and colleagues (1996) define presence as "the psychological sense of 'being there' in the virtual environment". According to Kalawsky (2000a), this ability to create a sense of "being-in" the virtual environment sets VE systems apart from other forms of media such as films and TV which are known to induce sense of presence in the environment. Other researchers suggested that presence should increase as a function of pictorial realism (Witmer and Singer 1998). While image realism encompasses the generation of accurate images with realistic behaviour, detailed discussion of image realism is restricted to the image quality of computer graphics scene as the overall aim of the thesis is to evaluate the computer-generated scenes. As such, in this thesis, image realism refers to the accurate and detailed geometric construction of computer-generated scenes that mimics accurate lighting effects of the real world. However, the lack of representation of other sensory information in the VE may reduce VE realism which may inadvertently influence participants' performances. The influence of these missing or reduced cues however is not directly investigated in this thesis but discussion of results will include their impact on participants' spatial awareness performance. In particular, the effect of sensory conflicts between visual and kinaesthetic cues will be highlighted.

The production of realistic images in static and dynamic forms is an endeavouring goal and challenge for computer graphics researchers. Recently, computer graphics techniques have

been shown to produce very high levels of realism (Vince 1995). While the creation of a VE is based on the computer graphics principles, for most applications not all the techniques used in computer graphics can be used for the real-time VE systems. In computer animation, a single photo-realistic image requires many minutes or hours to prepare, a full length movie could take days or weeks to render. The reasons for the long rendering time are the large database of polygons and the complex algorithms used to improve image realism such as various lighting models, shadow generations, texture mapping and anti-aliasing methods. This lengthy rendering time is acceptable as the process is done offline. For a VE, a system which operates in real-time, this rendering time is not acceptable as only 20 ms may be available to render an image (Vince 1995). Thus, while it is possible to create an image with high realism using computer graphics techniques, the long rendering time does not permit the use of the same techniques to create VE model with the same level of realism.

This chapter outlines the technological issues in creating and presenting a VE. Discussions of the issues are divided into two main sections. The first section describes the fundamental concept and issues in the modelling and rendering of the VE which includes discussion on the techniques and software algorithms to generate visual realism in the VE in real-time based on trade-offs between image realism and system performance are presented. The second section describes the types of VE systems, their advantages and their technological limitations. The three systems were also compared in terms of qualitative performance.

### 3.1 FUNDAMENTAL CONCEPT AND ISSUES IN VE MODELLING AND RENDERING

Basically a VE comprises of a database of modelled objects (and behaviours) and light sources. Input resulting from the user's interaction provided by the input devices will influence the state of the VE and its objects. Depending on the input, these changes are effected by algorithms such as animations and simulation procedures or collision detection algorithm. These changes are then reflected to the user via the output channels of the VE system such as the visual, audio and haptic display (Vince 1995).

A set of geometry can be used to describe an environment, its spatial relationship and interaction with users (Kessler 2002). Thus, because of its spatial nature, a VE is described and represented using a geometric database in the computer. This representation must provide enough details and contain many objects to provide the user a convincing illusion that they are in a realistic VE world (Kessler 2002). Aspects such as geometry accuracy as well as colour, texture and lighting may contribute towards realistic representation (Vince 1995) may



need to be represented. However, in order to create the VE and its virtual objects with great levels of details and realism would require significant efforts. Besides to present such a detailed environment to the viewer in real-time would require a computer with high processing power. In order to improve performance, several techniques have been developed to optimize the management and retrieval objects of a large and complex VE model database (which often comprises of millions of polygons). These techniques which include the use of a scene graphs and LOD are discussed in the next sub-sections.

In order to provide further understanding of these issues of developing and presenting VE models, details on how of VE models are created and the techniques used to improve visual realism is presented in the following sub-sections. This includes discussions on the trade-off between image-realism and system performance.

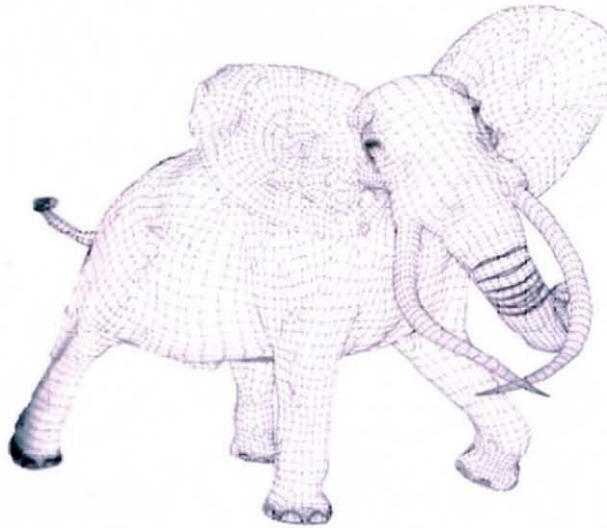
### 3.1.1 VE models

VE models are created using the techniques of 3D computer graphics. As described earlier, because of its spatial nature, a VE and its objects are usually described or constructed using geometric sets of polygons, lines and also text images. Text images are often treated as special objects but can also be represented using lines or polygons. Polygons, also known as faces, are flat surfaces which have at least three edges or lines. The corners are referred to as vertices. Each vertex has three coordinates: x, y, z. A polygon has two sides but only one side is visible unless specified otherwise. This has an advantage as it reduces the number of polygons to be rendered. Graphic systems may only be able to render convex polygons (triangles) as they are easy to process whereas concave polygon (polygons with four or more edges) are often converted into triangles before being rendered (Kessler 2002).

In addition to polygons and lines, most modelling systems also allow creating geometric objects such as spheres, cubes, cones or cylinders. These objects may be stored as their shape or converted into polygons. Many of these objects take longer to render compared to polygons and lines. As current computer graphics systems are capable of rendering millions of polygons per second, these sets of geometry is often decomposed into sets of polygons and lines by the rendering systems (Kessler 2002).

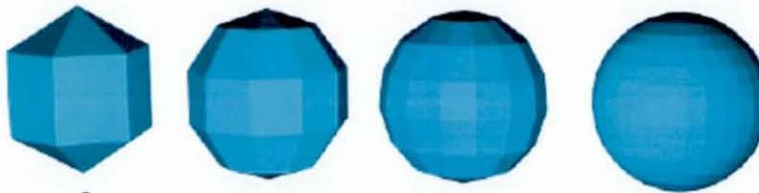
A very accurate representation of an object requires a high number of polygons and lines. As illustrated in Figure 3-1, the more complex and the more detail the object is the higher number of polygons counts needed to realistically model it. Thus, high image realism in object appearance comes at the price of more processing time to render the image due to the

high polygon counts requirements. This in turn may adversely affect systems performance in terms of update rate. However, using a technique (to be described in the later part of this section) called texture mapping, it is possible to create high realism in objects using texture mapping techniques with low polygon counts. This technique is based on projecting photographic images (textures) onto polygon-based objects.



**Figure 3-1 A polygonal 3D elephant**

(Image courtesy of Viewpoint Datalabs, Adapted from Vince 2000)

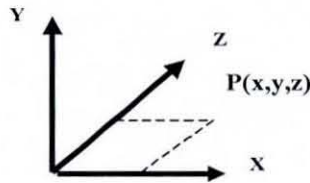


**Figure 3-2 Spheres of varying level realism based on the number of polygons used**

However, to make the surface appear smoother and continuous, as illustrated in Figure 3-2, more polygons must be used. But increasing the number of polygons would increase rendering time and consequently this will affect system performance in terms of update rate. Instead of increasing the number of polygons, alternative methods are available to model smooth objects or surface: *surface patch* and Constructive Solid Geometry (CSG). Surface patch is based on mathematical techniques to create a small smooth surface and these surfaces can be combined to form larger complex smooth surfaces. Two types of surface patches are Bézier-Spline and Non-Uniform Rational Bezier-Spline (NURBS). Both Bezier surfaces are difficult to render in real-time but in practice the model built from Bezier patches is converted into a mesh of triangles which can be rendered more quickly (Vince 1995). The CSG technique, a computer-aided design strategy, is based on the fact that some objects such as

sphere, cylinder, and ellipsoid can be described using mathematical equations. As such this technique can be used to form more complex objects based on these mathematical equations.

The VE model which is a representation of a scene (for example a room, building or town) often consists of a very large database collection of objects and its properties. While the VE model can be described using these geometric representation of polygons, lines and text, these representations need to be organized efficiently in order to facilitate the system managements and retrieval of the objects in the database. For example when an object moves in the environment, all polygons and lines related to the object must move together. This requires an efficient method of storing these representations to ensure the object's original structure is maintained. In computer graphics, one of the main principles is the use of the Cartesian coordinates system (Figure 3-3) to locate a point in space (Vince 1995).



**Figure 3-3 Cartesian coordinates systems**

This point can be used to represent the position of camera, light source, an object or a specific point on an object. Each object has its own coordinate system. The world coordinate system of a virtual world is shared among the objects in it. The camera coordinate system, having the eye or centre of projection as origin, defines the viewing volume space. It facilitates far and near clipping, to limit area where objects are visible in the scene. Thus, it represents the arbitrary position of the viewer in space.

In computer graphics, a 3D scene is organized into a data structure called coordinate system graph or scene graph (Malhorta 2002). Based on this principle, a VE model can be represented using a coordinate system graph or a scene graph. A scene graph is a collection of objects organized in a hierarchical tree-like form called directed-acyclic graph where objects are grouped according to location in the scene. Each node in the scene graph includes low-level descriptions of object geometry and their appearance, as well as the high-level spatial information such as specifying positions, animations, transformation and other application specific data.

For large complex scenes, most of the time a small portion of the model will be visible on the display at any one time. As such it is not necessary to render polygons which are not visible.

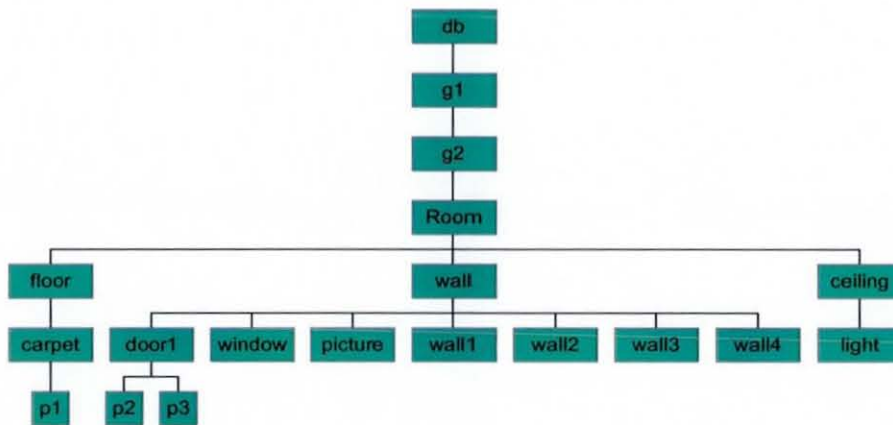


If a node in the scene graph is not visible, then all the sub-nodes of this node can be removed from the rendering pipeline, thus improving performance efficiency. Another related issue is to display parts or all parts of an object that are visible to the viewer only. The process of eliminating the parts of an object that are obscured by other objects is called hidden surface removal. Backface culling and view frustum culling are classic approaches to hidden surface removal (Murali 1999). In backface culling, any polygon whose normal is facing away from the viewpoint is considered not visible and therefore is not rendered. Since visibility of polygons is restricted to those in the viewpoint of the viewing frustum, view frustum culling approach renders only polygons that intersect with this viewing frustum direction. Other approaches for hidden surface removal include Z-buffer, Painter's algorithm and Warnock's algorithm (see for Foley, van Dam et al. 1995 for details). Since in these approaches some polygons are removed, thus their rendering can be avoided and this can improve rendering performance.

In many cases, a large complex scene may contain thousands or millions of polygons where the number of visible polygons still exceeds the rendering system capabilities and affects the interactive frame rate (Greenberg 1999). To improve the frame rate, one strategy is to create several levels of details (LOD) of an object in the database. It is not necessary to render distant objects are very small (often only a few pixels high) with very high resolution. However, when this object is near the viewer, it is still necessary to render the object with the highest resolution. Thus, an object can have several representations with varying level of details or resolution based on its distance from the viewer. Thus, for a distant object, a less detail representation of it will be rendered. While the switching of object at varying distance would add extra task for real-time systems but the overall benefits of improve rendering time is worth it (Vince 1995). This technique has been successfully implemented in flight and car simulators (Kemeny 1993).

In Multigen II Pro, a modelling software used in the author's work, structuring of the database is done hierarchically through the use of different modes (group, object, polygon, edge and vertex) when creating any element in the model, with the following defined order: groups are made of objects, objects made of polygons, and polygons are made of edges and vertices). In this software, the LOD technique can be done automatically. Figure 3-4 shows an example of scene representation of a room database. REALAX RXScene, another modelling software used in this thesis, also used a hierarchy tree structure with branch nodes includes further nodes such as light, sound, camera, LOD, and Dynamic Coordinate Systems (DCS) which is

used for animation. This software was used initially to introduce the author to the concept of VE modelling.



**Figure 3-4 A partial database representations for a room model in Multigen II Pro Software.**

### 3.1.2 VE Image realism

To realistically mimic objects in the real world, objects in the VE are assigned attributes or properties associated with them. Such attributes may include static or dynamic features, physical constraints or acoustic attributes, colour, lighting and texture.

Objects in the VE may be assigned static or dynamic features depending on whether they can move or not (with the exception of lighting properties). Floors or walls are examples of static objects while door and windows can be open and close thus have dynamic features. Additionally, some dynamic objects have physical constraints which limit their movements. For example, a door will only move within a certain degree of rotation.

Objects in the VE can be made to obey the physical laws of the real world. For example, objects falling at constant accelerations or object collisions that exhibit the impact of collisions such as surface distortion or movement changes (trajectory). Very accurate simulation of such objects in real time would require computers of very high processing power, which is beyond the capabilities of most current computers (Kessler 2002). For human computer interaction purposes, most systems just provide support for a small number of objects where objects may be given properties such as mass, velocity, acceleration and momentum. Newton's laws of motion provide the basis to describe the simulation of movements, collisions and force-interactions between objects (Vince 1995).

Some objects may have acoustic properties which may generate sound upon collision with other objects or a virtual radio in a virtual room may emit sound when switch on. In this thesis, the modelled objects developed by the author are static and have no acoustic



properties. However, the modelled objects were assigned other attributes such as colour, lighting and textures. These attributes are described in the next two subsections.

### 3.1.2.1 Object colours

In VE modelling, objects may be assigned colours from the colour space. A colour space is where a colour is defined such as using the three primary colours (red, blue and green (RGB)) or using the parameters of hue, saturation and value (HSV). In RGB colour space, a colour is specified using three numbers that range from 0 to 1. For example, (0,0,0) represent black and (1,1,1) represent white and other values represent other colours. Representing colour using RGB is not intuitive as it is difficult to search for a colour using this method. In HSV colour space, hue determines the colour; saturation controls the amount of light in the colour and value represents the lightness or darkness of a colour. Both RGB and HSV are closely related and many modelling software provide the user with both colour spaces to get the advantage of both.

Assigning objects in this manner gives the objects fixed colour for all surfaces, that is, when it is viewed from any angle the colour remained the same. But in the real world, this is not the case. Thus, in order to make the objects look more realistic the surfaces may be assigned with different colour shades so that it looks as if it is illuminated by some light source. To achieve this, accurate simulations of lighting effects is required. This involves complex simulations of light interaction with the coloured surfaces such as reflections, refractions, interference and interaction using mathematical equations. Several illumination models, reflection and shading models have been developed for this purpose (Vince 2000). A brief description of these models is presented next (see Vince (1995) for more detail descriptions of these models).

#### Illumination models:

The purpose of illumination models is to illuminate the virtual world by simulating the interaction of the light sources with the coloured surfaces of objects.

- **Point light source.** It radiates light in all direction, for example a light bulb. The intensity of the light can be specified in terms of the RGB or HSV colour space.
- **Directional light source.** As its name implies, it emit light from one direction and assumed to be located at a far distant (such as the Sun) and the lights rays are also assumed to be parallel.
- **Spot light source.** This simulates the characteristics of directed beam of light from a spot angle for angle of illuminations.



- **Ambient light source.** It is the background light level which has colour and intensity but no direction. It is included in the lighting calculation as a constant and typically accounts for 20-25% of the total illumination.

The above mentioned light sources are based on one light source only. Though it is possible to have multiple light sources as in the real world, but the demand on the computing power would be great due to the problem of light balancing where some surfaces are over illuminated and some surfaces are under illuminated. The increased in complexity of simulating these illuminations would in turn increase rendering time.

### **Shadows**

For most objects light cannot pass through them and because light travels in straight lines surfaces facing the light sources will be bright and surfaces away from the light sources will be in shadow (Coren, Ward et al.1999). Like other objects in the VE, shadows need to be represented and created. The presence of shadows in a VE would increase the perceived realism of the VE (Malhorta 2002). According to Slater, Usoh et al. (1995), shadows provides “alternative view of objects and provide direct information about their spatial relationships with the surfaces. It has been empirically shown that shadows were significant cues for certain performance tasks (Hu, Gooch et al. 2002, Hubona, Wheeler et al. 1999, Wanger, Ferweda et al. 1992).

Simple object shadows can be modelled using a set of polygons. This technique is easy to implement but the drawback is the shadow created is less realistic because the shadow has sharp edges.

In computer animation, various techniques have been developed to create shadows. One technique is called *ray-tracing* where a sharp shadow is produced. *Softbox lighting* technique (Vince 2000) creates a more realistic shadow compare with ray-tracing but it requires more computing time. While realistic shadows can be achieved using these techniques, however they are still difficult to implement in real time due to computational overhead (Vince 1995). Thus, a false shadow is used by creating a shadow polygon which can move with object movement but does not change shape with changes in surface geometry. In this thesis, this technique is employed in the modelling of objects' shadows because of its simple implementation. Moreover, Hubona, Wheeler et al. (1999) indicate that the presence of shadows aid in the performance of estimating object height and depth accurately but shadows' sharpness (accurate rendering) and shadows' shape (simple polygonal shape verses true shadows) does not influence perception of object size and position.

## ***Transparency***

In the real world, some objects, such as glass, are transparent. As such, this attribute may be assigned to such objects to create a better sense of realism. Transparency effects can be simulated with varying levels of realism. Besides showing that object can be seen through another, appropriate colour and intensity changes, reflections and refractions could also be simulated though the inaccurate optical effects may be noticeable (Vince 2000). In addition to modelling transparent media, variable transparency can also be used to fade out an object model description and bring in another object description (such as object's LOD described earlier). Instead of sudden removal of object, this allows for smooth transition between objects (Vince 1995).

## **Reflection models**

As the *illumination models* mentioned earlier is used to illuminate the VE, the *reflection models* are used to describe the reflective behaviour of the light in order to create more realism in the VE image.

- ***Diffuse reflection.*** The lights reflected from rough surfaces, such as carpets are reflected in all directions. These surfaces are called diffuses surfaces as it exhibit reflection properties which radiate lights in all directions. The brightness of the diffuse surface is independent of the viewing angle but it is proportional to the angle of the incident light. Thus, when the angle of incident is large the reflected light is dim but when the angle of incident is zero the reflected light is bright.
- ***Specular reflection.*** This reflection describes the reflection of lights from any polished or wet surfaces with specular highlights. The nature of specular reflection depends upon the reflective nature of the surface. It can be clear and precise (for example mirrors) or it can be less distinct and cover a small area (for example metallic surfaces). The specular highlights which are dependent upon the relative position of the observer to the surface is readily simulated and its size reflect the type of surfaces. These highlights could be simulated in computer graphics and show surfaces with gloss factors.
- ***Multiple diffuse reflections.*** In the real world, multiple diffuse reflections occurring between surfaces produce an effect call soft shadow effects. Additionally, the colour of objects can affect the colours of other objects. To create a VE that mimics these occurrences in the real world with more realism, a technique called *radiosity* is employed. It is a global illumination model that attempts to simulate multiple reflectors that occurs between surfaces. To compute the changes of illumination across a surface, the surface is converted to a mesh of small patches. A realistic scene is then created by computing light intensities for each patch. Higher degree of realism can be achieved by reducing the patch



size but at the expense of increasing the computational time thus affecting the update rates. As the light is diffused the rendered scene is independent of the observer. Thus, the processing of the *radiosities* can be done offline and rendered in real-time. But changes in object positions would require new processing of the *radiosities*.

## Shading models

Shading models or algorithms are used to describe how polygons or surface patches to be shaded to make more realistic image. Shading is done to adjust colour to depict accurate lighting (reflections and refractions) and textures of objects. This task is done polygon by polygon or pixel by pixel and can be a time consuming process. However, there are several shading techniques available which offer tradeoffs between rendering time and photorealistic image.

- **Flat shading or Lambert shading** is a simple shading algorithm which is based on the approach that an entire polygon is assigned a single colour. While it is simple to draw, the resulting object has a faceted appearance which reduces image realism. Increasing the number of polygon might compensate for this effect but at the expense of more computation time.
- **Gourand shading.** In this method, a colour calculation is made at each vertex of the polygons to get an average normal vector. These average normal vectors are then used by the illumination model to calculate reflected light and because neighbouring polygons share common average vertex normal, the boundary edge disappears giving the object the apparent smoothness. As such faceted objects in flat shading will look smooth using this algorithm. This algorithm is considered the fastest smooth shading algorithm but it is less realistic than Phong shading. However, using smaller polygons may make it approximate Phong shading. But this will inadvertently increase the number of polygons which in turn will increase demand on processing power. Thus, Gourand shading will create a realistic image of objects that consists of many polygons. For objects that is made up of few polygons it will not be rendered realistically.
- **Phong shading.** This algorithm assigns a colour for each pixel of a polygon by interpolating the angle of incidence and recalculating the correct colour for each pixel as done in Gourand shading. This technique results in a more smoother and realistic image but at the expense of additional computation overhead. Phong shading algorithms can handle texture mapping properly but it cannot handle real reflections and refractions.

While the images produced are not of outstanding realism, both Gourand and Phong are considered acceptable for most applications. Other shading techniques include ray tracing and radiosity (mentioned earlier). Ray tracing provides the most photorealistic image but it is the

slowest of all shading algorithms. While it is useful for rendering single image, the slow rendering speed makes it unsuitable for real-time applications. Similarly for radiosity algorithms, while it has been designed to imitate multiple diffuse light reflections in the real world, the computations speed is very slow. Radiosity technique has been successfully used to create the most impressive computer images of large interior architectural structures (Watt and Policarpo 1992, cited in Malhorta 2002).

### 3.1.2.2 Texture mapping

Surfaces of objects in the real world may have different texture such as rough, smooth, and bumpy. Some surface may reflect different kind of patterns such as brick wall, carpet or woven pattern. To model such surfaces realistically would take a considerable amount of effort and time but a technique called texture mapping makes this possible. Texture mapping is a quick way to increase image realism using texture maps (Catmul 1975, cited in Weinhaus and Devarajan 1997). Texture maps are 2-D images which can be taken from photographs or can be created using any paint program (Vince 2000). In texture mapping the surface of the object is covered with these images to create the realistic look. This method can also be used to realistically portray complex surface characteristics such as bumps, dimples, embossed or woven patterns without the need to model them.

It is important to match the size of the texture map to the projected polygons. If the texture map is smaller than the projected polygon, the maps can be repeated like a tile to cover the entire polygons. This method is used in this thesis to cover large area of objects (such as grass field and sky) entirely. The texture image size used in hardware accelerators is restricted to  $2^m \times 2^n$  texels (texture element) or sometimes  $2^m \times 2^m$ , where  $m$  and  $n$  are positive integers. The reason is to make efficient use of space available in texture memory. For some graphic accelerators there is a limit on the amount of the texture that can be used. Typically texture memory is limited to less than 100Mb (Costello and Bee 1997). As performance limitation can occur when textures are swapped in and out of the memory, the designer should consider the right size of texture to be used so that less memory is used but at the same time detail or resolution of the image is maintained. Thus, the trade-off here is between texture resolution and performance. In MultiGen II Pro software, to make efficient use of the texture memory and to prevent unwanted side effects (that is for texture to be properly displayed) the dimension of the texture need to be sized to power of 2 (Example: 2, 4, 8, 16, 32, 64, 128, ...).

Because objects' surfaces differ in shapes, several projections techniques are used to project the texture map to the objects' surfaces. Such techniques include cylindrical, spherical, and



radial. Thus, for a sphere a spherical projection can be used. Figure 3-5 illustrates the different texture mapping projection techniques.

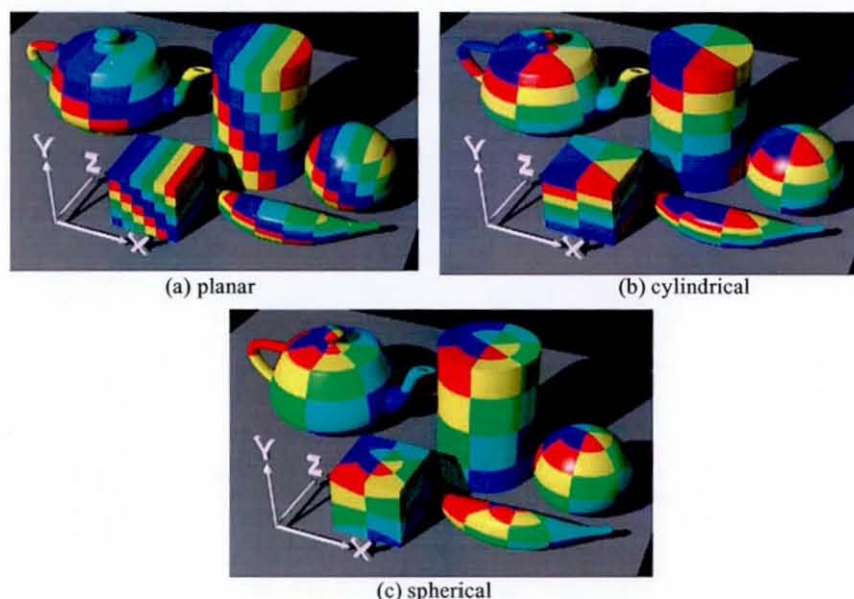


Figure 3-5 Examples of texture mapping. Adapted from Wolfe (1997)

The texture mapping described earlier is useful for a near object because it provides extra detail of the image to the viewer. However, for far objects this extra detail should not be reflected in the image as this would reduce realism. A technique called *MIP mapping* or *MIP texture* was proposed to solve this problem by Lance Williams (Vince 2000). In this strategy, instead of using a single texture map, a set of texture maps of different resolutions was used. An algorithm is then used to automatically select which resolution is used to cover objects' surfaces at different distances. This technique has a further advantage of reducing aliasing effects that occurs when texels are mapped onto screen pixels. Aliasing effects are a form of image degradation where edges (especially when there is high contrast) appeared jagged. Textured materials which contain fine regular details will also exhibit aliasing effects where a shimmering or swirling effect will occur when the texture moves. Aliasing effects can cause annoying effects and degrades image quality (Vince 2000).

Two other types of mapping which could be used to further enhance realism are environmental mapping and bump mapping. Environment mapping simulates the effect of polished surfaces that reflect their surroundings whereas bump mapping uses a texture map to modulate the way light is reflected pixel by pixel.

Texture mapping technique can be used to model objects in the distance by using an image of a scene. This technique, often called billboard geometry is only suitable for objects that are

very far away in which the absence of depth is not noticeable. This technique allows for a more realistic presentation of distant object without the need to model them. Thus, this reduces the number of polygons to model the objects which in turns reduce rendering time and improve system performance.

However, the technique described in the previous paragraph does not work for objects that can be viewed in all directions. For this, another strategy called billboarding is used. Similar to using billboard to represent an entire scene in the distant, this technique uses a picture of an object. The image of a picture is placed onto a planar surface (with background transparent effect) and the planar surface is then given rotational transformation properties so that it will always face the user giving it the impression of a solid object. This technique avoids the modelling of complex objects. Instead of using many polygons to model an object such as a tree, only one polygon is used with the texture map of the tree projected onto it. This results in not only huge saving in modelling time but also in memory and processing speed. As such both techniques were employed in this thesis to create realistic distant background scene and realistic models of trees in the VE models.

### 3.1.3 Viewing/simulation of the VE model

#### 3.1.3.1 The scene graph systems

The scene graph systems comprise of the scene graph itself and a set of scene graph software (Rahmat 2000). As defined earlier, a scene graph is a collection of nodes representing objects and its properties and other information organized in a hierarchical tree-like graph. Besides storing geometry for visual culling purposes, the scene graph needs to be managed in order to enable geometry to be extracted and created effectively and efficiently without compromising on the systems performance. Scene graph software refers to a set of software tools that are used to build and interact with the scene graph. It is designed to optimize for rendering performance.

The scene graph systems are based on two phases of operations. First the graphics application creates and loads the data into the scene graph, and then the system renders the contents of the scene graph into an image. Scene graph systems function to facilitate rapid applications development. In addition, the scene graph systems provides for the management of details in the database such as clipping planes, view-port controls and clearing of buffers.



### 3.1.3.2 The rendering pipeline

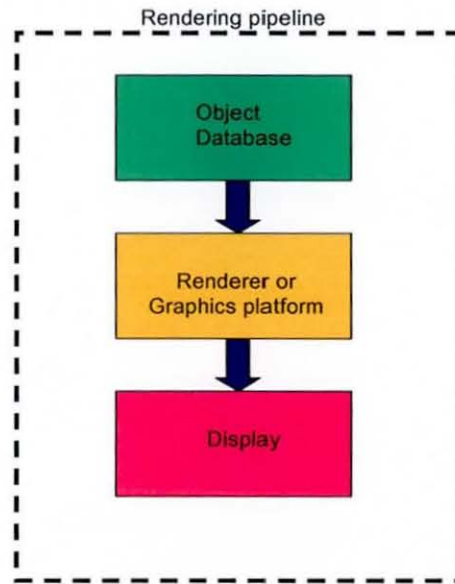


Figure 3-6 Rendering pipeline (Adapted from Bethel (1999))

All graphics applications have a number of common components: a graphics platform (rendering engine), object database and display device (Bethel 1999). The graphics platform transforms the mathematical descriptions of surfaces into array of pixels or images that can be viewed by the user. The object database (scene graph) is a data source for the rendering engine. Altogether, the rendering engine, the object database and the display device make up the rendering pipeline (Figure 3-6).

Thus, when the virtual database to represent the VE is created, a viewer software is used to load the database to be displayed to the VE model to the user. The graphics pipeline takes input (description of the scene) from the viewer software, perform rendering processes and finally output the scene on the display device. Generally, many graphics workstations have a high performance graphic pipeline built into the hardware architecture.

In this thesis, the Silicon Graphics Inc. (SGI) Performer PERFLY is one example of a viewer software used to view the VE models. PERFLY is a basic visual simulation application that can load, store, and display the scene databases. In Performer the description of the virtual scene database is represented by a tree of node called a scene graph. Each node is either an object or a set of objects. The nodes in the scene graph are arranged in a hierarchy. The hierarchy of the scene graph specifies the order in which the nodes are processed by a traversal. Rendering the virtual scene in Performer occurs in three stages (SGI Performer):

1. APP – updates the location and look of geometries and updates the viewing location and orientation
2. CULL – determines which geometries in the scene are visible (in the viewing frustum), taking occlusion into account
3. DRAW – renders all visible geometries

The rendering process is carried out once per frame. Even with very fast hardware system latency will always exist due to the need to process each of the above stage sequentially. However it is necessary to maintain a level such that real-time performance is not affected. In order to consider performance trade-offs designer need to examine the database model with respect to the graphic pipeline (Costello and Bee 1997).

### 3.1.4 Image Realism versus System Performance

According to Malhorta (2002), the key to realism is “the complexity of the scene in terms of the geometry of the model and in terms of how the interaction of light in the virtual world simulates its real-world environment.” However, the more complex the models the more polygons are required to model it and this means more computational effort is needed to generate the image (Costello and Bee 1997). Whilst the use of the illuminations, reflections and shading algorithms and other procedures would add more realism to the VE images, these would also increase demand on processing power to execute the algorithms. As put forward by Green and Sun (1995), “...in reality accuracy comes with a price, usually increased display time or memory usage.” While high image realism is certainly attractive and desirable, the success of VE application however does not depend only on the quality of the images presented but also in the naturalness of the simulation. As a VE allows user to interact with it in real-time, a prompt, fluent and synchronized response of the system would be essential for a naturalness of the interactive environment (Marzuryk and Gervautz 1996). As the user moves through the VE, the size and the perspective view of the scenes and objects must change accordingly. Thus, the process of rendering and response by the system must occur fast enough so that the user will not perceive the changes between images presented (Malhorta 2002), otherwise, the naturalness of the interaction will be compromised. Ideally, a refresh greater than 25 Hz is required but available systems have frame rates from 10 to 60 Hz. Refresh rate refers to the frequency which the display hardware can draw the image on the display surface. Another issue is frame rate, which is the problem of quick rendering of a complex model. It is the rate at which new updated scene is prepared for drawing to screen. Ideally we would want the frame rate to be the same as refresh rate (Helman 1993). According to Barfield (1995), frame rate of 15Hz seems sufficient to fulfil the sense of



presence in VE but higher values (up to 60Hz) are preferable (Deering 1993). Insufficient frame rate would result in artefacts such as jerky motion, reversal of motion, multiple images, shimmering edges and many others (Crow 1977, Watt 1989: cited in Pfautz 2000). These artefacts are present in cinema films whereby the frame rate is only one-third of the refresh rate. This means a movie frame is displayed to the projector frame three times before the next frame is available.

Frame rate depends on scene complexity (Pfautz 2000). The larger and more complex the database the more demand is on the computing power to render the scenes in real-time. Generally, the use of the illuminations, reflections and shading algorithms would provide very high realism in images but many of these are computationally expensive and result in an increase in rendering time (Weinhaus and Devarajan 1997). Long rendering time will affect image frame rates, which in turn may affect the refresh rate which is the smoothness of the simulation. Thus, for most real-time applications, it is not possible to implement all these algorithms to yield the high level of image realism. This implies frame rate is influenced by the scene complexity and detailed representation which includes polygons counts, image quality or resolution, use of algorithms. Thus, the decision of the choice of methods and algorithms used to improve visual realism is often a trade-off between computational cost and decrease in frame rate (Pfautz 2000; Marzuryk and Gervautz 1996). For most interactive applications, frame rate is more important than visual realism. As such, for these VE systems to achieve an acceptable frame rate in real time significant compromise in realism needs to be made. Generally, degradation in system performance is often unacceptable especially for real-time activities; as such slight decrease in image realism is acceptable.

Another issue is a varying frame rate. Depending upon the complexity of the scene to be rendered at any particular time, the graphics system might update the image at varying frame rates. If the scene is complex (contains objects with lots of texture and shadows), the update rate may be low and if the next scene to be rendered is less complex the update may be higher (Slater and Usoh et al. 1995). This might cause discomfort such as visual stress or fatigue to the user. Studies have shown that a slower constant frame rate rendering would be preferable to faster variable frame rate (Helman 1993).

As discussed earlier, the use of techniques such as texture mapping, LOD, scene graph may help improve system performance. Texture mapping helps reduce model complexity by reducing polygon counts. LOD techniques improve system performance by efficiently reduce polygon counts during the rendering process (see Section 3.1.1). A spatially organised scene

graph (that is objects are grouped according to their locations in the scene) helps facilitate the retrieval and management of the database. To further improve performance, one consideration would be to implement some of these algorithms in the hardware (Lastra 1995).

Another consideration is, while it is possible to generate very realistic image this does not guarantee that the images will be displayed with realistic visual appearance. This is because the display technologies have fundamental limits on the display process in terms of spatial resolution, absolute and dynamic luminance range, and color gamuts (Greenberg 1999). The spatial resolution of most displays does not reach the limits of the human visual acuity or spatial positioning ability unless the distance of the viewer from the display is increased (May and Badcock 2002). Increasing the viewer distance would result in reduce FOV. It has been suggested that wide FOV provides the user with additional information for visually guided behaviour and give the user the illusion of self-motion in the VE which in turn may increase user sense of presence (Hettinger 2002). Thus, the design decision is to consider trade-off between FOV and spatial resolution. The resolution for large projection screen must be increased and for HMDs the FOV size and resolution need to be increased. In terms of luminance the range producible are small relative to the range that can be measured in the real scenes and with regards to colour displays are limited with the range of reproducible colours (Greenberg 1999). However, despite the limitations, history has shown that display devices have succeeded in creating acceptable visual representations of scenes such as in pictures and cinema (Cutting and Vishton 1995).

In next section, the types of VE which correspond to the types of device used to present the VE model are described. The merits, drawbacks and limitations of each system are discussed.

## 3.2 TYPES OF VE

In 'Glimpses of Heaven, Visions of Hells', Meredith Bracken provides an illustrative description of the available VE systems:

"Viewing 3D graphics on a 2D screen is like looking into the ocean from a glass-bottom boat. We see through the window into the environment; we experience being on the boat. Looking into a virtual world on a stereographic screen is like snorkelling. We are at the boundary of a three-dimensional environment, seeing into its depths from its edge; we experience being on the surface of the sea. Using a 3D display with a computerised glove allows us to reach through the surface to touch objects within our grasp, while viewing our activity from outside the environment; our hands dabble in shallow water. Entering the multi-sensory world of VR is like wearing scuba gear and diving deep into the sea. By immersing ourselves in the underwater environment, moving among the reefs, listening to



the whale song, picking up shells to examine, and conversing with other divers, we participate fully in the experience of exploring the ocean. We're there."

Meredith Bricken (quoted in *Glimpses of Heaven, Visions of Hell* by Barrie Sherman and Phil Judkins, UK pbk, 1992).

This illustrative description could also be used to describe the various types of VE systems available. Based on Bricken's description, the VE can be categorized based on the various types of interface devices (input and output) which in turn provide the user the levels of immersion and interaction with the system into three categories. The first category is the non-immersive which is synonymous to viewing an ocean through a glass-bottom, where no interaction is allowed. The next level is the viewing of 3D images that allow object manipulations through the use of input devices such as gloves. In this second category, often referred to as semi-immersive system, users were not completely immersed in the environment. They are still aware of their surrounding and can interact with the virtual world with a glove from outside. Scuba diving comparison represents the third category of VE's fully immersive system where the user feels "being there" in the VE. In this system, users experience the feeling of being inside the virtual world and can interact with objects in it.

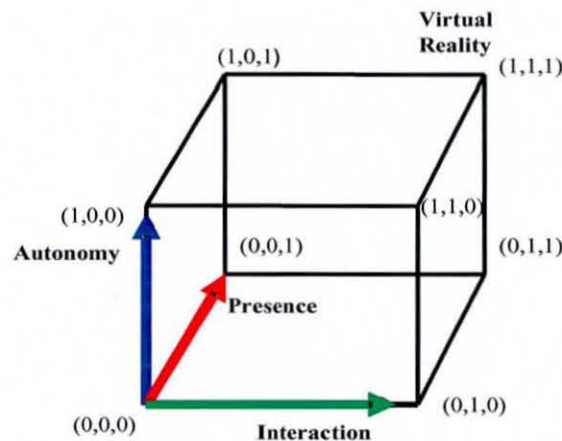


Figure 3-7 Zeltzer unit cube model (AIP cube) (1992)

Zeltzer (1992) proposed a unit cube model (AIP cube) to measure and compare a VE system (Figure 3-7). The proposed model is based on three basic properties to determine the level of a VE system: autonomy, interaction and presence.

- Autonomy – ability to react to events and stimuli (between objects, user and the environment)
- Interaction - the degree of access to the parameters or variables of an object
- Presence – number and fidelity of sensory input and output channels

Thus, according to this model, different VE systems may have different combinations of these properties, depending upon how advance it is. From this cube, it is implied that a 'perfect' VE system would have very high levels (1, 1, 1) on all three properties. Thus, a desktop system may be low in presence where as a very high end VE system may be high all three properties. This thesis uses this approach of describing VE as having varying levels of these properties to refer to various types and levels of VE systems available. However, this definition encompasses a wide range of systems; thus, we limit our definition to include computer generated images.

An ideal VR system should be able to provide all the human sensory cues. However, the current system is far from this ideal system. Due to the dominance of the visual sensory channel, most research has focused on the presentation of visual information to the user. It has been suggested an ideal visual display should have high resolution, high update rate, wide FOV, high brightness and contrast (Marzuryk and Gervautz 1996). However, most current available displays do not have a high rating on all of these features. The display of other senses such as sound, tactile and haptics is just recent. There is no taste display in existence yet. In this thesis, the discussion is restricted to the visual information display.

Different VR applications require different types of input and output devices to interface with users, thus the type of VR system may be classified based on the devices used (Isdale 1998). According to Kalawsky (1998), most VE systems fall into the following three main categories:

- Non-immersive system
- Semi-immersive system
- Full immersive system

Kalawsky (2000a) further refers to the term immersion as "the extent of the peripheral display imagery". Thus displays that present a full 360° information space are referred to as full immersive systems and displays that have an extent of less than this are grouped as semi-immersive, while desktop VR systems are refers to as non-immersive systems. Others researchers further suggest immersion includes the extent in which the computer displays are *extensive*, *surrounding*, *inclusive*, *vivid* and *matching* (Slater, Usoh et al. 1996). *Extensive* means the extent of how many is the sensory systems is accommodated, *surrounding* is the extent to which information is received by the sensory systems, *inclusive* to mean the extent that all external data are excluded, *vividness* means the variety and richness of the sensory information generated and *matching* to refer to the matching of user's proprioceptive feedback and the information generated on the displays. These systems are described in the



following sub-sections. However, even if the boundaries between categories are becomingly blurred, it is still a useful method of classifying all variations of VE systems.

### 3.2.1 Non-immersive systems

Based on Kalawsky (2000) categorization of VE systems, non-immersive systems, sometimes referred to desktop systems or Window on World (WoW) systems, would represent the least immersive of the three systems. The VE which is displayed on desktop monitors or projected displays may be presented through stereo or monoscopic viewing. User's interaction with these systems is by conventional means such as keyboard, mouse, trackball or spaceball. The advantages of such systems are cost effective as they do not require very high graphics performance; no special hardware required and can be implemented on high specification desktop PC. However, the drawbacks are the systems provide poor spatial interaction, it suffers from reduce FOV effects such as lack of peripheral vision (Pfautz 2000) and give the user less sense of scale due to the image size. Additionally, these types of systems provide almost no sense of *presence* or 'being there' and are restricted by the interface devices. Users who are highly present would experience more engaging reality and consider the displays as places visited not as images seen (Slater, Linakis et al.1996). Prothero and Hoffman (1995) found that subjects reported a significant higher sense of presence with wider FOV. Thus, desktop system may provide users with lower sense of presence because of the narrow FOV afforded. However, other researchers reported that the immersive factor do not influence participants performance on training transfer knowledge (Youngblut and Huie 2003). They further indicate that there is no difference in sense of presence during training in non-immersive desktop display and immersive projected display as reported by the participants. Thus, the latter study indicates that user sense of presence may not be affected by their sense of immersion.

### 3.2.2 Semi-immersive systems

Semi-immersive systems are typically projected VR characterised by a fixed, wide field of view, large display. Semi-immersive displays does not offer the user an all-encompassing display image but depending on which display system are being used it could provide a wide FOV of up to  $270^{\circ}$  (Kalawsky 2000). Panaromic projections or room systems (e.g Reality Center<sup>TM</sup>), Wall systems (e.g. Immersive Wall), desks systems (e.g. ImmersaDesk) are examples of semi-immersive systems (Figure 3-8). Santos, Bacoccoli et al. (2003) provides a comparison among these systems in terms of important features and potential applications. The desks systems, also known as workbenches, were developed to fit into an office or lab and are suitable for small group work application (2-5 persons). The wall systems allow

presentation in a big flat or curved screen display where models can be shown up to a 1 to 1 scale. These systems are often used to facilitate communication process in conferences, offices, public exhibits and laboratories. Two or more edge blended projectors could be used to provide a high-resolution seamless image. The room systems, primarily designed for collaborative works with massive data sets are normally driven by high power supercomputers. These systems, also known as reality centres, are common solution used by oil and gas companies. The number of viewers accommodated range from 10 to 120 persons and the screen types can be a single rear projection spherical or cylindrical plane.



**Figure 3-8 Examples of semi immersive systems: Immersadesk(left) located in SL 239 at Indiana University - Purdue University Indianapolis,USA; panorama systems (middle) developed by Panoramatech at [www.panoramatech.com](http://www.panoramatech.com); ImmersiveWall (right) located at AVVRC, Loughborough University, UK**

Due to the large FOV, semi-immersive systems give the user a greater sense of presence than non-immersive system and it also gives the user a better sense of scale because of the larger screen size. Another advantage of these types of systems is they allow sharing of virtual experience among a small group of users. However, despite this, transfer of control between users is one of the issues that must be considered (Costello 1997). Currently, the viewpoint of the VE is singly controlled by the leader of the group. Other users' (in the group) view of the scene is restricted to what the leader of the group see. Comparatively, the resolution of the semi-immersive systems can be far greater than fully immersive systems such as Head Mounted Displays (HMD); however, multiple projection systems are needed to achieve higher level resolutions. Better resolutions would determine the quality of image displayed in terms of the colours and textures.

There are several drawbacks to semi-immersive systems. With the exception of desk systems, due to the size, large space requirement is required to house the display systems. The projected image might need some distortion corrections to display the image correctly on the screen. Moreover, there are problems with choice of interaction devices for these systems.



Besides the more familiar devices such as joysticks, trackballs and 3D mouse, other available devices include wands and data gloves. Depending upon the types of applications, an interaction device used in one application may not be suitable for another. In addition to being more costly, setting up a projection is more difficult compared to a desktop system (Costello 1997). However, despite these drawbacks, semi-immersive applications (thus employing semi-immersive systems) represents one of the most interesting and cost-effective solutions for virtual reality (Persiani and Liverani 2000).

Although, it can be viewed monoscopically, stereo images are possible in semi-immersive systems using LCD shutter glasses (Figure 3-9). Stereoscopic effect is achieved when the graphics computer alternately display left and right view of the VE to both eyes respectively. When the left image is displayed, the glass blocked the image on the right eye, thus allowing only the left eye to view the image. When the right image is displayed, the left image is blocked (the lens is switched off) and the right eye lens is switched on to allow only the right eye to view the image. The switching of display between left and right images happens very fast (120Hz) that it is undetected by the user such that the resulting image is perceived as a single 3D image.

Whilst a shutter glass is less cumbersome (compared to HMD, presented in the next section), however, it is restricted in FOV and requires a very high frame rate for rendering of both left and right images.



**Figure 3-9 Shutter glasses**

As semi-immersive systems offer a more practical solution for VE applications, these types of systems will be employed in the three experiments (Chapter 5, 6 and 7) reported in this thesis. The semi-immersive systems used are based on large rear-projected walls systems. Rear-projection systems have the advantage of avoiding the projector to cast user's shadow on the screen especially when working at close range to the screen. This feature is particularly useful for our experiments as user may be placed at close distance to the screen. Further justifications for the choice of these systems are discussed in the Section 3.2.4 of this chapter. For interaction device, this thesis employed and compares two devices (a mouse and a trackball). These devices are described in more details in Section 4.1.2.4 of Chapter 4.

### 3.2.3 Fully Immersive Systems

Fully immersive systems are characterized by wide FOV of  $360^{\circ}$ . These systems provide the user with the most direct experience where users are immersed in the VE. Examples of fully immersive systems are Head-mounted display (HMD), head-coupled displayed such as the BOOM display, and the CAVE system.

A typical HMD (Kalawsky 1993) has two small display screens located a few centimetres from the viewers' eyes (Figure 3-10 (a)). The images displayed on these screens may be the same for binocular viewing or the images may be slightly different for each screen for stereo viewing. It is also possible to have monocular viewing using only a single display screen. A motion tracker is used to track the user's head and allows the computer to adjust the scene to the current view of the user. This gives the user the feeling of looking around and walking in the VE because the images presented to the user is based on his/her current position and orientation.



**Figure 3-10 Examples of fully-immersive systems**

The BOOM (Binocular Omni-Orientation Monitor) is a head-coupled stereoscopic display device (Bolas 1994). The high resolution displays and the optical systems are placed in a box attached to a counterbalanced arm (Figure 3-10 (b)). The user views the VE by looking into

the box and can move the box to any position. The BOOM display provides accurate head tracking but it is only a single user experience with restricted range of movements. One advantage of BOOM display over HMD is that it removed the weight of the HMD from the user's head to the mechanical arm.

The CAVE<sup>TM</sup> (Cave Automatic Virtual Environment) was developed at the University of Illinois at Chicago (Cruz-Neira, Sandin et al. 1993). It is a multi-persons, high resolution image and audio room which provides the illusion of immersion by rear projecting stereo images on the walls and floor of a room-sized cube (Figure 3-10 (c)). Several persons wearing lightweight stereo glasses can enter and walk freely inside the CAVE. However, the correct perspective and projections of the images were adjusted accordingly to one viewer's (the leader) movements who wears a head tracking system. While it allows multi-person views and non-encumbering, it needs space for the display systems and moreover it is very costly to acquire.

The fully immersive systems described in the previous paragraphs are characterised by a wide field of regard (360°) and visually coupled, that is the user's view is updated whenever he turn his head to look at any direction. The VE is often presented in full scale and relates to the human size. These give the user a sense of presence greater than non-immersive and semi-immersive systems. To further enhance the sense of immersion generally includes haptic devices such as datagloves to allow the user to feel the simulated objects, 3D tracking systems such as 'Flocks of Birds' or Fastrak system allows tracking of the user limbs and head, audio display for the sound effect and other non-visual devices. However, the sense of immersion provided depends on several factors such as FOV of HMD, resolution, update rate, contrast and illumination of the display (Costello 1997). For HMDS and BOOM, the trade-off is between large FOV and resolution in which large FOV would result in lower resolution display. It is noted that the CAVE systems is categorized as a semi-immersive system if less than six of the sides of the caves are used (Kjeldskov 2001). This is because the available field of regard is less than 360°.

While fully immersive systems provide users greater immersion and sense of presence in the VE, there are several drawbacks to these systems. As HMD are worn on the user's head, it should be lightweight and comfortable to wear but often this is not the case. HMDs are often heavy, this weight and position of the HMD might cause strain to the user's head, neck and spine (Stanney, Mourant et al.1998). As such lower weight and lower resolution HMDs are often used (Kalawsky 1993). However, this will lower image quality and image realism and



may subsequently affect performance. Though, the BOOM display avoids this weight and strain on the user, both HMD and BOOM is not easy to use. Additionally, both displays are limited to single person use only.

Another drawback of fully-immersive systems is that they often suffer lag from head and hand tracking to scene change caused by scene rendering time and handling of immersive input devices. Lags represent the time between the user initiating an action and the action actually occurring in the VE. Significant lag would results in slow update of display and users have to wait for images to appear. This may reduce realism and may subsequently affect the user's sense of presence and performance (Barfield 1995, Reddy 1994). This may also be disturbing to the user and may cause motion sickness.

Additional problems include tracking error and image flicker which further reduce sense of realism and immersion (LaViola 2000). Position tracking is the ability of the VE technologies to track the position of the head and limbs of the users in the real space so that an accurate representation of the user can be made in the VE. This depends on the accuracy of the trackers where inaccurate tracking (tracking error) would cause motion sickness (LaViola 2000). Image flicker is distracting and may cause eye fatigue (Harwood and Foley 1987). The peripheral vision is more sensitive to flicker than the fovea, thus the wider the FOV of the display the higher the tendency for flicker to be perceived (Boff and Lincoln 1988; cited in La Viola 2000). In order to remove the tendency to perceive flicker, the refresh rate of the system must be increased. A refresh rate of 30Hz is considered sufficient for vision in the fovea; however, this value must be increased for the vision in the periphery (LaViola 2000). An additional factor that affects flicker is phosphor persistence, which refers to the rate of fading after it has been energized. Long persistence phosphor will reduce flicker but this creates an image smear during motion, where the previous image is still in view (LaViola 2000).

In immersive display, stereo image presentation is often used (Pfautz 2000). Stereo image presentation is considered important as it provide the user with a sense of immersion (Hodges and Davis 1993, cited Pfautz 2000). Other researchers have similarly argued that stereo image presentation help increase user sense of immersion and also realism (Sadowski and Stanney 2002). The use of stereo images has been shown to improve performance on certain tasks (Yeh and Silverstein 1992). However, there are several drawbacks to stereo image presentation. Related technological issues were:

- Increased rendering time due to the need to process two images
- Additional hardware requirement for stereo viewing

- Ghosting effects. This occurs in time-multiplexed displays when the image intended for one viewpoint remains visible during presentation of the other viewpoint (Pfautz 2000)
- Complex hardware that need additional calibration. Improper calibration might lead to simulator sickness (Robinett and Rolland 1992)

Additionally, stereo displays caused more visual fatigue than monocular display (Okuyama 1999; cited in Pfautz 2000).

From these discussions, fully immersive systems require very high performance graphics software and hardware to achieve an acceptable level of realism in terms of image and interaction. This makes these systems very costly to acquire. However, future improvement in technology might produce inexpensive high refresh rate visual systems. Due to the drawbacks of fully-immersive systems, these systems will not be employed in the research presented in this thesis. As mentioned earlier, the semi-immersive systems will be used instead. Based on the comparison among the three types of VE systems, further justifications for the choice of such systems are given in the next section.

3.2.4 Comparison among the types of VR systems

In the following table (Table 3-1), Kalawsky (1996) provides a comparison of the three types of VR systems discussed in the previous sections based on the following features: resolution, scale (perception), sense of situational awareness (navigational skills), field of regard, lag and sense of immersion.

Table 3-1 Comparison between different VR systems implementations in terms of qualitative performance (Adapted from Kalawsky (1996))

Qualitative Performance			
Main features	Non-Immersive VR (desktop)	Semi-immersive VR (projection)	Fully-immersive VR (head-coupled)
Resolution	high	High	Low- Medium
Scale (perception)	low	Medium - high	High
Sense of situational awareness (navigational skills)	low	Medium	High
Field of regard	low	Medium	High
Lag	low	Low	Medium - high
Sense of immersion	None - low	Medium - high	Medium - high

Both non-immersive and semi-immersive systems can produce images of high resolution compared to fully-immersive system. Increased resolution is often associated with aesthetic

reason where increased resolution will make image clearer and sharper (Pfautz 2000). As reviewed in Chapter 2 increased resolution would increase participant's spatial performance.

Non-immersive systems are rated lowest on field of regard while fully immersive systems are rated highest and semi-immersive system field of regards falls somewhere in between. It is generally believed that a wide FOV can increase sense of immersion (Prothero and Hoffman 1995). A wide FOV is often associated with an increase in the user's sense of presence (Prothero and Hoffman 1995, Duh, Lin et al. 2002). A wide FOV as in a fully immersive system closely matched the HVS FOV, thus images displayed are often 1:1 scale. However, While wide FOV improves user sense of immersion but a wide FOV increase the likelihood of motion sickness (La Viola 2000).

Both non-immersive and semi-immersive systems are rated low on lag, but for fully immersive system lag is rated medium to high. As discussed earlier (Section 3.2.3), significant lag would reduce users' sense of realism and presence and may affect performance. Non-immersive systems are low in terms of perception of scale, sense of situational awareness and users may experience no or very low sense of immersion. In fully immersive systems users' experience higher perception of scale and situational awareness; though on sense of immersion it ranges from medium to high. Semi-immersive systems are rated medium or medium to high on these features.

As mentioned earlier, it should be noted that the boundaries between these systems are not clear and distinct. It is possible to convert a desktop system into a semi-immersive system by using a shutter glasses. With appropriate software, one can have a fully immersive system on desktop by including a HMD.

Whilst a fully immersive system is often perceived as advantageous in terms of increasing the user's sense of immersion and presence, some researchers indicate that for many applications the same effect is possible with proper 3D cues and interactive animation in non-immersive systems and less immersive systems (Robertson, Card et al.1993). Robertson and colleagues compare experience in non-immersive VE to a good video arcade game. They suggested as the user controls the animation and focuses on it, the user is drawn into the VE. They further suggest "mental and emotional immersion takes place, in spite of the lack of visual or perceptual immersion." Other researchers have indicated that that there is no difference in user sense of presence in transfer of knowledge task between non-immersive desktop and immersive display (Youngblut and Huie 2003). As mentioned earlier, fully immersive systems come with some negative attributes which may affect inhibit the sense of immersion,

performance and acceptability of the systems. Non-immersive systems users may also experience some side effects; however, these effects are common to normal computer system usage (Costello 1997). The desktop VR or non-immersive VR would be similar to a standard office computer where the visual side effects would be eyestrain and visual fatigue. Prolonged exposure to semi-immersive large projected display might also lead to eye strain and headaches. However, it is considered less visually taxing when viewing distances are set close to optical infinity (4 m or greater).

Based on the qualitative performance (Table 3-1), comparatively semi-immersive displays have the advantage of high resolution as in desktop but higher than desktop in terms of sense of immersion, situation awareness, scale and field of regard. It is low in lag, a negative attribute of fully immersive system. This suggests for better perception of image quality in terms of resolution, scale, immersion and situation awareness without the problems associated with lag, semi-immersive systems would be a better choice over the other two systems. In fact, according to Kalawsky (2000b), a flat screen semi-immersive display "is without doubt a cost-effective way of creating a compelling display environment." Additionally, due to the users' issues associated with fully immersive systems, the type of VE systems used in the experimental works reported in this thesis are non-immersive and semi-immersive systems.

Although, relative distance judgments based on motion parallax cues from head motions are very effective cues, almost as accurate as binocular disparity according to some researchers and more influential than accommodation cues according to others (see Section 2.3.2 of Chapter 2), no head tracking was used in our experiments. In addition to the problems related to head-tracking errors and lag which may reduce the realism experience and affect users' performance, motion parallax cues are removed by restricting participants' head movement in order to focus investigation on the factors to be examined.

### 3.3 CHAPTER SUMMARY

In this chapter the technological issues in a VE are presented. The fundamental concept of modelling a VE was initially introduced. In order to create a convincing simulation, the VE system should provide accurate simulation of the real world counterpart in terms of image and behaviour presentation to the viewer. Thus, a discussion on the issues of creating image realism in a VE model follows with focus on the algorithms and techniques of achieving high realism such as the use of illumination and reflection models, shading techniques and texture mapping. This includes discussion on trade-off between achieving image realism and maintaining acceptable system performance for interaction.

The types of VE system used for presenting the VE to the user were discussed next. These systems were compared in terms of the level of immersion each provides. The advantages and drawbacks of each system were also presented. The fully immersive systems were discussed in detailed with respect to the current limitations of such systems. Finally, a comparison among VE systems in terms of qualitative performance criteria was presented. The issues highlighted in this chapter (and from Chapter 2) provide knowledge for decisions on the basis for experimental approach and method taken in this thesis, particularly for the decisions taken in making the choice of techniques used in the modelling of the VE. The following chapter, Chapter 4, provides a discussion on the basis for experiment approach and method taken in this thesis.



# CHAPTER 4

---

## *BASIS FOR AN EXPERIMENTAL APPROACH FOR UNDERSTANDING SPATIAL AWARENESS*

### 4 OVERVIEW

In this chapter, the basis for the experimental work for the research presented in this thesis is outlined. The first section provides discussions and arguments for the basis for the experimental work. These are drawn upon the literature reviewed and issues discussed in the prior chapters of 1, 2 and 3. The resultant overall research aims, questions and research scope and assumptions will be highlighted. A summary of the experimental basis/approach taken is presented at the end of this section. This includes listing the research questions to be explored and stating the research scope and assumptions. The general research methodology employed and the arguments for the specific choice of experimental methods used to address the research questions in this thesis are presented in the second section. This includes methods for data collection and data analysis. Finally, a summary of the research methods and experiments is given at the end of the chapter.

## 4.1 BASIS FOR EXPERIMENTAL WORKS

It was highlighted in the earlier chapter (Chapter 1) that the VE technologies have been gaining a wide acceptance as an important tool in various areas of applications such as education and training, prototyping, medicine, data visualization, architecture and entertainment. Presenting a simulated experience of the real world with the flexibility to explore and view this virtual world from different perspectives interactively in real time has been reasons for its popular acceptance in these diverse fields of applications. As a VE enables a user to experience and explore this computer-generated 3D spaces, to be useful and effective users must be allowed to perceive the virtual 3D space and spatial relations in the VE in a similar way to the real world. Several researchers have stressed the importance of accurate space perception and distance estimation in VE as an essential prerequisite for the reliable use of VE applications (Wartenberg and Wiborg 2003). Similarly, others have argued that the utility of VE in any intended application is predicated upon the accuracy of spatial representation formed in the VE (Arthur, Hancock et al. 1997). Thus, knowledge and understanding on how to allow user's perception and performance in VE similar to the real world is essential for the effective and efficient design of VE related applications.

As a component of a much broader and an important concept of situational awareness, Endsley (2000) stressed that accurate perception (which refers to the user's spatial awareness of elements' status, attributes and dynamics) is essential. Incorrect perception would affect the next two levels of situational awareness of comprehension and projection, thus adversely affecting a person's overall situational awareness. This implies understanding spatial awareness, the spatial perception of 3D space which includes knowledge of objects' spatial relations and distances, is very critical and important for effective VE design. However, to date, the current VE technologies are still inadequate. Most available VE models do not provide the users with exact replicas of the real world places. As discussed in Chapter 3, there are several issues and constraints with regards to the modelling and presentation of VE that mimics real world places with high degree of realism. Often, the spatial properties (such as geometric constructions, lighting and textures) are not accurately modelled. With the exception of visual cues, most often other sensory cues (such as kinaesthetic and proprioceptive cues) are not available to the users, thus questions the VE technologies ability to provide similar experience and to be perceived similarly to the real world (Henry and Furness 1993, Ruddles and Paynes et al. 1998, Waller and Hunt et al. 1998, Yoon, Byun et al. 2000). Additionally, the user related issues associated with some VE systems exacerbate the problem of providing similar experience in both environments. The literature reviewed in Chapter 2 has indicated that VEs are frequently perceived differently from the real

environment. The exact reasons for the perceptual differences between the real and VE are still unknown (Willemson and Gooch 2002) and understanding factors that influence user's perception and spatial knowledge using VE technology is still limited (Cutmore, Hine et al. 2000). This indicates the need for further research. As the effective and efficient design of VE related applications depends on the user's ability to perceive VE similar to its real counterpart, it is thus essential to examine and understand factors that influence user's spatial awareness in the VE.

It has been suggested that comparing human task performance in the VE to a similar task performance in the real world can provide knowledge on which aspect of the VE technologies require improvements (Witmer and Sadowski 1998). As Kalawsky (2000a) said:

“There is considerable merit in being able to compare performance in the real world against performance in a virtual environment, especially if the virtual environment is mimicking the real world in some way.”

As such, examining the conditions in which spatial perceptions are systematically misrepresented in VE when compared to the real world would signify an essential move towards understanding the limit of VE (Waller 1999). Additionally, a comparison of real world task against a similar virtual world task would provide an objective baseline for the effectiveness of the performance in VE. Whilst it is not necessary to match virtual task to real world task especially for interactions techniques where this will limit the flexibility of methods interactions in the virtual world, a controlled comparison between the real and VE performance would still prove a useful benchmark (Mania 2001). An examination of factors that influence users' spatial task performance in the VE would contribute towards a more effective and efficient design VE, where the task performed in the VE is similar to task performed in the real world. Thus, the overall aim of the research presented in this thesis was to examine factors affecting a user's spatial awareness perception in the VE in comparison to similar perception in the real environments.

**Overall research aim:**

*To examine factors influencing spatial perception in the real and VE by comparing spatial performance in both environments*

In order to realize this aim, the research in this thesis explored the following key research questions.

1. *Is there a difference in spatial tasks (distance estimation and spatial memory task) performed in real and VE?*

2. *How does the display size (large and small) affect users' spatial task (distance estimation and spatial memory task) performance in real and VE?*
3. *How does the type of interface device (mouse and trackball) affect users' spatial task performance (distance estimation and spatial memory task) in VE?*
4. *How does the type of travel mode (drive and fly mode) affect user's spatial task performance (distance estimation and spatial memory task) in VE?*

The reasons for the exploration of these research questions, hence the basis for the experimental approach, are presented in the following subsections. Subsection 4.1.1 argued for the choice of performance measures employed in this thesis while subsection 4.1.2 reviewed and provided the rationale behind the factors selected for investigations in this research. In addition, both the scope and limitation of the research investigation are discussed and outlined. Subsection 4.1.3 gave arguments for techniques utilized for the modelling of the VE models used in this thesis. Finally, subsection 4.1.4 provides a summary of this experimental basis.

#### 4.1.1 Task performance measures

An important aspect towards understanding human performance in the VE is to identify tasks that will be performed in it (Arthur 2000). For testing training applications of VE, Lampton and the others (Lampton, Knerr et al.1994) developed Virtual Environment Performance Assessment Battery (VEPAB) as a move towards benchmarking VE performance. This includes description of tasks for performance evaluation: vision (acuity, colour, search, object recognition, size and distance estimation), locomotion, tracking, object manipulation and reaction time tasks. VEPAB uses simple tasks as opposed to complete training scenarios as these simple tasks formed the basics of other large tasks. Additionally, these tasks can be easily employed to other applications. Evaluation results showed that participants are sensitive to practice effects and as such in any task design, the user characteristics need to be taken into consideration. These results could provide a baseline for evaluation of VE implementation. As such in this thesis, in addition to collecting data on the task evaluated, participants' background information, practice time and test times were also collected as explanatory variables.

Spatial knowledge in the real world is often evaluated using performance measures such as map drawing (spatial representation), orientation judgment, navigation and distance estimation (McNamara 1986). These tasks are informative about certain aspects of spatial cognition and spatial behaviour. Thus, the choice of task used depends on the particular aspect

of spatial awareness being investigated. Kalawsky (2000a) suggested that the metrics developed for the real world case can also be used in the VE evaluation. In many VE applications (such as flight training, tele-operation of robots, visualization of complex data sets, product visualization and medical training) information about distance and depth of objects are of particular importance (Surdick, Davis et al. 1997). For spatial memory tasks, it has been argued that the utility of VE in any intended application is predicated upon the accuracy of spatial representation formed in the VE (Arthur, Hancock et al. 1997). In this thesis, the definition of spatial awareness encompasses objects' spatial relations and distances. Thus the evaluation of spatial awareness in terms of distance estimation task and spatial memory task would relate more towards the spatial awareness behaviour that this thesis intended to examine. Therefore, these two tasks are used as task performance measures in the research of this thesis and are described in the following sections.

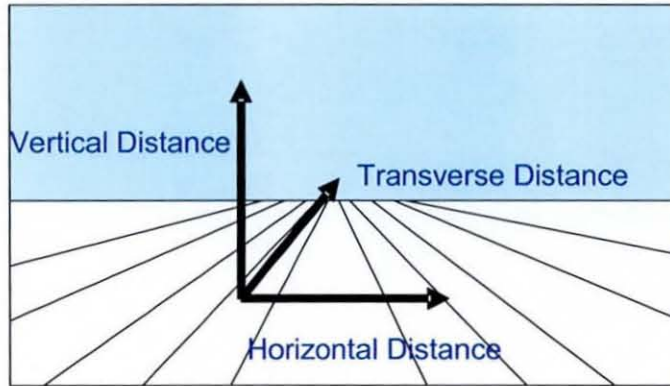
#### 4.1.1.1 Distance estimation tasks

The studies reviewed (in Chapter 2) on distance perception showed that distances estimated in the VE were less accurate than those found in the real world. Some results reported an overestimation and some studies reported an underestimation; while others reported that the differences were very small. These contradicting results may be due to the differences in the variables being investigated, such as egocentric distance versus exocentric distance, distances tested, methods of measuring distance (direct measures such as verbal measures and indirect measures of walking task), display types (desktops, HMD, head-tracked and non head-tracked) and other differences in experimental methods. As such, it is difficult to generalize findings from these studies. Various factors have been suggested to explain why distance is inaccurately perceived compared to the real world. However, the exact reasons for the perceptual difference between real and VE are still unknown (Willemson and Gooch 2002). Thus, more research is needed to understand contributing factors for distance misperception in the VE.

The focus of past studies has been investigations into egocentric distance in the VE (Witmer and Kline 1998, Sinai, Krebs et al. 1999, Witmer and Sadowski 1998, Lampton, McDonald et al. 1995, Lampton, Bliss et al. 1994). Only a few studies have examined exocentric distance (Caird and Hancock 1991, Waller 1999, Bigham 2000). Egocentric distance refers to the distance between the observer and the viewed objects while exocentric distance is the distance between objects or between points on the same object (Coren, Ward, et al. 1999). Lesser attention has been given by past researchers of similar studies to the examination of specific distance types such as vertical, horizontal and transverse (termed as asymmetrical distances see Figure 4-1). Vertical distance refers to the height, horizontal distance refers to distance



across the screen and transverse distance refers to distance into the screen. These distances are necessary for the perception of space and layout of a VE.



**Figure 4-1 Asymmetrical distances of vertical, horizontal and transverse**

These distances however have been intensively researched with respect to performance in the real world (Yang, Dixon et al. 1999, Dixon and Proffitt 2002, Higashiyama and Ueyama 1988, Higashiyama 1996). For example, the vertical-horizontal illusion theory (Yang and Dixon et al. 1999) indicated there is a difference in subject performance between vertical and horizontal distance. Higashiyama and Ueyama (1988) investigated the relationship between the perceived vertical and horizontal distances in a real outdoor setting. Participants were asked to adjust horizontal distance so that it appeared equal to vertical distance. Their study results showed that when vertical and horizontal distances are physically equal, vertical distance tends to be perceived larger than horizontal distance. Additionally, their study results showed that vertical distance of a building appears larger when viewed from far than at close viewing.

These conditions are referred to as vertical-horizontal illusion (VHI). Dixon and Proffitt (2002) defined VHI as a condition that “occurs when a physical vertical extent is overestimated in length relative to a comparable physical horizontal extent.” Yang, Dixon et al. (1999) provide a comprehensive review of theories related to VHI. The authors conducted a series of studies examining VHI in outdoors, pictures and VEs. Their second experiment reveals that observers who viewed outdoor poles yield greater distance overestimation compared to those who viewed pictures of the poles and this value increases with height. Their study further demonstrated that overestimation was greater for 3D environments compared to 2-D displays. Their results suggest that small projection causes small vertical overestimation. They proposed that vertical overestimation would increase if a picture is magnified (such as when projecting the picture to a large screen). This prediction was confirmed by a later study by Dixon and Proffitt (2002) who demonstrated that the differences

between VHI in 3D and 2-D environments were influenced more by the distance extent of the presentation rather than the dimensionality of the display. Their study result indicates that vertical overestimation increases with increased size of the virtual of pictured objects. The authors concluded that the larger the 2-D representation the more likely the visual system is to achieve a natural perception of the large depicted object in which vertical overestimation is more in the real environment than for pictures.

As reviewed in Chapter 2 several researchers have examined perception of room dimension in terms of the vertical (height) and horizontal (width and length) of rooms (Henry and Furness 1993, Yoon, Byun et al. 2000). Henry and Furness (1993) compared spatial dimension estimation in four viewing conditions: desktop monitor, stereoscopic head-tracked HMD, stereoscopic non head-tracked HMD, desktop monitor and real environment. It was found that vertical distances were estimated very accurately compared to the horizontal distance. The accuracy may be attributed to several reasons. Firstly, vertical height is often used as scale. Secondly, most rooms are based on standard height and thirdly the fact that their sample participants comes from the architectural background which makes height estimate more accurate compared to horizontal distance. All distances in the simulated conditions were smaller and less accurate compared to distance estimates in the real world.

Yoon, Byun et al. (2000) found no significant difference between distance estimate in the real and virtual room in terms of the width, height and length. However, both differed significantly from the actual distance. It was found that participants tended to make more errors in height estimation compared to width and length with the latter being estimated more accurately. Overall distance estimates were reported to be accurate even though users tended to overestimate distance.

The studies reviewed above suggested that different distance types yield different results. Examining exocentric distance tasks in terms of these asymmetrical distances (vertical, horizontal and transverse) would allow more detailed and systematic examination of distance estimation performance; hence provide more detailed understanding of these tasks. Essentially, the knowledge of distances between objects forms the basis of our understanding of the physical structure (Golledge 1991). It forms the basics for many other tasks such as navigation and wayfinding. It has been suggested that understanding the sub-tasks of a complex process leads to a better understanding of the system requirement (Wilson 1998, cited in Pfautz 2000). Complex processes might vary among application, thus understanding the basic sub-tasks that form the basis for many complex tasks provides a simpler approach toward understanding the higher-level process. As mentioned earlier, exocentric distance in

terms of asymmetrical distances has received much less research attention compared to egocentric distance. Thus, this aspect of spatial awareness is employed in the research presented in this thesis as one of the spatial task performance measures to compare spatial performance between real and VE in all experiments.

#### 4.1.1.2 Spatial memory task

Following a review of several studies, Arthur, Hancock et al. (1997) concluded, “a central issue for the use of VEs both as an interface and training tool is how users mentally represent that virtual space”. They further asserted that the utility of the VE for any applications for which they are being intended is dependent upon the accuracy of this spatial representation formed in the VE. As such, it is essential for a user to understand the space in which the tasks are to be performed. This implies that spatial memory tasks are considered important in terms of spatial representation of the VE. Therefore, understanding how people form cognitive maps or spatial memory of a VE is very important for effective VE design.

Caird and Hancock (1991) pointed out that information on how the user judges the actual layout of a simulated environment, spatial memory, has been very limited as most research efforts in simulation of physical environment have focused on questions of fidelity (realism) and perceived distance (spatial perception). Most studies that examined spatial memory task reviewed in Chapter 2 were concerned with comparisons of performance between display types (Johnson and Stewart 1999, Patrick, Cosgrove et al. 2000) or between map and VE (Richardson, Montello et al. 1999, Rossano and Moak 1998, Goerger, Darken et al. 1998). Though some researchers (Witmer, Bailey et al. 1996) have examined spatial memory task performance in the context of transfer of knowledge (training) in the VE to the real world, few studies have directly compared spatial memory task performance in the VE against similar performance in real world (Arthur, Hancock et al. 1997, Richardson, Montello et al. 1999).

Spatial memory tasks have been commonly used in usability studies and in assessing virtual interfaces (Mania 2001). As suggested in the earlier section, spatial memory tasks have been used as one of the performance measures in evaluating spatial knowledge. Spatial memory task has been compared in studies as a performance measure in the real and VE (Alfano and Michel 1990, Henry 1992, Arthur, Hancock et al. 1997, Wilson, Foreman et al. 1997, Goerger, Darken et al 1998, Rosano and Moak 1998, Wilson 1999, Patrick, Cosgrove et al. 2000). Empirical evidence has shown that sketch map is a valid measure of cognitive map or spatial memory task in VEs (Billinghurst and Weghorst 1995). Their study result which

showed positive correlation between orientation and a sketch map (map drawing) confirmed that sketch map is an acceptable measure of cognitive map in VE.

Because of the importance of accurate spatial representation formed in the VE, the spatial memory task was adopted as one of the metric to compare performance in the real and VE in the research presented in this thesis. Besides a commonly used task for measurement of spatial knowledge, this task is a more suitable measure because generally cognitive maps or spatial representation were formed by active interaction with the environment (Neisser 1996; cited in Billinghurst and Weghorst 1995). Thus, this metric would be appropriate in investigations of interactive images. As proposed in the next section, interactive images will also be examined in this thesis and spatial memory task would be appropriate for such investigation. Thus, spatial memory task is employed in this for Experiment 3 which examines spatial awareness in interactive images (reported in Chapter 7).

## 4.1.2 Selection of factors to be examined and scope of investigation

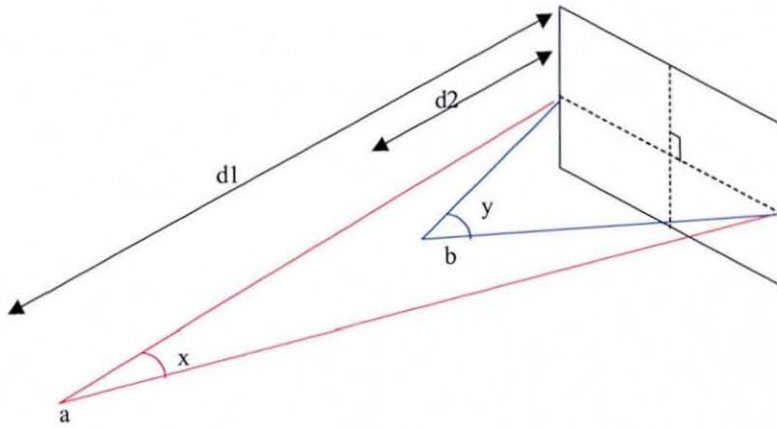
### 4.1.2.1 Display size

Whilst the exact reasons for perceptual differences between the real and VE are still unknown (Willemson and Gooch 2002), however various factors have been investigated and suggested. Related aspects of the display or the computer systems have been the main focus of intensively studied factors as potentially contributing to the misperception of distance in the VE (Waller 1999). In fact, the display system has been suggested as one of the probable causes of distance underestimation in a VE (Egglestons, Janson et al. 1996, Witmer and Kline 1998, Willemson and Gooch 2002). These factors include variation in display-types, FOV, image quality, image type, scene contrast, resolution, viewing modes, interface devices, modes of travel in VE, mismatch of cues. Other factors include, user's experience and user involvement with the VE (active or passive), physiological cues, pictorial cues, and motion factors.

As reviewed in Chapter 2, although some researchers indicated otherwise (Riley and Kaber 1999), several researchers have reported better performance on large projected display over desktop monitor (Patrick, Cosgrove et al. 2000). The latter have attributed the better performance of large projected display over desktop monitor to the larger display size inducing a greater sense of presence on the participants. They further claimed that the images on large display are large enough to appear real to the participants thus improving their performance. Similarly, other researchers have concluded that larger display affords better sense of presence on the user (Tan, Gergle et al. 2003) resulting in better performance on

large display. This suggests that the size of the display may exert an influence on the user's performance.

The size of the display or any objects is related to FOV and retinal image size in some manner. The retinal image size, which describes the size of the image on the retina, is usually measured in visual angle or FOV of the viewed scene or objects. The retinal image size is proportionally related to the FOV; that is larger FOV will result in larger retinal image size. However, the FOV is inversely related to the viewing distance of the observer from the viewed objects. The larger the viewing distance the smaller is the visual angle and the smaller the viewing distance the larger is the visual angle. This is also true when one views an image on a display (see Figure 4-2). The closer the observer ( $d_2$ ) is to a display, the larger the FOV ( $y$ ) and the retinal image. For larger distances the reverse is true.

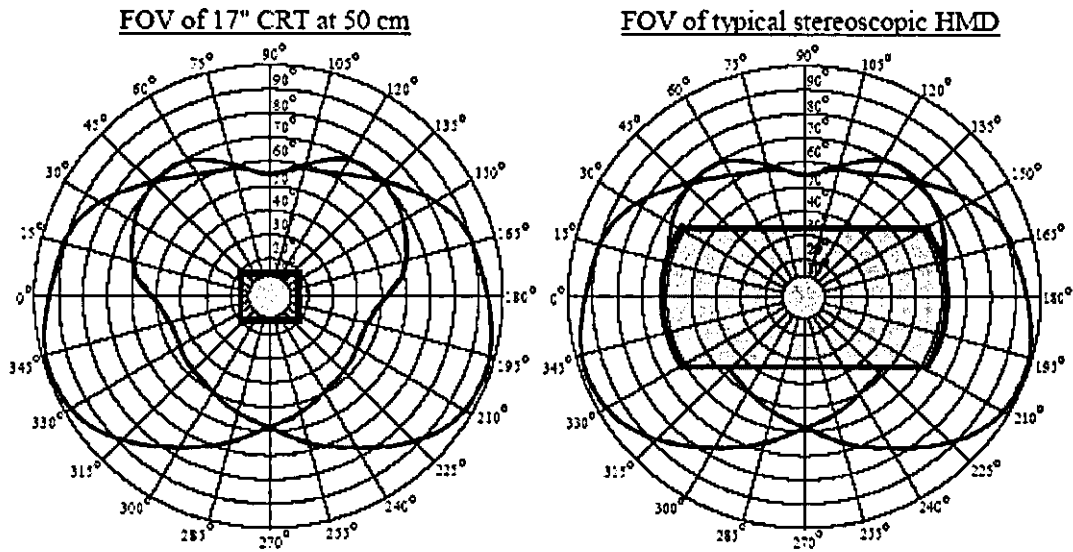


**Figure 4-2 Relationship between FOV and viewing distance**

Some researchers indicate that large display with wide FOV contributes to the improvement in participants' performance on navigation tasks (Czerwinski, Tan et al. 2002). It is generally believed that a wider FOV encourages a higher sense of presence (Prothero and Hoffman 1995). Prothero and Hoffman (1995) found that participants reported a significant higher sense of presence with wider FOV. The human FOV which span  $200^{\circ}$  (maximum FOV) horizontally and  $150^{\circ}$  vertically is very much larger when compared to the VE display. Inherently, a larger FOV for the VE display would closely match the human FOV and may yield similar performance in both environments. Figure 4-3 shows a resultant FOV from desktop monitor and HMD in comparison with the human FOV.

Studies have shown that participants reported that natural images were seen as more realistic with a larger FOV ( $100 \times 180$  degrees) than smaller FOV ( $30 \times 20$ ) (Hatada, Sakata & Kusaka 1980, cited in Pfautz 2000). They further report a positive relation between FOV and the "sensation of reality".





**Figure 4-3** The figures indicate comparison of display FOV with human FOV. The dark black lines indicate the left and right eye FOV. The box in the images represent the FOV of a 17" CRT monitor viewed at 50 cm (left image) and the FOV of typical HMD (right image). Adapted from Pfautz (2000).

Several studies also indicate that display size (FOV) influenced user's performance as well as sense of presence in the VE (Kline and Witmer 1996, Arthur 2000, Duh, Linh et al. 2002). The results of these studies showed that participants performed better and experienced a higher sense of presence and more realism in wide FOV images. Duh, Linh et al. (2002) reported that participants experienced more sense of presence and realism in wide FOV. They attributed this to participants receiving more peripheral information from a wide display. Arthur's (2000) study however revealed that reduced FOV influence participants' performance on search and walking tasks but it bore no effect on distance estimate tasks. In contrast, Kline and Witmer (1996) reported that participants' distance estimates were more accurate on wide FOV display than on small FOV display. However, for spatial representation task, results from Arthur (2000) and Johnson and Stewart (1999) reported no significant difference between wide and narrow FOV.

A real world study (Alfano & Michel 1990) that compares the effect of limiting FOV on user's performance on perceptual motor task and memory cognitive map test revealed that participants' performance was lower on the recall of objects locations in a room when the FOV was reduced. When asked to move rectangles of varying size onto their outlined counterparts, participants' performance improved with wider FOV. In an earlier study, Dolezal (1982) examined the effect of peripheral vision by wearing two 30cm long paper tubes of restricted FOV of  $12^\circ$  for six days. He found that he was unable to form a cognitive

map of a previously unseen room. He also reported under-reaching of objects because these objects appear smaller and nearer. Similarly, Hagen, Jones et al. (1978) also found that a truncated FOV lead to compression of distance in pictures.

Studies conducted in the real and virtual world described above indicate that display size (FOV) does influence user's performance in both environments. Generally, large FOV results in better user performance than narrow FOV on some tasks, while others indicate no difference. The inconclusive findings suggest that further research is necessary to determine if FOV affect distance estimate and spatial memory tasks.

Despite the research done in comparing FOV, very few scientists have directly examined the effect of physical display size and distance on task performance (Swaminathan and Sato 1997, Tan, Gergle et al. 2003). There are several other reasons to investigate the display size factor:

- In the world of perception, size matters a lot (Reeves and Nass 2000). It helps us to judge distance cues from size.
- In the case of displays, researchers in entertainment have shown that larger displays are more arousing and are preferred by user and they induce a greater sense of presence (Reeves and Nass 1996, Reeves and Nass 1999). These researchers' findings also showed that the higher the arousal, the better memory for the media experiences.
- In contrast to a television screen size, the IMAX giant flat screens, which could reach up to eight stories high, are large enough to encompass the viewers' peripheral visions and thus allow viewers to feel immersed in the scene (IMAX experience<sup>1</sup>).
- The increasing trend towards large display devices has raised a series of questions (Kasik 2002). One of the questions, which seek to understand the situations where such devices are beneficial, is of particular interest to VE technologies and consequently the research presented in this thesis. With regard to VE technologies, a related question would be whether larger display affords better spatial awareness performance than small display. The choice of display for VE presentation has cost and performance implications. For example, large panoramic display would improve user performance due to more sense of immersion and presence, however it is more costly to acquire compared to desktop PC. Similarly, other immersive displays such as HMDs and CAVEs are comparatively more expensive than desktop and large panoramic displays. Moreover, there are unwanted attributes that comes with fully immersive HMDs which questions its usage over other types of displays.

---

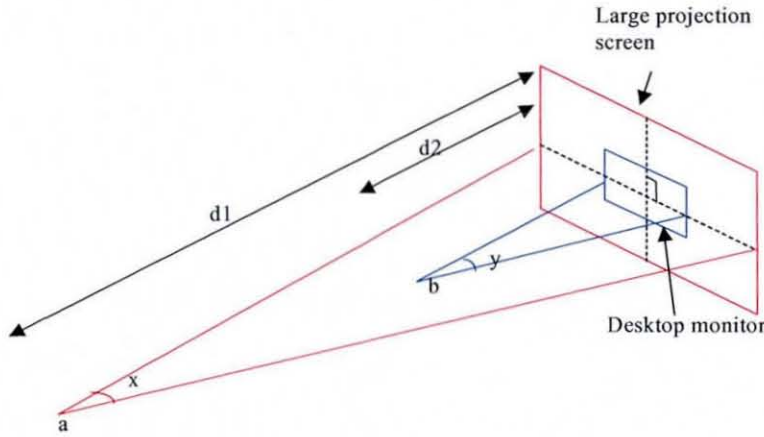
<sup>1</sup> The IMAX experience explained. Available at <http://www.bfi.org.uk/showing/imax/explained.html>

Some of the initial studies mentioned earlier reported a difference between participants' performances on a desktop monitor and a large projected display (Patrick, Cosgrove et al. 2000, Tan, Gergle et al. 2003). With regards to display size, both display types (desktop and large projected display) vary in physical size. Since, the FOV of both displays were equated, the physical display size may have been suggested to contribute to the performance difference between the large and small display. Whilst these studies indicate better participants' performance on larger display over small display on spatial orientation, spatial knowledge, mental rotations, and navigation, there exists empirical evidence to suggest that this is not true for reading, distance estimate and spatial memory task. From these results, it is shown that the better performance of large display over small display is task-dependent and it cannot be generalized for all tasks. Moreover, some researchers have reported that participants performed better on a desktop than on a large projected display (Riley and Kaber 1999). They attributed the results to the better resolution on the desktop monitor and participants' familiarity with desktop environment. The inconsistent findings suggest that further research is necessary in order to understand the role of physical display size in spatial perception of VE.

**Research Question:**

*How does the display size affect users' spatial task (distance estimation and spatial memory task) performance in real and VE?*

As described earlier, although display size is related to FOV, whereby large display is often associated with large FOV (and vice versa), it is possible to have similar FOV on both display sizes. Some of the studies mentioned earlier (Tan, Gergle et al. 2003, Patrick and Cosgrove et al. 2000) that reported better performance of large projected displays (large display) participants over desktop monitor (small displays) participants maintained a constant visual angle for different display sizes to isolate the effect of display size factor. However, in order to maintain similar visual angle ( $x = y$ ), the distance of the observer ( $d_1$  is larger than  $d_2$ ) from the display needs to be varied for both display size (see Figure 4-4). This experimental setup failed to account for other factors such as viewing distance and physiological cues which were varied in both the large and small display conditions. Besides the display size, these factors may also contribute to the better performance of large display over small display.



**Figure 4-4 Experiment setup of previous investigations: Similar visual angle ( $x = y$ ) and different viewing distances ( $d_1 > d_2$ )**

Viewing the image at different distances may have some impact on what the user may perceive. From the geometrical perception of pictures (see Section 2.8.1 of Chapter 2), the geometrically specified depths in picture are compressed and expanded when the viewing distance of the viewer from the display is decreased and increased respectively. Thus, close viewing will result in distance perceived being shorter than actual and viewing from a distant may result in the distance being perceived as much longer. Though, to realise such predictions the picture needs to contain strong linear perspectives.

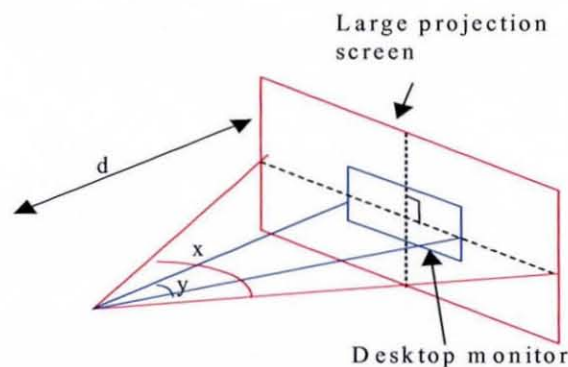
Some researchers have reported that an object viewed at greater distance portrayed large distances compared to an equivalent scene viewed at shorter distance (Gooding, Miller et al. 1991). The different viewing distances from the display may also result in different physiological cues acting at different distances, which in turn may affect performance.

Results from various psychological experiments carried out by NHK (Japanese Broadcasting Corporation) whose tasks were to foresee the performance required for next-generation TV systems in Japan, found that viewing distance of  $3H$  (where  $H$  is the height of the screen) gave the greatest sense of presence on the viewer (Oyama and Shiramatsu 2002). Since some researchers recommended to increase users' sense of presence in order to improve performance benefits in VE (Stanney, Kingdon et al. 2002), this further implies the need to investigate viewing distance as it might influence participants' performance in terms of sense of presence.

It has been reported that the distance of accommodation may influence the perceived size and the distance of an image (Iavecchia, Iavecchia et al. 1988). As reviewed in Chapter 2, even

though accommodation and convergence cues are limited in the range of distance for their effectiveness, some empirical evidence has revealed that our eyes converge and accommodate at varying distances in the picture. Thus, the better performance of large display over small display as reported by earlier studies may not be attributed to physical display size alone; other variables (such as viewing distance and physiological cues) that are not controlled by these researchers may also contribute to the results. Thus, it is important to consider and investigate the effect of display size by also considering the influence of other related factors such as viewing distance and hence these physiological cues as they might also contribute to the better performance of large display participants over small display participants.

In order to examine the effect of physical display size and at the same time consider the possible influence of the later factors, the following experimental approach was proposed. To investigate the effect of display size, two related experiments were proposed. The first experiment will investigate the effect of display size by controlling the effect of FOV that is by having similar FOV for both display size. This approach is similar to the approach taken by previous investigations, thus experimental setup would be similar to Figure 4-4. However, this setup which was employed by previous investigations fails to account for the effect of viewing distance and physiological cues. In this setup, fixing the FOV for both display results in varying viewing distance (and hence may varies the effect of physiological cues). From earlier discussions, these factors may also influence users' performance and contribute to the better performance of large display over small display. Thus, a second experiment was necessary to control the influence of the viewing distance and physiological cues. In the second setup, the viewing distance was fixed for both display size (see Figure 4-5). By comparing the results of both experiments, this approach, which considers the effects of other related factor, enables us to further explain the role of display size in influencing participants' performance.



**Figure 4-5 Experiment setup: Different FOV ( $x > y$ ) and similar viewing distance ( $d$ )**

Therefore, by considering the effect of other related factors (viewing distance and physiological cues) in examining the effect of display size, this thesis expands on the previous



research by further explaining the contribution of display size factor on participants' spatial task performance. Although display FOV is not directly investigated, as it is directly related to display in the experiments design, discussions of results will also include FOV.

#### 4.1.2.2 Image forms: static, dynamic and interactive images

A knowledge and understanding of perception of space in the real world, in photographs and in cinema is essential in the design of a useful and effective VE (Cutting 1997), which in turn assists in achieving the goal of faithfully representing the real world. Besides understanding perception in the real world, this implies the need to understand the perception of static images (photographs) and dynamic images (movies in cinema) too. Moreover, as suggested by several researchers (Stanney, Mourant et al. 1998), perception of depth in VE is very complex and not well understood. As such these researchers stressed the importance of conducting perceptual studies in both static and dynamic scenes as conclusions derived from the former might not be applicable to the latter.

Computer-generated images may be viewed as static images, as a video movie or computer animations (dynamic image) or as an interactive 3D VE. As mentioned earlier it is useful to compare performance in the real and VE if the VE is to simulate its real world counterpart. As one of the goals of VE is to emulate its real world counterpart, it would be essential to provide a comparative evaluation of different forms of VE image presentation with the respective real world correspondence. Most previous investigations reviewed in Chapter 2 involved the last forms of VE image, that is examining performance in interactive 3D VE (Johnson and Stewart 1999, Witmer and Kline 1998, Patrick, Cosgrove et al. 2000, Goerger, Darken et al. 1998, Arthur, Hancock et al. 1997, Heineken and Shultze 2000, Riley and Kaber 1999, Czerwinski and Tan et al. 2002, Tan and Czerwinski et al. 2003).

Yeh and Silverstein (1992) stated that "In normal 3D perception, depth information is often immediately available through motion of the observer and/or objects in the visual scene. The static imagery constitutes snapshots of the visual scene at any given instant in time that an observer could use to extract information about the spatial layout." The authors further argued that display applications such as graphical rendition of complex images for scientific visualization and situational awareness are typically static or have very slow update rates resulting from low information bandwidth and/or the complexity of the computation. The dynamic imagery would be an example of guided exploration or walkthrough of the visual scene such as in architectural applications. This suggests that when interacting with the VE, there are instances when the visual scene might be static (as in scientific visualization applications) and there are applications where user passively viewed a dynamic scene (as in

guided exploration of buildings in architectural applications). Thus, including the examination of static and dynamic images would represent a more comprehensive examination of spatial awareness in terms of the different context of how users work in the VE.

Thus, the approach taken in the research presented in this thesis is to provide a comparative evaluation of spatial awareness task in static images, dynamic images and interactive images of the real and VE. The results of these investigations would augment knowledge in existing literature on knowledge on spatial awareness in these types of VE image presentations.

**Research scope:**

*Examination of user's spatial awareness includes the following forms of image presentations: static, dynamic and interactive real and VE*

In this thesis the approach taken for comparing real and VE conditions in these forms of presentation is as follows. For static images, a photograph of the real world will be compared with picture of VE. For dynamic images, a movie of the real world will be compared to a simulation of the VE. Lastly for the interactive images, the physical real world environment will be compared with an interactive 3D VE.

For the third comparative evaluation, an investigation of interactive VE would in turn raise two related issues of interacting and exploring the VE: interface device and navigation. It has been suggested that the choice of interface device used in interacting with the VE would have an impact on participants' performance (Ruddles and Jones 2001). Similarly, as the interface device is related to interaction in VE, the methods used for navigation or exploring the VE might influence users' performance. As such, the influence of navigation and interface device on participants' performance would also be examined in this thesis. Both navigation and interface device are discussed next.

#### 4.1.2.3 Navigation

The task of navigation is one of the most prevalent user actions in interactive VE (especially in large scale 3D environment). There are two key aspects to navigation in VE: wayfinding and travel. Considerable research has been done on wayfinding but travel has received much less attention (Scott and Dalgarno 2001). Bowman, Koller et al. (1997) emphasized the importance of travel: "... is an important universal user interface task which needs to be better understood and implemented in order to maximize a user's comfort and productivity in a VE system". Travel refers to the control of viewpoint motion through a VE. In this study, investigation is limited to a first person view that is a simulation of what the user will see if he was in the environment (camera viewpoint). This approach is chosen due to its simple

implementation and less factors to control. Unlike the other type of viewpoint referred to as third person view, where movement is based on control (movement) of the representation of the person (such as avatar), the design of the avatar itself is constitute several research issues (Garau 2003). As such it is not within the scope of this thesis to investigate the influence of avatars (however, see Draper (1995) for a study on the influence of virtual body on spatial awareness and Garau (2003) for the influence of avatar design).

Various metaphors have been suggested for travel or motion control in VE. These include walking, flying, driving. The choice of travel modes (metaphor) used by the user might affect their sense of spatial judgment in the VE. For example, movement in the VE using drive mode is different from using fly mode. In drive mode, user viewpoint height above the ground is fixed (as in driving). Thus, the user needs to be concerned only with forward, backward, left and right movement. In addition to these movements, fly mode allows vertical movement. Thus, the user is allowed a 3D motion movement in fly mode compared to 2-D motion movement in drive mode. It would be of interest, if this extra dimension provided the user with extra benefit in terms of their spatial awareness in the VE.

Some researchers have suggested that constraint motion (less degree of freedom) to be an important navigation technique in many applications where users do not need the extra degree of freedoms (Bowman, Koller et al. 1997). They believe that this reduction in cognitive loading due to less degree of freedoms will allow participants to pay more attention to other tasks and features of the VE. Similarly, Gobel and Frendorf (2002) in their evaluation of the different 3D movement control during simulated navigation tasks in a medical application compared devices of varying degree of freedom (mouse, joystick, spaceballs and position trackers) and concluded that more degree of freedom does not necessarily produce better results than device with less degree of freedom.

The above studies indicate that more degree of freedom does not necessary afford better subject performance. Thus movement mode with less degree of freedoms would reduce mental workload of the user and allow them to focus more on the task required. In general, restricting the user's movement to less than 3D reduces the cognitive load and makes navigation easier. As indicated earlier, this thesis proposes to compare the spatial memory task performance of participants. Thus a travel method that helps reduce mental workload is necessary so that participants can focus their attention on remembering objects and spatial layout. Additionally, Stanney & Salvendy (1994) (cited in Mania 2001), argued that participants with low spatial ability are capable of mentally representing the structure of a complete system provided the system are well-organized, the task is clear on acquiring the

structure, and the workload is low. This implies that maintaining low task workload is important not only to focus participants' attention on the task required but also to minimize variance among participants in terms of spatial ability. As the degree of freedom for fly mode and drive mode is relatively minimal (3 and 2 degree of freedom respectively for fly and drive mode), the choice of these travel methods would be appropriate for the spatial task to be examined. As described earlier, the user is allowed a 3D motion movement in fly mode compared to 2-D motion movement in drive mode. It would be of interest, if this extra dimension provides the user extra benefit in terms of their spatial awareness in the VE. Thus, the use of these two travel modes is examined in this thesis.

**Research Question:**

*How does the type of travel mode (drive vs. fly mode) affect user's spatial task performance (distance estimation and spatial memory task) in VE?*

#### 4.1.2.4 Interface device

VE navigation can be implemented using a variety of input devices: mouse, trackball, joystick, position trackers, locomotive devices, eye tracking, haptic devices (see Baldis (1997) for an overview of these devices). The choice of interface device used for interaction in a VE can influence a user's performance (Ruddle and Jones 2001). Subsequently it is expected that participants' spatial awareness would be affected too. There are several reasons to suggest the interface device might influence users' navigation and hence their spatial judgment. First, different device types provide the user with various ways of using them. For example a mouse, some users may use short movements, long movements, or repeated short movements. Some users might drag the mouse, some might alternately drag and lift the mouse and the movement direction might be horizontally and vertically. There is no direct relationship between the cursor position on the display and the position of the device on the desk space as the mouse can be picked up and put at a different position without corresponding movement in the image. Similarly, a trackball too afford different ways of rolling the ball for movement in the VE. Unlike the mouse, the trackball requires no movement of the device; a user just needs to roll the ball to initiate relative movement in the image. This difference shows that this may create a different sense of where a user is in the environment and this might affect the user's spatial judgment.

Another reason is the scaling relationship between the input device and the image, that is, whether it is a relative device or absolute device. A relative device is one whose relative movement will create a relative movement on the image. For example, 1 cm movement of the device may create 2cm movement in the image. This is referred to as *gain*, that is, the control-

display ratio between physical movement of the device and control movement of the device on the image. Thus, in the previous example the control-display ratio was 1:2. The higher the control-display ratio, the greater the distance of movement but degrades fine control. Low ratio display results in rapid movement but allows for fine control. With an absolute device, the control-display ratio is 1:1; this means 1 cm movement of device would result in 1 cm movement in the image. Thus, the difference between a relative and absolute device may have an impact on the user's spatial awareness. However, it is not within the scope of this thesis to evaluate this variable (relative device versus absolute device). This, however, has been investigated by other researchers (Jacob and Sibert 1992).

Another reason for suggesting the choice of device might influence performance is the amount of proprioceptive feedback information received from the knowledge of movement of the body parts of the user. Different devices utilize different types of muscles for movement. For example, a mouse uses muscles of the wrist, forearms, arms and shoulder, while a trackball uses only the fingers (and/or palm of the hand). Proprioceptive feedback can provide powerful information of self-motion (Hlavacka and Mergner et al. 1996, cited in Harris, Jenkin et al. 2002). Thus, it is expected the types of interface used for navigation might affect users' performance and spatial judgment.

In his review of several studies, Baldi (1997) found conflicting evidence about the use of mouse. He found some studies indicated that some users often experience difficulties when navigating in a 3D environment using a mouse while several other studies have revealed that the traditional 2D device (keyboard and mouse) can be successfully used for 3D exploration. In another study, Jacob and Sibert (1992), compared a mouse to a Polhemus 3D space tracker on size and colour matching task. Results show participants performed better using a mouse on colour matching task compared to size matching. In contrast, participants performed better on size tasks using a 3D space tracker compared to the colour matching task. However, participants preferred the mouse to 3D space tracker on both tasks; they found it is easier to learn compared to the 3D space tracker. It should be noted that the mouse here was a relative device and the 3D space tracker was an absolute device. These inconclusive studies results provide further motivation for the examination of mouse (and trackball) utility in 3D space exploration.

Scott and Dalgarno (2001) conducted a comparative study on the usability of motion control interfaces among three 3D VE game. These games use a combination of keyboard, mouse, joystick and game console. They found that arrow keys to be most efficient and keyboard tools were rated highest. One possible explanation for this result is that most participants were



more familiar with keyboard keys. Alternatively, this may be due to the simplicity of pressing the arrow keys for left, right, forward and backward movement. Other researchers have compared several input devices (mouse, trackballs, touchscreen, touchpad, mousepen, and joysticks) on a performing star tracing task (Cohen, Meyer et al. 1993). They found that touch-screen and mouse were the best devices on speed and accuracy while joystick and touchpad were the worst.

In an unpublished work, Mueller, Bliss & Silver found no significant difference between mouse and trackball on a compensatory tracking task but both differed significantly from unmouse (a compact touch-sensitive tablet that perform the same tasks as a mouse). This study also revealed that subject performance for both devices did not differ significantly despite the more frequent use of the participants of the mouse and almost all the participants having either hardly or never used the trackball.

On a pointing and dragging task, three devices were compared (mouse, trackball, and stylus with tablet) (MacKenzie, Sellen et al. 1991). Results showed that the stylus displayed a higher rate of information processing than the mouse in pointing tasks but not during dragging. The trackball ranked the third for both tasks. However, the tasks examined in the above studies varied from pointing to dragging and drawing tasks.

Whilst the literature is abound on studies examining mouse and trackball in comparison to other devices, very few studies have actually compared mouse to trackball utility in the 3D VE navigation on spatial memory tasks. Tong and the others have compared mouse and HMD-bike in four conditions (mouse- monitor (non-immersive); HMD-bike (fully immersive, fully interactive); HMD-bike (limited interaction); HMD-bike (passive-guided movement). Participants were tested on spatial memory task and navigation tasks (Tong, Marlin et al. 1995). The results showed that mouse and fully-immersive conditions were significantly better than other conditions and that mouse was significantly better at object-location association than all other conditions. This result indicates that the mouse is a suitable interface device for the spatial memory task performance. Moreover, its performance is comparable to a fully immersive condition. As indicated earlier, the spatial memory task is one of the spatial task measure proposed to be examined in this thesis; thus, the choice of mouse as an interface device in this thesis would be appropriate for the task to be examined.

Ruddles and Jones (2001) suggested that the simpler the interface device to use, the greater the amount of cognitive resources that participants can devote to updating and maintaining their spatial memory task. They further suggested that the simplicity of an interface device is

affected by the mapping between the physical movements of interface device and movement in the VE and the number of degree of freedom being controlled. As one of the task being proposed to be examined in this thesis is the spatial memory task, devices that could provide reduce mental workload are required so that participants can focus on remembering objects and spatial layout. Both the mouse and the trackball meet this requirement. Bowman (2002) pointed out that interaction in VE is very complex for most users and one reason is the lack of familiar interface for interaction. This suggests that the use of a familiar device would reduce interaction complexity in a VE. The mouse is considered one of the common interface devices (Mueller, Bliss et al. (unpublished work), Zhai and MacKenzie 1998). Its utility is often synonymous with personal computers. It is also considered intuitive, direct and affords transfer of everyday motor skills (Zhai and MacKenzie 1998). The trackball, which is similar to a mouse (a mouse 'turned upside down'), is also another popular device. Even though the mouse and the trackball are relative device, it is expected that participants' familiarity with these devices make them simple to use so that users can focus on the given task of spatial memory. Participants' familiarity with both devices would minimize practice time by reducing learning time to use the device. Concurrently, this is hoped to reduce experiment trial times as longer trial times might affect participants' response (due to boredom or fatigue) which may indirectly confound the experimental results.

The proposed spatial memory task in this thesis involves interactions which are limited to movement in the VE with no object manipulation (Experiment 3 on interactive images). Therefore, the use of these two interface devices (mouse and trackball) is acceptable. In addition to the above arguments, ease of availability and cost factor make these devices an appropriate choice for evaluation in this thesis. Thus the influence of both devices on spatial task performance in interactive VE will be examined.

#### **Research Question:**

*How does the type of interface device (mouse and trackball) affect users' spatial task performance (distance estimation and spatial memory task) in VE?*

Additionally, an interface device questionnaire will be used to collect subjective responses from the participants on their comparative evaluation of these two devices.

#### **4.1.2.5 The use of non-stereo images**

Stereo image presentation provides viewers with a natural and intuitive viewing format of 3D environment (Hendrix and Barfield 1995). Besides giving viewers aesthetically pleasing presentations, it also provides viewers with more accurate perception of spatial layout in the

3D space (Yeh and Silverstein 1992). Some researchers stated that the use of stereo images provides the user with a greater sense of immersion and realism (Hatada, Sakata et al. 1980; cited in Pfautz 2000, Sadoswki and Stanney 2002). In his review of the literature, Pfautz (2000) found that stereo cues have been shown to improve performance in a variety of tasks: 3D tracking tasks, Fitt's Law and teleoperation tasks, distance estimation, relative depth judgments, azimuth and elevation judgments, path tracing tasks, 3D pointer positioning accuracy and detection of subtle features in medical images.

Despite the cited benefits of stereo image presentation in the preceding paragraph, there are several considerations for its non-use. About 10 percent of the population cannot make use of stereo cues to perceive depth (Wan and Mon-William 1996). Besides it has been suggested that the presence of stereo may not enhance performance when monocular cues present in the scene are as effective (Hendrix and Barfield 1995). Moreover, the effectiveness of stereo is limited to a small distance range (Hendrix and Barfield 1995). In fact, the effect of binocular cues diminishes at an increasing distance of the observer from the viewed objects (Cutting 1995).

In Chapter 2, it has been suggested that some people find it difficult to perceive depth in the presence of stereo cues alone, that is, when monocular cues are not present (Barbour and Meyer 1992). This highlights the importance of monocular cues. In fact, proper rendering and emphasis use of monocular cues may compensate for the absence of stereo cues. Additionally it has been suggested that motion parallax cues are almost as accurate as binocular disparity cues.

One of the issues highlighted in Chapter 3 is stereo image presentation. As a set of two images are required for stereo presentation, one for each eye, the requirement on the system resources is thus doubled because the scene has to be rendered twice. This might imply that, in the worst case, the frame rate would be reduced in half thus affecting system performance. Reduced frame rate may cause adverse effects such as jerky motion, reversal of motion, multiple images, and shimmering edges (Pfautz 2000) which in turn affect image realism and user's performance.

Most available VE displays are non-stereo (Wan and Mon-William 1996). This is largely due to the display characteristics which may hinder presentation of stereo images (Roberts, Slattery et al. 2000). It has been reported that stereo image presentations often result in more visual fatigue than monoscopic displays. People have reported eye strain and nausea when using stereo display (Howard and Rogers 2002). This has been attributed to the mismatch

between the physiological depth cues which focus on the display and converge at a different distance (Takeda, Hashimoto et al. 1999). In contrast to normal viewing where accommodation and vergence work in concert and are dependent upon the distance of the viewed objects, in stereo display, the viewer maintains fixed accommodation at the display plane with changing vergence focusing at varying distance in the virtual scene. This requirement for constant accommodation with changing vergence angle causes problems for the visual system such as visual stress and fatigue (Wann, Rushton et al. 1995). Mon-William and Wann (1998) demonstrated that a 10 minute viewing that requires constant ocular focus with changing vergence eye movement is enough to cause deficits in binocular visions. Even though improving image quality would provide a better stimulus for accommodation, this would further worsen the physiological cues problems (Wann and Rushton et al. 1995).

Moreover, stereo display requires careful calibration to provide accurate distance information. Even so, it has been reported that some individuals have difficulty in rapidly processing stereo depth cues although the observers have normal stereo ability (Surdick, Davis et al. 1994). As a user is often presented with a different image to each eye, rivalry between images may sometimes occur (Kalawsky 1993). This happens when the stimuli from one eye is dominant with a corresponding suppression of the stimuli from the eye. (Schiffman 1990). This may cause additional discomfort to the user (Stanney, Mourant et al. 1998).

Another drawback of stereo image presentation is the costly hardware and software requirements for stereo viewing. A powerful computer is required to render two images within acceptable frame rate. Thus, whether the increase in hardware cost and rendering time, visual and other related problems experienced by users, justifies the benefits of stereo images is a critical design decision. While it is most often technically possible to generate stereo images which look realistic, decision for its use must received serious considerations.

Zeltzer, referring to works done in his lab, reported that it has been shown that well-designed 2-D presentations have consistently lead to better performance than stereoscopic displays of 3D scenes for certain air traffic controller tasks (Lantz 1996). It has been suggested that performance may be as good as when stereoscopic information is present to when it is not present (Kim et al. 1987, cited in Howard and Rogers 2002). Other researchers found that stereoscopic presentations do not improve performance for altitude and depth judgments (Hendrix and Barfield 1995). They attributed this to the limitations of the current technology where inconsistent accommodation and vergence cues lead to eye fatigue and strain over longer periods of exposure. In his review of the literature, Pfautz (2000) however indicated that stereo is beneficial for a number of tasks. This may suggest that the benefits of stereo

images may be task dependent. These suggest that not all tasks require stereoscopic display to improve performance.

Stereo cues have been intensively researched compared to other cues (Cutting and Vishton 1995). Many previous researches incorporate stereo cues in their investigations (Henry and Furness 1993, Lampton, Bliss et al. 1994, Lampton, McDonald et al. 1995, Witmer and Sadowski 1998, Sinai, Krebs et al. 1999, Eggleston, Janson et al. 1996, Willemson and Gooch 2002). The use of stereo cues in computer graphics imagery is often questioned (Hsu, Pizlo et al. 1994); this may explain why it is the focus of much research more than other cues. This merits the investigation of the impact of other cues. Additionally, due to the hardware complexities and costs and user-related issues related to stereoscopic displays discussed earlier, the approach taken in this thesis is to conduct comparative evaluation on non-stereo images.

**Research scope:**

*The type of VE and real images used in the investigation is limited to non-stereo images only*

#### 4.1.2.6 Display type for image presentation

Another issue highlighted in Chapter 3 concerned with the displays used to present the VE model. While it is possible to generate an image with considerable realism, this does not guarantee the images will be displayed accurately. The reason for this is the display technologies are limited in terms of spatial resolution, absolute and dynamic luminance range and color gamut (Greenberg 1999). In Chapter 3, a review of the available displays for VE was presented. Based on the devices and the level of immersions, the display types are grouped into three categories: non-immersive systems, semi-immersive systems, and fully immersive systems. The display related factors have been the focus of past investigation into examining factors affecting spatial perception. Though some investigated these factors in the context of different display systems (Henry and Furness 1993, Johnson and Stewart 1999, Riley and Kaber 1999, Patrick, Cosgrove et al. 2000), other researchers conduct their studies based on fully immersive systems (Willemson and Gooch 2002, Heineken and Shultze 2000, Wright 1995). The literature suggests that a fully immersive system provides users with greater immersion and sense of presence (see earlier discussions in Sections 4.1.2.1). These immersions and sense of presence are often enhanced through the stimulation of other human sensory channels such as auditory, haptics and kinaesthetic. The greater immersion and sense of presence has been suggested to enhance the user's performance. However, there are several issues and problems related particularly to fully immersive systems which might hinder user



performance. As discussed in Section 3.2.3 of Chapter 3, the sense of immersion is influenced by several factors. For HMD, often there is a trade-off between FOV and image resolution. Large FOV which could increase users' immersion would result in lower image resolution. Additionally the use of HMD is limited to single users only. Other types of fully immersive systems which could accommodate a group of users are very expensive and few in number. In general fully immersive systems require very high performance graphics software and hardware to achieve acceptable realism in terms of image and interaction which make them costly to acquire. Additional problems include limitation of the current software and hardware. For example the lag and tracking error in the tracking systems; these not only reduce the user's sense of immersion and presence they also affect the user's health. Users of fully immersive systems were often subjected to health and safety problems (see Stanney, Mourant et al. 1998 for a review). The side-effects experience by non-immersive systems and semi-immersive systems are less severe; this is often limited to the problems associated with the use of normal desktop system. But prolonged exposure to large projected displays could lead to eye strain and headaches (Costello 1997). Considering the issues related to the immersive systems, the approach taken in this thesis is to conduct investigations into non-immersive systems and semi-immersive systems which correspond to small and large display respectively.

**Research scope:**

*The display types used for presentation images are limited to non-immersive and semi-immersive displays*

No head-tracking was employed in our experiments due to the problems related with head-tracking (such as head-tracking errors and lag) which may reduce realism and subsequently affect user's performance. Other reason includes to remove the confounding effect of these factors and to focus on the investigated factors. However, the importance of head motion parallax as an effective cue is acknowledged and interested readers are referred to the work of several researchers (Bakker, Werkhoven et.al. 1997; Bakker, Werkhoven et.al. 1999; Bakker, Werkhoven et.al. 2001; Bakker, Passenier et.al. 2003; Groen and Werkhoven 1998; Werkhoven and Groen 1998; Werkhoven and Groen 1998b) who investigated the effect of head-tracked (that is the effect of motion parallax) on perception and navigation in immersive VE.

### 4.1.3 VE Images modelling

The general aim of this thesis is to examine factors affecting spatial awareness in the real and VE by comparing spatial tasks performed in both environments. As mentioned earlier, one of the goals of VE technologies is to provide a synthetic experience indistinguishable from the

real world by matching the capabilities of human sensory channels (Durlach and Mavor 1995). As the visual sensory channel represents the most dominant sensory channel (Pfautz 2000) compared to other channels (such as auditory, tactile, haptics), the focus of this thesis is on matching the visual perception in the real and VE. In order to provide a convincing simulation, Kessler (2002) suggested that image presentation must provide enough detail to make the objects easily recognizable and enough objects to give the user the sense of 'being there' or sense of presence. Additionally, the VE systems must present the user's current view of the virtual scene in acceptable frame rate and to be useful VE must response to the user in a similar manner to the real world. Ideally it should be able to accurately emulate the real world's counterpart in terms of image and behaviour presentation with a high degree of realism. To emulate the real world environment with a high degree of realism involves an accurate simulation of all its aspects. This includes producing accurate geometry of objects as well as colour, texture and lighting (Vince 1995). However, current systems are still far from ideal. Current VE technology is incapable of replicating the real world environment with such a degree of realism. While it is possible to create high image realism using computer graphics techniques, VE technologies are constrained to the generation of such images in real-time. Real-time refers to the presenting and updating of images according to the observer's current view. To present a VE with a high degree of realism in real-time would require a very powerful computer workstation to process such an environment with an acceptable frame rate. Thus, the challenge of the VE design is the trade-off between system performance and image realism (Marzuryk and Gervautz 1996). In general, poor system performance is often unacceptable for real-time VE. As such, a slight decrease in image realism is often acceptable for most applications.

Various techniques are presented on how to improve image realism in Chapter 3. As image realism may influence a user's sense of presence, which may in turn influence his performance in the VE (Slater, Linakis et al. 1996), in this thesis, the construction of VE models takes into considerations some of these techniques by balancing the choice towards maintaining acceptable system performance. Slater et al. 2001 (cited in Garau 2003) breaks down realism in VE into three aspects: geometric realism, illumination realism and behavioural realism. To reduce modelling complexity and time, in this thesis the objects in the VE are static, as such the behavioural realism aspect is excluded from our VE modelling. Thus, the focus of modelling is on geometric detail and illumination realism. A detailed modelling of an object's geometry would increase its image realism but at the expense of the system performance, due to the increase in polygon counts. Moreover, the modelling process would not only be tedious but also very time-consuming. The use of texture mapping techniques (including the use of billboard geometry techniques) seems to be a viable and

attractive solution to incorporate high image realism without the compromise of increasing geometric modelling complexities. Many VEs have a cartoon-like appearance because they lack fine details such as texture (Witmer and Kline 1998). This non-realistic appearance might reduce the user's sense of immersion and thus his performance.

As such, in this thesis, the VE models developed incorporate digital images from real world scenes as texture maps to cover objects' surfaces. Such texture maps include images from grass, sky, trees, roads, and other objects' textures. Besides improving image appearance in terms of realism, texture mapping results in substantial reductions in modelling time, memory and processing speeds. Even though some researchers found no effect of textures (Witmer and Kline 1998), several studies have indicated that the presence of texture improves performance (Sinai, Krebs et al. 1999, Kline and Witmer 1996). Gibson (1950) (cited in Sekuler and Blake 1994) suggested that texture gradient provides information about distances and slant surfaces to the user as well as object size. Though, it is not the focus of this thesis to investigate the effect of texture on the user's performance, however the use of texture would improve image visual realism. As the approach taken in this thesis is to compare and evaluate performance in order to determine if the VE can be perceived similarly to the real environment, such images would provide a more reasonable comparison and evaluation. Moreover, some researchers have shown that there is no significant difference between computer-generated VE (with some form of realism) to the photographed-based VE (Willemson and Gooch 2002) but others who reported a difference used a less realistic or simple VE model.

It has been suggested earlier that to be convincing the VE should be populated with enough objects to increase the feel of being in the environment (Kessler 2002). However, the visual clutter may get in the way of the user performing the tasks. Ruddles and Jones (2001) found that users suffer from disorientation in small-scale cluttered VE thus hindering navigational tasks. This implies it might not be necessary to have many objects in the VE. This has the advantage of reducing scene complexity and increasing frame rate. As such the approach of this thesis would be to choose a real scene location that has less visual clutter so that the resultant VE model would also have less visual clutter.

Besides creating a more realistic representation, shadows provide depth and perspective cues to the viewer. Because this thesis involves the examination of space perception, the addition of such cues in the image would increase the accuracy of spatial judgement (Kunnapas 1968). Thus, shadow was also implemented in this thesis. Realistic shadows are still difficult to implement in real time due to computational overhead (Vince 1995). Moreover, there is empirical evidence to suggest that a shadow's shape (polygonal verses true realistic shadow)

has little influence over the user's performance on perception of object size and position (Hubona, Wheeler et al. 1999). Thus, due to these reasons and its simple implementation, shadow generation in this thesis is limited to a false shadow (created using static polygonal model).

Another technique, gourand shading, was adopted due to its less software complexities as compared to the more accurate models of Phong techniques, ray tracing and radiosity techniques. As such the lighting effects in the VE models described in this thesis were not modelled with high degree of accuracy. However, the lighting effects have been investigated by other researchers (Meyer, Rushmeier et al. 1986, Mania 2001, McNamara, Chalmers et al. 2000, Longhurst, Ledda et al. 2003, Lo, Chalmers et al. 2003). It has been demonstrated that when the lighting effect is closely modelled, subjective responses on a lighting questionnaire do not yield significant difference between real and virtual condition (Mania 2001). The author further showed that there is a positive correlation between presence and lighting for the virtual condition (HMD-monocular viewing) but this is not true for real, desktop and HMD-stereo conditions. Other researchers (Meyer, Rushmeier et al. 1986) have found that participants considered the match between a picture of a model and computer-generated picture of it (based on radiosity lighting model) is very similar. The modelling process in this thesis is limited to accurate geometric representation of the real world location with photo-realistic textured objects to create high realism in the image. However, some researchers (Willemson and Gooch 2002) have shown that when a photographed based VE is compared to a computer-generated VE, the difference is very small for spatial judgment tasks. Thus in this thesis, even though the lighting effect is not closely modelled, it is expected these techniques (such as accurate geometric representation, texture mapping and gourand shading) are minimally sufficient to yield similar spatial perception of the VE model to its real counterpart.

The details on how the VE models used in this thesis were created are discussed in Section 4.3 of this chapter. The research in this thesis aim to examine if resultant VE models created using the choice of techniques employed would allow the VE to be perceived similar to the real image/real environment counterparts in terms of the tasks measured.

**Research question:**

*Is there a difference in spatial tasks (distance estimation and spatial memory task) performed in real and VE?*

In the VE, the visual channel is often considered the most important (Pfautz 2000) compared to other channels such as audio, kinaesthetic and haptic cues and many believe that the visual

systems convey more accurate distance information compared to audition, tactile and kinaesthetic senses (Welch and Warren 1986; cited in Surdick, Davis et al. 1997). Comparatively, representation for the senses of sight, sound and odour are easy to develop compared to sensation of touch of any geometric objects (Kessler 2002). Moreover, as almost all VE applications provide a visual display, these cues will be incorporated in the VE models described in this thesis and others will not be represented. Initially, it was intended to include the impact of audio cues on spatial tasks performance but due to time constraints and to limit the number of controlled factors, the research's main focus was on visual factors. However, the absence of these cues may have influenced the users' performances in the VE. For instance the conflicting information given by the visual and kinaesthetic cues when a stationary user views a dynamic image may cause a user to experience cyber-sickness and this may affect the user's performance. The impact of these cues on spatial awareness will not be directly addressed in the research presented in this thesis but the impact of such cues deficiency and conflicts in the VE would be highlighted in the discussion of experiments' results.

#### 4.1.4 Summary of experimental basis/approach

The previous subsections have discussed and argued for the basis of the experimental approach adopted by the research presented in this thesis in terms of the choice of task performance measures employed, factors to be investigated and related issues to VE image modelling. The general aim of this thesis was to examine factors influencing the user's spatial awareness in the real and VE. Based on the VE modelled using the techniques described in Section 4.1.3, the first research question seeks to compare and examine the user's spatial awareness performance in this VE to its real world counterparts. It was presented earlier that comparing task performance in the VE to similar task performance in the real world can provide knowledge and understanding on the limits of the VE technologies (Witmer and Sadowski 1998, Walller 1999, Kalawsky 2000). Therefore the approach taken in thesis is to compare spatial task performance in both environments. Two commonly employed aspects of spatial awareness for task performance measures, distance estimation task and spatial memory tasks, were identified to explore spatial performance in the real and VE. Due to the dearth of studies concerned with exocentric distance, particularly those related to asymmetrical distance tasks, this thesis examines exocentric distances in terms of asymmetrical distances of vertical, horizontal and transverse. Spatial memory was employed as a task performance measure due to the few studies available performing direct comparisons of spatial memory between the real and VE. Additionally, due its appropriateness as a measure of spatial representation in interactive presentation it was used in this thesis as a task measure for the experiment in



interactive environments (Experiment 3, Chapter 7). Discussions in the prior subsections have suggested and argued for the investigations of the following main factors in this thesis: Image type (real versus VE), display size (large versus small), interface device type (mouse vs. trackball) and travel mode (drive vs. fly). From these factors, research questions 2, 3 and 4 were generated in the context of the overall aim of the thesis of investigating spatial awareness in the real and VE.

In summary, the following are the four research questions explored in this thesis:

- 1. Is there a difference in spatial tasks (distance estimation and spatial memory task) performed in real and VE?*
- 2. How does the display size (large and small) affect users' spatial task (distance estimation and spatial memory task) performance in real and VE?*
- 3. How does the type of interface device (mouse and trackball) affect users' spatial task performance (distance estimation and spatial memory task) in VE?*
- 4. How does the type of travel mode (drive and fly mode) affect user's spatial task performance (distance estimation and spatial memory task) in VE?*

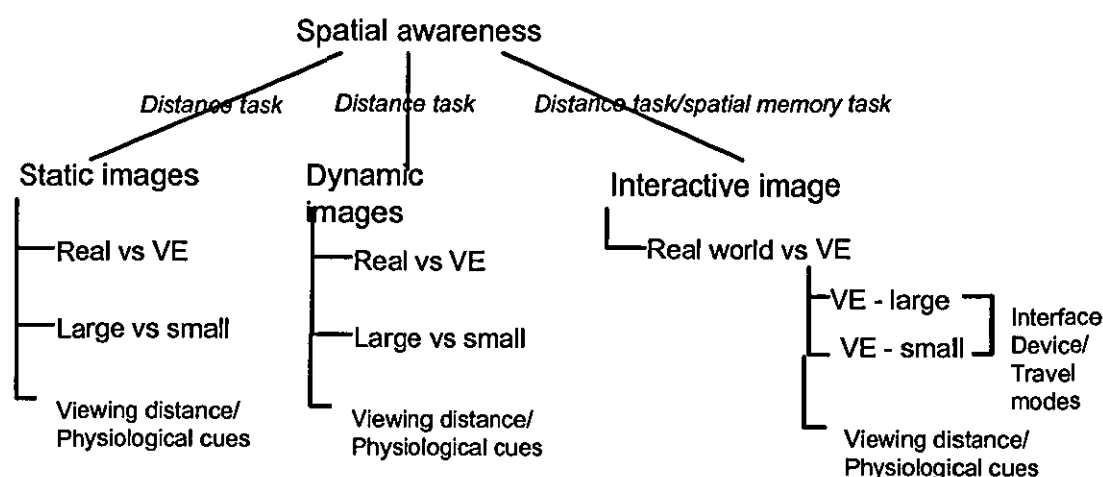
As argued in Section 4.1.2.1, the effect of other factors such as viewing distance and physiological cues would also be examined in order to explicate the ambiguity in the previous investigations regarding whether the better performance of large display over small display was due to the effect of display size factor.

The previous subsections have also discussed and defined the main assumptions and scope of research investigations in this thesis. In summary, the research assumptions and scope are:

- Examination of user's spatial awareness includes the following forms of image presentations: static, dynamic and interactive real and VE
- The type of real and VE images used in the investigation is limited to non-stereo images only
- The display types used for presentation images are limited to non-immersive and semi-immersive displays

As discussed in Section 4.1.2.2, in addition to the study of interactive VE, a common approach employed by previous investigations, it is also important to understand spatial perception in static images. This is because the results of the former may not extend to the latter. Due to these possible differences in user's spatial performance in the various contexts of image presentations, three types of image presentations (static, dynamic and interactive)

were examined in this research. Moreover, examination of the three types of image presentations would provide a comprehensive evaluation of spatial awareness in the context of how the users work in VE. Because three types of image presentations were considered in this thesis for the examining of factors of spatial awareness in the real and VE, three sets of experiments corresponding to these three types of image presentation were developed and conducted to explore the four main research questions stated earlier. The first experiment examined spatial awareness (in terms of distance estimate task) in the context of static real and VE images. The second experiment also examined spatial awareness in terms of distance estimate tasks but in the context of dynamic real and VE images. Finally, the third experiment will examine spatial awareness in an interactive real physical environment and an interactive VE. Additionally, the influence of an interface device and travel modes in the VE will also be investigated in the third experiment. These experiments are described further in the next section and in the experiment chapter of 5, 6 and 7. Figure 4-6 presents a summary of the overall research approach taken in this thesis, which includes tasks and factors examined.



**Figure 4-6 Factors and tasks performance investigated in this thesis**

Knowledge gained from the research in this thesis would further augment the existing literature and provide guidelines for designers and users of VE applications on factors which contribute towards cost effective use of VE and human performance efficiency in the VE.

## 4.2 EXPERIMENTAL METHODOLOGY

The methodology employed in the work presented in this thesis is based on the experimentation or hypothetico-deductive approach in which theories (general explanations of phenomena) are evaluated by generating and testing hypothesis (Coolican 2001). In examining factors affecting spatial awareness in the real and VE, the research presented in this thesis addressed four main research questions (described in the previous sections). These

research questions were addressed in the context of static, dynamic and interactive presentations. As proposed earlier, three sets of experiments of which correspond to each of these presentations were undertaken. The specific hypotheses generated from the research questions related to each of the three experiments are found in Table 4-4 in this chapter and in the experiment chapter of 5, 6 and 7.

The first experiment aims to examine factors affecting spatial awareness in the context of static images. Pertaining to research questions 1 and 2, Experiment 1 series seeks to examine the effect of image types (real and VE) and display sizes (large and small) on user's asymmetrical distances estimates. The literature reviewed is not clear on the effect of image type on user's perceptions. Some studies revealed that a VE is perceived differently from the real environment, while others suggested that it is possible to perceive the VE similarly to its real counterpart. However, since the VE models used in this research were closely modelled to the real world place in terms of geometric representations and textures and as demonstrated by several researchers (Willemson and Gooch 2002), it is expected that the difference between spatial task performance in the real and VE is small. Based on the experimental approach proposed earlier (see the last paragraph of Section 4.1.2.1), Experiment 1 comprises of two sub experiments, namely Experiment 1A and Experiment 1B. Experiment 1A investigates the effect of display size on distance estimate task while Experiment 1B examines the possible influence of viewing distance and physiological cues on distance estimate. Both studies aim to clarify the ambiguity of previous investigations regarding whether the display size factor is responsible for the better performance of large display over small display. Theoretical predictions and the literature suggested that both viewing distance and physiological cues might have also contributed an influence on spatial perception beside the display size factor.

Similarly, Experiment 2, which comprises two sub experiments (Experiment 2A and 2B), is based on the same premises and addresses research questions 1 and 2. However the effect of image types and display sizes on asymmetrical distances were examined in the context of dynamic images. As with Experiment 1A and 1B, Experiment 2A and 2B also investigates the effect of display size and the possible influence of viewing distance and physiological cues.

Finally, Experiment 3 comprehensively addresses all the four research questions in the context of interactive presentations. Experiment 3 addresses research question 1 and 2 by examining the effect of image (environment) types and display sizes factors on asymmetrical distance perception and spatial memory tasks in interactive real and VE. Similar to Experiment 1 and 2, Experiment 3 comprises of two sub experiments: Experiment 3A and 3B.

Additionally, since Experiment 3 series (Experiment 3A and 3B) was conducted in the context of interactive presentation, the related issues of interacting and exploring the VE was also examined. Therefore, in addition to image types and display size factors, the effect of device types and travel modes factors on spatial awareness in VE were also investigated. Examinations of the two later factors (device types and travel modes) seek to answer research question 3 and 4 respectively.

This research deals mainly with a quantitative approach to the examination of factors affecting spatial awareness in the real and VE. However, in addition to this quantitative approach, a qualitative approach (post-test questionnaires) was also employed. According to Kalawsky (2000a), a qualitative or subjective approach which reflects subjective opinions of the participants often yields important information which is not obtainable by other means. Thus, in this thesis, a qualitative approach is also included in order to gather additional information to help further explain the user's spatial perception and interaction with these images.

In the next sub-sections, the specific choices of experimental methods employed to address the research questions based on the three series of experiments proposed were described. Research methods here refer to the specific techniques used to collect and analyze data. The three sets of experiments share the same overall goal of examining factors affecting spatial awareness in the VE. Thus, they share some similarities in terms of experimental variables, setup and procedures, data collection and analysis. This is especially true for Experiment 1 and 2. As such, in the next sub-sections an overview of the experimental methods which are common to all experiments are described but the details on methods related to specific experiments will be presented in the respective experimental chapters (Chapter 5, 6 and 7). The first subsection (Section 4.2.1) describes the data collection process which includes description of experimental variables, experimental designs, participants, images/models, apparatus/room settings, experimental procedures and the post-test questionnaires. The second subsection (Section 4.2.2) deals with data preparation and analysis. Finally, the final subsection provides a summary of all experiments undertaken by the research in this thesis.

#### 4.2.1 Data collection

This section will cover aspects of experimental design and procedure common to all three experiments. Chapter 5, 6, and 7 describe other details that are specific to each experiment.

#### 4.2.1.1 Independent and dependent variables investigated

The overall aim of this research is to investigate factors affecting spatial awareness in static, dynamic and interactive images. The independent variable (IV) and dependent variables (DV) for all experiments are summarized in the Table 4-1.

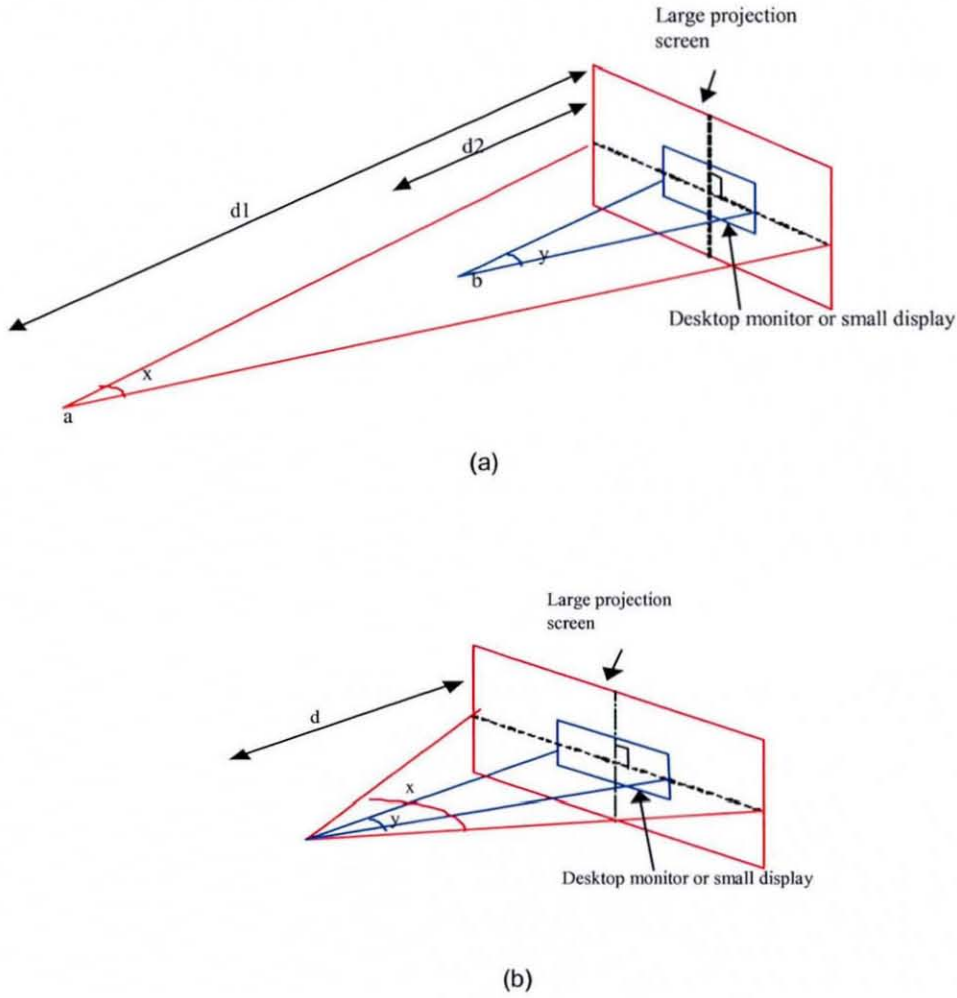
**Table 4-1 Summary of IV and DV variables examined in all experiments**

Experiment	IV	DV	Data collected
1A	Image type Display type	Estimated distance	1. Horizontal and Transverse distance 2. Post-test questionnaire (Self reported comments on distance estimate)
1B	Display type	Estimated distance	1. Horizontal and Transverse distance 2. Post-test questionnaire (Self reported comments on distance estimate, sports background)
2A	Image type Display type	Estimated distance	1. Vertical, Horizontal and Transverse distance 2. Post-test questionnaire (Self reported comments on distance estimate, sports background)
2B	Image type Display type	Estimated distance	1. Vertical, Horizontal and Transverse distance 2. Post-test questionnaire (Self reported comments on distance estimate, sports background)
3A	Display type Image type Device type Travel mode	Estimated distance Scores on spatial memory test Scores on questionnaire	1. Map test scores 2. Interface device questionnaire 3. Distance estimate (height, width and length of room which correspond to vertical, horizontal and transverse distance)
3B	Display type Device type Travel mode	Estimated distance Scores on spatial memory test	1. Map test scores 2. Interface device questionnaire 3. Display questionnaire 4. Distance estimate (height, width and length of room which correspond to vertical, horizontal and transverse distance)

While it is acknowledged that the effects of contrast and brightness are important (Adelson 1993; Tan, Gergle et al 2003), however, in order to focus investigation on the earlier mentioned factors, these factors are controlled. This is possible due to the setup of the experiment for both large and small display and the use of a rear-projection screen. For small display condition, the LCD projector is located directly behind the screen, yielding a very bright image. While greater contrast/brightness may result in object appear closer or more accurate but at certain point this accuracy drops. Thus, in the small condition, performance may be degraded. For large display condition, the LCD projector is not behind the screen and the projected image is less bright (or contrast) due to it large size, which also led to slight performance degradation. Thus, from both setups, it is expected the effect of contrast/brightness in small/large display condition is cancelled out. As such, the effects of these factors will not be further discussed.



### 4.2.1.2 Experiment setup



**Figure 4-7 Experiment setup for all experiments**

As discussed in Section 4.1.2.1, two experiment setups will be employed based on the variables that need to be controlled. These setups are as illustrated in Figure 4-7.  $d_1$ ,  $d_2$ , and  $d$  are viewing distances.  $x$  and  $y$  are FOVs. In setup (a) of Figure 4-7, the FOV (and retinal image size) of both displays size were fixed but the viewing distance (and physiological cues) was varied. This setup was employed by Experiment 1A, 2A, and 3B. For setup (b) of Figure 4-7, the viewing distance (physiological cues) were fixed but the FOV (and retinal image size) of the displays were varied. This setup was employed by Experiment 1B, 2B and 3A.

### 4.2.1.3 Experimental design

Two types of experimental design were employed in this thesis: between-subjects factorial design and mixed design. One advantage of between-subject design is that it avoids carry over

effects (interference from previous knowledge) or training bias. However, one major weakness of between-subject designs is the participant variables could be a possible source of variation among groups which may increase the chances of non-significant findings. It has been recommended that performing random allocation of participants to experimental conditions would reduce the likelihood of participants' variation (Coolican 2001).

In this thesis, because the same scene is used for all conditions, a between-subject design (that is different group of participants were used for each condition) was used for the design of Experiment 1 and 2 to avoid the carry over effects. Additionally to reduce the influence of participants' variation, participants were randomly allocated for each condition. Each factor in a between-subject design experiment represents a major IV under investigation, such as image type and display type. The factor may consist of different levels. For example, both image type and display are made up of two levels each. Thus, for experiment 1A, 2A and 2B, these two factors yield four different experiment conditions (Figure 4-8). The conditions refer to the different combinations of the levels of the factors. Experiment 1B investigated only one factor, display type. As such there are only two conditions representing each level of the display type (Figure 4-9).

		Factor 1	
		Level 1	Level 2
Factor 2	Level 1	Condition 1	Condition 2
	Level 2	Condition 3	Condition 4

**Figure 4-8 For Experiment 1A and 2: 2 x 2 factorial designs (2 factors: each factor has two levels; 4 experimental conditions). Factor 1 is image type (real and VE). Factor 2 is display type (desktop monitor and projected display)**

		Factor 1
Level 1	Level 2	Condition 1
		Condition 2

**Figure 4-9 Experiment 1B: One-factor design with two levels (2 experimental conditions). Factor 1 is display factor (large and small)**



		Factor 2		Factor 3	
		Level 1	Level 2	Level 1	Level 2
Factor 1	Level 1	Condition 1	Condition 2	Condition 5	Condition 6
		Condition 3	Condition 4	Condition 7	Condition 8
	Level 2	Condition 1	Condition 2	Condition 5	Condition 6
		Condition 3	Condition 4	Condition 7	Condition 8

**Figure 4-10 For Experiment 3: 2 x (2 x 2) mixed design (3 factors: each factor two levels; one between-subject and two within subject design; 8 experimental conditions for each between-subject factor level). Factor 1 is display factor (large and small). Factor 2 is interface device factor (mouse and trackball) and Factor 3 is travel mode factor (drive and fly). It should be noted that the fourth factor (image type) is not depicted here. This is because the real condition does not investigate factor 2 and 3.**

Experiment 3A used a mixed design, a combination of between-subject and within-subjects design (see Figure 4-10). The number of IV investigated in this experiment was four (image type, display type, interface device and travel modes), each comprising of two levels each. One reason for employing mixed design and not full between-subject design is that a full between-subject design requires a large number of participants (a group of ten participants would require about 160 participants). A between-subject is still used however for the display factor to maintain consistency with experiment 1 and 2. To reduce the number of participant's requirement, a within-subjects design is used for interface device and travel mode factors. This not only requires fewer participants but it allows participants to compare between interface devices used and travel modes used directly. This is possible because in within-subject design, the same participants will experience all the repeated factor conditions; thus, they can compare between the factors. However, one drawback of the within-subject design is the time for each experiment session would be much longer for each participant. For example, each participant needs to repeat all conditions for a fully within-subject design (that would be a total of 16 conditions for Experiment 3A design). A pilot study revealed that an eight-condition session requires about one and half to two hours, thus a fully repeated design would double the test session time. A longer time would make participants bored and tired and this may affect their overall performance. This is one reason why the mixed design which is a combination of between-subject and within-subject design was chosen. Another drawback of within-subject design is the order effect. This effect occurs from the order in which participants performed the conditions. For example, participants might improve on the later condition(s) because they had practice in the earlier condition(s) or they might perform worst in the later condition due to boredom or fatigue (Coolican 2001). As suggested by Coolican (2001), to reduce this effect a counterbalancing of the conditions was employed. For example, in this thesis, half of the participants used the mouse first, followed by the trackball while the other half used the trackball first, followed by the mouse. While the order effect is not

completely eliminated, the improvements due to practice (or low performance due to fatigue) in each condition would cancel out each other. Thus, this method yields results with the effect that is under investigation.

The design of Experiment 3B is, however, a fully within-subjects design. In Experiment 3A, the display factor is a between-subject factor and other factors are within-subject factors. In Experiment 3B we decide to make the display factor a within-subject factor. This was to allow participants to experience both display size conditions and compare them in a display questionnaire. In order to reduce possible effect of participants' boredom and fatigue, the experiment was conducted over a two-day period. A counter-balanced design was employed to reduce order effects.

#### 4.2.1.4 Participants

In their review of several studies, Richardson, Montello et al. (1999) found that there exist weak correlations between pencil-and-paper tests of spatial abilities and measures of environmental spatial ability such as learning the layout of a novel environment. This led them to conclude that there is currently no psychometric spatial abilities test that is a good predictor of environmental spatial ability. Other studies have shown weak correlation between spatial ability test results with performance (Riley and Kaber 1999). As such this thesis does not screen participants for their spatial ability.

The sample sizes (Experiment 1 & 2 – 40; Experiment 3 -32 &10) selected in this thesis was based on previous investigation for similar studies on distance estimate and spatial memory.

The pool of participants was taken mainly from the staff and students of Computer Science Department of Loughborough University. This was to reduce variance among participants in terms of computer knowledge and experience. The participants employed for the studies conducted in this thesis were either volunteers or paid volunteers. All participants either have normal or corrected-to-normal vision.

As reviewed in Chapter 2, participants' previous experience, knowledge and expectations on different display conditions may cause cognitive dissonance with regards to the experimental set-up of different display size and image type condition. However, it would be difficult to control all participants' beliefs, previous experiences and expectations and how they would react to the experimental setup. But it was expected that by having randomly selected samples and randomly assigning different groups of participants for each condition would minimize such effects.

#### 4.2.1.5 Real and VE image preparation/modelling

##### Scene locations

Most previous studies that compare real and VE using distance estimation tasks (Witmer and Kline 1998, Kline and Witmer 1996, Lampton, McDonald et al. 1995, Witmer and Sadowski 1998) compared environment based on a real world indoor-setting scenes or computer-generated scenes only with no real world counterpart (Waller 1999, Sinai, Krebs et al. 1999, Eggleston, Janson et al 1996, Heineken and Shultze 2000). The results of such studies do not necessarily extend to outdoor settings. Indoor settings (such as rooms, hallways or corridors) most often have standard heights and sizes (Henry 1992); while the features in an outdoor setting such as trees, hedges lampposts, roads, signposts are often of variable heights and sizes. As such, the differences between indoor and outdoor setting in terms of available information cues for distance estimation tasks might yield different results. Moreover, earlier studies by Tehgtsoonian and Tehgtsoonian (1969) and Tehgtsoonian and Tehgtsoonian (1970) indicated that there is a difference between perception of distance in an indoor and outdoor environment on distance perception using verbal report. They compare performance across varying distances. Their result showed that participants tended to overestimate in an indoor setting and underestimate (but more accurately) in outdoor setting. It is unknown if these findings will replicate in a VE. Thus, for these reasons, for the distance estimation experiments (Experiment 1 and 2), the scene locations chosen were outdoor settings.

It is noted however that there are two drawbacks to conducting the experiment in an outdoor setting. First, the environment might change due to physical or natural cause before the experiment is completed. This actually happened to our locations for Experiment 1 and 2. Secondly, conducting the experiment in an indoor setting gives the experiment more control of the real environment conditions, whereby the rooms or halls used in the experiments could be made inaccessible to others while the experiment is in progress. This is more difficult to do for an outdoor setting. The location itself might be subject to changes during the course of the experiment and it would be no longer comparable to the VE model of it. Additionally, possible undesirable distractions might occur during the course of the experiment which might have an effect on the results. As such, to avoid this problem, that is to allow possible comparison between real and VE, we drew upon an analogous situation in the crime scene investigation application where the investigating officer might take pictures and video movie of the crime scene locations<sup>2</sup>. Thus, for the outdoor setting experiments we made comparative

---

<sup>2</sup> George Shiro, Forensic scientist at Louisiana State Police Crime Laboratory, USA. Examination and Documentation of the Crime scene. Available at [www.crime-scene-investigator.net](http://www.crime-scene-investigator.net).



evaluations of pictures (real and virtual picture) and video movies (computer animations for the VE conditions). This method not only provided an alternative for comparison in an outdoor setting but also had practical implication in the application.

For Experiment 1 and 2, a location with few visual cues but with an adequate number of objects was required. An image with more objects in it might have provided more visual clues to participants in their estimation and have created more variance among participants and a less controlled experiment. This is because different participants might have used different objects to base their estimations on. Thus, the presence of many different objects would create more participants' variance, where different participant might not use the same objects to base their distance judgments. Few visual cues were thus necessary to reduce variance among participants and focus on the impact of the variables (image type and display type) under investigation. Additionally, as reviewed earlier, the visual clutter may get in the way of the user performing the tasks. Three separate locations in Loughborough University, which met this requirement, were identified. These locations are described in the respective chapters (Chapter 5 and 6).

However, for evaluation between the actual scenes with a virtual model, an indoor setting is the better choice due to more control of the experiment. This is especially necessary when conducting the real environment condition and to avoid the problem related to outdoor setting as discussed earlier. Thus, an indoor setting was selected for Experiment 3. The few visual cues requirement was also imposed. An additional requirement was a large empty space. A room in one of the university buildings was identified to meet such requirements and was chosen for Experiment 3.

### Real image preparation

For both Experiment 1A and 1B, a static image of the scene was taken using a digital camera. This image was placed on a Microsoft PowerPoint slide (in full-screen mode) for the real picture condition.

For the dynamic image condition of Experiment 2A and 2B, a video movie of the scene was taken using a digital camcorder. The movie was downloaded into a computer, edited using Adobe Premiere 6.0 and was saved in .AVI format.

For Experiment 3A, the real condition utilized the identified room earlier. Thus, no image preparation was necessary. However, to avoid interruption during experiment, the real condition experiment was conducted over the weekend only. Prior to the experiment, the

room was checked to ensure it was still similar to its modelled VE version. The real condition experiment was conducted over a period of three weekends.

## VE image modelling

For all experiments, a computer-generated VE model of the identified scene location was developed using REALAX RXScene (Experiment 1) and Multigen-Paradigm modelling software (Experiment 2 and 3). Prior to the modelling process, careful measurements of the objects size and locations were taken. These dimensions were used to create the VE models. The manual creation of VE models of real scenes presented a very tedious and time consuming task. This was one reason for not choosing a heavily cluttered scene location for the experiment. The following techniques were employed to create VE model of sufficient realism without compromising on system performance:

1. Texture maps were created from photographs of objects of the real scene and were projected onto the modelled objects to give the VE model more detail and realism. The photographs were edited using Micrografx Picture Publisher software. The respective texture of each object was exported to MultiGen II Pro software to generate the texture maps.
2. Billboard geometry and billboard techniques were employed. For background scenes, a picture of the background was taken to be used as the billboard. For objects such as trees, the picture of each tree was placed on the planar surface (with background transparent effect) and this surface was given a rotational transformation so that during simulation it would always face the user. As discussed in Chapter 3, both techniques increase image realism without corresponding increase in modelling time, memory and processing speeds.
3. Shadows were created using a set of polygons. Even though the resultant shadow was less realistic, this technique was easy to implement. As reviewed in Chapter 3, it is possible to create realistic shadow based on ray-tracing techniques but this result in increase in computational overhead.
4. For object shading, gourand shading was used. The realism provided by Gourand shading is considered acceptable for most applications. Other techniques mentioned in Chapter 3 include ray-tracing and radiosities do provide high realism but the slow rendering speeds make them unsuitable for real-time applications.

Due to the less complex scene and small size of the VE modelled in this thesis, the LOD technique (discussed in Chapter 3) was not implemented in the modelling process. Details and

issues related to the VE modelling process particularly for Experiment 2 and 3 VE models were discussed in their respective chapters of 6 and 7.

For presentation to viewers, a snapshot of the VE model was taken (using print screen command) and placed on a Microsoft PowerPoint slide to represent the VE picture condition in Experiment 1A. For Experiment 2A, a recorded movie of the simulation of movement through the VE was taken on the VHS tape and transferred to pc in .AVI format. However, for Experiment 2B, the actual simulation of the movement through the VE using the Silicon Graphics Inc. Performer PERFLY software was used. Similarly, the VE model in Experiment 3 was viewed using the same viewer software whereby the actual VE simulation was used instead of pictures or recorded movie of it.

### Stimuli used for distance estimation experiments

Most stimuli used for space perception studies employ a very narrow set of stimulus: thin poles, columns or cylinder, circles or discs were often used (Hecht, van Doorn et al. 1999). This is also true for an outdoor setting where often thin poles were used as target to estimate from. As mentioned earlier outdoor settings were proposed for distance estimation investigations (Experiment 1 and 2). Instead of selecting from previously employed stimuli, in this thesis a set of stimuli comprising of natural objects present in the natural scene such as trees, hedges, lampposts and roads were used. While the stimuli used was not typical of those used for space perception studies, Hecht, van Doorn et al. (1999) suggested these new set of stimuli used in the present study would allow the expansion of the list of stimuli used for visual perception studies and may allow reinterpretation of previous findings. Additionally, the authors conclude from previous evidence that different objects are likely to affect perception of subjective space differently and this indicates the need for further investigations.

#### 4.2.1.6 Display apparatus and room setting

With the exception of Experiment 1A which used a desktop monitor and projected display for small and large display respectively, other experiments (1B, 2 and 3) used a rear-projected screen for both large and small conditions with the projected image sizes adjusted according to large and small display conditions.

Since Experiment 1A was the initial exploratory experiment, the experiment was conducted under normal condition. However, for later experiments, a dark room setting was employed for all VE conditions. A dark room setting was necessary to reduce the peripheral view effects

from objects surrounding the screen. It had been suggested that these peripheral view effects might affect participants' distance estimates (Eby and Braunstein 1995, cited in Knapp 1999).

#### 4.2.1.7 Experimental procedures

In this section, the aspects of procedure which were common to all three experiments were described. However, further details of the procedure which were specific to each individual experiment were given in each respective chapter.

For all experiments, participants were first informed of the purpose of the experiment. They were also told that they could withdraw from the experiment at anytime without having to give any reason. Before the start of each experiment participants were first given a form to fill in about their personal information (name, age, gender, staff/student, etc). They were later given an instruction sheet describing the experimental procedure. They were encouraged to ask to clarify any question they had prior to the start of each experiment. They were also informed that all data collected would be confidential and would be used for data analysis and reporting only.

Before the start of each experiment, participants were asked if they were ready to begin the experiment trial. The participants were then ushered to the designated seat which was adjusted according to their height to ensure that their eye level is at centre of the image. For all experiments, participants were told to restrain from head and body movements during trials in order to remove the effect of motion parallax cues particularly from head motion. As reviewed in Section 2.3.2 of Chapter 2 showed the motion parallax cues are salient for spatial tasks such as distance estimates. Additionally participants' forward/backward movements would change the predefined FOV size. As such it was necessary to control for the confounding effect of these cues. However, the effect of these cues was acknowledged.

All three experiments in this thesis employed non-stereo images, where images were presented non-stereoscopically. After completion of the given experimental tasks, participants were asked to fill in a post-test questionnaire. Participants were reminded not to talk about the experiment to other potential participants as this might affect the latter's performance.

#### Methods of assessing distance perception

There are two broad categories of measuring distance perception: *direct and indirect method*. Knapp (1999) in his thesis dissertation provided a review and empirical comparisons of these

methods. A brief overview of these methods and results of his investigations are summarized in Table 4-2 and in the follow paragraphs.

**Table 4-2 Methods of measuring distance perception: comparison of direct and indirect methods**

Methods of measuring distance perception	Examples
<b>Direct Method</b> <ul style="list-style-type: none"> <li>participants were aware that they were asked to performed distance related task</li> </ul>	<ul style="list-style-type: none"> <li>Verbal report <ul style="list-style-type: none"> <li>-For verbal report, participants gave a verbal (or written) report of the distance in known units (such as metres or feet).</li> </ul> <i>Advantage:</i> <ul style="list-style-type: none"> <li>-face validity, clear what it intend to measure</li> </ul> <i>Disadvantage:</i> <ul style="list-style-type: none"> <li>-It assumed that the participant has internalized the metric of interest.</li> <li>-Influence by cognition</li> </ul> </li> <li>Visually-directed motoric behaviour <ul style="list-style-type: none"> <li>- participants were asked to walked either directly (the most common employed) or indirectly towards a previously seen targets</li> <li>- the distance walked represent the perceived distance.</li> <li>- variations of these methods includes combine waking with other tasks and triangulations by walking.</li> </ul> </li> </ul>
<b>Indirect Method</b> <ul style="list-style-type: none"> <li>participants were NOT aware that they were asked to performed distance related task</li> <li>Knowledge of performing the required task in direct methods may cause participants to perform differently and this may confound the results.</li> <li>this method avoids this problem by indirectly measuring distance through distance related behaviour.</li> </ul>	<ul style="list-style-type: none"> <li>verbal judgment of size,</li> <li>head motion procedure</li> <li>Judging apparent width of an aperture relative to perceived shoulder.</li> </ul>

In a comparison of these *direct and indirect methods*, Knapp (1999) found that verbal report of size gave the most accurate results, followed by verbal report of distance. Triangulation by walking, however, gave the worst results. Of all the methods presented above, several can be excluded for use in this thesis. The irregular size of target stimulus used in this research (trees, hedges, roads, and lamppost) made verbal judgement of size less suitable and of less interest. Due to the main use of pictures and computer-generated images as environment stimulus, visually directed motoric behaviour method can be eliminated. Additionally, this method gave the worst results. This left us with verbal report of distance. Due to its face validity and accuracy as a measure, the verbal report of distance method was employed in this thesis to assess the distance perception. However, as mentioned earlier this method assumed that the participant had internalized the metric of interests such as metre or feet. To ensure participants had similar internalized measure of metre or feet, participants were shown a metre ruler prior to the start of each trial.



## Methods of assessing spatial memory task

Spatial memory tests are often assessed using sketch maps. However, there are several variations in the implementations of sketch maps. Henry (1992) asked participants to sketch the plan of the gallery after the VE exploration. The maps were analyzed by giving the participants' sketch map a rating based on the number of missing rooms. Similarly, Billingham and Weghorst (1995) also used free map sketch method but more detailed analysis of the maps was conducted. Three sets of scoring methods were used: map goodness, object classes and object positioning. Besides comparison of resultant participants' sketch maps with the actual present a difficult task, these methods also indirectly assessed the drawing ability of the participants which may confound the results.

Other researchers avoid the latter by using a different strategy. Goerger and Darken et al. (1998) asked participants to place numbered magnets (to represent objects) on a metal whiteboard (to represent the room). They considered their method better than free recall and map drawing. Similarly, Rossano and Moak (1998) gave participants a blank sheet of paper (to represent the campus) and small posterboard rectangles (to represent the buildings). After placement of rectangles on the paper, they were required to trace it out and label them. Analyses were done by drawing vectors connecting the centres of the rectangles. These vectors were measured for angles and distance error.

Arthur, Hancock et al. (1997) however, gave maps to participants with scale and orientation information. Maps given contained two of the object position filled. Analyses were done by comparing distance between points (that represented object positions). This method not only avoids the assessment of drawing skills of the participants, which might confound results, but also reduces "performance demands" on the participants by giving them scale and orientation information. It has been suggested that experiment testing spatial representation should employ tasks that minimize performance demands so that the properties of the mental representation can be assessed accurately (Siegel 1981, cited in McNamara 1986).

As such, for the spatial memory task evaluation, the method employed by Arthur and Hancock et al. (1997) (but with slight modification is described in Chapter 7) was employed in this thesis.

### 4.2.1.8 Post-test questionnaires

For Experiment 1 and 2 a short post-test questionnaire was given to each participant at the completion of the experiment (Appendix A and B). The purpose of these questions was to

1. To understand how participants made their estimation
2. To identify which asymmetrical distance they felt was easier to estimate and their reasons for their choice
3. To survey their sport background activities

Information for question 3 was not gathered for the initial exploratory Experiment 1A & 1B. These questions were later added due to recommendations by anonymous reviewers. Participants were asked to rate their own estimation (that is their confidence about the accuracy of their estimate). Additional information on participants' personal background (age, gender and occupation (staff or students)) was also collected.

For Experiment 3, information on participants' personal background was also collected. Additional information such as VE experiment participations and how often they played computer games was also collected. In contrast to Experiment 1 and 2 where the participants were just passive observers of the real and the VE image, in Experiment 3 participants were required to interact with these images. Thus, a questionnaire (Interface device questionnaire) was administered to examine the users' experience and provide an evaluation of the interface devices and travel modes. For Experiment 3B, an additional questionnaire (Display questionnaire) was administered to evaluate the users' experience using both display size. Details of both questionnaires are presented in Chapter 7 and in Appendix C.

## 4.2.2 Data preparation and data analysis

### 4.2.2.1 Data preparation

Preliminary checks of data were conducted to ensure that there was no violation of assumption of parametric tests validity, that is, normal distribution, homogeneity of variance, interval data and independence. Data was checked for outliers by converting the data set into z-scores. Z-scores are a way of standardizing the data set. This was done using the following formula:

$$z = \frac{x - \bar{x}}{s}$$

Each score ( $x$ ) was subtracted from the mean of all scores ( $\bar{x}$ ) and this value is divided by the standard deviation of all scores ( $s$ ). These scores were then used to check which data falls within the limits. In a normal distribution, it is expected 5% to have absolute values greater than 1.96, 1% to have absolute values greater than 2.58, and none greater than 3.29 (Field 2000). Cases of data did not fall within these limits are thus classified as outliers and were removed prior to further analysis.

For distance estimate data, participants' performance accuracy was measured in terms of how close their estimated distance to the actual distance was. Because of the differences in the lengths of the distance type, the estimated distances were normalized as percentages of the actual distance as used by Henry (1992). The following formula was used to compute the percentage of estimation from the actual distance:

$$\% \text{ of Estimated Distance from actual} = \frac{\text{Estimated Distance}}{\text{Actual Distance}} * 100$$

This percentage format enabled comparisons between the results of the different lengths of all distances in all distance types. Values of above 100 imply overestimation and values of below 100 are underestimation. A value of 100 means estimated distance matches actual distance. As such, this method allows us to express estimated distance as underestimation or overestimation relative to the actual distance.

For spatial memory task data the details of data preparation are presented in Chapter 7 as these data are only relevant for interactive image experiments.

#### 4.2.2.2 Data analysis

For statistical analysis, the data in this experiment were analyzed using a statistical package called SPSS (version 11.0) and Microsoft Excel program. The following tests were used for the quantitative data analysis: SPSS General Linear Model (GLM) univariate (ANOVA), SPSS GLM Multivariate (MANOVA), ANCOVA (Analysis of covariance), SPSS t-tests and Microsoft Excel Student t-test.

ANOVA is useful for testing significance between several IVs. ANOVA provides information on how the IV interact with each other and what effects these interactions have on the DV. ANOVA compares the variance (variability in scores) between groups and variability within groups. An F-ratio represents the variance between groups divided by variance within groups. A large F-ratio indicates there is more variability between groups than there is within groups. Thus, a significant F-test means that we can reject the null hypothesis and accept that there is a significant difference between groups. Since the experimental design of Experiment 1A was a two-way between-group design, a two-way ANOVA between groups was used. As described earlier, the two-way design means there are two IVs. The advantage of this design (thus analyzing using 2-way ANOVA) is that the main effect of each IV can be tested and additionally it allows the exploration for possible interaction effect between the IV. An interaction effect occurs when the effect of one IV on the DV depends on the level of the

second IV. For example, the influence of display size on distance estimate may be different for the real and VE image. For the real image, distance may be underestimated on small display but overestimated on the large display. In this example there is an interaction effect.

MANOVA is used when there is more than one related DV. It creates a new summary DV which combines linearly the original DVs and provides information on whether there is significant difference between this composite DV and the IVs. Besides, it also provides the univariate test results for each of the DV separately. For Experiment 1B, 2A and 2B, as the interest is in the effect of the IVs on all the overall five distances of each distance type, MANOVA is used in the statistical analysis. While an alternative is to use ANOVA on each distance; however, the more ANOVAs conducted, the greater the chance for making a Type 1 error. This error is when we believe that our experimental manipulation is successful when it isn't (Field 2000). Thus, the advantage of using MANOVA is that it 'controls' for the risk of inflating Type 1 error (Pallant 2001). Additionally, conducting separate ANOVA on each distance would yield a separate result for each distance. On the other hand, conducting a MANOVA on all the related DVs will yield the effect of the IVs on the linear combination of all the DVs. As mentioned earlier, we are interested in the effect of the IVs on the estimated distances for each asymmetrical distance, hence the use of MANOVA. However, the results of the univariate tests results produced by MANOVA analysis were also reported.

Analysis of covariance (ANCOVA) was done to explore the differences between the IV groups while controlling for the effect of other variable or covariates. The purpose of including the covariate(s) was to investigate its influence on the DV scores. In SPSS, the regression procedure is used to remove the variation in the DV that is due to the covariate(s). After removal of the variance, the normal analysis of variance techniques (ANOVA or MANOVA) was then performed on the adjusted data. Thus by conducting ANCOVA the chances of detecting differences between the IV might be increased by removing the influence of the covariate variables. For example in Experiment 1 and 2 of this thesis, the covariate variable was the sport variable, that is whether participants were active in sport or not. Thus, to investigate the influence of sport variable, the results of a secondary analysis of variance which includes the sport variable as covariates was also reported. For Experiment 3, the covariates included sport background variable, computer games experience, practice time and map-test time. Computer games experience refers to the frequency of participants playing computer games per week. Practice time refers to the time taken by participants to practice using the interface device/travel mode prior to the actual test trial. Map-test time refers to the time taken for participants to complete the spatial memory test.

T-test (In Microsoft Excel) returns the probability associated with a Student's t-Test. This test is used to determine whether two samples are likely to have come from the same two underlying populations that have the same mean. This test allows comparison of two sets of data array to be compared. A two-tailed distribution was used and assumption of unequal variance (heteroscedastic) was made.

Significance level (or alpha ( $\alpha$ ) level) was set at .05; that is the null hypothesis was rejected when the probability that a result would occur was less than .05. The importance of the impact of the IV on the DV was evaluated by *eta squared* or *partial eta squared* provided by SPSS (Pallant 2001). This value, refers to the 'effect size' represented the proportion of the variance in the dependent variable that can be explained by the IV (Pallant 2001). To interpret the strength of eta squared values the following guidelines can be used (Cohen 1988; cited from Pallant 2001 p175):

- .01 = small effect
- .06 = moderate effect
- .14= large effect

The **observed power** of a test will allow interpretation of the chances of the test detecting a difference between groups. Power is often not a problem when the sample size is large (e.g.  $n = 100$ ), however for a small sample ( $n < 20$ ), a non significant result may be due to insufficient power (Steven 1996; cited in Pallant 2001). Steven further suggested that when the sample size is small, it is necessary to adjust the significant level to compensate (e.g. a cut-off of .10 or .15). 80 percent would be an ideal value for chances of detecting a relationship; a value of less than this for insignificant result may suggest insufficient power of the test instead of no significant difference between groups. As such a non significant result must be interpreted carefully.

In Table 4-3, a summary of the main statistical test used in analyzing the data for all experiments in this thesis is presented.

Table 4-3 Statistical test used to analyze data in all experiment

Experiment	Test used	IV	DV
1A (Static image)	2-way ANOVA	Image type Display type	Asymmetrical distance - horizontal - transverse
1B (Static image)	1-way MANOVA	Display type	Asymmetrical distance - horizontal - transverse - vertical  3 distances estimates for each distance type
2A (Dynamic image)	2-way MANOVA	Image type Display type	Asymmetrical distance - horizontal



	Analysis of covariance (ANCOVA) is also conducted to test influence of sport variable (a covariate) on participants distance estimate		<ul style="list-style-type: none"> <li>- transverse</li> <li>- vertical</li> </ul> 6 distances estimates for each distance type (but analysis done on five only)
2B (Dynamic image)	2-way MANOVA  Analysis of covariance (ANCOVA) is also conducted to test influence of sport variable (a covariate) on participants distance estimate	Image type Display type	Asymmetrical distance <ul style="list-style-type: none"> <li>- horizontal</li> <li>- transverse</li> <li>- vertical</li> </ul> 5 distances to estimate for each distance type
3A and 3B (Interactive image)  (distance estimate data)	One-way ANOVA and Student t-tests  Analysis of covariance (ANCOVA) is also conducted to test influence of sport variable, computer-game variable, practice time, map-test time (as covariate) on participants distance estimate	Environment type* (real and VE) For VE : <ul style="list-style-type: none"> <li>- Display type (Large and small)</li> <li>- Interface type (mouse and trackball)</li> <li>- Travel mode (drive and fly)</li> </ul>	Asymmetrical distance** <ul style="list-style-type: none"> <li>- horizontal</li> <li>- transverse</li> <li>- vertical</li> </ul> **Room size estimation: width, length and height of room respectively for each distance
3A and 3B (Interactive image) (spatial memory task data)	<u>Mixed between-within subjects MANOVA</u>  Analysis of covariance (ANCOVA) is also conducted to test influence of sport variable, computer-game variable, practice time, map-test time (as covariate) on participants distance estimate	Environment type (real and VE) For VE : <ul style="list-style-type: none"> <li>- Display type (Large and small)</li> <li>- Interface type (mouse and trackball)</li> </ul> Travel mode (drive and fly)	Number of correctly placed objects
3A and 3B (Interactive image) (questionnaire data - quantitative part)	Repeated measure ANOVA on questionnaire data	VE only: <ul style="list-style-type: none"> <li>- Display type (Large and small)</li> <li>- Interface type (mouse and trackball)</li> </ul> Travel mode (drive and fly)	7-point Likert scale – based on the median score
3A and 3B (Interactive image) (questionnaire data – participants comments)	Summarized and reported	-	-

\* real and VE is only for Experiment 3A only. Experiment 3B investigate VE condition only

### 4.2.3 Summary of all experiments in the thesis

An overview of all the three experiments conducted in this thesis is given in Table 4-4. The table provides an outline of each experiment to be presented in the next three chapters: Chapter 5 Experiment on static images, Chapter 6 on dynamic images and Chapter 7 on interactive images. Detailed experimental methods, procedure, results and discussions are presented in the respective chapters, as indicated in the table.

Table 4-4 Summary of the three experiments

EXPERIMENT	STATIC IMAGES	DYNAMIC IMAGES	INTERACTIVE IMAGES
Presented in	Chapter 5	Chapter 6	Chapter 7
Aims/objectives	<p><b>Experiment 1A:</b></p> <p>To investigate the effect of image type and display on participants' asymmetrical distance perception</p> <p>Hypothesis:</p> <ol style="list-style-type: none"> <li>1. There is no significant different between image type (real and VE image) on asymmetrical distance estimate tasks.</li> <li>2. There is no significance different between display type (large projected display and desktop monitor) on asymmetrical distance estimate tasks.</li> </ol> <p><b>Experiment 1B:</b></p> <p>To investigate the effect display size</p> <p>Major hypothesis:</p> <ol style="list-style-type: none"> <li>1. There is no significant difference between large and small on asymmetrical distance estimation tasks.</li> </ol> <p>Secondary hypotheses :</p> <ol style="list-style-type: none"> <li>2. There is no effect of viewing distance on asymmetrical distance perception</li> <li>3. There is no effect of physiological cues on</li> </ol>	<p><b>Experiment 2A:</b></p> <p>To investigate the effect of image type and display on participants' asymmetrical distance perception</p> <p>Hypothesis:</p> <ol style="list-style-type: none"> <li>1. There is no effect of image type (real and VE image) on asymmetrical distance perception (vertical, horizontal, transverse).</li> <li>2. There is no effect of display size (small and large size) on asymmetrical distance perception (vertical, horizontal, transverse).</li> </ol> <p><b>Experiment 2B:</b></p> <p>To investigate the effect of display size</p> <p>To investigate the effect of image resolution on distance on asymmetrical distance perception</p> <p>Main hypotheses:</p> <ol style="list-style-type: none"> <li>1. There is no effect of image type (real and VE image) on symmetrical distance perception.</li> <li>2. There is no effect of display size (small and large size) on asymmetrical distance perception.</li> </ol>	<p><b>Experiment 3A:</b></p> <p>Main hypotheses:</p> <ol style="list-style-type: none"> <li>1. The type of environment (real vs. VE model) has no effect on participants' distance estimation task (vertical, horizontal and transverse) performance</li> <li>2. The type of environment (real vs. VE model) has no effect on participants' spatial memory task performance</li> <li>3. The display type (small vs. large) has no effect or participants' distance estimation task (vertical, horizontal and transverse) performance in interactive VE</li> <li>4. The display type (small vs. large) has no effect or participants' spatial memory task performance in interactive VE</li> <li>5. The type of input device (mouse vs. trackball) has no effect on participants' spatial memory task performance in interactive VE</li> <li>6. The different modes of travel (drive, fly) have no effect on participants' spatial memory performance in interactive VE</li> </ol> <p>Secondary hypotheses:</p> <ol style="list-style-type: none"> <li>1. There is no effect of viewing distance on distance estimate task in interactive VE</li> <li>2. There is no effect of physiological cues on distance estimate task in interactive VE</li> <li>3. There is no effect of viewing distance on spatial memory task in interactive VE</li> </ol>

	asymmetrical distance perception.	<p>Secondary hypotheses:</p> <p>3. There is no effect of viewing distance on asymmetrical distance perception</p> <p>4. There is no effect of physiological cues on asymmetrical distance perception.</p> <p>5. Image resolution has no effect on asymmetrical distance perception</p>	<p>4. There is no effect of physiological cues on spatial memory task in interactive VE</p> <p><b>Experiment 3B:</b></p> <p>In this study the main aim is to understand the unexpected finding of Experiment 3A. As it VE condition, only Item 3-6 above were explored as hypotheses.</p>
<b>Factors investigated</b>	<p><b>Experiment 1A:</b></p> <ol style="list-style-type: none"> <li>Image type</li> <li>Display type</li> </ol> <p><b>Experiment 1B:</b></p> <ol style="list-style-type: none"> <li>Accommodation and vergence cues (viewing distance)</li> </ol>	<p><b>Experiment 2A:</b></p> <ol style="list-style-type: none"> <li>Image type</li> <li>Display type</li> </ol> <p><b>Experiment 2B:</b></p> <ol style="list-style-type: none"> <li>Accommodation and vergence cues (viewing distance)</li> <li>Image Resolution</li> </ol>	<p><b>Experiment 3A:</b></p> <ol style="list-style-type: none"> <li>Environment type</li> <li>Display type</li> <li>Interface device type</li> <li>Travel mode</li> </ol> <p><b>Experiment 3B</b></p> <ol style="list-style-type: none"> <li>Display type</li> <li>Interface device type</li> <li>Travel mode</li> </ol>
<b>Independent variables</b>	<p><b>Experiment 1A:</b></p> <ol style="list-style-type: none"> <li>Image type               <ol style="list-style-type: none"> <li>real image (real scene picture)</li> <li>VE image (VE scene picture)</li> </ol> </li> <li>Display type               <ol style="list-style-type: none"> <li>desktop monitor (small display )</li> <li>projected display (large display)</li> </ol> </li> </ol> <p><b>Experiment 1B</b></p> <ol style="list-style-type: none"> <li>Display type</li> </ol>	<p><b>Experiment 2A:</b></p> <ol style="list-style-type: none"> <li>Image type               <ol style="list-style-type: none"> <li>real image (real scene picture)</li> <li>VE image (VE scene picture)</li> </ol> </li> <li>Display type               <ol style="list-style-type: none"> <li>Projected display                   <ul style="list-style-type: none"> <li>projected image size adjusted according to small and large display condition</li> </ul> </li> </ol> </li> </ol> <p><b>Experiment 2B</b></p> <ol style="list-style-type: none"> <li>Image type - real and VE image (pictures of real and VE scene)</li> </ol>	<p><b>Experiment 3A:</b></p> <ol style="list-style-type: none"> <li>Environment type               <ol style="list-style-type: none"> <li>real scene</li> <li>Interactive VE</li> </ol> </li> <li>Display type               <ol style="list-style-type: none"> <li>Projected display                   <ul style="list-style-type: none"> <li>projected image size adjusted according to small and large display condition</li> </ul> </li> </ol> </li> <li>Interface device type               <ol style="list-style-type: none"> <li>Mouse</li> <li>Trackball</li> </ol> </li> <li>Travel Mode               <ol style="list-style-type: none"> <li>Drive</li> </ol> </li> </ol>

		2. Display type Projected display - projected image size adjusted according to small and large display condition	ii. Fly  <b>Experiment 3B:</b>  Item 2 – 4 only
<b>Dependent variables</b>	<b>Experiment 1A:</b>  Asymmetrical distance - horizontal - transverse  <b>Experiment 1B:</b>  Asymmetrical distance - vertical - horizontal - transverse	<b>Experiment 2A:</b>  Asymmetrical distance - vertical - horizontal - transverse  <b>Experiment 2B:</b>  Asymmetrical distance - vertical - horizontal - transverse	<b>Experiment 3A:</b>  1. Spatial map test - number correctly placed objects  2. Asymmetrical distance - vertical - horizontal - transverse  3. Interface device questionnaire  <b>Experiment 3B:</b>  Item 1- 3 and Display questionnaire
<b>Variables fixed</b>	<b>Experiment 1A:</b>  1. FOV 2. Retinal image size  <b>Experiment 1B:</b>  1. Viewing distance 2. Physiological cues	<b>Experiment 2A:</b>  1. FOV 2. Retinal image size  <b>Experiment 2B:</b>  1. Viewing distance 2. Physiological cues	<b>Experiment 3A:</b>  1. Viewing distance 2. Physiological cues  <b>Experiment 3B:</b>  1. Display size 2. FOV 3. Retinal image
<b>Variables varied</b>	<b>Experiment 1A:</b>  1. Display size 2. Image type 3. Viewing distance 4. Physiological cues	<b>Experiment 2A:</b>  1. Display size 2. Image type 3. Viewing distance 4. Physiological cues	<b>Experiment 3A:</b>  1. Display size 2. Image type 3. FOV 4. Retinal image size

	<b>Experiment 1B:</b> <ol style="list-style-type: none"> <li>1. Display size</li> <li>2. FOV</li> <li>3. Retinal image size</li> </ol>	<b>Experiment 2B:</b> <ol style="list-style-type: none"> <li>1. Display size</li> <li>2. Image type</li> <li>3. FOV</li> <li>4. Retinal image size</li> </ol>	<b>Experiment 3B:</b> <ol style="list-style-type: none"> <li>1. Viewing distance</li> <li>2. Physiological cues</li> </ol>
<b>Experimental conditions</b>	<b>Experiment 1A:</b> <ol style="list-style-type: none"> <li>1. Real/ Small</li> <li>2. Real/Large</li> <li>3. VE/Small</li> <li>4. VE/Large</li> </ol> <b>Experiment 1B:</b> <ol style="list-style-type: none"> <li>1. VE/Small</li> <li>2. VE/Large</li> </ol>	<b>Experiment 2A:</b> <ol style="list-style-type: none"> <li>1. Real/ Small</li> <li>2. Real/Large</li> <li>3. VE/Small</li> <li>4. VE/Large</li> </ol> <b>Experiment 2B:</b> <ol style="list-style-type: none"> <li>1. Real/ Small</li> <li>2. Real/Large</li> <li>3. VE/Small</li> <li>4. VE/Large</li> </ol>	<b>Experiment 3A:</b> <ol style="list-style-type: none"> <li>1. Real condition</li> <li>2. VE/Small/Mouse/Drive</li> <li>3. VE/Small/Mouse/Fly</li> <li>4. VE/Small/Trackball/Drive</li> <li>5. VE/Small/Trackball/Fly</li> <li>6. VE/Small/Mouse/Drive</li> <li>7. VE/Small/Mouse/Fly</li> <li>8. VE/Small/Trackball/Drive</li> <li>9. VE/Small/Trackball/Fly</li> </ol> <b>For Experiment 3B:</b> Item 2-9 only  Note: For Experiment 3A – the environment condition is real and VE For Experiment 3B – the environment condition is VE only
<b>Participant's tasks</b>	View image and estimate distance in image	View movie for 3 minutes and then estimate distance in image	Explore environment , then do <ul style="list-style-type: none"> <li>- spatial map test</li> <li>- asymmetrical distance estimate</li> <li>- interface device questionnaire</li> </ul>

### 4.3 CHAPTER SUMMARY

In this chapter the basis for the research conducted in this thesis was presented. First, the need and importance for VE to be perceived similarly to its real world counterpart were introduced, highlighting the importance of spatial awareness. Subsequently the inadequacy of current VE technologies to provide such accurate simulations was highlighted with regards to prior studies results. The paucity of knowledge on factors to provide similar perception in both environments was highlighted in relation to a review of related studies in Chapter 2.

The general research approach of comparing spatial awareness performance between real and VE was discussed next with focus on the distance estimate task and spatial memory task as task performance measures. This was followed by a discussion on factors and scope of the research presented in this thesis. This included display size, image presentation format and display type used, navigation and interface device. A discussion on image modelling was presented next, highlighting the need for image realism in VE models which includes discussion on the trade-off between generating image realism with system interactive performance.

The research methods on data collection and data analysis employed in this research were presented in the second section of this chapter. An overview of these methods was presented, leaving the details of methods related to specific experiment to be discussed in each respective chapter. In the last sub-section, an overview of all the experiments that outlines each experiment aims, factors, conditions and tasks was provided. This overview serves as a reading guide for the upcoming three chapters (Chapter 5, 6 and 7), which report on these experiments in greater details.



## PART II

### EXPERIMENTAL APPROACH, RESULTS AND ANALYSIS

- Chapter 5 – Experiment 1: Distance Perception in Static Images**
- Chapter 6 – Experiment 2: Distance Perception in Dynamic Images**
- Chapter 7 – Experiment 3: Distance Perception and Spatial Memory  
Tasks in Interactive VE**

# CHAPTER 5

---

## *EXPERIMENT 1: DISTANCE PERCEPTION IN STATIC IMAGES*

### 5 OVERVIEW

In this chapter, the experimental methodology and the results of the first set of studies (Experiment 1) that compares participants' spatial awareness in static images of the real and VE were outlined. The general aim of these studies was to compare participants' distance estimate performance between the real and VE images presented to them in a non-stereo mode. The effect of presenting the images on different display types was also investigated. Two studies were conducted. The basis for undertaking these studies was discussed in Chapter 4. The first study (Experiment 1A) compared participants' distance estimation in the pictures of the real and VE images displayed on a desktop and a projected display (Awang-Rambli and Kalawsky 2002) while the second study (Experiment 1B) was conducted to investigate the effect of viewing distance and physiological cues (factors that was not controlled in Experiment 1A). A discussion of the results from both studies and conclusions drawn were presented at the end of the chapter.

## 5.1 EXPERIMENT 1A: EFFECT OF IMAGE TYPE AND DISPLAY TYPE

### 5.1.1 Rationale

In order to be effectively applied to applications, particularly those that use VE to represent the real world counterpart, the VE technologies must allow users to perceive the real and VE similarly. The literature reviewed in Chapter 2 has been inconclusive on the users' performance in the real and VE. While some researchers reported an overestimation (Waller 1999, Yoon, Byun et al. 2000), generally distance perception in the VE has been found to be underestimated (Witmer and Kline 1998, Henry and Furness 1993). Numerous past studies have examined performance difference between a real environment and its 3D VE environment (Witmer and Kline 1998, Henry and Furness 1993, Lampton, Bliss et al. 1994, Lampton, McDonald et al. 1995, Witmer and Sadowski 1998, Yoon, Byun et al. 2000, Willemson and Gooch 2002, Messing 2004, Plumert, Kearney et al 2004). Very few studies found have compared pictures of the real and VE based on distance estimate tasks (Yang, Dixon et al. 1999). As such, in this study, we compare distance estimation task performance between picture of a real environment and picture of the VE model. Moreover, it is of interest to find out how static presentations influence one's spatial awareness as many applications use a static presentation at some point during viewing of the presentation (Kjeldahl and Prime 1995).

As discussed in Chapter 4, for some applications (such as scientific visualisation and crime scene investigations), there are instances when the visual scenes might be static to enable the viewer to extract spatial information (Yeh and Silverstein 1992). Other application such as crime scene investigation uses pictures (or static images) in the work process<sup>3</sup>. As part of documentation of the crime scene, pictures were taken from various viewpoints to be used for subsequent analysis and as evidence in the court of law. Viewpoints of the pictures taken depend on what the photographers thought were important and may not match what the investigators need. Most often crime scenes do not last very long. Usually, when the crime scene investigators have completed their job of examination and documentation of evidence, the scene must be released as soon as possible to return to its normal function. This is especially true if the scene is an area of commerce, investigators are often pressured to get the scene working and functional again (O'Connor 2004). One alternative is to "preserve" the crime scene through the creation of the computer generated VE model of it. This not only

---

<sup>3</sup> George Shiro, Forensic scientist at Louisiana State Police Crime Laboratory, USA. Examination and Documentation of the Crime scene. Available at [www.crime-scene-investigator.net](http://www.crime-scene-investigator.net)

allows investigators to revisit the crime scene at any time, it also allows unlimited number of snapshots to be taken at any viewpoints or angles they want. While the advantage of computer generated images is apparent in such application, limited studies are available to examine if the VE pictures convey similar information to the real pictures. Studies comparing 3D VE to real physical environment have revealed 3D VE are perceived differently from the real physical environment. As the results of such studies might not extend to pictures, this motivates current study to examine if pictures of the real and VE are perceived similarly or not.

### 5.1.2 Experimental aim and hypotheses

The overall aim of this initial study was to investigate participants' spatial awareness in terms of asymmetrical distances in static real and VE picture presented on a desktop monitor and projected display. The following hypotheses were investigated in this study:

H1: There is no significant different between image type (real and VE image) on asymmetrical distance estimate tasks.

H2: There is no significant different between display type (large projected display and desktop monitor) on asymmetrical distance estimate tasks.

The desktop monitor corresponds to a small display and the projected display corresponds to a large display and as such in this chapter these terms are used interchangeably. As described in Chapter 4 (Section 4.1.1.1), asymmetrical distance refers to vertical, horizontal and transverse. However, for this initial study only two of these distances were investigated: horizontal and transverse distance. Vertical distances will be investigated in Experiment 1B of this chapter, Experiment 2 in Chapter 6 and Experiment 3 in Chapter 7.

The use of static image as stimulus eliminates the effect of motion cues and thus allows us to examine the effect of the factors (image type and display type) under investigation in a more controlled situation. Similarly, excluding the stereo cues from the experimental design allows for the investigation of the impact of non-stereo cues (refer also to discussion in Section 4.1.2.5 of Chapter 4).

### 5.1.3 Methodology

#### 5.1.3.1 Participants

Forty volunteers participated in the study. Thirty-four of the participants were male. Participants' age ranged from 17 to 51 years with an average of 30. All participants either had normal or corrected-to-normal vision.

#### 5.1.3.2 Materials/Apparatus

##### Real picture

For this study, an image with very few objects in it was required, as more objects might have provided more clues to participants in their estimation and have created more variance among participants and thus a less controlled experiment. This is because different participants might have used different objects to base their estimations on. Thus, the presence of many different objects would have created more participants' variance where different participants might not use the same objects to base their distance judgments. A location on campus, which met this requirement, was identified. A photograph of this location (Figure 5-1) was taken using an Olympus Model C-920ZOOM digital camera in a standard, auto focus mode. The picture image vertical and horizontal resolution was 692 x 685. The image was placed on a Microsoft PowerPoint slide (in full-screen mode) for the real picture condition.



**Figure 5-1 Real picture**

##### Virtual picture

A VE model of the real scene was created using REALAX RXScene software on a Windows NT machine. Textures (trees, grass, road, sky, lamppost, hedges) from the real picture were used as textures for objects in the VE model. The viewpoint in the VE model was set to 1.5m above the ground, at the same point where the picture was taken in the real world. A snapshot

of the VE model was taken (using print screen command) and placed on a Microsoft PowerPoint slide to represent the VE picture condition (Figure 5-2). The picture image vertical and horizontal resolution was 791 x 769.



Figure 5-2 Virtual picture

#### 5.1.3.3 Display types

The images were displayed using a Windows NT machine with a 17" monitor display for the desktop condition. An LCD projector was connected to a Windows NT machine and was used to project the pictures (real and VE) onto a large white paper (135 x 95 cm) on the wall for the projected display condition. The resolution of display on the desktop was set at 1028 x 768, while the resolution for the projected display was set at 800 x 600 (the highest available on this LCD projector). These differences in screen resolution for both display types will be taken into consideration in the discussion of results.

#### 5.1.3.4 Experiment room

The projected display condition was set up in one of the room in the computer building of Loughborough University, while the desktop condition was setup in one of the computer labs in the same building. In order to examine if the differences in room locations may influence experiment result, an informal study was later conducted by setting up the desktop conditions in the same room used by the projected display conditions. Study results suggested that the results of the informal study were similar to those conducted in the computer labs, thus eliminating the room location as a potential variance in this experimental result.

The pictures were viewed under normal lighting conditions in the room. In hindsight a dark room setting would have been desirable because it has been suggested a dark room setting (which was employed in the later experiments) helps to reduce peripheral view effects from objects surrounding the screen. These peripheral view effects have been shown to affect



participants' distance estimates (Eby and Braunstein 1995, cited in Knapp 1999). However, this setting was considered acceptable in this initial exploratory study and it is believed that the impact of the effect was minimal because of the presence of few objects in the surrounding area of the monitor and projected screen. Additionally, the monitor screen was placed at a corner of the room and the projected display image was projected on a large paper which was pasted directly on a blank wall which further reduced the impact of peripheral view effects.

### 5.1.3.5 Experiment design and setup

Since Experiment 1A is the first experiment of a series of experiments undertaken by the research presented in this thesis, it is considered as an initial exploratory investigation. The experiment involved a 2 x 2 factorial between-subject design. A between-subject design was chosen to avoid carry-over effects due to the use of the same scene for all conditions. The two IVs were display type and the image type. The two levels of the display type were the desktop monitor and the projected display. The two levels of image type were the real picture and the VE picture. The DV was the estimated distance between objects. The participants were randomly assigned into the following four experimental conditions groups:

- real picture/desktop,
- real picture/projected display,
- VE picture/desktop
- VE picture/projected display.

To investigate the effect of physical display size on the distance estimation task, the same image was used on both display types. Additionally, to eliminate the effect of retinal image size on distance estimation a similar FOV was maintained by adjusting the viewing distance of the participants from the screen. Due to the room size constraint, the resultant retina image sizes on both display size differed slightly by a few degrees (Table 5-1). This very small difference was considered acceptable as this was an exploratory study. Figure 5-3 illustrates Experiment 1A setup.

**Table 5-1 Summary of experiment setup variables**

Display type	Distance from screen	FOV (in degree)	Display Resolution
Desktop	40 cm	20.54	1028x768
Large projected display	280 cm	22.03	800x600

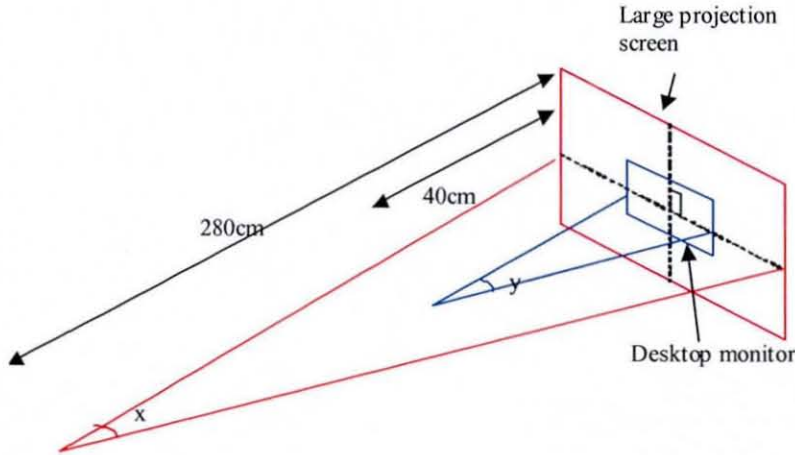


Figure 5-3 Experiment 1A setup

### 5.1.3.6 Procedure

Participants were given instructions verbally and as well as written on the computer monitor display or projected display. They were asked to estimate two types of distances: transverse distance and horizontal distance. Estimations were to be made in meters and to reduce differences of a meter length concept among participants; a meter long tape was shown to them prior to the start of the experiment.

To avoid participants changing their mind very often, participants were given 15 seconds to view the image and to report their estimation. This time was based on a similar study where participants were asked to rate the quality of display, when given more time (20 seconds), the participants change their mind very often (Storms and Zyda 2000). It was suggested the duration of human working memory (WM) is approximately twenty seconds with the rate of decay in WM changes depends on the amount of information (Peterson and Peterson 1959). The image used here has a reasonable amount of information in it, thus 15 seconds is within the WM constraints.

At the end of the trial, each participant was asked to complete a short post-test questionnaire to find out what strategies they used to estimate distances. The data sheet and post-test questionnaire forms can be found in Appendix A.

## 5.1.4 Results

### 5.1.4.1 Data preparation

Data was checked to ensure no violation of assumption of parametric test validity. Based on the z-scores outliers checking method described in Section 4.3.2 of Chapter 4, two cases of the data were identified as outliers and were removed prior to further analysis of the data.

Participants' performance accuracy was measured in terms of how close their estimated distance to the actual distance was. Due to the differences in the lengths of the distance type, the estimated distances were normalized as percentages of the actual distance as used by Henry (1992) (see Section 4.3.2, Chapter 4). Thus the estimated distance was converted to percentage value referred to as percentage of estimate from actual distance. This method allowed us to compare among distances of varying length and allows us to express estimated distance as underestimation (less 100) or overestimation (more than 100) relative to the actual distance.

A two-way between group analysis (ANOVA) was done (using SPSS version 11.0) separately for each distance type. For each analysis, the between group variables were image type (2 levels: real and VE image) and display type (2 levels: desktop and projected display). Significance level was set at .05. This means the null hypothesis will be rejected when the probability that a result is occurring is less than this value.

In the next sub-sections, the results of the asymmetrical distances were first presented (horizontal and transverse distance) followed by the post-test questionnaire results.

### 5.1.4.2 Horizontal distance

Figure 5-4 shows the mean estimated distance for the horizontal distance for the four experimental conditions. The actual distance for this distance was 8.73m. Examination of the Figure 5-4 suggests that there was a difference in horizontal distance estimation between the desktop and the projected display. The results of a two-way ANOVA showed that the main effect for display type [ $F(1,34)=4.059$ ,  $p=.052$ , partial eta squared=.107] approaches significance. While not significant, a comparison of the means (see Figure 5-4) showed that the large display participant' estimations were more accurate compared to the small display participants. From Figure 5-4, it appears that there was a difference between the real and VE images on the desktop and the projected display. However, the main effect for image type [ $F(1,34)=1.202$ ,  $p=.281$ ] did not reach significance indicating the difference between the



image types was small. The interaction effect [ $F(1,34)=.044, p=.836$ ] was not significant either, indicating no significant interaction between image types and display types.

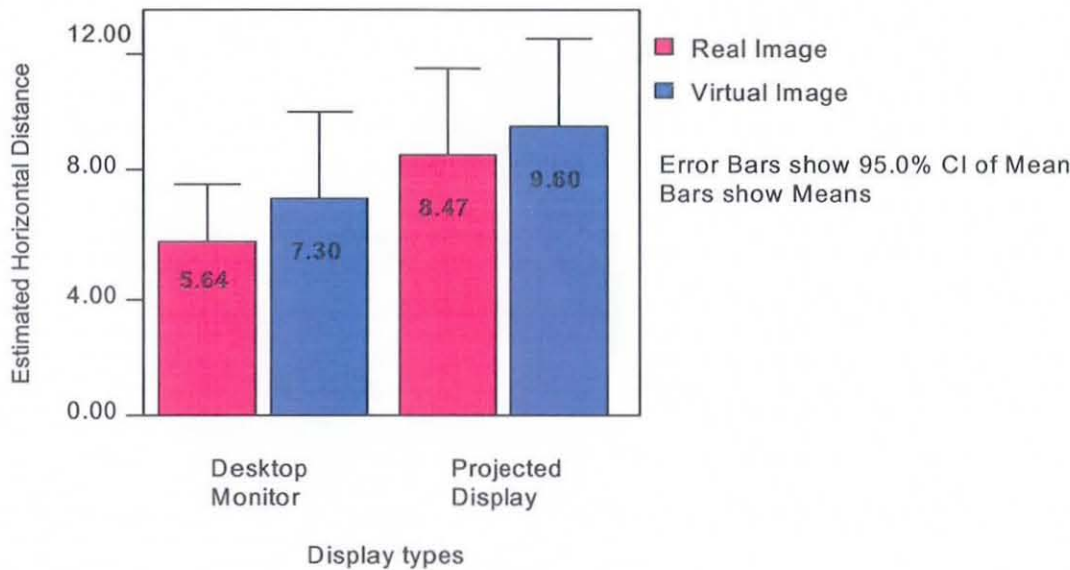
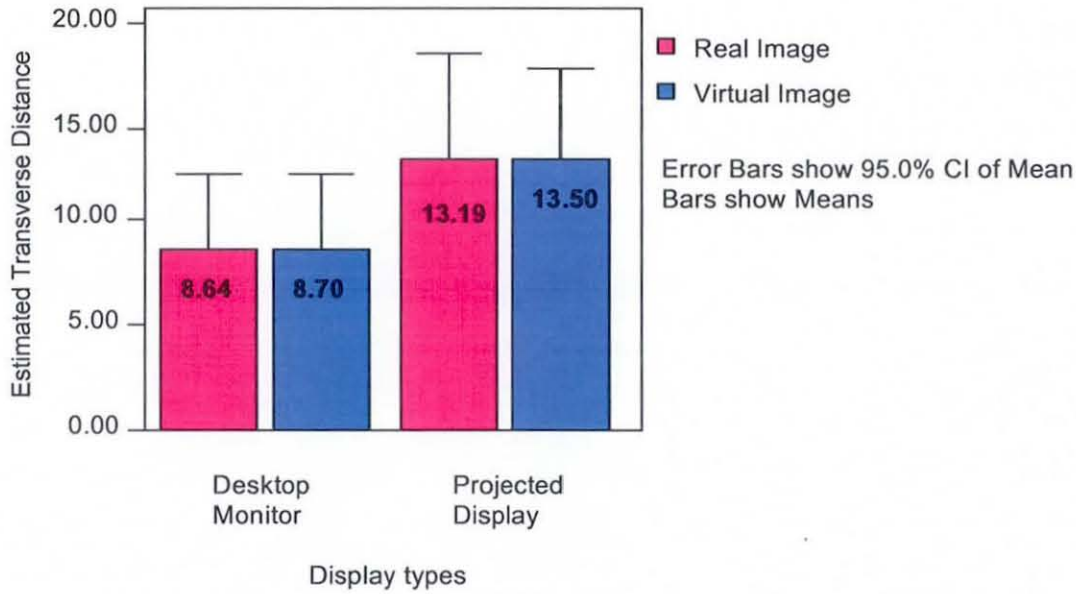


Figure 5-4 Mean estimated distance for horizontal distance

5.1.4.3 Transverse distance

The mean estimated distances for transverse distance in the four experimental conditions is shown in Figure 5-5. It is noted that the actual distance was 22.4m. Thus, from the figure the transverse distances were greatly underestimated in all conditions. Estimates were less than half on desktop monitor and slightly more than half on projected display. Examination of Figure 5-5 suggests that there was a large difference between distance estimated on a desktop and distance estimated on a projected display for both image types. The results of a two-way between-group ANOVA suggests that this difference was statistically significantly as indicated by the main effect of display type [ $F(1,34)=5.212, p=.029$ , partial Eta squared = .133]. From Figure 5-5, it was shown that the projected display participants' estimations were more accurate compared to the desktop monitor participants. The main effect for image type [ $F(1,34)=.008, p=.928$ ] and the interaction effect [ $F(1,34)=.004, p=.952$ ], however, did not reach statistical significance. As illustrated in Figure 5-5, both image types did not differ very much on either the desktop or the projected display.



**Figure 5-5 Mean estimated distance for transverse distance**

#### 5.1.4.4 Post-test questionnaire

In the post-test questionnaires, participants were asked to report how they made their distance estimation and to indicate which distance (transverse or horizontal) was easier to estimate. Generally, most participants (65 %) reported that their estimations were based on the features and locations of objects in the pictures: such as the trees, roads and lampposts. One participant based his estimation on objects that was not present in the scene. Some participants estimated distance by trying to imagine the real scene in the desktop/VE condition. Some participants based their estimates on everyday experience. Four participants expressed familiarity with the location in the scene. However, after examining their results, their estimations were not very accurate and were comparable with other participants. In all conditions, only three participants said that they guessed their estimations. Twenty-seven participants commented that horizontal distance is easier to estimate. The main reason was no perspective was involved in horizontal distance estimation. Only ten participants commented that transverse distance was easier to estimate and three participants commented that there was no difference between transverse and horizontal distance.

#### 5.1.5 Analysis

Generally, distances were underestimated in all conditions (with the exception of VE image on projected display) for both the transverse and horizontal distances. The present study results revealed that the difference between the real and VE picture on distance estimation task was small. It was shown that these differences were statistically insignificant on either

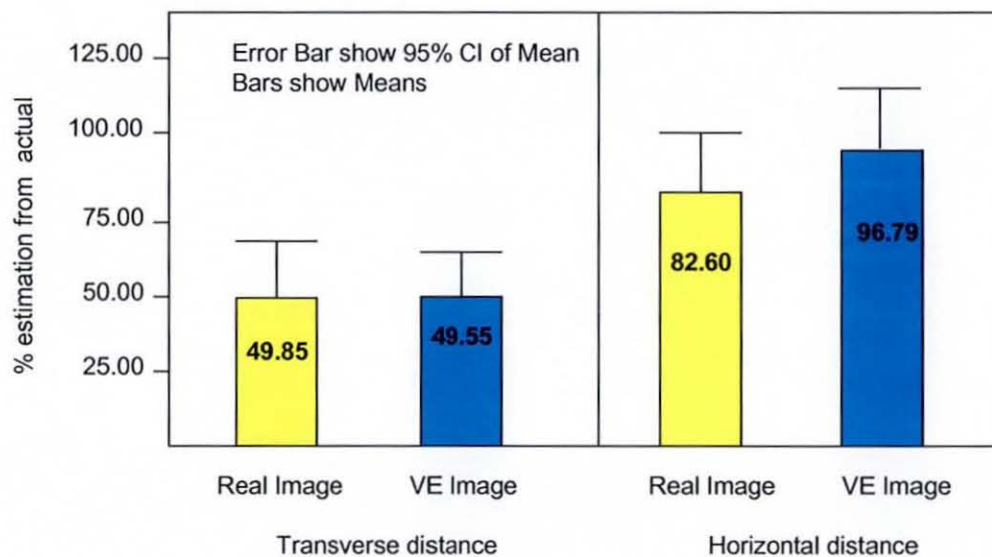
display type (desktop and projected display) for both horizontal and transverse distances (see Table 5-2).

**Table 5-2 Comparisons among display types for horizontal and transverse distances**

Distance type	Display type	Estimated Distance	
		Real Image	VE image
Horizontal distance	Desktop Monitor	5.64	7.3
	Projected Display	8.47	9.6
Transverse distance	Desktop Monitor	8.84	8.7
	Projected Display	13.19	13.50

For horizontal distance, a direct comparison of means revealed better estimates for the horizontal distance in VE picture/desktop condition but on the large projected display, estimates on the real image were more accurate compared to the VE image.

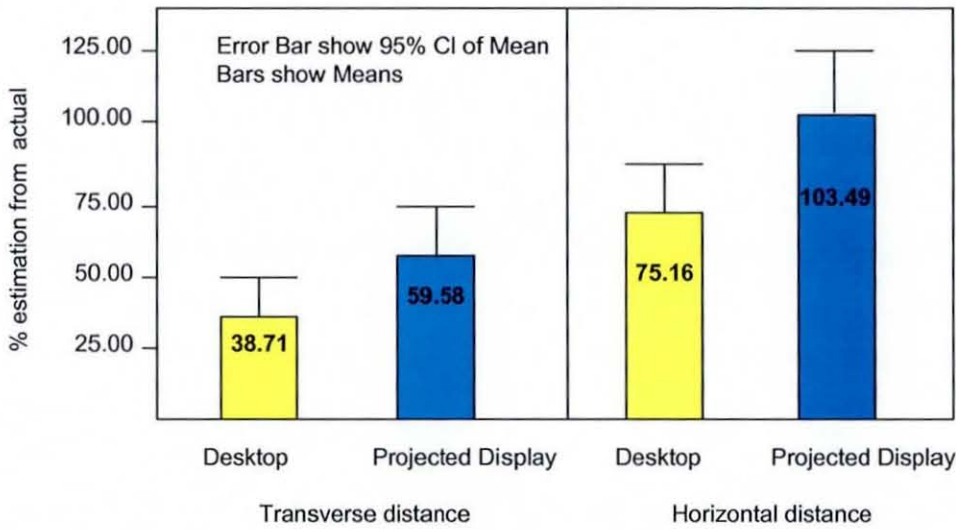
Based on percentage of estimate of distance from actual, Figure 5-6, illustrates that both the real and VE images did not differ very much for the transverse distance but for the horizontal distance estimates on the VE image was more accurate compared to the real image.



**Figure 5-6 Percentage of estimate from actual for transverse and horizontal distance – comparing image types**

From Figure 5-7, overall, both the real and VE images produced more accurate distance estimations on a projected display compared to the desktop monitor. Comparatively, it was shown that horizontal distance estimates were more accurate than transverse distance (82.6% on real image and 96.7% on VE image). Transverse distance was estimated at approximately half of the actual distance (Real image: 49.85%, VE: 49.55%).





**Figure 5-7 Percentage of estimate from actual for transverse and horizontal distance – comparing display types**

An important observation made in this experiment was the effect of display type on distance estimation: a projected display yielded more accurate estimation for transverse and horizontal distance perception (see Figure 5-7). For transverse distance, percentage of estimate from actual was approximately 59% on a projected display and 38% on desktop. Similarly, for horizontal distance, percentage of estimate from actual was approximately 103% on projected display and 75% on desktop. With the exception of horizontal distance on a projected display which showed an overestimation, generally distances tended to be underestimated. A direct comparison of the mean percentage of estimate values shows that transverse distance was greatly underestimated compared to the horizontal distance.

### 5.1.6 Discussion

Consistent with the findings of (Witmer and Kline 1998, Henry and Furness 1993, Lampton, Bliss et al. 1994, Lampton, McDonald et al. 1995, Witmer and Sadowski 1998), distances were generally underestimated in the real and VE for both transverse and horizontal distance. As the stimulus used was static pictures, this inaccuracy was expected. Lumsden (1980) indicated that inter-object distance distortion occurs when viewing a photograph of a three-dimensional scene. When viewing photographs of two or more objects which were viewed at increasing distances from the observers, Lumsden further suggested that an apparent decrease in distance between the objects would occur. Our results show similar occurrence on distance estimation for both the real and VE images.

### 5.1.6.1 Effect of image type

Despite the slight difference between image resolution, Experiment 1A results revealed similar performance between the real and VE condition. This finding is in contrast to Witmer and Kline (1998) who reported significant difference between the real world condition and the VE model. Witmer & Kline (1998) found that egocentric distances were underestimated more in the VE than in the real world. They reported estimates of 47% of true distance for VE and 72% of actual for real world condition. For transverse distance, the present study results reported an estimate of approximately 49% of actual distance for both the real and VE pictures. Witmer and Kline (1998) attributed the superior performance of the real world participants compared to VE to the difference in the depth cues available in the real and the VE model. Comparatively, fewer cues were present in their simple VE, compared to more cues present in the real physical environment. Their VE model of the hallway was not closely modelled upon the real hallway; most of the features in the real hallway were not present in the VE.

However, the results of non-significant difference between the real and VE conditions for distance estimate task for Experiment 1A is consistent with the results of Willemsens and Gooch (2002). In a study which compares distance perception in photographic-based VE of a hallway to a computer-generated VE version, Willemsens and Gooch (2002) found that in both images distances were underestimated and were significantly different from the distance perception in the real physical hallway. Similar to our study, their result showed that the difference between the photographic-based VE and computer-generated image was small even though comparatively the photographic image VE was more rich in visual information (such as shadows, inter-reflection and global illumination) which was not present in their computer-generated VE. For Experiment 1A, the VE was carefully modelled based on the actual physical real environment. The photographic VE model used by Willemsen and Gooch (2002) was created from stereo photographs pictures of the actual scene. However, in Experiment 1A, we created textures from the photographic pictures of the actual scene to give objects in our VE a similar appearance to the real picture. Although shadows were modelled, illumination was not accurate in our VE. Thus, similar to Willemsen and Gooch's study result, the non-significant difference between the real and the VE picture on distance estimation task suggested that our VE model must have provided the observer with visual cues necessary for distance perception similar to those available in the real pictures.

It was noted that other studies (Witmer and Kline 1998, Lampton, MacDonald et al. 1995, Lampton, Bliss et al. 1994, Henry and Furness 1993, Witmer and Sadowski 1998) that

reported a difference between the real and VE images have used VE models which were slightly or very different in terms of detailed information (such as texture, shadow, illumination) from the compared real world environment. Most of the VE models used contained less detailed information in terms of texture and other features that might be present in the real world counterpart. The presence of more cues or redundant pictorial cues would yield a more realistic and compelling sense of 3-D space similar to the physical real world condition. Kunnapas (1968) has demonstrated that increasing the number of cues increases the accuracy of distance judgment. As such, one of the possible contributing factors to the significant differences between these previous studies was due to the presence of fewer pictorial cues or depth information in the simple VE as compared to the real physical world condition.

The results of non-significant difference between distance perceived in the real and VE images in Experiment 1A also provide support for other previous research which showed that people can perceive horizontal and transverse distance in the VE similar to the real world (Waller 1999, Yoon, Byun et al. 2000, Yang, Dixon et al. 1999).

It was noted that the transverse distance estimated in our study was different from the egocentric distance estimate (see explanation in the paragraph related to Figure 5-8). In a study comparing estimated sizes and ratios between the real and the VE room, Yoon and colleagues (2000) concluded that estimates between the real and the VE were small but both were significantly different from the actual sizes (width and length) except for the height estimates.

In another similar but comparable experiment conducted by Yang, Dixon et al. (1999), it was reported that with regards to relative perception of horizontal and vertical extents, a snapshot of a VE scene on a desktop was similar to a picture. Yang, Dixon et al. (1999) further concluded that with angular sub-tenses of object equated, the lack of reliable differences between the real outdoor environments and the VE models and between pictures and the VE pictures suggests that whether observers viewed the real scene or the computer-generated VE images did not make a difference. Even though the vertical distance was not investigated in Experiment 1A, however as reported in the post-test questionnaire, most participants used object heights as visual cues for their distance estimate tasks. Thus, as demonstrated by Yang and colleagues, this might have resulted in the similar perception of the real and the VE picture in Experiment 1A.

In his studies of exocentric distances (distances between objects), Waller (1999) concluded that people can perceive distances in the VE world nearly as well as they can in the real world provided given proper feedback, wide FOV and ability to move around in the VE. Our present study did not allow for navigation in the VE as the stimulus used was a static picture and no feedback was given. However, our results still indicate similar performance between the real and VE image participants.

It was asserted earlier that the more accurate result of distance perception found in the real condition in Witmer and Kline (1998) and other studies (Lampton, MacDonald et al. 1995, Lampton, Bliss et al. 1994) was due to the use of the real physical environments for comparisons to the VE model. In contrast, our study compared pictures of the real world to the pictures of the VE. As reviewed in the literature, the perception of depth in pictures is less accurate when compared to perception of depth in the real physical world. The conflicting nature of picture perception of picture has been attributed to this less accurate perception. Additionally, Hagen, Jones et al. (1978) demonstrated that the truncation of the visual field in pictures which causes a shift in the localization of the visual field makes objects appear closer to the viewer than it actually is. This truncation causes an underestimation of size and distance in pictures. This suggests that the inaccurate perception in pictures may explain the less accurate result of distance perception in our real picture conditions compared to previous studies.

### Egocentric versus Exocentric distance

Comparatively, similar to the real world conditions result of Witmer and Kline (1998), several researchers (Hecht, Doorn et al. 1999) reported that more accurate estimates in picture for far distance than Experiment 1A results. In their study, Hecht and colleagues asked participants to report distance from self to corners of buildings using picture and real condition as stimuli. It is noted that these authors asked participants to estimate egocentric distance (distance from self to object) whereas in Experiment 1A participants were asked to estimate distance between objects. There is a difference between this two estimation tasks as illustrated in Figure 5-8.

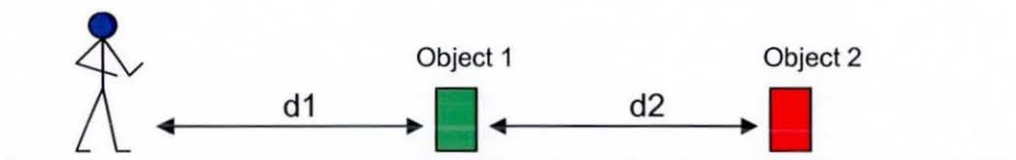


Figure 5-8 Egocentric distance versus exocentric transverse distance

Egocentric distance is represented by  $d_1$  while exocentric distance (for transverse distance) is represented by  $d_2$ . Egocentric distance is considered a straightforward estimation from self to object 1. This is not as simple for  $d_2$  estimation. One method would be to estimate how far Object 1 is and then estimate how far it is to Object 2. The additional error in estimates introduced by estimating  $d_1$  thus might have resulted in our transverse distance estimate being less accurate when compared to the real condition results of Witmer and Kline (1998) and the real picture condition of Hecht, Doorn et al. (1999).

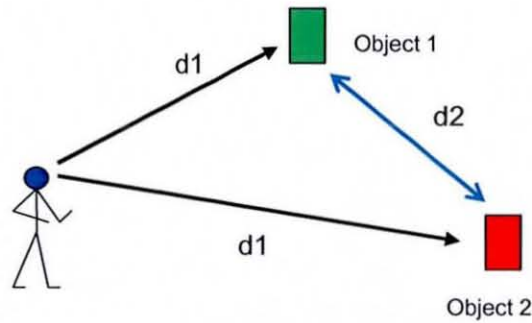
It was observed earlier that our VE condition result was similar to Witmer & Kline's (1998) VE condition. Since the distance estimate task used in our study was exocentric distance and theirs was egocentric distance, this suggests that it is possible to perceive egocentric distances and exocentric (transverse) distances in both the VE model and pictures of the VE model similarly.

### Comparison between horizontal and transverse distance

Experiment 1A results revealed that participants' estimates were more accurate when judging the horizontal distance when compared to the transverse distance (82.6% for real image and 96.79% for VE image). Correspondingly, Loomis, Da Silva et al. (1996) found that more estimation errors were made on the transverse distance than on the horizontal plane. They also found that the degree of perceptual distortion increases with distance from the observer. Comments reported by the participants in the post-test questionnaire revealed that more than two-thirds of the participants found that horizontal distance was easier to estimate. Only one quarter of the participants found that transverse distance was easier to estimate when compared to horizontal distance.

From Figure 5-9, for horizontal distance estimate, both objects (Objects 1 and 2) were of similar distance (if not equal) distance from the observer and both objects were clearly visible which made it easier to estimate. In contrast, for transverse distance, however, one of the objects receded into the distance (see Figure 5-8). When objects are viewed at increasing distances from the observer, Lumsden (1980) suggested that an apparent decrease in distance between the objects would occur. As mentioned earlier Loomis, Da Silva et al. (1996) also found that the further the object is from the observer the greater is the distortion in estimation. As such, this will result in more underestimation in transverse distance when compared to horizontal distance. Thus, this may offer one possible explanation why horizontal distance yields more accurate estimates when compared to transverse distance.





**Figure 5-9 Exocentric distance (Horizontal distance)**

In the real world, however, distance estimates on average ranges between 87-91% of actual distance (Wright 1995). Our horizontal distance estimates thus are more comparable to real world estimates.

As mentioned earlier the difference between the real and VE image was not significant for horizontal distance. However, it was observed that on average the VE participants' estimations were more accurate compared to the real image participants. This was unexpected as most previous investigations (Witmer and Kline 1998, Lampton, MacDonald et al. 1995, Lampton, Bliss et al. 1994, Henry and Furness 1993) have reported more accurate distance estimate in the real environment compared to the VE. This could be due, in part, to the large variability observed among participants which may have resulted from our between-group design method. However, previous studies based their conclusions on egocentric distance estimation (that is distance between self and object), while Experiment 1A was based on horizontal exocentric distance (distance between object). Horizontal distance is the distance across the screen as opposed to distance into the screen for egocentric distance. It should be noted again however that the difference is statistically insignificant; this difference could be due to random error.

#### 5.1.6.2 Effect of display type

It was demonstrated in Experiment 1A that there was a main effect of display type on distance estimation for both distance types (for horizontal distance significant at 10%). These results indicated that there was a statistically significant difference in participants' performance between the desktop monitor and the projected display. For both the transverse and horizontal distances, the results showed that a projected display yielded more accurate results when compared to desktop monitor. These results are consistent with previous research showing larger display resulted in better participants' performance (Patrick, Cosgrove et al. 2000, Tan, Gergle et al. 2003, Yang, Dixon et al. 1999, Dixon and Proffitt 2002, Czerwinski, Tan et al. 2002, Tan, Czerwinski et al. 2003) .



Yang, Dixon et al. (1999) found that vertical overestimation is reduced in desktop monitors compared to the real and VR conditions. They believed that our visual system is influenced by the perceived physical size of the projection. Thus, the reduced overestimation on a desktop display was due to the small size projection on the picture surface. They proposed that vertical overestimation would increase if a picture were distended such as projecting it onto a larger screen. This proposition was confirmed by a later study conducted by Dixon and Proffitt (2002) showing that the vertical overestimation was influenced more by the perceived distal object size rather than the dimensionality of the display (2D versus 3D). Although vertical estimation was not investigated in this study; most participants reported using objects' height in the scene to base their estimation. This may have accounted for the larger estimate values made when images were viewed on larger display.

In a comparable study, investigating spatial knowledge gained by navigating in VE viewed in three conditions (HMD, desktop monitor and large projection screen), it was found that performance on the large projection screen was more accurate than on the desktop monitor (Patrick, Cosgrove et al. 2000). Patrick and the others suggested that this better performance might be due to the physical image sizes that are large enough to induce more presence and a realistic appearance on the participants, thus resulting in better judgment of relative position was perceived. The more accurate result on projected display compared to the desktop monitor in our study may also be due to our participants having similar experience. In two separate studies (Czerwinski, Tan et al. 2002, Tan, Czerwinski et al. 2003) reported that large displays resulted in improved performance in 3D navigation, especially for females.

In a more recent study, Tan, Gergle et al. (2003) compared performance of users working on large display to that of users working on a standard desktop monitor. In their first study participants' performance on a reading performance task yielded no significant difference, however participants performed 26% better on a large display than on a small display for spatial orientation task. The results of their second study which compared two tasks (spatial orientation and shape test) still revealed better performance on a large display for spatial orientation task. In contrast, for shape test, where participants were asked to imagine themselves looking at a picture (as opposed to imagine themselves inside the picture for spatial orientation tasks), the results revealed no significant difference on display size. Results from this study suggest that the better performance on a large display is task dependence; that is not all tasks will result in superior performance on a large display. Similar to Patrick, Cosgrove et al. (1999), Tan and colleagues (2003) also attributed the better performance of their participants on a large display for spatial orientation task to the greater sense of presence

afforded by the large display. Thus, the better performance of the large display participants over the small display participants in Experiment 1A could have been also due to our participants experiencing more presence on a large display than on a small display.

However, it should be noted however that the display resolution used for the desktop monitor in Experiment 1A was much higher than on the projected display. The better performance of the projected display participants over the desktop monitor participants could further suggest that the difference in display resolution has a minor contribution to distance estimation tasks compared to the effect of physical display size. It is unlikely that lower resolution would lead to better performance (Duh, Lin et al. 2001, Kline and Witmer 1996). Duh and the others showed that better resolution leads to more sense of presence. Comparatively, the results of Experiment 1A indicated that the large projected display afforded more sense of presence than the desktop monitor even though it is of lower resolution than desktop monitor. As suggested by other researchers, a large display provides user a sense of presence, realism and scale (Patrick, Cosgrove et al. 2000). As such, one possible explanation for the higher sense of presence in large projected display must be contributed also by its large physical image size.

In this study and the earlier mentioned studies (Patrick, Cosgrove et al. 2000, Tan, Gergle et al. 2003, Yang, Dixon et al. 1999), the FOV of the display for all conditions are set to be equal (or similar as in our study). Equating the display FOV resulted in similar image size projected on the observer's retina. However, Experiment 1A results showed with similar FOV (and similar retinal image size), there is a significant difference between the large and the small displays. This suggests that the retinal image size as a cue to distance perception was less influential than the physical display size cues. This result provides support for previous investigations which suggest that FOV or the resulting image size is a weak cue because it is easily overridden by other cues (Beall, Loomis et al. 1995).

### 5.1.6.3 Examination of Experiment 1A setup

Figure 5-10 shows a different illustration of the experimental setup for Experiment 1A. X represent image projected on a desktop and Y represents the image projected on the large projected display. The perceived image size on the retina ( $x_1$  and  $y_1$ ) depends on the physical screen size and the FOV ( $\alpha$  and  $\beta$  subtended by both displays). As mentioned earlier, the FOV of both display types (desktop and projected display) were only slightly different (approximately  $23^\circ$  for large display and  $21^\circ$  for small display). Thus, the resulting perceived screen size (image size) on the observer's retina was only slightly different between both

display types. It was noted that the distance of the observer from the screen was different for both the large display and the small display conditions. In Figure 5-10,  $d_1$  and  $d_2$  represent the distance of the observer from the screen for the small and the large display respectively.

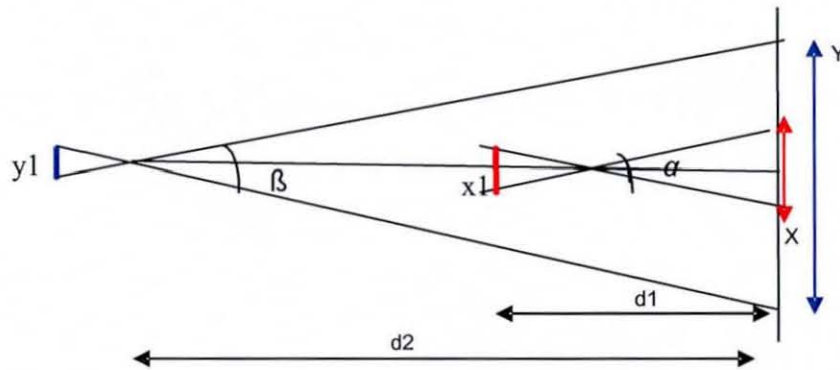


Figure 5-10 Experiment 1A Setup

As discussed in Chapter 4 (Section 4.1.2.1), the significant difference between the large and small display in previous investigations may not be attributed to the display size factor alone. Other factors such as the viewing distance and physiological cues (accommodation and vergence cues) may influence these results. This entailed the need to examine the effect of these later factors: viewing distance and physiological cues. The second study presented in the next section (Experiment 1B) was designed to investigate the effect of these factors. Basically, the distance of the observer from the screen for both display conditions were constant; by doing so, it was assumed that the physiological cues (accommodation and vergence cues) acting at the same distances would be similar. Equating the distance from display consequently changed the projected image size on the observer's retina. However, as indicated by Experiment 1A, even though the retinal image size was similar, there was a significant difference in distance estimation between the large and the small display. As discussed earlier, this suggested that the retinal image size was less influential as a visual cue. Since the difference in distance estimation was more influenced by the perceived distal size (small versus large) (Dixon and Proffitt 2002), it was expected that there should be no difference in distance estimation between both display sizes if the viewing distance and the physiological cues accounted as major factors for the difference in distance estimation.

It was noted that there was a difference in terms of display types used to present the image. A desktop monitor was used to represent the small display size condition while the projected display represented the large display condition. In the next study, this variance was controlled by using the same projected display but of different projected image size to represent the large

and small displays. A detailed description of the experimental design and method is further described in the next section (Section 5.2).

It was also noted that there was a difference in image resolution between the real and the VE picture. The image resolution of the VE picture was higher than the real picture. The use of a higher image resolution for the VE picture may have contributed to the results of similar distance perception in both the real and VE picture in our study. The difference in resolution between the real and VE image may offer one possible explanation for the better distance estimates on the VE picture for horizontal distance on desktop monitor. The effect image of image resolution on distance perception however was investigated in Experiment 2B (reported in Chapter 2).

### 5.1.7 Conclusion

Generally, findings were consistent with previous investigations showing that most distance judgments were underestimated. On average, current study revealed no significant difference between objects perceived in the real or VE image for transverse distance and horizontal distance. It was demonstrated that

- horizontal distance was estimated more accurately than transverse distance.
- transverse distance was perceived approximately 50% of the actual distance for both real and VE image
- horizontal distance estimate was more similar to the real world estimates.

On an average, results have shown that

- distance estimates on a large projected display produced significantly more accurate results compared to the same distance estimation on a small display or both horizontal and transverse distances.
- Generally estimates tend to be underestimated for both distance types (except for horizontal on large projected display).

While the results suggest that the difference may be due to the physical display size, other possible explanation could be due to the difference in the viewing distance and physiological cues. The contributions of these latter factors were examined in Experiment 1B, described in the next section.

Experiment 1A indicated that both the horizontal and transverse distances yielded different estimation values. For instance, the horizontal distance was estimated more accurately when compared to the transverse distance. While not statistically significant, a direct comparison

revealed that for the horizontal distance, estimation was more accurate on the VE image when compared to the real image. These differences thus further motivate us to include the investigation of vertical distance. All three distances (vertical, horizontal and transverse) are necessary and important for the perception of the 3D space; be it in the real world or the VE.

## 5.2 EXPERIMENT 1B: EFFECT OF DISPLAY SIZE

In order to investigate the effect of display size on distance estimation in Experiment 1A, it was necessary to maintain other factors such as retinal image size to be constant on both display types. To achieve this, the viewing distance of the observer was adjusted from the screen such that the FOV was equal (similar in our case for reasons explained earlier) for both display types. From Figure 5-10, if we make the ratio of  $x$  to  $d_1$  and the ratio of  $y$  to  $d_2$  equal we would have similar visual angle for both display size. Thus, if the values of  $x$ ,  $d_1$  and  $y$  were already set or known, the value of  $d_2$  could be easily obtained as follows:

$$\text{Since } x/d_1 = y/d_2, \text{ therefore } d_2 = (y * d_1) / x$$

However, varying the viewing distance introduces other variances beside display size in the design. Different accommodation and vergence cues might be present at different viewing distances. Thus, the main effect of display in Experiment 1A may not be attributed to the display size alone; viewing distance and physiological cues might have also contributed an effect. As reviewed in the prior chapters, both the accommodation and vergence cues (Section 2.4, Chapter 2) and the viewing distance (Chapter 2.8.1) may contribute an influence on distance perception. Both accommodation and vergence cues may provide accurate absolute depth information (Morrison & Whiteside, 1984; cited in Coren, Wards et al 1999).

It has been suggested in the previous chapter (Section 4.1.2.1 of Chapter 4), viewing an image at different distances may influence what the user may perceive. From a theoretical perspective, there is an effect of viewing a pictorial display at different distances (see Section 2.8.1 of Chapter 2 and Section 4.1.2.1 of Chapter 4). As we approach a picture, the geometrically specified depths in a picture are compressed proportionally to the closeness of our approach and as we move away from the picture, the depths are expanded proportionately. As the observer in our study is farther away from the screen for the large display compared to the observer in the small display, this might explain the larger estimations reported by the participants of a large display for transverse distance and smaller estimations for the small display. Sedgwick (1991) further explained that to realise the theoretical prediction, the picture need to contain strong linear perspective. A weak linear perspective may not reveal

this distortion. The picture in the present study has a strong linear perspective (the road and the hedge). However, empirical investigations have found such distortions in human perception but not to the predicted magnitude. Thus, if the distance of the observer from the display was fixed, the differences due to the effect of viewing distance on perception of depth especially for transverse distance were removed. Thus it is predicted that if these factors (the viewing distance and physiological cues) were to contribute a substantial effect on distance estimation, we would expect no significant difference between the large and the small displays. If this prediction is realised, we could therefore conclude that the significant difference between displays in Experiment 1A was more influenced by the viewing distance (and physiological cues) than by the display size factor.

The following major hypothesis is investigated in this study:

H1: There is no significant difference between the large and small displays on asymmetrical distance estimation tasks.

The secondary hypotheses are:

H2: There is no effect of viewing distance on asymmetrical distance perception

H3: There is no effect of physiological cues on asymmetrical distance perception.

As there was no main effect of image, that is, there was no significant difference between the real and VE image, only one image type was used. A real picture was used to test the condition in this study and no VE model picture was used.

## 5.2.1 Methodology

### 5.2.1.1 Participants

Twenty volunteers (10 females and 10 males), comprising of staff and students from Loughborough University took part in this study. The ages of the participants ranged from 18 to 41 years with an average age of 25.4. All participants either had normal or corrected-to-normal vision.

### 5.2.1.2 Materials/Apparatus

#### Real picture

For this study, a different picture was used. However, as in Experiment 1A, a location with few objects was required because more objects provide more cues to participants to base their



estimation thus creating more variance among participants. A location in campus was identified and this location is similar to Experiment 1A picture in terms of the presence of a road (perspective cues) and other objects such as trees. A photograph of a location on campus was taken using a digital camera and this picture is placed on Microsoft PowerPoint slide (Figure 5-11) as in Experiment 1A. The image resolution was 1280 x 960.



**Figure 5-11 Real picture**

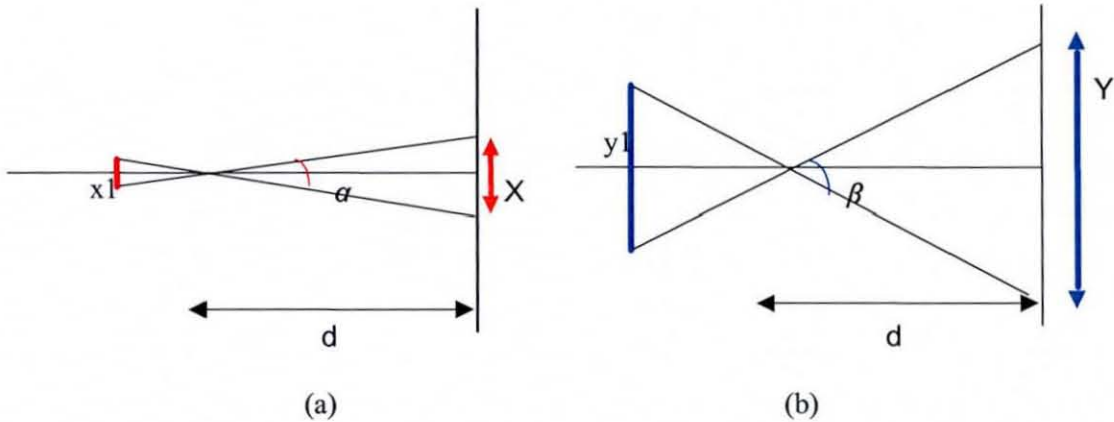
#### 5.2.1.3 Display apparatus and room setting

A large rear-projection screen was used to present the image. Two display area sizes were used: large display (156 x 208cm), small display (39 x 52cm). The distance from the screen was fixed at 100cm. As the observer was not allowed head or body movement, a closer distance may have caused the observer to fail to notice the lower end corners and top end corners of the projected image for the large screen condition. Thus, this distance was chosen to allow for the complete viewing of the projected image. The resulting FOV for the large display was  $92^{\circ}$  and  $29^{\circ}$  for the small display. The resolution of display was set at 1028 x 768. A dark room setting (except for the light from the projector screen) was employed here to reduce the peripheral view effects from the objects surrounding the projector screen that might affect the participants' distance estimation (Eby Braunstein (1995); cited in Knapp 1999).

#### 5.2.1.4 Experimental setup and design

The experiment was a between-subject design consisting of one IV (display size) and the DV (estimated distance). The two levels of display size were small and large. The three types of DVs were vertical, horizontal and transverse distance. Two experimental conditions were used for this study: Small display/real image and large display/real image. The participants were randomly assigned to each group of ten participants. Similar to Experiment 1A, different groups of participants were used for each condition to avoid training bias or interference from

previous knowledge. Additionally, only one picture, viewed in non stereo mode, was used for the study to ensure that the same visual information cues were available in both display conditions.



**Figure 5-12 Experiment 1B setup for small (a) and large display (b).  $X$  and  $Y$  represent the projected image size on the screen and  $\alpha$  and  $\beta$  represent the FOV for small and large display condition respectively. ' $d$ ' represent the fixed distance**

Figure 5-12 illustrates the setup of Experiment 1B. The left diagram represents the small display size condition and the diagram on the right represents the large display size condition. From this diagram, when the distance of the observer is the same for both display conditions (represented by  $d$ ), these consequently change the FOV of the display ( $\alpha$  and  $\beta$  for small and large display respectively), hence the projected image size ( $x1$  and  $y1$ ) on the observer's retina would be different. The eye level (centre of projection) was made similar for each participant. This was done by adjusting and positioning the chair equidistant from the edges of the picture. The chair height was also adjusted accordingly. A small weight hanging from the ceiling was used as a reference to locate the height of the eye position. The display screen (rear projected screen), display resolution (1028 x 768), and distance from the screen (100cm) were held constant in both experimental conditions. The FOV and projected image on the screen depended on the condition of the experiment (see Figure 5-12).

### 5.2.1.5 Procedure

The same procedure used for Experiment 1A was employed in Experiment 1B with one exception. Instead of estimating only two distances, participants in Experiment 1B estimated a total of nine distances (three distances for each vertical, horizontal, transverse distances).

## 5.2.2 Results

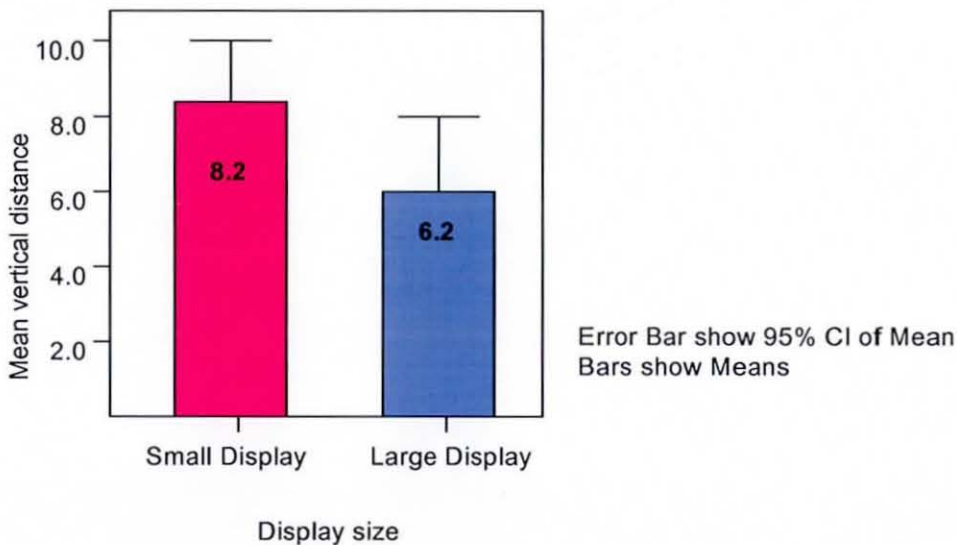
### 5.2.2.1 Data preparation

Preparation of the data was similar to Experiment 1A. Examination of the data revealed two outliers (based on z-scores). These data were removed prior to analysis. Another two cases of data whose estimate was more than twice the actual estimates were also removed.

A one-way between-group MANOVA was performed separately for each distance type (vertical, horizontal, and transverse) to explore the effect of display size (small versus large) on distance estimation. The three DVs were the three different estimated distances for each distance type. The IV was display type (large versus small screen). Significant level was set at 0.05.

In the next section, the results for vertical, horizontal and transverse distance are first presented. This is followed next by a section on comparison among distance types.

### 5.2.2.2 Vertical distance



**Figure 5-13 Comparison of mean estimated distance between display size for vertical distance**

Figure 5-13 shows a comparison of the mean estimated vertical distance on the large and small displays. Inspection of Figure 5-13 indicates that there is a difference between distance estimation on both display sizes. The results of MANOVA analysis revealed that the difference between the large and the small display was not significant for all the three distances; that is the main effect of display did not reach statistical significance [ $F(3,12)=1.440, p=.280, \text{Pillai's trace}=.265, \text{partial eta squared}=.265$ ]. It was noted that the



magnitude of the effect size was considered large ( $\eta^2=.265$ ). This means about 26% of the variance in vertical distance estimation was explained by display size. From Figure 5-14, estimated distances were generally larger than the actual distance for the large display (as indicated by the line position above the actual distance line) and smaller for the small display (as indicated by the line position below the actual distance line). The figure also indicates that, overall, distance estimates on the small display were more accurate when compared to the large display.

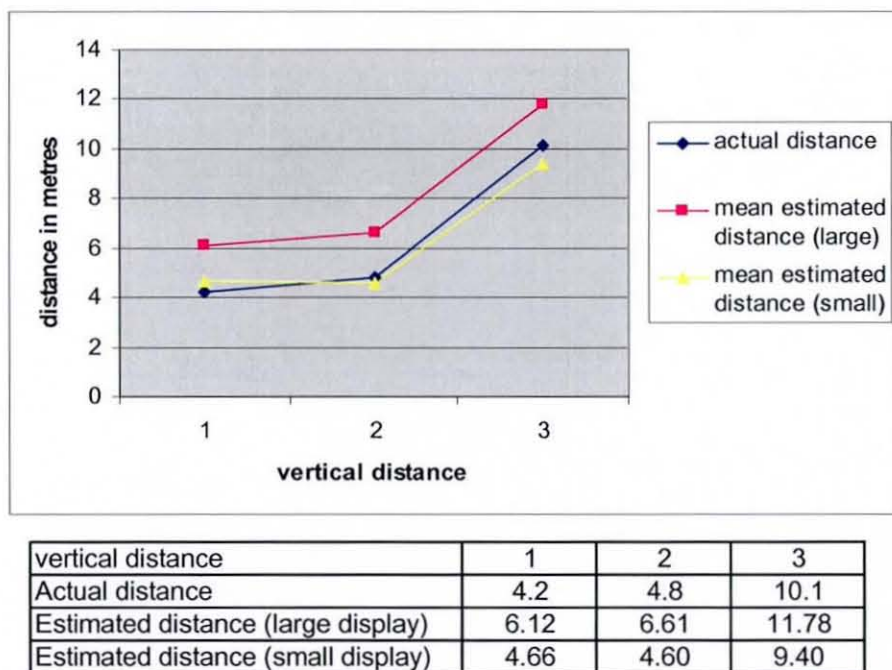
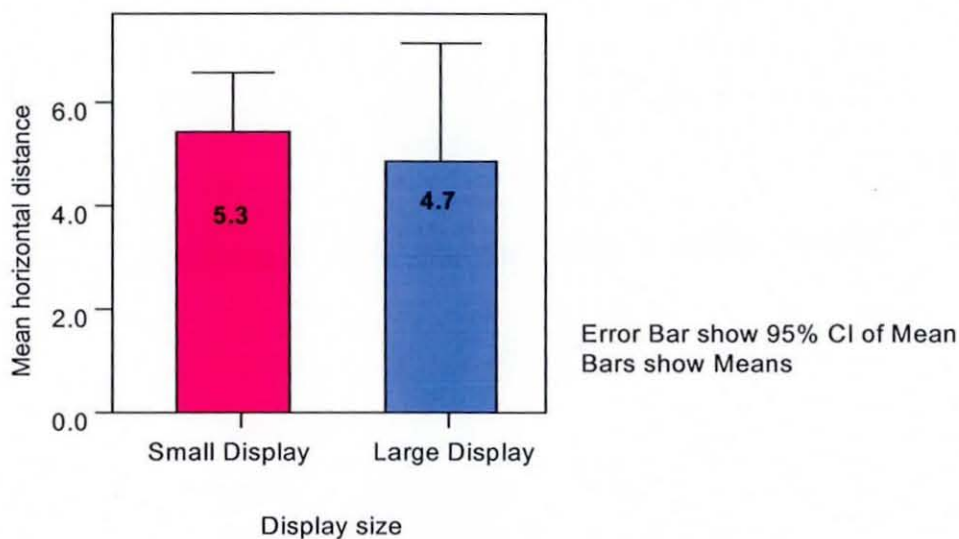


Figure 5-14 Comparison of mean of estimated vertical distance to the actual distance between the large and small displays. Accompanying table indicates points on the graph

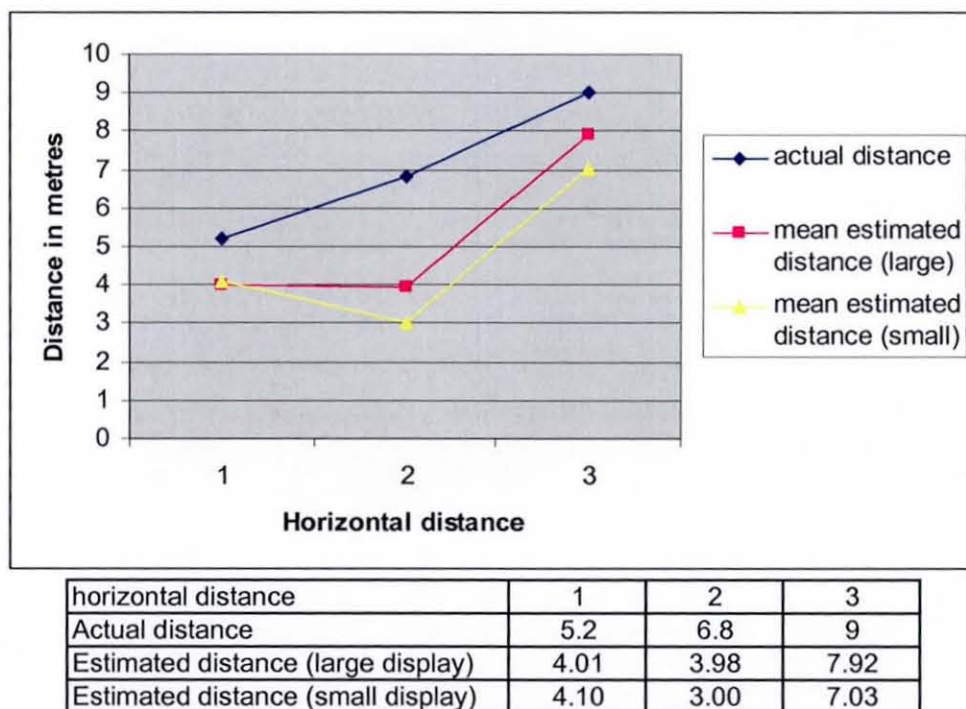
### 5.2.2.3 Horizontal distance

A comparison of mean estimated horizontal distance between the large and the small display is illustrated in Figure 5-14. Inspection of Figure 5-14 indicates that the difference between estimation on the large and the small displays was small. The results of a one-way between groups MANOVA analysis supported this observation. The main effect of display [ $F(3,12)=2.62$ ,  $p=.099$ ; Pillai' trace=.396; partial  $\eta^2=.396$ ] did not reach statistical significance, that is, there was no significant difference between the large and the small display on the combined three distances. However, it was noted that the magnitude of the effect size was considered large ( $\eta^2=.396$ ), that is about 39% of the variance in horizontal distance estimation was explained by display size factor.



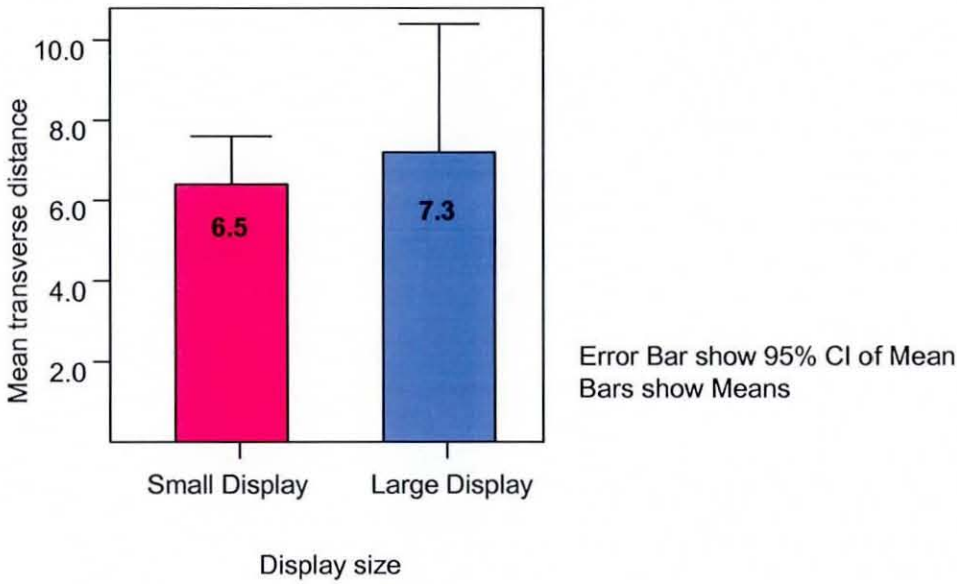
**Figure 5-15 Comparison of mean estimated distance between display size for horizontal distance**

From Figure 5-16, distances were generally underestimated for both display conditions (as indicated by the lower position of the estimated distance lines than the actual distance line). Overall, from Figure 5-16 estimates for the large display were more accurate when compared to the small display.



**Figure 5-16 Comparison of mean of estimated horizontal distance to the actual distance between the large and small displays. Accompanying table indicates points on the graph**

5.2.2.4 Transverse distance



**Figure 5-17 Comparison of mean estimated distance between display size for transverse distance**

Figure 5-17 shows a comparison of mean estimated distance between the large and small displays for transverse distance. Figure 5-17 indicates that the difference of estimated distance between the large and small displays was very small. However, the results of a MANOVA analysis on this dataset revealed the main effect of display [ $F(3,12)=3.339$ ,  $p=.056$ ; Pillai's trace=.455; partial eta squared=.455] did approach statistical significance. The magnitude of the effect size was considered large (eta squared=.413), that is about 41% of the variance in transverse distance estimation can be explained by display factor.

From Figure 5-18, the transverse distance was greatly underestimated especially for large distances when compared to the actual distance. The figure further shows that distance estimates on the large and small displays did not differ very much for all the three distances.



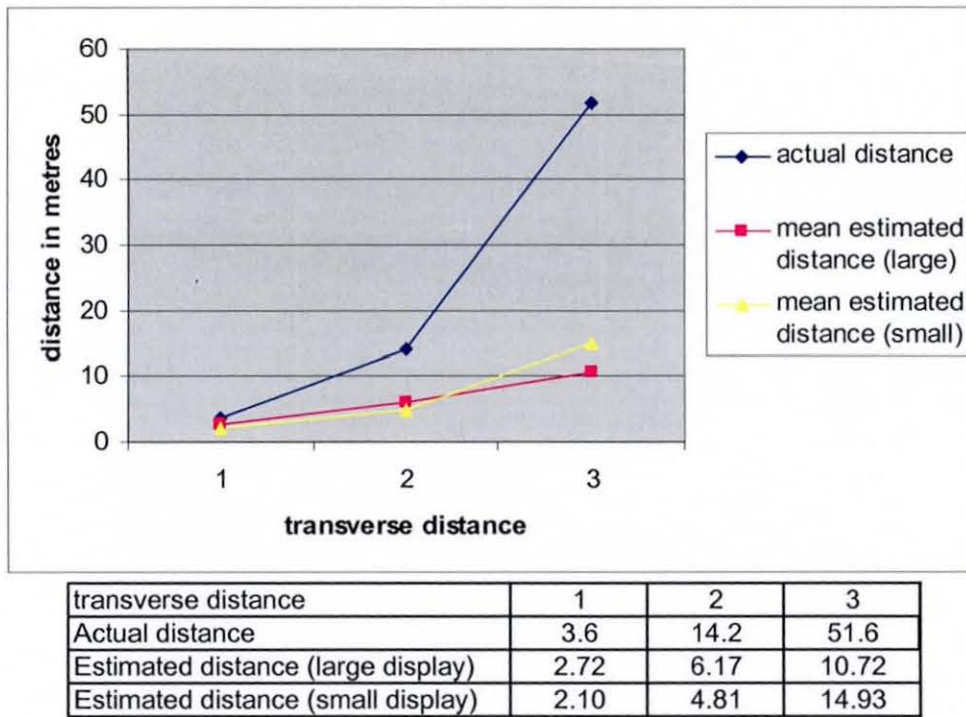


Figure 5-18 Comparison of mean of estimated transverse distance to actual distance between large and small display. Accompanying table indicates points on the graph.

### 5.2.2.5 Comparison among asymmetrical distances

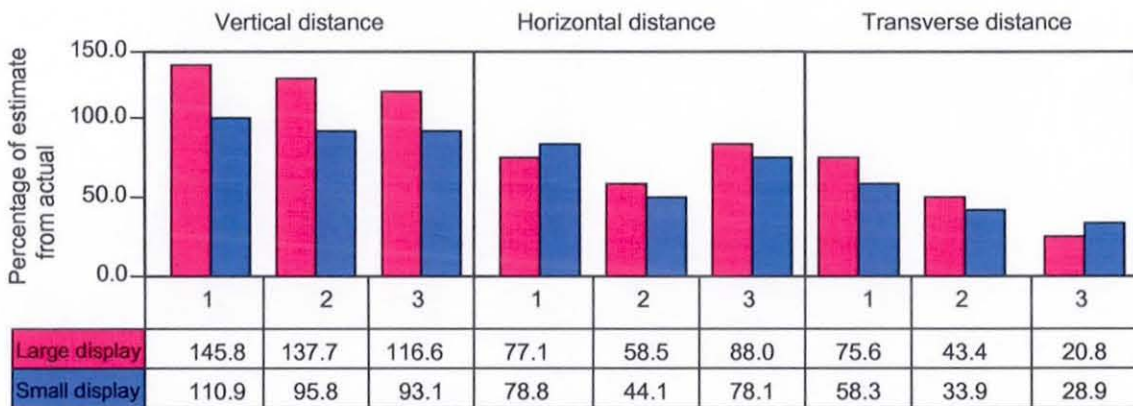


Figure 5-19 Comparison among asymmetrical distances

Figure 5-19 provides a comparison among the three asymmetrical distances. From the figure, vertical distances were generally overestimated on a large display and underestimated on a small display. Generally, distances were underestimated for both the horizontal and transverse distances regardless of display sizes. However, underestimations were much greater for the transverse distance compared to the horizontal distance on both display sizes. On average, estimates were better on a large display than on a small display for the horizontal and transverse distances. In contrast, estimates were better on a small display than on a large display for the vertical distance.

### 5.2.3 Analysis

Findings from this experiment revealed that there was no significant main effect of display for the three asymmetrical distances. However, the main effect of display did approach significance for the transverse distance. The large magnitude of effect size for all distances suggested that a large percentage of the variance in distance estimation was explained by the display size. This was also indicated by the large differences between the mean percentages of estimate scores.

Generally, distances were overestimated on the large display and underestimated (but more accurate) on the small display for the vertical distance. In contrast, larger error was reported on the small display compared to the large display for the horizontal and transverse distances. Overall, distances were underestimated for both the horizontal and transverse distances. For the vertical distance, the magnitude of effect size was considered large indicating a large percentage of the variance in the vertical estimation was explained by the display size factor. A similar observation was noted for the horizontal and transverse distances.

### 5.2.4 Discussion

It was predicted earlier that there would be no significant difference between the large and small displays if the viewing distance and the physiological cues do cause a variation in the participants' distance estimation. The findings of this study confirmed this prediction. The difference between the large and the small displays was small, thus the information provided by the accommodation and vergence cues did contributed a large influence than the display factor in the distance estimation task. This result provides support for previous studies which showed that when we look at pictures our eyes converge and accommodate as if we are looking at objects at various distances in responses to the pictorial depth cues found in the pictures (Enright 1987a; Enright 1987b; cited in Coren and Ward et al. 1999). It is generally accepted that the judgment of distance is based, to some extent, upon these physiological process (Swenson 1932). However, their range of effectiveness is limited to short distances (Sekuler and Blake 1994). For accommodation cues, its effectiveness is up to 2m (Schiffman 1990) and for vergence cues it is useful for a distance of up to 6m (Howard and Rogers 2002). As the viewing distance was set at 1m (100cm) from the display in Experiment 1B, 0.4m (40cm) and 2.8m (280cm) from the display in Experiment 1A, these distances are within the range of effectiveness of both cues. Thus, both cues were available to participants as information for their distance perception.

Additionally, from the geometrical theory of perception, Sedgewick (1991) showed that as we approach a picture the geometrically specified depths in a picture are compressed proportionally in accordance with the closeness of our approach and as we move away from the picture, the depths are expanded proportionately. This is especially true when the picture contains strong linear perspective as in our picture (the narrowing of the road at the far distance). Thus, as the magnitude of effect size for display factor is considered large (vertical distance: 26.5%; horizontal distance: 39.6%; transverse distance: 45.5%); this suggested that the display factor still accounts for a large percentage of the variation in distance estimation.

These results are consistent to Experiment 1A, where large errors were reported on the small display compared to the large display for both the horizontal and transverse distances. This implies that distance estimations on a large display are more accurate compared to a small display. Similar to Experiment 1A, the effect size in this study was also large suggesting that the large variation in distance estimation is still explained by the display factor.

Distances were largely underestimated in transverse distance especially for larger and farther distances where estimates were less than 30% from actual. Comparatively, nearer and shorter distances were estimated more accurately. Consistent with the findings from Experiment 1A, the horizontal distances were estimated more accurately than the transverse distances. In Experiment 1A, the horizontal distances were overestimated on a large display. However, in Experiment 1B, the horizontal distances were underestimated on both display sizes.

Experiment 1A used a desktop monitor for the small display condition and a projected display for the large display condition, while Experiment 1B used a projected display (with adjusted display area for the small and large conditions). Consistent distance estimation results in both experiments reflect that the variation in distance estimation between small and large display in Experiment 1A was not due to the display type (desktop versus projected display) but the display size. This result supports findings from an earlier study (Patrick, Cosgrove et al. 2000) which suggests that the large image size in the projected display induce realistic experience in participants, thus giving them better judgment of distance on the projected display compared to desktop participants.

In Experiment 1B, vertical distance was also investigated. Findings from this study suggest that vertical distances which are nearer to the observer tend to be overestimated more than those located farther away from the observer. The result shows that vertical distances were generally overestimated in the large display condition and underestimated (but more accurate)

in the small display condition. This finding is consistent with the VHI condition found in pictures where vertical distance tends to be overestimated. It is also in line with past research (Dixon & Proffitt, 2002; Yang, Dixon et al. 1999; Higashiyama and Ueyama 1988) where vertical distance tends to be overestimated. Higashiyama and Ueyama (1988) found that when the vertical and horizontal distances were physically equal, the vertical distance tended to be perceived larger than the horizontal distance. Yang, Dixon et al. (1999) found that vertical overestimation is reduced in desktop monitors compared to the real and VR conditions. They believed that the reduced overestimation on the desktop condition was due to the small size projection on the picture surface. This implies our visual system is influenced by the perceived physical size of the projection. Yang, Dixon et al. (1999) proposed that vertical overestimation would increase if a picture was distended such as projecting it onto a larger screen. This prediction is confirmed by a later study by Dixon and Proffitt (2002). Consistently, the results of Experiment 1B also revealed that more distance overestimation was found on a large display compared to on a small display.

Witmer and Kline (1998) showed that estimates were more accurate to a small cylinder (stimulus used in their experiment) than to a large cylinder. As vertical estimates are more accurate compared to horizontal and transverse distances, one possible implication from Witmer and Kline (1998) study is that participant may base their estimations on objects' height. Thus, this may explained the more accurate result for the small cylinder compared to the large cylinder in Witmer and Kline (1998) study. Similarly, Experiment 1B showed that vertical distance estimations on a small display were more accurate compared to the large display. These estimations were in contrast to estimation in horizontal and transverse distances. This dissimilarity provided the motivation to further include vertical distance in our subsequent studies.

### 5.2.5 Conclusion

Consistent with the findings from Experiment 1A, the transverse distance was underestimated more when compared to the horizontal distance. Similarly, while not significant, a direct comparison of means indicated that the large display yielded more accurate estimates than the small display.

Additionally, despite the use of different display types (desktop monitor- for small display and projected display- for large display) in Experiment 1A and the use of similar projected display in Experiment 1B, the consistent results in both studies suggest that the display type is less influential than the physical display size.

It is asserted that, in addition to the physical display size which might induce realism and greater sense of presence on the user, other factors which contribute to the difference between display sizes are the viewing distance and the physiological cues of accommodation and vergence.

### 5.3 OVERALL CONCLUSIONS

This chapter presents a set of two experiments which examined user's spatial awareness in static images of the real and VE. A description of experimental methodology, results, discussion and conclusions for Experiment 1A and 1B was provided. Experiment 1A aimed to investigate the effect of image type and display type on asymmetrical distances (horizontal and transverse distance) while Experiment 1B extend this investigation by examining the effect of viewing distance and subsequently physiological cues on user's distance estimates.

The results of Experiment 1A revealed that there was no significant difference for distance estimation between the real and virtual picture. This suggests that it is possible to perceive distances in the real and VE picture similarly within the constraint of this experiment.

However, the main effect of display was significant, suggesting the physical display size factor has a significant effect on distance estimate tasks. It was shown that more accurate estimates were found on the large projected display compared to the desktop monitor.

The results of non-significant effect of display size in Experiment 1B further suggest that the viewing distance and the physiological factors also contribute largely towards the significant effect of display type in Experiment 1A. However, the large effect size for display size indicates that display size still constitute major factor of influence in distance estimation tasks.

In the next two chapters, investigation into spatial awareness in dynamic real and VE images (Chapter 6) and in interactive real and VE (Chapter 7) are presented. An overall analysis of the results of experiments presented in this chapter will be further discussed in Chapter 8, along with the results from experiments reported in Chapter 6 and 7.



# CHAPTER 6

---

## *EXPERIMENT 2: DISTANCE PERCEPTION*

### *IN DYNAMIC IMAGES*

## 6 OVERVIEW

The results of the set of studies in Experiment 1 on static images reported in Chapter 5 revealed that the physical display size, viewing distance and physiological cues contributed significantly towards participants' distance estimate performance. The results further showed that the participants' distance estimate performance on the real and VE picture did not differ greatly. This indicates that it was possible to perceive distances similarly on a non-stereo real and VE static picture within the given experimental constraint.

It has been suggested that the extendibility of these conclusions to dynamic images is questionable (Stanney, Mourant et al. 1998, Peruch, Vercher et al. 1995). In dynamic images such as movies, video sequences or animations, the viewing perspective of the viewer dynamically changes due to movement or motion. This created effects that are not experienced by viewers of static or static images. In static images the viewing perspective will always stay the same and would not change at all. But in dynamic images, the view of the

spatial environment may change dynamically based on movement; for example, the size of objects may expand or contract depending upon whether the viewer is approaching or moving away from the objects respectively. This extra information is not available in static images. As such, the conclusion derived from the investigation of spatial awareness based on static images (Experiment 1) might not be valid when dealing with dynamic images.

Therefore, in this chapter, the experimental methodology and the results of investigating user's spatial awareness in dynamic images are outlined. The general experimental approach was similar to that of Experiment 1. However, video images of the real environment and its computer generated VE model were used to represent the real and VE. The general aim of Experiment 2 is to compare participants' asymmetrical distance estimates performance between the real and VE image presented in non-stereo and non-immersive/semi-immersive mode. The effect of display size and viewing distance was also investigated. The first study (Experiment 2A) investigated the effect of image type and display size while the second study (Experiment 2B) investigated the effect of viewing distance, physiological cues and image resolution. Discussion of the results and conclusions are presented at the end of the chapter.

## 6.1 EXPERIMENT 2A: EFFECT OF IMAGE TYPE AND DISPLAY SIZE

### 6.1.1 Rationale

The extendibility and validity of conclusions derived from Experiment 1 which was based on static images are questionable when dealing with dynamic images such as movies, video sequences and animation. Stanney, Mourant et al. (1998) have noted that depth perception in dynamic scenes are complex and not well understood and thus suggested that "it is important to conduct depth perception studies in both static and dynamic scenes as the results from the former may not generalized to the later".

There are differences between the static and dynamic scenes or images. In static images, the viewing perspective always stay the same; that is the relationships among objects will always be the same regardless of the observer's viewing positions. In dynamic images, the presence of motion or movement changes the viewing perspective of the viewer which creates effects that are not experienced by static images viewers. Additionally, moving closer towards objects or moving away from the objects may results in the respective expansion or contraction of the retinal image sizes on the viewer's retina. As reviewed in Section 2.3.2 of Chapter 2, in addition to static cues, cues to motion (such as motion parallax and motion perspective) were available to a moving observer. As such, these differences in information or

cues provided by the static and dynamic images provided motivation for the investigation of dynamic images. However, in this research the effect of motion parallax cues resulting from the head and body motion were eliminated by fixing the viewer's head and body movements in the experimental trials. While the relative effectiveness of motion parallax cues was acknowledged, it was necessary to remove these cues in order to avoid their confounding effects on the results of investigations for the intended factors. Furthermore, the investigations in this research are limited to non-head tracked conditions.

An investigation by Willemson and Gooch (2002) which compared egocentric distance perception in a real image-based VE and a computer-generated VE (both viewed stereoscopically on a HMD) revealed small differences; though, the image-based VE was shown to perform slightly better. An earlier study by Yanagisawa and Akahori (1999), found that following a virtual tour of the VE campus, their participants formed a more accurate spatial representation of the computer-generated virtual campus compared to a photographed-based VE campus. The authors suggested that the photographed-based virtual campus contains more detailed information in the image compared to the computer-generated virtual campus thus imposing more cognitive load on the photograph-based virtual campus participants. Consequently, less mental effort was available for the acquisition of the survey knowledge of the virtual campus for the photograph-based virtual campus participants, thus, resulting in their poor performance compared to the computer-generated VE campus participants. These contradicting evidences from these two studies provided further motivation for investigation of the differences in spatial awareness performance in the real and VE.

### 6.1.2 Experiment aims and hypotheses

The overall aim of this study was to examine participants' spatial awareness in video representation movies of a real and VE. Exocentric distance estimate in terms of vertical, horizontal and transverse distance as proposed in Chapter 4 was employed as performance task measure.

The two major hypotheses investigated in this study were:

- H1: There is no effect of image type (real and VE image) on asymmetrical distance perception (vertical, horizontal, transverse).
- H2: There is no effect of display size (small and large size) on asymmetrical distance perception (vertical, horizontal, transverse).

Similar to Experiment 1, images were presented in a non-stereo mode in order to exclude stereo cues and to examine the effects of other cues on distance perception. The small and large displays corresponded to the non-immersive and semi-immersive display respectively.

### 6.1.3 Methodology

#### 6.1.3.1 Participants

Forty volunteers (equal number of males and females) participated in this study. The average age of the participants was 36.15 with age ranging from 23 to 50 years. The forty volunteers were randomly allocated to each of the four groups comprising of ten members each. All participants had normal or corrected-to-normal vision.

#### 6.1.3.2 Materials/Apparatus

##### Real environment



**Figure 6-1 Real Environment**

For this experiment, a location with few visual cues but with an adequate number of objects was required. Few visual cues were necessary to reduce variance among participants and to focus on the impact of the variables (image type and display size) under investigation. A football practice field on campus was identified to meet this requirement. Similar to Experiment 1, an outdoor setting was employed in this study.

For the real world condition, a video movie of the practice football field was used to represent the real world condition (see Figure 6-1). The movie was taken by capturing the scene while walking forward along a predefined path from one corner of the field to its opposite end using a digital camcorder. This provided the user with a forward view of the scene only. The

forward movement was chosen as it is a more natural for viewing than sideways (lateral) motion even though the later contains more depth information. The movie was downloaded to a computer and was edited using the Adobe Premiere software. The movie was later saved as an AVI file format. The video image resolution was 720 x 576.

## Virtual Environment



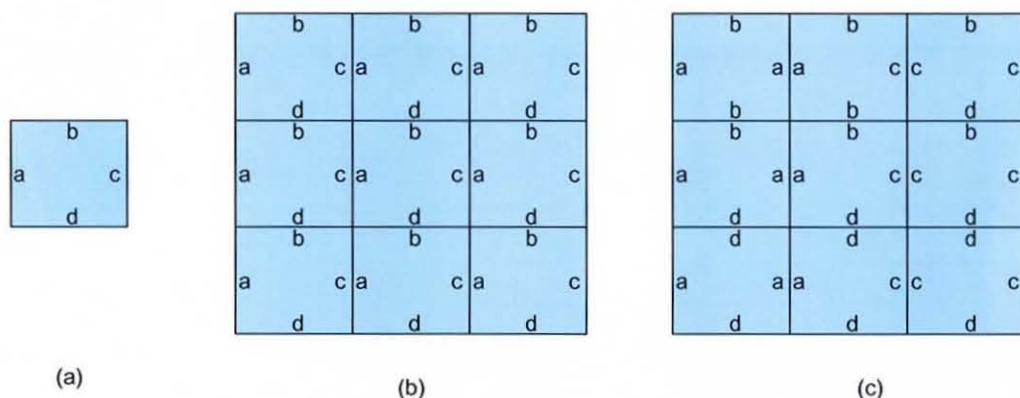
**Figure 6-2 Virtual Environment**

For the VE (Figure 6-2), the scene was modelled using the MultiGen II Pro software which runs on a SGI workstation. Detailed measurements of the location and its objects were carefully taken before the modelling process. Pictures of the objects at the location were taken using a digital camera. Appropriate textures from these pictures (e.g. grass, trees, and road textures) were used as texture maps in the modelled scene in order to match the VE model as close as possible to the real world. Preparations of the texture maps were done using an image editing software called Micrografx Picture Publisher 8. For objects such as trees and litterbin and lamps, a billboard technique (Section 3.1.2.2 of Chapter 3) was employed. Outlines of the image were first created from the pictures. The images of the objects themselves were then extracted from the pictures. These were then placed on a transparent background in the Adobe Photoshop editing software and a special function in this software was used to export the transparent image (in GIF format) to the SGI computer. These images (or texture maps) were used to create corresponding objects in the VE model. In Multigen II Pro Software, these images were placed on a billboard (polygonal faces that always facing the viewpoint).

Objects such as hedges, grass, roads and sky have continuous and repeated textures. Textures patterns from the objects were initially taken by extracting part of the image and saving it in JPEG format. In the VE model, the textures were placed onto the object surface using one of the texture projection tools which depends on the shape of the objects. The repetition of the



textures for small objects was not obvious; however for large objects such as the grass, the road and the sky, the lines of repetitions of the textures were clearly visible. Moreover, when the image was viewed using a viewer software (SGI Performer PERFLY), these repetitions of textures created an undesirable shimmering waves effect. The initial method was to use only one copy of the texture (see Figure 6-3 (a)).



**Figure 6-3 (a) original texture copy (b) repetition of the same texture pattern (c) copies opposite edge of the pattern to make it similar to the edges of the centre pattern**

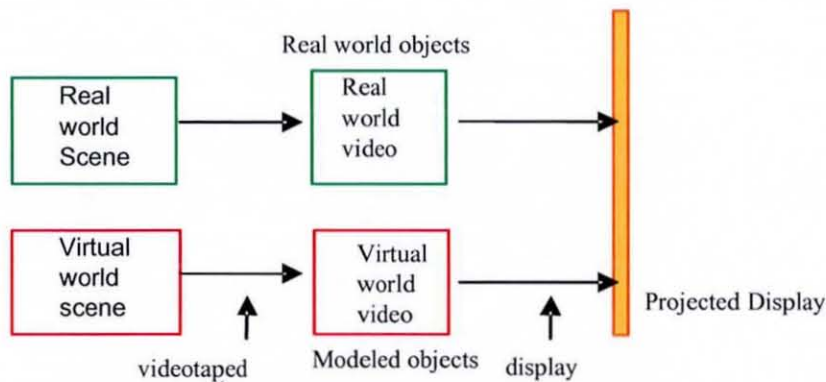
When this pattern was repeated many times across the object surface the repetition lines were clearly visible. This was especially true when there were some differences in the textural pattern at the texture's edges (a,b,c,d) (see Figure 6-3(b)). That is, when the copies of texture pattern are placed next to each other, a 'line' seems to divide between each repeated texture pattern. It was realized that the problem was the texture's size was too small and a large surface area needed to be covered, thus more repetition of texture patterns across the object's surface resulted in more repetitions of the 'lines'. To resolve this problem a bigger texture size was required. First eight copies of the extracted textures were made and arranged as in Figure 6-3(b). Using the cloning and painting function of the retouch tools of Micrografx software, the textures from the opposite edges of the outside squares were copied to match the edges of the centre square. The results were shown in Figure 6-3(c). The edges between squares were then blurred to reduce the effect of the 'lines' between the squares. This method yielded a bigger and more continuous texture pattern, thus less repetition of the texture pattern was needed. This method greatly reduced the shimmering effect especially for the grass (which has the largest surface area). This process was done for the grass and the sky texture pattern.

The shadows of objects were also approximately modelled in correspondence to the shadows in the real video movie. Due to its simple implementation, the shadows were modelled using a set of polygons. Employing this method has the advantage of reducing the demand on the

computational resources compared to the creation of more realistic shadows, thus improved system performance. Moreover, as reviewed in Chapter 3, it has been shown that a shadow's shape (polygonal shadows verses true shadows) has no effect on the perception of the object position and size.

For the sky effect in the VE, a very large hemisphere model was created in MultiGen II Pro. The inner surface was textured using a sky texture taken from the pictures of the real scene. The VE model was strategically placed in the hemisphere so that the lines of the polygons that made up the hemisphere were not visible. This was possible to do because the scene was viewed in one direction only.

Movements in the VE model were simulated and recorded similar to the movements in the video movie using the SGI Performer PERFLY. The viewpoint in the VE model was set to 1.4m, the height at which the actual scene was taken. This simulation was captured onto video tape, and then converted to the AVI file format. Figure 6-4 below describes the process of acquiring the real and VE movies.



**Figure 6-4 Images - Method of acquisition and display**

There were several reasons for converting the original simulated PERFLY movie to the AVI file format. The main reason was to enable the user to control the flow of the movie. The Windows Media Player which displays the real video movie provides control buttons such as pause, stop, and play buttons to do this. This option however was not available in the PERFLY software. Other reasons include maintaining consistency for both image types in the following:

- the process of image acquisition (see Figure 6-4)
- the use of a viewer software. In this study, the Windows Media player was used to run the AVI format of both movies on a PC. The VE movie otherwise would be viewed using a

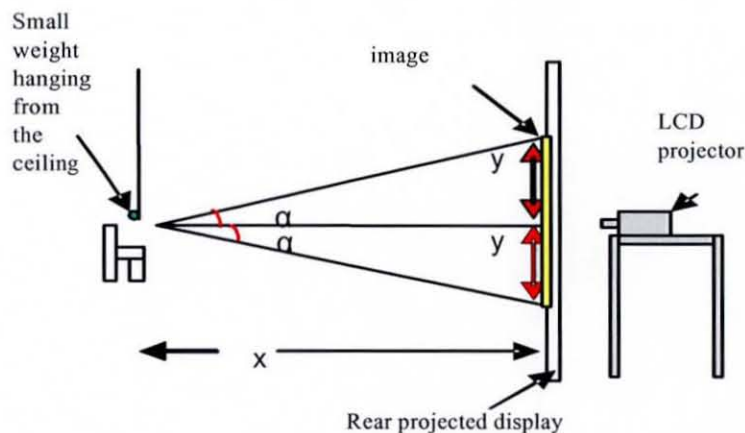


PERFLY viewer software on a SGI machine, while the real movie would be viewed on a PC using the Windows Media Player

Additionally it was not possible, however, to save the simulated PERFLY movie directly to the AVI file format. Thus, the simulation was first recorded on a VHS tape, and then transferred to a PC. The resultant image resolution was 200 lines (resolution of the VHS of the tape).

### 6.1.3.3 Display apparatus and room setting

The movies (real and VE) were displayed using an LCD projector connected to a computer. A large rear-projected flat screen was used to view the images. The display area size on the screen was adjusted to two size conditions: small display (30 x 40cm) and large display (136 x 179cm) condition. Due to the room size constraints, for the large screen condition, a large mirror was used to reflect the images on the screen to increase the image area size (Figure 6-5).



**Figure 6-5 Experiment 2A display setup.**  $x$  is the viewing distance.,  $2\alpha$  is the display FOV and  $2y$  is the image size.

The experimental room had no window thus giving it a dark condition when the lights were switched off. A dark setting was desirable here to reduce peripheral view effects from objects surrounding the projector screen which might have affected participants' distance estimation (Eby & Braunstein (1995), cited in Knapp (1999)).

### 6.1.3.4 Experiment setup and design

The experiment involved a 2 x 2 between-subject design with two IV (image type and display size) of two levels each, thus yielding the following experimental conditions:

- real world movie/ small display

- real world movie/large display,
- VE movie/small display and
- VE movie/large display

The DVs were the three types of distances: vertical, horizontal and transverse. The following variables were held constant: display resolution, display used (projected display only), FOV, eye level, textures of images, shadows, viewing and movement methods, the paths through the scene and the room setting (dark room). The movement methods in the movies were restricted to play, forward and pause only. The movement speed through both scenes was set at 1.08 m/s, matching the walking speed in the real scene. The eye level was set at the centre of the projected display. Similar to Experiment 1A, the FOV of both display sizes were equated by adjusting the distance of the viewer from the display size (see Figure 6-6). Table 6-1 provide a summary of the main experiment variables.

Table 6-1 Summary of experiment variables

Display type	Distance from screen	FOV	Display Resolution
Small	60cm	28degrees	1024 x 768
Large	272cm	28degrees	1024 x 768

6.1.3.5 Procedure

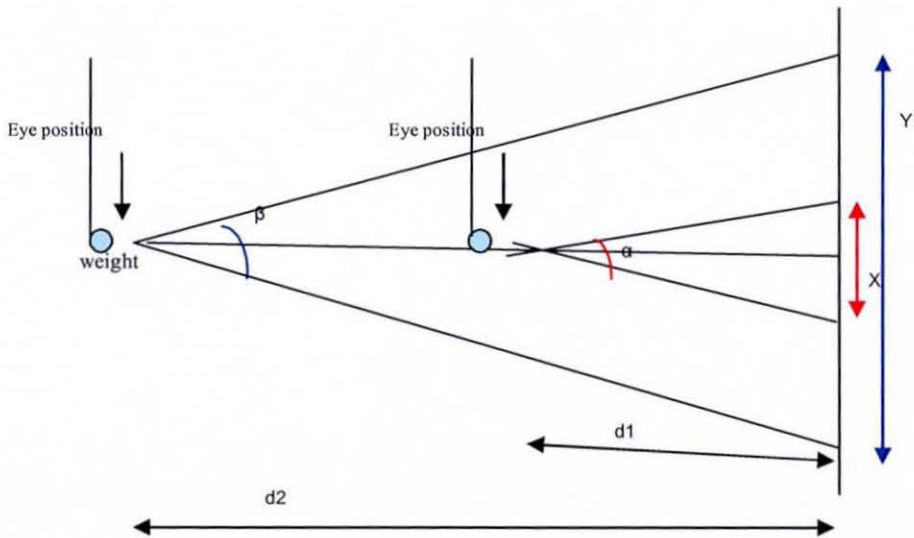


Figure 6-6 Experiment 2A setup. Small display condition indicated by viewing distance d1, FOV  $\alpha$  and image size X. Large display condition is indicated by viewing distance d2, FOV  $\beta$  and image size Y. The eye level is set to be at the centre of projection

Participants were initially briefed on the purpose and the procedure of the experiment. Based on the experiment condition, participants were seated at the assigned distance for each display condition. The effect of head motion parallax for each participant was held constant by restricting head and body movements. The eye level for all participants was kept constant by



adjusting the height of the seat of each participant. A small weight hanging from a ceiling, set to the eye level height was used as a reference (see Figure 6-5 and Figure 6-6).

Prior to making estimations, participants were allowed to view the movies to familiarize themselves with the environment and the objects in it. Participants were reminded that movement was restricted to play, forward and pause only using the mouse buttons. As the movement tasks were simple play/forward/pause of the movies, practice using the mouse to do so was not necessary. However, the participants were informed of the respective functions of the mouse buttons. Participants were allowed to view and review the movie three times (about four minutes). Participants were then informed when the time was up. The experimenter then set the scene at a preset viewpoint in the movie. Participants were then told what distance to estimate based on what they saw earlier. They viewed the static scene from this viewpoint for another 15 seconds before reporting their estimates. This was repeated for each of the eighteen distances, that is, six distances for each of asymmetrical distances (horizontal, vertical and transverse distances). All estimations were made in meters. A meter long ruler was shown to participants (vertically and horizontally) as an aide memoire. Each participant then completed a short post-test questionnaire. As recommended by an anonymous reviewer (see Section 2.7.2, Chapter 2), the post-test questionnaire also collected information on participants' sporting background. The datasheet and instructions can be found in Appendix B.

## 6.1.4 Results

### 6.1.4.1 Data preparation

A preliminary report of an initial analysis of this data set was done and reported (Awang-Rambli and Kalawsky 2003). Preparation of the data was similar to Experiment 1A in Chapter 5. Any outliers identified by the z-scores checking method (Field 2000) was removed prior to further processing of the data.

Participants' performance accuracy was measured in terms of how close was their estimated distance to the actual distance. Similar to Experiment 1, the estimated distances were normalised as percentages of the actual distance using the same formula (Section 4.2.2 of Chapter 4). As indicated in the Chapter 5, the conversion into percentage format was necessary to allow statistical comparisons and analysis of the different lengths of the asymmetrical. The percentage of estimate from actual allowed us to express estimates as an overestimation (more than 100) or an underestimation (less than 100).

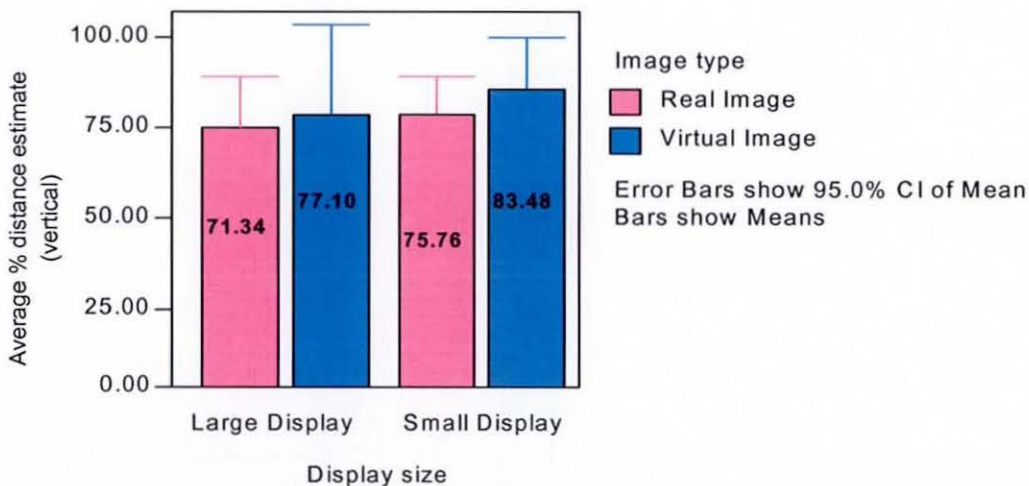


A separate two-way between-groups MANOVA was performed to investigate the effect of image type and display size on each of the asymmetrical distances (vertical, horizontal, and transverse). Conducting a MANOVA on all the related DVs would yield the overall effect of the IVs on the linear combination of all the DVs. As the concern of this experiment was on the effect of the IVs on the six estimated distances of each asymmetrical distance, the MANOVA analysis was performed on the dataset. The DV were the six different distances for each of the distance type. However, the analysis was done on five of six distances to maintain consistency with Experiment 2B which collected only five distances for each asymmetrical distance. The IVs were image type (real and VE) and display size (large and small). Significant level was initially set at 0.05. The results of the univariate tests for each of the distances were also reported

To investigate the effect of sporting background of the participants an ANCOVA analysis was conducted on each of the data set (vertical, horizontal and transverse) using the sport variable as a covariate. A Student t-test (in Microsoft Office Excel) was used to compare among the four experimental conditions.

In the next subsections, the results for each of the asymmetrical distances are first presented. This was followed by a comparison among these distances, the results of post-test questionnaire and examination of the effect of participants' sport background on distance estimate task.

#### 6.1.4.2 Vertical distance



**Figure 6-7 Vertical distance: Mean percentage of estimate for the real and VE on a large and small display**

Figure 6-7 shows the means of the estimated distance for the real and VE images on a large and a small display. It was noted from the figure that distances were generally underestimated. Estimates tended to be more accurate on a small display than on a large display for both image types.

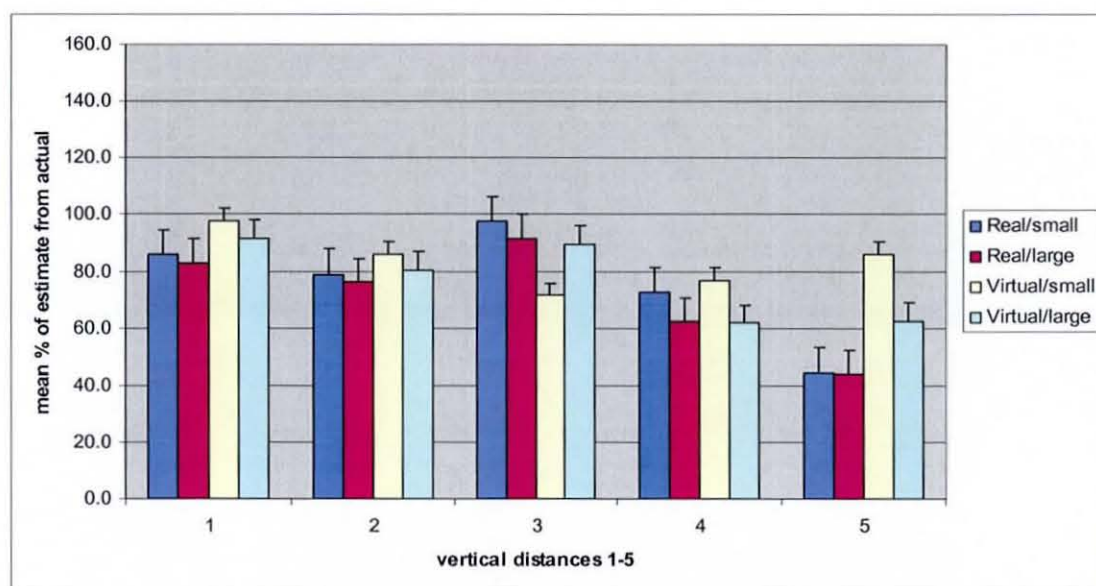
The results of a two-way between-group MANOVA revealed a violation of equality of covariance. However, the effect of this violation was unclear and Field (2000) suggested this test is highly unstable and the Hotelling's and Pillai's Trace statistics can be assumed to be robust. Walker (1998) stated that Pillai's Trace statistics is the most robust when assumption was not met (such as covariance not homogeneous) and it is particularly useful when the sample size was small and the cell sizes were unequal. As this was the case in our data, the results from the Pillai's Trace statistics were used. From the multivariate analysis, there was a statistically significant difference between the real and the VE on the combined five vertical distance estimations:  $F(5,24)=4.805, p=.003$ ; Pillai's Trace=.500, partial eta squared=.500, observed power=.942. The magnitude of the effect size was considered large. 50.0% (partial eta squared multiply 100) of the variance (effect + error) in distance estimation was explained by the image type. Generally, estimates tended to be more accurate on a VE image than on a real image.

The main effect of display size [ $F(5,24)=1.626, p=.286$ , Pillai's Trace=.253, partial eta squared=.253, observed power=.469], however, did not reach statistical difference, that is the difference between estimations on a large and a small display was considered small. The magnitude of effect size [25.3%] was considered large, implying the variance explained by display size was considered large. On average, estimates were more accurate on a small display than on a large display for both the real and the VE images (Figure 6-7).

No interaction effect between the image and display size was revealed:  $F(5,24)=1.474, p=.235$ , Pillai's Trace=.235, partial eta squared=.235, observed power=.427.

The univariate test on individual distances however revealed no main effect except for distance 5 (see Figure 6-8, distance 5 refers to the 5<sup>th</sup> vertical distance). With the exception of distance 3, it was shown generally that on a large display the VE image participants tended to perform better than the real image participants. Similarly, on the small display, estimates by the VE participants were better than the real image participants (except for distance 3). On average, with the exception of distance 3, estimations were more accurate on a small display compared to a large display for both the real and VE (Figure 6-8).





**Figure 6-8 Mean percentage of estimated distances for all the five vertical distances among all experimental conditions**

The results of the t-test comparisons among the experimental conditions are shown in Table 6-2. No significant difference was reported on any combinations of the comparisons. These indicated that the differences among the experimental conditions were small.

**Table 6-2 Results of comparisons among experimental conditions using Students' t-test**

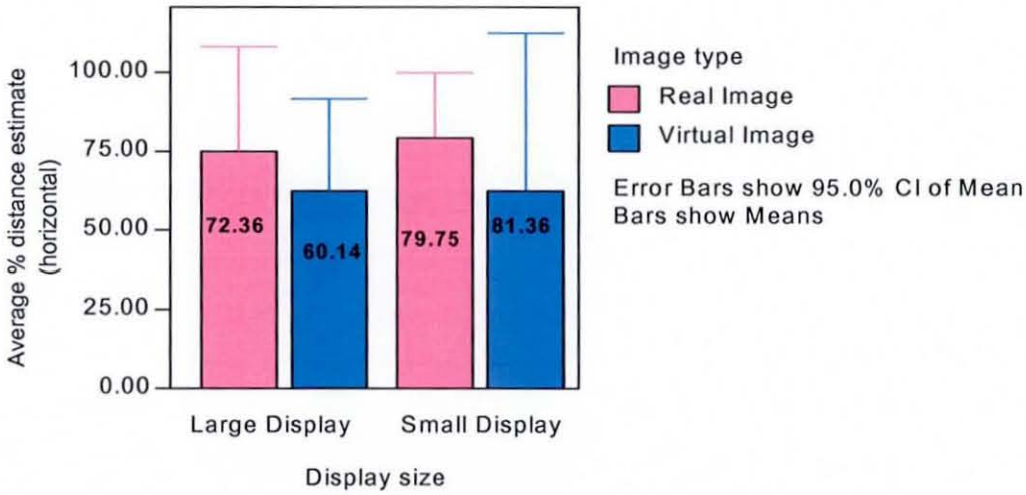
Conditions compared	Student T-test values*
	vertical
Real/Small vs Real/Large	0.4566
Virtual/Small vs Virtual/large	0.3601
Real/Small vs Virtual/Small	0.1895
Real/Large vs Virtual/large	0.4136

\*Two tailed distribution and assume unequal variance

#### 6.1.4.3 Horizontal distance

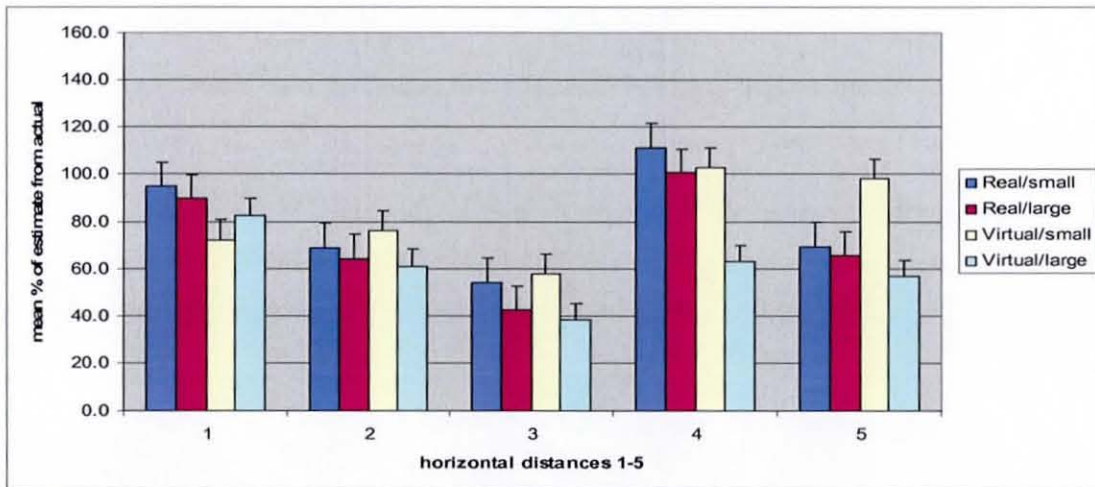
Figure 6-9 shows that horizontal distances were generally underestimated. From Figure 6-9, it can be seen that participants for both the real and the VE performed better on a small display than on a large display. It was indicated also that overall on a large display, the real image participants performed better than the VE image participants. In contrast, on the small display, participants on a VE image performed slightly better than on a real image. The results of the MANOVA analysis revealed that the covariances were not equal. As recommended earlier, the more robust Pillai's Trace statistics was reported. There was a main effect of image [ $F(5,24)=2.830$ ,  $p=.038$ , Pillai's Trace=.371, partial eta squared=.371, observed power=.740] which indicated that there was a significant difference on the horizontal distance estimation between the real and the VE on the combined five distances. The real image participants tended to be more accurate than the VE image participants on a large display but

on a small display estimates on a VE image were slightly better than the estimates on a real image.



**Figure 6-9 Horizontal distance: Mean percentage of estimate for the real and VE on a large and small display**

No main effect of the display size was reported [ $F(5,24)=.683, p=.641$ , Pillai's Trace=.173, partial eta squared=.125, observed power=.205]. It was indicated that on average, distances were underestimated more on a large display compared to a small display. The interaction effect also did not reach significant level [ $F(5,24)=.914, p=.488$ , Pillai's Trace=.160, partial eta squared=.160, observed power=.269].



**Figure 6-10 Mean percentage of estimated distances for all the five horizontal distances among all experimental conditions.**

Further examination of the univariate tests results however indicated that no significant effect for each of the five distances (Figure 6-10). For the real image, distance estimates on a small display tended to be more accurate than on a large display. Similarly for the VE image



(except for distance 1), performance was better on a small than on a large display. On a large display, estimates by real image participants were better compared to the VE image participants but on a small display the opposite was true, that is, estimates on a VE image were more accurate compared to estimates on the real image.

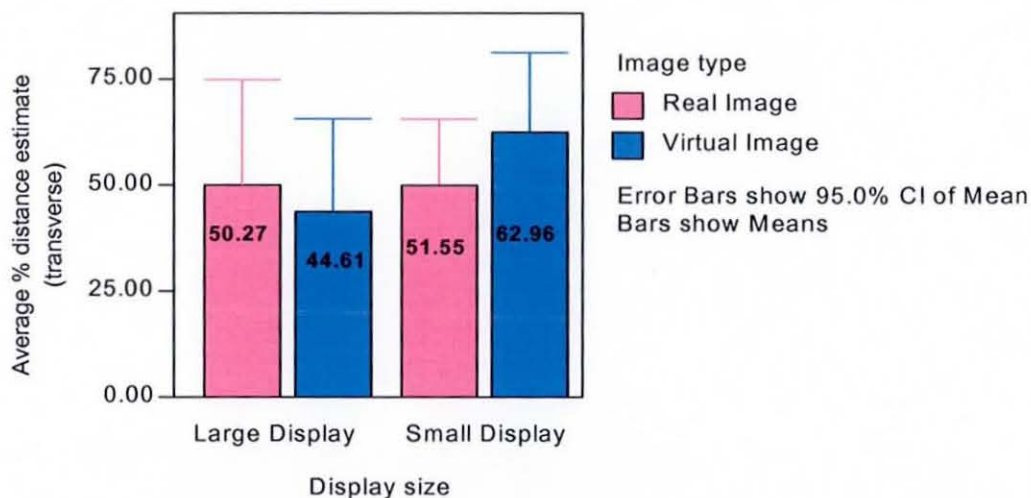
In Table 6-3, the results of comparisons among the experimental conditions using several t-tests revealed that only the VE image pair of comparisons reached significant difference. It was shown that there was a significant difference between the VE/large and the VE/small conditions. No other significant difference was reported.

**Table 6-3 Results of comparisons among experimental conditions using Students' t-test**

Conditions compared	Student T-test values*
	horizontal
Real/Small vs Real/Large	0.4788
Virtual/Small vs Virtual/large	0.0215
Real/Small vs Virtual/Small	0.8606
Real/Large vs Virtual/large	0.238

\*Two tailed distribution and assume unequal variance

#### 6.1.4.4 Transverse distance



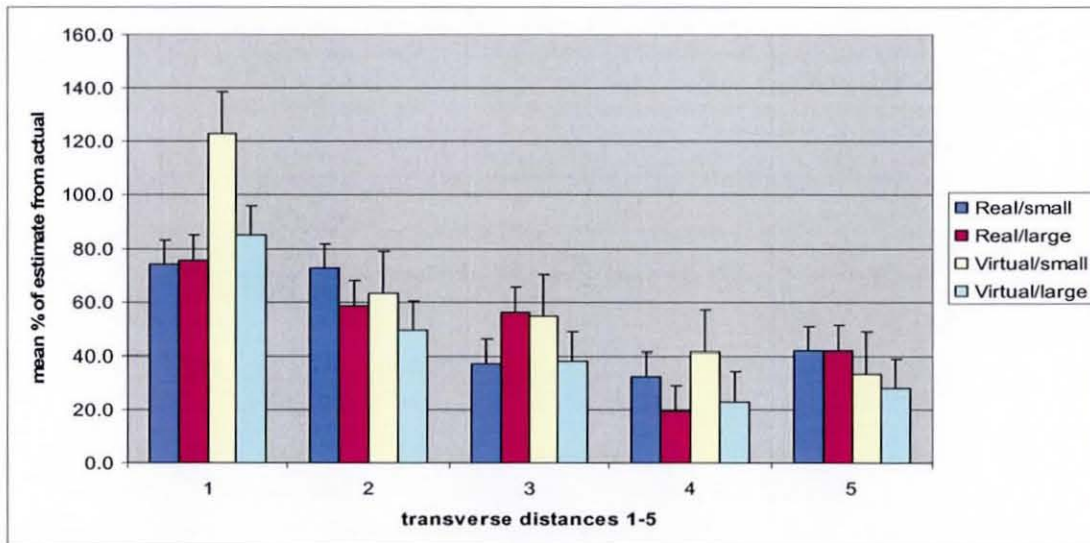
**Figure 6-11 Transverse distance: Mean percentage of estimate for the real and VE on a large and small display**

Figure 6-11 depicts the mean percentage of estimate for transverse distance in the four experimental conditions. From the figure, it was shown that distances were generally underestimated. Similar to vertical and horizontal analysis, the results of MANOVA analysis indicate the covariances were not similar. As such the Pillai's Trace was reported. The results of the analysis showed that there was a significant effect of image type [ $F(5,24)=4.110$ ,  $p=.008$ ; Pillai's trace=.461, partial eta squared=.461, observed power=.898] on the combined



five transverse distance estimations. The effect of display size however did not reach significant level: [ $F(5,24)=2.003$ ,  $p=.115$ ; Pillai's trace=.294, partial eta squared=.294, observed power=.566]. Examining the effect size for both the image and the display size, the magnitude of effect size was considered large. This indicates that the variances explained by the IV were large. Overall it was shown that both the real and VE image participants performed better on a small display compared to a large display, though for the real image, estimates on a small display were only slightly better than on a large display. Similar to the horizontal distance estimates, the real image participants performed better than the VE image participants on a large display but on a small display the VE image participants performed better than the real image participants. However, this interaction effect did not reach significant level [ $F(5,24)=1.351$ ,  $p=.278$ ; Pillai's trace=.220, partial eta squared=.220, observed power=.393].

A closer examination of the univariate test results revealed no significant main effect or interaction for each of the five distances (except distance 5, main effect of display at 5% and distance 1 with main effect of image at 10%). With the exception of distance 1, all distances were generally underestimated (Figure 6-12). For the VE image, more accurate estimates were found on a small display compared to a large display (except for distance 1). But for the real image, this was only true for distance 2 and 4. For distance 3, performance on a large display was better than on a small display. For distance 1 and 5, there was only a slight difference in estimates between the large and the small display.



**Figure 6-12 Mean percentage of estimated distances for all the five transverse distances among all experimental conditions**

The results of comparisons among the experimental conditions for the transverse distances were similar to the horizontal distances (Table 6-4). The only pair of comparison to reach statistical significance was for the VE image on a large and a small display. Other combinations were not statistically significant.

**Table 6-4 Results of comparisons among experimental conditions using Students' t-test**

Conditions compared	Student T-test values*
	transverse
Real/Small vs Real/Large	0.8712
Virtual/Small vs Virtual/Large	0.0491
Real/Small vs Virtual/Small	0.1646
Real/Large vs Virtual/Large	0.5289

\*Two tailed distribution and assume unequal variance

6.1.4.5 Comparisons among asymmetrical distances

From Figure 6-13, it was observed that the transverse distance consistently gave the worst estimates under all conditions when compared to the vertical and horizontal distance. For all experimental conditions, the vertical distance tended to be estimated more accurately when compared to the transverse distance. A series of t-test conducted statistically confirmed this difference (see Table 6-5). It was noted that t-test comparisons of horizontal and transverse distance yielded significant differences on all conditions except for the VE /small conditions. However, both the vertical and horizontal distances did not show any consistent relationship. The performance of the participants appeared to be dependent on the distance to be estimated (that is, whether it was distance 1, 2, 3, 4 or 5). The t-test results, however, revealed that there was a significance difference between the vertical and the horizontal distance under the VE /large condition only.

**Table 6-5 Results of t-test values for comparison among distance types under the four experimental conditions**

Condition	vertical-horizontal	vertical-transverse	horizontal-transverse
Real/Small	0.588733	0.000091	0.000361
Real/Large	0.907872	0.009453	0.040489
Virtual /Small	0.791903	0.013055	0.059182
Virtual/Large	0.040394	0.000158	0.074353



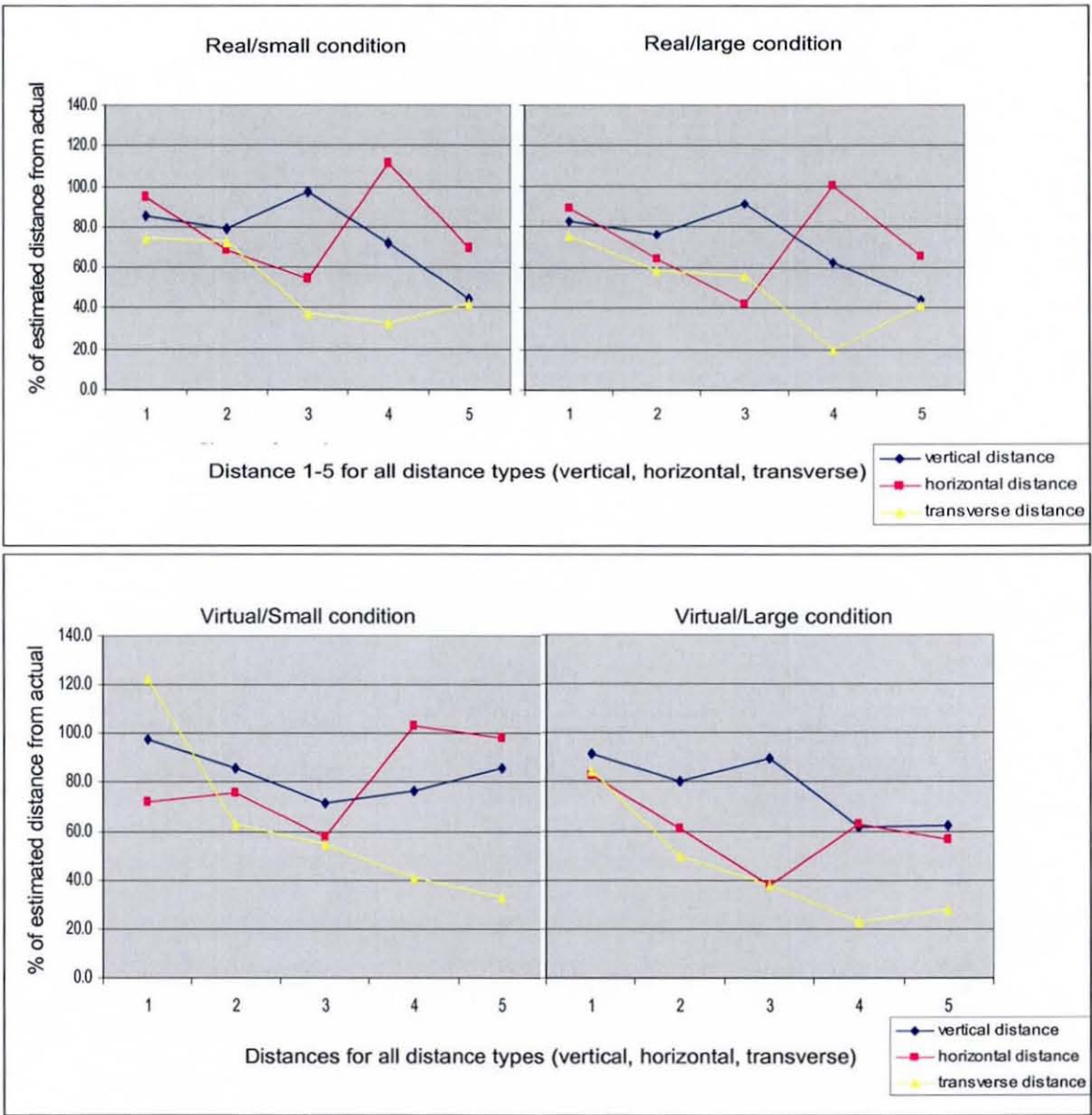


Figure 6-13 Comparison among distance types under the four experimental conditions

6.1.4.6 Post-test questionnaires result

Participants were asked to rate their distance estimation performance on the scale of 1 to 7 (7 represents very accurate). The average response was 4. Three participants felt confident of their estimation (rating =6) while four participants were very uncertain of their estimation (rate = 2). More than half of the participants found the transverse distance as most difficult to estimate (33 out of 40) and the vertical distance as most easy to estimate (31 out of 40). A survey of participants' sport background revealed that only nine did not play any sports; others play at least one of the following sports: tennis, badminton, squash, netball, hockey,

cricket, and cycling. However, none of the participants were professional players. All participants reported played the indicated sports as part of their leisure activities.

Only three participants did not find viewing the movie had assisted them in their estimation, the rest found it allowed them to make a better estimation especially for distant objects. Generally, most participants reported using familiar objects in the scene (such as trees, lamppost, and goal posts) as a basis for their estimation. Others used methods such as using their own height as a guide, imagined themselves walking in the scene, or calculated distance based on the speed of the camera moving through the scene.

### 6.1.4.7 The effect of sport variance on distance estimates

As mentioned earlier, the participants' sports background might have exerted an influence on their distance estimates. This data was collected as a categorical variable where participants indicated whether they played any kind of sport or not. The results of an ANCOVA analysis which investigated the influence of the sport variable on participants' distance estimates was summarized in Table 6-6.

**Table 6-6 Summary of result from ANCOVA analysis using sport variable as covariate for all distance type (vertical, horizontal and transverse)**

Distance type	Effect	F(5,23)	p value	Partial eta squared	Observed power
Vertical	Sport	0.117	0.987	0.025	0.071
	Image	4.126	0.008	0.473	0.896
	Display	1.594	0.202	0.257	0.456
	Image*Display	1.255	0.317	0.214	0.362
Horizontal	Sport	0.405	0.84	0.081	0.134
	Image	2.649	0.049	0.365	0.702
	Display	0.645	0.668	0.123	0.194
	Image*Display	0.945	0.454	0.175	0.284
Transverse	Sport	0.793	0.565	0.147	0.234
	Image	4.442	0.006	0.491	0.919
	Display	2.047	0.109	0.308	0.572
	Image*Display	1.696	0.176	0.269	0.483

Similar to the results of MANOVA analysis, this analysis revealed that the effect of image type was significant for all asymmetrical distances but the effect of display size and the interaction effect were not significant. The similar results from both analyses suggested the influence of sport variable was minimal in these data sets. This indicated that the effect of the sport variable was highly insignificant for all asymmetrical distances. However, the observed power of the test was considered low. As such, the non-significant difference may suggest insufficient power of the test to detect a difference (Type II error: A belief that there is no difference when actually there really is a difference between groups). Thus, careful interpretation of this non-significant result was required. Other explanatory information such

as types of sports (tennis, football, badminton, etc) and their ability (such as an amateur, professional or not) were also collected. A review of participants' sport ability revealed that all participants' played sports as part of their leisure activity; none played sports as a professional or an amateur. As such, their similarity in sporting background level might account for the insignificant influence of the sport variable. No further analysis was performed using the sport variable since the results yield no significant difference between groups.

### 6.1.5 Analysis

When users were allowed to view a movie of a scene prior to making their distance estimate, Experiment 2A results showed that there was a significant difference between participants' performances in the real and VE. Regardless, of the type display used, there was a main effect of image for all asymmetrical distances. On average, the VE image participants tended to perform better than the real image on both the large and small display for the vertical distance. For both the horizontal and transverse distance, generally the better performance of the VE participants over the real image participants was reflected on the small display only (though, this differs only slightly for the horizontal distance). However, on a large display, the real image participants tended to perform better than the VE image participants.

The effect of display size on distance estimation tasks revealed no significant difference for all asymmetrical distances. Numerical comparison of the means of percentage of estimates however revealed that distance estimation on a small display was better than on a large display for all asymmetrical distances (with the exception of the real image condition in transverse distance, this difference was very small).

When the individual distances were examined, no significant effect of image was revealed (except for distance 5 in vertical distances). Similarly, there was no significant effect of display (except for distance 5 in transverse distances) or interaction effects were reported. The range of estimates for the vertical distance was from 44% to 97%. Similarly, for horizontal distance (with the exception of distance 4) and for transverse distance (with the exception of distance 1) distances were underestimated in both image types. The range of estimates for the horizontal distance was 38% to 94% and the range of estimates for the transverse distance was 23% to 84%. For each asymmetrical distance there were five distances to be estimated and these distances varied in lengths and were located at different positions in the environment. These differences might account for the great differences in estimates accuracy between each distance.



A direct comparison among the four experimental conditions yielded no significant difference for the vertical distance. For the horizontal and transverse distances, the difference between the estimates for the VE image presented on a large and a small display reached statistical significance.

For all viewing conditions, the vertical distance and horizontal were estimated more accurately when compared to the transverse distance. The vertical distance estimate was statistically more accurate than the transverse distance; however, it did not differ significantly from the horizontal distance. It was shown that the transverse distance was statistically less accurate than the vertical and horizontal distances. The results of the post-test questionnaire showed similar observations: the vertical distance was easy to estimate and the transverse distance was most difficult to estimate. Experiment 2A results also showed that participants sporting background (that is whether they play sport or not) did not influence their distance estimate.

## 6.1.6 Discussion

### 6.1.6.1 Image types

Consistent with the results of previous findings (Henry and Furness 1993, Lampton, Bliss et al. 1994, Witmer and Sadowski 1998, Witmer and Kline 1998), distances were generally underestimated in the real and VE images. Several previous investigations (Henry and Furness 1993, Witmer and Kline 1998, Witmer and Sadowski 1998, Lampton, McDonald et al. 1995) showed that there was a significant difference between the real and VE conditions while others (Willemson and Gooch 2002, Waller 1999, Yoon, Byun et al. 2000) reported that these differences were small. Corroborating the findings of the former investigations, Experiment 2A results revealed that there was a significant difference between the real and VE image for asymmetrical distances. However, the previous investigations mentioned earlier reported better performance by the real world participants compared to VE participants while Experiment 2A results vary depending on which asymmetrical distance was investigated:

- For horizontal and transverse distance estimates:
  - on a large display -similar results to these previous investigations; that is, the real image participants performed better than the VE image participants.
  - on a small display -the VE image participants yielded more accurate estimates compared to the real image participants.
- For the vertical distances,
  - on both display sizes -VE participants performed better than the real image participants

The better performance on VE image provide supports the investigations done by Yanagisawa and Akahori (1999) who reported better performance on a VE image (computer-generated VE) compared to a real image (photographed-based VE) but the task investigated by these researchers was spatial representation of the visited scene. Yanagisawa and Akahori (1999) suggested that the photographed-based VE contains more detailed information compared to the computer-generated VE. This detailed information imposes more cognitive load on the real image participants thus degrading their spatial representation task performance. Thus, the better performance of the VE participants in the Experiment 2A might be due to the less cognitive load imposed on the participants on a small display.

Another possible explanation is the image quality. On a large display, the image resolution for both image types appeared to be reduced particularly for the VE image. But on a small display, the degradation in image resolution was less obvious. Studies have shown that the use of a high resolution image improves participant's performance on distance judgment task (Duh, Linh et al. 2002, Kline and Witmer 1996, Jää-Aro and Kjell Dahl 1997). Thus, image resolution factor might account for the better performance of the real image participants over the VE participants on a large display for the horizontal and transverse distances. It was noted however, for the vertical distance, image quality appeared to have less impact on distance estimates performance as the VE image participants tended to perform better than real image participants on both display sizes.

#### 6.1.6.2 Display sizes

While the effect of display size was not statistically significant, Experiment 2A results showed that numerical comparisons of the mean percentage of estimates scores revealed that generally the small display participants tended to yield better estimations compared to the large display participants for all asymmetrical distances. In contrast, the results of past studies (Patrick, Cosgrove et al. 2000, Tan, Gergle et al. 2003, Dixon and Proffitt 2002, Czerwinski, Tan et al. 2002, Tan, Czerwinski et al. 2003) revealed that performance on a large display was significantly better than on a small display. Patrick, Cosgrove et al. (2000) suggested that the large image size might induce realistic experience in their participants thus giving better judgment of relative distances. Tan and colleagues (2003) suggested that the better performance of their large display participants over the small display participants was due to the large display affording a greater sense of presence. They further suggested that users were more effective when they felt more presence in the VE. However, the large images viewed in Experiment 2A failed to induce similar experiences. The creation of the VE model in Experiment 2A was based on careful measurements of the real world scene. Textures from the

real world scene pictures were used for the objects in the VE model. To further reduce the variances between both movies, the process of producing both movies were made similar (see Figure 6.4), the file formats were made the same and both movies were run from the same computer. However, in hindsight, there was a clear difference between the movies when presented on a small display and a large display especially for the VE. When viewed on a large display, the image appeared less clear compared to when presented on a small display. The process of transferring the simulated VE movie from the SGI machine to a PC via recording on a VHS tape had resulted in a very low resolution image: 200 lines of resolution (video format). As mentioned earlier, studies have shown that high resolution image improve performance on distance judgment. Thus, image resolution factor might have influenced Experiment 2A results particularly for the VE image as it appeared less clear when projected on a large display. But on a small display, this reduction in image quality was less obvious. Thus, this might explain the better performance of the VE image participants over the real image participants on a small display for all asymmetrical distances. The reduced image quality might have accounted for the poor distance estimate performance on a large display compared to a small display for all asymmetrical distances for the VE image participants. For consistency and to avoid reduced image quality, the next study (Experiment 2B) employed the original simulated PERFLY movie of the VE.

### 6.1.6.3 Individual distances

An examination of individual distance estimates in all asymmetrical distances revealed that not all distances yield similar effects or direction of effects:

- For most distances, no main effect of image and display or interaction effect was shown but for some distances there was a significant main effect.
  - For example, there was a main effect of image for vertical distance number 5 and there was a main effect of display for transverse distance number 3.
- For all distance types, not all distances yielded better performance on the VE image compared to the real image
- Not all distances yielded better performance on a small display compared to a large display.

The types of stimulus employed in this study might be partially responsible for these differences. In an attempt to expand the limited list of stimulus used by past studies, Experiment 2A employed objects that were present in the scene as stimuli. These included trees, hedges, signpost, lampposts, roads, bins and goalposts (see Figure 6-1 and 6-2). Some objects may be easy to estimate (such as the roads and goalposts) while others may be difficult to estimate such as trees and hedges (not all trees are of the same or of a particular height). As such the variations in the stimulus types might explain Experiment 2A results of large variability among distances estimates.

The position of the objects in the scene may have similar effects on this large variation among distances. In Experiment 2A, participants' view was limited to forward movement along a single line only, that is, a straight line of movement from one corner of the football field to the opposite corner of the football field. Objects that were located to the far right or the far left of this line of movement may be difficult to estimate compared to objects located along or near this line of movement. As objects located in the peripheral visual field are viewed with low acuity compared to objects located in the central of the visual field, thus objects' positions in the scene may offer another explanation for Experiment 2A results' differences.

Additionally, differences in the distances might also contribute to the differences in estimation accuracy among distances whereby shorter distances were often estimated accurately compared to longer distances.

#### 6.1.6.4 Comparison among asymmetrical distances

A direct comparison among asymmetrical distances, showed that these results were in line with the findings of Experiment 1 (on static images); users yielded more accurate results when estimating the vertical distances compared to the horizontal and transverse distances. However, the difference between the vertical and horizontal distance was not significantly different on both display sizes and both image types. Consistently, the transverse distance yielded the worst performance. This was further supported by the post-test questionnaire results. Participants' comments revealed that the vertical distance was the easiest to estimate while transverse distance was the most difficult.

The results from Experiment 2A are consistent with the findings of Henry & Furness (1993), who found subjects' performance were almost veridical on vertical distance compared to horizontal distance. This result was expected, as people are generally more familiar with their own height as a scale to other objects. This was further supported by our post-test questionnaire results which revealed that participants did actually use their heights to base their estimations from. Very accurate performance in Henry and Furness' (1993) study might also be attributed to the difference in the type of stimulus used. Their participants estimated height of rooms in a museum while our participants estimated vertical distances of objects in an outdoor setting. Interior spaces usually have standard heights and the fact that their participants came from the architectural background may have accounted for the almost perfect estimations in their study.

Experiment 2A showed that the transverse distances gave the worst performance. Similar findings by Loomis, Da Silva et al. (1996) showed that more estimation errors were made on the transverse distance than on the horizontal plane and this error was magnified when this distance was increased. For transverse distance, our participants reported less than half of the actual distance. This inaccuracy was more pronounced for larger distances. A similar observation by Witmer & Kline (1998) was reported for egocentric distance estimation. They found distance perception in VE to be less than half (47%) of the actual distance.

#### 6.1.6.5 Influence of sport background factor

It was suggested that participants' sporting background might influence their distance estimates (anonymous paper reviewer in Awang-Rambli and Kalawsky 2002, Coren, Ward et al. 1999). However, the results of the analysis showed there was no significant effect of sport background on the current data set. Examination of participants' sport ability data revealed that all participants played sports as part of their leisure activity; none of the participants were professional players. As such, the non-significant effect of sport background on distance might be due to the similar sport background. It was expected however that a professional sportsman to perform better than the non-professionals (such as those who play sport as a leisure activity) as their distance judgement would be fairly accurate due to frequent training.

#### 6.1.7 Conclusion

Generally, distances were underestimated for all asymmetrical distances in both the real and VE images. Participants' performances on the distance estimation task differed significantly between image types for all asymmetrical distances (vertical, horizontal and transverse).

For vertical distance, results showed that more accurate estimations were observed in the VE image compared to the real image on both the large and small displays. However, for the horizontal and transverse distances, better performance was noted for the VE participants on a small display but on a large display the real image participants' estimates were more accurate compared to VE image participants' estimates. The resultant poor quality of the VE image might have accounted for the poor estimations of the horizontal distance and transverse distances on a large display.

For the vertical distances however, distance estimates seemed not to be influenced by the poor VE image quality. Surprisingly, the VE participants tended to perform better than the real image participants on both display sizes.



While not significantly different, overall, distances perceived in both image types presented on a small display produced less estimation error compared to when viewed on a larger display. This result does not agree with other findings that reported better spatial perception on a large display. A more likely explanation for these results was the low image resolution used in our study, particularly for the VE image whereby the viewed image was less clear, especially when presented on a large display. As such, in the next study (Experiment 2B), the original, high resolution simulated PERFLY movie of the VE was employed.

For most individual distance estimates in all asymmetrical distances, the effect of image and display or interaction was not statistically significant (except for a few of the distances). Moreover, the individual distance estimates in all asymmetrical distances revealed that not all distances yielded similar effects or direction of effects. The use of different objects at various positions in the scene might have explained the differences in estimates among distances in each asymmetrical distance.

On average, the vertical distances were perceived more accurately when compared to the horizontal and transverse distances. Transverse distance was perceived less than half of the actual distance. More compression of distance estimates was observed for larger distances.

## 6.2 EXPERIMENT 2B: EFFECT OF IMAGE TYPE AND DISPLAY SIZE

### 6.2.1 Rationale

The results of Experiment 2A revealed that there was a main effect image on distance estimation, that is, there was a significant difference in users' performance between the real and the VE image. On average, for the horizontal and transverse distances, the real image participants tended to perform better than the VE image on a large display but the reverse was true for a small display. However, for the vertical distance, the VE image participants outperformed the real image participants on both display sizes. In evaluating these results, it was noted that the image quality of the VE movie was of a very low resolution. While this was not noticeable on a small display, on a large display the image was not sharp and clear. The image resolution must have been degraded during the process of transferring the VE movie from the SGI machine to the PC via a video tape, which has 200 lines of resolution. The original simulated PERFLY movie was not used in Experiment 2A because it was intended to maintain consistency for both image types in the followings:

- similar process of image acquisition (see Figure 6-4)
- similar use of a viewer software.

- Lastly, the main reason was that the Windows Media Player which was used to view the real movie provides control buttons [pause, stop, play buttons] to allow the user to control the movie, an option which was not available in PERFLY software.

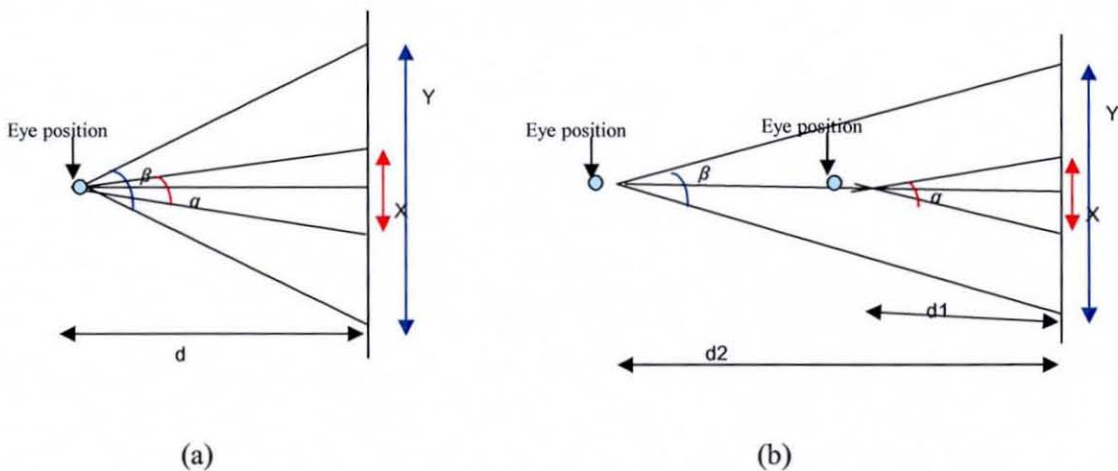
As such, it was initially decided to convert the simulated PERFLY movie to an AVI file format for viewing on a PC. However, as mentioned earlier, it was not possible to record the simulated PERFLY movie of the VE directly to an AVI file format. Thus, the available option was to record the simulation of the movie on a video tape and then transfer this to the PC to AVI file format.

As discussed earlier, the lower performance of the users in the VE movie (especially on a large display) might be due this poor image resolution. Thus, to determine whether the poor performance of VE users, (at least for the horizontal and transverse distance) was attributed to the poor image resolution of the VE, it was necessary to use the original, high resolution simulated PERFLY movie of the VE for Experiment 2B. Additionally, informal observation during Experiment 2A trials revealed that most of the participants tended to just watch movie and not use the control buttons even though it was instructed prior to the viewing that they can control the flow of the movie. This might be due to the slow pace of the movie (walking pace of 1.08m/s). Thus, for Experiment 2B, the option of controlling the movie was not included. Thus, using PERFLY software to view the original simulated movie was made possible because the movie control option (which is not available on PERFLY software) was no longer needed.

While not significantly different, numerical comparisons of the mean percentage of estimates scores indicated that participants tended to perform better on a small display for all asymmetrical distances. Again, a more likely explanation was the image quality of the movies. It was observed that when projected on a large display, the image tended to be less clear and this was especially true for the VE movie.

For the real movie, the difference was less obvious. Thus, for Experiment 2B, the original, high resolution simulated PERFLY movie was used instead of the converted AVI file format. It was not possible, however, to recapture the movie of the real world scene again using a higher resolution camera because the original site of the scene was no longer available. The practice football field is now the site of a new building. The remaining option was to use the same video movie as used in Experiment 2A. As the resolution of the VE movie was set to a higher resolution [1280x1024] than the real movie and if image resolution significantly affects distance estimation accuracy, we therefore would expect a main effect of image type favouring the VE image in the Experiment 2B results.

The setup of Experiment 2B employed the setup of Experiment 1B (Chapter 5). Similar to Experiment 1B, this setup allowed us to investigate the effect of the viewing distance and physiological cues of accommodation and vergence. In Experiment 2A (Figure 6-14 (b)) the FOV was equated on both display sizes but in Experiment 2B setup (Figure 6-14 (a)) the FOVs for the large and the small display were different. Accordingly, the retinal image size for the users would be the same in Experiment 2A and different in Experiment 2B for both display sizes. In Experiment 2A, equating the FOV for each display size revealed no significant difference between the large and the small display. A comparison of means however, revealed that the small display yielded better performance than on a large display. As suggested earlier, the low image resolution might have accounted for the poor performance of the large display participants. Thus, if the low resolution of the image was to result in the low distance estimate performance on a large display in Experiment 2A, using a high resolution image in Experiment 2B we would expect the opposite result, that is, estimates on a large display would be better than on a small display.



**Figure 6-14 Experiment setup for Experiment 2B (a) and Experiment 2A (b).  $X$  and  $Y$  represent the physical image size for small and large display respectively.  $\alpha$  and  $\beta$  are the corresponding FOV for both display size**

In Experiment 2A, the FOV was fixed but the viewing distance was varied. Thus, the physiological cues acting at these different distances were different. However, in Experiment 2B setup, the FOV was varied but the viewing distance of the user was fixed for both display sizes. As such, the retinal image size for both display sizes would be different, but the physiological cues acting at this fixed distance would be the same. In Experiment 2A, when the FOVs of both display sizes were equated, this produced similar image size on the observers' retina. While the effect of display size was not significant (it was suggested earlier due to the low resolution of the image when projected on a large display), however, on average, the participants' distance estimate performance on a small display was better than on a large display, even though the retinal image was similar. This implies that discounting the effect of accommodation and vergence cues, the difference in performance between the small

and the large display was influenced more by the physical image size rather than the retinal image size. Additionally, as mentioned earlier the retinal image size was a weak cue to distance (Beall, Loomis et al. 1995), thus, if the physiological cues and viewing distance were to influence the distance estimated, we would expect no significant effect of display size. As discussed earlier it was expected that there would be a main effect of image favouring VE movie over real movie if the image resolution did contribute an effect in distance estimation task.

## 6.2.2 Experiment aim and hypothesis

The aim of Experiment 2B was to investigate the effect of display size on asymmetrical distance perception in the real and VE. The following hypotheses were explored in this investigation.

The main hypotheses were:

H1: There is no effect of image type (real and VE image) on asymmetrical distance perception.

H2: There is no effect of display size (small and large size) on asymmetrical distance perception.

The secondary hypotheses were:

H3: There is no effect of viewing distance on asymmetrical distance perception

H4: There is no effect of physiological cues on asymmetrical distance perception.

H5: Image resolution has no effect on asymmetrical distance perception.

## 6.2.3 Methodology

### 6.2.3.1 Participants

Four groups of 10 participants each were used for the study. Forty volunteers (20 males) comprising of staff and students participated in the study. The ages of the participants ranged from 18 to 44 years with an average of 27.9. All participants have normal or corrected-to-normal vision.

### 6.2.3.2 Materials/Apparatus

#### The real and VE images

This study used the same real movie for the real environment condition as Experiment 2A. The real movie was viewed using the Windows media player from a PC. However, for the VE movie, the original MultiGen II Pro Flight file model was used (see Figure 6-15). The image



resolution was 1280 x 1024. The model was viewed using the SGI Performer PERFLY viewer software running on a SGI computer.



**Figure 6-15 The original high resolution VE movie**

#### 6.2.3.3 Display apparatus and room setting

The images were displayed on a rear-projected display screen. Two display area sizes were used: large display (156 x 208 cm), small display (39 x 52cm). These sizes corresponded to the approximate largest and smallest possible display area at the current room setup. The distance from the screen was fixed at 100cm. Initially, 60cm was chosen for comfortable viewing especially for the small display. However, for the large display, at this distance viewers might fail to notice objects that were located especially at the lower part of the image when viewing the movie. The FOV for the large display was approximately  $92^{\circ}$  and  $29^{\circ}$  for the small display. The resolution of display was set at 1028 x 768. Similar to Experiment 2A, a dark room setting was also employed here.

#### 6.2.3.4 Experiment setup, design and procedure

The experiment setup of and design of Experiment 2B was similar to Experiment 2A with one exception. The number of distances to estimate for each asymmetrical distance was reduced to five from the total of six. Based on observations and the results of Experiment 2A, some distances presented ambiguity and was difficult to see for some viewers. These distances were not included in Experiment 2B. Thus the total number of estimated distances was fifteen instead of sixteen. For consistency, these distances were also excluded from analysis in Experiment 2A.

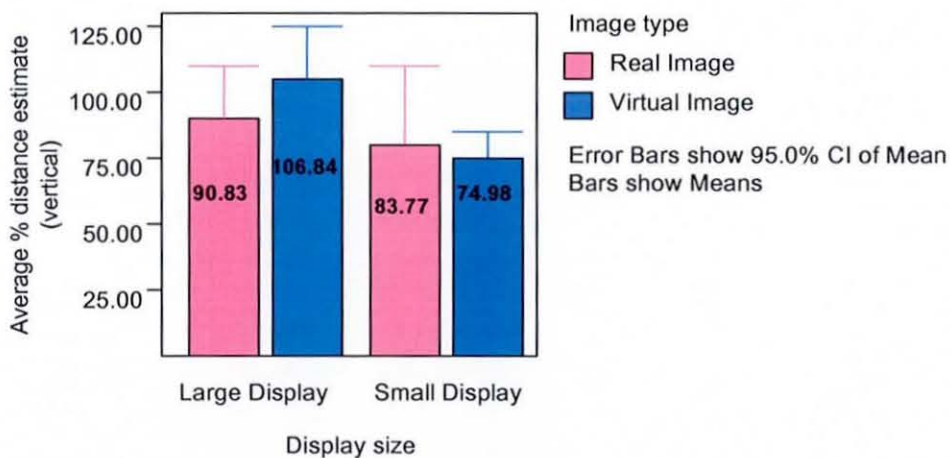


## 6.2.4 Results

### 6.2.4.1 Data preparation

In terms of data preparation, this section was similar to Experiment 2A. The results for vertical, horizontal and transverse were presented first, followed by comparisons among asymmetrical distances, post-test questionnaire results and effects of participants' sport background on distance estimates.

### 6.2.4.2 Vertical distance



**Figure 6-16 Vertical distance: Mean percentage of estimate for the real and VE on a large and small display**

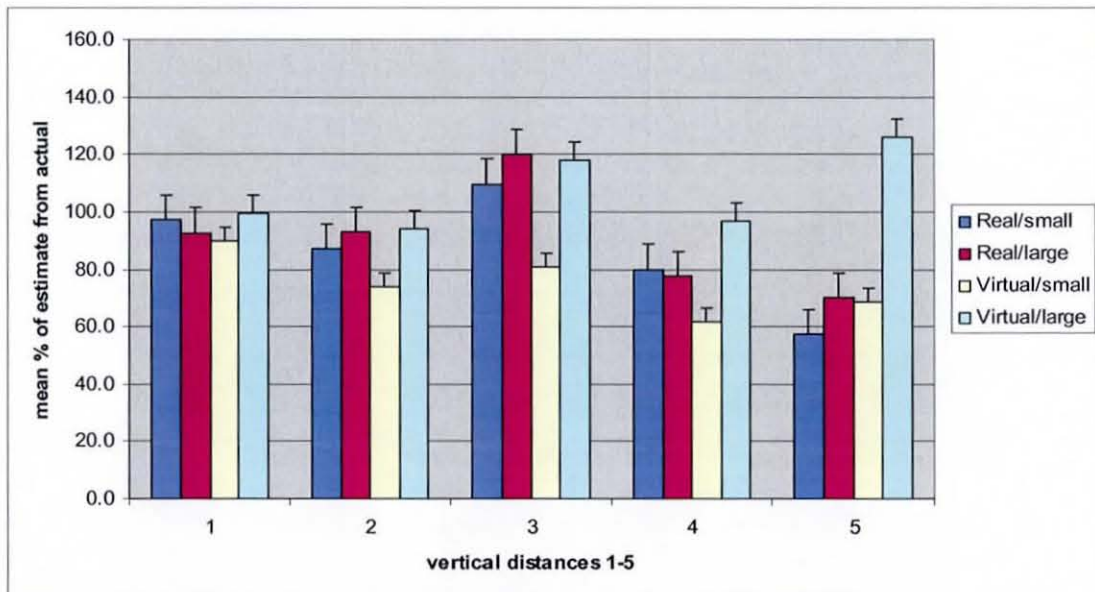
Figure 6-16 shows a comparison of mean percentage of distance estimates for the four experimental conditions. It can be inferred from the figure that

- for VE image, the distance estimate performance of the large display participants was better than the small display participants
- for the real image, estimates on a large display were better than on a small display.
- overall, VE/small condition yielded the lowest performance.
- for large display, distance estimates on the VE image were larger compared to estimates on the real image.
- on a small display, estimates on the real image were larger than on the VE image.

The results of MANOVA analysis showed that there was a statistically significant difference between the real and the VE for vertical distance estimations:  $F(5,25)=2.765, p=.040$ ; Pillai's Trace=.356, partial eta squared=.356, observed power = .732. This implied that regardless of display size, there was a significant difference for vertical distance estimates between the real and VE images.

The main effect for display size [ $F(5,25)=.123, p=.483$ , Pillai's Trace=.281, partial eta squared=.280, observed power=.555] and the interaction effect [ $F(5,25)=.826, p=.543$ , Pillai's Trace=.142, partial eta squared=.142, observed power=.247] did not reach statistical difference. On average, the estimates on a large display were more accurate compared to the estimates on a small display.

The results of the univariate tests for all the five vertical distances revealed no main effect of image and display (except for distance 5) or interaction. Figure 6-17 shows for the real image, estimates on a large display were not consistently better than on a small display. Distance 2, 3 and 5 show better estimates on a large display than on a small display but the reverse was true for distance 1 and 4. For the VE image, distance estimates on a large display were better than on a small display for all distances. Overall, the figure also indicates that on a large display the VE image yielded less error than the real image (except for distance 3) but on a small display the reverse was true (except for distance 5).



**Figure 6-17 Mean percentage of estimated distances for all the five vertical distances among all experimental conditions**

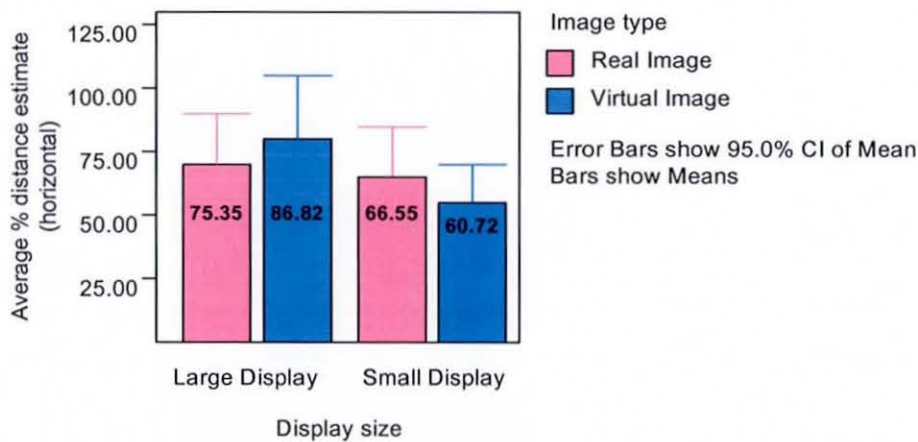
Table 6-7 shows the results of t-test comparisons between the experimental conditions. From the table it was shown that for the real image, the difference between the large and small display was not significant but for the VE image this difference reached significance level ( $p > .05$ ). On the small display both the real and VE images did not differ significantly but on the large display the difference was statistically significant.

**Table 6-7 Results of comparisons among experimental conditions using Students' t-test**

Conditions compared	Student T-test values*
Real/Small vs Real/Large	0.6171
Virtual/Small vs Virtual/Large	0.0003
Real/Small vs Virtual/Small	0.3265
Real/Large vs Virtual/Large	0.03

\*Two-tailed distribution and assume unequal variance

### 6.2.4.3 Horizontal distance



**Figure 6-18 Horizontal distance: Mean percentage of estimate for the real and VE on a large and small display**

Figure 6-18 shows the differences for horizontal distance estimates between the image types and between the display sizes. On average, the horizontal distance estimates on a large display was more accurate compared to on a small display. On a large display, generally the VE image yielded less estimation error when compared to the real image but the reverse was true on a small display.

The results of a two-way between-group MANOVA analysis on this dataset revealed that there was no main effect of image [ $F(5,25)=1.628, p=.189$ , Pillai's Trace=.246, partial eta squared=.246, observed power=.473] or display [ $F(5,25)=1.274, p=.306$ , Pillai's Trace=.203, partial eta squared=.203, observed power=.374] on horizontal distance estimation, that is there was no significant differences between the real and the VE and between the small and the large display for the horizontal distance estimations.

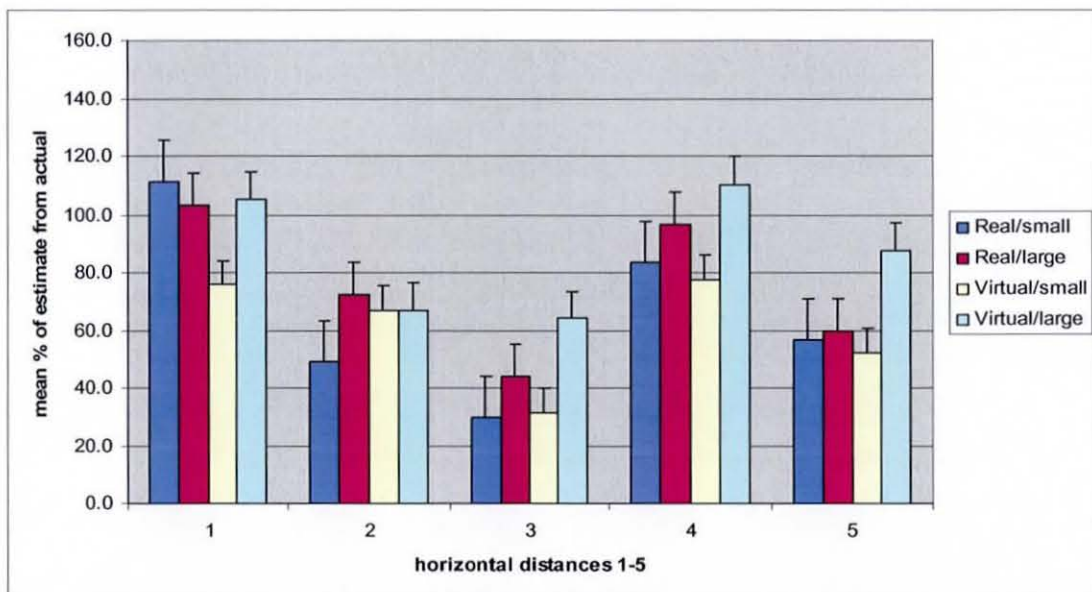
However, the interaction effect between the image type and display type was statistically significant:  $F(5,25)=1.787, p=.039$ , Pillai's Trace=.358, partial eta squared=.358, observed



power=.736). This indicates that both the image type and the display size have different effects on horizontal distance estimation tasks. For the VE image overall, the large display participants tended to perform better than the small display participants. Similarly for the real image, on average, estimates on a large display were better than on a small display.

The univariate test results for each horizontal distance revealed no main effect of image type, display size (except for distance 3) and interaction.

For the VE image, Figure 6-19 indicates that the VE /large participants tended to perform better than the VE /small participants (for distance 2, this difference was very small). For the real image, with the exception of distance 1, performance on the small display was better than on the large display. On the large display, the VE image participants yielded more accurate estimates compared to the real image participants (except for distance 2) but on the small display estimates on the real image were better than on the VE image (except for distance 2 and 3).



**Figure 6-19 Mean percentage of estimated distances for all the five horizontal distances among all experimental conditions**

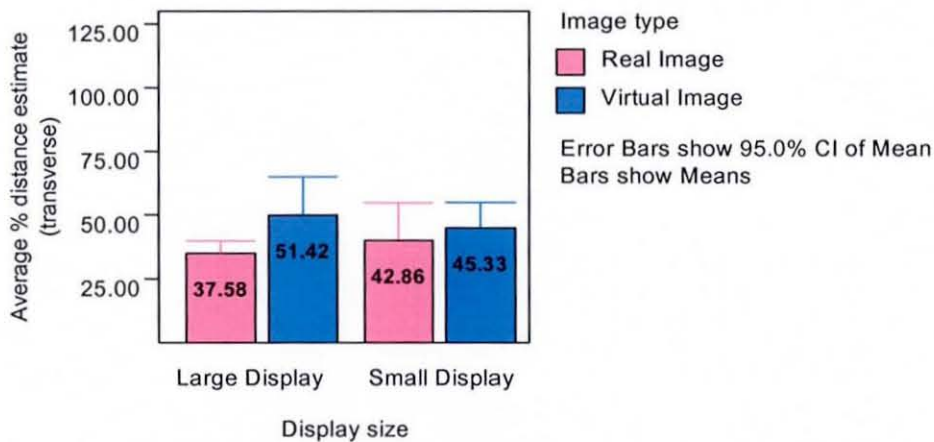
From Table 6-8, similar to the vertical distance results, it was indicated that for the real image the difference between the large and the small display was not significant but for the VE image this difference was statistically significant. On a large display, the difference between the real and the VE image was significant but this was not so on a small display.

**Table 6-8 Results of comparisons among experimental conditions using the Students' t-test**

Conditions compared	Student T-test values*
	horizontal
Real/Small vs Real/Large	0.88709
Virtual/Small vs Virtual/Large	0.00195
Real/Small vs Virtual/Small	0.4617
Real/Large vs Virtual/Large	0.0275

\*Two-tailed distribution and assume unequal variance

#### 6.2.4.4 Transverse distance



**Figure 6-20 Transverse distance: Mean percentage of estimate for real and VE on large and small display**

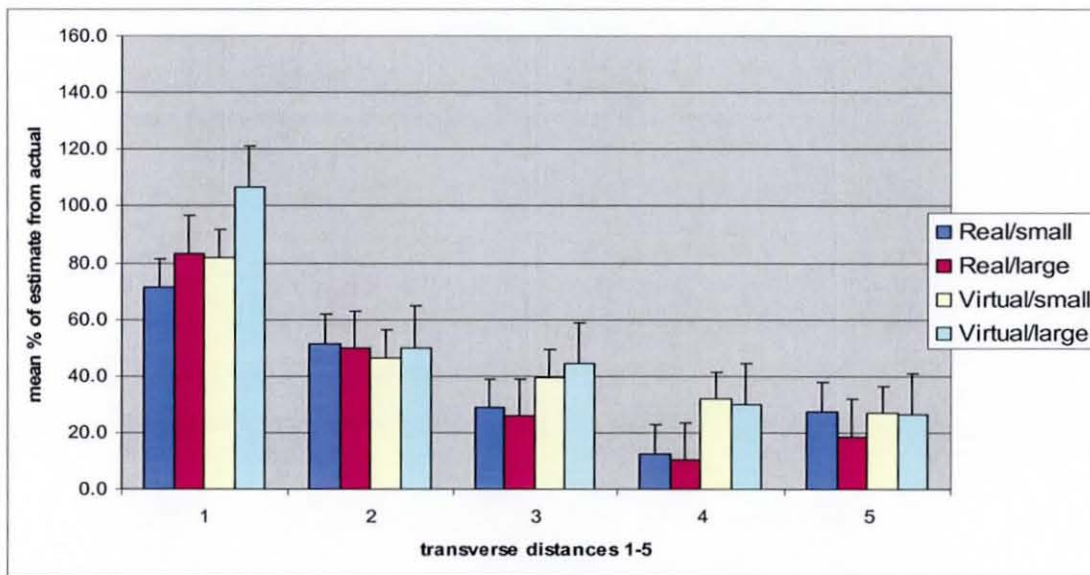
Figure 6-20 shows the differences between image types and between display sizes for transverse distance estimates. The figure indicates that overall, distances were largely underestimated. The mean percentage of estimates from actual scores indicated that on average, transverse distances were underestimated more in the real image compared to the VE image on both display sizes.

The results of a two-way between groups MANOVA performed on this dataset revealed no interaction effect:  $F(5,25)=.554, p=.734$ , Pillai's Trace=.100, partial eta squared=.100, observed power=.173). This indicated that there was no difference in the effect of image type on transverse distance estimates for the large and small display (see Figure 6-20). However, there was a statistically significant difference between the real and the VE on transverse distance estimation:  $F(5,25)=4.330, p=.006$ ; Pillai's Trace=.464, partial eta squared=.464, observed power=.917. Regardless of the display size, this indicated that there was a significant difference between the real and the VE image on the transverse distance estimates. From Figure 6-20, the real image was significantly less accurate than the VE image.



There was no main effect of display [ $F(5,25)=1.183, p=.346$ , Pillai's Trace=.191, partial eta squared=.191, observed power=.173] on transverse distance estimation, that is there was no significant differences between the small and the large display on transverse distance estimations. A comparison of the mean of the percentage estimate scores showed that for VE image, transverse distance estimates were better on a large display than on a small display. However, for the real image, the reverse was true.

The univariate tests results yielded no main effect of image (except for distance 3 and 4), no main effect of display size and no interaction effect ( $p > .05$ ). In Figure 6-21, for the real image, generally, distance estimates were better on a small display than on a large display (except for distance 1). For the VE image, the better performance on a large display was not reflected for all distances. Distance 1, 2, and 3 showed that estimates were better on a large display compared to a small display but the reverse was true for distance 4 and 5. For distance 3 and 4, estimates on the VE image was better than on the real image. As indicated by univariate tests, this difference was significant. Generally, on a small display distance estimates on the VE image were better than on the real image (except for distance 2 and 5). Similarly, on a large display, distance estimates on the VE image were more accurate compared to distance estimates on the real image (except for distance 2).



**Figure 6-21 Mean percentage of estimated distances for all the five transverse distances among all experimental conditions**

Table 6-9 indicates that the only comparison to reach statistical significance was between the real and the VE image on the large display. No other significance results were reported for other comparisons.

Table 6-9 Results of comparisons among experimental conditions using Students' t-test

Conditions compared	Student T-test values*
	transverse
Real/Small vs Real/Large	0.3672
Virtual/Small vs Virtual/Large	0.1618
Real/Small vs Virtual/Small	0.7295
Real/Large vs Virtual/Large	0.0682

\*Two-tailed distribution and assume unequal variance

6.2.4.5 Comparisons among asymmetrical distances

Figure 6-22 shows comparison among asymmetrical distances under the four experimental conditions. It was indicated that the vertical distance tended to be estimated more accurately when compared to the horizontal and transverse distance in all conditions; the transverse distance yielded the worst estimates.

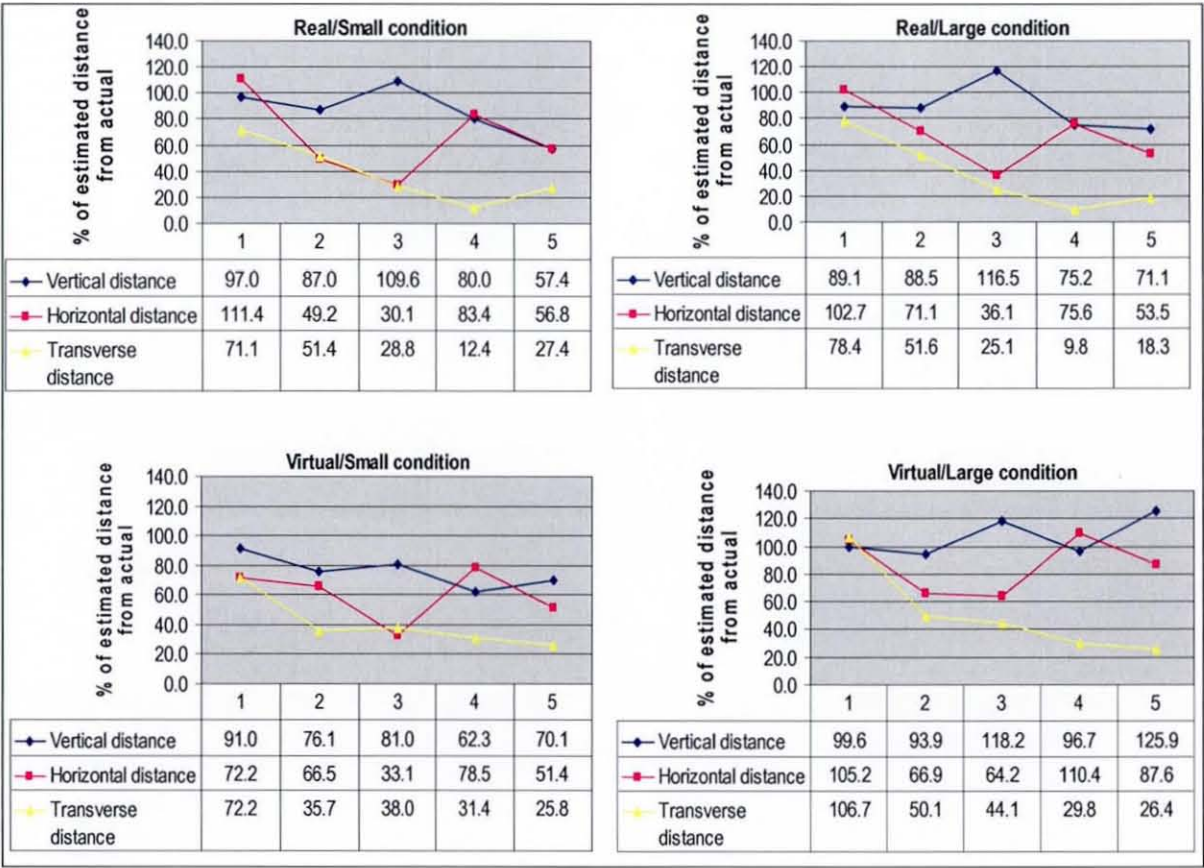


Figure 6-22 Comparison among asymmetrical distances under the four experimental conditions

A series of t-tests (Student t-tests) showed that the difference between the vertical and the transverse was highly significant in all conditions (Table 6-10). Similarly, the difference between the horizontal and the transverse distance was also significant. For the vertical and horizontal distances, all conditions reached statistical except in the real/small condition.

**Table 6-10 Results of t-test values for comparison among distance types under the four experimental conditions**

Conditions	Student T-test values*		
	vertical-horizontal	vertical-transverse	horizontal-transverse
Real/ Small	0.088	0.000	0.010
Real/Large	0.005	0.000	0.000
Virtual/Small	0.006	0.000	0.001
Virtual/Large	0.050	0.000	0.005

\*Two-tailed distribution and assume unequal variance

6.2.4.6 Post-test Questionnaire

Participants were asked to rate their estimation accuracy on the scale of 1 to 7 (7 represents very accurate). Generally, participants were more confident of their estimates on the real image compared to the VE image. About half of the participants (21 out of 40) rated themselves less than 4. Thirteen expressed slight confidence in their estimation; though, none felt very confident of their estimations (6 and 7 is zero). Five participants were not confident of their estimations (score = 2).

Most participants found the transverse distance very difficult to estimate (36 out of 40), while 25 found the vertical distance the most easy to estimate and 13 found the horizontal distances easy to estimate. Only two participants found the transverse distance too easy to estimate. A survey on their sports background revealed that 17 participants did not play any sports while the remainder played at least one of the following sports: tennis, badminton, squash, netball, hockey, and cricket. However, none of them were professional players. Generally, all participants indicated that they play sports as one of their leisure time activities only. Eight participants did not find viewing the movie had assisted them in their estimation, the rest found that the movie allowed them to make a better estimation especially for distant objects.

6.2.4.7 The effect of sport variance on distance estimates

To investigate the effect of sporting backgrounds of participants, an ANCOVA analysis was performed on the data using the sport background variable as a covariate. A summary of the results is given in Table 6-12.

From the table, it was shown that these results were similar to the earlier MANOVA analysis. There was a main effect of image for the vertical and transverse distances but an interaction effect for the horizontal distance. Similar to the earlier analysis no other significant effect was reported. The effect of sport variable was not significant for all asymmetrical distances. This result suggested that the contribution of sport background as a factor of influence was

minimal in these data sets. The observed power of the test was low. As such, careful interpretation of insignificant results is necessary.

A review of participants' sport ability revealed that all participants played sports as one of their leisure activities. No participant reported being professional players. Thus, all participants were similar in terms of their sporting background. Similar to Experiment 2A, this might explain the non-significant effect of sport background on these results. Therefore, these results do not allow us to generalize the effect of sport background on distance judgments for dynamic images.

**Table 6-11 Results of MANOVA analysis (Multivariate tests) using sport variables as covariate for all distance types (vertical, horizontal and transverse)**

Distance type	Effect	F(5,24)	p value	Partial eta squared	Observed power
Vertical	Sport	1.254	0.316	0.207	0.365
	Image	2.652	0.048	0.356	0.707
	Display	2.173	0.091	0.312	0.606
	Image*Display	0.792	0.566	0.142	0.235
Horizontal	Sport	1.471	0.236	0.235	0.426
	Image	1.671	0.18	0.258	0.481
	Display	1.466	0.238	0.234	0.425
	Image*Display	3.184	0.024	0.399	0.796
Transverse	Sport	0.863	0.52	0.152	0.255
	Image	4.302	0.006	0.473	0.913
	Display	0.909	0.492	0.159	0.268
	Image*Display	0.519	0.759	0.098	0.163

6.2.5 Analysis

On average, distances were underestimated for all asymmetrical distances. However, there was an exception to this; the vertical distance on a large display was generally overestimated.

Findings from this experiment revealed that there was a main effect of image for the vertical and transverse distance. This suggested that regardless of the display size used, there was a clear difference between the real and the VE on distance estimation for the vertical and transverse distances.

On average, for both vertical and horizontal distances, the VE image participants were more accurate compared to the real image participants on the large display. However, on a small display the reverse was true.

No significant effect of image type was reported for the horizontal distance; however, the interaction of image and display factors reached significant level ( $p > .05$ ). This indicated that the effect of image type for horizontal distance estimates is dependant upon the type of display used.

Comparison of mean percentage estimate scores suggested that the participants tended to perform better on a large display than on a small display for both the real and VE. No interaction effect was reported for the vertical and transverse distances.

Experiment 2B showed that there was no main effect of display for all asymmetrical distances. With the exception of the real image in transverse distance, numerical comparison of the mean percentage of estimate scores revealed that more distance underestimations were made on a small display than on a large display.

Examination of individual distances revealed that for most vertical distances the VE image participants were more accurate than the real image participants but on the small display the real image participants were more accurate. For the transverse distance, most distances reflected more accurate estimates in the VE image than in the real image. For most horizontal distances, more accurate estimates were reflected for the VE image compared to the real image for the large display but on the small display the real image participants tended to perform better than the VE image participants. Most of the horizontal and vertical distances indicated that estimates on the large display were more accurate than on a small display. For most transverse distances, this is true for VE image only; for the real image, estimates were more accurate on a small display.

It was mentioned earlier there were five different distances to estimate for each asymmetrical distance and these distance varied widely in terms of length, types and positions. The inconsistency in the direction of effects for all distances may be attributed to these differences.

Results showed that the vertical distance was estimated more accurately compared to the horizontal and transverse distances. This result was supported by the post-test questionnaire result where most participants found vertical distance easier to estimate compared to the transverse distance. Study results showed that the transverse distance yielded the worst estimate and this was consistent with most participants self-report comments which indicated that it was difficult to estimate compared to other asymmetrical distances. The participants' similar sport background level might account for the insignificant effect of the sport variable on their distance estimates.



## 6.2.6 Experiment 2B Discussion

Consistent with the results of Experiment 2A and the results of previous studies (Henry and Furness 1993, Witmer and Kline 1998, Witmer and Sadowski 1998, Wright 1995, Eggleston, Janson et al. 1996), generally distances were underestimated (with the exception of the vertical distance on a large display where distances were overestimated). It has been suggested (Yang, Dixon et al. 1999) and later confirmed (Dixon and Proffitt et al. 2002) that vertical overestimation will increase with an increase in the physical extent of the display size. This might explain the Experiment 2B results of overestimation for the vertical distance on a large display.

### 6.2.6.1 Image type

It was predicted that if image resolution was to play a role in the better performance of a real image over VE image (particularly on a large display), the use of a higher resolution VE image would result in the better estimates on VE image compared to the real image, especially for the horizontal and transverse distances. It was noted that in Experiment 2A, the vertical distance did not appear to be influenced by the low image resolution since on both display sizes whereby the VE participants performed better than the real image participants.

The results of current investigations confirmed this prediction at least for all asymmetrical distance types on a large display. The results indicated that the VE image participant estimates were significantly better than the real image participants estimates (except for the horizontal distance, the main effect was not significant but the interaction effect was significant).

For the small display, this was only true for the transverse distance. For the vertical and horizontal distances, it was shown that on a small display the real image participants performed better than the VE image participants. These results suggested that better image resolution does contribute an influence on participants' distance judgments for all asymmetrical distances on both display sizes. This result provides support for previous investigations whereby high image resolution results in improved distance judgments (Kline and Witmer 1996, Duh, Linh et al. 2002). However, this result was in contrast to the results of Thompson, Willemson et al. (in press) who reported no influence of image quality on distance judgment.

### 6.2.6.2 Display size

In Experiment 2A, while there was no significant difference between the large and small display, overall distance estimate performance was better on a small display than on a large display. It was suggested that this result may be partially influenced by the image resolution whereby the low image resolution may cause performance degradation especially on a large display. The use of a higher image resolution (at least for the VE image) in Experiment 2B revealed that the participants performed better on a large display compared to a small display. These results confirmed the earlier assertion that image resolution has some effect on distance estimation. That is, higher image resolution results in improved distance judgments. It was demonstrated that both high image resolution (Duh, Lin et al. 2001) and large display size (Patrick, Cosgrove et al. 2000, Tan, Gergle et al. 2003) would lead to a greater sense of presence and realism thus better performance. Thus, Experiment 2B results provided support for these findings whereby both factors do affect participants' distance judgments with some exceptions.

For the vertical and horizontal distances presented on a small display higher image resolution was used for VE image compared to real image. As such it was expected that the VE image participants would outperform the real image participants similar to the results of Experiment 2A. However, this was not observed in Experiment 2B results. Instead, the real image participants performed better than VE image participants. Participants' variations offered a more likely explanation for these inconsistencies.

It was shown that the difference between the large and the small display for the real image was not significant for both the vertical and horizontal distance. Some researchers have shown that a high resolution image of wide FOV offers more realism than a low resolution image (Duh, Lin et al. 2002). Others have reported that a wide FOV or a large display would result in better spatial performance over a small display (Patrick, Cosgrove et al. 2000, Tan, Gergle et al. 2003, Czerwinski, Tan et al. 2002). As such, it was expected that if a high image resolution was used for the real image in Experiment 2A, performance on the large display would be better than on the small display.

Similar to the arguments in Experiment 1B, the insignificant difference between the large and the small display in Experiment 2B suggests that the viewing distance and the physiological cues contributed an influence on distance judgment performance. In Experiment 2B, the retinal image size was different for both display sizes (due to the difference in FOV) (See Figure 6-14). When the retinal image size for both the large and small display was similar as

in Experiment 2A, the results showed that the small display participants yielded better estimation than the large display participants. For Experiment 2B, when the retinal image size was different, there was no difference between display sizes. Therefore, the results from both experiments suggested that the retinal image size was less influential on distance judgment task in dynamic images. This is because regardless of retinal image size, the effect of display was not significant.

The results of Experiment 2A showed that when a low resolution image (for a VE image) was presented on the small display, the viewed image appeared sharper and clearer compared to when viewed on the large display. The drop in image quality was not noticeable when presented on a small display but on a large display this is visible. This implied that image quality was less influential on a small display. This may also explain why better estimates were found on a small display compared to on a large display for all asymmetrical distances in Experiment 2A. Experiment 2B showed that when a higher image resolution was used for the VE image, the large display participants yielded better performance than the small display participants. These results suggested that the level of image resolution plays a significant role in affecting distance perception. The display size, viewing distance and physiological cues were also factors that influence distance judgments tasks. As argued earlier, besides the physical display size factor, the viewing distance and physiological cues also contributed towards explaining the better performance of the large display over the small display.

#### 6.2.6.3 Examination of individual distances

Similar to the results of Experiment 2A, Experiment 2B results indicated that an examination of individual distances revealed inconsistency in terms of effect types and direction of effects. Although overall results indicated that there was a main effect, examination of individual distances did not reveal this effect for all asymmetrical distances. Similarly for the direction of effect, not all distances in each asymmetrical distance showed a similar direction of effect. As the same image was used in Experiment 2A and Experiment 2B, similar to the explanation offered for Experiment 2A's results, these inconsistencies might be due to the variations in objects used as stimulus, objects positions their lengths. Similarly, the result of non-significant influence of sport background on distance estimates may be attributed to the homogeneity in the data set with regards to participants' sport ability.

#### 6.2.6.4 Comparison among asymmetrical distances

Comparison of participants' estimates among asymmetrical distances provided support for previous investigations (Henry and Furness 1993, Loomis, Da Silva et al. 1996). Vertical

distance was estimated more accurately compared to the horizontal and transverse distances. Similar to past studies' results, the transverse distance yielded the worst estimates with most distances were estimated on average less than 50% from the actual distance.

### 6.2.7 Experiment 2B Conclusion

When a higher resolution of VE image was used compared to the real image to replace the low resolution VE image in Experiment 2A, Experiment 2B results showed that overall the VE image participants performed better than the real image participants on a large display. These results provided support for the prediction that distance judgment was influenced by the image resolution especially for the horizontal and transverse distances on a large display.

For the vertical distance, the better performance of low resolution VE image participants over the real image participants in Experiment 2A and the better performance of high resolution VE image over real image suggested that for vertical distance, the quality of image does not appear to have an impact on the vertical distance judgments.

On a small display, large variability among participants might also account for inconsistent performance of the low and high resolution VE images. Moreover, on the small display, the difference in image quality between the real and the VE image (regardless of whether low or high resolution VE image) was very small or less noticeable. This implied that on a small display, image quality was less influential on distance judgment. But on a large display, poor image quality might constitute an important factor that affect user' distance judgment performance.

The non-significant difference between distance judgment performance on a large and a small display may be attributed to the similar viewing distances and, subsequently, similar physiological cues acting at the same distance from the screen. Thus, both cues contributed an influence on the user's distance estimation judgment.

The results of Experiment 2A and 2B suggested that the influence of retinal image size was very small, at least for the current experimental setup.

Similarly, the slightly better estimates of large display participants over the small display participants (in Experiment 2B) indicate that distance estimation was influenced more by the physical display size of the image rather than the retinal image size. From the examination of the individual distance estimates, it was suggested that object types, object positions in the

scene and object lengths were other factors that might affect participants' distance estimate accuracy.

The small variations in sporting ability among participants offered possible explanation for the non-significant impact of sporting background on distance estimates. Current results provide support for past studies that vertical distance was estimated more accurately compared to horizontal distance. The transverse distance being the most difficult to estimate gave the worst estimates.

### 6.3 OVERALL CONCLUSIONS

This chapter described a set of two related experiments (Experiment 2A and 2B) which investigated users' spatial awareness in terms of asymmetrical distance perceptions in dynamic images. The experimental methods, results, discussions, and conclusions for each experiment were presented. Basically, the experiment approach for Experiment 2A and 2B were similar to Experiment 1A and 1B (reported in Chapter 5) respectively.

Experiment 2A examined the effect of image type and display size on asymmetrical distance estimates while Experiment 2B examined the effect of viewing distance (hence physiological cues) and image resolution on asymmetrical distance estimates.

The results of Experiment 2A showed a main effect of image type for all asymmetrical distances. However, the direction of effect varies depending upon the image type and display size used. The effect of display size was not significant but surprisingly distances were more accurate on a small display compared to a large display. The use of a low resolution was suggested for this unexpected finding.

Experiment 2B results suggested image resolution played a significant role in influencing asymmetrical distance perceptions. Generally, distance perceptions in VE image were significantly better than in the real image. The non-significant effect of display size in Experiment 2B indicated that besides the display size, both the viewing distance and the physiological cues partially explained the better performance of the large display over the small display for asymmetrical distance perceptions in dynamic images.

A further discussion of these experimental results will be presented in the overall analysis of all experimental results in Chapter 8. Prior to that, in the next chapter (Chapter 7) the experimental analysis of user's spatial awareness in interactive images is presented.



# CHAPTER 7

---

## *EXPERIMENT 3: DISTANCE PERCEPTION AND SPATIAL MEMORY TASK IN INTERACTIVE IMAGES*

### 7 OVERVIEW

This chapter outlines the experimental methodology and the relevant results of Experiment 3 which examined user spatial awareness in the interactive real and VE where users were allowed to freely explore and navigate in these environments. In addition to distance estimation tasks, in Experiment 3, spatial memory task was also evaluated. Moreover, as the users were allowed to interact with the VE, the effect of different interface devices and the navigation method used for interactions on the user's spatial memory were also examined. Similar to Experiment 1 and 2, the VE images were presented to the participants in a non-stereo viewing mode and the effect of presenting the images on different display sizes was also examined.

Initially, the approach was to conduct only one major study for Experiment 3 using the setup of Experiment 1B and 2B in which the viewing distance of the observer was constant and the FOV of both display sizes were varied. Most previous investigations (based on interactive images) used the setup of Experiment 1A and 2A to investigate the effect of display size.

Thus, Experiment 3A's experimental setup was based on the setup of Experiment 1B and 2B. Employing the setup of Experiment 1B and 2B allowed us to compare the results of the first study (Experiment 3A) with those of the previous investigations. However, the unexpected findings from this study provided further motivation to conduct another study (Experiment 3B). Thus, two sets of studies for Experiment 3 (Experiment 3A and 3B) are reported in this chapter. In the following sections, the experiment aims, hypothesis, experimental methodology, results and discussions from both studies were presented.

## 7.1 EXPERIMENT 3A: EFFECT OF VIEWING DISTANCE, PHYSIOLOGICAL CUES, INTERFACE DEVICE AND TRAVEL MODES

In this experiment, spatial awareness in the interactive real and VE was examined. In the previous chapters (5 and 6), the investigations of spatial awareness were conducted on static and dynamic images. The results from such situations may not generalize to interactive VE where users were allowed to explore and interact with the VE. Additionally, when users were allowed to interact with the VE, issues such as the choice of interface device for interaction and method of navigation in the VE would warrant further investigations as these factors may influence a user's spatial performance. As such, in this experiment, the effect of interface devices (mouse and trackball) and navigation methods (drive mode and fly mode) were examined and compared. They were chosen because they are most likely to be used in low cost VE applications and represent interface types that are necessary familiar or intuitive to most users. Further rationale for the investigation of these devices and navigation or travel method were presented in Chapter 4. Similar to Experiment 1 and 2, participants viewed the VE images presented in a non-stereo mode. The effect of display size (that is presenting the VE images on a large and a small display) was explored.

### 7.1.1 Experimental Aims & Hypotheses

The overall aim of this study was to investigate the user's spatial awareness in an interactive VE presented on varying display size in comparison to similar task performance in the real environment. The spatial tasks evaluated were spatial memory task and distance estimation tasks. The influence of interface device type (a mouse and a trackball) and travel mode (drive and fly mode) used for interacting with the VE on distance estimate and spatial memory tasks in the VE was also examined. The following main hypotheses were explored in this study:

1. The type of environment (real vs. VE model) has no effect on participant's distance estimation task (vertical, horizontal and transverse) performance

2. The type of environment (real vs. VE model) has no effect on participant's spatial memory task performance
3. The display size (small vs. large) has no effect on participant's distance estimation task (vertical, horizontal and transverse) performance in interactive VE
4. The display size (small vs. large) has no effect on participant's spatial memory task performance in interactive VE
5. The type of input device (mouse vs. trackball) has no effect on participant's spatial memory task performance in interactive VE
6. The different mode of travel (drive, fly) has no effect on participant's spatial memory performance in interactive VE

The secondary hypotheses were:

1. There is no effect of viewing distance on distance estimate task in interactive VE
2. There is no effect of physiological cues on distance estimate task in interactive VE
3. There is no effect of viewing distance on spatial memory task in interactive VE
4. There is no effect of physiological cues on spatial memory task in interactive VE

In this study, the VE model used was based on a room, thus the terms height, width and length were often used interchangeably to refer to vertical, horizontal and length respectively. Additionally, a survey on the users' evaluation of both interface devices and both travel modes were conducted using a post-test questionnaire. The questionnaire was undertaken to provide more information on users' experience using the interface devices and travel modes, which may provide explanatory information on the spatial task performance results. The objective of the questionnaire was to survey which interface device and travel mode was preferred by the user based on the set of criteria defined in the questionnaire (to be described in next section).

## 7.1.2 Methodology

### 7.1.2.1 Participants

A total of thirty-four paid volunteers, comprising of staff and students, participated in the study. The ages of the participants ranged from 21 to 44 years with an average age of 31.8. Twenty-four (seven females and seventeen males) participated under the VE conditions, while the remaining ten (1 female and 9 males) participated in the real environment condition. For the VE conditions, two groups were required (VE/large and VE/small), thus twelve participants were randomly allocated to each group. All participants had normal or corrected-to-normal vision. A summary of participants' background in terms of gender, sport

background, frequency in playing computer games and participations in VE experiments is presented in Table 7-1. It is noted that the sporting background indicates whether participants play any sport such as football, hockey, tennis, badminton and cricket. From the table it was shown that most participants reported playing sport for leisure activities; only two reported that they were amateur players.

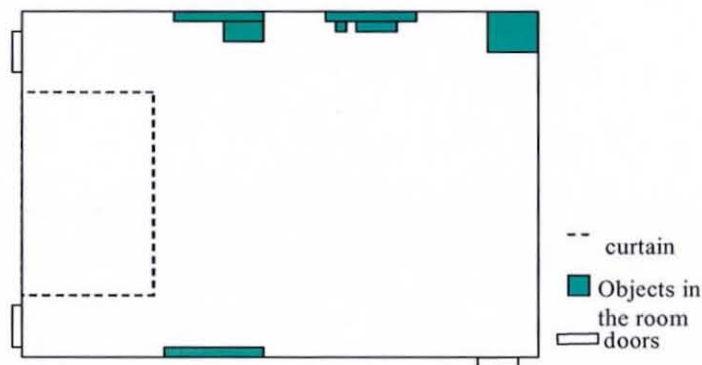
**Table 7-1 Summary of Real and VE participants' background**

Condition	Gender		Play sport		Play Computer games / per week			Participate in VE experiment		
	Male	Female	Yes	No	0	1-4	5 or more	0	1-4	5 or more
Real	9	1	8	2	Not relevant					
VE/Large display	8	4	7	5	3	6	3	6	6	0
VE/Small display	9	3	9	3	5	5	2	6	5	1

### 7.1.2.2 Materials/Apparatus

#### Real environment

In order to undertake this experiment it was necessary to employ a room that has the following characteristics: spaciousness and uncluttered. Spaciousness in this study means 'larger in extent or capacity, in length and breadth'<sup>4</sup>, while non-cluttered means 'contains few objects or almost vacant space.' These characteristics were necessary in order to carry out the spatial memory test, where the objects for recall were to be placed in the room. Moreover, as reviewed in Chapter 4, a cluttered environment may hinder a user's navigational tasks. A room in one of the university's buildings was chosen as it met these requirements.



**Figure 7-1 Layout of the experimental room**

Figure 7-1 depicts the layout of the room and location of objects in the room. The room was approximately 15m x 8m in dimension and was fully carpeted. All of the objects

<sup>4</sup> Merriam-Webster online dictionary definition. Available at <http://www.m-w.com>

(bookshelves, notice-boards, computer and computer tables) were located on the walls or at corners of the room, thus creating a large vacant room. Figure 7-2 shows a picture of the real room with the objects for spatial memory test placed on the floor of the room.



**Figure 7-2 Picture of the Real Environment with objects for the spatial memory test placed on the floor**

### Virtual environment

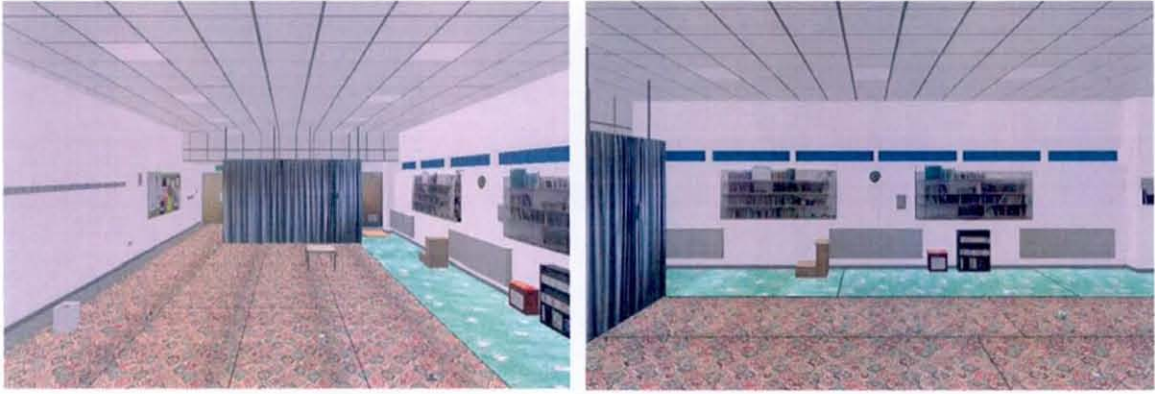
A 3-D model based on the real room was created using MultiGen II software, running on a Silicon graphics workstation. Detailed measurements of the room, the objects and their locations were carefully taken before the modelling process. Pictures of objects in the room were taken using a digital camera which captured the images at a resolution of 1280 x 960. Appropriate textures from these pictures (e.g. carpets, bookshelves, notices, doors, table) were used as textures in the modelled scene to match the model as close as possible to the real environment. Figures 7-3 shows two different views of the 3D VE model. The resolution of the VE model was 1280 x 1028.

### Practice environment for the VE condition

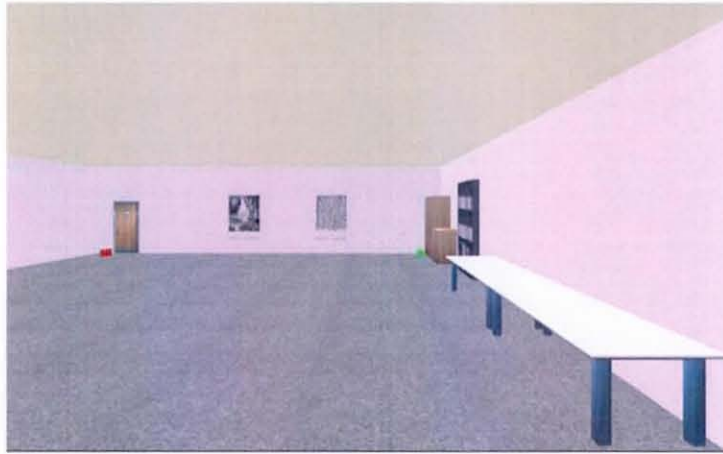
A different 3-D model of a room was created using MultiGen II Pro software for the practice sessions. This room had no real world equivalent and was much larger the test VE. However, it is similar to the test VE in terms of the following: it is uncluttered and fully carpeted. Additionally this room had pictures on the walls, and some tables and cupboards at corners of the walls (Figure 7-4). It was observed that during the pilot sessions some participants just made a few movements or moved at one or two corner of the room and assumed they were already familiar using the device and travel mode. As a result, during the trial sessions, when participants had to look around for objects, they realised they had not enough practice using the device or travel. To ensure that participants moved around the room and practiced using



the interface device, the participants were asked to look for different coloured cubes placed at each corner of the room.



**Figure 7-3** Different views of the 3D VE model



**Figure 7-4** Practice environment for VE conditions

### ***Problems encountered during development and testing of VE***

Two major software problems were encountered during the VE model developments. The first problem was with the viewer software when viewing the VE model in a 'drive' mode. During viewing, with collision detection, the viewpoint jumps rapidly up and down. Careful examination of the source code of the SGI Performer PERFLY software and some experimentation, it was found that the rapid up and down movements of the viewpoint was due to the bounding box of the objects in the VE. During viewing (that is movement through the VE), the collision detection algorithm detected the bounding box of the ceiling and the floor, causing the viewing point to shift up and down between each object's bounding box. When the ceiling of the room was removed from the VE, the bouncing of the viewpoint stopped. However, removing the ceiling failed to make the VE comparable to the real environment condition. One solution was to increase the scale of the room but the bouncing of the viewpoint still occurred. After several trials of using small test models, it was found that reversing the face of the objects (a function in MultiGen II Pro software) removed the

bouncing problem of the viewpoint. Initially, this solution was tested on the ceiling's objects. The ceiling was made up of several objects: the vertical and horizontal lines, ceiling tiles, lamps and heating ventilation. The drive mode was tested by presenting these objects one by one and checking if they caused the bouncing problem, if they did, the command 'reverse face' was used on the object. After many trials, the bouncing viewpoint problem was eliminated completely.

The second problem was when the VE model was transferred to a different platform for the experiment. The VE model was created on a Silicon Graphic workstation and, for the experiment, the VE model was transferred to a PC. This transfer was necessary to solve the problem of switching from a mouse to a trackball for each experiment condition during experiment trials. The switching was more easily done when on a PC. Moreover, a PC version of the viewer software was available. However, this transfer resulted in the objects in the model either not being displayed or being displayed without texture. The missing texture problem was resolved by changing the path address option to 'relative to current database'. Examination of the VE model database revealed that the missing objects were all externally referenced objects. As there was no option to change the object path address to "relative" as in texture, the first solution was to convert the externally referenced objects to be part of the main database. But the objects' locations were still incorrectly referenced by the main database. After several trials, the problem was resolved through a tedious method: that is by changing the path of each referenced object to the similar path used in the target machine (PC) before transferring the VE model to the PC. This was done individually for each externally referenced object. For a large database, this would be very time-consuming, thus a faster method would be to write a C program to change the addresses of the objects to the target machine address. However, since the number of objects in the VE model was relatively few, the changes were done manually.

### Objects for the spatial memory test

Nine objects were identified to be used for the spatial memory test: a book, an alarm clock, an umbrella, a telephone, a pencil, a trashcan, a mug, a camera and a small table. These items were considered highly familiar items and they represented similar items used by other investigators of object location studies (Postma, Izendorn et al. 1998, Arthur, Hancock et al. 1998). The pencil was later replaced by a ball. This decision was made after placing the objects in the VE and the pencil was hardly visible due to its size and the colour of the textured carpet. The number of objects considered here was based on the limits of the human capacity for processing information (Miller 1956). The normal memory span is seven to nine



items, a number which represents the capacity of the short-term memory (Greene and Hicks 1984). Nine was chosen to avoid ceiling effect which may occur when all results are perfect scores or floor effect which may occur when all scores are too low (Johnson and Stewart 1999).

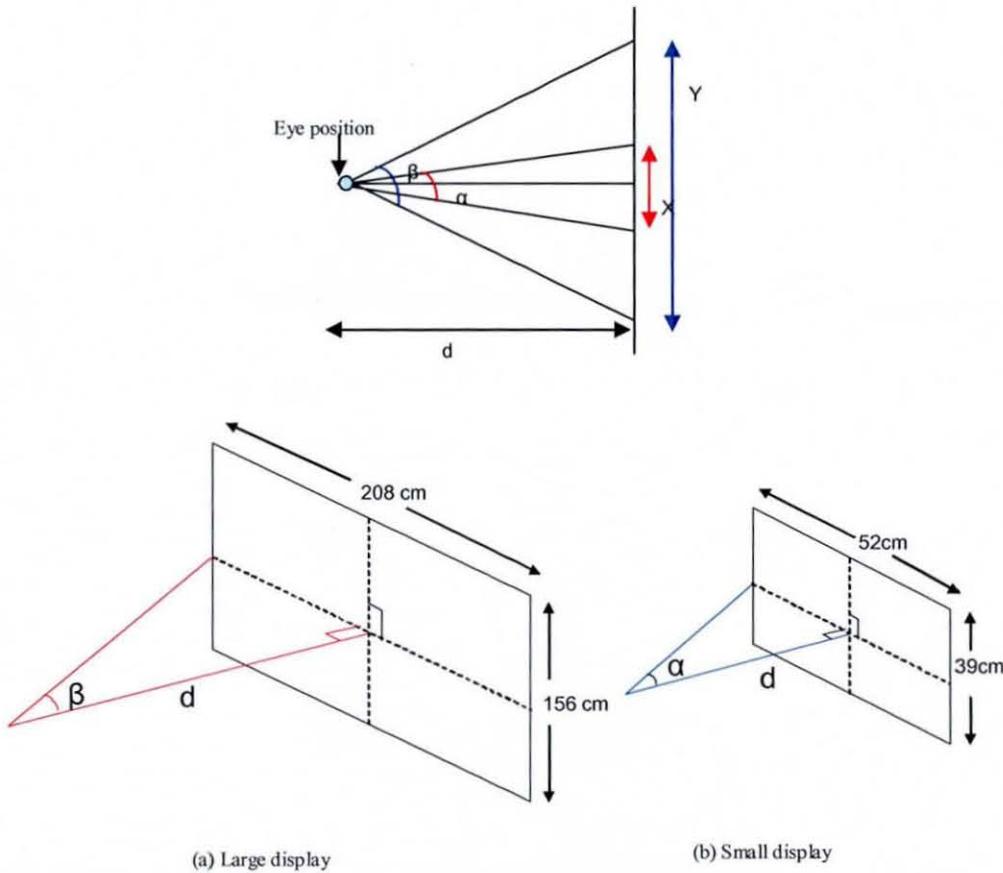
The objects were randomly placed on the floor of the room. Initially, it was planned to place some objects on the walls and on the floor. After careful consideration of the experiment process, placing objects on the walls would limit the type of objects to be used for the spatial memory (objects that could be hang or posted on the walls only). Additionally it was not practical to relocate object positions for the real world conditions for each test trial (such as putting on new nails on the wall for hanging objects), as such changes would not have been allowed by the university.

The random locations of the objects were generated using Microsoft Excel random number generator function. For each object, two sets of random values were generated to represent its  $x$  and  $y$  value. No  $z$  value was required since objects were placed on the floor ( $z$  – refers to the height and it is assumed to be zero, while  $x$  and  $y$  refers to the width and length respectively). The objects were randomly located for each of the test trial conditions to reduce carryover effect or learning effect.

### Display apparatus

The VE model was displayed on a rear-projected display for both the large and small display conditions. The Silicon Graphics Inc. Performer PERFLY (PC version) was used to view the VE model. The computer used to run the software is a Pentium III 2.66GHz with 500Mb RAM. The video card is based on NVIDIA GFORCE 4.

Waller (1999) found that a GFOV value of between  $50^{\circ}$ - $80^{\circ}$  yields more accurate estimates. A pilot session revealed that a GFOV value of  $50^{\circ}$  made it difficult for the user to move around the room and look for objects especially in drive mode. Reducing the GFOV made the field of regard smaller and this is not suitable especially for drive mode where all the user could see was the wall of the room. However, using a large GFOV value of  $100^{\circ}$  makes the VE appear distorted and compressed (see Waller 1999). Thus, the GFOV of the VE was fixed at  $80^{\circ}$  for all conditions. As previously reviewed in Chapter 4, a dark room setting was necessary to reduce peripheral view effects from the objects surrounding the display screen.



**Figure 7-5 Experiment 3A setup. Viewing distance ( $d$ ) was fixed for both the large and small displays but the FOV was varied ( $\beta > \alpha$ )**

Similar to Experiment 1B and 2B, the participant's distance from the display screen was equated for both the small and large display conditions (see Figure 7-5), thus, controlling the effect of viewing distance and physiological cues of accommodation and vergence cues. This setup was selected as most previous researchers (Yang, Dixon et al. 1999, Patrick, Cosgrove et al. 2000, Tan, Gergle et al. 2003) employed the setup of Experiment 1A and 2A where the visual angles were equated for both display sizes and the viewing distance was varied. Employing the former setup allowed us to compare Experiment 3A results with the results of these previous investigations.

The viewing distance was set at 100cm. The projected image size for the large display condition was 208 x 156 cm, while the projected image size for the small display condition was 39 x 52 cm. The resulting FOV for the large and small displays were  $92^\circ$  and  $22^\circ$  respectively.

Similar to Experiment 1B and 2B, if the viewing distance and the physiological cues were to contribute to the better performance of a large display over a small display in the previous



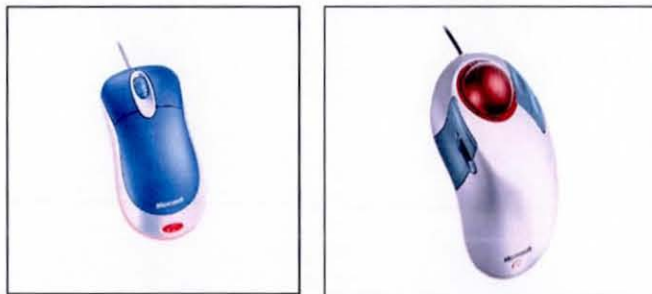
investigations, it is expected that the results of Experiment 3A will show a non-significant effect of display size.

### Interface devices

As discussed in Chapter 4, the interface devices chosen for navigating through the VE were a mouse and a trackball. Both the mouse and the trackball were considered the most commonly used input device available and both were expected to considerably reduce practice session time for the participants (and less learning time to navigate the VE). Additionally, since interaction in this study was limited to movement or navigation only with no object manipulation, these two input devices were considered adequate for the tasks under investigation.

A Microsoft Optical Mouse Blue (Figure 7-6), which is based on optical technology was chosen for this experiment. The mouse has two buttons (left and right) and one wheel for scrolling. The functions of these buttons can be changed easily to suit requirements. For the purpose of this experiment, the left and right button was set for forward and backward movement respectively. The wheel allowed movement according to where the cursor was pointed to by pressing it down and moving the mouse accordingly. Alternatively, the user could also use the left button or right button (instead of the middle button/wheel) for this purpose.

For the trackball, The Microsoft Trackball Explorer (Figure 7-6) was used. This device was also based on optical technology for precise cursor movement and accuracy. The trackball however has 6 buttons, whose functions could also be easily changed. In this experiment, only three buttons were used to allow for consistency with the mouse. The leftmost and the rightmost buttons were set for forward and backward movement respectively. Rolling the ball allowed movement according to where the cursor was pointed to which was similar to moving the mouse device. For the trackball, the user could roll the ball for movement in any direction but for the mouse the user need to move the mouse to accomplish the same function.

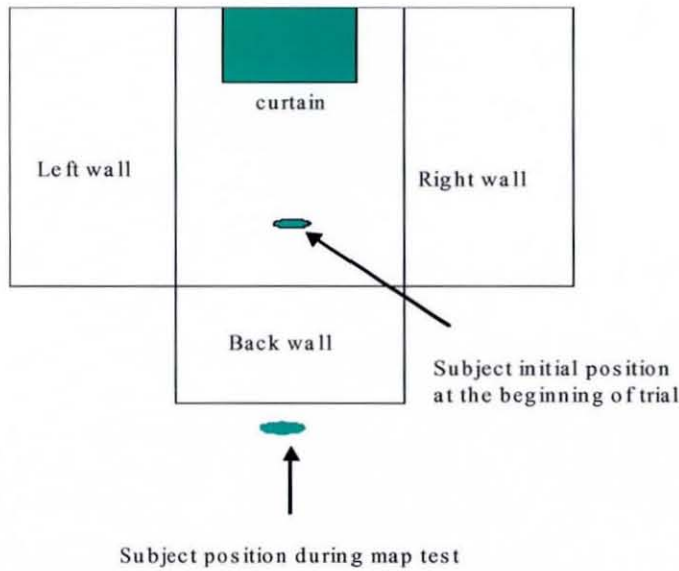


**Figure 7-6 Mouse ( left image) & Trackball (right image)**



### Spatial memory test

For the spatial memory test assessment, the method used by Arthur, Hancock et al. (1997) was employed. However, a slight modification to this method was made. Arthur and Hancock et al. (1997) gave a map of the room with a scale and orientation information (the map contained two of the object positions already filled). Similarly, in Experiment 3A the participants were also given a scaled map of the room. However, for the orientation of the room, participants were informed of their orientation in the room with respect to their initial position before exploration of the environment. Additionally, during the spatial memory test the map was placed in front of the participants similar to their initial position in the room or VE (see Figure 7-7).



**Figure 7-7 Participants' initial position in the Real and VE conditions for exploration**

During the spatial memory test, participants were asked to recall object positions by identifying the correct object to its correct position (absolute placement). This placement process method was chosen because some empirical evidence suggested that this method yield no difference among gender performances (Postma, Izendorn et al. 1998). Thus employing this method would reduce the influence of performance due to gender differences. To represent the object positions on the map (Figure 7-7), the participants were asked to draw a cross (X) for the centre of each of the object location and to label it using the given object number. Since the objects tested were of different sizes, this method allowed for precise identification of the object locations. Arthur, Hancock et al. (1997) used a point instead of a cross. A cross was considered more precise since a point may allow for error due to the different sizes of points drawn by different participants.

## Post-test questionnaire

The overall purpose of the interface device questionnaire was to survey the participant's evaluation of their experience when using both the interface device and both the travel modes. It was expected that the subjective responses from this questionnaire would yield useful information to serve as explanatory information on the spatial memory tasks performance results. Additionally information on participants' familiarity with the use of the interface device was also collected.

The questionnaire was divided into two major sections. The first section consisted of four parts. The first part gathered information about participant's background (familiarity) on the use of the interface device. The second part was concerned with the mode of travel used in the test (drive mode and fly). This part consisted of four questions (Table 7-2). The purpose of this part was to identify which mode of travel is preferred by the participants based on the interface device being used. The criteria were based on the followings:

- 1.1. Ease of movement in the VE
- 1.2. Control of movement in the VE
- 1.3. Assist them in the task required (that is object recall (spatial memory task))
- 1.4. Overall preference of the travel mode for each interface device

**Table 7-2 Question 2 on Mode of travel in VE**

No.	Question	Rating
2(i)	In your opinion, which mode of travel helps you to move easily in the environment?	1 = Difficult, 7 = Easy
2(ii)	In your opinion, which mode of travel helps you to control your movement in the environment?	1 = Less control 7 = Most control
2(iii)	In your opinion, which mode of travel helps you to easily recall object position?	1 = Difficult , 7 = Easy
2(iv)	In your opinion, which mode of travel do you prefer to use?	1 = Least preferred 7 = Most preferred

The third part was concerned directly with the interface device used and this comprises of five questions (Table 7-3). The purpose of this part was to identify which interface device is preferred (regardless of travel mode) by the participants based on the similar criteria asked in Question 2. An additional question was based on which device helped them to position themselves in the VE.

**Table 7-3 Question 3 on Interface device use in VE**

No.	Question	Rating
3(i)	In your opinion, which interface device do you find easy to	1 = Easy



	use?	7 =Difficult
3(ii)	In your opinion, which interface device allows you to move and position yourself easily in the environment?	1 = Easy 7 =Difficult
3(iii)	In your opinion, which interface device allows to control your movement in the environment?	1 = Most control 7 = Least control
3(iv)	In your opinion, which interface device do you feel makes it easier to recall object position?	1 = Easy 7 =Difficult
3(v)	In your opinion, which interface device do you prefer to use?	1 = Most preferred 7 = Least preferred

The choice of criteria used for the above questionnaire was adapted from VRUSE questionnaire (Kalawsky 1999). The purpose of the VRUSE questionnaire is to measure the usability of a VR system according to the attitude and perception of the users. As the purpose of the Interface device questionnaire was on the evaluation of the interface device and the travel modes, only *Part 2 User Input* of the VRUSE questionnaire was referenced. However, not all the questions in this part were used. As we were interested on the user's evaluation of the interface device and navigation, the choice of criteria was limited to the relevant questions as assessed by the following criteria: *ease of use* and *appropriateness* (Questionnaire 8, 9, 10, 11, 12, 17 and 18). But these questions were reworded to suit the current experiment requirements (see above). Furthermore, the purpose of the questionnaire in Experiment 3A was to compare between interface devices and between navigation modes, thus questions were delineated and presented along this line. Another modification was the scale, while the VRUSE was based on a five-point Likert scale; we use a seven-point Likert scale. This was to maintain consistency with the questionnaire in Experiment 1 and 2.

The fourth part of the Interface device questionnaire deals with participant's performance accuracy rating. The purpose was to survey how confident they were on the spatial memory tasks. It has been suggested that asking participants to provide confidence ratings of their magnitude of estimate is one direct method of assessing their certainty (Radvansky, Carlson-Radvansky et al. 1995).

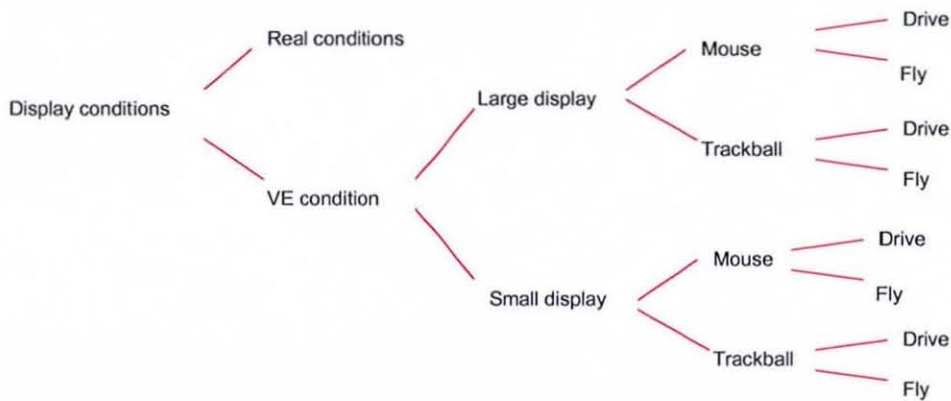
Section two of the interface questionnaire was concerned with the VE model itself and it comprises of only one question (Question 5). Participants were asked if they recognized the room or not and if they did, they were asked to rate how much this knowledge helped them in the recall process. It should be noted that even though the model was based on an actual room, the VE model differed from the actual room in terms of the arrangement and the positions of the objects. It also differed in terms of the presence and the absence of some objects. This is because the room was a common room and it is frequently used by students

and as such, changes on a daily basis were expected. For ease of modelling and to maintain consistency of the VE model with the real room conditions, prior to conducting the real room condition, it was necessary to ensure that the setting of the real room in terms of objects number and positions were the same as the VE model. Thus, it was expected that the participant familiarity with the room would have less influence. Furthermore the test objects used in the spatial memory test were objects which were not originally present in the room.

For each question in each section, participants were asked to make comments on their choice. Finally, participants were asked to make overall comments with the experiment in general. The participants' comments provided additional information towards understanding the responses they provided.

### 7.1.2.3 Experiment design

In this study, a mixed design was employed. The real and VE conditions were between-subject variables while the interface device (mouse and trackball) and travel modes (drive and fly modes) were within-subject variables. A group of ten participants experienced the real environment condition. For the VE conditions, 12 participants were assigned to each large and small display condition. Since the interface device and travel mode were within subject variables, all participants condition experienced both interface device and both travel modes. Figure 7-8 provides an overview of the variables examined in this study.



**Figure 7-8 Summary of experimental variables for Experiment 3A**

Three types of DVs were collected: room size estimation and spatial memory recall (object placement) test, post-test questionnaire ratings. For the room size estimation, participants were asked to estimate the height, width and length of the room. For the spatial memory recall (object placement) test, the number of correctly placed objects was collected. For the post-test questionnaire, participants' ratings on the interface device and travel modes and their subjective responses gathered.



#### 7.1.2.4 Procedure

##### Real environment

Participants first undertook a short questionnaire to capture their background (age, gender, sports background). They were initially briefed on the purpose and the procedure of the experiment. Additionally, they were also given an instruction sheet which explained the details of the experiment (see Appendix C for instruction sheet). Prior to entering the test room, participants were given a list of objects they needed to recall. They were then told to close their eyes and were then led to the test room. They were positioned at the initial position facing the curtain (see Figure 7-7). This gave the participants an orientation of the room (which will be later used in the spatial memory test). They were informed that they were to move about in the room after being told to open their eyes. Participants were asked to remember objects and their locations and were told they were to recall them later in the test. Participants were told not to worry about the names of the objects as the list of objects will be provided during the test later (Arthur, Hancock et al. 1997). Participants were reminded that all objects to be recalled were located on the floor only. Participants were encouraged to ask if they were not certain of an object's name during the test trial. As the experimental room was a single, simple and non-cluttered room, participants were allowed 3 minutes to browse in the room. When their time was up, participants were asked to immediately close their eyes again to prevent further viewing of the room. They were then led to out of the test room to undertake the spatial memory test.

##### **Spatial memory test:**

Participants were given an A3 size paper (Figure 7-7) showing the basic layout of the room. They were told that the map sheet represents a scaled drawing of the virtual room. The paper was placed in front of the participants similar to their initial position when they started viewing the VE, that is facing the curtain (refer to Figure 7-7).

A list of 9 objects was given to the participants. They were told to mark a cross using a pencil at the location they thought was the centre of the each object's location and label it. Subjects were given as much time needed to complete this map test. The time taken to take the spatial memory test was recorded for each participant.

When the participants had completed the spatial memory test, the participants were asked to estimate the size of the room by estimating the height, width and length of the room in metre unit. They performed this estimation without being in the room. A layout of the room was given to indicate to the participant which parts of the room constitute the length and width of the room. Figure 7-9 shows pictures of a participant in the real condition.





**Figure 7-9 Real condition test trial**

### Virtual environment

As with the real condition, participants first undertook a short questionnaire on their background (age, gender, sports background, computer games experience and VE experience,) and then were briefed on the purpose and procedure of the experiment. They were also given an instruction sheet which explained the details of the experiment (see Appendix C for instruction sheet for the VE condition).

Depending upon the experiment condition, participants were either exposed to a VE model presented on a large display or a small display. However, all participants experienced a total of four conditions each: mouse/drive, mouse/fly, trackball/drive, and trackball/fly. The following represents the sessions each participant went through for each of these four conditions:

**Practice session:** Participants were first given a practice environment to familiarize themselves with movement in the VE using an interface device and travel mode depending on the condition they were assigned based on the counterbalanced design. In the practice environment, participants were asked to approach six coloured cubes located at each corner of the room. This was to ensure that participants practice moving around the room. Participants were given as much time needed to familiarize with navigation in the VE practice environment. Practice time for each participant was recorded.

**Test session:** Participants were seated about 100cm from the screen and their eye level (centre of projection) was made similar for each participant. This was done by positioning the participants at equidistant from the edges of the pictures. The chair height was adjusted accordingly. Participants were then given the experimental VE to navigate. All participants start at an initial position of the VE room model facing the curtain similar to the real environment condition (see Figure 7-7). They were then asked to move about the VE and were told to

remember the objects and their locations in the room. They were told they would be required to recall them later. Movement in the VE was not restricted; they were free to move about. Similar to the real world condition, participants were allowed 3 minutes to browse in the room. Participants were allowed to ask questions if they were not certain of an object's name during the test trial.

**Spatial memory test session:** After completion of the test session, participants were given a spatial memory test similar to the spatial memory test in the real world condition.

Participants were given 5 minutes break between each session. After completing the four conditions, the participants were asked to estimate the volume of the room by estimating the height, width and length of the room. Similar to the real condition, the participants did not view the VE room again when making this estimate. Finally, each participant was asked to complete the interface device questionnaire. Figure 7-10 shows pictures of participants during the test and practice sessions in the large and small VE conditions.

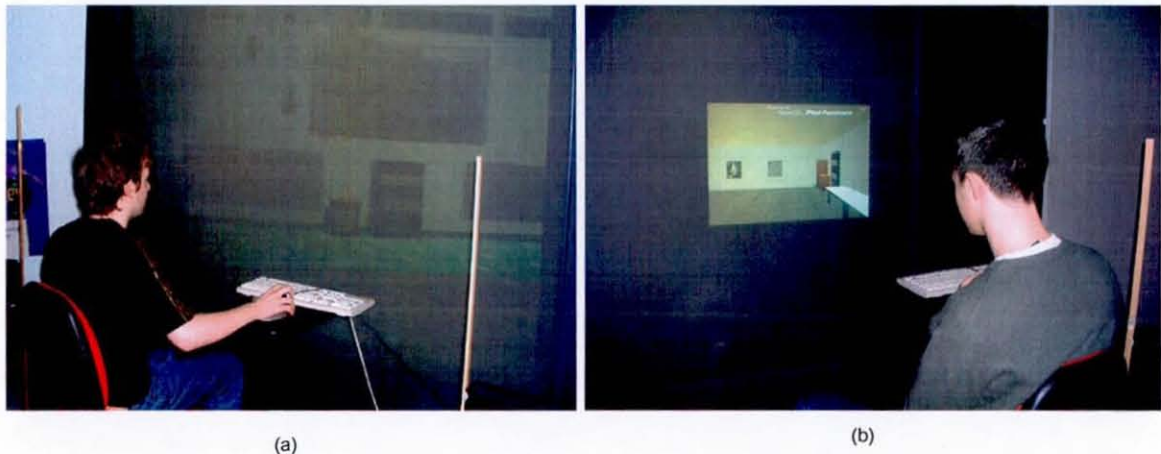


Figure 7-10 VE conditions: (a) Large display (test trial) (b) Small display (practice session)

### 7.1.3 Results

#### 7.1.3.1 Data preparation

##### Room size estimation data

For room size estimation performance accuracy was based upon the percentage of estimate from actual distance based on the formula used in Experiment 1 and 2 (see Section 4.2.2 of Chapter 4). As in Experiment 1 and 2 a value of more than 100 shows an overestimation from the actual distance, while a value of less than 100 indicate an underestimation. A value of 100 indicates a veridical estimate of the actual distance. The result from the real condition



was used as a comparison to the results of the VE conditions. Data was also checked for outliers.

### Spatial memory data

For the spatial memory test, participants were required to indicate (by drawing an X, followed by the object number) on piece of blank paper the locations of objects seen in the VE. As a precautionary measure additional experimental sessions were run and any sessions with irregularities were discarded prior to analysis (see also Garau 2003). The data was analyzed using a real world method employed by Alfano and Michel (1990) with some slight adjustment. In their method, the position of object was correctly placed if it falls within a square grid, otherwise it was considered as an incorrect placement. However, the authors considered the 1-inch square criterion they used for correct/incorrect placement decision was too strict. Thus, instead of using a 1-inch square criterion for correct/incorrect placement, a slightly larger and less strict criterion was employed in Experiment 3A (that is, 3cm square criterion was used – 1 inch is equal to 2.54cm). Thus, based on this method data was collected based on the number of correct objects placements for each participant. Using this method of analysis would enable comparison of the real world conditions results to the real world study results of Alfano and Michel (1990).

The time taken to practice in the practice VE (practice time) and the time taken to take the spatial memory test (map test time) were also recorded for each participant in each condition. Unit measure for these times was in second.

### Post-test questionnaires data

The post-test questionnaire data was collected in two forms: rating on seven point Likert scale and subjective comments from the subjects. The reliability of the scales of the post-test questionnaire was checked on its internal consistency. The internal consistency here refers to the degree of the items that make up the scale measure the same underlying construct. The Cronbach's alpha coefficient is a common indicator of internal consistency. This value should be above .7; for scales with less than 10 items a low Cronbach value may be used (Pallant 2001). A reliability analysis for the scale was performed. However, prior to that, it was necessary to reverse "negatively worded" items. In questionnaire, question 2(i), 2(ii), 3(i), 3(ii) and 3(iv) were reversed (see Appendix C for the method to reverse). The reliability analysis revealed that the Cronbach's coefficient value was .7278, which was above the ideal value and reflected the internal consistency of the scales of our questionnaire.

Coolican (1999) recommended that data produced using the Likert scale method is best treated as ordinal type of data. To describe the average of the score, Coolican (1999) further recommended using the median score because the mean score should only be used for the interval level data. Therefore, in this thesis, analysis of the questionnaire data was based on the median score. Prior to analysis, the scale of the negatively worded items [Q2(i), 2(iii), 3(i), 3(ii), 3(iv) ] in the post-test questionnaire were reversed to enable comparability among items. In order to determine if there was a significant difference between groups, however, a repeated measure ANOVA was carried out on the data.

In the following sub-sections, the room size estimations results are first reported, followed by the spatial memory test result and post-test questionnaire.

7.1.3.2 Room size estimation

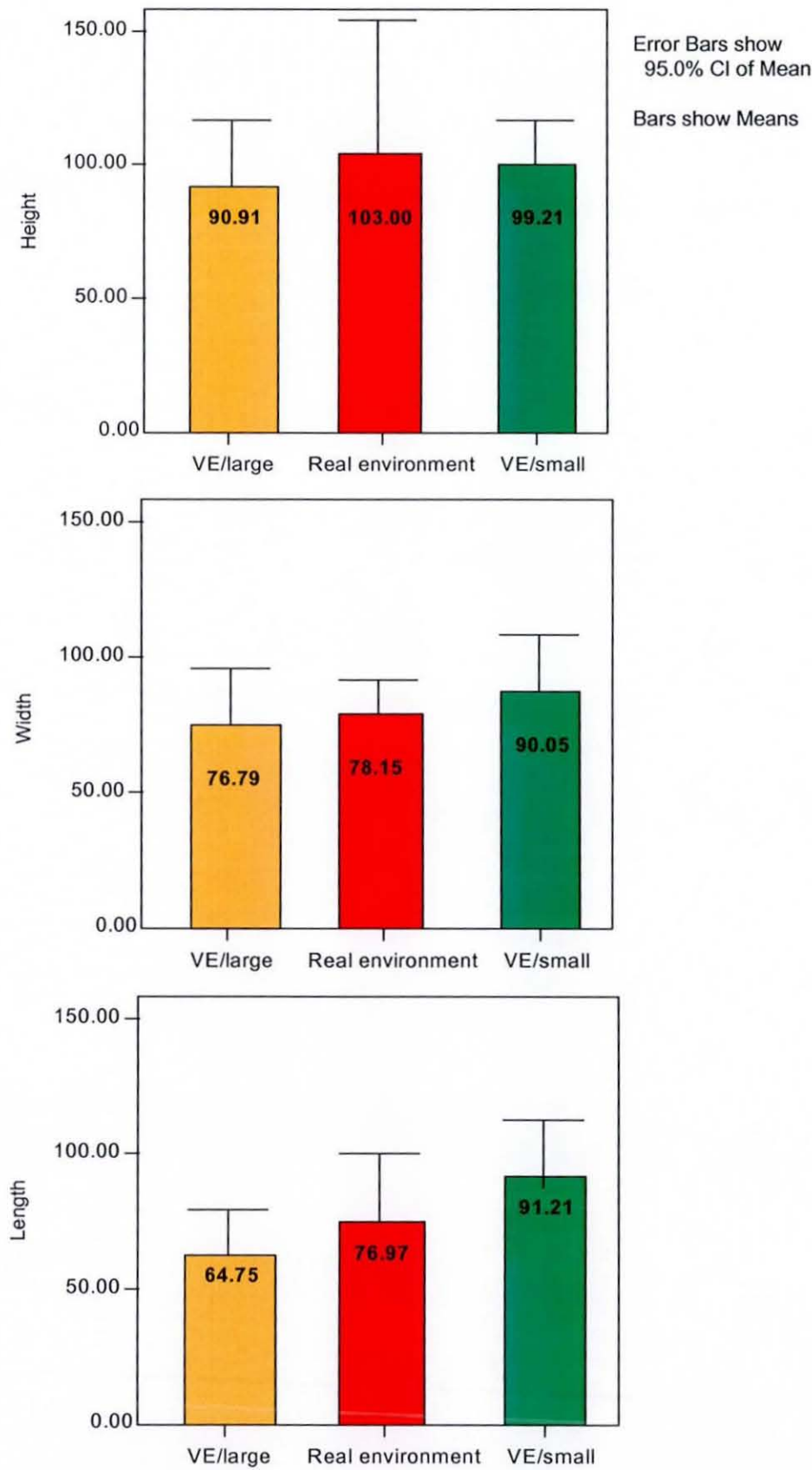


Figure 7-11 Comparison among real and VE conditions for height, width, and length



Figure 7-11 shows a comparison of mean percentage of estimate scores among display conditions (real, VE/large and VE/small) for each of the asymmetrical distances (height, width and length). For the height, an overestimation was revealed for the real condition while both VE conditions showed an underestimation. From the graphs (Figure 7-11), it was shown that performance on the VE/small condition tended to be better than on the VE/large condition for all asymmetrical distances. Additionally, for the width and length, estimates on the small display tended to be better than on the real and VE/large conditions. Also from the graphs it was shown that the VE/large condition yielded the worst performance.

### Analysis based on all experimental conditions

The results of MANOVA analysis revealed that the effect of display condition on distance estimates was not significant ( $F(6,48)=1.047$ ,  $p=.408$ ). This indicates that the difference among display conditions for distance estimation task was not large enough of practical significance. When each of the asymmetrical distances was considered separately, the results also showed no significant different among display conditions (Real, VE/large and VE/small):

- Height :  $F(2,28) = .253$ ,  $p = .778$
- Width :  $F(2,28) = .850$ ,  $p = .439$
- Length:  $F(2,28) = 2.553$ ,  $p = .098$

For the length distance, the difference among display types was significant at 10%. A post-hoc comparison revealed that for the length distance, the significant difference was between the VE conditions only; other comparisons revealed no significant difference.

### *The influence of sport variable*

A secondary multivariate analysis (ANCOVA) on these data which included the *sport* variable (that is participants' sporting background) as a covariate revealed similar results (Table 7-4). This indicates, even with the influence of *sport* variance removed, the difference between display conditions was still small. The effect of *sport* variable on this data set was not significant.

**Table 7-4 Summary of effect of sport variable and display on distance estimate**

Source	Dependent Variable	F	Sig.
SPORT VARIABLE	height	.428	.519
	width	.365	.551
	length	.714	.406
DISPLAY	height	.137	.873
	width	.754	.481
	length	2.169	.136

In hindsight, the time allocated for viewing the room for the real and VE was actually different. Even though, the viewing time was limited to 3 minutes for both environments, due to the repeated measure design of the experiment, the VE participants viewed the VE model for four times (a total of 12 minutes). Thus, the non-significant difference between the real and the VE conditions may be due to the long viewing time which may have resulted in improved performance for the VE participants. It is well-known that more practice results in improved learning (Stanney, Mourant et al. 1998). As indicated by some researchers, more experience in the VE might improve performance on route's findings, direction and relative distance estimate accuracy (Ruddles, Paynes et al. 1998). In contrast, some researchers found that experience only improves landmark direction but not on distance estimation accuracy (Allen and McDonald 1997). However, our current finding provides support for the results of the former study.

### Analysis based on VE conditions only

Because of the difference in viewing time between the real and the VE conditions, a second MANOVA analysis was conducted on the VE conditions data only. The results however still showed that the effect of display on distance estimates did not reach significant level ( $F(3,18)=1.845$ ,  $p = .175$ ). Similar to earlier analysis there was no significant difference between display size for the height and the width distances (Height:  $F(1,21) = .349$ ,  $p = .561$ ; Width:  $F(1,21)=1.275$ ,  $p=.272$ ). However, for the length distance the difference between display sizes reached significant level ( $F(1,21)=4.969$ ,  $p=.037$ ). This indicates that for the length distance, there was a difference in distance estimates between the large and small display.

### *The influence of sport variable and computer game variable*

When the influence of *sport* variable and *computer games* variable (that is how often participants play computer games) were removed, an ANCOVA analysis revealed no significant effect of display size on overall distance estimates ( $F(3,16) = 1.074$ ,  $p=.388$ ). The analysis also indicates that there was no significant difference between the large and the small display for the height, width and length [Height: $F(1,21)=.015$ ,  $p=.903$ ; Width: $F(1,21)=.372$ ,  $p=.550$ ; Length: $F(1,21)=2.605$ ;  $p=.124$ ].

A comparison of the adjusted mean scores revealed that the estimates on a small display were better than on a large display especially for the width and length distances (Table 7-5).

Table 7-5 Comparison of adjusted mean score between display size for all distance type

Dependent Variable	display size	Mean	Std. Error
height	large display	94.147	10.147
	small display	95.972	10.147
width	large display	79.703	8.412
	small display	87.134	8.412
length	large display	68.529	8.084
	small display	87.435	8.084

The effects of the *sport* variable and *computer games* variable were not significant on the distance estimates data [*sport* variable:  $F(3,16)=.175, p=.899$ ; *computer games* variable:  $F(3,16)=1.217, p=.336$ ].

Comparison among asymmetrical distances

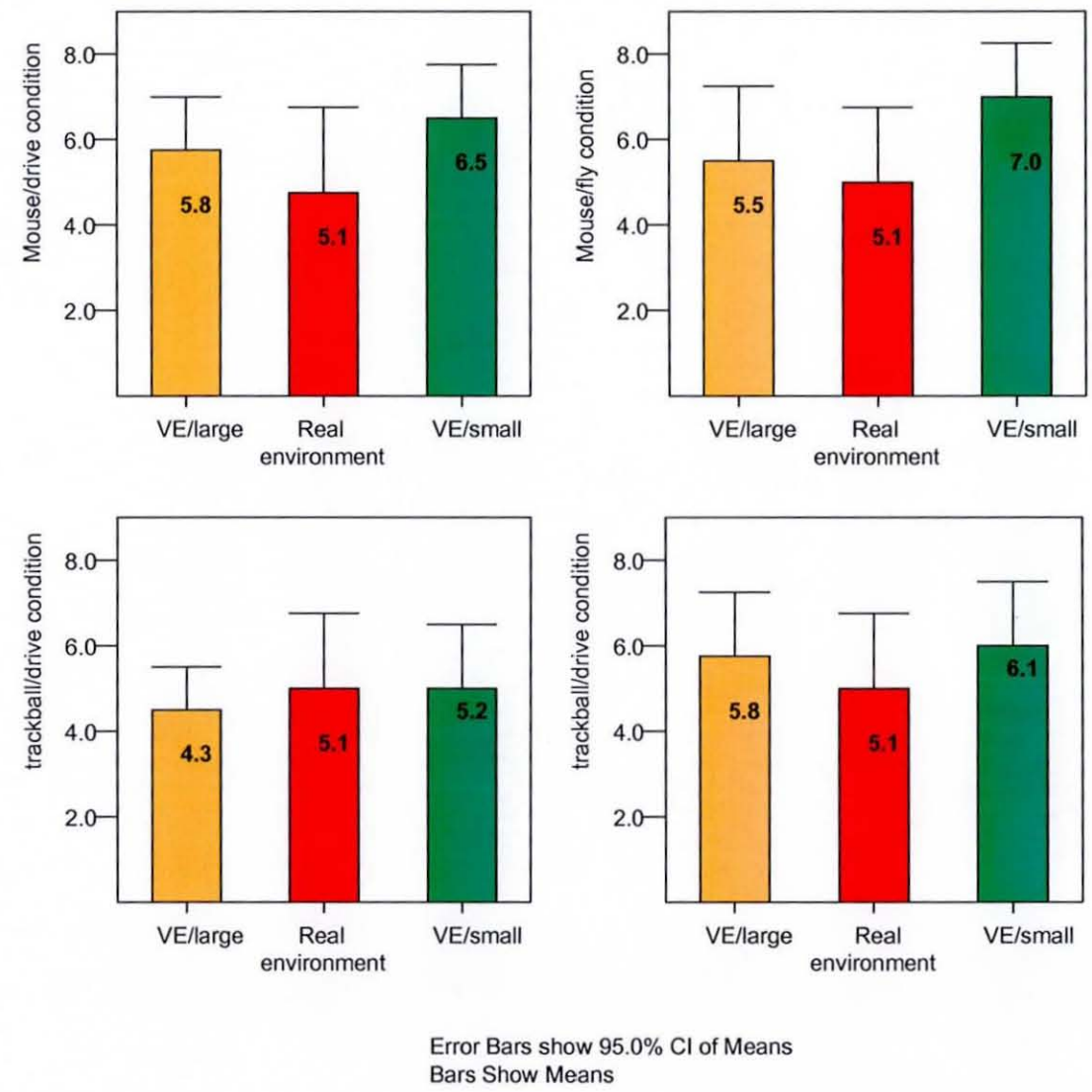
Similar to the results of Experiment 1 and 2, overall, the height distance was estimated more accurately compared to the width and the length distances (Mean percentage for Height = 97%, Width = 81.66% and Length = 77%). As mentioned earlier these distances correspond respectively to vertical, horizontal and transverse distances. The results of the Student t-test comparisons revealed that the difference among asymmetrical distances was highly significant ( $t < .0000$ ) which provide support for previous investigations (Henry and Furness 1993, Loomis, Da Silva et al. 1996) and the results of Experiment 1 and 2.

7.1.3.3 Spatial memory data (map test data)

Table 7-6 Mean, standard deviation and standard error for all conditions among display types

		N	Mean	Std. Deviation	Std. Error
mouse-drive	Real	10	5.100	2.3310	.7371
	Large	12	5.750	2.0505	.5919
	Small	12	6.500	1.9771	.5708
	Total	34	5.824	2.1245	.3643
mouse-fly	Real	10	5.100	2.3310	.7371
	Large	12	5.500	2.5045	.7230
	Small	12	7.000	1.6514	.4767
	Total	34	5.912	2.2746	.3901
trackball-drive	Real	10	5.100	2.3310	.7371
	Large	12	4.333	1.8257	.5270
	Small	12	5.167	1.9462	.5618
	Total	34	4.853	2.0020	.3433
trackball-fly	Real	10	5.100	2.3310	.7371
	Large	12	5.750	2.2613	.6528
	Small	12	6.083	2.3916	.6904
	Total	34	5.676	2.2926	.3932

Table 7-5 shows the mean, standard deviations and standard error for all experimental conditions (mouse/drive, mouse/fly, trackball/drive and trackball/fly) among the display conditions of real, VE/large and VE/small for the spatial memory data which were based on the number of correctly placed objects. No consistently very high or very low number of correctly placed objects (ceiling or floor effects) indicated that the number and the type of objects used in the spatial memory test were not too difficult or too easy for the participants (Johnson and Stewart 1999). This provides us with more confidence with our test method.



**Figure 7-12 Comparison among real environment, VE/large and VE/small conditions in terms of the mean number of object correctly placed for all interface device/travel conditions**

In Figure 7-12, results from the real world condition was compared with each of the experimental VE. From the figure, the real world results showed the percentage of correctly placed item was about 56.66% (that is 5.1 out of the total of 9 items).

A direct comparison of the mean scores revealed that for the mouse/drive, mouse/fly and trackball/fly conditions the VE participants (large and small display) outperformed the real world participants. However, for the trackball/drive condition, the real world participants did not differ very much from the VE/small participants but both scores were higher than the VE/large participants. For all conditions, the number of correctly placed objects was higher on VE/small display compared to VE/large display.

While the graphs in Figure 7-12 indicate differences among display conditions (Real, VE/large, VE/small), however, a one-way ANOVA analysis revealed that these differences were not significant for all conditions (Table 7-7).

**Table 7-7 Results of one-way ANOVA for comparison among groups (Real, VE/large and VE/small ) for all conditions**

CONDITIONS	F VALUE	P VALUE
Mouse/drive	F(2,33) = 1.211	p = .312
Mouse/fly	F(2,33) = 2.393	p = .108
Trackball/drive	F(2,33) = .613	p = .548
Trackball/fly	F(2,33) = .496	p = .614

**Table 7-8 Post-hoc comparisons among display types**

Dependent Variable				Mean Difference (I-J)	Std. Error	Sig.
		(I) display code	(J) display code			
mouse-drive	Games-Howell	Real	Large	-.650	.9454	.774
		Small	Real	1.400	.9323	.314
			Large	.750	.8223	.639
mouse-fly	Games-Howell	Real	Large	-.400	1.0325	.921
		Small	Real	1.900	.8778	.109
			Large	1.500	.8660	.219
trackball-drive	Games-Howell	Real	Large	.767	.9062	.680
		Small	Real	.067	.9268	.997
			Large	.833	.7703	.535
trackball-fly	Games-Howell	Real	Large	-.650	.9846	.789
		Small	Real	.983	1.0099	.602
			Large	.333	.9501	.935

Similarly, based on the more conservative Games–Howell assumption of unequal variance, none of the post-hoc comparisons among the display conditions reached significant level (Table 7-8). This means the difference between display conditions (real, VE/large, VE/small) was considered small. The only comparisons to reach statistical significant was between the mouse/fly and the trackball/drive conditions on a small display (see Table 7-9 (a) and (b)).



**Table 7-9 Comparisons among experimental conditions**  
(a) Large display condition

**Spatial memory test – Paired Samples Test on Large display**

	Paired Differences			t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean			
mouse-drive - mouse-fly	.250	3.1370	.9056	.276	11	.788
mouse-drive - trackball-drive	1.417	2.9375	.8480	1.671	11	.123
mouse-drive - trackball-fly	.000	2.0889	.6030	.000	11	1.000
mouse-fly - trackball-drive	1.167	2.6912	.7769	1.502	11	.161
mouse-fly - trackball-fly	-.250	3.7203	1.0740	-.233	11	.820
trackball-drive - trackball-fly	-1.417	3.3428	.9650	-1.468	11	.170

(b) Small Display condition

**Spatial Memory test --- Paired Samples Test on small display**

	Paired Differences			t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean			
mouse-drive - mouse-fly	-.500	3.0896	.8919	-.561	11	.586
mouse-drive - trackball-drive	1.333	2.6742	.7720	1.727	11	.112
mouse-drive - trackball-fly	.417	3.3428	.9650	.432	11	.674
mouse-fly - trackball-drive	1.833	2.4802	.7160	2.561	11	.026
mouse-fly - trackball-fly	.917	2.1933	.6332	1.448	11	.176
trackball-drive - trackball-fly	-.917	2.9683	.8569	-1.070	11	.308

The effect of device types, travel modes and display sizes

To compare the effect of device types, travel modes and display sizes only the data from the VE conditions were used. The reason for the exclusion of the real data was comparability. As mentioned earlier in the previous section, the VE participants spend more viewing time in the VE compared to the real world participants (12 minutes in the VE and 3 minutes in the real condition). As such the VE participants have more practice compared to the real participants and yielded better performance (as shown in Figure 7-13). Thus, for the next analysis, data from the real condition was not included.

The results of a repeated measure ANOVA showed that there was a main effect for device type but not for travel mode [device type:  $F(1,22) = 5.839, p=.024$ ; travel mode:  $F(1,22) = 2.364, p=.138$ ].

The overall score showed that the use of a mouse device yielded a more accurate result compared to a trackball [Mouse: mean score = 6.188, Trackball: mean score = 5.33].

While not significant, using a fly mode tended to yield more accurate object recall scores compared to using a drive mode [Drive mode: mean score = 5.438; Fly mode: mean score = 6.053].

An examination of interaction effects revealed that none of them reached significant difference (that is,  $p > .05$ ).

The effect of display sizes approached significant level:  $F(1,22) = 3.732$ ,  $p = .066$ . Similar to earlier analysis a comparison mean scores revealed that performance on a small display performance was better than on a large display [small display: mean score = 6.188; large display: mean score = 5.333].

### ***The influence of sport variable and computer games variable***

When ANCOVA analysis was performed by including the *sport* variable and *computer games* variable as covariates, the effect for display size was significant [ $F(1,20)=4.726$ ,  $p = .042$ ]. Therefore, when the influences of these variables were removed, there was a statistically significant difference between the display sizes. This indicates that besides display size factor both covariates did contribute an influence on spatial memory tasks.

Results showed that performance was better on a small display compared to a large display [Large display: Adjusted Mean score = 5.267; Small display: Adjusted Mean Score: 6.254]. The effect of the *sport* and *computer games* variables did not reached significance level [*sport* variable:  $F(1,20)=2.392$ ;  $p=.138$ ; *computer games* variables:  $F(1,20)=.016$ ,  $p=.899$ ].

Similar to earlier analysis, a pair-wise comparison on the adjusted mean scores revealed a significant difference for interface device ( $p=.029$ ) but not for travel mode. For the interface device, the adjusted mean scores revealed performance using a mouse device was better than using a trackball [Mouse: Adjusted mean score=6.188; Trackball: Adjusted mean score=5.333]. While not statistically significant, a comparison of the adjusted mean scores indicated that the fly mode yielded more accurate results compared to the drive mode [fly mode=6.080; drive mode=5.38]. No other effects were statistically significant. A summary of the mean and standard errors for all conditions is given in Table 7-10.

Table 7-10 Summary of mean and standard error

DISPLAY SIZE \* DEVICE \* TRAVEL MODE

Measure: MEASURE\_1

display size	DEVICE	TRAVELM	Mean	Std. Error
VE/large display	1	1	5.790	.615
		2	5.407	.634
	2	1	4.346	.569
		2	5.523	.666
VE/small display	1	1	6.460	.615
		2	7.093	.634
	2	1	5.154	.569
		2	6.310	.666

*The influence of practice time and map-test time*

To examine the influence of the practice time and the map-test time, a separate univariate analysis was conducted on each of the experimental conditions by including both the time variables as covariates in the analysis. A summary of the results of the analysis is presented in Table 7-11.

Table 7-11 Summary of analysis on the effect of practice time and map-test time

CONDITIONS	DISPLAY	PRACTICE TIME	MAP TEST TIME
Mouse/drive	$F(1,24)=.645, p=.431$	$F(1,24)=.073, p=.790$	$F(1,24)=.006, p=.940$
Mouse/fly	$F(1,24)=2.669, p=.118$	$F(1,24)=.014, p=.908$	$F(1,24)=.000, p=.987$
Trackball/drive	$F(1,24)=1.018, p=.325$	$F(1,24)=.013, p=.911$	$F(1,24)=2.761, p=.112$
Trackball/fly	$F(1,24)=.463, p=.504$	$F(1,24)=.707, p=.410$	$F(1,24)=5.755, p=.026$

From the table, the results showed that there was no significant difference between the display sizes for all conditions for spatial memory data except for the trackball/fly conditions whereby the effect of map-test time was significant indicating that the map-test time constituted a variance on the spatial memory data score. In fact, the map-test time explained 22.3% (partial eta squared value multiplied by 100) of the variance on the spatial data score for the trackball/fly condition. From Figure 7-13, the map test time was comparatively higher on a large display compared to on a small display for other conditions; the reverse was true for the trackball/fly condition. However, for practice time, the difference between the large and the small display was small.

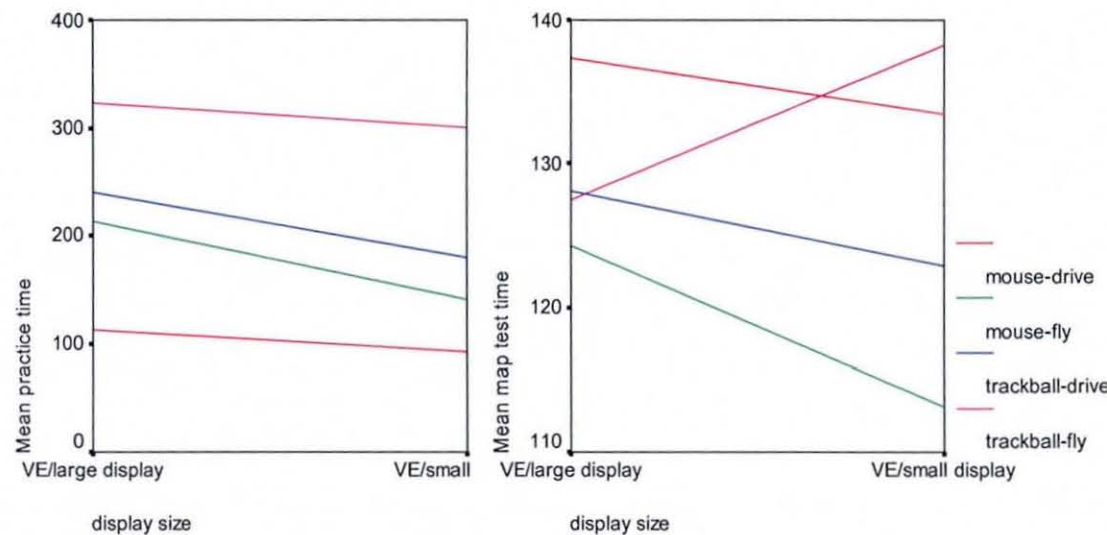


Figure 7-13 Comparison of practice time (left) and map test time (right) among conditions

### 7.1.3.4 Interface device questionnaire

The Interface device questionnaires were administered to the VE groups only: VE/large display and VE/small display. The first part of the Interface device questionnaire gathered information on the participants' familiarity with the interface device and revealed that most of the participants (23 out of 24) used the mouse at least once a day; only one person used it once a week. This supports the general beliefs that a majority of the population are familiar with the mouse device. However, for the trackball, all participants either hardly used (14 out of 24) or never used the device at all (10). The second and third part of the Interface device questionnaire asked to participants to evaluate the travel modes and the device types respectively. These results are described next.

### Travel modes

#### *Question 2(i) Ease of movement in the VE*

From Figure 7-14 (Q2i), the ease of movement in the VE was rated progressively better from trackball/fly to mouse/drive condition. Overall, drive mode using a mouse helped participants moved easily in the VE compared to other conditions, while fly mode using a trackball was rated the most difficult among the four conditions.

In both devices, participants found the drive mode helped them to move easily than the fly mode. These results were similar for both displays size conditions.

A direct comparison revealed that the mouse/drive and the trackball/fly conditions were equally rated on both display size conditions. However, for the mouse/fly and trackball/drive

conditions, both were rated easier to move in the VE on a large display compared to on a small display. Overall, a mouse device was rated easier to move in the VE than the trackball device.

### ***Question 2(ii) Control of movement in the VE***

Participants found that the drive mode allowed them to control their movement in the VE better than the fly mode (Figure 7-14 - Qii). Overall, the participants reported that the mouse/drive condition offered more control of their movements in the VE compared to other conditions. Participants found the trackball/fly mode condition as the most difficult to control movement in the VE compared to other conditions. The results were similar for the large and small display conditions.

Participants did not differ on their rating for the large and small display for the mouse/drive condition. However, for other conditions, more control over movement was afforded by the large display compared to the small display. Overall, a mouse allowed more control of user's movement in the VE than a trackball.

### ***Question 2(iii) Ease of recall for object positions***

The participants rated the fly mode better than the drive mode for the mouse device on both display on ease of recall for object position (Figure 7-14 -Q2iii). For the trackball on a large display both travel modes were rated equally but on a small display the participants tended to rate the fly mode higher.

Overall, for the drive mode, the mouse device was rated better than the trackball. Similarly for the fly mode, the mouse was rated better than the trackball. The results were similar for the large and small display conditions.

A direct comparison revealed that for all conditions (except for the trackball/fly condition), the large display condition was rated easier to recall object compared to the small display condition. Overall, the mouse device allowed participants to recall objects' positions more easily than the trackball device.

### ***Question 2(iv) Usage preference***

As indicated in Figure 7-14 (Q2iv), overall, the mouse/drive condition was preferred most compared to other conditions, whilst the trackball/fly condition was the least preferred. A direct comparison of the median score revealed that the mouse/drive condition was equally



rated on the large and small display conditions. However, for other conditions, the participants tended to rate the large display condition better than the small screen condition.

Overall, participants preferred the drive mode to the fly mode in the mouse condition. In the trackball condition, this was only true on a large display; on a small display both travel modes were rated equally. Generally, from the graph (Figure 7-14(Q2iv)) the ratings for the mouse conditions were higher than the trackball conditions.

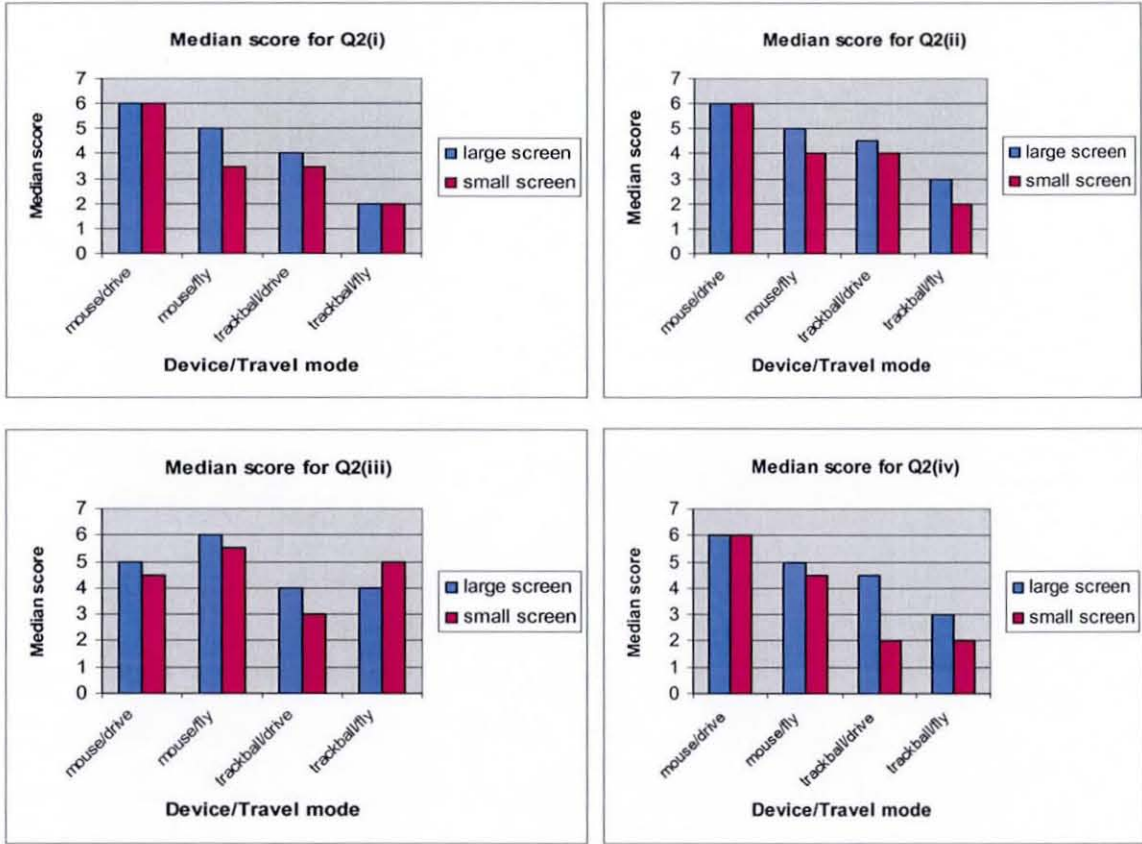


Figure 7-14 Median score for Q2i, Q2ii, Q2iii and Q2iv of interface device questionnaire

A repeated measure ANOVA was performed on Question 2 to test the significant difference between groups. A summary of the main effects and the interactions effects is presented in the Table 7-12. From Table 7-12, it was indicated that the difference between the devices was highly significant for all questions (Q2 (i) – Q2 (iv)). Similarly for travel modes, the difference between the drive mode and the fly mode was highly significant (with the exception of Question 2(iii), the value approaches significance). However, participants' ratings between displays did not reach significant level ( $p > .05$ ) for all questions. Other interaction effects were also reported as not significant for all questions.

**Table 7-12 Summary of main effect and interaction for Q2(i) –(iv)**

QUESTION NO.	EFFECTS	F	P VALUE
Q2(i) Ease of movement in the VE	Device Travel	F(1,22) = 33.366 F(1,22) = 25.858	.000 .000
Q2(ii) Ease of control in the VE	Device Travel	F(1,22) = 10.700 F(1,22) = 9.18	.003 .006
Q2(iii) Ease of recall for object position	Device Travel	F(1,22) = 20.613 F(1,22) = 4.081	.000 .056
Q2(iv) Usage preference	Device Travel	F(1,22) = 18.873 F(1,22) = 8.475	.000 .008

## Interface devices

### ***Question 3(i) Ease of use***

From the graph (Figure 7-15 - Q3i)), it was shown that the participants rated the mouse better than the trackball. Participants tended to rate the devices similarly in either display size, suggesting that the display size factor did not influence participants' preference for devices in terms of ease of use.

### ***Question 3(ii) Ease of movement and positioning in the VE***

With regards to moving and positioning themselves in the VE, participants again rated the mouse better than the trackball (see Figure 7-15 – Q3ii). However, a direct comparison revealed that the mouse on a large display was rated better than on a small display. In contrast, the trackball was rated similarly on both display sizes.

### ***Question 3(iii) Control of movement***

The participants found that the mouse allowed them to have more control of their movements than the trackball in the large and small display conditions (see Figure 7-15 – Q3iii). However, the mouse was rated similarly on both display sizes but for the trackball, ratings on a small display were only slightly higher than on a large display.

### ***Question 3(iv) Ease of recall of object positions***

In both display conditions, the participants found that it was easier to recall objects' position when using a mouse than when using a trackball (Q3iv). Additionally, from the graph (Figure 7-15– Q3iv), it was indicated that it was easier to recall objects' positions on a large display for both devices.

**Question 3(v) Usage preference**

In terms of usage preference (Q3v), subjects preferred the mouse more to a trackball in both display size conditions. From the graph (Figure 7-15 – Q3v), the mouse was rated slightly better on a large display than on a small display. However, the trackball was rated similarly on both display sizes.

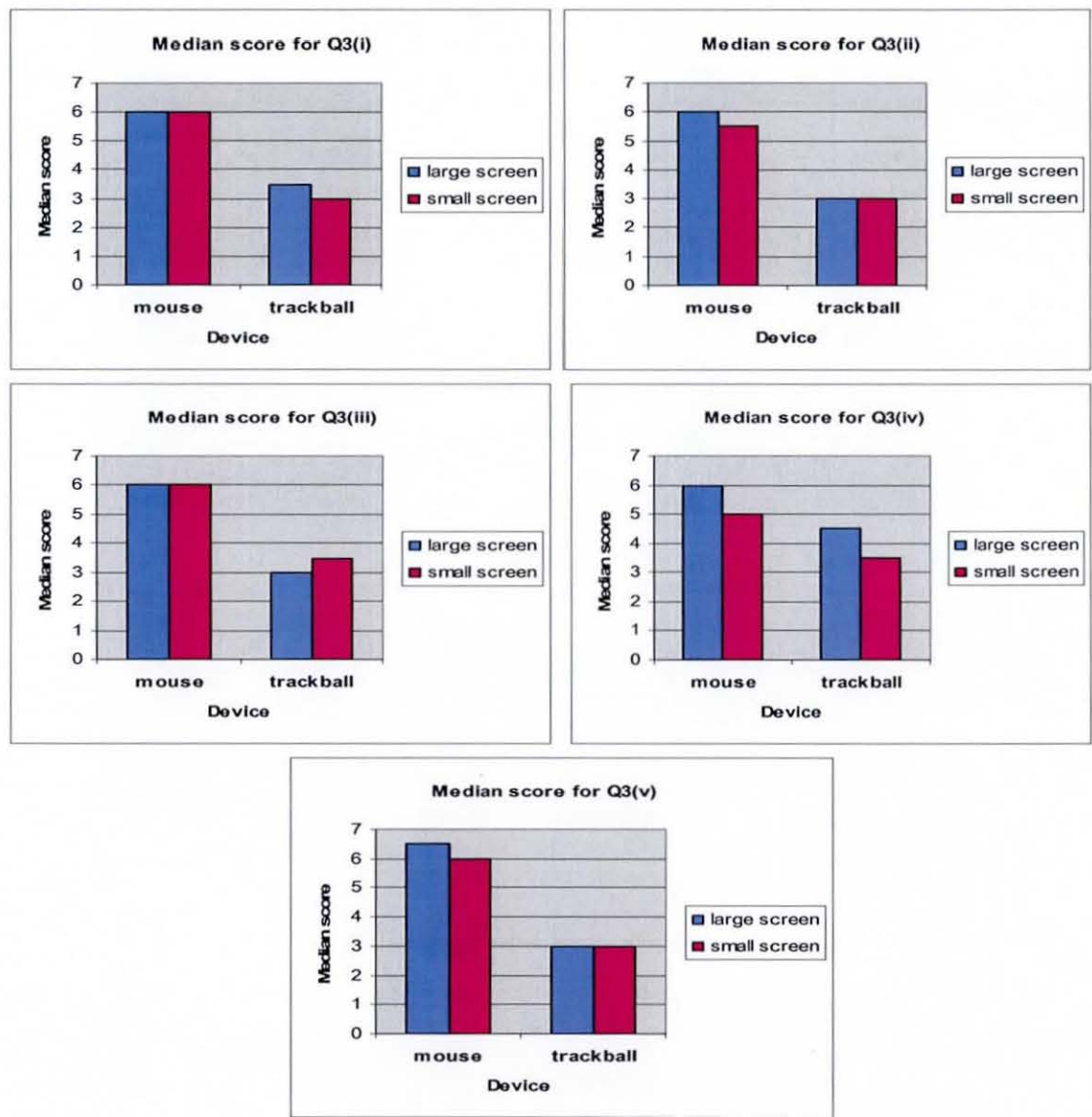


Figure 7-15 Median scores for Q3i, Q3ii, Q3iii, Q3iv, Q3v of interface device questionnaire

The results of a repeated measure ANOVA analysis on Question 3 are presented in Table 7-13. The effect of the device was significant for all questions. However, none of the interaction effect reached significant level. Participants’ ratings on both display sizes did not differ very much as indicated by the non-significant effect of the display size.



**Table 7-13 Summary of main effect and interaction for Q3(i)-(v)**

QUESTION NO.	EFFECTS	F	P VALUE
Q3(i) Ease of use	Device	$F(1,22) = 49.841$	<b>.000</b>
	Device*Display	$F(1,22) = .000$	1.00
	Display	$F(1,22) = .391$	.538
Q3(ii) Ease of movement and positioning in VE	Device	$F(1,22) = 29.59$	<b>.000</b>
	Device*Display	$F(1,22) = .041$	.842
	Display	$F(1,22) = .907$	.351
Q3(iii) Control of movement	Device	$F(1,22) = 11.803$	<b>.002</b>
	Device*Display	$F(1,22) = .201$	.658
	Display	$F(1,22) = .289$	.596
Q3(iv) Ease of recall of object position	Device	$F(1,22) = 13.070$	<b>.002</b>
	Device*Display	$F(1,22) = .011$	.919
	Display	$F(1,22) = 5.554$	.028
Q3(v) Usage preference	Device	$F(1,22) = 35.602$	<b>.000</b>
	Device*Display	$F(1,22) = .366$	.551
	Display	$F(1,22) = .185$	.671

### Accuracy rating (Q4)

When participants were asked to rate how accurate they were on the spatial memory test, the results showed that the participants' ratings tended to fall in the mid range (score of 3 - 4.5).

The results of ANOVA analysis revealed there was a main effect of device on accuracy ratings (Table 7-14).

**Table 7-14 Summary of main effects and interaction for Q4**

QUESTION NO.	EFFECTS	F	P VALUE
Q4 Accuracy rating	Device	$F(1,22) = 7.895$	<b>.010</b>
	Device*Display	$F(1,22) = 1.072$	.312
	Display	$F(1,22) = .013$	.910

Generally, participants tended to feel more confident of their recall when using a trackball than when using a mouse on both display size conditions. Whilst not statistically significant ( $p > .05$ ), the participants felt that they were more accurate on a large display compared to a small display. Comparatively, it was indicated that using a trackball on a large display gave better recall accuracy compared to other conditions; using a mouse on a small display, however, gave the worst rating.

### Familiarity with location

The purpose of the second section of the Interface device questionnaire was to gather information on whether the participants were familiar with the actual model itself. If so, they were asked to rate how much this prior knowledge of the actual room helped them in their recall. The survey results showed that only ten out of twenty-four participants were familiar with the location. When asked to rate how much their knowledge of the room helped them in

the object location recall test only two found that this knowledge helped them in their recall. The rest found that the knowledge did not greatly assist them.

### Participants' subjective responses for Interface device questionnaires

Participants were encouraged to write additional comments they have besides the rating they gave on each question in the questionnaire. Not all the participants provided additional comments, but those that did their comments are presented in Appendix C.

These comments further explain the participants' rating scores. Even though the participants found that the drive mode easy to control over fly mode, providing participants an overview of the room gave the fly mode an edge over the drive mode. Several reasons were given for participants' preference of the mouse over the trackball:

- The mouse was easy to use,
- The mouse required fewer fingers to control the buttons
- Participants more familiar with a mouse
- Participants felt more sense of movement when using a mouse.

For the trackball, the static position made it easy to control but the need to use three fingers (as opposed to two in a mouse) to control the buttons made it difficult to use. The rolling of the ball appeared to distract the user from the intended tasks, forget the buttons function and did not provide the user with the sense of movement because control was provided by the fingertips only. In support of the common belief, the participants found that recall was easier with more practice. Interestingly, one participant found that the type of travel mode affected his recall more than the type of device.

## 7.1.4 Discussion

In the next subsections, discussions of Experiment 3A results are presented. Discussions are organized into three separate subsections based on the results from the followings: distance estimate task, spatial memory task and interface device questionnaire.

### 7.1.4.1 Distance estimate task

#### Effect of Image types (Real and virtual conditions)

Results from the distance estimation tasks revealed that there was no significant difference among display conditions (real, VE/large and VE/small) for all asymmetrical distances. However, removing the effect of covariate (*sport* variable) did not change the results of non-significance. This result is inconsistent with the results of most previous investigations who



reported a significant difference between the real and VE (Henry and Furness 1993, Lampton, McDonald et al. 1995, Waller, Hunt et al. 1998, Witmer and Kline 1998, Sinai, Krebs et al. 1999, Willemson and Gooch 2002, Loomis and Knapp 2003).

However, the result of a very recent investigation (Plumert, Kearney et al. 2004) provides support for Experiment 3A's result. Plumert and colleagues' investigation revealed no significant difference between the real and VE on distance estimate tasks using verbal report, a method also employed in Experiment 3A. They attributed the similar performance between the real and VE conditions to the display type used. Previous investigations which found significant difference between the real and VE conditions compared VE presented on HMDs to the real world condition. In contrast, Plumert and colleagues compared a VE viewed on a large projected display to the real world condition. They argued that while recent investigations have ruled out the effect of restricted viewing condition of the HMD (Messing 2004, Creem-Regehr, Willemson et al. 2003, Knapp and Loomis in press) and the image rendering quality (Willemson and Gooch 2002, Thompson, Willemson et al. in press), one study indicates that the mechanical aspects of the HMD (mass and moments of inertia) were partially responsible for the inaccurate distance judgment in VE (Willemson, Colton et al. 2004). Therefore, they attributed the similar performance between the real and VE condition to the use of the projected display. A similar argument could be used to explain the result of Experiment 3A.

Alternatively, one possible explanation is the use of only a single room for the distance estimate tasks. Yoon, Byun et al. (2000) who compared a single real room to a VE model of it found similar results to Experiment 3A. In contrast, where multiple rooms (rooms - the museum gallery) were used for distance estimation tasks, Henry and Furness (1993) reported a significant difference in performance between the real and VE conditions. Therefore, similar to Yoon, Byun et al. (2000), the use of a single room might account for the similar performance between the real and VE conditions of distance estimate tasks. However, the room (real and VE) used by Yoon, Byun et al. (2000) was very simple and contained less details compared to the one used in Experiment 3A.

A more plausible explanation is practice effects. As mentioned earlier, the viewing time between the real and VE participants was different due to the repeated measure design of the experiment. The VE participants viewed the room for 12 minutes while real world participants viewed the room for 3 minutes only. The VE participants may have improved performance due to practice effects. It has been shown in the literature that increased

experience in the VE could improve participants' performance which includes distance estimation tasks (Ruddles and Paynes et al. 1998). While other researchers (Allen and McDonald 1997) indicate otherwise, given the widely accepted belief that practice improves learning (Stanney, Mourant et al. 1998) and findings from the other researchers (Ruddles, Paynes et al. 1998), it seems more plausible that more experience may have improved the VE participant's performance in Experiment 3A. Moreover, Lampton, Knerr et al. (1994) showed that participants were sensitive to practice effects.

While not significantly different, interestingly a direct comparison of the means score revealed that the VE/small display yielded more accurate results compared to the real and VE/large conditions for all asymmetrical distances. In addition to practice effects mentioned in the previous paragraphs, the details in the real condition may have imposed more cognitive demand on the real participant, thus degrading their performance (Yanagisawa and Akahori 1999).

### Effect of Display size (Large and Small)

The results of the analysis revealed that there was no significance difference between display sizes for all asymmetrical distances (except for the length estimates). Removing the influence of *sporting background* and *computer games* variables however did not change the picture. No significant difference was shown for all asymmetrical distances. This indicates that the influence of both variables were minimal on this data set.

It should be noted that the experiment setup for Experiment 3A was similar to Experiment 1B and 2B where the viewing distance and physiological cues were fixed and the FOV of the display were varied. It was argued (in Chapter 4) that the better performance of a large display over a small display for most previous investigations may not be attributed to the display size alone, other factors (such as viewing distance and physiological cues) which were not controlled may also account for the variance.

The experiment setup for previous investigations was similar to Experiment 1A and 2A. By comparing the results of these previous investigations and using similar arguments in Experiment 1B and 2B, the non-significant difference between the large and small display confirmed prediction that other factors (such as viewing distance and physiological cues) beside display size may have contribute to the main effect of display size. However, a direct comparison of means revealed that estimates on a small display was more accurate compared to the large display for all asymmetrical distances. This was unexpected given the results of

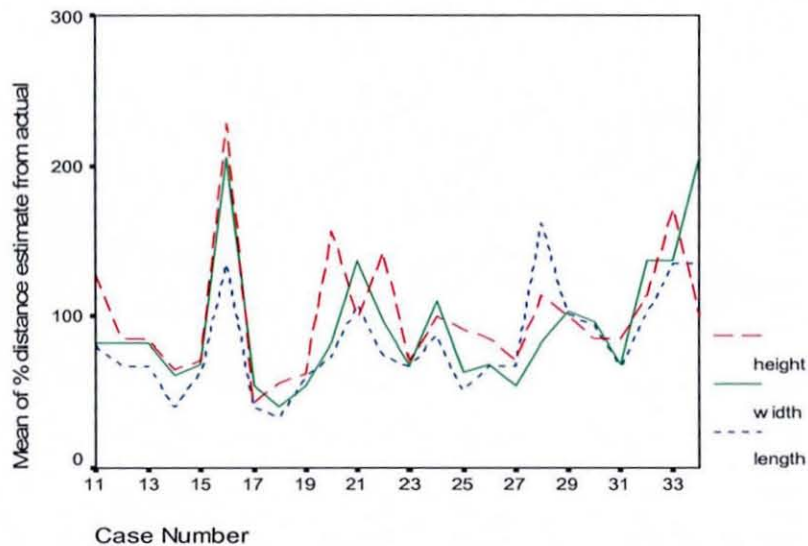
previous findings (Kline and Witmer 1996, Patrick, Cosgrove et al. 2000, Czerwinski, Tan et al. 2002, Tan, Gergle et al. 2003) and theoretical considerations.

In Experiment 3A, the FOV of the display was varied: a large display with large FOV and a small display with a small FOV. It has been suggested that a wide FOV would provides a user with more sense of presence (“being there”) (Prothero and Hoffman 1995) and more sense of realism (Hatada and Sakata 1980, cited in Pfautz 2002, Arthur 2000, Duh, Lin et al. 2002) in the VE. Larger FOV which closely matches the human FOV may yield similar performance to the real world (Prothero and Hoffman 1995). Thus, it would be expected that performance on a large display to be better than on a small display.

However, investigations by Arthur (2000) failed to show that reduced FOV influence the distance and memory tasks. He attributed these results to the large variability among participants, showing that the FOV had different effects on different participants. He also explained the non-significant effect of the FOV on these two tasks was attributed to his test methods. He suggested that a room size estimate or a matching size task would reveal a difference. However, the use of a room size estimate test method in Experiment 3A also did not reveal a significant difference.

In contrast, a study by Kline and Witmer (1996) that utilized a high-resolution non-head-tracked HMD display revealed that distance estimates on a large FOV display tended to be better than on a small FOV display. These authors also reported that the narrow FOV participants tended to overestimate distances which were similar to those found in Experiment 3A. However, the large FOV participants’ estimated distances were more accurate compared to those of the narrow FOV participants.

As reviewed in Chapter 4, previous studies which compared a large projection screen to a desktop monitor revealed better performance on large projection screen for spatial orientation task (Tan, Gergle et al. 2003), spatial memory task (Patrick, Cosgrove et al. 2000) and navigation tasks (Czerwinski, Tan et al. 2002, Tan, Czerwinski et al. 2003) but not on shape test and reading tasks (Tan, Gergle et al. 2003). Thus, another possible explanation is that for interactive images, the better performance of a large display over a small display is task dependent. However, it remains to be investigated why the small display was better than the large display as demonstrated by Experiment 3A. Large variability among participants as indicated by Figure 7-16 may have partially accounted for the results of Experiment 3A.



**Figure 7-16 Distance estimates for each participants in the VE conditions (Case 1-10 is real world conditions data which was not shown)**

### Comparisons among asymmetrical distances

For height estimates, on the average, the mean percentage estimate score for the real conditions were overestimated while for the VE conditions the scores were underestimated. This result provides support for the results of previous investigations (Yang, Dixon et al. 1999, Dixon and Proffitt 2002) whereby the height estimates in the real condition were overestimated compared to those in the images (in this case the VE conditions). Generally, estimates for the height distance was significantly more accurate compared to the width and length distance, with the length distance giving the least accuracy. These results supports the results of Experiment 1 and 2 (of Chapter 5 and 6 respectively) where the height, width and length estimates corresponded to the vertical, horizontal and transverse.

#### 7.1.4.2 Spatial memory tasks

##### Effects of Image Type (Real and virtual conditions)

Analysis of variance results revealed that there was no significant difference between the real and the VE for the spatial memory tasks for all experimental conditions. Some researchers (Arthur, Hancock et al. 1997) suggested that spatial representation resulting from interaction with a small scale VE is comparable to the real world experience. Corroborating these findings, Experiment 3A results revealed similar observations. Both studies shared some similarities in terms of experimental methods: the use of a single room, free navigation of the experimental room, participants to focus on object locations instead of names and the recall of nine object locations. However, in Experiment 3A the viewing time was restricted to 3

minutes while Arthur, Hancock et al (1997) gave their participants as much time they needed to explore the room. Given the small number of objects to recall, the single room and the fact that our participants repeated the experimental conditions four times should make little difference and should still make reasonable comparisons.

In a different study, Richardson, Montello et al. (1999) tested participants' acquisition of spatial representations of an environment via a map, a real and VE found that participants' performance using a simple single floor was similar for all conditions. However, the use of a complex building, the VE learners was shown to yield the worst performance. This implies that the use of a simple environment may yield similar performance between the real and VE participants. Therefore, the results of Experiment 3A provide support for previous works which indicate that it is possible to use VE to perceive spatial relation similarly to real world conditions when a simple environment was used.

Comparatively, the participants in the real condition physically walked in the room while participants in the VE used the mouse (trackball) to control user movements in the VE. While in this experiment, both the mouse and the trackball functioned to control users' viewpoints in the VE, Gaunet, Vidal et al. (2001) pointed out that active exploration with a joystick shared some important aspects with walking in the real world. The authors suggested that "there is a tight linkage between visual self-motion and motor-activity, just as in the real world." Thus, the process of gathering visual information was similar in both conditions. Similarly this may be true when the mouse and trackball were used; thus, provide explanation for the similar results between the real and VE conditions in Experiment 3A.

However, one important difference between the movements in the real and VE conditions was the presence of the proprioceptive cues. In the VE conditions, there was a mismatch between the visual/vestibular cues to the motion perceived. Visual cues (optic flow) from the display indicated there was motion but the vestibular cues indicated a stationary position (Richardson, Montello et al. 1999). This mismatch will often result in the user feeling nausea (May and Badcock 2002). Thus, it is expected that the VE participants may perform poorly compared to the real world participants. However, none of our participants reported such effects; this may be due to the use of the non-stereo and non-immersive/semi-immersive display. This nausea effect is often experienced by the users of stereoscopic and immersive display such as HMDs (Wan and Mon-Williams 2002). Moreover, some form of proprioceptive feedback was given by muscle movements of the wrist, arm and shoulder for the mouse device and fingers for the



trackball. This might alternatively explained the similar performance between the real and the VE conditions.

The similar performance between the real and VE conditions in Experiment 3A further suggested that the use of these input devices may be minimally sufficient to provide the proprioceptive feedback necessary to elicit the necessary information to indicate movement. This is consistent with the results of other researchers who reported that the use of a more natural walking interface which is similar to the real world performed no better than a conventional input device such as a joystick (Witmer and Kline (1998). A study by Grant and Magee (1998) also revealed that the presence of proprioceptive cues from the use of walking interface was not beneficial on an orientation task compared to the use of a joystick; though the walking interface did assisted on transfer of spatial knowledge. Moreover, the flexibility of the human sensory system might partially account for this effect. In fact, a slight movement of the head or without even physically moving, information from the visual sense is enough to provide the user with a compelling sense of movement (Harris, Jenkin et al. 2002).

A direct comparison of means revealed that, on average, the number of correctly placed objects in the VE was shown to be slightly higher than those of the real condition. One possible explanation for the slightly better performance of the VE participants over the real world participants is practice effects. As mentioned earlier in the previous section, in the VE conditions, participants get to view the VE model for four times compared to only once for the real world participants. It has been suggested that the acquisition of spatial knowledge (mental representation) of an environment is increased with an increase in exploration time and/or increase of the observer displacements (Peruch, Vercher et al. 1995). Thus, more viewing time and practice effects might have improved the overall performance in the VE for Experiment 3A.

Another possible reason is that the VE participants viewing area was confined to the screen so participants can focus on objects' locations. It has been suggested that viewing from a single orientation might yield more accurate results as it allowed user to focus on the spatial layout of objects (Arthur, Hancock et al. 1997).

In Experiment 3A, participants in the real conditions actually needed to physically move about in order to look for objects (some of the objects were small). This physical act of movement may have imposed greater mental demand on the real participants, thus less focus was available on the object locations. Moreover, the real participants might have spent more

time looking for objects than remembering where objects were located. With less time given compared to the VE participants, together this might influence the real world participants' slightly poor recall of object locations in the spatial memory test compared to the VE participants.

### Effects of Display size (Large and small)

Consistent with the results of Johnson and Stewart (1999) and Arthur (2000), Experiment 3A's results showed no significant difference between the large and small displays on spatial memory tasks. Johnson and Stewart (1999) studies revealed that their participants were equally able to develop spatial representation of a virtual heliport in a wide and narrow FOV HMD. Their participants scores were not significantly different in both HMDs (wide FOV= 76%, narrow FOV = 78 %). Similarly, Experiment 3A results showed that participants' scores did not differ significantly on a large and a small display. While not significantly different, similarly to the results of Johnson and Stewart (1999), comparatively for all conditions performance on a small display was slightly better than on a large display.

However, when the effects of covariates (*sport* and *computer games* variables) were removed from the data, results showed that the difference between the large and small display was statistically significant. Surprisingly, the recall of objects' positions was more accurate on a small display compared to a large display. With regards to the experimental setup, in Experiment 3A, the FOV of both display sizes was varied whereby the large display provided a large FOV and the small display provided a small FOV. It has been suggested that a wide FOV induces more sense of presence (Prothero and Hoffman 1995, Arthur 2000, Duh, Lin et al. 2002) and sense of realism (Duh, Lin et al. 2002). Inherently a large FOV would be similar to the human FOV compared to a small FOV. Additionally, the narrow FOV in a small display eliminates most of the peripheral vision which have been suggested necessary for the development of the survey knowledge (Alfano and Michel 1990). Therefore, it would be expected that the performance on a large FOV display would be better than on a small FOV display. However, the results of Experiment 3A were contrary to expectation. It was noted that the viewing distance and the physiological cues for both display sizes were fixed for Experiment 3A which was similar in construct to Experiment 1B and 2B in Chapter 5 and 6 respectively. From these studies it was shown that when the viewing distance was made similar, the results showed no significant difference between display sizes. The non-significant difference between display sizes found in Experiment 1B and 2B allowed us to conclude that the better performance on a large display over a small display in Experiment 1A and 2A were also attributed to the viewing distance and the physiological cues factors. As

initially intended, comparing Experiment 3A's results with the previous investigations (Kline and Witmer 1996, Patrick, Cosgrove et al. 2000, Czerwinski, Tan et al. 2002, Tan, Gergle et al. 2003) which employed similar experiment setup to Experiment 1A and 2A would allow us to draw similar conclusions. However, the unexpected results of Experiment 3A suggested that this is not possible. This is because instead of revealing a non-significant effect of display, Experiment 3A's results showed a statistically significant main effect of display. Moreover, performance on a small display was shown to be better than on a large display. Accepting this result suggested that the acquisition of spatial representation in interactive VE is not necessarily better using a large display.

A small display was more comparable to the more familiar desktop monitor. Since most participants were familiar with a small size display (desktop monitor), this might explain the better performance on a small display compared to a large display. Additionally, while a large display was more realistic and induces more sense of presence, however, similar to the real world situation the large display participants may require more mental effort compared to the small display participants. Therefore this may adversely affect the large display participants' performance. Moreover, on a large display, the textured carpet which comprises of fine details showed some signs of aliasing effects, where shimmering effects occurs when the user moves in the VE. These effects may have degraded the image quality of the objects in terms of visibility especially for small objects. Since all of the objects for the spatial memory test were located on the carpeted floor, these effects may have affected the user's view of the objects and positions on a large display. It was noted however, this aliasing effect was not prominent in a small display. Thus, the lower performance on a large display may be attributed to the aliasing effects of the carpet texture which affects the user's view of the objects and positions. As suggested in Chapter 3, one technique of reducing such effects was to use the MIP mapping algorithm where a set of texture maps of different resolution was used on the objects for different viewing distances. Thus, eliminating such effects might have improved user performance on large display.

As discussed earlier, the unexpected findings of the better performance of a small display over a large display indicated that it was not possible to compare the result of Experiment 3A results to the results of previous investigations. As such, a comparison to a study similar to Experiment 3A but similar in setup to previous investigations (or Experiment 1A and 2A) may yield reasonable comparison and provide useful information to understand this unexpected findings.

### Effects of Travel modes (Drive and Fly)

Whilst not statistically significant, the fly mode yielded more accurate spatial memory test results compared to the drive mode. The fly mode may be difficult to use due to the extra degree of freedom which may in turn incur more cognitive demand on the user (Ruddle and Jones 2001), however, by allowing the vertical up and down movement gave fly mode the extra advantage of having an overview of the whole room and overall object positions compared to the drive mode. Having an overview of the room is similar to a map view of all objects' locations, whereby the participants can see the overall objects' spatial relations which could translate easily to the spatial map test. Thus, this may explain the slightly better performance of the fly mode over the drive mode. Moreover, some researchers (Richardson, Montello et al. 1999) have shown that "maps are powerful for acquiring quick and accurate spatial knowledge."

### Effects of Interface devices (Mouse and Trackball)

For the spatial memory task performance, the results indicated there was a significant difference between device types with performance using a mouse was better than using a trackball. With respect to familiarity with the interface device, about 96% of the participants were familiar with a mouse and more than half had either hardly or never used a trackball before. This indicated that the better performance on a mouse device over a trackball could have been affected by the participants' prior experience with the mouse device.

In a study comparing a mouse/monitor condition to other immersive conditions (HMD/bike conditions), Tong, Marlin et al. (1995) demonstrated that the mouse/monitor participants was significantly better at object locations association compared to other conditions. Thus, Experiment 3A provides support for the results of Thong and colleagues in terms of better mouse performance on objects location test.

Moreover, as explained in Chapter 4, there was a difference between the mouse and the trackball in terms of proprioceptive cues provided. The mouse relies on muscles of the wrist, forearms, arms and shoulder while the trackball only uses the muscles from the fingers (and/or the palm of the hand). Thus, the mouse device provides more proprioceptive cues to the user. It has been suggested that the proprioceptive cues can provide powerful information of motion (Hlavacka and Mergner et al 1996, cited in Harris and Jenkin et al 2002).

Christou and Bulthoff's (1999) review of previous studies found that the participants' perception of the spatial layout was better for participants who received proprioceptive cues

from moving around than those who remained static and did not received this cues. Thus, this extra information (proprioceptive cues) might have affected user's performance when using a mouse in the VE and hence his spatial judgment.

From the participants' self-reported assessments, one participant commented that the mouse device allowed him to feel directly the translation of the movement of the mouse on the screen. But for trackball, rolling of the ball, gave him less sense of this feeling. Additionally, even though both devices were relative devices, the post-test questionnaire results showed that participants found the trackball as more difficult to use compared to the mouse.

It has been suggested a simple interface device allowed users to devote more of their cognitive resources towards the task at hand rather than on the device itself (Ruddles and Jones 2001). Thus, the mouse device must have allowed more time to focus on learning the objects' locations while the trackball participants spend more of their cognitive resources on the device itself. However, some participants commented that given more practice and frequent use, the trackball would be easier to use. Though not reliably different, participants' confidence rating of their accuracy of estimate (as shown by the post-test questionnaire) indicated more confidence of their recall accuracy when using a trackball even though the spatial memory test results suggested otherwise.

No main effect and interaction effect for all conditions (except in the trackball/fly condition) when the practice time and the map test time were included in the analysis as covariates. This indicates that both times explained insignificant variances in the spatial memory data with the exception of trackball/fly condition. For the trackball/fly conditions, on average, participants tended to spend more time on a map-test when on a small display compared to the large display. However, for other conditions this difference was very small. Similarly for practice time, participants did not differ very much on either display sizes.

#### 7.1.4.3 Interface device questionnaire

##### Effects of Travel modes

In terms of ease of movement in the VE, participants rated the mouse/drive the highest followed by mouse/fly, trackball/drive and trackball/fly. For both devices, participants found the drive mode allowed them to move easily than the fly mode. While both the large and small displays were rated equally for the mouse/drive and trackball/fly conditions. However,



for other conditions (mouse/fly and trackball/drive) the large display allowed the participants to move easily than on a small display.

A similar pattern of results to ease of movement was observed in terms of movement control in the VE and usage preference with some exceptions. For movement control, both display sizes was rated equally in the mouse/drive condition.

In terms of usage preference, both travel modes were rated equally on a small display for the trackball/fly condition.

In contrast for ease of object recall, most participants rated the fly mode better than the drive mode for a mouse on both display sizes. For trackball, fly mode was rated higher than drive mode on a small display but on a large display both drive and fly mode were ranked equally. For both travel modes, the mouse was rated better than the trackball in terms of ease of object recall.

It was noted that for all conditions in the interface device questionnaire, the differences between the device types and between the travel modes were statistically significant. However, for display sizes, the differences between the large and the small display were not significant for all conditions.

Drive mode provided the participants less degree of freedom (2D) compared to the fly mode (3D). The extra degree of freedom in the fly mode may have adversely influenced participants' ease of movements in the VE and control of movements in the VE. As suggested in the literature (Ruddles and Jones 2001), the extra degree of freedom may increase cognitive demand on the user.

The results of the spatial memory data were not consistent with the overall questionnaire results on travel modes as shown by the slightly better performance on a fly mode compared to a drive mode. As explained earlier, the extra degree of freedom gave the fly mode a map view advantage of the objects' locations which translated to better results in the spatial memory test. Thus, even though the participants found it difficult to move and control their movement in the VE using fly mode, however, using this mode yielded better performance on the spatial memory test. Consequently, this was also reflected in their choice of the fly mode over the drive mode in terms of ease of recall of objects' positions.

## Effects of Interface devices

The interface devices were evaluated directly in part 3 of the Interface device questionnaire. In general, most participants ranked the mouse device significantly higher than the trackball in terms of the following:

- ease of use
- ease of movement and self-positioning in the VE
- afford more control of movement
- ease of recall for object positions

Similarly, in terms of overall usage preference, the mouse device was preferred more than the trackball. For both devices, ratings on both displays were not statistically significant, indicating that both devices were rated similarly on both display sizes. This means display size did not affect participants' ratings.

The better rating of the mouse over the trackball agrees with the results of the spatial memory test. The ease of use, ease of movement and control afforded by the mouse might have influenced the participants' better performance. As suggested by Ruddles and Jones (2001), a simple interface allows users to devote more of their cognitive resources towards updating their spatial memory.

As reviewed in Chapter 4, some evidence indicated that it was easy to navigate a 3-D space using a mouse and some evidence indicated that users found it difficult to use a mouse for 3-D navigation. The results of Experiment 3A study provide support for the former studies. Additionally, the participants' familiarity with the mouse more than the trackball may further accounted for this result. As commented by one of the participants "As I am used to mouse, I find it easy to use. If I have been using the trackball it would have been easier too ...." In contrast, one study indicated no performance difference between a mouse and a trackball even though most participants were familiar with the mouse but not with the trackball (Mueller, Bliss et al, unpublished work). However, the task (tracking task) investigated was different from the current task. Several studies reviewed in Chapter 4 revealed that the mouse performed significantly better than other devices such as Polhemus tracker (Jacob and Sibert 1992) and HMD-bike (Tong, Marlin et al. 1995). A number of researchers have compared a wide range of input devices including the mouse and trackballs but on different tasks (Cohen, Meyer et al. 1993, MacKenzie, Sellen et al. 1991). However, when compared to trackball, the result of Experiment 3A revealed that the mouse yielded significantly more accurate result on spatial memory test.

### Accuracy rating

In general, when asked to rate the accuracy of their spatial memory test, participants tended to take the central score. As put forward by Coolican (1999), the central scores or 'on-the-fence' positions of the participants is one of the drawbacks of Likert scales method, where the scores reflect 'undecided answers'. As such these scores could not be interpreted conclusively. But participants' ratings on their accuracy were slightly higher on a large display compared to a small display.

On average, participants were least confident when using a mouse on small screen. Surprisingly, using a trackball on a large display was rated highest in terms of recall accuracy. However, these results contradicted the participants' results in the spatial memory test where participants' scores were significantly higher when using a mouse than using a trackball. Similarly, the spatial memory test results also showed that performance on a small display was better than on a large display.

In terms of prior knowledge of the location, less than half of the participants had prior knowledge of the room and out of this number only two participants found that their knowledge of the room had assisted them in their spatial memory test. However, an examination of their score did not revealed very accurate result. This was expected as the objects used for the spatial memory test were not objects originally present in the actual room. Additionally due its use as a common room, it was expected there would be some changes in the presence and absence of objects and their arrangement in the room and this make it slightly dissimilar from the VE model. However, it was stressed earlier that for the real condition, the objects and its spatial arrangement were made similar to the modelled VE prior to the conduct of the real world condition study.

### 7.1.5 Conclusions

Experiment 3A aimed to investigate user's spatial awareness in terms of distance estimates and spatial memory task in an interactive VE compared to similar tasks performed in the real condition.

Study results showed that the participants performed similarly on the real and VE condition on distance estimation tasks. Given the well-established findings that user underestimated distance significantly different from the real world, these findings were quite unexpected. However, the result of a very recent study (Plumert, Kearney et al. 2004) is consistent with Experiment 3A results. The extra viewing time (more than three fold) in the VE condition

compared to the real condition was also suggested to improve the VE participants' performance to be more comparable to the real participants.

It was shown that there was no significant difference between the large and small display for the height and the width distances. In contrast, there was a main effect of display for the length distance. However, the introduction of the *sport* variable and *computer games* variables in the analysis changed the picture; no main effect for all asymmetrical distances was revealed. While not statistically significant, the small display participants tended to perform better than large display participants. Similar to the results of previous investigations, the height distance was estimated more accurately compared to the width and length distances.

With regards to the spatial memory tests, the difference between the real and the VE for all experimental conditions was not statistically significant. As such it was possible to perceive spatial relations in the real and VE conditions similarly. With the exception of trackball/drive condition, generally the VE participants tended to perform slightly better than the real participants on the spatial memory test. Extra viewing time, practice effects and less demand on the cognitive resources in VE condition have been suggested to improve the VE participants' performance to be slightly better performance than those of the real participants.

Experiment 3A result also indicated that the difference between the large and small display on spatial memory test was not large enough to reach statistical significant for all experimental conditions. However, when the VE data was considered only in the analysis the main effect of display approached significance level and when the covariates (sport and computer games background) were introduced in the analysis this main effect reached statistical significance level.

On average, performance on the small display was slightly better than on the large display. Reduced image quality, due to the effects such as aliasing on the large display, may have degraded performance on the large display. The unexpected better performance of a small display over a large display indicated that it was not possible to make comparisons with the previous investigations' results.

Participants' performances on a spatial memory test were better when using a mouse than when using a trackball. Participants' familiarity with a mouse over a trackball and the extra proprioceptive cues derived from the movement of the mouse may have improved

performance over the trackball. Additionally, results from the post-test questionnaire also suggested that overall participants preferred the mouse to the trackball in terms of ease of use, ease of movement, control of movement and on object recall. Participants' self-reported comments that the movements of the mouse resulted in a better sense of corresponding visual movement on the screen compared to the trackball further explained the results.

For travel modes, the difference between the drive and fly modes on the spatial memory test was very small. On average participants performed slightly better using the fly mode. Consistent with the result of the spatial memory test results, participants tended to rate drive mode better than fly mode in terms of ease of movement and control of movement but in terms of ease of recall fly mode was rated higher than drive mode. The map view provided by vertical movement of the fly mode may have contributed to the slightly better performance of fly mode over drive mode. This implies that the more familiar method of movement of the drive mode does not necessarily resulted in a better performance in the VE. The unnatural movement of the fly mode could be more beneficial in the VE compared to the drive mode.

The non-significant effect of display size on the post-test questionnaire indicates participants' ratings were similar on both display size. This implies that the display size did not reliably influence participants' decision on the preference of travel modes and device types.

## 7.2 EXPERIMENT 3B EFFECT OF DISPLAY SIZE

The results of Experiment 3A suggested that there was no reliable difference between the display sizes on the participants' performance in the spatial memory test and distance estimate tasks but when the effect of covariates (sport background and computer games experience) were removed, the effect of display size became significant for spatial memory test performance. The main effect of display was unexpected.

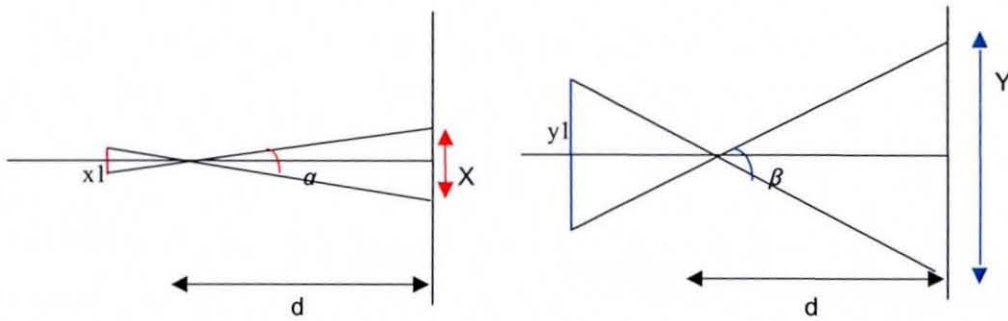
The setup of Experiment 3A followed those of Experiments 1B and 2B whereby the viewing distance and physiological cues were fixed and the FOV of display was varied. It was initially stated that the results of previous investigations which showed better performance on the large display over the small display would be compared to the result of Experiment 3A (see Section 4.1.2 of Chapter 4). When the FOV was fixed (and viewing distance and physiological cues were varied), previous investigators revealed a significant effect of display size. It was also initially argued that the better performance of the large display over the small display was not due to the physical display size alone, other factors (such as viewing distance and physiological cues) which were not controlled by these previous investigators may have



also explained the difference. Thus, when the viewing distance (and physiological cues) was fixed and the FOV was varied (as in Experiment 3A), the results of non-significant difference between the display sizes would allowed us to conclude that the better performance of the large display over the small display was not attributed to display size alone but was also influenced by the viewing distance and the physiological cues. However, the results of Experiment 3A showed a main effect of display which did not allow us to draw such conclusion. But these results may suggest that the viewing distance (and physiological cues) and FOV have marginal influence on the spatial memory test performance. This is because the study results (from Experiment 3A and previous investigation results) showed a main effect of display regardless of whether viewing distance (and physiological cues) and FOV were fixed or varied. This may suggest that the main effect of display in the previous investigations were largely contributed by the physical display size.

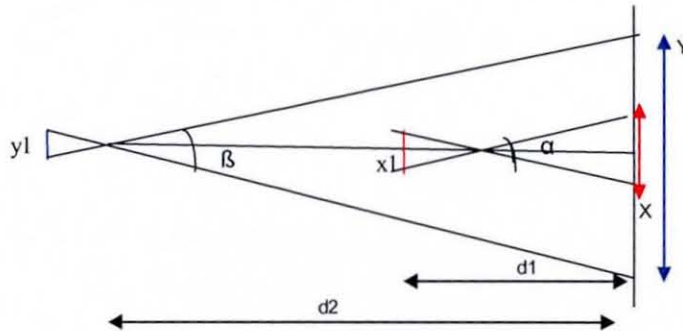
However, a direct comparison of the means indicated that the small display participants outperformed the large display participants. This was unexpected given the theoretical considerations and the results of previous investigations reviewed in Chapter 4 which suggested that performance on a large display would be better compared to a small display. Experiment 3A, however, did not allowed participants to compare between the display sizes. Thus it was necessary to conduct another experiment which gave the participants such opportunity. This would allow us to examine why the small display participants exhibited better performance over the large display participants in Experiment 3A. For the next experiment, a short display questionnaire (to be described later) was designed to compare participants' ratings on both display sizes. Examination of participants' ratings and self-reported comments may reveal more information on why performance was better on a small display compared to a large display. It was expected that asking participants to rate both display sizes on some criteria would provide more information on why performance is better on a small display compared to a large display.

Due to the differences in experimental methods and stimulus used in the previous studies, analysing the results of Experiment 3A with those of previous studies may not yield reasonable comparisons. Moreover, the results of better performance of a small display over a large display were unexpected. Thus the experiment setup for the next experiment (Experiment 3B) followed the setup of the previous studies and those of Experiment 1A and 2A (see Figure 7-17).



**Figure 7-17 Experiment 3A setup.** Left diagram represent small display setup and right diagram represent the large display setup

Thus, for Experiment 3B, the FOV of the display was fixed and the viewing distance was varied (see Figure 7-18). Similar to the results of the previous investigations, it was expected that there would be a significant difference between the large and the small display condition if display size were to contribute significantly towards spatial memory and distance task performances.



**Figure 7-18 Experiment 3B setup.** Retinal image size ( $x1 = y1$ ). FOV ( $\alpha = \beta$ ). Viewing distance ( $d1 \neq d2$ ). Physical image size ( $x \neq y$ )

A full within-subjects design was employed whereby each participant experienced all experiment conditions. In addition to the device types and travel modes, this approach allowed the same participant to experience both display sizes and to make comparison between them. Thus, the display size, device type and travel modes were within-subject factors.

Experiment 3B was similar to Experiment 3A in terms of material/apparatus and experimental procedure, with the following exceptions:

- Experiment design
  - As mentioned earlier, a within-subject design was employed. In Experiment 3A, only the device and travel mode factors were within-subject factors while the display size was a between-subject factor. In the current study, all these factors were within-subject factors in which all participants will experience all experimental conditions. As in Experiment 3A, a counterbalanced design was employed to remove the order effect.

- Period of experimental trial
  - The length of the experimental trial in the Experiment 3B was twice of the length of time used in Experiment 3A. To avoid participants' fatigue and boredom, the experimental was conducted over a two consecutive day period. Half of the participants experienced the large display on the first day, followed by the small display on the second day. The order was reverse for the second half of the participants.
- Display questionnaire
  - In addition to the interface device questionnaire, a display questionnaire was administered to the participants on the second day of the experimental trial to enable participants to make direct comparison between the two display sizes.

Eight (7 male and 1 female) were recruited for this study. None of them had participated in Experiment 3A. All participants had normal or corrected-to-normal vision. The average age was 27.86. With regards to their background in terms of

- Sport background – two did not play any sport at all, the rest played at least one of the following games (football, basketball, tennis, badminton, volleyball, golf)
- Computer games – two never played, three played between 1-4 times per week and three played at least 5 times a week
- VE experiment participations – only one had participated in a VE experiment before, the others had no experience in any VE experiment
- Familiarity with the interface device- all participants used the mouse at least once a day but more than half (5 out of 8) had never used the trackball
- Knowledge of the modelled room – none of the participants recognized the model room

Since the setup of Experiment 3B followed the setup of Experiment 1A and 2A, the following hypotheses will be examined:

1. The display type (small vs. large) has no effect on participants' distance estimation task (vertical, horizontal and transverse) performance in interactive VE
2. The display type (small vs. large) has no effect on participants' spatial memory task performance in interactive VE
3. The type of input device (mouse vs. trackball) has no effect on participants' spatial memory task performance in interactive VE
4. The different modes of travel (drive, fly) have no effect on participants' spatial memory performance in interactive VE

No real condition was compared as the main aim of Experiment 3B was to understand the unexpected finding of the better performance of small display over large display particularly for spatial memory tasks performance.



### 7.2.1 Display questionnaire

The display questionnaire was administered in order to examine directly whether participants' ratings and subjective comments matched their objective ratings of better performance on small display over large display as reported in Experiment 3A. It was expected that the display questionnaire would yield useful information that would aid understanding of the better performance of the small display over the large display in Experiment 3A. The questionnaire was based on three questions, designed to directly asked participants to rate their display size preference in terms of

- i) Ease of object recall
- ii) Overall preference
- iii) Confidence rating

These criteria represented a subset of the criteria in the Interface device questionnaire. The questionnaire was based on a 7-point rating scale, similar to the scale used in Interface device questionnaire. Participants were also encouraged to make additional comments as to their choice of rating. For analysis, 7 were considered a high rating while 1 was considered a low rating. To enable comparability among items, prior to analysis, Q(i), a negatively worded item was reversed using the SPSS transform recode function.

### 7.2.2 Results

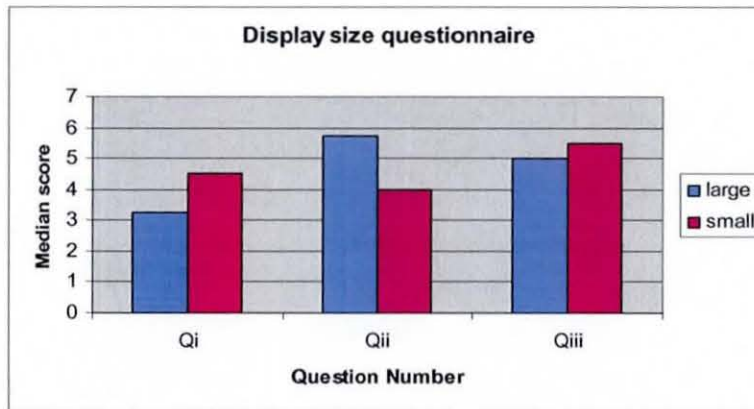
Due to the similar method of data collection, the similar methods of analysis and assumptions used in Experiment 3A was employed in Experiment 3B. However, there were two exceptions to this:

- The Display size factor was treated in Experiment 3B as a within subject factor as opposed to a between-subject in Experiment 3A.
- The display questionnaire, which was not administered in Experiment 3A, will be analyzed using the similar method to interface device questionnaire

Since the main aim of Experiment 3B was to examine why participants' performed better on a small display over a large display in Experiment 3A, the results of the questionnaire were presented and discussed first. This is followed by the results of the room size estimate data, spatial memory test and interface device questionnaires.

#### 7.2.2.1 Display questionnaire

Figure 7-19 indicates that participants rated the small display higher than the large display on ease of object position recall (Qi) and in terms of confidence rating (Qiii). However, on preference (Qii) they preferred the large display over the small display.



**Figure 7-19 Comparison of median score between large and small display for each question of display size questionnaire**

The results of repeated ANOVA analysis revealed no significant difference ( $p > .05$ ) between the display sizes for all questions. Two out of eight participants did not provide additional comments. For those who commented, half of them provided positive comments on small display and half provide positive comments on large display. On a small display, some of the participants' comments were:

- "Easier to find out size of room, got used to small display"
- "Used to look at small screen"
- "as screen is small, easier to recognize object location"
- "Large display: Not being used to the size. Being used to a small screen I think"
- "Large display : Seems ok but prefer the smaller one"

These comments suggested that the participants' familiarity with the small display may have influenced their preference and their performance. Interestingly, on a small display, one participant commented that the small display "seemed most natural" and another commented that small display "does seem real". This indicated that VE model used in Experiment 3 appeared to invoke a sense of realism on the participants.

On a large display, some participants commented that

- "Large display: got a better look at room. Small: too compact"
- "Gives better perspective"
- "Image is clearer"
- "Clearer, better perception"
- "Easier to see larger objects- more time spent looking at locations"

Additionally, one participant felt that he was immersed in the environment when using a large display. In his words: "with small screen I was still aware of the edges in my peripheral vision, with big screen I found I was drawn into the environment and less aware of the surroundings." Although the display used in this study was a semi-immersive and non stereo



display, viewing the image on a large display did provide the participants with the feelings of being immersed in the ‘virtual room’. This is similar to those reported by other researchers (Robertson, Card et al. 1993). Richardson and colleagues argued it was possible to induce a similar a sense of immersion when proper 3D cues and interactivity are available in a non-immersive VE. A user may feel drawn into the 3D world when he has control of the animation and focuses on it. As a result, the user may feel a sense of mental and emotional immersion even though the display is non-immersive. They compared this experience to similar experiences in playing a video arcade game. The VE model used Experiment 3 must have provided enough 3D cues to induce the same feeling on our participants. Additionally, allowing participants to control their movements have resulted in our participants having the same experience as suggested by Richardson and colleagues.

7.2.2.2 Room size estimate

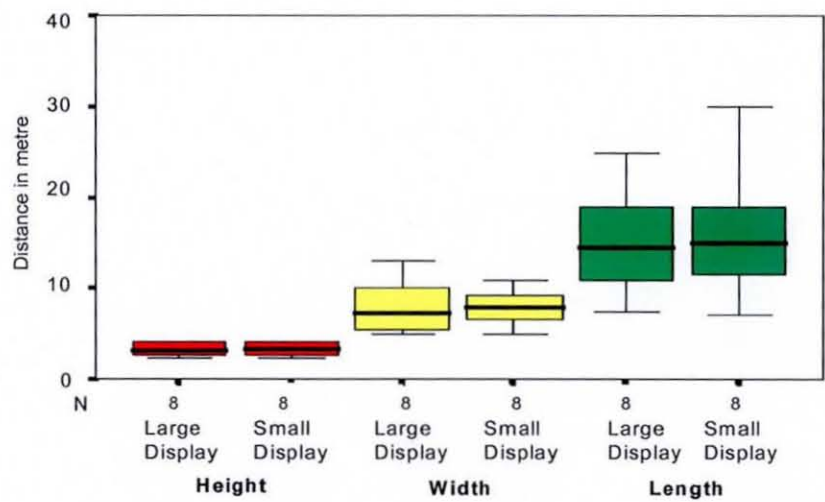


Figure 7-20 Comparisons between large and small display in terms of asymmetrical distances

Table 7-15 Comparison of means among experimental conditions.

Experimental conditions					
Display	Distance	N	Mean	Std. Deviation	% of dist. estimate from actual (estimate/actual *100)
large	height	8	3.28	0.68	93.71
	width	8	7.94	2.85	109.07
	length	8	15.19	5.72	102.64
small	height	8	3.29	0.68	94.00
	width	8	7.94	1.94	109.07
	length	8	16.00	7.13	108.11

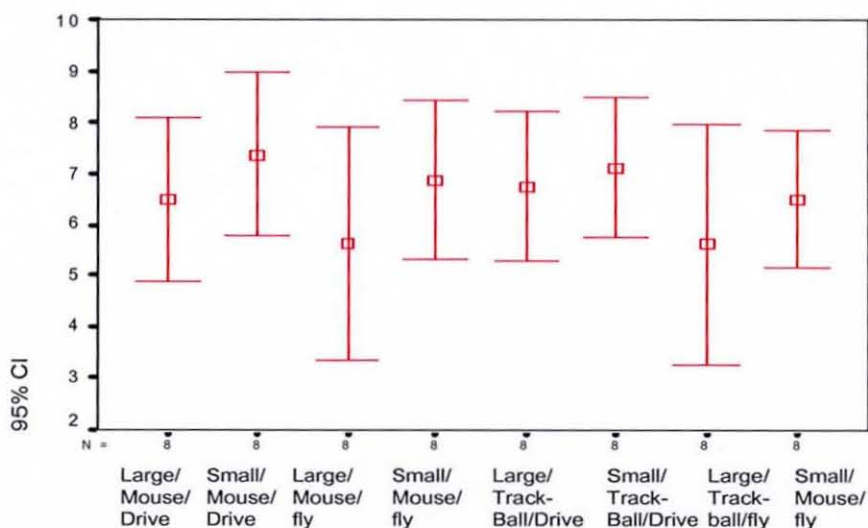
A repeated ANOVA analysis on distance estimate data revealed that no significant difference between the large and the small display: ( $F(1,7)=.062, p=.811$ ). This is evidenced from Figure 7-20 and from comparisons of means in Table 7-16.

The results showed there was a main effect of distance ( $F(2,6)=15.981, p=.004$ ), indicating there was a difference between asymmetrical distances (height, width, and length). A paired sample t-test also revealed no reliable difference between the large and small display for each asymmetrical distances ( $p > .05$ ).

The results of an ANCOVA analysis, which included *sport* variable and *computer games* variables (either alone or both) as covariates, remained similar to the results of the first ANOVA analysis. This indicated that both covariates did not constitute a large variance in this dataset on distance estimate tasks. These results were similar to the results of Experiment 3A whereby analysis (with and without covariates) revealed no significant difference between the display sizes.

Based on the means comparisons, current study results indicated that the small display estimates were large (except for the width distance) compared to the large display estimates which were consistent to Experiment 3A results. In contrast to Experiment 3A results, the length and width distances were generally overestimated. As indicated by the large standard deviation value (Table 7-16), there was a large variability for the width estimates and in particular for the length estimates compared to the height estimates. This may have accounted for the inconsistency of the Experiment 3B results to those of Experiment 3A.

### 7.2.2.3 Spatial memory test



**Figure 7-21** Comparison among large and small display, mouse and trackball and travel modes

A repeated ANOVA analysis spatial memory data revealed no main effects or interaction effects ( $p > .05$ ) and as indicated by Figure 7-21. However, the inclusion of the *sport* variable as covariates in an ANCOVA analysis revealed the following interaction effects:



- Device \* sport :  $F(1,6) = 15.322, p=.008$
- Display \* device \* travel :  $F(1,6) = 6.425, p=.044$
- Display \* device \* travel \* sport :  $F(1,6) = 8.960, p=.024$

When the *computer games* variable was included in the analysis as a covariate, no main effect or interaction effect was reported. However, including both the *sport* and *computer games* variables as covariates revealed only one interaction effect:

- device\*sport :  $F(1,5) = 12.373, p=.017$
- display\*device\*travel\* sport :  $F(1,5)=5.428, p=.067$

In contrast, when the effect of *sport* variable was removed in the ANCOVA analysis of Experiment 3A data, results showed a main effect for display and device type only. No other effects were reported. Although in terms of main effect and interaction effect there were slight differences between the results of Experiment 3B and Experiment 3A, comparisons of means among experimental conditions revealed similar trends of results.

**Table 7-16 Comparison of means among experimental conditions**

Display	Device	Travel mode	Mean	Std error
large	mouse	drive	6.5	0.74
		fly	7.375	0.64
	trackball	drive	5.625	0.92
		fly	6.875	0.65
small	mouse	drive	6.75	0.66
		fly	7.125	0.50
	trackball	drive	5.625	1.07
		fly	6.5	0.61

From Table 7-16, generally, spatial memory performance using the fly mode was more accurate compared to using the drive mode in all experimental conditions. Similarly to Experiment 3A results, this was true for all conditions except for the mouse conditions on a large display whereby the drive mode was better than the fly mode. A direct comparison of means indicated generally, using a mouse was better than using a trackball.

Similar to the results of Experiment 3A, comparisons of means indicated that the mouse/drive condition was more accurate than the trackball/drive condition. This was true for both display sizes. Similarly, the mouse/fly condition was better than the trackball/fly condition on both display sizes. For Experiment 3A, this was only true on a small display; on a large display the reverse was true.

Consistent with the results of Experiment 3A, the mouse/drive on small display was slightly better than mouse/drive on large display. In contrast to Experiment 3A result, the fly mode on a large display was slightly better than the fly mode on a small display. The results of Experiment 3A showed that for both travel modes, performance was better on a small display

compared to a large display. Experiment 3B results showed this was true for the trackball using the drive mode but for the fly mode, performance was slightly better on a large display compared to a small display.

7.2.2.4 Interface device questionnaire

Travel modes

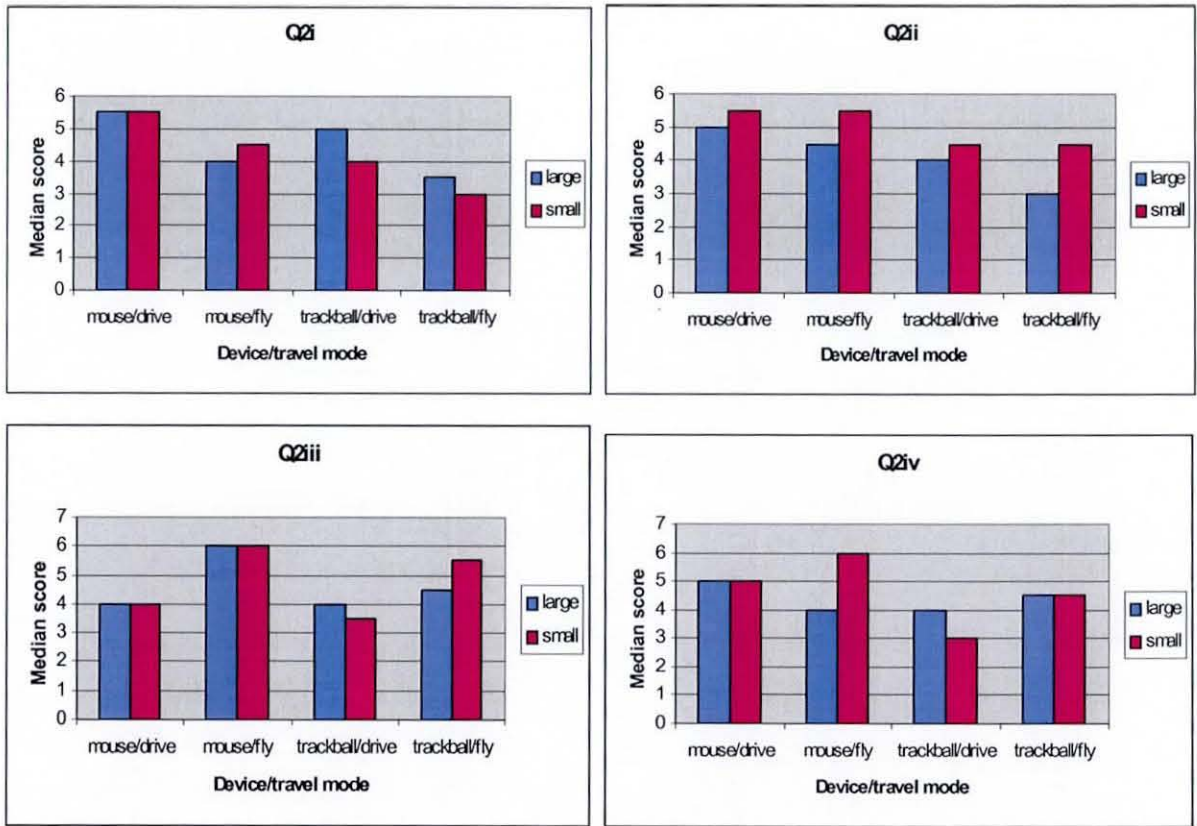


Figure 7-22 Median score for Q2i, 2ii, 2iii, 2iv of the Interface device questionnaire

The results of participants' ratings for Q2i-2iv are shown on Figure 7-22. A similar trend of ratings was observed for Q2i, 2ii and Q2iii to those reported in Experiment 3A:

- For Q2i (ease of movement), the mouse/drive condition was rated highest and the trackball/fly was rated lowest
- For Q2ii (control of movement), the mouse/drive was rated highest and the trackball/fly was rated lowest. For Experiment 3B, this is especially true on a large display
- For Q2iii (easy recall of object position), for both display devices the fly mode was rated higher than the drive mode

However, for Q2iv (on overall preference), Experiment 3B results showed that the drive mode was preferred over the fly mode on a large display only whereas in Experiment 3A participants preferred the drive mode better than the fly mode on both display sizes.



Interface device

For part 3 of the Interface device questionnaire in which participants were asked to make a direct comparison between device type, examination of median scores (Figure 7-23) revealed a similar trend of rating were given for all questions (Q3i, Q3ii, Q3iii, Q3iv, and Q3v) to those reported in Experiment 3A. Similar to Experiment 3A results, participants tended to rate the mouse higher than the trackball on both displays sizes ( $F(1,7) = 7.00, p = .033$ , but when the covariates were removed no main effect were revealed). The difference of scores between display sizes however is not significant ( $p > .05$ ). These results gave us more confidence on the earlier presented results of Experiment 3A.

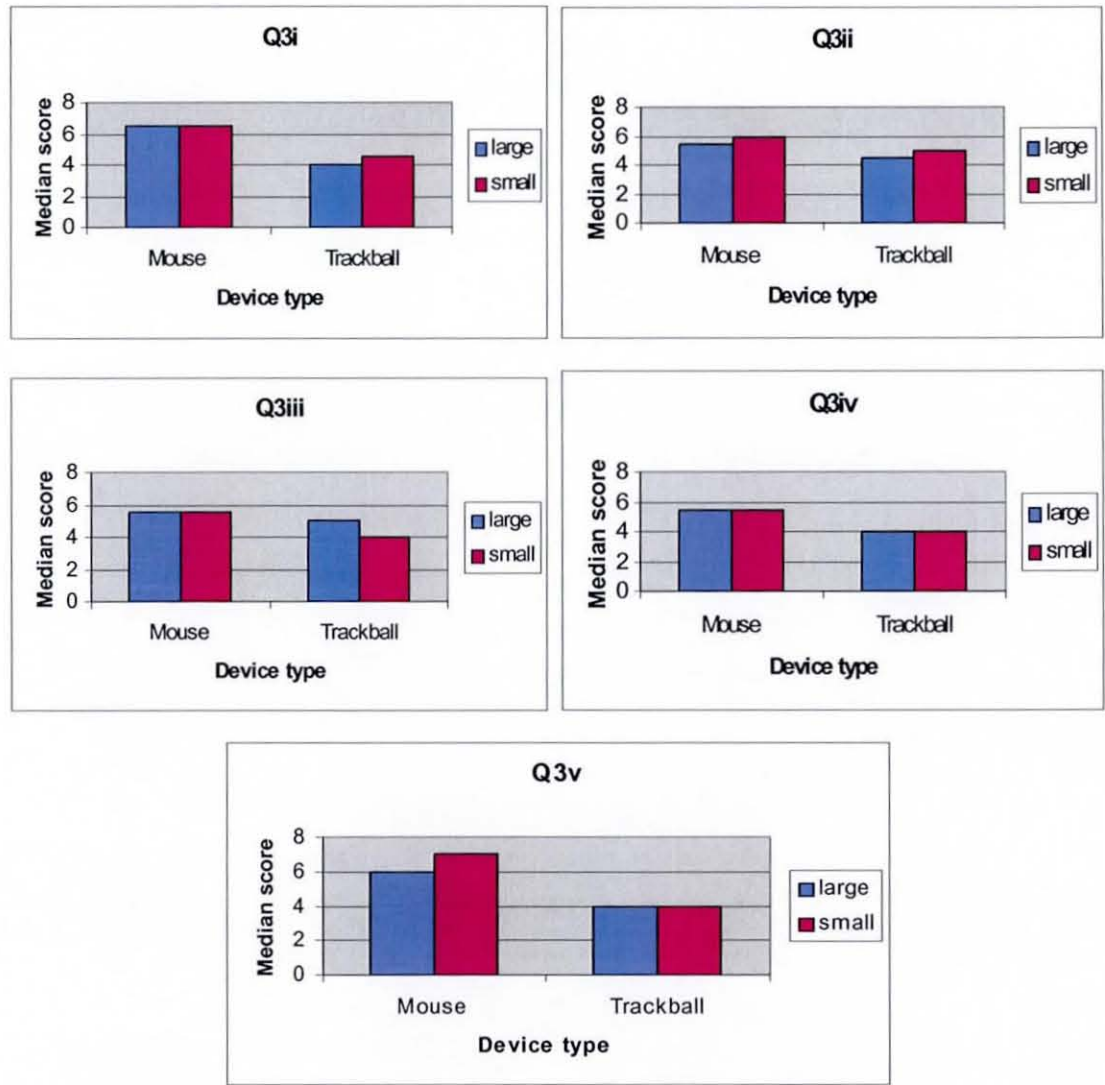


Figure 7-23 Median score for Q3i, 3ii, 3iii, 3iv, 3v of the Interface device questionnaire

Question 4 Recall accuracy

In terms of recall accuracy, participants did not differ very much on their ratings between display sizes for both device types. Similar to the results of Experiment 3A, the large display



was rated slightly higher than the small display. Whilst Experiment 3A results showed that the participants were more confident when using a trackball compared to using a mouse, Experiment 3B results showed that the reverse was true.

Participants' additional comments on interface device questionnaire were not presented in this section as these comments were similar to those given in Experiment 3A. These comments were however available in Appendix C.

## 7.2.3 Discussion

### 7.2.3.1 Display questionnaire results

The results of the display questionnaire indicated that even though the participants marginally preferred the large display to the small display, however, in terms of ease of recall and confidence on their spatial memory test accuracy, the small display was rated slightly higher than the large display.

The self-reported comments from the participants further suggests that a large display does not necessary improve participants' performance; for some participants the small display may yield better performance over the large display. A common reason given by the participants was that they were used to small display. This was expected as all participants (staffs and students from the Computer Science Department) reported using the mouse device at least once a day. Thus, they were more frequently exposed to using a desktop monitor (notebook or laptop screen). Comparatively, a large display exposure would be less frequent.

In a study comparing display types (desktop monitor, HMD, large screen projection) and navigational aids on participants' navigation performance, presence, and workload during exploration of a virtual office using a tele-robotic vehicle, Riley and Kaber (1999) found that participants performed significantly better on a desktop monitor compared to two other displays. Beside the image resolution, the authors also suggested that the participants' familiarity with a desktop might have contributed to their better performance in the desktop condition. Therefore similar reasons may explain the unexpected findings of the better performance of a small display over a large display in Experiment 3A and 3B.

Some participants commented that the small display "seems real" and "most natural" to them. These comments suggested that a VE model presented on a small display may appear realistic to some viewers. Theoretically, an image presented on a large display would appear natural

and real as it is relatively more similar to the real world in terms of scale. Moreover, past studies have shown that the participants performed better on a large display compared to a small display for some tasks (Kline and Witmer 1996, Patrick, Cosgrove et al. 2000, Czerwinski, Tan et al. 2002, Tan, Gergle et al. 2003). However, as indicated by the participants' subjective responses it was also possible to induce a sense of realism on a small display, at least within the constraint of Experiment 3A and 3B.

However, in support of previous works (Kline and Witmer 1996, Patrick, Cosgrove et al. 2000, Czerwinski, Tan et al. 2002, Tan, Gergle et al. 2003), Experiment 3B results also provide support for claims that a large display could invoke a sense of immersion on the viewers as reported by one of the participants. Thus, consistent with theoretical prediction and as reported by several researchers (Patrick, Cosgrove et al. 2000, Hatada, Sakata & Kusaka 1980, cited in Pfautz 2000), an image on a large display may appear natural and real as it was more similar to the real world in terms of scale.

#### 7.2.3.2 Effects of display size on distance estimate task

Similar to the results of Experiment 3A, the results of Experiment 3B showed that there is no significant effect of display size on distance estimate tasks (even with covariates' effects removed). Both results showed that regardless of whether the FOV and viewing distance were fixed or varied, the difference of distance estimate tasks on both display sizes was very small. This suggests that the physical display size did not contribute largely towards distance estimate tasks in interactive images. Similarly, the influence of FOV, viewing distance and physiological cues on distance estimate tasks in interactive images were also minimal.

#### 7.2.3.3 Effects of display size on spatial memory task

For spatial memory tasks, however, no significant effect of display was revealed when the FOV was fixed and the viewing distance was varied (Experiment 3B) but the effect of display was significant when the FOV was varied and the viewing distance was fixed (Experiment 3A). This implies that the better performance of a small display over a large display was influenced more by the FOV of the display for spatial memory task.

#### 7.2.3.4 Effects of device type, travel modes, sporting background on spatial memory task

Contrary to the results of Experiment 3A, for spatial memory task there was no significant main effect of device type. However, the significant interaction effect of the display size,

device type, travel modes, and sporting background (at 10%) suggested that all these factors contributed some influence to spatial memory task performance in interactive images.

With some exceptions, the similarity of Experiment 3B results for distance estimation tasks, spatial memory tests and interface device questionnaire to those of Experiment 3A provide support for the results of the latter and gives us the confidence in the methods employed in both studies. Large participants' variance might partially account for the slight difference in results.

### 7.2.4 Conclusion

While it was not clearly indicated from the second study why the participants performed better on the small display over a large display, the results did indicate that a large display does not necessary have the same impact on all viewers.

The subjective responses from the participants suggested that not all participants would rate the large display better than the small display. As reported by display questionnaire, some viewers rated the small display higher than the large display on ease of recall and confidence rating. The common reason given by the participants include more familiar with the small display compared to the large display. This might have explained the better performance on the small display over the large display. Although theoretically, small display is less similar to the real world in terms of scale compared to the large display, for some viewers, it was still possible to induce a sense of visual realism on a small display.

Consistent with previous investigations, the large display provided the viewer with a sense of immersive feeling and this sense of immersive feeling may be experienced even if the display is only semi-immersive and non-stereo.

## 7.3 OVERALL CONCLUSIONS

The overall aim of the experiments presented in this chapter was to examine user spatial awareness in interactive real and VE. The goal was to examine the effect of type of environment (real verses VE), display size (large verses small), input device (mouse verses trackball) and travel modes (drive verses fly) on distance estimate and spatial memory tasks. Two separate, but related, studies were described: Experiment 3A and Experiment 3B. Details of experimental methods employed in these studies were first described prior to the presenting of the experimental results, discussions and finally on the conclusions derived.

The results of Experiment 3A suggested that it was possible to perceive distance similarly in the real and VE. The results of non-significant difference between the large and the small display implies that the viewing distance and physiological were also contributing factors to the effect of display size factor.

Contrary to the results of previous investigations, performance on a small display was slightly better than on a large display. For spatial memory task, the results of Experiment 3A indicated that it was also possible to perceive spatial relations similarly in the real and VE. A main effect of display size was shown when the influence of the sporting background and computer experience variable was removed, revealing better performance of the small display participants over the large display participants. This unexpected finding provided motivation for the undertaking of Experiment 3B.

Spatial memory tasks performed using a mouse was significantly more accurate compared to using a trackball. For travel mode, spatial memory performance using a fly mode was only slightly better than drive mode.

The results of Experiment 3B provided further clarification on the unexpected findings of Experiment 3A. The results of the display questionnaire suggested that a large display does not necessary have the same impact on all users for spatial memory tasks. Users' familiarity with a particular display size may have influence their performance.

For distance estimate tasks, the results of Experiment 3B suggested that the influence of physical display, viewing distance and physiological cues were minimal.

For spatial memory tasks, the results suggested that the better performance of small display participants over large display participants were influenced more by the FOV of the display.

While no main effect was reported, however, a significant interaction (at 10%) among display size, device type, travel modes and sporting background factors indicated that all of these factors contributed significantly to the performance of spatial memory tasks in interactive images.

In the next chapter, Chapter 8, overall analyses of findings from all experiments (Experiment 1, 2 and 3) are further discussed.

## PART III

### IMPLICATIONS DRAWN FROM THE LITERATURE AND EXPERIMENTS UNDERTAKEN

**Chapter 8 – Overall Analysis of Results**

**Chapter 9 – Implication for Spatial Awareness Perception in VE,  
Final Conclusions and Research Contributions**

**Chapter 10 – Recommendations/Future Work**



# CHAPTER 8

---

## *OVERALL ANALYSIS OF RESULTS*

### 8 OVERVIEW

In this chapter, an overall analysis of results from the experiments on static images, dynamic images and dynamic images is presented. The overall conclusions and implications of findings from research however are presented in the next chapter, Chapter 9.

Prior to the presentation of the overall analysis of the results of Experiment 1, 2 and 3, the main findings from each experiment are summarized in a tabular format to serve as a guide to the reader. The overall analysis of the results is presented in two separate sections based on the task performance measures examined. The first section (Section 8.2) is based on distance estimate tasks measures which were examined in all three experiments (Experiment 1, 2 and 3). The second section is based on the spatial memory task measure which was examined only in Experiment 3. The results of the post-test questionnaires are also discussed in the second section.

## 8.1 SUMMARY OF ALL EXPERIMENTS' RESULTS

The overall aim of the research presented in this thesis, hence the aim of the three series of experiments presented in this thesis, was to examine user's spatial awareness performance in the VE in comparison to similar performance in the real world. These experiments examined factors that affected spatial awareness in the real and VE in the context of static, dynamic and interactive environment presented to the participants in non-stereo, non-immersive and semi-immersive displays. Factors related to display were the main focus of this thesis with an emphasis on display size factor, viewing distance and physiological cues. An overview of the experimental goals and designs for these experiments was presented earlier in Table 4-4 of Chapter 4.

In this chapter, Table 8-1 and Table 8-2 provide a summary of the research findings from these experiments. Table 8-1 provides a summary of the main findings on distance estimates task performance in static, dynamic and interactive images for Experiment 1, 2 and 3 respectively while Table 8-2 provides a summary of the main findings on spatial memory task performance in interactive images and questionnaire results from Experiment 3. Both tables serve to guide the reading of the discussion in Section 8-2 and Section 8-3 respectively.

**Table 8-1 Summary of results for distance estimates for all experiments (1,2, and 3)**

<b>Main Data Analysed</b>	<b>Experiment 1A:</b> Asymmetrical Distance - Horizontal - Transverse  <b>Experiment 1B:</b> Asymmetrical Distance - Vertical - Horizontal - Transverse	<b>Experiment 2A:</b> Asymmetrical Distance - Vertical - Horizontal - Transverse  <b>Experiment 2B:</b> Asymmetrical Distance - Vertical - Horizontal - Transverse	<b>Experiment 3A:</b> Asymmetrical Distance (Room Size Estimates)* - Vertical - Horizontal - Transverse  <b>Experiment 3B:</b> Asymmetrical Distance (Room Size Estimates) - Vertical - Horizontal - Transverse  *vertical refers to height, horizontal refers to width and transverse refers to length
<b>Main Findings</b>	<b>Experiment 1A:</b>  Image type:  No significant difference was found between real and VE image for horizontal and transverse distance  Display size:  A main effect of display size for horizontal (at 10%) and	<b>Experiment 2A:</b>  Image type:  A main effect ( $p < .05$ ) of image type (real and VE) for all asymmetrical distances. For vertical, VE is better than real image on both display sizes. For horizontal and transverse distance, estimates in the VE are better than in the real image for small display. For large display estimates in the	<b>Experiment 3A:</b>  Image type (Environment type):  No main effect of image for all asymmetrical distance types: height, width and length. Generally, performance on VE is better than on real image for all asymmetrical distance types.  Display size (for VE only):

	<p>transverse distance (at 5%) was revealed. Results showed that performance on a large display is better than on a small display.</p> <p>Generally, horizontal estimates were more accurate than transverse estimates</p> <p><b>Experiment 1B:</b></p> <p>Image type:</p> <p>The effect of image type was not investigated.</p> <p>Display size:</p> <p>No main effect of display size was found for all asymmetrical distances. Performance was slightly better on a large display compared to a small display for horizontal and transverse distance. For vertical distance the reverse is true. The non significant effect of display size factor implies physiological cues and viewing distance contribute an influence in asymmetrical distance estimates. Retinal image cues are found to be less influential.</p> <p>Generally, vertical estimate is more accurate than horizontal and transverse distance</p>	<p>real image is better than in the VE.</p> <p>Display size:</p> <p>No main effect of display for all asymmetrical distances. Performance is slightly better on small display compared to large display.</p> <p>Generally, vertical estimate is more accurate than horizontal and transverse distance</p> <p><b>Experiment 2B:</b></p> <p>Image type:</p> <p>Main effect of image for vertical and transverse distance (<math>p &lt; .05</math>). On a large display, performance in the VE is better than in the real image. On a small display, performance in the real image is better than VE for vertical and horizontal only. For transverse distance, performance in the VE is more accurate than in the real image.</p> <p>Display size:</p> <p>No main effect of display for all asymmetrical distances. With the exception of transverse estimates in the real image, performance was slightly better on a large display compared to a small display. Similar to Experiment 1B, the non significant effect of display size factor implies physiological cues and viewing distance contribute an influence in asymmetrical distance estimates.</p> <p>Results from experiment 2A and 2B enables us to conclude that image resolution contributes an influence on distance estimation</p> <p>Generally, vertical distance estimate is more accurate than horizontal and transverse</p>	<p>No main effect of display for all distance types (after the removal of covariates effects). Performance is slightly better on small display over large display.</p> <p>Generally, the height distance is more accurate compared to width and length distance</p> <p><b>Experiment 3B:</b></p> <p>Image type (environment type):</p> <p>The effect of image type (environment type) was not investigated.</p> <p>Display size:</p> <p>No main effect of display size was found for all asymmetrical distance (after removal of covariates effect). Generally, performance is slightly better on a small display compared to a large display (except for width distance).</p> <p>Generally, the height estimate is more accurate compared to the width and length estimates</p>
--	---	--	--

Table 8-2 Summary of results for spatial memory task and post-test questionnaire for Experiment 3A and 3B

Analysis	Experiment 3A Results	Experiment 3B results
Main Data Analysed	<p>1. Spatial Memory Test</p> <p>- Number of correctly placed object</p> <p>2. Interface Device Questionnaire</p> <p>- participants' rating scores</p>	<p>1. Spatial Memory Test</p> <p>- Number of correctly placed object</p> <p>2. Interface Device Questionnaire</p> <p>- participants' rating scores</p>

		3. Display Questionnaire - participants' rating scores
<b>Spatial memory test</b>	<p>Overall,</p> <ol style="list-style-type: none"> <li>1. No difference between real and VE</li> <li>2. No difference between large and small display</li> </ol> <p>When effect of covariates are removed:</p> <ul style="list-style-type: none"> <li>- Main effect of display (performance on small display is better on large display)</li> </ul> <p>Notes: Covariates refers to <i>sport background</i> and <i>computer games experience</i></p>	<p>Overall,</p> <ol style="list-style-type: none"> <li>1. Image type was not investigated</li> <li>2. No difference between large and small display</li> </ol> <p>When effect of covariates are removed:</p> <ul style="list-style-type: none"> <li>- Interaction effect of display size, device type, travel mode and sport background approach significant (<math>p = .067</math>)</li> </ul> <p>Notes: Covariates refers to <i>sport background</i> and <i>computer games experience</i></p>
	<p>Interface device:</p> <ul style="list-style-type: none"> <li>- A main effect of device was found (performance using a mouse is better than using a trackball)</li> </ul>	<p>Interface device:</p> <ul style="list-style-type: none"> <li>- No main effect of device (but a direct comparison of means reveals performance using a mouse is better than using a trackball)</li> <li>- BUT the interaction effect of display size, device type, travel mode, and sport background approached significance when the covariates were removed</li> </ul>
	<p>Travel mode</p> <ul style="list-style-type: none"> <li>- No main effect of travel mode was found (but performance using a fly mode is better than using a drive mode)</li> </ul>	<p>Travel mode</p> <ul style="list-style-type: none"> <li>- No main effect of travel mode was found (but a direct comparison of means reveals performance using a fly mode is better than using a drive mode)</li> <li>- BUT the interaction effect of display size, device type, travel mode and sport background approached significance when the covariates were removed</li> </ul>
<b>Post-test questionnaire</b>	<p>Interface device:</p> <ul style="list-style-type: none"> <li>- generally mouse is rated significantly better than trackball on all questions (see Table 7-13)</li> </ul>	<p>Interface device:</p> <ul style="list-style-type: none"> <li>- generally mouse is rated better than trackball on all questions (significant for Qi and Qv only, but when covariates removed, none is significant)</li> </ul>
	<p>Travel mode:</p> <ul style="list-style-type: none"> <li>- Drive mode is rated better than fly mode for all questions except Q2(iii) on object recall (<math>p &lt; .05</math>) (see Table 7-12)</li> </ul>	<p>Travel mode:</p> <ul style="list-style-type: none"> <li>- Drive mode is rated better than fly mode except for Q2(iii) on object recall (see Figure 7-24) (not significant)</li> </ul>
	<p>Recall accuracy:</p> <ul style="list-style-type: none"> <li>- trackball is rated better than mouse (<math>p &lt; .05</math>)</li> <li>- No difference between large display and small display (large display is better than small display)</li> </ul>	<p>Recall accuracy:</p> <ul style="list-style-type: none"> <li>- mouse is rated slightly better than trackball (not significant, <math>p &gt; .05</math>)</li> <li>- No difference between large and small display (but a direct comparison of means showed large display is better than small display)</li> </ul>
<b>Display questionnaire</b>	Not investigated	<ul style="list-style-type: none"> <li>- No significant difference (<math>p &gt; .05</math>) was found between display size for all questions</li> <li>- Small display is rated better than large display for Q(i) (ease of recall) and Q(iii) (confidence rating)</li> <li>- for Q(ii) (overall preference), large display is rated better than small display</li> </ul>

## 8.2 OVERALL ANALYSIS BASED ON DISTANCE ESTIMATE TASK PERFORMANCE

### 8.2.1 Comparison based on environment types (Real and VE)

The results of the first experiment on static images (Experiment 1) revealed that participants' distance estimates did not differ significantly on both image types (real and VE picture) for horizontal and transverse distances. This suggests that participants' spatial awareness in terms of distance estimation task performed using the real picture was similar to those performed using the VE picture, within the constraint of the experiment.

Findings from this study are inline with previous researchers who showed that it is possible to perceive the real and VE similarly (Waller 1999, Willemson and Gooch 2002, Yoon, Byun et al. 2000, Yang, Dixon et al. 1999). With regard to relative perception of vertical and horizontal extents, Yang, Dixon et al. (1999) demonstrated that whether an observer viewed a snapshot of the VR scene on a desktop or a picture of the scene, it did not make any difference; both were perceived similar in terms of horizontal and vertical illusion. Yoon and colleagues (2000) reported that distance estimates between the real and VE room were not significantly different but both were different from the actual sizes in terms of width and length but not for the height estimates. Willemson and Gooch (2002) who compared panoramic photographic-based VE to computer-generated VE model of the same scene found that the difference is quite small even though the photographic image is richer in visual information (such as shadows and global illumination) than the computer-generated VE. Similar to Willemson and Gooch's (2002) investigation in terms of stimulus (photographic based image versus computer generated model) and for textures of the computer-generated model, photographic images were utilized to create the textures for objects in the VE. Experiment 1A's results suggests that the VE picture used provides the visual information necessary for the perception of distance similar to those available in the real pictures.

With the exception of Yoon, Byun et al. (2000), most of the earlier researchers mentioned and reviewed in Chapter 2 who reported a difference between the real and VE based their conclusions on comparisons between the real physical environment and a 3D VE model of it and not between pictures as compared in Experiment 1. For example, Henry and Furness (1992) compared a museum and a VE model of it, Lampton, McDonald et al. (1995) compared a virtual corridor of an office building to a real corridor and Witmer and Sadowski (1998) compared a virtual hallway to a real hallway. Witmer and Kline (1998) used a real hallway and a VE model of it. More accurate estimates were found by these studies in the real physical environment because it was richer in terms of visual information when compared to



the VE model (Witmer and Kline 1998). However, Yoon and colleagues (2000) used a real room and a VE model of it but they reported no significance difference. One possible explanation for Yoon and colleagues (2000) results is the use of a very simple room (with one window, one door and one chair) for the stimulus. Additionally, their participants had practiced estimations in 3 different rooms prior to the actual trial which may have improved their participants' estimations.

In contrast to the results of Experiment 1 on static images, the results of Experiment 2 on dynamic images (where participants are passive viewers of the images) showed a significant difference between the real and the VE images in terms of distance estimates task for all asymmetrical distances (vertical, horizontal and transverse). These results imply that the conclusions drawn from static images do not extend to dynamic images.

These results are consistent with the results of previous findings (Henry and Furness 1993, Witmer and Kline 1998, Witmer and Sadowski 1998, Lampton, McDonald et al. 1995). However, the directions of results in Experiment 2 were partially influenced by the image resolution. When a low VE image resolution was used compared to the resolution of the real image (as in Experiment 2A),

- the VE image participants tended to perform better than the real image participants on a small display for all distances.
- However, on a large display the real image participants tended to perform better than the VE image participants for horizontal and transverse distances only.
- For vertical distance, a low image resolution did not seem to affect VE image participants' performances. Results showed that the VE image participants tended to perform better than the real image participants on both display sizes.

As expected when a high image resolution was used for the VE image (as in Experiment 2B) the results showed that the VE image participants tended to perform better than the real image participants on a large display for all asymmetrical distances. On the small display, however, this is only true for transverse distance. For vertical and horizontal distance, the real image participants performed better than the VE image participants. This result is statistically significant ( $p < .05$ ) for vertical and transverse distances.

Experiment 2 results suggest that the influence of image quality appears to be less influential on distance estimation task performed on a small display but on a large display the quality of image matters as the use of a low image resolution may degrade distance estimate performance. However, as mentioned earlier low image resolution does not appear to affect vertical distance estimates on either display sizes.

However, when participants were allowed to explore the VE as examined in Experiment 3 on interactive images, similar to the results of static images, our results showed no significant difference between the real and VE conditions for all asymmetrical distances. This result is inconsistent with the results of most previous investigations (Henry and Furness 1993, Witmer and Kline 1998, Witmer and Sadowski 1998, Lampton, McDonald et al. 1995) which reported a significant difference between the real and VE conditions.

It was noted that the real conditions in Experiment 3 were based on the real physical environment similar to those used in these previous investigations. Given the well-established findings of a significant difference between the real and VE conditions, the results of our findings were unexpected. However, the results of a very recent investigation (Plumert, Kearney et al. 2004) were consistent with our findings. Plumert, Kearney et al. (2004) and Experiment 3 share some similarities in terms of method of study:

- The use of verbal report method for distance estimate task
- The use of a large projection screen, instead of a HMD as the display type

Plumert, Kearney et al. (2004) suggested that the significant difference between real and VE conditions in the previous investigations were due to the use of HMD. Though, several studies have ruled out the effect of restricted viewing conditions and image quality of HMD, a more recent study (Willemson, Calton et al. 2004) indicates that the mechanical aspects of the HMD (such as mass and inertia) were partly responsible for the inaccurate performance in the VE using a HMD as the display type. Thus, the use of a HMD instead of a large projection screen may have yielded a difference in results between real and VE conditions for the previous investigations. Consequently, similar arguments could be used to explain the results of non-significant difference between the real and VE in Experiment 3.

Two other possible explanations might have accounted for the results of Experiment 3. First is the use of a single room. Yoon, Byun et al. (2000) reported no difference in performance on distance estimate task between the real and VE when they used a single VE and a real room in their study. In contrast, Henry and Furness (1993) used multiple rooms as their real and VE conditions and they reported a significant difference in distance estimate performance between the real and VE. Similarly, based on spatial representation task, Richardson, Montello et al. (1999) reported no significant difference between the real and VE conditions when using a simple single floor environment but when a complex building is used the authors found a significant difference in distance estimates performance between the real and VE. The use of a single room may have simplified the participants' tasks in both conditions, thus may have accounted for the similar performance between the real and VE.

A second and more likely explanation is practice effects. Due to the experimental design, there is a difference between the real and VE conditions in terms of viewing time. The VE participants have more viewing time compared to the real world participants. Given the common belief that more practice improves learning (Stanney, Mourant et al. 1998) and the results of studies that indicate more experience in the VE improve participants' distance estimation tasks (Ruddles, Payne et al. 1998), it seems more likely that more practice may have accounted for the improved performance of VE participants over the real world participants.

Surprisingly, Experiment 3 results showed that distance estimate performance in the VE/small condition is slightly better than in the real conditions for all asymmetrical distances. Beside the practice effects, the details in the real condition might have imposed more cognitive load on the real condition participants thus slightly degrading their performance (Yanagisawa and Akahori 1999).

### 8.2.2 Comparison based on display size (Large and Small)

Examination of distance estimate performance for static images presented on large and small display (Experiment 1A) revealed a significant difference for horizontal ( $p < .052$ ) and transverse distance ( $p < .029$ ). The results showed that participants' estimation were more accurate on a large display compared to a small display. These results are consistent with findings of previous researchers (Patrick, Cosgrove et al. 2000, Tan, Gergle et al. 2003, Czerwinski, Tan et al. 2002, Tan, Czerwinski et al. 2003). It is noted that the tasks investigated in these studies were different from Experiment 1A. The better performance of participants in the previous studies on different tasks such as navigation, spatial orientation and spatial memory tasks has been attributed to the sense of presence and realism induced by the larger display. As such, the better performance of the participants in Experiment 1A could be due to our participants having similar experiences. Additionally, larger display provides participants with a better sense of scale much closer to the real world (Patrick, Cosgrove et al. 2000), thus the better estimates on large display.

As discussed in Section 4.1.1.1 of Chapter 4, the results of better performance of large display over small display participants in previous investigations may not be due to the large physical display size alone, other factors such as viewing distance and physiological cues might have also influenced the results. The reasons lie in the experimental setup of Experiment 1A and previous investigations. In order to factor out the effect of FOV (and retinal image size) on both display sizes, the FOV value was fixed for both display sizes. However, fixing the FOV

values resulted in different viewing distances for each display size and in such conditions, the accommodation and vergence cues acting at these different distances would be different. Thus, the significant difference between display sizes might not be attributed to the physical display size alone; the viewing distance and the physiological cues (accommodation and vergence cues) might also exert an influence on the results. Experiment 1B was conducted to examine the possible effect of these factors (viewing distance and physiological cues) by fixing the viewing distance and physiological cues for both display size. No significant difference between the large and small displays was observed in Experiment 1B. The non-significant result could be explained by the similar viewing distance and similar accommodation and vergence cues acting at the same distance from the display screen for both display conditions. Thus, the results from Experiment 1A and Experiment 1B suggest that, besides the display factor, viewing distance and physiological cues (accommodation and vergence cues) do contribute a large influence on distance estimation task.

Despite the non-significant effect of display in Experiment 1B, the large magnitude of effect size still suggests that a large variation of the distance estimation was explained by display size factor. This is also indicated by the difference of percentage mean of estimate between the large and small screen conditions whereby larger error was reported on small display condition for horizontal and transverse distance compared to on large display. This implies that distance estimation is more accurate on a large screen compared to a small screen for both asymmetrical distances. This result is consistent with the results from the first experiment (Experiment 1A) where distance estimation was more accurate on a large display compared to a small display for horizontal and transverse distances.

While not significant, vertical distances were estimated more accurately on a small display compared to a large display. Vertical distance tended to be overestimated more on a large display than on a small display which is consistent with previous findings whereby it was indicated that the larger the display the larger is the overestimation of the vertical extent (Yang, Dixon et al. 1999).

From the results of Experiment 1A, there was a main effect of display even though the image size projected on the observers' retina is similar; suggesting the physical display size is a contributing factor on distance estimation task. But from Experiment 1B there was no significant difference between the large and small display even though the image sizes projected on the observers' retina were different (that is, large and small image were projected for large and small display respectively). This implies that the retinal image size (or the FOV)

is a weak cue and easily overridden by other cues as suggested by previous researchers (Beall, Loomis et al. 1995).

However, when the experiment was conducted employing dynamic images (Experiment 2A and 2B) the results showed no significant difference for asymmetrical distance tasks performed on large and small display. This result is inconsistent with previous findings.

It should be noted the setup of Experiment 2A was similar to Experiment 1A where the FOV of the display was fixed for both display sizes. Similarly, the setup of Experiment 2B was the same as Experiment 1B where the FOV of the display was varied and the viewing distance (and the effect of accommodation and vergence cues) was made the same. This result appears to indicate that the conclusions drawn from the results of experiment on static images in which a large display was better than a small display do not extend to the dynamic images. However, the results of the latter might have been influenced by the resolution of the image of the stimulus (as argued in Chapter 6):

- When a low image resolution was used for VE condition (as in Experiment 2A), though not significant, performance was better on a small display compared to a large display for all asymmetrical distances. Similarly, this was also true for the real image condition.
- When a high image resolution for VE condition was used (as in Experiment 2B), performance was better on a large display compared to a small display for all asymmetrical distances.
  - For real image conditions (which used similar image resolution as in Experiment 2A), the results were consistent with those of Experiment 2A that is performance on a small display was slightly better compared to a large display for transverse distance.
  - However, for vertical and horizontal distance, performance on a large display was slightly better than in the small display.

The better performance of a large display over a small display in the second study (Experiment 2B) has two possible implications:

- First, this confirmed the prediction that image resolution does influence performance. While this conclusion contradicts the conclusion drawn by some researchers (Willemson and Gooch 2002, Thompson, Willemson et al. in press), it supports the claims by others (Jää-Aro and Kjelldahl 1997, Kline and Witmer 1996, Duh, Lin et al. 2002, Eggleston, Jansen et al. 1996, Loomis and Knapp 2003).
- Second, because the viewing distance and physiological cues were fixed and the FOV (and retinal image size) was varied, applying similar arguments used in Experiment 1B, this implies that in addition to display size, viewing distance and physiological cues may



also contribute an influence on asymmetrical distance estimate task at least for all asymmetrical distances in the VE condition and vertical and horizontal distances in the real conditions.

Thus, it is argued if higher image resolution for VE condition was used in Experiment 2A, the results might have been similar to the results of previous investigations (Patrick, Cosgrove et al. 2000, Tan, Gergle et al. 2003, Czerwinski, Tan et al. 2002, Tan, Czerwinski et al. 2003) whereby better performance of a large display was observed over a small display. This indicates that similar to the results of experiment on static images (Experiment 1), in dynamic images (Experiment 2) the better performance of a large display over a small display is influenced not only by the physical display size but also by the viewing distance and the physiological cues. Additionally, the image resolution is also another factor which may influence the distance estimate performance. As expected, the use of a low image resolution may yield contradicting results where performance on a small display is better than on a large display.

For experiment on interactive images (Experiment 3), the setup of Experiment 3A was similar to those of Experiment 1B and 2B where the viewing distance and the physiological cues were fixed and the FOV (and retinal image size) were varied. Results showed that there was no difference in distances estimate performance on a large and a small display (with the exception of length distance). Similarly, when the effects of covariates (sport background and computer games experience) were removed, the results showed no significant difference between the large and the small display for all asymmetrical distances.

It should be noted that the setup of previous investigations (Patrick, Cosgrove et al. 2000, Tan, Gergle et al. 2003, Czerwinski, Tan et al. 2002, Tan, Czerwinski et al. 2003) was similar to Experiment 1A and 2A setup whereby the FOV of the display (and retinal image size) was fixed and the viewing distance (and physiological cues) were varied. Thus, applying similar arguments used by Experiment 1B and 2B, this may imply that besides display size factors, other factors such as viewing distance and physiological cues might have also explained the results of previous investigations which indicated that the better performance of a large display over a small display was due to physical display size. However, contrary to expectation and the previous investigation results, performance on a small display was better than a large display.

In order to understand why performance is better on a small display, Experiment 3B was undertaken by including a display questionnaire which directly asked participants to rate both display sizes based on some criteria. Experiment 3B employed the setup of Experiment 1A

and 2A where the FOV was fixed (retinal image size) and the viewing distance (and physiological cues) was varied. Similar to the result of Experiment 3A but in contrast to the result of the previous investigations which employed the setup of Experiment 3B, Experiment 3B results showed no significant difference between performance on a large and a small display. Contrary to the conclusions drawn from previous investigations (Patrick, Cosgrove et al. 2000, Tan, Gergle et al. 2003, Czerwinski, Tan et al. 2002, Tan, Czerwinski et al. 2003), current study results suggest that

- the physical display size does not contribute largely towards distance estimate task performance in interactive images at least within the current experiment constraint.
- the contribution of viewing distance (and physiological cues) and FOV were considered minimal for distance estimate task in interactive images.

This is because both Experiment 3A and 3B's results showed no main effect of display regardless of whether FOV or viewing distance (and the physiological cues) were fixed or varied. One possible explanation is the fact that participants were allowed to interact with images which might have influenced performance. Another explanation is that participants' estimates were based on experience on the viewed environment. Although during the experiment participants were asked to estimate based on what they have viewed, it is possible some participants drew upon knowledge from past experience of room sizes, particularly for height estimates. It has been suggested that most interior spaces come with standard heights (Henry and Furness 1993). Thus, estimates based on past experience may have resulted in a similar distance estimates performance on a small display and on a large display.

Examination of results indicates large variability in the data, particularly for transverse or length estimates. Although the display questionnaire did not ask participants to rate the display size based on the distance estimate tasks, participants' comments indicated that not all participants preferred large display over small display. Some indicated they rated small display higher than large display because there were more familiar with small display (more likely desktop monitor since participants were student and staffs of computer science department). Thus, participants' familiarity with the display size might have influenced their performance and their results of distance estimates in interactive images.

### 8.2.3 Comparison among asymmetrical distances

Consistent with previous findings (Witmer and Kline 1998, Henry and Furness 1993, Lampton, Bliss et al. 1994, Witmer and Sadowski 1998), distances were generally underestimated in the real and VE images with some exceptions. For example, distances tended to be overestimated for

- vertical and horizontal distance (large display) in static images
- vertical distance (large/ real condition) in dynamic images
- horizontal and transverse distance (large and small display condition) in interactive images

Consistent with results of previous investigations, vertical tended to be overestimated (Dixon and Proffitt 2002, Yang, Dixon et al. 1999, Higashiyama and Ueyama 1988). Yang and colleagues (1999) suggested that vertical overestimation would increase when projected on a large display. Current results provide support for this assertion. Whilst these results may suggest that different distance type may yield different results, large variability among data (particularly for interactive images) may also explain current results.

Consistent with the results of Henry and Furness (1993), across all distances, the general trend was that vertical distance was estimated more accurately compared to horizontal and transverse distance. Transverse distance yielded the worst performance. This is supported by the post-test questionnaire results whereby participants found vertical distance was easier to estimate compared to transverse distance. Henry and Furness (1993) suggested veridical estimates for vertical distance were in part due to the fact that participants were more familiar, using their own height as scale for estimates compared to other objects. This is also supported by our post-test questionnaire results where participants commented using their height to assist them in their estimates. Findings by Loomis, Da Silva et al. (1996) indicated that more estimation error was made on transverse distance compared to horizontal distance and this error is magnified when distance length increases. Current results also indicate that estimation error increases with increases in distance length.

It should be noted that for Experiment 1 and 2, the stimulus was based on outdoor setting and the objects used for distance estimates were different from the commonly employed objects used for distance estimate. The use of familiar objects might have influenced participants' estimates as they may have relied on knowledge from past experience to perform the distance estimates (see next paragraph). However, this is not necessarily true, objects such as trees and hedges may differ in sizes and heights, thus participants may not be able to draw upon knowledge from past experience.

Examination of individual distance results in Experiment 2 suggest that object types, object position in the scene and distance from the viewers are potential contributing factors to distance estimate task performance. However, further investigation is necessary to support this hypothesis.

For distance estimate task, the literature suggests that in the real world, distance estimates conducted in indoor setting was different from those conducted in outdoor environments (Teghtsoonian and Teghtsoonian 1969, Teghtsoonian and Teghtsoonian 1970). The results of these studies which were conducted in the real world environment showed an overestimation for indoor settings regardless of range and an underestimation for outdoor setting. A more recent investigation (Messing 2004) revealed that there was a reliable difference between distance estimates in indoor setting and in outdoor setting. Messing (2004) showed that distance estimates in indoor setting are more accurate than in outdoor setting, though both were underestimated.

Comparisons of the range of estimates for all asymmetrical distances between Experiment 1 and 2 (outdoor settings) to those of Experiment 3 (indoor settings) revealed that generally distance estimates were more accurate in Experiment 3 (indoor setting) (see Table 8-3). This provides support for Messing (2004) study's results. With respect to Teghtsoonian & Teghtsoonian's results, a similar pattern of results in terms of distance estimate size were found, whereby larger estimates were found in indoor settings compared to outdoor settings. This implies that the trends of results for distance and spatial memory performance in the VE model of an outdoor setting and indoor setting was similar to the real world outdoor and indoor setting performance. Consistently, this trend of results is also found in the findings of Messing (2004).

**Table 8-3 Comparison of range of estimates among experiments in outdoor and indoor setting**

Distance type	Experiment					
	Outdoor settings				Indoor setting	
	1A	1B	2A	2B	3A	3B
Vertical	-	93-145%	71-83%	74-106%	90-103%	93-94%
Horizontal	82-96%	44-88%	60-81%	60-86%	76-90%	~109%
Transverse	~ 49%	20 -75%	44-62%	37-51%	67-91%	102-108%

### 8.3 OVERALL ANALYSIS BASED ON SPATIAL MEMORY TASK

#### 8.3.1 Comparison based on environment types (Real and VE)

Corroborating the findings of several investigators (Arthur, Hancock et al. 1997), in terms of spatial memory performance, Experiment 3A's results revealed that there is no significant difference between the real and VE conditions. Arthur and colleagues suggested that spatial representation formed from interaction with small scale VE is comparable to real world experience. A similar finding by Richardson, Montello et al. (1999) also indicates participants' performance between the real and the VE is similar when a simple single floor

environment is used compared to a more complex building. However, on the latter, Richardson and colleagues found a significant difference between the real and VE conditions. Thus, the non significant difference in performance between the real and VE conditions found in Experiment 3 may also due to the use of a small and simple VE.

Several researchers (Bakker, Werkhoven et. al 1999; Chance, Gaunet et al. 1998) indicated that spatial orientation abilities largely deteriorated when the non-visual sensory modalities (such as vestibular or proprioceptive cues) are not or insufficiently simulated in the VE.

Other researchers (Gaunet, Vidal et al. 2001) however have pointed out that active exploration with a joystick share some important aspects with walking in the real world. They further suggested as in the real world, there is a tight connection between visual self-motion and motor-activity when using joystick. Thus, the process of gathering information may be similar in both the real and VE conditions. Similarly, this may be true when a mouse and a trackball are used. Alternatively, this might explain the similar performance between the real and VE conditions in Experiment 3A.

In the real world, the proprioceptive cues were provided by walking. However, in the VE, there is a mismatch between the visual and vestibular cues, where the visual cues indicate movement and the vestibular cues indicate that the participants were stationary. It has been suggested this mismatch may result in users feeling nausea which may affect their performance (May and Badcock 2002). However, none of the participants reported such feelings. This may be due to the use of non immersive and semi immersive projected display.

Moreover, the result of a recent investigation (Willemson, Calton et al. in press) suggests that the mechanical aspects of the HMD (an immersive display) may explain the inaccurate user's spatial perception of the VE. Similar argument have been used by Plumert, Kearney et al. (2004) to explain their results of non-significant difference between real and VE conditions on distance estimate task.

Another possible explanation is some form of proprioceptive feedback given by the muscle movement of the wrist and arm and shoulder for mouse and trackball might compensate for the missing cues. This suggests that the use of these input devices may be minimally sufficient to provide proprioceptive feedback necessary to indicate movement. Several studies have shown that the use of a more natural walking interface is no better than using a joystick (Witmer and Kline 1988, Grant and Magee 1998). Moreover, the flexibility of the human sensory system might partially account for these results. In fact the visual sense (without even



moving) is enough to provide the user with a compelling sense of movement (Harris and Jenkin, et al. 2002). Thus, users may have adapted to the movement perceived in the VE to represent their movement.

Whilst not significant, performance in the VE condition (except trackball/drive) tended to be better than in the real world condition. The longer viewing time and thus more experience in the VE compare to the real world condition was suggested as one possible explanation. It has been suggested the acquisition of spatial knowledge increased with increase exploration time or displacement in the VE (Peruch, Vercher et al. 1995). Alternatively, as the viewing area of the VE conditions is confined to the screen, this allows VE participants to focus on object locations. This is not the case for the real world participants; they need to move about to find objects. The need to move about might have imposed more mental demand on the real world participants, thus less cognitive resources is available for remembering object locations. All these might contribute to the slightly better performance of VE participants over real world participants.

With regards to the exception case of track/drive condition, as will be discussed later in Section 8.3.3, trackball was rated lower compared to mouse in the interface device questionnaire. While drive mode was easier to use and control but fly mode was rated higher on ease of object location recall. Despite the extra viewing time and experience in the VE, the combination of trackball and drive mode may explain VE participants' poor performance over real world participants.

### 8.3.2 Comparison based on display size (Large and Small)

The results of Experiment 3A revealed no significant difference for spatial memory task performance between a large and a small display. However, after removing the effect of covariates (sport background variable and computer games experience variables) it was shown that there was a significant effect of display size factor ( $p = .042$ ) with performance on a small display is better than on a large display. Since the experiment setup of Experiment 3A was similar to those of Experiment 1B and 2B, it was expected that there would be no significant difference between the large and small displays. Thus, by comparing Experiment 3A's result with the results of previous investigations (Patrick, Cosgrove et al. 2000, Czerwinski, Tan et al. 2002, Tan, Gergle et al. 2003), it is not possible to conclude that the main effect of display size in previous investigations was partially influenced by viewing distance and physiological cues. These unexpected findings had motivated the undertaking of Experiment 3B which replicate the setup of previous investigations in terms of fixed FOV and

varying viewing distance (and physiological cues). This setup was also similar to those of Experiment 1A and 2A. This setup allowed us to determine whether the results of the previous investigations could be reproduced by Experiment 3B.

However, the results of Experiment 3B showed no significant difference between the large and small display. Removing the effects of covariates also did not change the picture. When the FOV was fixed (viewing distance, viewing distance and physical display size were varied) results showed no main effect of display (Experiment 3B) but when the FOV was varied (viewing distance, physiological cues and physical display size was fixed) results showed there was a main effect of display (Experiment 3A). The results of Experiment 3A and 3B suggest that

- the better performance of the small display over the large display was influenced more by the FOV of the display.
- Moreover, when the effects of covariates (sport background and computer games experience) were removed, results showed that the interaction effect of display size, device type, travel mode and sport background factors approached significant level. This indicates other factors such as the device types, travel modes and sport background do contribute some influence to spatial memory task performance in VE.

The better performance of the small display over the large display is unexpected given the results of previous investigations which reported better performance on a large display compared to a small display (Patrick, Cosgrove et al. 2000, Czerwinski, Tan et al. 2002, Tan, Gergle et al. 2003). Some researchers have indicated that a wide FOV enhances a user's sense of presence as well as performance (Prothero and Hoffman et al 1995, Kline and Witmer 1996, Arthur 2000, Duh, Linh et al. 2002). The results of these studies also indicate that more sense of realism is experienced in wide FOV images compared to a narrow FOV.

Theoretically, a wide FOV display would closely match the human FOV compared to a narrow FOV display. Therefore, it is expected that performance in the VE would be similar to the real world when a wider FOV display is used. Thus, the better performance of a small display over a large display in Experiment 3 is contrary to expectation.

Fortunately, the results of display questionnaire from Experiment 3B yielded some useful and important information which provide explanation for Experiment 3 results. The display questionnaire results showed that while participants generally preferred a large display over a small display, surprisingly in terms of *ease of object recall* and *confidence rating*, participants rated a small display better than a large display. In terms of subjective comments, half of the participants positively commented on the small display and the other half positively

commented on the large display (see Section 7.2.2.1 of Chapter 7). The participants' opinions reflected that a large display does not necessarily yield the same impact on all viewers. Some viewers might perform better on a small display, while others may perform better on a large display. While it is not clear why this is so, subjective comments from the participants may suggest that participants' familiarity with small display (possibly the desktop monitor since participants were students and staff of the Computer Science Department) may have influenced their performance.

### 8.3.3 Interface device

Consistently, the results of Experiment 3A and 3B showed that participants' spatial memory task performance was better when using a mouse than using a trackball. For Experiment 3A, the difference is significant ( $p = .024$ ) but for Experiment 3B the interaction effect of display size, device type, travel modes and sport background factors which approached significance ( $p = .067$ ) indicated that spatial memory task performance in interactive images was partially influenced by these factors. The results of the interface device questionnaire provide support for the better performance of a mouse over a trackball. The questionnaire results showed that participants ranked mouse higher than trackball in terms of the following:

- ease of use
- ease of movement and self-positioning in the VE
- afford more control of movement
- ease of recall for object positions
- usage preference

The extra proprioceptive cues derived from the movement of the mouse compared to the static position of the trackball might have accounted for the better performance of the mouse over the trackball. Although there is a conflict between the visual cues (which indicate there is movement in the VE) and vestibular cues (which indicate the participant is in a stationary position), the participants appeared to adapt their movement based on what they saw. Additionally, participants reported that movements using a mouse resulted in a better sense of corresponding visual movements on the screen compared to when using a trackball. This indicates that the participants were better at relating their movements on the screen using a mouse compared to using a trackball.

However, the difference in spatial memory performance resulting from the use of a mouse and a trackball may suggest that the missing sensory cues not simulated in the current VE model may have contributed some influence on participants' performance to some degree in

the VE. In terms of display size, the ratings for both display sizes were not statistically different.

### 8.3.4 Travel mode

In terms of travel mode, the difference between a drive and a fly mode was very small for spatial memory task performance. Whilst not statistically significant, both Experiment 3A and 3B showed that a fly mode allowed participants to perform slightly better than a drive mode.

Participants' comments in the interface device questionnaire provide support for the better spatial memory task performance using fly mode over drive mode. Even though participants commented that the drive mode allowed them to move and control movements easily in the VE in the interface device questionnaire, and they even choose the drive mode over the fly mode for overall preference, however on *ease of object recall* they rated the fly mode higher than the drive mode.

The extra degree of freedom afforded by the fly mode allowed participants to have an overview of the room and the objects spatial relations. Even though this extra degree of freedom might have incurred more cognitive demand on the user (Ruddles and Jones 2001), Experiment 3's results indicate that the map view provided by vertical movement resulted in overall better performance in the spatial memory test for fly mode. This implies that the more familiar method of movement of drive mode does not necessary result in a better spatial memory task performance in the VE. The "unnatural" movement of the fly mode (for human locomotion) in the real world is more beneficial in the VE compared to the drive mode.

For Experiment 3B, results showed the interaction effect of display size, device type, travel mode and sport background factors approached significant. This suggests that for interactive images, besides display size other factors such as device type, travel modes and participants' sport background were also contributing factors toward spatial memory task performance.

## 8.4 CONCLUDING REMARKS

In this chapter an overall analysis of the results from the three experiments conducted in the research presented in this thesis is presented. To facilitate comparisons and discussions of results, the overall analysis were based on the task performance measures examined in the experiments: distance estimates tasks and spatial memory tasks (including interface device

and display questionnaire results). A summary of results based on distance estimate task (Table 8-1) and spatial memory task (Table 8-2) were given to guide reading.

For distance estimate task, overall analysis was based on comparisons of distance estimates performance in image types and in display types for all experiments (Experiment 1, 2 and 3). The discussion was related to research questions 1 and 2 (see Section 4.1.4 of Chapter 4) of this thesis with respect to distance estimate tasks. Additionally, comparisons of performance among distance types (vertical, horizontal and transverse) were also presented.

For spatial memory tasks, analysis was based on comparison of spatial memory task performance in image types, display types, interface device and travel modes. The discussion was related to research questions 1, 2, 3 and 4 (see Section 4.1.4 of Chapter 4) of this thesis with respect to spatial memory tasks. Discussion of these results includes the results of post-test questionnaire (Interface device and display questionnaire) which provide additional information to support and explain the findings.

In the next chapter (Chapter 9), the major findings and contributions of the research are highlighted. The implications of these findings towards spatial awareness perception in the VE and VE applications are also highlighted.



# CHAPTER 9

---

## *FINAL CONCLUSIONS, RESEARCH CONTRIBUTIONS AND IMPLICATIONS FOR SPATIAL AWARENESS PERCEPTION IN VE*

### 9 OVERVIEW

This chapter is organized into two major sections. The first section presents the overall conclusions and research contributions. This includes discussions on the major findings from the three experiments and the impact of image modelling on the conclusions drawn. These results are considered with respects to the key research questions being proposed. The methodological contributions in terms of the approach to investigate the display related factors examined in this thesis are also highlighted.

The second section provides discussions on the implications of these experimental results on spatial awareness perception in VE. This includes a discussion on the associated impact on VE related applications.

## 9.1 FINAL CONCLUSIONS AND CONTRIBUTIONS

The research presented in this thesis has examined users' spatial awareness performance in the real and VE by evaluating factors influencing spatial performance in both environments. Factors related to display such as display size, viewing distance, physiological cues, interface device and navigation methods were investigated. These factors were examined in the context of static, dynamic and interactive environments presented to the users in non-stereo, non-immersive and semi-immersive display. Distance estimate tasks (in terms of asymmetrical distance) and spatial memory tasks were identified as task performance measures. The key research questions addressed in this research were

1. *Is there a difference in spatial tasks (distance estimation and spatial memory task) performed in real and VE?*
2. *How does the display size (large and small) affect a user's spatial task (distance estimation and spatial memory task) performance in real and VE?*
3. *How does the type of interface device (mouse and trackball) affect a user's spatial task (spatial memory task) performance in VE?*
4. *How does the type of travel mode (drive vs. fly mode) affect a user's spatial task performance (spatial memory) in VE?*

These research questions were empirically explored through the testing of a series of hypotheses in three sets of experiments (see Chapter 5, 6 and 7).

In the following subsections, the major results from these experiments were organized and presented based on performance task measures used: distance estimate and spatial memory task. First, the major results from the experiments on distance estimates tasks in static, dynamic and interactive images are presented. The major results from the studies on spatial memory tasks in interactive images were presented next. These results were then highlighted in relations to the four key research questions proposed in this thesis. The effect of image modelling on the results of real and VE comparisons was presented next. Finally, the methodological contributions of the research are also highlighted.

### 9.1.1 Distance estimates tasks

#### 9.1.1.1 Experiment on static images using distance estimate tasks

With regards to the distance estimate tasks (horizontal and transverse distance), Experiment 1's results suggested that it is possible within the constraint of this experiment to perceive real

and VE images similarly based on asymmetrical distance estimate tasks (horizontal and transverse) indicating that both image types provide similar information for distance estimate tasks to the observer. Although this result contradicts the conclusions of some researchers who reported a significant difference in distance estimate performance between real and VE (Henry and Furness 1993, Lampton, McDonald et al. 1995, Waller, Hunt et al. 1998, Witmer and Kline 1998, Sinai, Krebs et al. 1999, Loomis and Knapp 2003), this result is consistent with those who reported no significant difference of distance estimate performance in real and VE (Waller 1999, Yoon, Byun et al. 2000, Yang, Dixon et al. 1999).

In terms of display size, the results from Experiment 1A on static images showed distance estimate performance on a large display were better than on a small display for both horizontal (significant at 10%,  $p=.052$ ) and transverse distances (significant,  $p < .05$ ).

Interestingly, for horizontal distance, a direct comparison of means indicated that the VE image participants performed better than the real image participants on a small display but for real image distance estimate performance was better on a large display.

In contrast to horizontal and transverse distance, Experiment 1B revealed that vertical distance was more accurate on a small display than on a large display. Vertical distance tended to be largely overestimated on a large display, which confirmed the prediction of Yang, Dixon et al. (1999) and provide support for others (Dixon and Proffitt 2002).

The results from both Experiment 1A and 1B suggest that besides display size, other factors such as viewing distance and physiological conditions also contribute to the better performance of large display over small display. However, the large variance attributed by the display size suggested that the display size also constitutes a major influence on distance estimate tasks in static images.

#### 9.1.1.2 Experiment on dynamic images using distance estimate tasks

Contrary to the results of experiment on static images but consistent with the results of previous findings (Henry and Furness 1993, Witmer and Kline 1998, Witmer and Sadowski 1998, Lampton, McDonald et al. 1995), the results of experiments on dynamic images (Experiment 2) showed that the distance estimate between the real and the VE images were statistically difference for all asymmetrical distances ( $p < .05$ ).

The use of a low image resolution for the VE image on a small display compared to the real image resulted in better performance of the VE image participants over real image

participants but the opposite was true on the large display. However, this constraint appeared to affect horizontal and transverse distance only.

For vertical distance, the use of a low image resolution for the VE image compared to the real image have not effect on performance on displays size. On average, performance of VE image participants was better than the real image participants on both display sizes.

Employing a high image resolution for the VE image appeared to be more beneficial on large display; results showed that distance estimate task performance were more accurate on a large display compared to a small display for all asymmetrical distances (vertical, horizontal, and transverse distance).

The use of a low image resolution for VE results in a slightly better performance on a small display over a large display for all asymmetrical distances. This was also true for real image conditions.

However, in terms of display size factor the use of a high image resolution for the VE condition yielded more accurate distance estimates on a large display compared to on a small display for all asymmetrical distances. For the real image conditions, this was true for vertical and horizontal distances only. However for transverse distances, more accurate distance estimates were found on the small display compared to on the large display.

The results from Experiment 2 enable us to conclude that other factors besides display size factor such a viewing distance and physiological cues also contributed to the result of better performance of large display over small display. Moreover, for dynamic images, image resolution was indicated as another important factor affecting distance estimate performance:

- A very low image resolution would degrade distance estimate performance when presented on a large display and would be better presented on a small display for improved performance. However, this is true for horizontal and transverse distance only.
- Vertical distance does not appear to be affected by low image resolution for the VE image, distance estimate performance for the VE image was better than real image on either display sizes.
- On a large display, an image of high resolution is necessary to elicit better distance estimate performance. It was shown that the difference in image resolution between the real and VE in Experiment 2A and 2B was sufficiently large to promote a difference in perception between both environments but the minimal level of image resolution

necessary for the VE image before performance degrades would require further investigations.

The varying effects of factors for each asymmetrical distance (vertical, horizontal and transverse) highlighted the importance of examining each of these distances. Generally, consistent with previous investigations, vertical distance is estimated more accurately compared to horizontal and transverse distance. Transverse distance yields the worst estimate. The impact of these results was discussed in Section 9.1.1 (Chapter 9).

It was mentioned in Chapter 4 that the type of objects used for distance estimates in Experiment 1 (static images) and 2 (dynamic images) were not typical of those commonly employed in previous investigations. Instead of employing the commonly used stimulus such as poles, columns or discs, we employed the naturally occurring objects in the scene. It was expected that user's familiarity with the object may provide the participants with extra information for distance estimate task. In retrospect however this may be true for some objects only. Other objects such as trees and hedges and even lampposts, which are of various sizes and lengths, could not have provided the participants with any clue to distance estimate unless participants assume that certain objects (such as lampposts, signpost or road) are of certain distances. As such, current findings would still present a reasonable comparison with past investigations. It is noted that for transverse distance (comparable to egocentric distance, see Section 5.1.6 of Chapter 5), additional information from familiar objects does not seem to affect distance estimate performance as reflected by the gross underestimation (up to 20% of actual for larger distances).

### 9.1.1.3 Experiment on interactive image using distance estimate tasks

Whilst similar to the results of experiment on static images but inconsistent with the results of most previous investigations (Henry and Furness 1993, Witmer and Kline 1998, Witmer and Sadowski 1998, Lampton, Bliss et al. 1994, Lampton, McDonald et al. 1995) that reported significant differences between real and VE, the results from interactive image experiments (Experiment 3A and 3B) revealed no significant difference for distance estimate task performance between the real and the VE conditions. However, the results of a very recent investigation (Plumert, Kearney et al. 2004) do agree with our findings. The results from Experiment 3A and the result from Plumert, Kearney et al. (2004) suggest that it is possible to perceive real and VE conditions similarly in terms of distance perception when the VE is presented on a large projected display. The current results also imply that the use of a simple environment (such as a single room) may account for the similar performance between real and VE conditions. Additionally, the results also suggest that more practice and experience in



the VE may have helped improved participants' performance to be similar to the real conditions.

One unexpected finding from Experiment 3A investigation was the slightly better performance of VE/small conditions over real conditions for all asymmetrical distances. As explained in the previous paragraph, practice effects may contribute to this difference. Alternatively, another possible explanation is the influence of more cognitive demand on the real participants who need to focus more on physically moving about in the environment (Ruddles and Jones 2001) rather than on the assigned task. Additionally, the presence of more details in the real environment may imposed more cognitive workload on the real participants (Yanagisawa and Akahori 1999), thus less cognitive resources were available to the real participants for the required task. Subsequently, this may affect real image participant's distance estimate performance.

Contrary to the conclusions drawn from previous investigations (Henry and Furness 1993, Witmer and Kline 1998, Witmer and Sadowski 1998, Lampton, McDonald et al. 1995), the analysis of Experiment 3A and 3B results suggests that the physical display size does not appear to contribute largely towards distance estimate task performance in interactive images at least within the experiments' constraints. Similarly, the results imply that the contribution of viewing distance (and physiological cues) and FOV were also considered minimal for the distance estimate task in interactive images.

Another unexpected finding is that on average, distance estimate performance on a small display is slightly better than on a large display. Participants' reliance on previous knowledge to estimate instead of basing judgment on the viewed image has been suggested as a possible explanation. The results from the display questionnaire indicate that for interactive images a large display may not necessary yield better performance; for some users, familiarity with small display may similarly improve their performance on a small display over a large display.

### 9.1.2 Spatial memory tasks in interactive images experiment

Spatial memory tasks were only investigated in the interactive images experiment (Experiment 3). Unlike experiments on static and dynamic images, in this study participants were allowed to control their movement in the VE and as such the spatial memory task represented a suitable measure of spatial representation (see Section 4.1.1.2 of Chapter 4).

The results of Experiment 3A revealed no significant difference in spatial memory task performance between the real and the VE conditions suggesting that it is possible to perceive

real and VE in terms of spatial representation. Thus the spatial representation knowledge formed in the VE is similar to those formed in its real counterpart at least within the constraint of this study. This result is consistent with those reported by other studies (Arthur, Hancock et al. 1997, Richardson, Montello et al. (1999).

Interestingly, general performances in the VE conditions (except for trackball/drive condition) were slightly better than in the real conditions. By using the same arguments to explain the better performance of VE conditions over real conditions for distance estimate tasks in interactive image, practice effects, more experience in the VE, and a reduction in cognitive demand for the VE participants compared to the real participants may also explain the better performance of VE conditions over real conditions on spatial memory tasks.

The results from Experiment 3A showed there was a statistically significant difference ( $p < .05$ ) between the large and small display for spatial memory tasks in which performance on a small display was better than on a large display. Given the theoretical considerations and the results of previous studies (see Section 4.1.2.1 of Chapter 4), the result of better performance of a small display over a large display was quite unexpected. However, further insights provided by the subjective responses and comments from the participants in the display questionnaire may suggest this is possible. It was also indicated from the display questionnaire that participants' familiarity with a small display may favourably influence their performance on this display size. Further experimentation however is necessary to confirm this assertion.

Additionally, the results from Experiment 3B suggested that beside display size, the better performance of small display over large display was influenced by other factors such as device types, travel modes and participants' sport background. The results from Experiment 3A and 3B suggested that the better performance of a small display over a large display was contributed largely by FOV.

Generally, participants' spatial memory task performance using a mouse was significantly ( $p < .05$ ) better compared to using a trackball. Results from Experiment 3B suggest that the better performance of mouse over trackball was influenced by other factors such as display size, travel mode and participants' sport background. Computer experience did not appear to have a large influence. A more likely explanation is that all the participants were staffs and students of the Department of Computer Science who used the computer daily and in this sense they are "equal" in terms of using the computer (and most likely using a mouse too).

The subjective responses from the interface device questionnaire provided support for the better performance of mouse over trackball:

- Participants ranked mouse higher than trackball in terms of ease of use, ease of movement and self-positioning in the VE, more control of movement, ease of recall for object positions and overall preference.
- Participants' comments indicated they could relate their movement in the VE to the movement of the mouse device. This may suggest that the extra proprioceptive cues derived from the movement of the mouse (compared to the static trackball) could offer another plausible explanation for this result. This implies that the missing sensory cues not present in the VE model (such as vestibular and proprioceptive cues) do contribute an influence on participants' spatial memory performance.

Additionally, participants familiarity with the mouse compared to the trackball may have also contributed to the result.

The effect of travel mode on spatial memory task performance is not significant in interactive VE. However, performance using the fly mode was slightly better than using the drive mode. The significant interaction effect suggests that this result was influenced by other factors such as display, device and participants' sport background.

The extra degree of freedom (vertical movement) afforded by the fly mode which provide participants with an overview of the room have been suggested to improve participants' spatial memory task performance in the fly mode. The slightly better performance when using fly mode was further supported by the subjective responses in the post-test questionnaire whereby participants rated the fly mode higher than the drive mode on *ease of object recall*.

The results of Experiment 3 suggest that the more familiar method of movement of drive mode does not necessary yield better spatial memory performance; the unnatural movement of fly mode may yield better results. This may further imply that, in terms of movement method, it is not always necessary to closely mimic real world movement to improve performance.

### 9.1.3 Effect of image modelling on real and VE comparisons

The results of similar distance and spatial memory performance between the real and the VE at least for static and interactive experiment may in part be influenced by the VE model used. The techniques employed to model the VE appear to be minimally sufficient to at least yield some level of realism to invoke on the viewer similar spatial perception (distance estimate and spatial representation) to its real image/real world counterpart. The conclusions drawn from

these studies may not extend to other VE models modelled using different techniques. Considering the need to make trade-offs between image realism and computing resources for real-time VE, it would be of interest to compare the effect of using different techniques to create different levels of realism on spatial awareness.

#### 9.1.4 Key Research Questions Addressed

As mentioned earlier, there are four main research questions directed towards understanding spatial awareness in the real and VE in this thesis. These questions were examined in three series of experiments in the context of static, dynamic and interactive presentations.

The first question addressed the underlying premise whether it is possible to perceive spatial awareness in terms of distance perception and spatial memory tasks in the VE similar to its real world counterpart.

- Findings from the first experiment on distance perception in static images indicated that it is possible to perceive static pictures of real and VE in terms of these asymmetrical distances of horizontal and transverse distance.
- Results from the experiment on distance perception in dynamic images however suggested that image resolution played a significant role in user's distance perception performance in both the real and VE. When the real and VE image differ largely in terms of image resolution, there is a significant difference between the real and the VE. This may provide explanation for previous investigations' results that showed a difference in user's distance estimate performance between these environments.
- Contrary to the results of previous investigations, findings from the experiments in the context of interactive presentations revealed no significant difference for distance estimate and spatial memory tasks performed between the real and VE conditions. A more recent investigation (Plumert, Kearney et al. 2004) however provides support for this result. Results further showed that other factors such as more practice, more experience and low cognitive workload may have contributed towards the improved user's distance estimate and spatial memory performance in the VE over the real conditions.

The second research question examined the impact of display size factor on user's spatial task performance in the real and VE. It was postulated (see Section 4.1.2.1 of Chapter 4) that the results of better performance of large display over small display by previous investigations were also influenced by other factors such as viewing distance and physiological cues.

- The result of the first sub-experiment (Experiment 1A) on distance perception in static images indicated that display size contributed a major influence on user's better

performance of the large display over the small display for horizontal and transverse distances. This result provides support for previous investigation findings.

- However, the result of the second sub-experiment (Experiment 1B) suggests that viewing distance and physiological cues also contributed to this result.
  - In contrast to the horizontal and transverse distances, findings further indicated that the vertical distance was estimated slightly better on the small display compared to the large display.
  - In support of previous investigation, vertical distance tends to be overestimated on a large display.
- Findings from the experiment on distance perception in dynamic images indicated that the image resolution played a significant role in user's distance perception performance on both display size (large and small).
  - Result shows that the use of a higher resolution image for the VE condition compared to the real condition produced more accurate distance estimates on the large display for all asymmetrical distances.
  - Results further suggested that a low resolution real (or VE image) is better presented on a small display for improve asymmetrical distance estimate performance and a high resolution image is necessary to improve asymmetrical distance performance on large display. However, this is valid only for horizontal and transverse distance; for vertical distance, low image resolution does not influence performance in either display size.
  - Findings confirmed the results of previous investigation which showed better performance of a large display over a small display was influenced by display size factor.
  - However, current study also indicated that other factors such as viewing distance and physiological cues also contributed to these results.
- Result showed an unanticipated finding of the non-significant effect of display size on the distance estimate tasks in the interactive images.
  - It was also indicated that the contribution of FOV, viewing distance and physiological cues were considered small for distance estimate tasks in the interactive image.
  - It was shown that distance estimate performance was slightly better on a small display over a large display. Subjective responses from the display questionnaire appeared to suggest that users' familiarity with a small display may improve users' distance estimate performances on a small display over a large display.



- Findings from the experiments on spatial memory in the interactive image revealed an unexpected finding that spatial memory performance is more accurate on a small display compared to on a large display.
  - It was shown that a combination of factors such as device types, travel modes and sporting background all contributed to this result.
  - Result further indicated that the better performance of the small display over the large display was largely influenced by the FOV factor.
  - Subjective responses from the display questionnaire suggested that a large display does not necessarily results in a better spatial memory performance. It showed that familiarity with a display size may be partially responsible for the better user's spatial memory performance on a small display over a large display.

The third and fourth research questions explored the impact of interface device and travel mode on user's spatial memory performance in the VE respectively. These questions were investigated in the third experiment which examined spatial awareness in interactive images.

- Findings indicated that using a mouse resulted in a better spatial memory performance than using a trackball. This was also reflected in the interface device questionnaire data; implying a parallel between the objective and subjective responses data.
- Results also suggested that a familiar method of movement such as a drive mode does not necessary yield better spatial memory performance. It was shown that the unnatural movement method of flying yielded slightly better spatial memory performance over drive mode. Subjective response produced by the interface device questionnaire provides support for this result in terms of *ease of object recall*.

### 9.1.5 Scope of conclusions

Findings from these investigations are limited to within the experiment's scope and constraints only, thus should be considered and interpreted within the controlled conditions:

- The image is presented to the viewer in non-stereo mode only, thus the results are valid for monoscopic vision only
- The image is also presented in a non-head-tracked, non-immersive and semi-immersive conditions, thus the influence of motion parallax cues on the result is not investigated.
- It should also be noted that the values of the variables (display size, viewing distance) investigated were limited to two sizes and thus the conclusions drawn from these research are limited to these values only. These results may not necessary apply to other sizes (but see Chapter 10 for future work).

However, these investigations could benefit from further improvements and much wider interpretations. For future studies, several recommendations which include expanding on the research scopes and limitation and improvement on experiment methods are proposed in the next chapter (Chapter 11).

### 9.1.6 Method contributions

Methodological contributions in this thesis are concerned with the proposed research approaches or methods employed in examining factors affecting spatial awareness in the real and VE. The following highlight some of the main contributions in terms of approach and methods in this thesis:

- The results of previous investigations suggested that the better users' spatial performance on a large display over a small display was due to the physical display size. However, a review of the literature and theoretical considerations suggest that other factors (such as viewing distance and physiological cues) may have influenced these results and thus conclusions drawn from these investigations. As such in this thesis the use of two related studies to examine the effect of display size and the possible influence of these other factors was proposed. The experimental setups for both studies to enable such investigations were described in Section 4.1.2.1 and Section 4.3.1.2 of Chapter 4.
- Most related studies on display size factor focused on objective evaluation only. It was shown in this thesis that further insights and understanding of unexpected findings in this thesis (such as in Experiment 3 on interactive images) would not have been possible if the experimentation was based solely upon objective evaluation. Thus this thesis highlights the importance to include subjective evaluation in addition to objective evaluation in experimentation.
- This thesis provides detailed examinations of distance estimate tasks in terms of asymmetrical distances (vertical, horizontal and transverse). Instead of investigating distances in terms of egocentric or exocentric distance as typically done previously, the detailed breakdowns into these three individual distances yielded more important information about the different effects of the factors on each of the distances as shown by empirical results presented in this thesis and as suggested by the literature.

## 9.2 IMPLICATION FOR SPATIAL AWARENESS PERCEPTION IN VE

The research work presented in this thesis has examined several factors affecting users' spatial awareness in the VE through a series of empirical investigations. In this section, the implications of the findings from this research work in VE in terms of the two tasks examined in this thesis are presented in two subsections. The first subsection discusses the implication

of experimental results from investigations on distance estimate tasks while the second subsection discusses the implications of results from investigations on spatial memory tasks.

### 9.2.1 Implications of results from distance estimate task performance

Information about depth and distance about objects are very important for some applications such as flight training, tele-operation of robots, visualization of complex data sets, product visualization, medical training and crime scene applications (Surdick, Davis et al. 1997). In fact the effective use of such applications relies on the VE technologies to provide such accurate information. The implications are discussed based on the key variables investigated: image type, display size, interface device and travel modes

#### 9.2.1.1 Image type

The results from the experiments which examined distance estimation on static and interactive images (Experiment 1 and 3 respectively) suggested that there is no significant difference between real and VE conditions. Thus, at least within the constraint of these experiments it is possible to perceive VE similar to its real counterparts in terms of distance estimate perception. These results may provide assurance for current and potential application users of VE technologies in terms of similar distance perception in both static and interactive environments.

The results of distance perception in dynamic images (Experiment 2) however showed that when there is a sufficiently large difference between both environments in terms of image resolution, distances are perceived differently in both environments whereby the better performance of which image (real or VE) is dependence upon the types of asymmetrical distances (see Section 8.2.1 of Chapter 8). The implications from these results is that the significant difference of distance perception in previous investigations may be partially attributed to the use of low image resolution or the use of less realistic VE models compared to the real world. Indeed, Witmer and Kline (1998) suggested the large perceptual difference between the real and VE performance may be due to the difference between the VE model from its actual real world space. For example some of the features in the real world were not modelled in their VE model. Other researchers (Jää-Aro and Kjell Dahl 1997) have indicated that a poor image may degrade distance judgment performance while others (Willemson and Gooch 2002) have shown otherwise. These contradicting conclusions may suggest the difference in results could be due to the level of image resolution or realism used in these studies. There may be a minimal level of image resolution before performance degrades in

VE. Although, findings from Wright (1995) suggest that to improve image quality alone might not result in more accurate distance estimate, suggesting other factors may also be involved. However, identification of this minimal image resolution requirement would be beneficial to guide VE designers' decisions.

As expected, additional implication from the results of Experiment 2 showed that low resolution VE images would benefit from small display presentations. For higher resolution VE images, presentation on large displays would benefit from the combined benefit of large display and better image quality. However, as mentioned earlier, our investigations show that this is only true for horizontal and transverse distance. For vertical distance, distance estimate performance is better on VE both display sizes for low image resolution but for high image resolution VE image performance is better on large display. The results from Experiment 2 additionally showed that image resolution is less influential on distance estimate performance when presented on a small display as indicated by the better performance of the low resolution VE image (when compared to the real image) over real image.

Although Experiment 1 and 2's results showed that a VE is perceived similarly to the real environment, however, the range of estimates indicates that distances were not accurately perceived when compared to the actual distance. Similar to the results of previous investigations (Henry and Furness 1993, Witmer and Kline 1998, Witmer and Sadowski 1998, Wright 1995, Eggleston and Janson et al. 1996), distances were generally underestimated. This implies users may make considerable distance judgment errors in VE. This inaccuracy raises a major concern especially for applications which rely on very accurate distance judgments for their success. As mentioned earlier, this encompasses a number of applications such as flight training, tele-operation of robots, visualization of complex data sets, product visualization, medical training, and crime scene applications. For training applications, these results imply that the distance judgment skills may not transfer well to the real world. For visualization applications such as product design and architectural design the virtual design may not translate accurately when the actual product is designed; it may be smaller (or larger) than expected.

#### 9.2.1.2 Display size

Consistent with the results of previous investigations (Patrick, Cosgrove et al. 2000, Czerwinski, Tan et al. 2002, Tan, Gergle et al. 2003), the results from experiments on static images (Experiment 1) suggest that distances on larger displays are perceived more accurately compared to small displays. This difference is statistically significant ( $p < .05$ ). This implies for static image presentation the size of display matters whereby larger display would results

in more accurate distance perception. The results from Experiment 1 further suggest that other factors such as viewing distance (and hence physiological cues) were also implicated as contributing factors to this better performance of a large display over a small display. This further suggests that viewing distance (hence the physiological cues) also influence distance perception in static images.

However, the results of Experiment 2 on dynamic images were partially influenced by image resolution. When a low image resolution was used (Experiment 2A), more accurate estimates were found on a small display than on a large display but when a high image resolution was used, distance estimate performance on a large display was better than on a small display (with the exception of transverse distance on real image whereby performance on a small display is better than on a large display). These differences however were not statistically significant ( $p > .05$ ). These results are expected because a low resolution image presented on a large display would result in a coarser and grainier image compared to when presented on a small display. However, the insignificant difference is unexpected. It is predicted if a high image resolution was used in Experiment 2A a significant result may be yielded whereby estimates on a large display is better than on a small display. Similar to Experiment 1, the results of Experiment 2A and 2B suggest that display size as well as viewing distance (and hence physiological cues) were also contributing factors to distance estimate performance in dynamic images. However, as with the impact on image type, the results of Experiment 2 also suggested that the direction of effect of display size was influenced by the levels of image resolution.

Contrary to the findings of some researchers (Willemson and Gooch 2002, Thompson, Willemson et al. in press) but consistent with the findings of other researchers (Kline and Witmer 1996, Jää-Aro and Kjell Dahl 1997, Duh, Lin et al. 2002), the results of Experiment 2 imply that the role of image resolution is important in influencing user's distance estimate performance on different display sizes for dynamic images.

For Experiment 3A and 3B on interactive VE, the results seemed to suggest that physical display size, viewing distance and physiological cues do not appear to have a significant influence on distance judgment performance, at least within the constraint of these studies. In line with the results of other researchers (Johnson and Stewart 1999, Arthur 2000), the results showed that distance judgments performed on a large display does not differ very much from distance judgment performed on a small display for interactive images. Although the effect of navigation was not directly investigated in this research, comparing the results of static and dynamic images to interactive VE appear to suggest that allowing participants to navigate in



the VE may reduce the differences of perceptual judgment between a large and a small display. If this assertion is accepted, this may imply that the size of display does not matter if participants were allowed to navigate in the VE.

The slightly better performance of the small display over the large display as reported by experiments on interactive images (Experiment 3) may suggest that large display does not necessarily result in better performance. These findings may be good news for application users whereby small and less expensive display may only be necessary for effective performance or presentations. However, subjective responses from the participants further indicate that user's familiarity using small display may have influenced their performance on such display. These results suggest that if a set of users were used to or frequently exposed to a small display, providing them with a large display may not help improve their performance. However, other factors may be involved in producing these results of better performance on a small display compared to a large display, thus implying the need for further investigations.

### 9.2.1.3 Asymmetrical distances

The results from Experiment 3 also showed that the influence of the factors such as display size, viewing distance, physiological cues, and image resolution on the perception of distance may vary depending upon the type of asymmetrical distances: vertical (or height), horizontal (or width), transverse (or length). Thus, designers should take specific account of these factors on different asymmetrical distances into consideration in their design.

Generally, vertical distance is significantly more accurate compared to horizontal and transverse distance where transverse distance is most often largely underestimated for longer distances. This variation of differences should receive careful considerations.

- This result may suggest that in VE, the height of objects may be perceived more accurately compared to its width or depth. In applications such as architecture, product and scientific visualizations, the objects or space may not be perceived accurately as intended whereby object's height may be perceived accurately but its width and length may be overestimated or underestimated. This may have more critical implications on other applications such as flight simulation applications whereby the altitude of planes may be perceived accurately but horizontal (lateral) distances and transverse (forward) distance may not be as accurate.
- Underestimation in transverse (forward) distance may suggest that a pilot may thought an object (such as runway) is near when it is still very far away.
- Overestimation of horizontal distance may lead pilots to think that another plane is still far away when actually it is near. Thus, the transfer of skill to the real world may not be

as intended. Fortunately, in flight simulators applications, pilots also received training in actual aircrafts.

Even though, no equal lengths of vertical and horizontal distances were compared directly in all experiments, the results from static, dynamic and interactive images seem to suggest that VHI also occurs in VE. Consistent with prediction by Yang, Dixon et al (1999) and confirmed later by Dixon and Proffitt (2002), our experiments' results indicated that vertical distance tended to be overestimated more on a large display compared to on a small display. Similar to the findings of these researchers, vertical distance tended to be estimated more accurately on a small display compared to a large display. Similar to the results of previous findings (Henry and Furness 1993, Loomis, Da Silva et al. 1996), vertical distance yielded more accurate estimates compared to other asymmetrical distances of horizontal and transverse distances.

Overall, the results of studies in the research presented in this thesis suggest that whilst spatial awareness in VE is similar to real world counterparts in terms of distance judgment, the inaccurate distance judgment in VE should raise concerns about the utility of VE technologies in applications particularly those relying on very accurate distance judgment in the VE.

## 9.2.2 Implications of results from spatial memory task performance

The importance of accurate spatial representation perceived from interacting with VE have been emphasised by several researchers (Arthur, Hancock et al. 1997). They argued that the utility of VE in any intended applications is predicated upon the accuracy of spatial representation formed in the VE.

### 9.2.2.1 Image type

The results of Experiment 3 which examined spatial memory performance in interactive images suggested that spatial knowledge in VE was similar to that acquired in the real world. Thus, in terms of spatial memory, it is possible to perceive spatial relations between objects similarly in both environments. Whilst this result confirms the results of past studies (Arthur, Hancock et al. 1997, Richardson, Montello et al. 1999), it further suggests that the possible reason for the similar performance between real and VE conditions is practice effects. The longer viewing time and more experience in the VE compared to the real world condition may have improved the VE participants' performance to be similar to the real world participants. The implication from this is that more practice in the VE may improve participants' spatial knowledge acquisition to be similar or even better than in the real world. These results provide support for the common belief that practice improves learning (Stanney, Mourant et

al. 1998) and previous investigations that more experience increase spatial knowledge acquisition tasks (Ruddles, Payne et al. 1998). This could mean for training applications, more exposure time and more practice could improve trainees' spatial judgment skills.

Other possible explanation for the better spatial memory performance in VE condition over real condition is the impact of more mental or cognitive workload on real world participants. The details of information available in the real world compared to the VE image (Yanagasiwa and Akahori 1999) and the need to focus on physically moving in the real environment may imposed more cognitive workload on the real world participants. Thus for the real world participants less cognitive resources are available for the required tasks and this may subsequently degraded their performance. Accepting this argument may imply that working with VE models would have the advantage of having more cognitive resources devoted towards the assigned tasks. One possible application that may benefit from this is crime scene investigations. Investigators working with VE models can focus more of their cognitive resources on necessary tasks compared to when working in the actual physical environment.

The similar performance between the real and VE conditions may also suggest that the use of input device may be sufficient to provide information about movement similar to the real world. The implication of this is that the use of input devices (particularly the mouse) may be minimally sufficient to compensate for the missing sensory cues (such as vestibular cues) and provide some proprioceptive feedback necessary to indicate movement. This may also explained why previous investigations (Witmer and Kline 1998, Grant and Magee 1998) found that distance judgment performance using a treadmill (a walking interface device) was similar to when using a more traditional device, a joystick.

Whilst not significant, Experiment 3's results also indicated that spatial representation formed in VE was slightly better than those formed in the real world. Following the argument of practice effects discussed earlier, more practice would yield significant result. Though, this yet has to be further empirically proven. However, accepting these results implies that VE could be used to improve users' spatial skills especially for training applications which require spatial judgment. Providing trainees with more practice and experience in the VE could improve their spatial knowledge acquisitions. Applications that could benefit directly from these would be military training, fire fighting training and other application that requires spatial skills trainings. Furthermore, in addition to cost factor, the advantage of training in a VE is that trainees can practise in a safe VE instead of training in actual places or situations which are rare, remote or dangerous (Waller, Hunt et al. 1998a, Sinai, Krebs et al. 1999).

### 9.2.2.2 Display size

Contrary to the results of previous investigations (Patrick, Cosgrove et al. 2000, Czerwinski, Tan et al. 2002, Tan, Gergle et al. 2003) but consistent with findings of other researchers (Johnson and Stewart 1999, Arthur 2000), in terms of display size factor, results from the experiment on interactive images (Experiment 3) indicated that spatial memory resulting from a large display did not differ significantly from a small display. However, spatial memory performance was significantly different in both display sizes when the variances from the covariates (sports background and computer games experience) were removed from the data. This result suggested that both factors also contribute towards the effect of display size factor.

Surprisingly, spatial memory performance on a small display was better than on a large display. Results from both Experiment 3A and 3B suggested that the better performance of small display over large display was influenced more by FOV rather the display size factor. Moreover the interaction effects of the display size, interface device, travel modes, sport background, suggested that all these factors contribute to the better performance of the small display over the large display. The better performance on a small display than on a large display was unexpected as it contradicted theoretical considerations and the results of past studies (Patrick, Cosgrove et al. 2000, Czerwinski, Tan et al. 2002, Tan, Gergle et al. 2003). However, similar to distance estimates in interactive images (see the previous section, Section 9.1.1), the subjective responses from the display questionnaire could partially explain this results. Thus, similar implications from the effect of display size on distance estimate tasks would apply for spatial memory tasks.

### 9.2.2.3 Device types

Results from Experiment 3A and 3B in interactive VE indicated that the choice of interface device significantly ( $p < .05$ ) affected participant spatial memory performance whereby performance using a mouse was better than using a trackball. However, this result was significantly (at 10%) affected by the other factors such as display size, travel modes and participants' sport background. The use of familiar device may be more beneficial to the user as they are already used to it and do not have to relearn the skill of using this device. It is noted that the use of a mouse is limited in terms of functionality and may be beneficial for simple tasks such as free exploration of small VE space as investigated in Experiment 3. Complex interactions (such as objects manipulations) and large VE space would yield different results and may require other devices such as space-balls, data-gloves and trackers. Space-balls provide three translational and three rotational degree of freedom and are often used in CAD and robotic applications. For data-gloves, tracking sensors are used to sense

user's hand in 3D space which is used to control applications. Similar to data-gloves but instead of glove, trackers have to be held in the hand. Both data-gloves and trackers are often used in combination with HMDs.

#### 9.2.2.4 Travel modes

The effect of travel modes on spatial memory task performance was marginally better for fly mode over drive mode. Similar to interface device, this result was significantly (at 10%) influenced by the size of display, interface device, travel modes and participants' sporting background. The slightly better performance of fly mode over drive mode implies that the more natural type of movement available in the real world does not necessary yield better performance. Another implication is, in terms of travel mode, it is not necessary to replicate natural real world movement in the VE. However, training applications which require transfer of corresponding skills in the VE to the real environment may not benefit from this "unnatural movement". Other applications such as architectural design and crime scene investigations do not have such constraints. These applications may benefits from this unnatural movement in terms of performance improvements. It should be stressed that the benefits of mouse over trackball and fly mode over drive mode may apply for simple exploration task of the VE and a small scale VE as examined in the current studies (Experiment 3). Different and more complex tasks and large scale VE may yield different conclusions from these studies.

In this section, the implications of experimental results on user's spatial awareness (in terms of distance and space perception) in VE applications were considered. However, it should be noted that the results and conclusions drawn from our studies should be interpreted within the research scopes and limitations described in the earlier chapters.

### 9.3 CONCLUDING REMARKS

In this chapter, the major research findings and contributions presented in this thesis were outlined in the first section. The main results from the three experiments on static, dynamic and interactive presentations were presented in terms of the task performance measures investigated. The influence of the VE models created using the techniques described in this research on experiments' results were also presented. Next these results were highlighted in relations to the four key research questions explored. Finally, some of the thesis main contributions in terms of research methods were also highlighted. Some of the research work reported in this thesis have been presented and published (Awang-Rambli and Kalawsky 2002, Awang-Rambli and Kalawsky 2003).



Discussions of experiment results and its implication for spatial awareness (in terms of distance and space perception) were presented in the second section. The impacts of these results on users' spatial awareness in VE were considered. Direct impacts on specific VE applications were also highlighted. The implications drawn from the result of experiments examining distance estimate tasks were initially described and this is followed by the implications from the experiments investigating spatial memory tasks.

This thesis has contributed towards knowledge and understanding of the effect and influence of the investigated factors on spatial awareness in the real and VE. It has expanded on investigations by previous researchers by explaining the contribution of display size factor on participants' spatial task performance through the investigations of the effects of other related factors (such as viewing distance, physiological cues, image resolution, interface device and travel modes).

This thesis has also presented findings from the investigations of the effect of these factors on users' distance estimate and spatial memory tasks in the context of static, dynamic and interactive real and VE presentations. Whilst in this chapter, several contributions in terms of empirical results have highlighted some important findings and implications, careful interpretations of these findings should be made within the constraint of the experiments' limitations and scopes.

Based on these results and its research scopes and limitations, the final chapter (Chapter 10) will provide some recommendations and potential areas for future research work.

The results presented in this thesis will be of particular relevance to anyone wanting to apply a VE system to support training or applications where VE surrogates of real world scenarios are employed. Consequently, the research provides strong evidence to suggest transferring training or task characteristics from a VE to a real world should be undertaken with care.

# CHAPTER 10

---

## *RECOMMENDATIONS & FUTURE WORKS*

### 10 OVERVIEW

This chapter outlines several recommendations and directions for future works based on the research work conducted in this thesis. By re-examining some of the main constraints and assumptions of the research, some recommendations and areas for further research are identified and presented.

#### 10.1 RECOMMENDATIONS AND FUTURE WORKS

The research reported in this thesis expands on previous investigations and makes several important contributions to knowledge of spatial awareness perception in VE, particularly on factors affecting spatial awareness perception in VE, as reported in the previous chapter (Chapter 10). However, the results of these studies are constrained by the research scope and limitation, thus indicate more work is still required. Based on the results and assumptions in the research work, the following are some recommendations and potential avenues for further research:

- The distance estimate tasks in Experiment 1 and 2 were compared based on images of real and VE. Similar tasks could also be conducted in the real physical world. The result of the real world performance could be used as a baseline comparison for performance in the real and virtual image.
- The results of Experiment 2 suggest that image resolution may influence the participants' spatial performance. Considering the trade-off between image fidelity and computing resources, future research could extend investigations to examine the effect of varying level of image resolution on spatial tasks and determine the minimal level of image resolution necessary for VE image before spatial performances degrade.
- Experiment 2A could benefit from re-investigation using a high image resolution for VE conditions based on the same experimental setup that is fixing the FOV and varying the viewing distance. The results could provide a more direct comparison with Experiment 2B as initially planned.
- Objects' located on the far left and far right of the image or screen would be located in the viewer's peripheral visions, which the eyes viewed with low acuity compared to the centrally located objects which were viewed with high acuity. Thus, for the distance estimate tasks, it would be of interest to investigate the influence of objects' positions on distance estimate performance. Furthermore, for Experiment 1 and 2, different objects were used for the distance estimates. Due to familiarity factors, different objects may have different effects on distance estimates. The effect of the type of objects could be investigated by using the same objects at different positions or different objects at the same positions. A difference would suggest that the type of object is another factor which influence distance estimates. It is expected that more familiar objects would yield more accurate estimates.
- The use of a simple and single room environment was suggested as one possible reason for similar perception between real and VE conditions in Experiment 3. Future investigation could employ a more complex environment which consists of several rooms or a building. Additionally, the VE model in Experiment 3 was "uncluttered". The effect of a cluttered environment on spatial performance would be another factor that may influence performance due to its potential impact on users' navigation. Several researchers have commenced investigations on the influence of different movement interfaces, different levels of cluttered environment, collision response algorithms and FOV on search tasks (Ruddles and Jones 2001) but not on spatial performance.

- Replicating Experiment 3 with some adjustment to the viewing time where the number of viewing time in the real conditions is matched to those of VE conditions will help clarify the contribution of practice effects in these results.
- Experiment 3's result indicates that a large display does not necessary have the same impact on all viewers. Some viewers may perform better on small displays while others may perform better on large displays. The subjective comments imply that the reason for this appears to be users' familiarity with small displays. However, other factors may also be involved. Thus, future investigations could examine the possible effects of related factors such as work experiences (such as computer-related jobs verses non-computer-related jobs) and gender factors.
- The VE image in Experiment 3 suffers from anti-aliasing effects especially when presented on a large display. One possible solution would be to use the MIPS technique (see Section 3.1.2.2 of Chapter 3) which was based on LOD techniques. Instead of using a set of objects, the MIPS technique uses a set of texture maps of varying resolution corresponding to the set of distances of the objects from the viewer.
- As suggested by an anonymous reviewer, the sporting background may influence participants' performance. In Experiment 1 and 2, we tried but failed to recruit volunteers from professional sportsmen. The results of Experiment 3 suggest user's sports background does contribute some influence on spatial memory task performance. Empirical investigations could be further conducted to examine the contribution of sport background (the types of sport played) on spatial performance.
- The effect of viewing distance on the sense of presence when viewing VE was not investigated. Some psychological studies have indicated viewing distance has some influence on TV viewers' presence (Oyama and Shiramatsu 2002). As it was recommended to promote a user's sense of presence to improve performance benefits in VE (Stanney, Kingdon et al. 2002), examining the effect of viewing distance on sense of presence when viewing VE images would present another useful avenue for further investigations.
- The impact of different levels of realism on a user's spatial awareness was not investigated in this thesis. Considering the need to make trade-offs between image realism and computing resources for real-time VE, it would be of interest to compare the effect of using different techniques to improve or create a different level of realism on spatial awareness.

- The effect of navigation on spatial judgment was not directly investigated in this research. However, a comparison of results from static and dynamic images to interactive environments seems to suggest that navigation may have an impact on spatial judgment. Thus, further studies were required to support this assertion.
- The input device investigated in this thesis was mouse and trackball. It would be useful to extend the investigation to include other devices such as a walking interface (such as treadmill), joystick, handheld trackball and gloves. Similarly, the travel modes compared were limited to driving and flying. This could be expanded to other forms of travel modes such as teleportation or other movements supported by vehicles modelled in the VE.
- The spatial memory test was based on the paper and pencil method. Another useful test method would be to present the VE model with all objects to be located placed at one corner of the screen and then to ask participants to place objects at their correct locations. Software could be developed to record these results immediately for each participant and this would certainly save the time taken in analysing the map test results manually. An alternative method to analysing the map test data would be to collect information based on offset errors in the x and y (and possibly z) directions. These results could then be compared with the method used in this thesis.
- A spatial ability test was not conducted on the participants due to questions of its relevance and usefulness. While randomization of the participants helps reduce participants' variance, this test could have been used to determine if the participants were similar in terms of spatial ability, thus further reducing variance among participants. A spatial ability test could also be used to determine if the differences in performance between factors could be due to or influenced by spatial ability differences.
- The use of a questionnaire in experimental studies is important to elicit certain information which may not be obtainable by other methods. For example in Experiment 3, the display questionnaire provided further information on understanding the better performance of small display over large display. However, questionnaires usually provide information based on the set of criteria dictated by the researcher. Other criteria not identified or overlooked by the researcher may be of importance. Potential future research could include expanding on the list of criteria and conducting objective studies. The results of display questionnaire indicated more work is needed. The display questionnaire in this thesis could be formally developed, structured and verified to serve a guideline for designers.



- Before the VE models were created using modelling software, information about the scales and measurements of the objects and the location itself needed to be gathered. The techniques (measuring tapes, rulers and photographs) employed in this research to gather such information were not only tedious but also time consuming and error prone especially for models based on an outdoor setting. In fact considerable amounts of time were spent collecting such information in order to produce the accurate models used in the experiments in this thesis. Various other software and techniques (such as photogrammetry and laser technology) are available in the market to enable more quick, efficient and accurate measurement and modelling of 3D objects and locations. Photogrammetry is a technique of measuring objects (2D or 3D) from photographs or digital images. PhotoModeler is one example of a software that take measurements and models 3D from photographs (more information can be obtained from [www.photomodeler.com](http://www.photomodeler.com)). The 3D laser scanning technology allows cost savings and avoids labour intensive methods of collecting dimensions data with tape measurements and it also provides a safe way to collect the geometric dimensions which are unsafe and difficult to reach (Thigayagarajan 2003).
- The approach in this thesis was purposely limited to non-stereo presentations. As reviewed in Chapter 2, 3 and 4, it was argued that stereo presentation provides the user with more sense of immersion and realism. Additionally, it is beneficial for some spatial tasks especially for near distance, but the negative attributes (such as complex and costly hardware and viewers' related issues) that come with stereo presentations have dissuaded some decisions for its use. However, the use of auto-stereo displays could help overcome some of the viewers' related problems. This is because auto-stereo systems do not require the viewer to wear special eyewear such as shutter glass or other head gear for stereo presentations (Dodgson 1997). Although most currently available auto-stereo displays are relatively small, recent developments have seen some larger displays such as Autostereo 3D display wall by QinetiQ (Moseley 2004). This may present another potential avenue for future research on spatial awareness perception which includes stereo cues.
- In this thesis, the main focus was on visual modality. As discussed in Chapter 4, the absence of other modalities such as audio, kinaesthetic, proprioceptive and haptic may influenced user's performance in the VE. Thus, for future research the influence of the presence or absence these modalities on spatial awareness may be another area for investigation
- In this thesis, investigation is limited to the first person view, that is, a simulation of what the user see if he is in the VE. The other type (third person view) includes a

representation of the user (called avatars) in the VE. In a multi-users VE, multiple users (or avatars) are present simultaneously in the VE. One potential application of multi-users VE is in video conferencing. Another area for further investigation which is examining the effect of perceived distance (or other tasks such as spatial orientation) on communication between avatars, that is, how it affect factors such as communication constructs, conversational appropriateness and social interactions.

- Another interesting area to consider for future research is subconscious perception. Conscious perception is when we know what we see (or hear, taste, smell, feel) which can be accounted for. However, subconscious means below the level of consciousness. Subconscious perception means the perceiver is not aware of what he see (or other senses) and it is observable through a change in behaviour as a respond subconscious stimuli. In addition to the investigation of the conscious stimuli, future work of spatial perception should include the examination of the influence of subconscious perception on users' navigation or actions in the VE.
- Since the results of the experiments suggest that display size does play an important role in influencing perception of space and distance, it would be useful to determine the optimal display size that would yield accurate perception and increase the users' sense of presence. The latter is necessary as some researchers suggested that increase of sense of presence in user would lead to better their performance (Stanney, Kingdon et al. 2002).
- In this research, only two dimensions of the display size and FOV were examined, for future work more values of the variables could be used to know whether the relations between variables are monotonic or if an optima actually exist
- Finally, analysis of the data in this research is based on the hypotheses proposed. Even though, analysis of the data yield some interesting results as discussed in the respective chapters, because of the amount of information in the experimental data further analysis could again be conducted. To get more out of the existing data, future work could include for example examining for the longitudinal effects and for differences in results close and far transverse.

## 10.2 CONCLUDING REMARKS

The aim of this thesis was to examine spatial awareness perception in the real and VE. The research presented described investigations into factors affecting spatial awareness in terms of distance estimate and spatial memory tasks in the real and VE in the context of static, dynamic and interactive presentation.

Based on the experiments' results, scopes and limitation, several recommendations and potential areas for further work were proposed. Some of the recommendations include proposition for new methods and improvement of the methods employed by the research in this thesis. These recommendations would provide further clarification, enhancement and support for some of the findings from the research presented in this thesis. Several potential areas for further research work were also highlighted. These focus primarily on suggestions for the investigations of other aspects and factors that are related and could affect a user's spatial awareness in VE.

## REFERENCES

- Adelson, E.H. (1993). "Perceptual organization and the judgment of brightness". Science 262: 2042-2044
- Alfano, P. L. and G. F. Michel (1990). "Restricting the field of view: Perceptual and performance effects." Perceptual and Motor Skills 70: 35-45.
- Allen, R. C. and D. P. McDonald (1997). Landmark Direction and Distance Estimation in Large Scale Virtual Environments. Human Factors and Ergonomics Society 41st Annual Meeting - 1997.
- Allison, R. S., L. R. Harris, et al. (2002). "Simulating self-motion II: A virtual reality tricycle." Virtual Reality 6: 86-95.
- Arthur, E. J., P. A. Hancock, et al. (1997). "The Perception Of Spatial Layout In Real And Virtual Worlds." Ergonomics 40(1): 69-77.
- Arthur, K. (2000). Effects of Field of View on Performance with Head-Mounted Displays, PhD Thesis, University of North Carolina at Chapel Hill.  
Available at [www4.tomshardware.com/column/20000110/index.html](http://www4.tomshardware.com/column/20000110/index.html).
- Arthur, K., K.S. Booth, et al. (1993). Evaluating 3D task performance for fish tank virtual worlds. ACM Transactions on Information Systems 11(3):239-265.
- Awang Rambli, D. R. and R. S. Kalawsky (2003). Effects of varying display size on user's asymmetrical distance perception in the real and virtual environment. Virtual Concept 2003, Biarritz-France.
- Awang Rambli, D. R. and R. S. Kalawsky. (2002). The Effect of Display and Image Type on Inter-Object Distance Estimation in Virtual and Real Environment. SIGCHI-NZ Symposium on Computer-Human Interaction, Hamilton, New Zealand.
- Baird, J. W. (1903). "The influence of accommodation and convergence upon the perception of depth." American Journal of Psychology 14: 150-200.
- Bakker N.H., Werkhoven P.J., et. al. (1999). "The effects of proprioceptive and visual feedback on geographical orientation in virtual environments." Presence. 8(1);36-53.
- Bakker N.H., Werkhoven P.J., & Passenier P.O. (1997). Orientation in virtual environments. Proceedings of the 16th European Annual Conference on Human Decision Making and Manual Control, Kassel, Germany, pp.208-209.
- Bakker N.H., Werkhoven P.J., et.al. (2001). "Calibrating Visual Path Integration in VE." Presence 10 (2): 216-224.
- Bakker, N.H., Passenier, P.O, et.al.. (2003). "The Effects of Head-Slaved Navigation and the use of Teleports on Spatial Orientation in Virtual Environments (VE)." Human Factors, 45(1): 160-169.
- Baldis, S. F. (1997). Interface Device for 3-D Interaction. Master's Thesis. Electrical Engineering, University of Washington, Washington, USA.
- Barbour, C. G. and G. W. Meyer (1992). "Visual cues and pictorial limitations for computer-generated photo-realistic images." The Visual Computer 9(3): 151-165.

- Barfield, W. and C. Hendrix (1995). "The effect of update rate on sense of presence within virtual environment." Virtual Reality: Research, Development and Application 1(1): 3-16.
- Barfield, W. and T. A. Furness (1995). Introduction to Virtual Environments and Advanced Interface Design. In Virtual Environments and Advanced Interface Design. W. Barfield and T. A. Furness (Eds.). New York, Oxford University Press: 3-13.
- Bass, E. J., J. P. Zeng, et al. (1996). A context-based approach to training situational awareness. HICS '96 -Third Annual Symposium on Human Interaction with Complex Systems, Los Alamitos, CA, USA, IEEE Computer Society.
- Bastos, R., K. Hoff, et al. (1999). Increased photorealism for interactive architectural walkthrough. Symposium on Interactive 3D Graphics, Atlanta, Georgia USA.
- Beall, A. C., J. M. Loomis, et al. (1995). Absolute motion parallax weakly determines visual scale in real and virtual environments. SPIE Proceedings on Human Vision, Visual Processing, and Digital Display VI.
- Bengston, J. K., J. C. Stergios, et al. (1980). "Optic Array Determinants of Apparent Distance and Size in Pictures." Journal of Experimental Psychology: Human Perception and Performance 6(4): 751-759.
- Bethel, W. (1999). RM Scene Graph Technical white paper. Novato, CA, USA, R3vis Corporation.
- Bideau, B., R. Kulpa, et al. (2003). "Real handball goalkeeper vs. virtual handball thrower." Presence: Teleoperators and Virtual Environments 12(4): 411-421.
- Bigham, D. D. (2000). Distance Perception and Visualization Using Virtual Environment. Master's Thesis, Naval Postgraduate School, Monterey, California, USA.
- Billinghurst, M. and S. Weghorst (1995). The use of sketch maps to measure cognitive maps of virtual environments. Virtual Reality Annual International Symposium '95, North Carolina, USA.
- Bingham, G. P., A. Bradley, et al. (2001). "Accommodation, Occlusion, and Disparity Matching are used to guide reaching: A comparison of actual versus virtual environment." Journal of General Psychology: Human perception and Performance 27(6): 1314-1334.
- Bloomer, C. M. (1990). Principles of Visual Perception. London, The Herbert Press.
- Bolas, M. T. (1994). "Human Factors in the Design of an Immersive Display." IEEE Computer Graphics and Applications 14(1): 55-57.
- Boring, E. G. (1964). "Size constancy in pictures." American Journal of Psychology 77: 494-499.
- Bowman, D. (2002). Principles for the design of performance-oriented interaction techniques. Handbook of Virtual Environments: Design, Implementation and Applications. Stanney K. M.(Ed.) London, Lawrence Erlbaum Associates: 277-300.
- Bowman, D., A. Datey, et al. (2002). Empirical Comparison of Human Behavior and Performance with Different Display Devices for Virtual Environments. Human Factors and Ergonomics Society Annual Meeting, 2002.



- Bowman, D., D. Koller, et al. (1997). Travel in Immersive Virtual Environments: An Evaluation of Viewpoint Motion Control Techniques. Virtual Reality Annual International Symposium (VRAIS)-1997.
- Brooks, B. M., E. A. Attree, et al. (1999). "The specificity of memory enhancement during interaction with a virtual environment." Memory 9(7): 65-78.
- Brooks, F. P. Jr. (1999). "What's real about virtual reality?" IEEE Computer Graphics and Applications Nov/Dec 1999: 16-27.
- Bruce, V., P. R. Green, et al. (1997). Visual Perception: Physiology, Psychology and Ecology, Psychology Press Ltd.
- Caird, J. K. and P. A. Hancock (1991). Perceived Spatial Layout of a Simulated Scene as a Function of Experience. Human Factors Society 35th Annual Meeting-1991.
- Carr, K. T. and R. D. England (1993). The Role of Realism in Virtual World. Virtual Reality International '93. Third Annual Conference on Virtual Reality, London, Meckler: Westport & London.
- Catling, S. (2000). Presentation notes by GeoVisions Project Working Group at the LSE on July 8th 2000 by Simon Catling, Deputy Head, Institute of Education, Oxford Brookes University.
- Chance, S.S. Gaunet F., et al. 1998. "Locomotion mode affects the updating of objects during travel: The contribution of vestibular and proprioceptive inputs to path integration." Presence-Teleoperators and Virtual Environment 7(2):168-178
- Christou, C. G. and H. H. Bulthoff (1999). The Perception Of Spatial Layout in a Virtual World. Technical Report TR 75, Max-Planck –Institut, Tubingen, Germany.
- Cobb, S., S. Nichols, et al. (in press). "Virtual Reality Induced Symptoms and Effects (VRISE)." To appear in Presence: Teleoperators and Virtual Environments.
- Cohen, O., S. Meyer, et al. (1993). Studying the movement of high tech rodentia: pointing and dragging. INTERACT '93 and CHI '93 conference companion on Human factors in computing systems, Amsterdam, The Netherlands.
- Coolican, H. (1999). Research Methods and Statistics in Psychology, Hodder & Stoughton.
- Coren, S., L. M. Ward, et al. (1999). Sensation and Perception, Harcourt Brace College Publishers.
- Costello, P. J. (1997). "Health and safety Issues associated with Virtual Reality - A review of Current Literature", AGOCG 37 Report Series, Available at <http://www.agocg.ac.uk/virtual.html>.
- Costello, P. J. and S. Bee (1997). "3D Model Databases: availability of 3D models on the WWW", AGOCG 36 Report Series. Available at <http://www.agocg.ac.uk/virtual.htm>.
- Cremer-Regehr, S. H., P. Willemson, et al. (2003). The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual environments. Technical Report UUCS-03-016, University of Utah, USA.
- Cruz-Neira, C., D. Sandin, J, et al. (1993). Surround-screen projection-based virtual reality: the design and implementation of the CAVE. International Conference on Computer Graphics and Interactive Techniques, New York, USA, ACM Press.

- Crvarich, G. (1995). An Exploration Of Techniques To Improve Relative Distance Judgments Within An Exocentric Display, Master Thesis, University of Washington. USA.
- Cutmore, T., H. R Hine, et al. (2000). "Cognitive And Gender Factors Influencing Navigation In A Virtual Environment." International Journal Human-Computer Studies 53: 223-249.
- Cutting, J. E. (1987). "Rigidity in Cinema Seen From the Front Row, Side Aisle." Journal of Experimental Psychology: Human Perception and Performance 13(3): 323-334.
- Cutting, J. E. (1997). "How the Eye Measures Reality and Virtual Reality." Behavior Research Methods, Instrumentation, and Computers 29: 29-36.
- Cutting, J. E. (in press). Reconceiving perceptual space. In Perceiving pictures: An interdisciplinary approach to pictorial space. H. Hecht, M. Atherton and R. Schwartz (Eds.), MIT press.
- Cutting, J.E. (1986). Perception with an eye for motion. MIT Press, Cambridge, Mass.
- Cutting, J. E. and P. M. Vishton (1995). Perceiving layout and knowing distances: The integration, relative potency, and context use of different information about depth. Handbook of perception and cognition: Perception of space and motion. W. Epstein and S. Rogers (Eds.). San Diego, CA, Academic Press. 5: 69-117.
- Czerwinski, M., D. S. Tan, et al. (2002). Women Take a Wider View. CHI 2002 Conference on Human Factors in Computing Systems, Minneapolis, Minnesota, USA.
- Deering, M. (1993). Data complexity for virtual reality: Where do all the triangles go? VRAIS 1993.
- Dees, J. W. (1966). "Accuracy of absolute visual distance and size estimation in space as a function of stereopsis and motion parallax." Journal of Experimental Psychology 72: 466-476.
- Dixon, M. W. and D. R. Proffitt (2002). "Overestimation of heights in virtual reality is influenced more by perceived distal size than by the 2-D versus 3-D dimensionality of the display." Perception 31: 103-112.
- Dodgson, N. A. (1997). Autostereo displays: 3D without glasses. EID 97(Electronic Information Displays), Esher, Surrey.
- Dolezal, H. (1982). Living in a world transformed: perceptual and performatory adaptation to visual distortion. New York, Academic Press.
- Draper, M. (1995). Exploring the Influence of a Virtual Body on Spatial Awareness. Master's Thesis, University of Washington. USA.
- Draper, M. (1996). Can Your Eyes Make You Sick? Investigating the Relationship Between the Vestibulo-Ocular Reflex and Virtual Reality. Technical Report R-96-3, Human Interface Technology Laboratory, University of Washington, Seattle, USA.
- Duh, B. L., J. W. Lin, et al. (2002). "Effects of characteristics of image quality on balance in an immersive environment." Presence: Teleoperators and Virtual Environments 11(3): 324-332.
- Durlach, M. and A. Mavor (1995). Virtual Reality: Scientific and Technical Challenges. Washington D.C., National Academy Press.

- Egglestons, R. G., W. P. Janson, et al. (1996). Virtual Reality System Effects On Size-Distance Judgments In A Virtual Environment. IEEE Annual Virtual reality International Symposium, Santa Clara, California.
- Ellis, S. R. (1994). "What are Virtual Environments?" IEEE Computer Graphics & Applications 14(1): 17-22.
- Endsley, M. R. (2000). Theoretical underpinning of situational awareness: A critical review. Situational Analysis and Measurements. M. R. Endsley and D. J. Garland. Mahwah, NJ, Lawrence Erlbaum Associates.
- Ferris, S. H. (1972). "Motion parallax and absolute distance." Journal of experimental Psychology 95(2): 258-263.
- Festinger, L. (1957). A Theory of Cognitive Dissonance. Stanford, CA: Stanford University Press.
- Field, A. (2000). Discovering Statistics using SPSS for windows, Sage Publication.
- Foley, J. D., S. K. van Dam, et al. (1995). Computer Graphics: Principles and Practice. Boston, MA, Addison-Wesley
- Garau, M. (2003). The Impact of Avatar Fidelity on Social Interaction in Virtual Environment, PhD Thesis, University College London, UK.
- Göbel, M. and W. Friesdorf (2002). Evaluation of input devices for 3D-navigation in medical application. 6th International conference Work with Display Unit (WWDU) 2002, Berchtesgaden-Germany.
- Goerger, S. R., R. P. Darken, et al. (1998). Spatial Knowledge Acquisition From Maps And Virtual Environments In Complex Architectural Spaces. 16th Applied Behavioral Sciences Symposium, U.S.Air Force Academy, Colorado Springs, CO,USA.
- Goldstein, E. B. (1991). Perceived orientation, spatial layout and the geometry of pictures. Pictorial Communication In Virtual and Real Environment. S. R. Ellis, M. K. Kaiser and A. J. Grunwald, (Eds.), Taylor & Francis Ltd.
- Goldstein, E. B. (1996). Sensation and Perception. USA, Brooks/Cole Publishing Company, ITP.
- Golledge, R. G. (1991). Cognition of Physical and Built Environments. In Environment, Cognition, and Action : An Integrated Approach. G. Garling and G. W. Evans. NY (Eds.), Oxford Press: 35-62.
- Gooding, L., G. E. Miller, et al. (1991). "The effect of viewing distance and disparity on perceived depth." Stereoscopic Displays and Applications II, SPIE Proceedings 1457: 259-266.
- Grant, S. C. and L. E. Magee (1998). "Contributions of proprioception to navigation in virtual environments." Human Factors 40(3): 489-497.
- Green, M. and H. Sun (1995). Computer Graphics Modeling for Virtual Environments. In Virtual Environments and Advanced Interface Design. Barfield, W and Furness T.A (Eds.), 63-101, Oxford University Press, New York.
- Greenberg, D. P. (1999). "A framework for realistic image synthesis,." Communications of ACM 42(8): 45-53.
- Greene, J. and C. Hicks (1984). Basic cognitive processes, Open University Press.
- Groen J. & Werkhoven P. (1998). "Visuo-motor adaptation to virtual hand position in interactive virtual environments." Presence, 7(5): 429-446.

- Hagen, M., R. J. Jones, et al. (1978). "On a neglected variable in theories of pictorial perception: Truncation of the visual field." Perception & Psychophysics 23(4): 3326-330.
- Harris, L. R., M. R. Jenkin, et al. (2002). "Simulating self-motion I: Cues for the perception of motion." Virtual Reality 6: 75-85.
- Harris, L., M. Jenkin, et al. (1999). Vestibular cues and virtual environments: choosing the magnitude of the vestibular cue. IEEE Virtual Reality conference, Houston, Texas.
- Harwood, K. and P. Foley (1987). "Temporal resolution: An insight into the Video Display Terminal (VDT) Problem". Human Factors 29(4): 447-452.
- Hecht, H., A. v. Doorn, et al. (1999). "Compression of Visual Space in Natural Scenes and in their Photographic counterparts." Perception & Psychophysics 61(7): 1269-1286.
- Helman, J. L. (1993). "Designing Virtual Reality Systems to meet Physio and Psychological Requirements", SIGGRAPH Course Number 23: Applied Virtual Reality.
- Hendrix, C. and W. Barfield (1995). "Relationship between monocular and binocular depth cues for judgments of spatial information and spatial instrument design." Display 16(3): 103-113.
- Henry, D. (1992). Spatial Perception In Virtual Environments: Evaluating An Architectural Application, Master's Thesis. University of Washington, USA.
- Henry, D. and T. Furness (1993). Spatial Perception In Virtual Environments: Evaluating An Architectural Application. IEEE Virtual Reality Annual International Symposium, Seattle, Washington.
- Hettinger, L. J. (2002). Illusory self-motion in Virtual Environments. In Handbook of Virtual Environments: Design, Implementation, and Applications. Stanney, K. M.(Ed.), Lawrence Erlbaum Associates: 471-491.
- Higashiyama, A. (1996). "Horizontal and Vertical distance perception: The discorded-orientation theory." Perception & Psychophysics 58(2): 259-270.
- Higashiyama, A. and E. Ueyama (1988). "The perception of vertical and horizontal distances in outdoor setting." Perception & Psychophysics 44(2): 151-156.
- Howard, I. P. and B. J. Rogers (2002). Seeing in depth, Vol.2, I. Porteous Publishers.
- Howard, T. L. J., A. D. Murta, et al. (2000). "Virtual environment for scene of crime reconstruction and analysis." Proceedings of SPIE/IS&T 3960: 41-48.
- Howarth, P. A. and P. J. Costello (1996). Studies into the effects of Immersion n Virtual Environments. VISERG Report 9603, Department of Human Science, Loughborough University, UK.
- Hsu, J., Z. Pizlo, et al. (1994). "Design of studies to test the effectiveness of stereo imaging truth or dare: Is stereo viewing really better?" Proceedings of SPIE -Stereoscopic Displays and Virtual Reality Systems 2177: 211-222.
- Hu, H. H., A. A. Gooch, et al. (2002). "Visual cues for perceiving distances from objects to surfaces." Presence : Teleoperators and Virtual Environments 11(6): 652-664.
- Hubona, G. S., P. N. Wheeler, et al. (1999). "The role of objects shadows in promoting 3D visualization." ACM Transactions on Computer-Human Interaction 6(3): 214-242.

- Iavecchia, J. H., H. P. Iavecchia, et al. (1988). "Eye accommodation to head-up virtual images." Human Factors 30(6): 689-702.
- Isdale, J. (1998). What is Virtual Reality? Available at <http://www.cse.dmu.ac.uk/~cph/VR/whatisvr.html>. 5 Oct 1998.
- Jää-Aro, K.-M. and L. Kjelldahl (1997). "Effects Of Image Resolution On Depth Perception In Stereo And Non-Stereo Images." Stereoscopic Displays and Applications VIII, SPIE Proceedings 3012: 319-326.
- Jacob, R. J. K. and L. E. Sibert (1992). The preceptual structure of multidimensional input device selection. CHI 1992.
- James, K. R. and J. K. Caird (1995). The effects of optic flow, proprioception, and texture on novice locomotion in virtual environments. Human Factors and Ergonomics Society 39th Annual Meeting - 1995.
- Johnson, D. M. and J. E. Stewart II (1999). "Use of Virtual Environments for the Acquisition of Spatial Knowledge: Comparison Among Visual Displays." Military Psychology 11(20): 129-148.
- Kalawsky, R. S. (1993). The Science of Virtual Reality and Virtual Environment, Addison Wesley.
- Kalawsky, R. S. (1996). "Exploiting Virtual Reality Techniques in Education and Training: Technological Issues," SIMA Report Series.
- Kalawsky, R. S. (1998). "The Impact of Spatial Immersive VR on Smart Schools and Colleges". Keynote Address Virtual Reality in Education 1998 (VRET 1998), July, City University, London, UK.
- Kalawsky, R. S. (1999). "VRUSE - a computerised diagnostic tool: for usability evaluation of virtual/synthetic environment systems." Applied Ergonomics 30: 11-25.
- Kalawsky, R. S. (2000a). The Validity Of Presence A Reliable Human Performance Metric In Immersive Environment. Presence 2000, 3rd International Workshop on Presence, Delft, Netherlands.
- Kalawsky, R. S. (2000b). JTAP Project 3-5 Human Factors Aspects of Virtual Environments Design in Education: Project Report. Available at <http://sgi-hursk.lboro.ac.uk/~avrrc/JTAP305FinalReport.pdf>.
- Kasik, D. J. (2002). Graphic in the Large: Is Bigger Better? Discussion Panel. Siggraph 2002, 21-26 July. L. Carpenter, B. Fisher, R. A. May and N. Streitz. San Antonio. TX., USA.
- Kemeny, A. (1993). A cooperative driving simulator. International training equipment conference and exhibition, London, UK.
- Kennedy, R. S., D. S. Lanham, et al. (1995). "A Comparison of Cybersickness Incidences, Symptom Profiles, Measurement Techniques, and Suggestions for Further Research." Presence: Teleoperators and Virtual Environments 6(6): 638-644.
- Kessler, G. D. (2002). Virtual Environment Models. Handbook of Virtual Environments: Design, Implementation, and Applications. Stanney, K. M. (Ed.), Lawrence Erlbaum Associates: 255-276.



- Kjeldskov, J. (2001). Interaction: Full and partial Immersive virtual reality displays. IRIS24, Bergen, Norway.
- Kjelldahl, L. and M. Prime (1995). "A Study on how depth perception is affected by different presentation methods of 3D objects on a 2D display." Computers & Graphics 19(2): 199-202.
- Kline, P. B. and B. G. Witmer (1996). Distance perception in virtual environments: effects of field of view and surface texture at near distances. Human Factors and Ergonomics Society 40th Annual Meeting -1996.
- Knapp, J. M. (1999). The Visual Perception of Egocentric Distance in Virtual Environments. Doctoral Theses Dissertation. Psychology Department. Santa Barbara, California, USA., University of California.
- Knapp, J. M. and J. M. Loomis (in press). "Limited field of view of head-mounted displays is not the cause of distance estimation in virtual environments." To appear in Presence: Teleoperators and Virtual Environments October 2004.
- Kraft, R. N. and J. F. Patterson (1986). "Distance perception of photographic displays of natural settings." Perceptual and Motor Skills 62: 179-186.
- Kunnapas, T. (1968). "Distance perception as a function of available visual cues." Journal of Experimental Psychology 77(4): 523-529.
- Lampton, D. R., B. W. Knerr, et al. (1994). "The Virtual Environment Performance Assessment Battery (VEPAB): Development and Evaluation." Presence 3(2): 145-157.
- Lampton, D. R., D. McDonald, et al. (1995). Distance estimation in virtual environment. Human Factors and Ergonomics Society 39th Annual Meeting.
- Lampton, D. R., J. P. Bliss, et al. (1994). Object Recognition, Size estimation, and distance estimation in real-world and virtual environments. Human Factors and Ergonomics Society 38th Annual Meeting -1994.
- Lantz, E. (1996). The Future of Virtual Reality: Head Mounted Displays versus Spatially Immersive Displays. SIGGRAPH 23rd Annual Conference on Computer Graphics and Interactive Techniques, ACM Press.
- Lastra, A. (1995). Technology for Virtual Reality. SIGGRAPH 1995 course notes.
- LaViola, J. J. (2000). "A discussion of cybersickness in virtual environments." SIGCHI Bulletin 32(1): 47-56.
- Lo, C. H. and A. Chalmers (2003). Stereo vision for computer graphics: the effect that stereo vision has on human judgments of visual realism. 19th Spring Conference on Computer Graphics, Budmerice, Slovakia.
- Longhurst, P., P. Ledda, et al. (2003). Psychophysically based artistic techniques for increased perceived realism of virtual environments. 2nd International conference on Computer graphics, Virtual Reality, Visualization and Interaction, Cape Town, South Africa.
- Loomis, J. M. and J. M. Knapp (2003). Visual perception of egocentric distance in real and virtual environments. Visual and Adaptive Environments. L. J. Hettinger and M. W. Haas (Eds.). Hillsdale NJ, Erlbaum.
- Loomis, J. M., J. A. Da Silva, et al. (1996). "Visual perception of location and distance." Current Directions in Psychological Science 5(3): 72-77.

- Lumsden, E. A. (1980). Problem of magnification and minification: An explanation of the distortion of distance, slant, shape, and velocity. The perception of Pictures. M. E. Hagen (Ed.). New Jersey, Academic Press: 91-135.
- MacKenzie, I. S., A. Sellen, et al. (1991). A comparison of input devices on elemental pointing and dragging tasks. CHI 91 Conference on Human Factors in Computing Systems.
- Malhorta, P. (2002). Issues involved in real-time rendering of virtual environments. Master's Thesis. College of Architecture and Urban Studies. Blacksburg, Virginia, Virginia Polytechnic Institute and State University. USA.
- Mania, A. (2001). Fidelity Metrics for Virtual Environment Simulations based on Human Judgments of Spatial Memory Awareness States. PhD Thesis. Department of Computer Science, University of Bristol, UK.
- Marzuryk, T. and M. Gervautz (1996). Virtual Reality: History, applications, technology and future. Technical Report: TR-186-2-96-06, Institute of Computer Graphics, Vienna University of Technology Austria.
- May, J. G. and D. R. Badcock (2002). Vision and Virtual Environment. Handbook of Virtual Environments: Design, Implementation, and Applications. Stanney K. M.(Ed.), Lawrence Erlbaum Associates.: 29-63.
- McCandless, J. W., S. R. Ellis, et al. (1999). The effect of accommodative demand, motion parallax and age on virtual object localizations. Human Factors and Ergonomics Society 43rd Annual Meeting, Houston, Texas.
- McNamara, A., A. Chalmers, et al. (2000). Comparing real and synthetic scenes using human judgments of lightness. Eurographics Workshop on Rendering Techniques 2000.
- McNamara, T. P. (1986). "Mental representations of spatial relations." Cognitive Psychology 18(87-121).
- Messing, R. (2004). Distance perception and cues to distance in virtual reality. First Symposium on Applied perception in Graphics and visualization- August 1994., Los Angeles, CA, USA.
- Meyer, G. W., H. E. Rushmeier, et al. (1986). "An experimental evaluation of computer graphics imagery." ACM Transactions on Graphics 5(1): 30-50.
- Miller, G. A. (1956). "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for processing information." The Psychological Review 63: 81-97.
- Mon-Williams, M. and J. P. Wann (1998). "Binocular virtual reality displays: when problems do and don't occur." Human Factors 40(March 01 1998).
- Morar, S. S. (2002). The perception of depth in desktop virtual environments. PhD Thesis abstract, Brunel University, UK.
- Morley, S. (2002). "Virtual reality crime scene." Forensic Science 184: 20-21.
- Moseley, C. (2004). QinetiQ demonstrates autostereo 3D display. Electronic newsroom. Available at [www.QinetiQ.com](http://www.QinetiQ.com).
- Mueller, J. K., J. P. Bliss, et al. Comparison of mouse, trackball and unmouse in a compensatory tracking task. Unpublished report.

- O'Connor, T. (2004). An Introduction to crime scene analysis. Lecture Notes. Criminal Investigation course. Available at <http://faculty.ncwc.edu/toconnor/315/315lect04.htm>
- Oyama, H. and N. Shiramatsu (2002). "Smaller and bigger displays." Displays 23: 31-39.
- Pallant, J. (2001). SPSS Survival Manual: A step by step guide to data analysis using SPSS, Open University Press.
- Patrick, E., D. Cosgrove, et al. (2000). Using a Large Projection Screen as an Alternative to Head-Mounted Displays for Virtual Environments. CHI 2000.
- Patterson, Robert. (1992). "Human stereopsis." Human Factors 34(2) 669-692.
- Persiani, F. and A. Liverani (2000). Semi-immersive synthetic environment for cooperative air traffic control. ICAS 2000, 22nd International Congress of Aeronautical Sciences, Harrogate, UK.
- Peruch, P., J. L. Vercher, et al. (1995). "Acquisition Of Spatial Knowledge Through Visual Exploration Of Simulated Environments." Ecological Psychology 7(1): 1-20.
- Peterson, L. R. and M. J. Peterson (1959). "Short term retention of individual verbal items." Journal of Experimental Psychology 58(3): 193-198.
- Pfautz, J. D. (2002). Depth perception in computer graphics, PhD Thesis, University of Cambridge, UK.
- Plumert, J. M., J. K. Kearney, et al. (2004). Distance perception in Real and Virtual Environment. First Symposium on Applied perception in graphics and visualization (APGV 04), August 2004, Los Angeles, CA, USA.
- Postma, A., R. Izendoorn, et al. (1998). "Sex Differences in Object Location Memory." Brain and Cognition 36: 334-335.
- Predebon, J. (1994). "Convergence responses to monocularly viewed objects: implications for distance perception." Perception 23: 303-319.
- Prothero, J. and H. Hoffman (1995). Widening the field of view increases the sense of presence in the immersive Virtual Environment. Technical Report R-95-5. Seattle, Washington, University of Washington: Human Interface Technology Laboratory.
- Radvansky, G. A., L. A. Carlson-Radvansky, et al. (1995). "Uncertainty in estimating distances from memory." Memory & Cognition 23(5): 596-606.
- Rahmat, O. (2000). Primer: Scene graphs explained. Available at [www4.tomshardware.com/column/20000110/index.html](http://www4.tomshardware.com/column/20000110/index.html)
- Reddy, M. (1994). Reducing Lags in Virtual Reality Systems using Motion Sensitive Level of Detail. Second UK VR\_SIG Conference, Theale, Reading, UK
- Redlick, F. P., H. L. Jenkin, et al. (2001). "Humans can use optic flow to estimate distance of travel." Vision Research 41: 213-219.
- Reeves, B. and C. Nass (1996). The Media Equation, Cambridge University Press.
- Reeves, B. and C. Nass (1999). The Media Equation: How people treat computers, television, and new media like real people and places. CLSI Publications: Stanford, CA. USA.

- Reeves, B. and C. Nass (2000). "Perceptual Bandwidth." Communications of the ACM 43(March): 65-70.
- Reiser, J. J., D. H. Ahmead, et al. (1990). "Visual perception and the guidance of locomotion without vision to previously seen targets." Perception 19: 675-689.
- Richardson, A. E., D. R. Montello, et al. (1999). "Spatial Knowledge Acquisition From Maps And From Navigation In Real And Virtual Environments." Memory & Cognition 27(4): 741-750.
- Roberts, J. and O. Slattery (2000). "Display characteristics and the impact on usability for stereo." Proceedings of SPIE -Stereoscopic Displays and Virtual Reality Systems VII 3957: 128-137.
- Robertson, G. G., S. K. Card, et al. (1993). "Non immersive virtual reality." IEEE Computer 26(2): 81-83.
- Robinett, W. and J. P. Rolland (1992). "A Computational Model for the Stereoscopic Optics of Head-Mounted Display." Presence 1(1): 45-62.
- Rock, I., J. Shallo, et al. (1978). "Pictorial depth and related constancy effects as a function of recognition". Abstract. Perception 7(1): 3-19.
- Roscoe, S. N. (1984). "Judgments Of Size And Distance With Imaging Displays." Human Factors 26(2): 617-629.
- Rossano, M. J. and J. Moak (1998). "Spatial Representations Acquired From Computer Models: Cognitive Load, Orientation Specificity And The Acquisition Of Survey Knowledge." British Journal of Psychology 89: 481-497.
- Ruddle, R. A. and D. M. Jones (2001). "Movement in Cluttered in Virtual Environments." Presence: Teleoperators and Virtual Environments 10(5): 511-524.
- Ruddle, R. A., S. J. Payne, et al. (1998). "Navigating large -scale "Desk-top" Virtual Buildings: Effects of orientation aids and familiarity." Presence: Teleoperators and Virtual Environments 7(2): 179-192.
- Rushton, S. and J. Wann (1993). Problems in perception and action in virtual worlds. Virtual Reality International '93, Third Annual Conference on Virtual Reality, London, Meckler; Westport & London.
- Rushton, S., M. Mon-Williams, et al. (1994). "Binocular vision in a bi-ocular world: New generation head-mounted displays avoid causing visual deficit." Displays 15: 255-260.
- Sadoswki, W. and K. M. Stanney (2002). Presence in Virtual Environments. Handbook of Virtual Environments: Design, Implementation, and Applications. K. M. Stanney (Ed.), Lawrence Erlbaum Associates: 791-806.
- Santos, C. L., G. Bacocoli, et al. (2003). The collaborative immersive visualization environments to oil companies. Virtual Concept 2003.
- Schiffman, H. R. (1990). Sensation and Perception: An Integrated approach, John Wiley and Son.
- Schiffman, H. R. (2000). Sensation and perception. An integrated Approach., John Wiley & Sons, Inc.
- Scott, J. and B. Dalgarno (2001). Interface Issues for 3D Motion Control. OzCHI 2001, Perth.

- Sedgwick, H. A. (1991). The effects of viewpoint on the virtual space of pictures. In Pictorial Communication In Virtual and Real Environment. S. R. Ellis, M. K. Kaiser and A. J. Grunwald (Eds.), Taylor & Francis Ltd.
- Sekuler, R. and R. Blake (1994). Perception, McGraw-Hill International Editions Psychology Series.
- Sherman, B. and P. Judkins (1992). Glimpses of Heaven, Visions of Hell: Virtual Reality and its implications. London, Hodder and Stoughton.
- Siegel, A. W. and S. H. White (1975). The development of spatial representation of large-scale environments. Advances in child development and behaviour. H. W. Reese (Ed.), New York, Academic Press.
- Sinai, M. J., W. K. Krebs, et al. (1999). Egocentric Distance Perception In A Virtual Environment Using A Perceptual Matching Task. 43rd Annual Meeting Human Factors and Ergonomics Society.
- Slater, M., M. Usoh, et al. (1995). The influence of shadows on presence in immersive virtual environments. Virtual Environments '95, New York, Springer Computer Science.
- Slater, M., V. Linakis, et al. (1996). Immersion, presence and performance in virtual environments: An experiment using tri-dimensional chess. ACM Virtual Reality Software and Technology (VRST)
- Smith, O. W. (1958). "Comparison of Apparent Depth in a Photograph Viewed From Two Distances." Perceptual and Motor Skills 8: 79-81.
- Smith, O. W. and H. Gruber (1958). "Perception of Depth in Photographs." Perceptual and Motor Skills 8: 307-313.
- Smith, O. W. and P. C. Smith (1961). "Ball throwing responses to photographically portrayed targets." Journal of Experimental Psychology 62(3): 223-233.
- Stanney, K. M. and M. J. Zyda (2002). Virtual Environment in the 21st Century. Handbook of Virtual Environments: Design, Implementation and Applications. Stanney, K.M.(Ed.) London, Lawrence Erlbaum Associates: 1-14.
- Stanney, K. M., K. S. Kingdon, et al. (2002). "Human performance in immersive virtual environments: effects of exposure duration, user control, and scene complexity." Human Performance 15(4): 339-366.
- Stanney, K. M., R. R. Mourant, et al. (1998). "Human Factors Issues In Virtual Environments: A Review Of The Literature." Presence:Teleoperators and Virtual Environments 7(4): 327-351.
- Storms, R. L., M. J. Zyda, et al. (2000). "Interactions in Perceived Quality of Auditory-Visual Displays." Presence: Teleoperators and Virtual Environments 9(6): 557-580.
- Surdick, R. T., E. T. Davis, et al. (1997). "The perception of distance in simulated visual displays: A comparison of the effectiveness and accuracy of multiple depth cues across viewing distances." Presence: Teleoperators and Virtual Environments 6(5): 513-531.
- Swaminathan, K. and S. Sato (1997). "Interaction design for large display." Interactions 4(1): 15-24.
- Swenson, H. A. (1932). "The relative influence of accommodation and convergence in the judgment of distance." Journal of General Psychology 7(360-380).



- Takeda, T., K. Hashimoto, et al. (1999). "Characteristics of accommodation toward apparent depth." Vision Research 39: 2087-2097.
- Tan, D. S., D. Gergle, et al. (2003). With Similar Visual Angles, Larger Displays Improve Spatial Performance. CHI 2003, Ft Lauderdale-Florida.
- Tan, D. S., M. Czerwinski, et al. (2003). Women go with the (Optical) Flow. CHI 2003, Ft Lauderdale, Florida, USA.
- Teghsoonian, M. and R. Teghsoonian (1969). "Scaling of apparent distance in natural indoor settings." Psychonomics Science 16(6): 281-283.
- Teghtsoonian, R. and M. Teghtsoonian (1970). "Scaling apparent distance in a natural outdoor setting." Psychonomic Science 21(4): 215-216.
- Thiyagarajan, R. (2003). Use of 3D Laser Scanning and Close Range Photogrammetry for Shipbuilding. 2003 Ship Production Track of the World Maritime Conference, San Francisco, California, USA.
- Thompson, W. B., P. Willemson, et al. (in press). "Does the quality of the computer graphics matter when judging distances in visually immersive environments?" Presence: Teleoperators and Virtual Environments, 2004.
- Tong, F. H., S. G. Marlin, et al. (1995). Cognitive Map Formation in a 3D Visual Virtual World. IRIS/PRECARN Workshop 1995.
- Vince, J. (1995). Virtual Reality systems, Addison-Wesley.
- Vince, J. (2000). Essential computer animation: How to understand techniques and potential of computer animation. London, Springer-Verlag.
- Wade, N. J. and M. T. Swanston (2001). Visual Perception: An Introduction, Psychology Press:Taylor & Francis Group.
- Walker, M. D. (1998). Course notes lectures. Available at <http://www.colorado.edu/epob/epob4640mwalker/lect6.html>.
- Waller, D. (1999). "Factors Affecting the Perception of Interobject Distances in Virtual Environments." Presence: Teleoperators and Virtual Environments 8(6): 657-670.
- Waller, D., E. Hunt, et al. (1998a). Measuring spatial knowledge in a virtual environment: Distances and angles, 39th Annual Meeting of the Psychonomic Society, Dallas, TX, USA.
- Wanger, B. (1992). The effect of shadow quality on the perception of spatial relationships in computer generated imagery. 1992 ACM SIGGRAPH Symposium on Interactive 3D Graphics, New York, ACM.
- Wanger, B., J. A. Ferweda, et al. (1992). "Perceiving spatial relationships in computer-generated image." IEEE Computer Graphics & Applications 12(3): 44-58.
- Wann, J. and M. Mon-Williams (1996). "What does virtual reality NEED?:Human factors issues in the design of three-dimensional computer environments." International Journal Human-Computer Studies 44: 829-847.

- Wann, J. P. and M. Mon-Williams (2002). Measurement of visual aftereffects following virtual environment exposure. In Handbook of Virtual Environments: Design, Implementation, and Applications. K. M. Stanney (Ed.), Lawrence Erlbaum Associates.
- Wann, J. P., S. Rushton, et al. (1995). "Natural problems for stereoscopic depth perception in virtual environments." Vision Research 35(19): 2731-2736.
- Wartenberg, C. and P. Wiborg (2003). "Precision of exocentric distance judgment in desktop and cube presentation." Presence: Teleoperators and Virtual Environments 12(2): 196-206.
- Weinhaus, F. M. and V. Devarajan (1997). "Texture mapping 3D models of real-world scenes." ACM Computing Surveys 29(4): 325-365.
- Werkhoven P. & Groen J. (1998a). "Manipulation performance in interactive virtual environments." Human Factors. 40(3): 432-442.
- Werkhoven P. & Groen J. (1998b). "Natural interactions in virtual environments. Modern Simulation & Training" 5: 4-13.
- Willemson, P. and A. A. Gooch (2002). Perceived Egocentric Distance in Real, Image-based, and Traditional Virtual Environments. IEEE VR2002 (IEEE Virtual Reality Conference), March 2002, Orlando, Florida, USA.
- Willemson, P., M. B. Colton, et al. (2004). The effects of head-mounted display mechanics on distance judgments in virtual environments. First SIGGRAPH Symposium on Applied Perception in Graphics and Visualization, August 2004, L.A. USA.
- Wilson, P. N. (1999). "Active Exploration of a Virtual Environment Does Not Promote Orientation or Memory for Objects." Environment and Behavior 31(6): 752-763.
- Wilson, P. N., N. Foreman, et al. (1997). "Active Versus Passive Processing Of Spatial Information In A Computer-Simulated Environment." Ecological Psychology 9(3): 207-222.
- Witmer, B. G. and P. B. Kline (1998). "Judging Perceived And Traversed Distance In Virtual Environment." Presence 7(2): 144-167.
- Witmer, B. G. and W. J. J. Sadoski (1998). "Non-visually guided locomotion to a previously viewed target in real and virtual environments." Human Factors 40(3): 478-488.
- Witmer, B.G. and M. Singer (1998). "Measuring presence in virtual environments: A Presence Questionnaire." Presence: Teleoperators and Virtual Environments 7(3): 225-240.
- Witmer, B.G., J. H. Bailey, et al. (1996). "Virtual space and real world spaces: transfer of route knowledge." International Journal Human-Computer Studies 45: 413-428.
- Wolfe, R. (1997). "SIGGRAPH 97 Technical Slide Set Credits: Mapping Techniques." Computers & Graphics 31(4): 61-63.
- Wright, R. H. (1995). Virtual Reality Psychophysics: Forward and Lateral Distance, Height and Speed Perceptions with a Wide-Angle Helmet Display. ARI Technical Report 1025, U.S. Army Research Institute for the Behavioral and Social Sciences.. Alexandria, VA, USA.
- Yanagasiwa, M. and K. Akahori (1999). "The Influence of Photographic Reality in Virtual Environment on Spatial Representation." Japan Journal of Educational Psychology 22(4): 239-249.

- Yang, T. L., M. W. Dixon, et al. (1999). "Seeing Big Things: Overestimation Of Heights Is Greater For Real Objects Than Objects In Pictures." Perception 28: 445-467.
- Yeh, Y.-Y. and L. D. Silverstein (1992). "Spatial judgments with monoscopic and stereoscopic presentation of perspective displays." Human Factors 34(5): 583-600.
- Yoon, J., E. Byun, et al. (2000). Comparison of space perception between a real environment and a virtual environment. IEA 2000/HFES 2000 Congress.
- Youngblut, C. and O. Huie (2003). The relationship between presence and performance in virtual environments: Results of a VERTS study, IEEE Virtual Reality 2003 (VR'03). IEEE Computer Society.
- Zhai, S. and I. S. MacKenzie (1998). Teaching old mice new tricks: Innovations in computer mouse design. Ergon-Axia'98 - the First World Congress on Ergonomics for Global Quality and Productivity.

## APPENDIX A

### EXPERIMENT 1 ON DISTANCE PERCEPTION IN STATIC IMAGES: SUMMARIES, TEST MATERIALS & COLLECTED DATA

#### A.1 Experiment 1 Hypotheses

Experiment	Hypotheses
1A	<p>H1: There is no significant different between image type (Real and VE image) on asymmetrical distance estimate tasks (horizontal and transverse).</p> <p>H2: There is no significant different between display type (large projected display and desktop monitor) on distance estimate tasks (horizontal and transverse).</p>
1B	<p>Primary hypothesis are:</p> <p>H1: There is no significant different between large and small on asymmetrical distance estimation tasks (vertical, horizontal, transverse).</p> <p>Secondary hypotheses are:</p> <p>H2: There is no effect of viewing distance on asymmetrical distance perception (vertical, horizontal, transverse)</p> <p>H3: There is no effect of physiological cues on asymmetrical distance perception (vertical, horizontal, transverse)</p>

#### A.2 Summary of Factors controlled in Experiment 1

Factors/variables	Experiment 1A	Experiment 1B
VE image resolution	High (1028 x 986)	Not applicable (only Real image)
Real image resolution	High (1280 x 960)	High (1280 x 960)
Physical image size	Different for large and small display	Different for large and small display
FOV(horizontal and vertical)	Same for large and small display	Different for large and small display
Retinal image size	Same for large and small display	Different for large and small display
Viewing distance	Different for large and small display	Same for large and small display
Physiological cues	Different for large and small display	Same for large and small display

#### A.3 Summary of Display setup in Experiment 1

	Experiment 1A		Experiment 1B	
	Small display	Large display	Small display	Large display
Image size	30 x 40cm	136 x 179 cm	39 x 52 cm	156 x 208 cm
Viewing distance	40cm	280 cm	100cm	100cm
Vertical FOV	21°	22°	22°	36°
Horizontal FOV	18°	18°	29°	92°

**A.4 Summary of results for Experiment 1**

Analysis	Distance type	Experiment 1A Results	Experiment 1B results
Overall analysis	Vertical distance	Not investigated	1. generally overestimated 2. NO main effect of display VE – proj. disp > small
	Horizontal distance	1. generally underestimated 2. No main effect of image Large display - RE > VE Small display - VE > RE 3. Main effect of display approach significance (p = .052) RE – proj. disp > desktop VE – proj. disp > small 4. NO interaction effect	1. generally underestimated 2. Not investigated 3. NO main effect of display VE – proj. disp > small 4. Not applicable
	Transverse distance	1. generally underestimated 2. NO main effect of image Proj. disp - very small difference Desktop - very small difference 3. Main effect of display RE – proj. disp > desktop VE – proj. disp > desktop 4. NO interaction effect	1. generally underestimated 2. Not investigated 3. NO main effect of display (approach significant, p = .056) VE – proj. disp > desktop 4. Not applicable
Examine individual distance	Vertical distance	Not applicable as only one distance is involved	1. No main effect of display 2. Generally, Proj. disp > desktop for all distances
	Horizontal distance	Not applicable as only one distance is involved	1. No main effect of display 2. Generally, Proj. disp > desktop for all distances (except #1)
	Transverse distance	Not applicable as only one distance is involved	1. No main effect of display 2. Generally, Proj. disp > desktop for all distances (except #3)
Comparison among experimental conditions	Vertical distance	Not investigated	Significant difference between VE/large and VE/small
	Horizontal distance	No significant difference among all pairs of comparison	No significant difference between experimental condition
	Transverse distance	No significant difference among all pairs of comparison	No significant difference between experimental condition

Note: '>' refers to 'performed better than', '<' refers to 'performed less than',  
'proj. disp' refers to 'projected display', RE = Real, VE = Virtual environment



A.5 Test Materials Collected Data and Statistical Analysis of Data for Experiment 1

A.5.1 Experiment 1A

A.5.1.1 Experiment 1A Data sheet

Experiment No.: ..... Date: .....  
Group No. : ..... Conditions: .....

ID / Name: ..... Staff / Student :  
Age: ..... Gender: .....

Data to be collected : Distance Estimation(in metres), X and Y.

Please write your estimations in the following table.

	Estimated Distance (in metres)
Distance between the two lampposts (X)	
Distance between lamppost and the hedge (Y)	

Post-test questionnaires

Please answer the following questions:

1. Briefly describe how do you (or any strategies you used to) estimate the distances.  
.....  
.....

2. Which one of the following is easier to estimate? (please circle ONE only)  
  
a. X – distance between the two lampposts  
b. Y – distance between lamppost and the hedge  
c. Both  
d. No difference

3. Provide reasons to support your answer in 2.  
.....  
.....

NOTE: Please note that all information here will be dealt with confidentiality and will be only be used for data analysis and reporting purposes.

**A.5.1.2 Experiment 1A -Instruction sheet****Instructions to participants:**

You will be presented with a picture.

- a. Your task is to estimate two distances:
  - i. X - the distance between the two lamp posts in the pictures. Lamppost 1 ( nearest to you). lamp post 2 (farthest from you).
  - ii. Y – the distance between the lamppost 1 and the hedge on your right .
  - iii. Estimation is to be made in metre unit
- b. You will be given only fifteen seconds for each estimation. You will be reminded when the time is up and you are to write down your answer immediately on the sheet provided.

Questions may not be asked during experiments, so please clear up any questions before we begin the experiment.

A.5.1.3 Experiment 1A – Collected data

	Image Type	Display type	Transverse	Horizontal
			Distance X	Distance Y
1	Real Image	Desktop	2.5	3
2	Real Image	Desktop	100	180
3	Real Image	Desktop	15	7
4	Real Image	Desktop	8	5
5	Real Image	Desktop	20	30
6	Real Image	Desktop	6	5
7	Real Image	Desktop	10	7
8	Real Image	Desktop	15	11
9	Real Image	Desktop	8	4
10	Real Image	Desktop	4.6	3.1
11	Virtual Image	Desktop	10	15
12	Virtual Image	Desktop	8	4
13	Virtual Image	Desktop	2	3
14	Virtual Image	Desktop	5	6
15	Virtual Image	Desktop	4	4
16	Virtual Image	Desktop	3	3
17	Virtual Image	Desktop	20	10
18	Virtual Image	Desktop	15	8
19	Virtual Image	Desktop	10	10
20	Virtual Image	Desktop	10	10
21	Real Image	Large Screen	7.5	6
22	Real Image	Large Screen	12	10
23	Real Image	Large Screen	30	15
24	Real Image	Large Screen	10	12
25	Real Image	Large Screen	20	5
26	Real Image	Large Screen	6.2	4
27	Real Image	Large Screen	5	4
28	Real Image	Large Screen	20	15
29	Real Image	Large Screen	15	10
30	Real Image	Large Screen	6.2	3.7
31	Virtual Image	Large Screen	8	6
32	Virtual Image	Large Screen	18	9
33	Virtual Image	Large Screen	15	7
34	Virtual Image	Large Screen	18	15
35	Virtual Image	Large Screen	12	8
36	Virtual Image	Large Screen	15	10
37	Virtual Image	Large Screen	20	15
38	Virtual Image	Large Screen	4	6
39	Virtual Image	Large Screen	20	15
40	Virtual Image	Large Screen	5	5

A.5.1.3 Experiment 1A Statistical analysis results

- I. Comparison among experiment conditions using Microsoft office Excel Student t-test

Conditions compared	transverse distance
	student t-test values
Real/small vs Real/large	0.1528
Virtual/small vs Virtual/large	0.0820
Real/small vs Virtual/small	0.9795
Real/large vs Virtual/large	0.9236

Conditions compared	horizontal distance
	student t-test values
Real/small vs Real/large	0.1182
Virtual/small vs Virtual/large	0.2137
Real/small vs Virtual/small	0.3053
Real/large vs Virtual/large	0.5612

II. Comparison among distance types

Experimental conditions	horizontal-transverse
real/small	0.0668
real/large	0.0737
virtual/small	0.0165
virtual/large	0.0101

A.5.2 Experiment 1B

A.5.2.1 Experiment 1B Data sheet

Experiment No.: .....Date: .....

Group No. : .....Conditions: .....

ID / Name: .....Occupation: ... Staff / Student

Age: .....Gender: .....

**Data to be collected:** Distance Estimation (in meters)

Please write down ESTIMATED DISTANCE only in the column provided.

No.	Distance to estimate	Estimated Distance (in meters)
	1 VERTICAL DISTANCE	
1	Height of Building	
2	Height of Tree	
3	Height of lamppost	
	2 HORIZONTAL DISTANCE	
4	Distance of from edge of building to edge of footpath (the grassy area)	
5	Distance from left edge of footpath to stop sign on the road	
6	Distance of the roof the main entrance of the building	
	3 TRANSVERSE DISTANCE	
7	Distance of concrete part of the footpath	
8	Distance of the two rails in front of the main entrance of the building	
9	Distance of black square on the road to the arrow sign at the end of the road	

A.5.2.2 Experiment 1B -Instruction sheet

Instructions to participants:

1. The purpose of the study is to investigate participants' distance perception of distance in still images.
2. Your task is to estimate distances of and between objects. There will 9 distances to estimate.
3. You will be presented with a picture. Please remain seated at the designated chair. The experimenter will adjust the position and height of your seat so that your eye level is at the centre of the display. Please do not lean forward or backward.
4. Your task is to estimate distances and heights of certain objects in the picture.
5. You will be presented with a picture for fifteen seconds prior to each estimation. You will be told what distances to estimate.
6. You will be reminded when the time is up and you are to write down your answer immediately on the sheet provided.
7. Steps 4,5,6 will be repeated for each of the nine distances.
8. All estimations are to be made in meter unit. Participants are shown a meter long tape to remind them the length of a meter prior to the start of the experiment.
9. You will be required to fill in a short questionnaire after the test.
10. Participants should be advised that they could withdraw from the experiment at any time without having to give reason.

Thank you for your participation.

A.5.2.4 Experiment 1B - Post-test Questionnaire

Post-test questionnaires

Please answer the following questions:

1. Briefly describe how do you (or any strategies you used to) estimate the distances.  
.....  
.....
2. i. Which distance (vertical, horizontal or transverse) you find most easy to estimate? Please provide reasons.  
.....  
.....  
ii. Which distance (vertical, horizontal or transverse) you find most difficult to estimate? Please provide reasons.  
.....  
.....
4. How accurate do you feel is your estimations? (Please tick one)  
Uncertain    ☐ 1    ☐ 2    ☐ 3    ☐ 4    ☐ 5    ☐ 6    ☐ 7    Very certain  
.....

NOTE: Please note that all information here will be dealt with confidentiality and will be only used for data analysis and reporting purposes



A.5.2.4 Experiment 1B –Collected Data

No		v1	v2	v3	h1	h2	h3	t1	t2	t3
	actual data	4.8	10.1	4.2	6.8	9	5.2	3.6	14.2	51.6
1	Large Screen	7	14	8	4	7	4	3.5	6	9
2	Large Screen	7	10	4.5	3	20	12	2.5	7	22
3	Large Screen	9	18	12	2.5	3	2.5	3	2.8	5.5
4	Large Screen	7	16	5	5	15	4	4	7	10
5	Large Screen	9	16	5	4	7	4.5	2	5	8
6	Large Screen	9	11.5	10	4.5	13	4	2.5	6.5	17
7	Large Screen	4	6.5	3.1	3.8	6.8	4.6	3	5.5	12
8	Large Screen	3	5	2.5	2	6	3	2	7	8
9	Large Screen	5.5	9	6.5	5	8.5	4.5	3	7.7	15
10	Large Screen	6	10	3	5	5	5	1.5	8	12
11	Small Screen	4	6	3	2.5	5	2.5	2	3.5	6.5
12	Small Screen	6	12	2.5	3	7	3	1	2	20
13	Small Screen	6.2	10.8	4.6	2.5	7.7	3.7	2.2	6.2	18.5
14	Small Screen	1.5	3	2	2	2.5	1	1	2	2.5
15	Small Screen	15	25	15	2	12	8	1.5	2.5	20
16	Small Screen	5	12	10	5	15	10	4	8	24
17	Small Screen	5	7	4.5	2.5	6	3.5	2	6	13
18	Small Screen	4.5	15	6	3.5	6	5	2.5	6	20
19	Small Screen	3.5	7	6	2	15	6.5	2	12	40
20	Small Screen	15	25	12.5	6.5	6.5	7	3.5	5.7	23

A.5.2.4 Experiment 1B – Statistical Analysis Data

I. Comparison among experiment conditions using Microsoft office Excel Student t-test

	vertical
Conditions compared	student t-test values
Virtual/small vs Virtual/large	0.0042

	horizontal
Conditions compared	student t-test values
Virtual/small vs Virtual/large	0.4828

	transverse
Conditions compared	student t-test values
Virtual/small vs Virtual/large	0.4015

II. Comparison among distance types

Distance types compared	VE/large	VE/small
vertical-horizontal	0.0000	0.0000
vertical- transverse	0.0000	0.0000
horizontal-transverse	0.0002	0.0178

## APPENDIX B

### EXPERIMENT 2 ON DISTANCE PERCEPTION IN DYNAMIC IMAGES: SUMMARIES, TEST MATERIALS & COLLECTED DATA

#### B.1 Experiment 2 - Hypotheses

Experiment	Hypotheses
2A	<p>H1: There is no effect of image type (real and VE image) on asymmetrical distance perception (vertical, horizontal, transverse).</p> <p>H2: There is no effect of display size (small and large size) on asymmetrical distance perception (vertical, horizontal, transverse).</p>
2B	<p>Primary hypothesis</p> <p>H1: There is no significant different between large and small on asymmetrical distance estimation tasks (vertical, horizontal, transverse)</p> <p>Secondary hypotheses are:</p> <p>H2: There is no effect of viewing distance on asymmetrical distance perception (vertical, horizontal, transverse)</p> <p>H3: There is no effect of physiological cues on asymmetrical distance perception (vertical, horizontal, transverse)</p>

#### B.2 Summary of Factors controlled in Experiment 2

Factors/variables	Experiment 2A	Experiment 2B
VE image resolution	Low (200 lines of resolution)	High (1280 x 1028)
Real image resolution	Same image with the same resolution is used for both experiment	
Physical image size	Different for large and small display	Different for large and small display
FOV(horizontal and vertical)	Same for large and small display	Different for large and small display
Retinal image size	Same for large and small display	Different for large and small display
Viewing distance	Different for large and small display	Same for large and small display
Physiological cues	Different for large and small display	Same for large and small display

#### B.3 Summary of Display setup in Experiment 2

	Experiment 2A		Experiment 2B	
	Small display	Large display	Small display	Large display
Image size	30 x 40cm	136 x 179 cm	39 x 52 cm	156 x 208 cm
Viewing distance	60cm	272 cm	100cm	100cm
Vertical FOV	28°	28°	22°	36°
Horizontal FOV	18°	18°	29°	92°

## B.5 Summary of Results for Experiment 2

Analysis	Distance type	Experiment 2A Results	Experiment 2B results
Overall analysis	Vertical distance	1. generally underestimated (except on VE/large condition)  2. Main effect of image Large display - VE > RE Small display - VE > RE  3. NO main effect of display RE - small > large VE - small > large  4. NO interaction effect	1. generally underestimated  2. Main effect of image Large display - VE > RE Small display - RE > VE  3. NO main effect of display RE - large > small VE - large > small  4. NO interaction effect
	Horizontal distance	1. generally underestimated  2. Main effect of image Large display - RE > VE Small display - VE > RE  3. NO main effect of display RE - small > large VE - small > large  4. NO interaction effect	1. generally underestimated  2. NO main effect of image Large display - VE > RE Small display - RE > VE  3. NO main effect of display RE - large > small VE - large > small  4. YES interaction effect Large > small for real and VE
	Transverse distance	1. generally underestimated  2. Main effect of image Large display - RE > VE Small display - VE > RE  3. NO main effect of display RE - small > large VE - small > large  4. NO interaction effect	1. generally underestimated  2. Main effect of image Large display - VE > RE Small display - VE > RE  3. NO main effect of display RE - small > large VE - large > small  4. NO interaction effect
Examine individual distance	Vertical distance	1. No main effect of - image (except #5) - display No interaction effect  2. Generally, Large display - VE > RE (except #3) Small display - VE > RE (except #3)  3. Generally, RE - small > large VE - small > large (except #3)	1. No main effect of a. image (except #5) b. display No interaction effect  2. Generally, Large display - VE > RE (except #3) Small display - VE > RE (except #5)  3. Generally, RE - large > small (except #1,4) VE - large > small

	Horizontal distance	<p>1. No main effect of</p> <ul style="list-style-type: none"><li>- image</li><li>- display</li></ul> <p>No interaction effect</p> <p>2. Generally,</p> <p>Large display - RE &gt; VE (except #3)</p> <p>Small display - VE &gt; RE (except # 1,3)</p> <p>3. Generally,</p> <p>RE – small &gt; large</p> <p>VE – small &gt; large (except #1)</p>	<p>1. No main effect of</p> <ul style="list-style-type: none"><li>- image (except #5)</li><li>- display</li></ul> <p>No interaction effect</p> <p>2. Generally,</p> <p>Large display - VE &gt; RE (except #2)</p> <p>Small display - VE &gt; RE (except #2,3)</p> <p>3. Generally,</p> <p>RE – small &gt; large (except #1)</p> <p>VE – large &gt; large (except #2)</p>
	Transverse distance	<p>1. No main effect of</p> <ul style="list-style-type: none"><li>a. image</li><li>b. display (except #3)</li></ul> <p>No interaction effect</p> <p>2. Generally,</p> <p>Large display - RE &gt; VE (except # 1,4)</p> <p>Small display - VE &gt; RE (except # 2,5)</p> <p>3. Generally,</p> <p>RE – small &gt; large (except # 2,4)</p> <p>VE – small &gt; large (except #3)</p>	<p>1. No main effect of</p> <ul style="list-style-type: none"><li>- image (except #3,4)</li><li>- display</li></ul> <p>No interaction effect</p> <p>2. Generally,</p> <p>Large display - VE &gt; RE (except #2)</p> <p>Small display - VE &gt; RE (except #2,5)</p> <p>3. Generally,</p> <p>RE – small &gt; large (except #1)</p> <p>VE – small &gt; large (except #3)</p>
Comparison among experimental conditions	Vertical distance	No significant difference for all pair of comparison	Significant difference between VE/large and VE/small RE/large and VE/large
	Horizontal distance	Significant difference between VE/large and VE/small only	Significant difference between VE/large and VE/small RE/large and VE/large
	Transverse distance	Significant difference between VE/large and VE/small only	No significant difference for all pair of comparison RE/large and VE/large (approach significance)

Note: '>' refers to 'performed better than', '<' refers to 'performed less than',  
'proj. disp' refers to 'projected display', RE = Real, VE = Virtual environment

B.6 Test Materials and Collected Data for Experiment 2

B.6.1 Experiment 2A

B.6.1.1 Experiment 2A Data sheet

Experiment No.: ..... Date: .....  
Group No. : 1.....2.....3.....4..... Conditions: ... ..  
ID / Name: ..... Occupation: .....  
Age: ..... Gender: .....  

Data to be collected : Distance Estimation(in metres).

Please write down ESTIMATED DISTANCE only in the column provided.

No.	Distance to estimate	Estimated Distance (in metres)
	1 VERTICAL DISTANCE	
1	Height of goal post (yellow color)	
2	Height of lamppost.4	
3	Height of Tree closest to lamppost 2 (tree 1)	
4	Height of lamppost 2	
5	Height of the signpost	
6	Height of the hedge on the left	
	2 HORIZONTAL DISTANCE	
7	Distance from the lamppost 2 to lamppost 3	
8	Distance from the right edge of the goal post to the road	
9	Distance from lamppost 4 to the hedge on the right	
10	Distance from lamppost 2 to the signpost	
11	Distance of the left edge of the goalpost to the hedge on the left	
12	Distance between the legs of the goalpost	
	3 TRANSVERSE DISTANCE	
13	Distance from the litter box to the black plastic path	
14	Distance from lamppost 1 to lamppost 4	
15	Distance from the right edge of the goalpost to the tree on the right (Tree2)	
16	Distance from lamppost 4 to signpost	
17	Distance from tree1 and tree2	
18	Distance from the litter box to lamppost 1	

B.6.1.2 Experiment 2A Post-test Questionnaire

Post-test questionnaires

Please answer the following questions:

1. Briefly describe how do you (or any strategies you used to) estimate the distances.
2. Did you find the initial viewing of the movie useful for making the estimate? Provide reasons to support your answer.
3. i. Which distance(s) (vertical, horizontal or transverse) you find most easy to estimate? Please provide reasons.
- ii. Which distance(s) (vertical, horizontal or transverse) you find most difficult to estimate? Please provide reasons.
4. How accurate do you feel is your estimations? (Please tick one )  
Uncertain 1 2 3 4 5 6 7 Very certain
5. Do you play any kind of sports? If yes, please indicate.

NOTE: Please note that all information here will be dealt with confidentiality and will be only be used for data analysis and reporting purposes.



**B.6.1.3 Experiment 2A -Instruction sheet****Instructions to participants:**

1. You will be presented with a movie. Please remain seated at the designated chair and do not lean forward or move backward. The position and height of your seat will be adjusted by the experimenter so that your eye level is at the center of the display. This will be indicated also by the use of a ping-pong ball hang from the ceiling.
2. Your task is to estimate distances and heights of certain objects in the movie
3. There will be given eighteen distances and heights to estimate
4. Before making the estimates, you will be allowed to view the movie, using the FORWARD, STOP, PLAY and PAUSE button. You are encouraged to make notes of objects (trees, lamppost, goalposts, hedges, road, litterbin, etc. ) and notice the distances between objects in the movie. You are only allocated about 4 minutes to view the movie. You will be informed when the time is up.
5. The experimenter will then set a view position for you to make the estimation from.
6. You will then be told what distances/height to estimate
7. From the given view position you will be given 15 seconds to view the scene before writing down your answer on the data sheet provided.
8. You will be reminded when the time is up and you are to write down your answer immediately on the sheet provided.
9. Steps 6,7, and 8 will be repeated for each of the eighteen distances.
10. All estimations are to be made in metre unit. Participants are shown a metre long tape to remind them the length of a metre prior to the start of the experiment.
11. Participants are reminded not to move their head/body forward and backward during the estimation.
12. You will be required to fill in a short questionnaire after the test.
13. Participants should be advised that they could withdraw from the experiment at any time without having to give reason.

Thank you for your participation.

B.6.1.4 Experiment 2A – Collected data

Estimated Data

		Vertical Distance						Horizontal Distance						Transverse Distance					
	Actual distance	2.8	6.19	7.88	14.3	2.6	3.6	18.94	11.7	3.36	16.72	11.4	7.4	6.28	42.8	19.1	42.6	6.02	4.6
1	Real Small Screen	2.5	9.0	11.0	9.0	2.0	3.0	25.0	22.5	2.5	18.0	5.0	5.0	3.0	12.0	5.0	13.0	4.0	3
2		2.5	6.0	5.5	7.0	2.0	3.0	10.0	9.0	2.5	8.0	6.0	6.0	3.0	15.0	8.0	25.0	3.0	1.5
3		2.7	5.8	4.0	6.3	1.7	2.7	10.0	8.3	2.5	8.3	5.0	4.0	2.7	10.0	3.3	5.0	3.3	2.7
4		2.8	8.0	12.0	13.0	3.0	4.0	20.0	20.0	3.5	18.0	6.0	8.0	3.5	18.0	8.0	12.0	8.0	1.5
5		2.5	6.0	5.0	6.0	2.0	2.6	12.0	8.0	4.0	7.5	5.0	5.0	5.0	20.0	6.0	20.0	5.0	3
6		2.5	6.0	6.0	6.0	2.7	2.8	20.0	17.0	3.5	15.0	8.0	5.0	6.0	20.0	9.0	18.0	5.0	5
7		2.5	6.0	4.5	6.0	2.0	2.8	25.0	25.0	3.0	20.0	12.0	5.5	4.0	28.0	17.0	19.0	4.0	2
8		1.7	5.0	4.0	5.0	1.5	1.5	12.0	20.0	3.0	10.0	3.0	4.0	7.0	25.0	7.0	5.0	3.0	5
9		2.5	4.5	5.0	4.5	3.0	3.0	10.0	6.0	4.0	8.0	2.5	4.0	2.0	8.0	4.0	2.5	2.5	1
10		2.5	5.0	4.5	6.0	2.0	2.5	10.0	8.0	3.5	8.0	6.0	6.0	5.0	30.0	4.0	7.5	3.0	3.5
11	Real large screen	2.5	7.0	5.0	7.0	2.5	3.0	15.0	12.0	2.5	14.0	6.0	6.0	5.0	20.0	20.0	18.0	8.0	2.5
12		2.5	6.5	5.0	6.5	2.3	3.0	10.0	9.0	3.3	9.0	3.0	4.0	3.0	15.0	7.5	5.0	2.5	2
13		2.8	30.0	25.0	30.0	5.0	5.0	230.0	65.0	14.0	130.0	50.0	6.5	20.0	220.0	69.0	12.0	22.0	30
14		2.5	7.5	14.0	8.0	2.5	6.0	30.0	25.0	5.0	25.0	7.0	7.0	12.0	45.0	20.0	30.0	14.0	5
15		1.0	2.5	3.5	5.0	2.0	2.5	6.5	5.0	2.0	7.0	2.0	3.0	3.0	5.0	3.0	3.0	2.5	1.5
16		2.5	8.0	10.0	8.5	2.5	3.0	50.0	30.0	5.0	25.0	15.0	5.5	5.0	50.0	25.0	8.0	5.0	3
17		2.5	6.5	4.0	9.0	2.5	3.0	13.0	12.0	3.0	11.0	4.0	8.0	3.0	17.0	8.0	10.0	5.0	1
18		2.0	4.0	3.5	4.0	1.8	2.5	10.0	7.5	4.5	8.0	2.0	4.0	2.5	15.0	10.0	8.0	2.5	0.5
19		2.0	5.0	3.5	5.0	1.5	2.3	6.0	6.5	1.5	5.0	3.0	3.0	1.5	5.0	3.0	7.0	2.5	1
20		2.0	3.7	3.2	4.0	1.7	2.0	5.5	5.0	1.3	4.6	0.7	2.5	1.6	3.0	1.3	2.4	1.1	1
21	Virtual Small Screen	2.3	5.0	10.0	15.0	5.0	3.0	20.0	15.0	2.5	30.0	15.0	5.0	3.5	10.0	20.0	20.0	7.0	2
22		2.0	4.0	4.0	8.0	1.0	2.5	5.0	5.0	2.0	3.0	1.0	2.0	1.0	2.0	2.0	2.0	1.5	1
23		2.5	5.0	4.0	8.0	2.0	3.0	20.0	20.0	2.0	25.0	10.0	5.0	3.0	25.0	5.0	20.0	5.0	1
24		2.5	5.0	6.0	11.0	2.5	3.0	13.0	10.0	2.5	11.0	8.0	7.5	4.0	20.0	15.0	35.0	7.5	1
25		2.5	4.5	5.0	10.0	2.0	3.0	25.0	22.0	3.0	30.0	10.0	9.5	5.0	20.0	20.0	22.0	10.0	4
26		2.3	4.5	5.0	25.0	1.8	2.8	15.0	10.0	3.0	20.0	4.0	6.0	4.0	12.0	7.0	18.0	7.0	2
27		2.5	3.3	8.3	11.7	3.0	3.0	8.3	6.0	2.3	7.7	3.7	3.7	2.0	6.7	6.7	6.7	4.0	1.7
28		2.5	4.0	6.0	10.0	3.0	3.5	20.0	15.0	3.0	15.0	5.0	8.0	5.0	20.0	10.0	25.0	10.0	5
29		2.0	4.0	4.0	8.0	3.0	2.0	15.0	8.0	3.0	15.0	8.0	3.0	2.0	12.0	10.0	17.0	4.0	1
30		2.5	4.5	6.0	12.0	2.0	3.5	7.0	5.0	1.5	6.0	2.5	3.5	2.5	12.0	8.0	9.0	3.5	2.5
31	Virtual large screen	2.0	8.0	4.0	10.0	1.8	2.3	10.0	8.0	2.0	9.0	2.0	5.0	2.0	8.0	6.0	8.0	4.0	2
32		1.5	2.2	1.7	3.0	1.5	1.8	3.0	2.2	1.6	2.0	0.7	1.7	1.3	2.5	1.8	1.8	1.5	0.5
33		3.2	6.5	9.0	15.0	3.5	4.2	12.0	1.0	4.0	10.0	6.0	6.0	5.0	30.0	5.5	7.5	4.5	2.5
34		2.3	10.5	7.5	30.0	2.5	3.0	25.0	23.0	5.0	15.0	4.0	5.0	4.5	25.0	15.0	8.0	7.0	1
35		3.0	3.6	3.6	6.0	2.3	2.3	6.6	5.3	1.6	6.0	1.0	2.6	1.6	4.0	4.0	5.0	2.6	0.3
36		3.0	10.0	7.0	15.0	4.0	4.5	20.0	15.0	4.0	20.0	10.0	5.0	3.0	20.0	12.0	15.0	6.0	2
37		2.0	4.0	5.0	10.5	2.0	2.1	10.0	8.0	2.5	8.0	2.5	5.0	2.0	6.0	5.5	8.0	3.5	1.8
38		4.0	10.0	7.0	12.0	3.0	5.0	20.0	15.0	3.0	12.0	4.0	5.0	2.0	7.0	8.0	7.0	4.0	0.5
39		2.5	6.0	5.0	6.0	2.0	3.0	20.0	15.0	5.0	17.0	10.0	7.5	5.0	20.0	18.0	27.0	10.0	3.5
40		2.0	4.0	3.5	6.0	1.5	2.5	5.0	4.0	1.5	3.5	2.5	3.0	1.0	5.0	5.0	6.0	2.0	1

Summary of participants' information

1	Average age (Range)	36.15 (23- 52)
2.	Staff	18
3.	Student	22
4	Male	25
5	Female	15

Summary of post-test questionnaire information (Q2, 3, 4, and 5)

Question 2	Yes – 37 No – 3 No is present in Real/SS, Real/SS, Virtual/LS condition
Question 3	i. V ii. T -----24 person ----- i. H ii. V ----- 4 person ----- i. H ii. T -----4 person ----- i. V ii. T & H – 3 person ----- i. T ii. H --- 1person ----- i. V & H ii. T – 2 person ----- i. V ii. H ---2 person
Question 4	Average = 4, Highest = 6 (3 person) Lowest = 2 (4 person)
Question 5	Yes = 31 No = 9  Sports: Tennis, cycling, gym, squash, table tennis, football, badminton, jogging, netball, volleyball, ice-hockey, cricket, tennis, marathon , canoeing

Note: V = vertical, H = Horizontal, T = transverse distance

**B.6.2 Experiment 2B****B.6.2.1 Experiment 2B Data sheet**

Experiment No.: .....		Date: .....
Group No. : 1.....2.....3.....4.....		Conditions: .....
ID / Name: .....		Occupation: .....
Age: .....		Gender: .....

**Data to be collected : Distance Estimation(in metres)**

Please write down **ESTIMATED DISTANCE** only in the column provided.

No.	Distance to estimate	Estimated Distance (in metres)
	<b>VERTICAL DISTANCE</b>	
1	Height of lamppost 4	
2	Height of Tree closest to lamppost 2 (tree 1)	
3	Height of lamppost 2	
4	Height of the signpost	
5	Height of the hedge on the left	
	<b>HORIZONTAL DISTANCE</b>	
6	Distance from the right edge of the goal post to the road	
7	Distance from lamppost 4 to the hedge on the right	
8	Distance from lamppost 2 to the signpost	
9	Distance of the left edge of the goalpost to the hedge on the left	
10	Distance between the legs of the goalpost	
	<b>TRANSVERSE DISTANCE</b>	
11	Distance from the litter box to the black plastic path	
12	Distance from lamppost 1 to lamppost 4	
13	Distance from the right edge of the goalpost to the tree on the right (Tree2)	
14	Distance from lamppost 4 to signpost	
15	Distance from tree1 and tree2	

N.B.: It should be noted that the number of distance to estimate is reduce for Experiment 2B to fifteen.

B.6.2.2 Experiment 2B Instruction sheet

The instruction for Experiment 2B is similar to Experiment 2A.

B.6.2.3 Experiment 2B Post-test questionnaire

The post-test questionnaire for Experiment 2B is similar to Experiment 2A

B.6.2.4 Experiment 2B Collected Data

		Vertical					Horizontal					Transverse				
	Actual distance	6.19	7.86	14.32	2.54	3.63	11.65	3.36	16.72	11.4	7.35	6.28	42.8	19.1	42.66	6.02
1	Real Small Screen	5.0	4.0	5.5	2.0	3.0	7.0	5.0	8.0	4.0	4.0	3.0	50.0	7.0	20.0	6.0
2		1.5	2.5	3.5	1.3	2.0	3.0	2.0	3.0	1.0	2.0	1.5	1.0	2.5	3.0	1.5
3		7.0	4.5	7.0	2.0	3.5	6.0	3.0	7.5	2.5	3.0	4.0	10.0	5.0	2.5	3.0
4		3.5	2.0	4.0	1.5	1.5	6.0	1.5	5.0	2.0	3.0	1.0	7.0	3.0	1.5	2.0
5		12.0	13.0	15.0	2.0	2.5	10.0	5.0	7.0	2.0	3.5	2.0	4.0	3.0	2.0	3.0
6		9.0	10.0	9.0	4.0	5.0	10.0	3.5	10.0	7.0	3.5	2.0	20.0	7.0	15.0	6.0
7		9.0	10.0	17.0	3.0	4.0	25.0	2.5	20.0	12.0	8.0	4.0	20.0	30.0	20.0	6.0
8		7.5	7.0	12.0	2.5	4.0	13.0	3.7	14.0	5.5	5.3	3.5	15.0	8.0	6.0	5.5
9		7.0	5.0	7.0	4.0	3.0	20.0	7.5	20.0	4.0	5.0	5.0	25.0	10.0	7.0	4.0
10		15.0	10.0	15.0	3.0	4.0	25.0	10.0	35.0	20.0	10.0	10.0	40.0	20.0	20.0	10.0
11	Real large screen	9.0	12.0	9.0	7.0	6.0	12.0	6.0	12.0	6.0	5.0	2.0	6.0	10.0	8.0	10.0
12		9.0	6.8	9.5	1.5	2.0	4.5	4.3	5.5	2.9	3.8	2.5	4.6	6.3	3.7	2.5
13		8.0	5.0	8.0	2.0	3.5	10.0	3.0	8.5	4.0	7.0	3.0	6.0	4.0	5.0	4.0
14		5.0	4.5	12.0	2.5	3.0	8.0	3.5	6.5	4.0	4.5	1.5	6.0	5.0	2.5	3.0
15		7.0	6.0	8.0	1.8	3.0	9.0	2.3	9.0	3.0	6.0	4.0	9.0	4.0	3.0	5.0
16		6.7	6.0	8.0	3.0	3.5	9.0	4.5	11.0	6.0	7.0	2.8	8.0	5.0	5.0	4.0
17		4.0	4.0	5.0	1.5	2.0	7.0	3.0	7.0	3.0	3.5	2.0	4.0	4.0	6.0	5.0
18		9.2	7.7	9.2	3.1	4.6	31.0	3.7	18.5	12.3	6.2	1.8	9.2	6.2	6.2	6.2
19		6.0	7.0	11.0	1.8	3.0	15.0	3.0	15.0	5.0	6.0	3.0	15.0	7.0	5.0	5.0
20		12.0	8.0	20.0	4.0	5.0	8.0	4.0	9.0	5.0	4.0	3.0	10.0	3.0	3.0	3.0
21	Virtual Small Screen	4.5	5.0	9.5	2.5	2.5	10.0	3.0	7.5	1.0	2.5	0.8	10.0	4.0	7.0	3.0
22		7.1	6.2	12.3	3.7	3.1	7.7	3.1	10.8	6.2	4.6	1.8	15.4	9.2	23.1	2.2
23		8.0	5.0	10.0	2.0	3.0	9.0	3.0	8.0	3.0	5.0	3.0	10.0	4.5	12.0	3.0
24		3.5	4.0	7.0	1.8	2.2	7.5	2.0	6.5	3.0	4.0	1.5	6.0	3.3	4.8	2.5
25		4.0	3.5	7.5	2.0	2.2	7.0	2.5	8.0	2.0	4.5	2.5	7.0	6.0	10.0	3.8
26		4.5	7.0	14.0	3.0	3.0	14.0	2.5	12.0	4.0	9.0	2.0	15.0	15.0	20.0	6.0
27		5.0	4.5	8.0	2.0	2.0	8.0	3.5	10.0	2.0	5.0	7.0	15.0	10.0	15.0	8.0
28		5.5	6.0	13.0	1.5	2.5	12.0	1.8	12.0	8.0	6.0	2.0	15.0	11.0	18.0	5.0
29		7.0	10.0	22.0	2.0	3.0	20.0	5.0	18.0	20.0	6.0	2.0	30.0	12.0	40.0	10.0
30		3.0	2.5	7.0	2.0	3.0	6.0	1.5	4.0	3.0	3.5	1.5	10.0	5.0	12.0	3.5
31	Virtual large screen	15.0	8.0	23.0	6.2	5.0	25.0	3.0	18.0	4.0	8.0	4.0	12.0	6.0	9.0	6.0
32		15.0	12.0	20.0	2.0	4.0	20.0	4.0	15.0	2.5	4.0	2.0	10.0	5.0	4.0	3.0
33		7.0	7.0	15.0	3.0	4.0	8.0	4.0	8.0	6.0	5.0	3.0	20.0	10.0	22.0	6.0
34		4	4.0	6.0	2.0	3.0	5	3.3	6.0	3.0	3.0	2.0	4.0	3.0	2.8	2.0
35		5.5	6.0	8.3	1.5	1.5	12.0	5.5	30.0	12.0	5.0	10.0	15.0	10.0	30.0	8.0
36		4.7	7.2	14.2	3.0	3.2	12.0	3.5	13.5	5.7	5.5	2.5	7.0	8.0	10.0	5.5
37		7.5	6.0	35.0	2.5	3.0	25.0	4.0	27.0	10.0	7.4	2.0	10.0	15.0	25.0	6.0
38		1.5	7.0	15.0	3.0	3.0	8.4	5.0	14.0	10.0	8.0	3.0	20.0	30.0	30.0	20.0
39		8.0	7.0	18.0	3.2	3.0	10.0	2.0	13.0	10.0	5.0	4.0	18.0	10.0	15.0	5.0
40		5.0	10.0	18.0	2.0	3.0	10.0	4.0	20.0	14.0	4.5	3.0	10.0	8.0	10.0	10.0

Summary of participants' information

1	Average age (Range)	27.9 (18-44)
2.	Staff	3
3.	Student	37
4	Male	20
5	Female	20

Summary of post-test questionnaire information for Q2, 3, 4, and 5

2	Question 2 initial viewing of the movies useful for estimates	Yes – 31 No – 9 No is present in Real/SS, Real/SS, Virtual/LS condition
3	Question 3 ---i. easy to estimate, ii. Most difficult	i. V ii. T ----24 person ----- i. H ii. V ----2 person ----- i. H ii. T ---- 10 person ----- i. V,H ii. T – 1 person ----- i. T ii. V --- 1 person ----- i. T ii. – 1 person ----- i. V ii. H --- 1 person
4	Question 4 – accuracy rating	Average = 3.8 . Median = 3 Highest = 5 Lowest = 2
5	Question 5 – do you play sport	Yes = 23 No = 27 Sports: football, hockey, cricket, badminton, tennis, squash, bowling,

Note: V = vertical, H = Horizontal, T = transverse distance



## APPENDIX C

### EXPERIMENT 3 ON DISTANCE PERCEPTION AND SPATIAL MEMORY IN INTERACTIVE IMAGES: SUMMARIES, TEST MATERIALS & COLLECTED DATA

#### C.1 Experiment 3 – Hypothesis

Experiment	Hypotheses
3A	<p>Main hypotheses:</p> <ol style="list-style-type: none"> <li>1. The type of environment (real vs. VE model) has no effect on participants' distance estimation task (vertical, horizontal and transverse) performance</li> <li>2. The type of environment (real vs. VE model) has no effect on participants' spatial memory task performance.</li> <li>3. The display type (small vs. large) has no effect on participants' distance estimation task (vertical, horizontal and transverse) performance in interactive VE</li> <li>4. The display type (small vs. large) has no effect on participants' spatial memory task performance in interactive VE</li> <li>5. The type of input device (mouse vs. trackball) has no effect on participants' spatial memory task performance in interactive VE</li> <li>6. The different modes of travel (drive, fly) have no effect on participants' spatial memory performance in interactive VE</li> </ol> <p>Secondary hypotheses:</p> <ol style="list-style-type: none"> <li>1. There is no effect of viewing distance on distance estimate task in interactive VE</li> <li>2. There is no effect of physiological cues on distance estimate task in interactive VE</li> <li>3. There is no effect of viewing distance on spatial memory task in interactive VE</li> <li>4. There is no effect of physiological cues on spatial memory task in interactive VE</li> </ol>
3B	In this study the main is to understand the unexpected finding of Experiment 3A. Item 3-6 above is explored as hypotheses, since the experiment involved VE condition only

#### C.2 Summary of Factors controlled in Experiment 3

Factors/variables	Experiment 3A	Experiment 3B
VE image resolution	1280 x1028	1280 x 1028
Real image resolution	Not applicable since real physical environment is used as real condition	
Physical image size	Different for large and small display	Different for large and small display
FOV(horizontal and vertical)	Different for large and small display	Same for large and small display
Retinal image size	Different for large and small display	Same for large and small display
Viewing distance	Same for large and small display	Different for large and small display
Physiological cues	Same for large and small display	Different for large and small display

### C.3 Summary of Display setup in Experiment 3

	Experiment 3A		Experiment 3B	
	Small display	Large display	Small display	Large display
Image size	39 x 52 cm	156 x 208 cm	30 x 40cm	136 x 179 cm
Viewing distance	100cm	100cm	60cm	272 cm
Vertical FOV	22°	36°	28°	28°
Horizontal FOV	29°	92°	18°	18°

### C.4 Summary of results for Experiment 3

Analysis	Experiment 3A Results	Experiment 3B results
Distance estimate	Overall,  1. No difference between RE and VE 2. No difference between large and small	Overall,  1. Not investigated 2. No difference between large and small
	Height: - No difference between large and small  Width: - No difference between large and small  Length: - Main effect of display (small > large)	Height: - No difference between large and small  Width: - No difference between large and small  Length: - No difference between large and small
	Remove effects of covariates: - No difference between large and small for all distance type - small better than large	Remove effects of covariates: - No difference between large and small for all distance types - Small better than large (except for width)
	Significant difference between distance type  - Height is more accurate compared to width and length	Significant difference between distance type  - Height is more accurate compared to width and length
Spatial memory test	Overall,  1. No difference between RE and VE 2. No difference between large and small  Remove effect of covariates: - Main effect of display (small > large)	Overall,  1. Not investigated 2. No difference between large and small  Remove effect of covariates: - Interaction effect of display * travel display * travel * sport
	Interface device:  - mouse better than trackball	Interface device:  - mouse better than trackball
	Travel mode  - fly mode better than drive mode	Travel mode  - fly mode better than drive mode (with some exception)
Interface device questionnaire	Interface device:  - generally, mouse better than trackball for Q3(i) – Q3(v)	Interface device:  - generally, mouse better than trackball for Q3(i) – (v)

	Travel mode: Q2(iii) object recall – fly is better than drive Q2(i), (ii), (iv) -- drive is better than fly	Travel mode: Q2(iii) object recall – fly is better than drive Q2(i), (ii), (iv) -- drive is better than fly (some are similar – see Figure 7-30)
	Recall accuracy: - Trackball better than mouse (significant) - No difference between large and small (but large better than small)	Recall accuracy: - Mouse better than trackball - No difference between large and small (but large better than small)
Display questionnaire		- no significant difference between display size for all questions - small display is rated higher than large display for Qi and Qiii for Qii – large display is rated higher than small display

## C.5 Test Materials and Collected data for Experiment 3

### C.5.1 Instruction sheet for Real condition

#### Instruction to Real condition participants:

1. The purpose of the experiment is to investigate subject spatial awareness in the real world condition.
2. First, you will be asked to fill in a short questionnaire on your background information
3. The experiment will be divided into 3 phases:

#### **Phase 1:** test session

You will be given a list of nine objects found in the room.

There will be nine objects on the floor. You are to remember objects and their locations in the room, as you will be required to recall them later. If you have any question about the name of object, you may ask the experimenter.

You will be asked to close your eyes before entering the test room and you will be told when to open your eyes.

When ready, you will be told to move about in the room for about 3 minutes from the initial starting position.

You will be told when the time is up and you are to close your eyes immediately.

You will be escorted out of the test room.

#### **Phase 2:** Spatial recall test.

You will be given an A3 size paper showing the basic layout of the room.

The diagram represents a scaled drawing of the walls and floor of the test room.

You will also be given a list of nine objects found in the test room.

You are to mark a cross on the paper using a pencil given; a position you think is the center of the each object's location and label it with the object's name.

You can take as much time needed to complete this test.

#### **Phase 3:**

In the last phase, you will be asked to complete a post-test questionnaire.

4. You are advised that you could withdraw from the experiment at any time without having to give reason.

## C.5.2 Instruction sheet for VE condition

### Instruction for VE condition

1. The purpose of the experiment is to investigate subject spatial awareness in the VE.
2. First, you will be asked to fill in a short questionnaire on your background information
3. The experiment will be divided into 6 phases:

#### Phase 1: Practice session

You will be given a practice environment to familiarize yourself with movement in the VE using an interface device. A travel mode will be chosen for the subject.

You will be given, as much needed time to familiarize yourself with movement using the interface device.

Subject is reminded that the practice environment will be different from the test environment.

Please indicate to the experimenter when you are ready to start the test session.

#### Phase 2: Test session

You will be given 2 minute to rest before the start of the test trial.

You will be seated at designated chair in front of the projected display. Experimenter will adjust the seating height for you.

Subject is reminded to make no head/body movement during the navigation of the VE.

When ready, subject will be told to move about in the VE model for about 3 minutes from the initial starting position.

There will be nine objects on the floor (list given). You are to remember objects and their locations in the room, as you will be required to recall them later. If you have any question about the name of object, you may ask the experimenter.

You will be told when the time is up.

#### Phase 3: Spatial recall test.

You will be given an A3 size paper showing the basic layout of the room.

The diagram represents a scaled drawing of the walls and floor of the virtual room.

You will also be given a list of nine objects found in the virtual room.

You are to mark a cross on the paper using a pencil given; a position you think is the center of the each object's location and label it with the object's name.

You can take as much time needed to complete this test.

Phase 4: After completion of the spatial recall test, participants were asked to repeat phase 1 – 3 again using a different travel mode. The objects positions will be different for each test session

Phase 5: After phase 4, you will be given 5 minutes break before repeating Phase 1-4 again using a different interface device.

Phase 6: In the last phase, you will were asked to complete a posttest questionnaire.

You are advised that you could withdraw from the experiment at any time without having to give reason

C.5.3 Practice session instruction for VE

This session will allow you to familiarize yourself with movement in a VE using the interface device provided: a trackball. The interface device only allows you to move around in the VE but would not allow you to pick up or manipulate objects. The VE used will be different from the test VE. To ensure that you can navigate around the VE, you are to find and approached six coloured cubes (red, green, blue, purple, yellow and orange) found in the environment.

However, you will be allowed as much time needed to practice using the interface device.

Please indicate to the experimenter when you are ready to start the test session.

Interface device: Mouse

**Left button:** Move forward

**Right button:** Move backward

**Middle button (wheel):** This button allows you move according to where you point the cursor. Press this button continuously and move the mouse accordingly. Alternatively, you could also use the left button or right button (instead of the middle button/wheel) for this purpose

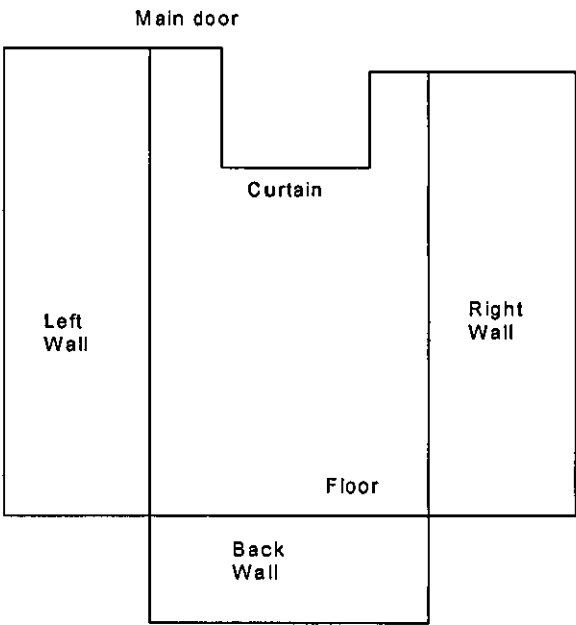
Interface device: trackball

**Left (below) button:** Move forward

**Left (above) button:** Move backward

**Wheel:** This button allows you move according to where you point cursor. Press this button continuously and roll the ball accordingly. Alternatively, you could also use the any of the left buttons (instead of the wheel) for this purpose.

C.5.4 Data sheet for the spatial memory test in the Real and VE conditions



Note: This blank map is given to participants to fill object locations. Map shown here is not drawn to scale. Actual map is drawn to scale





**C.5.6 Data sheet for participants' information**

1.	Name/ID:.....
2.	Age: .....
3.	Gender: Female/Male
4.	How often do you play computer games in a week? Please circle one.
	a. 0                      b. 1-4                      c. 5 or more
5.	How often have you participated in a VE experiment before?
	a. 0                      b. 1-4                      c. 5 or more
6.	i. Do you play any kind of sports? If yes, please indicate.
	.....
	.....
	ii. If your answer to 6(i) is yes, do you play any of the sport as
	(a) A professional
	(b) An amateur
	(c) part of a leisure activity
	(d) Others (please indicate:.....)
	Please circle one of the above choices.

C.5.7 Interface Device questionnaire

1. PLEASE CHECK ONE AND WRITE ANY COMMENT YOU MAY HAVE IN THE SPACE GIVEN.

1. Familiarity with interface device

How often do you use this input device?

I usually used it at least

	Mouse	Trackball
once a day	<input type="checkbox"/> .....	<input type="checkbox"/> .....
once a week	<input type="checkbox"/> .....	<input type="checkbox"/> .....
once a month	<input type="checkbox"/> .....	<input type="checkbox"/> .....
hardly used	<input type="checkbox"/> .....	<input type="checkbox"/> .....
never used	<input type="checkbox"/> .....	<input type="checkbox"/> .....

2. Mode of travel in the environment

i. In your opinion, which mode of travel helps you to move easily in the environment?

Mouse

Drive: Easy (1) (2) (3) (4) (5) (6) (7) Difficult

☐.....

Fly : Easy (1) (2) (3) (4) (5) (6) (7) Difficult

☐.....

Trackball

Drive: Easy (1) (2) (3) (4) (5) (6) (7) Difficult

☐.....

Fly : Easy (1) (2) (3) (4) (5) (6) (7) Difficult

☐.....

ii. In your opinion, which mode of travel helps you to control your movement in the environment?

Mouse

Drive: Less (1) (2) (3) (4) (5) (6) (7) Most control

☐.....

Fly : Less (1) (2) (3) (4) (5) (6) (7) Most control

☐.....

Trackball

Drive: Less (1) (2) (3) (4) (5) (6) (7) Most control

☐.....

Fly : Less (1) (2) (3) (4) (5) (6) (7) Most control

☐.....

iii. In your opinion, which mode of travel helps you to easily recall object position?

Mouse

Drive: Easy (1) (2) (3) (4) (5) (6) (7) Difficult

☐.....

Fly : Easy (1) (2) (3) (4) (5) (6) (7) Difficult

☐.....

Trackball

Drive: Easy (1) (2) (3) (4) (5) (6) (7) Difficult

☐.....

Fly : Easy (1) (2) (3) (4) (5) (6) (7) Difficult

☐.....

iv. In your opinion, which mode of travel do you prefer to use?

Mouse

Drive: Least (1) (2) (3) (4) (5) (6) (7) Most prefer

☐.....

Fly : Least (1) (2) (3) (4) (5) (6) (7) Most prefer

☐.....

Trackball

Drive: Least (1) (2) (3) (4) (5) (6) (7) Most prefer

☐.....

Fly : Least (1) (2) (3) (4) (5) (6) (7) Most prefer

☐.....

3. INTERFACE DEVICE.

i. In your opinion, which interface device do you find easy to use?

( )

Mouse: Easy ① ② ③ ④ ⑤ ⑥ 7 Difficult

□.....

Trackball: Easy ① ② ③ ④ ⑤ ⑥ ⑦ Difficult

□.....

ii. In your opinion, which interface device allows you to move and position yourself easily in the environment?

Mouse: Easy ① ② ③ ④ ⑤ ⑥ ⑦ Difficult

□.....

Trackball: Easy ① ② ③ ④ ⑤ ⑥ ⑦ Difficult

□.....

iii. In your opinion, which interface device allows to control your movement in the environment?

Mouse: least ① ② ③ ④ ⑤ ⑥ ⑦ most control

□.....

Trackball: least ① ② ③ ④ ⑤ ⑥ ⑦ most control

□.....

iv. In your opinion, which interface device do you feel makes it easier to recall object position?

Mouse: Easy ① ② ③ ④ ⑤ ⑥ ⑦ Difficult

□.....

Trackball: Easy ① ② ③ ④ ⑤ ⑥ ⑦ Difficult

□.....

v. In your opinion, which interface device do you prefer to use?

Mouse: Least ① ② ③ ④ ⑤ ⑥ ⑦ Most prefer

□.....

Trackball: Least ① ② ③ ④ ⑤ ⑥ ⑦ Most prefer

□.....

4. RECALL ACCURACY

How accurate do you feel on your object location recall test? (Please tick one)

Mouse: Not accurate ① ② ③ ④ ⑤ ⑥ ⑦ Very

accurate

☐.....

Trackball: Not accurate ☒1 ☐2 ☐3 ☐4 ☐5 ☐6 ☒7 Very accurate

☐.....

B. Virtual Environment model:

5. Familiarity with the location.

i. Do you recognize this room? Yes / No. .... (Please circle one)

ii. If your answer to (i) is yes, how much does this knowledge of the room assist you in your recall of objects' locations? (Please tick one)

Not helpful ☒1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 Very helpful

☐.....

6. ADDITIONAL COMMENTS.

Please write down below any additional comment that you may have with regards the experiment as a whole.

.....  
.....  
.....

~~~~~Thank you for your participation in this study~~~~~

NOTE: Please note that all information here will be dealt with confidentiality and will only be used for data analysis and reporting purposes. All reported data would be anonymous.

C.5.8 Display Device questionnaire

Subject No.:

Display size preference questions:

i.

In your opinion, which display size do you feel is easier for you to recall object position?

Large: Easy

1

2

3

4

5

6

7

Difficult

☐

.....

Small: Easy

1

2

3

4

5

6

7

Difficult

☐

.....

ii.

In your opinion, which display size do you prefer to use?

Large: Least

1

2

3

4

5

6

7

Most preferred

☐

.....

Small: Least

1

2

3

4

5

6

7

Most preferred

☐

.....

iii.

How accurate do you feel is your object location recall test on the following display size?

Large: Not accurate

1

2

3

4

5

6

7

Very accurate

☐

.....

Small: Not accurate

1

2

3

4

5

6

7

Very accurate

☐

.....



**C.5.9 To transform negatively reworded questions for questionnaire**

For analysis purposes the scale values for Question 2 (i), (iii), 3(i), 3(ii) and 3(iv) in Interface device questionnaire were reversed so that all the questions have positively worded and all have 7 as the positive response and 1 as the negative response. The transformation of data is done in SPSS using the *transform-recode* command where the following changes are made:

1 → 7  
2 → 6  
3 → 5  
4 → 4  
5 → 3  
6 → 2  
7 → 1

These transformations however do not affect the original value of the data.

For Display questionnaire, the negatively worded item was Q(i)

C.6 Experiment 3 – Collected Data

C.6.1 Experiment 3A – Collected data

C.6.1.1 REAL CONDITION DATA

I. Participants’ information

| Subject | gender | age | sport-i | sport-ii                       | sport-iii |
|---------|--------|-----|---------|--------------------------------|-----------|
| 1       | M      | 42  | N       |                                |           |
| 2       | M      | 30  | Y       | FOOTBALL,BADMINTON             | C         |
| 3       | M      | 40  | Y       | FOOTBALL                       | C         |
| 4       | M      | 27  | Y       | FOOTBALL,GOLF,TENNIS,BADMINTON | C         |
| 5       | M      | 39  | Y       | FOOTBALL                       | C         |
| 6       | M      | 40  | Y       | GOLF, BADMINTON                | C         |
| 7       | M      | 40  | N       |                                |           |
| 8       | M      | 34  | Y       | FOOTBALL                       | C         |
| 9       | F      | 26  | Y       | NETBALL,BADMINTON,BASKETBALL   | C         |
| 10      | M      | 43  | Y       | VOLLEYBALL                     | C         |

sport-i - play sport or not ?  
sport-ii- list of sport  
sport-iii - A= professional, B=amateur, C=leisure, D= others

Question on familiarity with room

| Subject | A.i | A. ii | B |
|---------|-----|-------|---|
| 1       | Y   | 4     | 4 |
| 2       | Y   | 5     | 5 |
| 3       | Y   | 6     | 6 |
| 4       | Y   | 5     | 6 |
| 5       | Y   | 4     | 3 |
| 6       | Y   | 7     | 6 |
| 7       | Y   | 6     | 6 |
| 8       | Y   | 3     | 5 |
| 9       | Y   | 4     | 5 |
| 10      | Y   | 6     | 5 |

|      |                             |
|------|-----------------------------|
| A.i  | Familiar with room?         |
| A.ii | Does A.i help in estimation |
| B    | How accurate is estimation  |

II. Room size data

|               | Display | height | width | length |
|---------------|---------|--------|-------|--------|
| 1             | Real    | 3      | 6     | 10     |
| 2             | Real    | 2      | 4     | 5      |
| 3             | Real    | 6.5    | 6.8   | 14.5   |
| 4             | Real    | 9      | 10    | 14     |
| 5             | Real    | 3.5    | 6.24  | 13.7   |
| 6             | Real    | 2      | 10    | 30     |
| 7             | Real    | 20     | 6     | 17     |
| 8             | Real    | 2.13   | 6.1   | 15.2   |
| 9             | Real    | 4.5    | 5     | 10     |
| 10            | Real    | 24     | 11    | 12.2   |
| Actual length |         | 3.5    | 7.28  | 14.81  |

## III. Spatial memory test data

| subject | No. of correct objects | Map test time (s) |
|---------|------------------------|-------------------|
| 1       | 3                      | 270               |
| 2       | 6                      | 173               |
| 3       | 5                      | 281               |
| 4       | 9                      | 75                |
| 5       | 5                      | 206               |
| 6       | 4                      | 135               |
| 7       | 1                      | 204               |
| 8       | 6                      | 172               |
| 9       | 8                      | 225               |
| 10      | 4                      | 107               |

## C.6.1.2 VE CONDITION DATA –EXPERIMENT 3A

## I. Participations' Information

| Subject | Display | gender | co-games | VE-particip | sport-i | sport-ii                    | sport-iii |
|---------|---------|--------|----------|-------------|---------|-----------------------------|-----------|
| 1       | large   | F      | A        | A           | N       |                             |           |
| 2       | large   | F      | B        | B           | Y       | TABLE TENNIS                | C         |
| 3       | large   | M      | C        | B           | Y       | FOOTBALL                    | C         |
| 4       | large   | M      | B        | B           | Y       | VOLLEYBALL                  | C         |
| 5       | large   | M      | B        | B           | Y       | SQUASH                      | B         |
| 6       | large   | M      | C        | A           | Y       | TABLE TENNIS                | B         |
| 7       | large   | M      | B        | B           | N       |                             |           |
| 8       | large   | F      | A        | A           | Y       | BADMINTON,SQUASH            | C         |
| 9       | large   | M      | C        | A           | N       |                             |           |
| 10      | large   | M      | B        | B           | Y       | FOOTBALL, BASKETBALL        | C         |
| 11      | large   | M      | A        | A           | N       |                             |           |
| 12      | large   | F      | B        | A           | N       |                             |           |
| 13      | small   | M      | A        | A           | N       |                             |           |
| 14      | small   | M      | C        | B           | Y       | FOOTBALL, BADMINTON         | C         |
| 15      | small   | M      | B        | A           | N       |                             |           |
| 16      | small   | M      | C        | B           | Y       | FOOTBALL, BADMINTON         | C         |
| 17      | small   | M      | B        | A           | N       |                             |           |
| 18      | small   | F      | A        | A           | Y       | VOLLEYBALL                  | C         |
| 19      | small   | B      | B        | Y           | Y       | FOOTBALL BASKETBALL. TENNIS | C         |
| 20      | small   | M      | B        | A           | Y       |                             | C         |
| 21      | small   | M      | A        | B           | Y       | CRICKET                     | C         |
| 22      | small   | F      | A        | A           | Y       | BADMINTON                   | C         |
| 23      | small   | F      | A        | B           | Y       | TABLE TENNIS                | C         |
| 24      | small   | M      | B        | B           | Y       | FOOTBALL                    | C         |

co-games ----how often play computer games, A= 0, B= 1-4, C= 5 or more

VE-participation --- how often participate in VE experiment, A= 0, B= 1-4, C= 5 or more

sport-i - play sport or not ?

sport-ii- list of sport

sport-iii - A= professional, B=amateur, C=leisure, D= others

II. Room Size data

| Subject       | Display | height | width | length |
|---------------|---------|--------|-------|--------|
| 1             | Large   | 4.5    | 6     | 12     |
| 2             | Large   | 3      | 6     | 10     |
| 3             | Large   | 3      | 6     | 10     |
| 4             | Large   | 2.3    | 4.5   | 8      |
| 5             | Large   | 2.5    | 5     | 9.5    |
| 6             | Large   | 8      | 15    | 20     |
| 7             | Large   | 1.5    | 4     | 6      |
| 8             | Large   | 2      | 3     | 5      |
| 9             | Large   | 2.2    | 4     | 9      |
| 10            | Large   | 5.5    | 6     | 11     |
| 11            | Large   | 3.5    | 10    | 16     |
| 12            | Large   | 5      | 7     | 11     |
| 13            | Small   | 2.5    | 5     | 10     |
| 14            | Small   | 3.5    | 8     | 13     |
| 15            | Small   | 3.2    | 4.6   | 7.6    |
| 16            | Small   | 3      | 5     | 10     |
| 17            | Small   | 2.5    | 4     | 10     |
| 18            | Small   | 4      | 6     | 24     |
| 19            | Small   | 3.5    | 7.5   | 15     |
| 20            | Small   | 3      | 7     | 14     |
| 21            | Small   | 3      | 5     | 10     |
| 22            | Small   | 4      | 10    | 15     |
| 23            | Small   | 6      | 10    | 20     |
| 24            | Small   | 3.5    | 15    | 20     |
| Actual length |         | 3.5    | 7.28  | 14.81  |

III. Spatial Memory data

| Device  |         | mouse      |     |     |          |     |     | trackball  |     |     |          |     |     |
|---------|---------|------------|-----|-----|----------|-----|-----|------------|-----|-----|----------|-----|-----|
| Mode    |         | drive mode |     |     | fly mode |     |     | drive mode |     |     | fly mode |     |     |
| subject | display | md         | mdp | mdm | mf       | mfp | mfm | td         | tdp | tdm | tf       | tfp | tfm |
| 1       | large   | 7          | 286 | 300 | 7        | 350 | 83  | 4          | 287 | 147 | 3        | 294 | 165 |
| 2       | large   | 2          | 224 | 140 | 8        | 659 | 110 | 8          | 674 | 53  | 1        | 541 | 226 |
| 3       | large   | 5          | 113 | 173 | 8        | 155 | 100 | 2          | 119 | 138 | 5        | 99  | 102 |
| 4       | large   | 5          | 186 | 271 | 2        | 565 | 331 | 4          | 346 | 130 | 7        | 520 | 129 |
| 5       | large   | 6          | 93  | 96  | 6        | 204 | 130 | 6          | 71  | 109 | 8        | 157 | 83  |
| 6       | large   | 2          | 95  | 99  | 3        | 182 | 198 | 4          | 82  | 117 | 5        | 132 | 62  |
| 7       | large   | 7          | 80  | 90  | 9        | 110 | 99  | 4          | 784 | 115 | 9        | 445 | 119 |
| 8       | large   | 5          | 68  | 73  | 4        | 86  | 81  | 2          | 135 | 124 | 5        | 199 | 71  |
| 9       | large   | 7          | 110 | 150 | 4        | 112 | 136 | 2          | 195 | 250 | 8        | 395 | 117 |
| 10      | large   | 7          | 38  | 93  | 5        | 37  | 70  | 5          | 46  | 93  | 5        | 182 | 161 |
| 11      | large   | 8          | 35  | 57  | 8        | 61  | 65  | 6          | 83  | 163 | 6        | 533 | 122 |
| 12      | large   | 8          | 33  | 106 | 2        | 35  | 88  | 5          | 66  | 99  | 7        | 392 | 173 |
| 13      | small   | 3          | 60  | 176 | 7        | 107 | 206 | 6          | 84  | 124 | 6        | 174 | 186 |
| 14      | small   | 7          | 115 | 235 | 8        | 67  | 100 | 5          | 190 | 212 | 7        | 266 | 162 |
| 15      | small   | 3          | 90  | 88  | 8        | 36  | 67  | 7          | 46  | 60  | 8        | 50  | 56  |

|    |       |   |     |     |   |     |     |   |     |     |   |     |     |
|----|-------|---|-----|-----|---|-----|-----|---|-----|-----|---|-----|-----|
| 16 | small | 9 | 158 | 56  | 6 | 283 | 124 | 7 | 121 | 59  | 8 | 158 | 92  |
| 17 | small | 6 | 38  | 149 | 9 | 70  | 110 | 6 | 28  | 121 | 6 | 52  | 112 |
| 18 | small | 7 | 46  | 176 | 9 | 246 | 116 | 6 | 232 | 96  | 9 | 845 | 138 |
| 19 | small | 6 | 102 | 118 | 7 | 73  | 93  | 3 | 390 | 130 | 8 | 249 | 90  |
| 20 | small | 5 | 66  | 82  | 8 | 98  | 69  | 1 | 230 | 79  | 4 | 95  | 91  |
| 21 | small | 8 | 107 | 174 | 7 | 172 | 122 | 5 | 186 | 205 | 8 | 308 | 220 |
| 22 | small | 8 | 40  | 140 | 3 | 45  | 96  | 4 | 84  | 132 | 4 | 440 | 107 |
| 23 | small | 8 | 171 | 67  | 6 | 267 | 176 | 8 | 315 | 130 | 1 | 375 | 285 |
| 24 | small | 8 | 114 | 140 | 6 | 232 | 78  | 4 | 250 | 127 | 4 | 602 | 121 |

IV. Questionnaire data

a. Interface device questionnaire

Familiarity with interface device

| subject | mouse | trackball |
|---------|-------|-----------|
| 1       | 1     | 4         |
| 2       | 1     | 5         |
| 3       | 1     | 4         |
| 4       | 1     | 5         |
| 5       | 1     | 4         |
| 6       | 1     | 4         |
| 7       | 1     | 5         |
| 8       | 1     | 4         |
| 9       | 1     | 4         |
| 10      | 1     | 5         |
| 11      | 1     | 4         |
| 12      | 1     | 5         |
| 13      | 1     | 5         |
| 14      | 1     | 4         |
| 15      | 1     | 5         |
| 16      | 1     | 4         |
| 17      | 1     | 4         |
| 18      | 1     | 4         |
| 19      | 1     | 4         |
| 20      | 1     | 5         |
| 21      | 1     | 5         |
| 22      | 1     | 5         |
| 23      | 2     | 4         |
| 24      | 1     | 4         |

- 1= once a day
- 2= once a week
- 3= once a month
- 4=hardly used
- 5=never used

Question 2 –travel mode

|         | M |   | T |   | M |   | T |   | M |   | T |   | M |   | T |   |
|---------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| subject | D | F | D | F | D | F | D | F | D | F | D | F | D | F | D | F |
| 1       | 3 | 4 | 5 | 7 | 6 | 3 | 3 | 2 | 2 | 1 | 6 | 7 | 5 | 6 | 3 | 1 |
| 2       | 2 | 5 | 3 | 6 | 5 | 4 | 5 | 3 | 2 | 4 | 2 | 5 | 5 | 4 | 5 | 4 |
| 3       | 4 | 3 | 5 | 6 | 4 | 5 | 5 | 4 | 4 | 2 | 4 | 3 | 4 | 6 | 3 | 2 |
| 4       | 2 | 7 | 2 | 6 | 6 | 3 | 6 | 3 | 3 | 4 | 2 | 2 | 7 | 3 | 7 | 5 |
| 5       | 2 | 3 | 4 | 6 | 4 | 6 | 3 | 2 | 3 | 2 | 5 | 6 | 6 | 6 | 3 | 1 |
| 6       | 3 | 4 | 4 | 6 | 6 | 4 | 3 | 1 | 2 | 4 | 5 | 6 | 7 | 1 | 5 | 1 |
| 7       | 1 | 2 | 4 | 6 | 6 | 5 | 4 | 5 | 5 | 2 | 6 | 3 | 7 | 5 | 4 | 6 |
| 8       | 5 | 5 | 4 | 4 | 6 | 6 | 5 | 4 | 4 | 5 | 4 | 5 | 5 | 6 | 5 | 6 |
| 9       | 1 | 1 | 3 | 5 | 7 | 5 | 3 | 1 | 4 | 2 | 4 | 2 | 6 | 3 | 6 | 2 |
| 10      | 1 | 2 | 2 | 3 | 7 | 6 | 5 | 6 | 3 | 2 | 4 | 3 | 7 | 5 | 5 | 3 |
| 11      | 2 | 2 | 6 | 7 | 7 | 6 | 5 | 3 | 2 | 1 | 3 | 4 | 6 | 7 | 3 | 4 |
| 12      | 2 | 3 | 5 | 5 | 2 | 2 | 4 | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 3 | 3 |
| 13      | 4 | 6 | 5 | 7 | 5 | 2 | 4 | 1 | 4 | 2 | 6 | 4 | 6 | 5 | 2 | 1 |
| 14      | 2 | 3 | 7 | 7 | 6 | 6 | 2 | 2 | 2 | 2 | 6 | 6 | 7 | 6 | 2 | 2 |
| 15      | 2 | 5 | 3 | 6 | 6 | 4 | 5 | 2 | 4 | 2 | 5 | 3 | 8 | 7 | 2 | 1 |
| 16      | 1 | 2 | 3 | 3 | 6 | 5 | 5 | 5 | 1 | 2 | 2 | 3 | 6 | 5 | 5 | 5 |
| 17      | 1 | 4 | 1 | 3 | 7 | 4 | 7 | 5 | 5 | 5 | 5 | 5 | 4 | 2 | 5 | 2 |
| 18      | 2 | 7 | 4 | 7 | 6 | 2 | 6 | 1 | 1 | 4 | 6 | 7 | 6 | 1 | 2 | 1 |
| 19      | 3 | 2 | 6 | 5 | 5 | 6 | 2 | 2 | 3 | 1 | 6 | 4 | 4 | 7 | 1 | 1 |
| 20      | 2 | 5 | 6 | 2 | 6 | 3 | 3 | 6 | 2 | 4 | 5 | 2 | 8 | 4 | 2 | 6 |
| 21      | 3 | 6 | 5 | 7 | 4 | 4 | 6 | 6 | 4 | 3 | 6 | 5 | 3 | 3 | 6 | 6 |
| 22      | 3 | 4 | 5 | 6 | 6 | 7 | 4 | 2 | 5 | 3 | 4 | 6 | 6 | 5 | 4 | 2 |
| 23      | 2 | 4 | 4 | 5 | 5 | 3 | 4 | 2 | 4 | 2 | 4 | 5 | 5 | 3 | 2 | 1 |
| 24      | 2 | 6 | 3 | 7 | 2 | 5 | 3 | 6 | 2 | 4 | 4 | 5 | 7 | 4 | 5 | 4 |



Question 3- interface device

|         | 3i |   | 3ii |   | 3iii |   | 3iv |   | 3v |   |
|---------|----|---|-----|---|------|---|-----|---|----|---|
| subject | M  | T | M   | T | M    | T | M   | T | M  | T |
| 1       | 1  | 6 | 1   | 5 | 7    | 1 | 1   | 6 | 7  | 1 |
| 2       | 2  | 4 | 4   | 4 | 5    | 4 | 2   | 4 | 5  | 4 |
| 3       | 2  | 6 | 3   | 6 | 6    | 3 | 2   | 3 | 6  | 2 |
| 4       | 2  | 5 | 2   | 5 | 5    | 2 | 2   | 2 | 7  | 3 |
| 5       | 1  | 4 | 2   | 6 | 6    | 2 | 2   | 4 | 6  | 2 |
| 6       | 2  | 2 | 2   | 2 | 4    | 3 | 2   | 3 | 7  | 3 |
| 7       | 2  | 4 | 3   | 5 | 2    | 4 | 4   | 4 | 6  | 3 |
| 8       | 5  | 5 | 5   | 5 | 6    | 5 | 4   | 5 | 4  | 5 |
| 9       | 1  | 5 | 1   | 5 | 6    | 3 | 2   | 5 | 7  | 2 |
| 10      | 1  | 3 | 2   | 3 | 2    | 4 | 2   | 3 | 7  | 5 |
| 11      | 1  | 6 | 2   | 7 | 7    | 2 | 2   | 3 | 7  | 3 |
| 12      | 2  | 4 | 4   | 6 | 6    | 3 | 3   | 3 | 5  | 3 |
| 13      | 2  | 5 | 5   | 5 | 4    | 5 | 3   | 4 | 6  | 2 |
| 14      | 2  | 6 | 2   | 6 | 6    | 2 | 2   | 6 | 6  | 1 |
| 15      | 1  | 4 | 2   | 5 | 1    | 5 | 1   | 3 | 7  | 3 |
| 16      | 2  | 3 | 4   | 3 | 6    | 5 | 5   | 4 | 5  | 6 |
| 17      | 1  | 1 | 3   | 3 | 4    | 3 | 5   | 5 | 6  | 5 |
| 18      | 1  | 6 | 1   | 7 | 6    | 1 | 2   | 7 | 7  | 1 |
| 19      | 1  | 7 | 2   | 6 | 7    | 1 | 1   | 6 | 7  | 1 |
| 20      | 6  | 5 | 6   | 5 | 6    | 5 | 3   | 2 | 6  | 5 |
| 21      | 2  | 5 | 4   | 6 | 6    | 4 | 2   | 5 | 2  | 5 |
| 22      | 2  | 6 | 2   | 6 | 6    | 2 | 4   | 4 | 6  | 3 |
| 23      | 3  | 5 | 3   | 5 | 3    | 5 | 3   | 5 | 5  | 2 |
| 24      | 2  | 4 | 2   | 5 | 6    | 3 | 6   | 4 | 6  | 4 |

Question 4 – recall accuracy and Question 5 - familiarity with room

Question 4 - recall accuracy

| subject | mouse | trackball |
|---------|-------|-----------|
| 1       | 5     | 3         |
| 2       | 5     | 4         |
| 3       | 4     | 4         |
| 4       | 4     | 7         |
| 5       | 4     | 3         |
| 6       | 5     | 3         |
| 7       | 4     | 4         |
| 8       | 4     | 4         |
| 9       | 3     | 3         |
| 10      | 6     | 5         |
| 11      | 5     | 3         |
| 12      | 3     | 3         |
| 13      | 3     | 2         |
| 14      | 6     | 4         |
| 15      | 5     | 4         |
| 16      | 5     | 5         |
| 17      | 5     | 5         |
| 18      | 5     | 1         |
| 19      | 5     | 3         |

Question 5 - familiarity with environment

| subject | Recognize | helpful |
|---------|-----------|---------|
| 1       | Y         | 7       |
| 2       | N         | 0       |
| 3       | N         | 0       |
| 4       | Y         | 7       |
| 5       | N         | 0       |
| 6       | Y         | 4       |
| 7       | N         | 0       |
| 8       | N         | 0       |
| 9       | N         | 0       |
| 10      | N         | 0       |
| 11      | N         | 0       |
| 12      | N         | 0       |
| 13      | N         | 0       |
| 14      | Y         | 6       |
| 15      | N         | 0       |
| 16      | Y         | 1       |
| 17      | N         | 0       |
| 18      | Y         | 6       |
| 19      | N         | 0       |

|    |   |   |    |   |   |
|----|---|---|----|---|---|
| 20 | 5 | 5 | 20 | Y | 6 |
| 21 | 4 | 5 | 21 | N | 0 |
| 22 | 3 | 3 | 22 | Y | 5 |
| 23 | 4 | 2 | 23 | Y | 2 |
| 24 | 6 | 4 | 24 | Y | 5 |

b. Participants' comments

Note: Missing participants' number implies no comment was given

| Subject | Comments                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|---------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1       | Trackball in fly mode - less control (sense of control - finger tips) - you could not estimate movement<br>fast moving when you have less control your concentration will more on to control rather than the placement of objects, therefore disturb capacity to recall<br>mouse in fly mode - very useful -fly mode - you can see everything from above - your overview of the whole room and that recall better object position<br>mouse in drive mode - obstacle - we eye level - walls , curtains, view span limited, therefore placement of objects not that accurate, prefer more to trackball<br>trackball in drive - more control - slower - view span forward - cannot see from above - predicts position of object from distance |
| 2       | I think the speed of fly mode is fast so that I almost cannot control the movement. I've never used the trackball. So if good control, it is necessary to give train to use trackball. In addition because of the fast movement, I can't remember the position of the object for the first time because i am not familiar with the environment. so I can't recall the location of each object. I think being familiar with the environment is important for recalling the location of each object<br>fast movement makes user tired and it also affect the recall of location                                                                                                                                                              |
| 3       | Didn't feel that the interface type help determine object recall, just that the trackball was harder to use than the mouse                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 4       | Drive environment is easy to control, fly environment is difficult. Mouse is easy to use . Trackball is difficult. The first recall is not clear and then the objects are are easy to remember. Added rolling is difficult in trackball to remember which button to use. Moving the mouse can feel it translate the move in picture, but when using trackball, the rolling ball doesn't feel the translate movement in the picture because the movement is only the tip of the finger                                                                                                                                                                                                                                                      |
| 5       | later tasks were easier as I became mor practiced t recalling positions of objects                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 6       | The speed of the movement is fast that makes my eyes feel tired                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 7       | As I am used to the mouse, I find it easy to use. I I have been using the trackball it would have been easier too. I find it difficult to control the flying mode                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 8       | drive mode - easy to move forward and back ward. Fly mode - good at the adjustment of height so that I can see the objects position clearer, a bit dizzy trackball a bit hard for new user but it is more convenient to move in 3-d spatial environment. After practice it became easier to use. Quicker time to use to I have more time to remember the location of objects                                                                                                                                                                                                                                                                                                                                                               |
| 9       | The acceleration and movement is deferent t a lot of modern game, so navigation was more difficult at first. To help with the placement of objects I used distinctive points in the room such as the curtains and the door, placing objects relative to this points                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 10      | Assuming that I did better in the recalling test using the mouse, I think this is due to the fact that I could control the movement of the mouse easily and locate objects better. This gave me the opportunity to explore the environment more efficiently than using a trackball                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 11      | The mouse interface may have an added advantage because the user is already familiar with the experiment, having done it with trackball first time                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

| Subject                               | Comments                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|---------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| N<br>o<br>t                           | When I tried to back to the wall, I could not do that. It was a little bit difficult to recognize object with the texture of the floor. It was very smooth and natural feel when I turned around in the room                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| E<br>:<br>M                           | I don't think my results are accurate because I don't play computer games which you need to be in control of the mouse movement                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 14<br>s<br>s<br>i<br>n<br>g           | Mouse/drive easy to control forward and backward movement. Mouse fly easy to look at object from top view<br>trackball drive - hard to control the device to move the desired direction. Trackball fly - difficult to get the accurate angle for each object. Trackball drive /fly can't get accurate position of object. Mouse/drive easy to control and aim at the position of object. Mouse is easy to use because just use two fingers to control the device. trackball need to use three fingers and your mind too                                                                                                                                                                                                   |
| 15<br>p<br>a<br>r<br>t<br>i<br>c<br>i | I felt less prepared for the first trial and spent a majority of the time identifying which item is which I feel this may affect the first 'map' I drew. I personally felt that using a mouse was easier then the trackball for the sole reason I have used a mouse for along time for both games and work. because I had never used a trackball before simply 'getting a feel' for the device was a challenge in itself. Regarding flying/driving I felt that although I'd prefer to drive, the vertical movement allowed me to position the view in such a way that I could see a large proportion of the room and the use that static view to memorize the contents location without worrying about the control device |
| 16<br>a<br>n<br>t<br>s                | mouse/drive easy because maintain same height. Mouse fly need practice. Trackball need to use at least three fingers. Mouse drive easy to control. Easy to stay in one position. Trackball difficult to stay in one position. Mouse easy to control but larger space would be better. .Trackball is better when the space is small                                                                                                                                                                                                                                                                                                                                                                                        |
| 17<br>n<br>u<br>m                     | drive - no need to worry about up and own . Fly - very hard to stay at same level. Trackball - the fact that your hand don't move makes control easier. Mouse drive is easy but cannot approach object closely. Mouse fly - it harder to move around but it easy to get close to objects. Trackball control is nicer. Trackball flying is irritating. Mode of travel affect more the object position rather than interface device                                                                                                                                                                                                                                                                                         |
| 18<br>b                               | I prefer to use the mouse to control my movements                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 20<br>r                               | Even though, it's not easy to recall object when using mouse, but this is under my control and I can use more time to remember the exact location of the objects.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 21                                    | Even though, it's not easy to recall object when using mouse, but this is under my control and I can use more time to remember the exact location of the objects.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

## C.6.2 VE CONDITION DATA –EXPERIMENT 3B

## I. Participants' Information

| subject | age | gender | co-games | VE-participation | sport-I | sport-II                        | sport-III |
|---------|-----|--------|----------|------------------|---------|---------------------------------|-----------|
| 1       | 20  | M      | C        | A                | Y       | FOOTBALL, BASKETBALL            | C         |
| 2       | 54  | M      | A        | B                | N       |                                 |           |
| 3       | 19  | M      | B        | A                | Y       | TENNIS, TABLE TENNIS, BADMINTON | C         |
| 4       | 33  | M      | C        | A                | Y       | GOLF, BOWLING                   | C         |
| 5       | 27  | F      | A        | A                | Y       | BADMINTON, VOLLEYBALL           | C         |
| 6       | 19  | M      | B        | A                | Y       | TABLE TENNIS, BADMINTON         | C         |
| 7       | 23  | M      | C        | A                | Y       | FOOTBALL, BADMINTON             | C         |
| 8       | 28  | M      | B        | A                | N       |                                 |           |

co-games —how often play computer games, A= 0, B= 1-4, C= 5 or more

VE-participation — how often participate in VE experiment, A= 0, B= 1-4, C= 5 or more

sport-I - play sport or not ?

sport-II- list of sport

sport-III - A= professional, B=amateur, C=leisure, D= others

## II. Room size data

| subject | large  |       |        | small  |       |        |
|---------|--------|-------|--------|--------|-------|--------|
|         | Height | Width | Length | Height | Width | Length |
| 1       | 3      | 10    | 25     | 3      | 9     | 15     |
| 2       | 4      | 8     | 15     | 4      | 8     | 15     |
| 3       | 2.4    | 13    | 21     | 3.5    | 11    | 22     |
| 4       | 2.4    | 5.5   | 17     | 2.4    | 9.5   | 30     |
| 5       | 4      | 10    | 14     | 2.5    | 7     | 12     |
| 6       | 3.2    | 5     | 7.5    | 4.00   | 8     | 16     |
| 7       | 4      | 5.5   | 11     | 4      | 5     | 7      |
| 8       | 3.2    | 6.5   | 11     | 2.9    | 6     | 11     |

## III. Spatial memory data

| Subject | LARGE |      |      |     |      |      |           |      |      |     |      |      |
|---------|-------|------|------|-----|------|------|-----------|------|------|-----|------|------|
|         | MOUSE |      |      |     |      |      | TRACKBALL |      |      |     |      |      |
|         | LMD   | LMDp | LMDm | LMF | LMFp | LMFm | LTD       | LTDp | LTDm | LTF | LTFp | LTFm |
| 1       | 4     | 127  | 358  | 7   | 108  | 245  | 6         | 174  | 180  | 9   | 94   | 63   |
| 2       | 5     | 132  | 75   | 3   | 580  | 130  | 2         | 263  | 110  | 8   | 263  | 57   |
| 3       | 7     | 38   | 69   | 9   | 37   | 91   | 2         | 59   | 168  | 6   | 105  | 150  |
| 4       | 8     | 68   | 117  | 8   | 132  | 190  | 5         | 113  | 208  | 5   | 260  | 151  |
| 5       | 8     | 68   | 75   | 8   | 93   | 64   | 5         | 100  | 69   | 8   | 72   | 72   |
| 6       | 4     | 27   | 76   | 8   | 94   | 228  | 7         | 60   | 72   | 6   | 60   | 166  |
| 7       | 7     | 17   | 80   | 9   | 18   | 100  | 9         | 41   | 95   | 4   | 31   | 104  |
| 8       | 9     | 102  | 220  | 7   | 117  | 190  | 9         | 136  | 230  | 9   | 108  | 197  |

LMD – Small/Mouse/Drive  
LMDp - Small/Mouse/Drive practice time  
LMDm - Small/Mouse/Drive map test time

LMF – Small/Mouse/ Fly  
LMFp - Small/Mouse/ Fly practice time  
LMFm - Small/Mouse/ Fly map test time

LTD – Small/Trackball/Drive  
LTDp - Small/Trackball/Drive practice time  
LTDm - Small/Trackball/Drive map test time

LTF – Small/ Trackball / Fly  
LTFp - Small/ Trackball / Fly practice time  
LTFm - Small/ Trackball / Fly map test time

| Subject | SMALL |      |      |     |      |      |           |      |      |     |      |      |
|---------|-------|------|------|-----|------|------|-----------|------|------|-----|------|------|
|         | MOUSE |      |      |     |      |      | TRACKBALL |      |      |     |      |      |
|         | SMD   | SMDp | SMDm | SMF | SMFp | SMFm | STD       | STDp | STDm | STF | STFp | STFm |
| 1       | 8     | 54   | 148  | 9   | 100  | 92   | 1         | 75   | 220  | 5   | 81   | 130  |
| 2       | 8     | 89   | 88   | 4   | 500  | 110  | 6         | 111  | 108  | 8   | 125  | 108  |
| 3       | 9     | 30   | 76   | 6   | 28   | 80   | 7         | 30   | 63   | 6   | 38   | 78   |
| 4       | 8     | 120  | 160  | 7   | 50   | 125  | 8         | 80   | 185  | 9   | 220  | 90   |
| 5       | 6     | 183  | 148  | 7   | 341  | 132  | 2         | 110  | 127  | 7   | 520  | 60   |
| 6       | 4     | 80   | 185  | 7   | 220  | 90   | 5         | 120  | 160  | 4   | 50   | 125  |
| 7       | 8     | 30   | 82   | 8   | 57   | 118  | 7         | 117  | 176  | 7   | 108  | 116  |
| 8       | 5     | 169  | 160  | 9   | 180  | 180  | 9         | 171  | 114  | 6   | 185  | 181  |

SMD – Small/Mouse/Drive  
SMDp - Small/Mouse/Drive practice time  
SMDm - Small/Mouse/Drive map test time

SMF – Small/Mouse/ Fly  
SMFp - Small/Mouse/ Fly practice time  
SMFm - Small/Mouse/ Fly map test time

STD – Small/Trackball/Drive  
STDp - Small/Trackball/Drive practice time  
STDm - Small/Trackball/Drive map test time

STF – Small/ Trackball / Fly  
STFp - Small/ Trackball / Fly practice time  
STFm - Small/ Trackball / Fly map test time

IV. Questionnaire data

a. Display questionnaire

| subject | Qi    |       | Qii   |       | Qiii  |       |
|---------|-------|-------|-------|-------|-------|-------|
|         | large | small | large | small | large | small |
| 1       | 4     | 1     | 4     | 7     | 3     | 5     |
| 2       | 7     | 6     | 3     | 4     | 1     | 1     |
| 3       | 4     | 2     | 4     | 6     | 4     | 6     |
| 4       | 3     | 2     | 6     | 5     | 4     | 5     |
| 5       | 1     | 3     | 7     | 3     | 6     | 4     |
| 6       | 3     | 4     | 6     | 4     | 5     | 6     |
| 7       | 3     | 5     | 6     | 4     | 5     | 6     |
| 8       | 4     | 6     | 5     | 3     | 5     | 4     |

## b. Interface device questionnaire

## Familiarity with interface device

| subject | M | T |
|---------|---|---|
| 1       | 1 | 5 |
| 2       | 1 | 4 |
| 3       | 1 | 4 |
| 4       | 1 | 3 |
| 5       | 1 | 5 |
| 6       | 1 | 5 |
| 7       | 1 | 5 |
| 8       | 1 | 5 |

subject 1-4 is large-small condition  
 subject 5-8 is small- large condition

M - mouse  
 T - trackball

1= once a day  
 2= once a week  
 3= once a month  
 4=hardly used  
 5=never used

## Question 2 – Travel modes

---- M = mouse, T = trackball, D = drive mode, F = Fly mode

|         | Large |   |   |   |     |   |   |   |      |   |   |   |     |   |   |   |
|---------|-------|---|---|---|-----|---|---|---|------|---|---|---|-----|---|---|---|
|         | 2i    |   |   |   | 2ii |   |   |   | 2iii |   |   |   | 2iv |   |   |   |
|         | M     |   | T |   | M   |   | T |   | M    |   | T |   | M   |   | T |   |
| subject | D     | F | D | F | D   | F | D | F | D    | F | D | F | D   | F | D | F |
| 1       | 5     | 4 | 1 | 2 | 5   | 2 | 6 | 4 | 4    | 4 | 4 | 4 | 4   | 4 | 7 | 6 |
| 2       | 3     | 4 | 5 | 7 | 5   | 2 | 5 | 2 | 4    | 4 | 4 | 4 | 6   | 2 | 6 | 2 |
| 3       | 4     | 2 | 5 | 6 | 5   | 6 | 4 | 3 | 5    | 1 | 4 | 4 | 5   | 7 | 3 | 3 |
| 4       | 2     | 4 | 3 | 6 | 6   | 4 | 3 | 2 | 4    | 2 | 4 | 2 | 5   | 4 | 4 | 3 |
| 5       | 2     | 2 | 3 | 3 | 7   | 5 | 6 | 2 | 2    | 1 | 2 | 3 | 3   | 2 | 4 | 6 |
| 6       | 2     | 2 | 2 | 2 | 5   | 5 | 4 | 5 | 4    | 2 | 4 | 2 | 5   | 6 | 4 | 6 |
| 7       | 5     | 6 | 4 | 5 | 4   | 6 | 4 | 5 | 5    | 5 | 4 | 5 | 5   | 6 | 4 | 5 |
| 8       | 2     | 5 | 3 | 4 | 6   | 3 | 4 | 3 | 4    | 2 | 4 | 2 | 6   | 3 | 4 | 4 |

|         | Small |   |   |   |     |   |   |   |      |   |   |   |     |   |   |   |
|---------|-------|---|---|---|-----|---|---|---|------|---|---|---|-----|---|---|---|
|         | 2i    |   |   |   | 2ii |   |   |   | 2iii |   |   |   | 2iv |   |   |   |
|         | M     |   | T |   | M   |   | T |   | M    |   | T |   | M   |   | T |   |
| subject | D     | F | D | F | D   | F | D | F | D    | F | D | F | D   | F | D | F |
| 1       | 2     | 7 | 1 | 3 | 6   | 2 | 7 | 6 | 4    | 2 | 4 | 2 | 5   | 1 | 7 | 7 |
| 2       | 3     | 4 | 4 | 7 | 5   | 4 | 4 | 3 | 7    | 7 | 7 | 7 | 5   | 3 | 4 | 2 |
| 3       | 3     | 2 | 4 | 3 | 5   | 6 | 4 | 5 | 4    | 1 | 5 | 2 | 4   | 7 | 3 | 5 |
| 4       | 2     | 4 | 3 | 4 | 6   | 5 | 5 | 4 | 3    | 2 | 4 | 3 | 6   | 5 | 5 | 4 |
| 5       | 1     | 3 | 5 | 7 | 7   | 6 | 5 | 2 | 1    | 1 | 4 | 2 | 5   | 7 | 3 | 2 |
| 6       | 4     | 1 | 5 | 4 | 3   | 6 | 3 | 5 | 4    | 2 | 5 | 3 | 3   | 6 | 2 | 5 |
| 7       | 2     | 4 | 4 | 6 | 4   | 6 | 4 | 7 | 6    | 3 | 4 | 1 | 2   | 6 | 3 | 6 |
| 8       | 3     | 2 | 4 | 6 | 7   | 5 | 5 | 2 | 5    | 2 | 5 | 3 | 5   | 6 | 3 | 3 |



Question3 - interface device

| subject | Large |   |     |   |      |   |     |   |    |   | Small |   |     |   |      |   |     |   |    |   |
|---------|-------|---|-----|---|------|---|-----|---|----|---|-------|---|-----|---|------|---|-----|---|----|---|
|         | 3i    |   | 3ii |   | 3iii |   | 3iv |   | 3v |   | 3i    |   | 3ii |   | 3iii |   | 3iv |   | 3v |   |
|         | M     | T | M   | T | M    | T | M   | T | M  | T | M     | T | M   | T | M    | T | M   | T | M  | T |
| 1       | 1     | 1 | 5   | 2 | 3    | 6 | 4   | 4 | 6  | 7 | 1     | 1 | 2   | 1 | 4    | 6 | 4   | 4 | 7  | 7 |
| 2       | 4     | 1 | 4   | 1 | 4    | 7 | 4   | 1 | 4  | 7 | 2     | 5 | 3   | 5 | 2    | 4 | 7   | 7 | 7  | 4 |
| 3       | 1     | 4 | 1   | 4 | 6    | 3 | 1   | 4 | 7  | 3 | 1     | 3 | 1   | 3 | 6    | 4 | 1   | 3 | 7  | 4 |
| 4       | 2     | 4 | 2   | 3 | 6    | 5 | 2   | 3 | 6  | 5 | 2     | 3 | 2   | 3 | 6    | 5 | 2   | 3 | 6  | 5 |
| 5       | 1     | 5 | 2   | 5 | 6    | 5 | 1   | 4 | 7  | 3 | 1     | 4 | 1   | 3 | 7    | 4 | 2   | 4 | 7  | 4 |
| 6       | 1     | 2 | 2   | 2 | 5    | 4 | 2   | 2 | 6  | 4 | 1     | 3 | 1   | 3 | 6    | 4 | 2   | 4 | 7  | 3 |
| 7       | 5     | 4 | 6   | 6 | 5    | 5 | 5   | 5 | 6  | 4 | 2     | 4 | 5   | 3 | 3    | 6 | 4   | 3 | 5  | 4 |
| 8       | 2     | 5 | 3   | 4 | 6    | 3 | 3   | 4 | 6  | 4 | 2     | 6 | 2   | 4 | 5    | 2 | 3   | 4 | 6  | 3 |

Question 4 – Recall accuracy

| subject | large |   | small |   |
|---------|-------|---|-------|---|
|         | M     | T | M     | T |
| 1       | 5     | 5 | 4     | 4 |
| 2       | 4     | 7 | 1     | 1 |
| 3       | 5     | 3 | 5     | 4 |
| 4       | 5     | 4 | 5     | 5 |
| 5       | 6     | 4 | 6     | 4 |
| 6       | 4     | 6 | 5     | 3 |
| 7       | 5     | 4 | 5     | 6 |
| 8       | 5     | 4 | 4     | 3 |

Question 4 – Familiarity with environment

| subject | large |   | small |   |
|---------|-------|---|-------|---|
|         | R     | H | R     | H |
| 1       | 0     |   | 0     |   |
| 2       | 0     |   | 0     |   |
| 3       | 0     |   | 0     |   |
| 4       | 0     |   | 0     |   |
| 5       | 0     |   | 0     |   |
| 6       | 0     |   | 0     |   |
| 7       | 0     |   | 0     |   |
| 8       | 0     |   | 0     |   |

R =recognize  
H = helpful in recall  
7    HELPFUL  
     NOT  
1    HELPFUL

0 NOT FAMILIAR - do not recognize room  
1 Familiar -recognize room

## c. Participants' comments – display questionnaire

--A participants' number missing implies no comment was made

| Subject | Q   | Display | Comments                                                                                                                                                                                                |
|---------|-----|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1       | i   | Large   | Made the room too big to recall objects                                                                                                                                                                 |
|         |     | Small   | Easier to find out size of room, got used to small display                                                                                                                                              |
|         | ii  | Large   | as above                                                                                                                                                                                                |
|         |     | Small   | Used to look at small screen                                                                                                                                                                            |
|         | iii | Large   | Guessing some of the object location                                                                                                                                                                    |
|         |     | Small   | as screen is small, easier to recognize object location                                                                                                                                                 |
| 3       | i   | Large   | Not sure, not being used to the size                                                                                                                                                                    |
|         |     | Small   | Being used to a small screen I think                                                                                                                                                                    |
|         | ii  | Large   | Seems ok but prefer the smaller one                                                                                                                                                                     |
|         |     | Small   | Seemed the most natural                                                                                                                                                                                 |
|         | iii | Large   | ----                                                                                                                                                                                                    |
|         |     | Small   | -----                                                                                                                                                                                                   |
| 4       | i   | Large   | -----                                                                                                                                                                                                   |
|         |     | Small   | Small screen allows you to see all object without moving your position                                                                                                                                  |
|         | ii  | Large   | More immersive, less like a game, with small screen I was still aware of the edges in my peripheral vision, with big screen I found I was drawn into the environment and less aware of the surroundings |
|         |     | Small   |                                                                                                                                                                                                         |
|         | iii | Large   |                                                                                                                                                                                                         |
|         |     | Small   |                                                                                                                                                                                                         |
| 5       | i   | Large   | Got a better look at the room                                                                                                                                                                           |
|         |     | Small   | Bit too compact                                                                                                                                                                                         |
|         | ii  | Large   | Gives better perspective                                                                                                                                                                                |
|         |     | Small   | Does seem real                                                                                                                                                                                          |
|         | iii | Large   | --                                                                                                                                                                                                      |
|         |     | Small   | --                                                                                                                                                                                                      |
| 7       | i   | Large   | Image is clearer                                                                                                                                                                                        |
|         |     | Small   | --                                                                                                                                                                                                      |
|         | ii  | Large   | Clearer, better perception                                                                                                                                                                              |
|         |     | Small   | ---                                                                                                                                                                                                     |
|         | iii | Large   | Fairly accurate                                                                                                                                                                                         |
|         |     | Small   | Fairly accurate                                                                                                                                                                                         |
| 8       | i   | Large   | Easier to see larger objects- more time spent looking at locations                                                                                                                                      |
|         |     | Small   | --                                                                                                                                                                                                      |
|         | ii  | Large   | --                                                                                                                                                                                                      |
|         |     | Small   | --                                                                                                                                                                                                      |
|         | iii | Large   | --                                                                                                                                                                                                      |
|         |     | Small   | --                                                                                                                                                                                                      |

## d. Participants' comments – Device questionnaire

--A subject number missing implies no comment was made

**Question 2**

| Subject | Q   | Device | Mode  | Comments                                            |
|---------|-----|--------|-------|-----------------------------------------------------|
| 1       | i   | Mouse  | Drive | Easier to control with mouse                        |
|         |     |        | Fly   | Very difficult to get used to it with mouse control |
|         |     | T/ball | Drive | Easy to control using trackball                     |
|         |     |        | Fly   | Easier to use compare to mouse                      |
|         | ii  | Mouse  | Drive | mouse more interact with program                    |
|         |     |        | Fly   | Control not very good with program                  |
|         |     | T/ball | Drive | More interact with program and accurate control     |
|         |     |        | Fly   | As above                                            |
|         | iii | Mouse  | Drive | Less view to look at the room                       |
|         |     |        | Fly   | more different angle /point of view                 |
|         |     | T/ball | Drive | Less view to look at the room                       |
|         |     |        | Fly   | More different angle/point of view                  |
| 3       | iv  | Mouse  | Drive | Easy to get used to on controlling                  |
|         |     |        | Fly   | Hard to control                                     |
|         |     | T/ball | Drive | Easy control                                        |
|         |     |        | Fly   | Easy control, more view to look at                  |
|         | i   | Mouse  | Drive | Easy to use                                         |
|         |     |        | Fly   | Easy to use and can go anywhere                     |
|         |     | T/ball | Drive | Easy but occasionally fiddly                        |
|         |     |        | Fly   | As above                                            |
|         | ii  | Mouse  | Drive | Very natural to use                                 |
|         |     |        | Fly   | Very natural like a flight simulator                |
|         |     | T/ball | Drive | Quite controllable when you get used to it          |
|         |     |        | Fly   | Good but easy to get muddled                        |
|         | iii | Mouse  | Drive | Quite easy but hard to see the floor                |
|         |     |        | Fly   | Very easy to get a good vintage spot                |
|         |     | T/ball | Drive | Ok but again hard to see the objects sometimes      |
|         |     |        | Fly   | Easy once a good bird's eye view is found           |
|         | iv  | Mouse  | Drive | Easy but looks maneuverability                      |
|         |     |        | Fly   | Easiest and most natural to use                     |
|         |     | T/ball | Drive | Not as natural as mouse as good as fly mode         |
|         |     |        | Fly   | Good but a mouse is preferred                       |

**Note:** T/ball = trackball

**Question 3**

| Subject | Q   | Device    | Comments                                                                                   |
|---------|-----|-----------|--------------------------------------------------------------------------------------------|
| 1       | i   | Mouse     | Use mouse everyday                                                                         |
|         |     | Trackball | Interact well between user and computer                                                    |
|         | ii  | Mouse     | It is easy but the interaction with program is not as good as trackball                    |
|         |     | Trackball | Very easy to locate/position as the interaction between trackball and program is very good |
|         | iii | Mouse     | Not as easy as trackball                                                                   |
|         |     | Trackball | Easy to control                                                                            |
|         | iv  | Mouse     | No difference between mouse and trackball                                                  |
|         |     | Trackball | --                                                                                         |
|         | v   | Mouse     | Got used to mouses                                                                         |
|         |     | Trackball | Just as good as mouse or may be better on this experiment                                  |
| 3       | i   | Mouse     | Easy due to lots of experience                                                             |
|         |     | Trackball | Not hard but not much experience                                                           |
|         | ii  | Mouse     | Very intuitive and natural                                                                 |
|         |     | Trackball | A fiddly sometimes                                                                         |
|         | iii | Mouse     | Most control due to experience                                                             |
|         |     | Trackball | Not bad but mouse is easier                                                                |
|         | iv  | Mouse     | Can concentrate on objects and not on mouse                                                |
|         |     | Trackball | Occasionally distracted by the interface being fiddly                                      |
|         | v   | Mouse     | I am most used to it                                                                       |
|         |     | Trackball | More interesting perhaps but less easy                                                     |
| 3       | i   | Mouse     | Because of years of experience                                                             |
|         |     | Trackball | Lack of experience but not too hard                                                        |
|         | ii  | Mouse     | Like using a flight sim                                                                    |
|         |     | Trackball | Quite weird due to using the thumb                                                         |
|         | iii | Mouse     | Very intuitive                                                                             |
|         |     | Trackball | Not very natural to use if new to the idea                                                 |
|         | iv  | Mouse     | Very easy to position yourself                                                             |
|         |     | Trackball | Quite hard to get in the right position                                                    |
|         | v   | Mouse     | Have used them before                                                                      |
|         |     | Trackball | Not much experience with them                                                              |

Additional comments by participants

| Subject | Comments                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|---------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2       | I prefer not to travel but rather rotate whilst not moving then I can look at the objects better. Actually fly was better for this because I could see all objects by rotating. With drive I had to travel and rotate to see all objects                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 4       | Better coordination with trackball today, mouse better still. Become generally easier with practice. Again in fly mode looking yourself I atop corner allows the best viewing angle drive mode give you poor viewing close up because of only 2 dimension of freedom. Trackball, I found a little oversensitive and was finding myself compensating my movement. With mouse this was less the case probably due to better familiarity with the drive. Flying is generally more difficult with the trackball because of the above reasons, but ultimately the better tool for completing the task. I guess part of this is due to the fact that most people are use to controlling machinery in 2 dimension that is a car forward, backward, left and right, but when you add the up and down 3 <sup>rd</sup> dimension it makes things difficult |

## APPENDIX D

### PUBLICATIONS

1. Awang Rambli, D. R. and R. S. Kalawsky (2003). Effects of varying display size on user's asymmetrical distance perception in the real and virtual environment. Virtual Concept 2003, Biarritz-France.
2. Awang Rambli, D. R. and R. S. Kalawsky. (2002). The Effect of Display and Image Type on Inter-Object Distance Estimation in Virtual and Real Environment. SIGCHI-NZ Symposium on Computer-Human Interaction, Hamilton, New Zealand.



# EFFECTS OF VARYING DISPLAY SIZE ON USER'S ASYMMETRICAL DISTANCE PERCEPTION IN THE REAL AND VIRTUAL ENVIRONMENT

D.R.Awang Rambli <sup>1</sup>, R.S.Kalawsky <sup>2</sup>

(1) : Department of Computer Science  
Loughborough University, LE11 3TU,UK  
+44(0)1509222097/+44(0)1509211586  
E-mail : D.R.Awang-Rambli@lboro.ac.uk

(2) : Advanced Virtual Reality Research Centre  
Loughborough University, LE11 3TU,UK  
+44(0)1509222097/+44(0)1509211586  
E-mail : R.S.Kalawsky@lboro.ac.uk

**Abstract:** Recent investigations into perception in the Virtual Environment (VE) have suggested display as one of the probable cause of perceptual difference between the real and virtual environment, in particular with respect to distance perception. In this paper, we report a study that investigates user's perception of asymmetrical distances in the real and VE presented on display of varying sizes. Video images of the real world and of its virtual model were used to represent the real and VE. A monoscopic viewing of the image was employed to eliminate stereo-acuity problems. Restricting participants' head and body movements reduced the effects of motion-parallax cues. Other variables controlled include display FOV and resolution and the viewing conditions. The asymmetrical distances estimated were vertical, horizontal and transverse distance. Our results showed the differences in distance perception between real and VE were small. Vertical distance estimations were better than horizontal and transverse distance in both real and VE; a main effect of distance was revealed. On the average, participants' performances were better on small display compared to large display. Findings from this study would have implication for applications that require spatial judgment tasks; the choice of display size might have an impact on users' performance.

**Key words:** Virtual environment, distance perception, asymmetrical distances, projected display, video movies.

## 1- Introduction

VE are computer-generated environments typically designed to represent and provide experience of places or locations in a real world or even a non-existent world. The success of applications that use VE to represent its real-world counterpart depends on VE technologies to provide similar perception and experience in both worlds. Users must be allowed to perceived spatial relations in the VE in the equivalent way as they would in the real world. Spatial awareness refers to our awareness of

the elements within an environment. It includes knowledge and understanding of object locations and relative position in the 3D space. Spatial knowledge in VE is often evaluated in VE using performance measures that include distance estimation [1]. Essentially, the knowledge of distance between objects forms the basics of our understanding of the physical structure [2]. While some researchers reported an overestimation [1][3], generally distance perception in VE has been found to be underestimated [4] [5] [6]. The reason for these differences in distance perception in the VE are still unknown [6].

The display system has been suggested as one of the probable cause of distance underestimation in VE [6]. In their studies comparing perceived egocentric distances in three types of environment (real environment, stereoscopic photographic panorama, and virtual stereoscopic computer model), Willemson & Gooch found small difference between the photographic panorama VE and the computer model VE, leading them to suggest that the display device plays a role in affecting the distance judgment in VE [6]. Roscoe suggested that the basic problem with all computer animated, sensor-generated, and optically generated display is that they produce systematic errors in size and distance judgments [7]. He concluded that spatial information on computer display requires modification for it to appear normal. Most studies on display aspects of VE had focused on comparing spatial performance on various display types: desktop monitor & Head mounted display (HMD) with tracked and non tracked condition [5]; HMD & desktop monitor [3][8][9]; HMD, desktop monitor and projected display [10]. Few studies have explored the effects of display size on spatial performance [11]. Recent investigations have reported better subject's performance in VE presented on larger display [10][11][12]. Several variations of spatial task have been investigated in these studies: orientation, mental rotations, navigation and distance estimation. The aim of the current study is to examine the effect of varying display size

but similar visual angle on participants' distance perception performance in the video images of real and VE. The focus of past studies has been on perception of distance between observer and objects (egocentric distance), while few studies have examined exocentric distance (distance between objects or points). Our study investigated the later distance by specifically examining the following asymmetrical distances: vertical, horizontal and transverse. This experiment was designed to extend the investigations of our initial study that compare distance perception in still images of real and VE. It is also part of a series of experiments in our research to investigate similarity of user's perception of the VE to the real world.

## 2- Experiment

The goal of the study is to focus on users' perception of distance in real and virtual environments: is distance perceived in a video of real environment similar to distance perceived in a video its virtual model? Is there any performance difference when these images are presented on different display size? Specifically, this study explored the effect of varying display size on participants' distance estimation task performance in the video images of real and virtual in a more controlled manner. The field of view (FOV) and the display resolutions of both displays were equated. Movement path through both the video and VE model were set to be similar and predefined. Stereo-acuity problems that might be experienced by users were eliminated by allowing monoscopic viewing of image. The effect of motion parallax cues (due to head movement) was reduced by requiring the participants to fix their head and body movement during the study. Participants were asked to estimate asymmetrical distances (vertical, horizontal and transverse distance) from a video movie of a real scene and a video movie of the simulated VE. Asymmetrical distance in this study refers to the following type of distances: vertical, horizontal and transverse distance. Vertical relates to heights of objects, or vertical extent in a scene. Horizontal (lateral) distance involves distance on a horizontal plane, while transverse distance is the distance going into the horizon, that is forward distance into the picture. These distances are necessary for the perception of space and layout of a VE. Figure 1 illustrates the three types of distances measured.

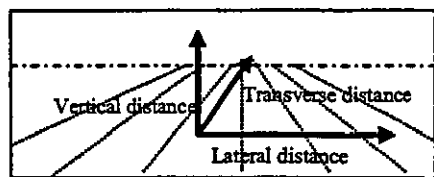


Fig. 1: Vertical, lateral and transverse distance (Adapted from Awang-Rambli & Kalawsky [12])

## 3- Methodology

### 3.1 - Participants

Forty volunteers, comprising staff and students (25 males and 15 females), participated in the study. Participants' age range

from 23 to 50 years with an average of 36.15. All participants have normal or normal corrected vision.

### 3.2 - Material and Apparatus

#### 3.2.1 - Real Environment

A suitable location on campus was chosen as the real world environment. The location was a football practice field, chosen for its few visual cues but with an adequate number of objects for users to make estimation from. For the real world condition, a digital camcorder was used to videotape the movie of the location. This was done by capturing the scene while walking forward along a predefined path from one corner of the field to its opposite end. This provides the user with forward view of the scene only. The movie was then edited using Adobe Premiere software and saved as AVI format for viewing on the projected display.

#### 3.2.2 - Virtual Environment

The virtual environment scene was modeled using MultiGen Pro software, running on a Silicon Graphic computer. Detailed measurements of the field and objects and their locations were carefully taken before the modelling process. Pictures of objects on the field were taken using a digital camera. Appropriate textures from these pictures (e.g. grass, trees, road textures) were used as textures in the modelled scene to match the virtual model as close as possible to the video of the real world. Shadows of objects were also approximately modelled. Movement in VE model was simulated similar to movements in the video movie using OpenGL Performer viewer software called PERFLY. The viewpoint in the virtual model is set to 1.4m, the height at which the actual scene is taken. The simulated movie will be run on a Windows NT computer, which makes it necessary to convert the simulation movie format to .AVI format. However, it is not possible to record the simulated movie by PERFLY directly. Thus, the simulation was first captured onto a VHS tape, and then converted to AVI format.

#### 3.2.3 - Display Apparatus And Room Setting

The movies (real and virtual) were displayed using an LCD projector connected to a computer. A single rear-projected display screen was utilized to allow viewing at close range without casting the shadow of the observer on the display screen. The display area size on the screen was adjusted to two size conditions: small display (0.3 x 0.4 m) and large display (1.36 x 1.59m) condition.

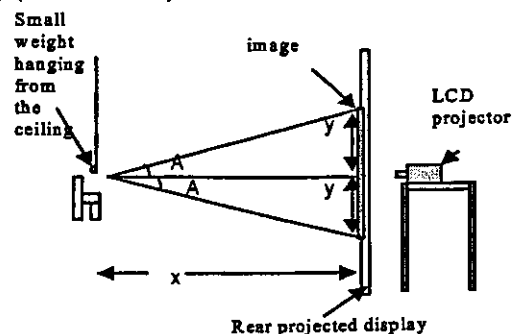


Fig. 2: Experimental set-up

The experimental room has no window thus giving it a dark condition when the lights are switched off. A dark setting is desirable here to reduce peripheral view effects from objects surrounding the projector screen, which might affect participants' distance estimations (Eby & Braunstein (1995) cited in [13]).

### 3.3 – Experimental Setup

The experiment involved a 2 x 2 factorial design. The two independent variables (IV) were image type and display type. Two levels of image type IV are video movies of the real and VE. The two levels of display type IV were small display and large display. The dependent variable (DV) is the estimated distance. Three levels of DV measured are vertical, horizontal, and transverse distances of objects in the environments. Four experimental conditions were used for this study: Real world movie (small display), Real world movie (large display), Virtual movie (small display), and Virtual movie (large display). As the same scene was used for all conditions, different group of participants were used for each condition to avoid training bias or interference from previous knowledge. Thus, four groups of ten participants were required for the study. The forty participants were randomly assigned to each group. Variables that were held constant between conditions include the followings: display resolution, display used (projected display only) FOV, eye level (centre of projection), textures of images, shadows, viewing and movement methods and paths through the scene and room setting (dark room). Resolutions of the display for all conditions were set to the same resolution (1024 X 768). The FOV of both display sizes were equated at the same angle: ~28 degrees. This was done by placing the viewer at a distance of 0.6m from the projected screen (for the small display), and 2.72m from the projected screen (for the large display). These distances were calculated as follows: Distance from display ( $x$ ) =  $y/\tan A$  (refer to Figure 2). Speed of movement through both scenes is set at 1.08 m/s, matching the speed of walking pace taken when real scene was captured.

### 3.4 – Procedures

Participants were initially briefed on the purpose and the procedure of the experiment. To ensure that participants were presented with the same FOV for each display type ( $x$  = small,  $y$  = large), they were seated at distances ( $d_1$  for small display) and  $d_2$  (for large display) from the projected display such that the angle subtended by the display size is the same ( $\alpha = \beta$ ) when they viewed the projected display under the small and large display condition (Figure 3). To reduce the effect of motion parallax cues, subjects were told not to move their head and body forward/backward and sideways during the experiment [10]. The eye level for all participants was set at the centre of the image projection height. This was done by adjusting the seat of each participant. A small weight hanging from a ceiling, set to the eye level height was used as a reference (see Figure 2 & 3).

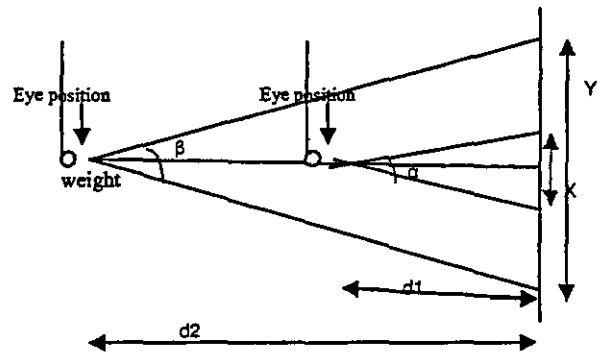


Fig. 3: Setting of eye level to be at the centre of projection

Prior to making estimations, participants were allowed to view the movies to familiarize themselves with the environment and the objects in it. Movement was restricted to play, forward and pause button only using a mouse. As the movement tasks were simple play/forward/pause of the movies, practice using the mouse to do this was not necessary. However, participants were informed of the respective functions of the mouse buttons. Participants were allowed to view the movie for three times and were informed when the time is up. The experimenter then set the scene at a preset viewpoint in the movie. Participants were informed of what distance to estimate. They were allowed to view the static scene from this viewpoint for up to 15 seconds before reporting their estimates. This was repeated for each of the eighteen distances, that is, six for each distance type. During estimations, participants were reminded not to move their head and body forward and backward or sideways to reduce motion parallax cues due to head movements. All estimations were made in meters (a meter long ruler was shown to participants before viewing the stimulus as an aide memoire). Each participant then completed a short post-test questionnaire.

## 4- Results

Initial examination of the data revealed one extreme value for one case of the data. This value occurred far from the middle of distribution (i.e. more than 75<sup>th</sup> percentile in a box plot) and was removed prior to further data analysis.

The estimated distances for each of the experimental condition (taken as average) were first compared to the actual distance. This was done separately for each of the 6 distances. Figure 4, 5 and 6 illustrate these comparisons for each distance type.

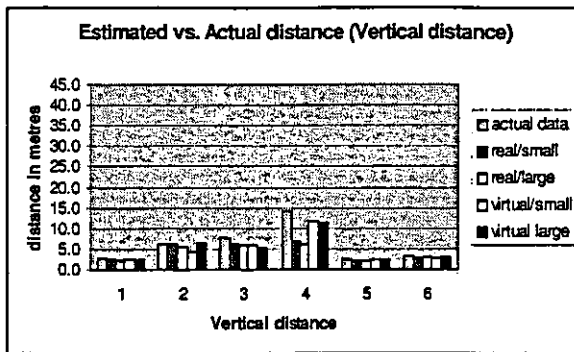


Fig. 4: Comparison of estimated distance to actual distance for each of the 6 vertical distances.

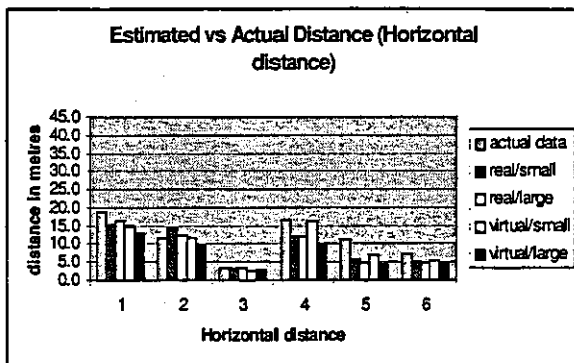


Fig. 5: Comparison of estimated distance to actual distance for each of the 6 horizontal distances

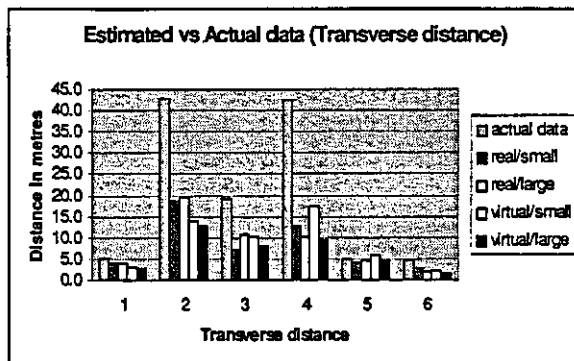


Fig. 6: Comparison of estimated distance to actual distance for each of the 6 transverse distances

Figure 4, and 5 show that vertical and horizontal distances were generally underestimated. Overall, it was revealed that differences among the four conditions (real image/small screen, real image/large screen, virtual image/small screen, and virtual image/large screen condition) were quite small.

For transverse distance, Figure 6 shows that distances were also underestimated but it was greater for larger distances (distance 2 and 4). Participants' estimations for transverse distance were generally smaller on virtual/large condition

compared to other conditions. A similar observation was noted for horizontal distance.

For further analysis, the raw data were transformed into a percentage format, that is, each estimated distance was calculated as a percentage of its actual distance. This allows us to statistically combined the results of different length of distances [14]. The following formula is used to make the conversion: % of estimated distance to actual = (estimated distance/actual distance) \* 100.

For each type of distance, an average of these percentage values was taken to represent each of the vertical, horizontal and transverse distances. To avoid overestimation values offsetting under-estimation values, the direction of error was ignored. Thus, values over 100 were adjusted by subtracting them from 200 and values over 200 were assigned zero, prior to averaging. These data were further analysed using general ANOVA/MANOVA. Image type (real versus VE) and display size (small versus large) were the independent variables. Estimated distances were used as the dependent variable. Three sets of analyses were performed for each distance type: vertical, horizontal and transverse. Significant level was initially set at 0.05.

#### 4.1 – Vertical Distance

A 2 (small versus large) x 2 (real versus virtual) ANOVA revealed no significant effect of image on vertical distance estimation ( $F(1,35) = .03$ ,  $p > 0.854372$ ; ). The results showed that all participants underestimate distance. A direct comparison between real and virtual image showed that the difference was small ( $M_{\text{real}} = 69.522$ ,  $M_{\text{virtual}} = 70.11$ ). However, the effect of display did approach significant ( $F(1,35) = 3.90$ ,  $p = 0.056105$ ). ). Both real and VE participants' performance were better on a small screen than on a large screen ( $M_{\text{small}} = 72.98402$ ,  $M_{\text{large}} = 66.65403$ ).

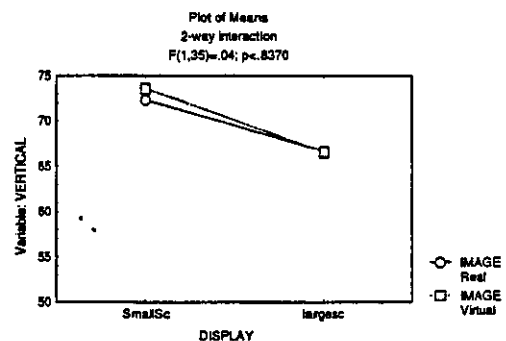


Fig. 7: Plot of means (2-way interactions) for display and image

A plot of means (2-way interactions) revealed no interactions of display and image type ( $F(1,35) = .042987$ ,  $p = .836952$ ) for vertical distance (Figure 7). From the plot, virtual image participants tend to perform better than real image participants on small display. No difference was observed on large display.

#### 4.2. Horizontal Distance

Similar to vertical distance judgment, a small difference was observed between real and virtual image ( $M_{real}=58.86211$ ,  $M_{virtual}=57.66484$ ,  $F(1,35)=.051956$ ,  $p=.821020$ ). Distances were generally underestimated in both environments, though real image participants performance were slightly better than virtual image participants especially on small display. On average participants were more accurate on small displays compared to large displays, however, no significant difference is observed ( $M_{small}=62.54536$ ,  $M_{large}=53.98160$ ;  $F(1,35)=2.658185$ ,  $p=.111986$ ).

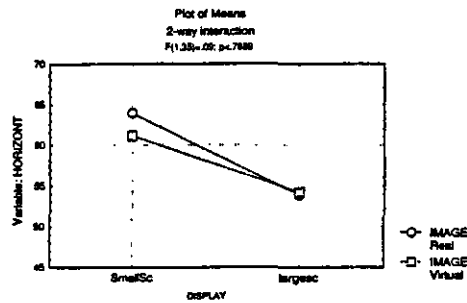


Fig. 8: Plot of means (2-way interactions) for display and image

A plot of means (2-way interactions), revealed no significant interaction of display and image type ( $F(1,35)=.087700$ ,  $p=.768871$ ) for horizontal distance (Figure 8).

4.3. Transverse Distance

A similar observation to horizontal distance result was noted for transverse distance. A 2 (small versus large) x 2 (real versus virtual) ANOVA showed no significant effect of image or display on transverse distance perception (Image:  $F(1,35)=.671762$ ,  $p=.417985$ ; Display:  $F(1,35)=1.804926$ ,  $p=.187764$ ). Underestimation for transverse distance is notably large; the percentage of estimation to actual is less than half. The result showed that percentage of estimation to actual for real image is higher than virtual image ( $M_{real}=47.87539$ ,  $M_{virtual}=43.48404$ ). Participants tend to perform slightly better on real image compared to virtual image. Similarly, distance perception is more accurate on a small screen than on a large screen ( $M_{small}=49.27899$ ,  $M_{large}=42.08065$ ).

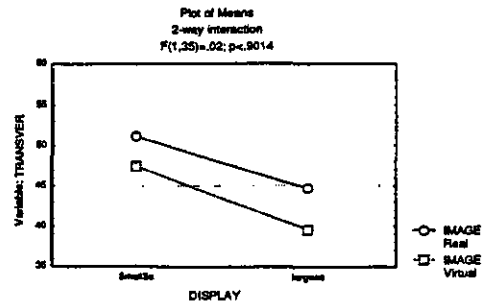


Fig. 9: Plot of means (2-way interactions) for display and image

A plot of means (2-way interaction) for transverse distance also revealed no significant interaction of image type and display type ( $F(1,35)=.015583$ ,  $p=.901369$ ) (Figure 9).

4.4. Comparison among distance types

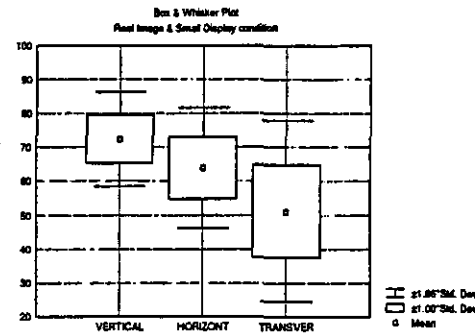


Fig.10 Real image/small screen condition

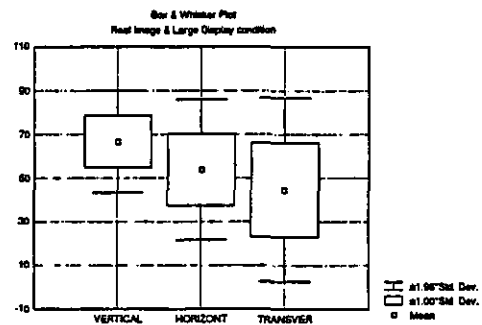


Fig. 11: Real image/ large screen condition

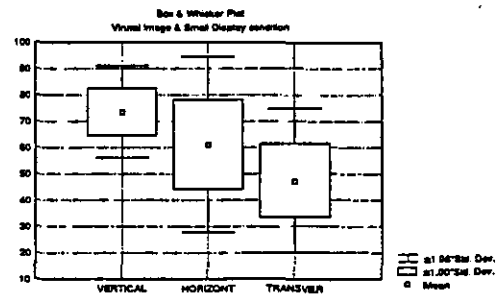


Fig. 12: Virtual image/small screen condition

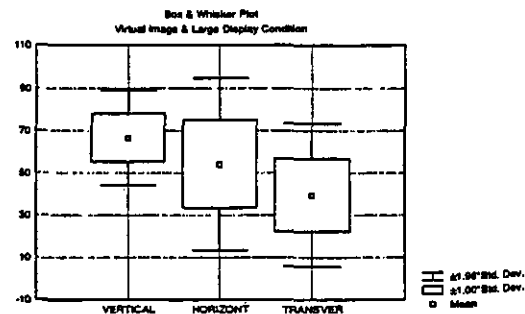


Fig. 13: Virtual image/large screen condition

Figure 10, 11, 12, and 13 compares the three distances in each of the four conditions: real image/small screen, real image/large screen, virtual image/small screen, and virtual image/large screen condition respectively. From the plots, the performance of the participants generally follow the same trend in all conditions; less estimation error was made in vertical distance followed by horizontal and transverse. In real/small and virtual/large condition, a t-test was used to compare the means between distance types yielded highly significant p-values ( $p < .001$ , Real/Small:  $M_{\text{vertical}}=72.35$ ,  $M_{\text{horizontal}}=63.92$ ,  $M_{\text{Transverse}} = 51.14$ ; Virtual/Large:  $M_{\text{vertical}} = 66.61$ ,  $M_{\text{horizontal}}=54.16$ ,  $M_{\text{Transverse}}=39.55$ ). However, in the real/large condition, only the horizontal-transverse comparison was not significant, other comparisons (vertical -horizontal and vertical-transverse) were found to be significant ( $M_{\text{vertical}}=66.68$ ,  $M_{\text{horizontal}}=53.80$ ,  $M_{\text{Transverse}}=44.61$ ). For the virtual/small condition, both vertical and horizontal estimations differ significantly from transverse distance but the difference between vertical and horizontal distances approaches significant ( $M_{\text{vertical}}=73.61$ ,  $M_{\text{horizontal}}=61.16$ ,  $M_{\text{Transverse}}=47.41$ ).

#### 4.5. Distance Estimation in Real and VE

By regarding the distance variable as a repeated factor a second ANOVA/MANOVA was performed on the dataset. The results showed the difference between real and VE is small. Overall, real image participants performed slightly better than virtual image and small display participants' estimation is better than large display participants. The effect of display on estimation does approach significant ( $p = .07$ ).

#### 4.6. Post-test Questionnaires Result

Participants were asked to rate their estimation on the scale of 1 to 7 (7 represent very accurate). The average response was 4. Only three felt confident of their estimation (rating =6). Four participants were very uncertain of their estimation (rate = 2). Most participants found transverse distance difficult to estimate (33) and vertical distance most easy to estimate (31). Survey on their sports background, only nine do not play any sports, the remainder are active in at least one of the following sports: tennis, badminton, squash, netball, hockey, cricket, and cycling. Only three participants did not find viewing the movie assisted them in their estimation, the rest found it allows them to make better estimation especially for distance objects. Generally, most participants reported using familiar objects in the scene (such as trees, lamppost, goal posts) to base their estimations. Others used their own height, imagined walking in the scene, and calculated distance based on the speed of the camera moving through the scene.

### 5- Analysis

Distances were generally underestimated in the real and virtual image for all distance types. This compression was more pronounced for large transverse distances, where estimation was less than half the actual distance. The difference in performance between real and virtual image participants for all distance types was small. Overall, subjects' performances were

better on small display compared to large display. For all viewing conditions, vertical distance was estimated significantly better than horizontal and transverse distances. Our post-test questionnaire result showed that most participants found vertical distance easy to estimate and transverse distance, most difficult.

### 6- Discussion

Consistent with [4][5][6] our present study indicates that distances were generally underestimated in the real and virtual environment. Previous investigations into distance estimation in the real world and virtual world reported differences with the VE producing larger error [4][5]. While contrary to these studies, our result is consistent with [6] who found no significant difference between an image-based panorama VE and computer modelled VE. In their study, Willemson & Gooch utilized picture images of the real scene to create the image-based environment. Similarly, our VE model used pictures of the real scene as textures for objects in the VE. Other studies have indicated that it is possible to perceive VE is similarly to the real world [3][15], however the VE used in these studies were simple and impoverished. It has been suggested that under impoverished conditions, the difference between both environments is small [16].

Prior studies showed that subjects' performance on large display is significantly better than on small display [10] [11] [12]. Our study however yields contrasting results. The present results revealed that subjects performed better on small display compared to large display for both images. In a related study, which compared distance perception on desktop and large projected display, Awang-Rambli & Kalawsky found that their participants performed significantly better on large display than small display [12]. Their subjects, however, performed distance estimation task on static pictures of real and virtual images. Patrick et al suggested that large image size might induce realistic experience in the participants in their study thus giving better judgment of relative distances [10]. However, their participants were allowed exploration of the test environment and were tested on a cognitive map test. These authors reported larger values on larger display, which correspond to better estimation results. Similarly, Tan and the others [11] reported that, with visual angles of the large display and desktop monitor equated, their subjects performed 26% better on large display compared to small display for on spatial orientation tasks. Results from their second experiment suggested this might be due to large display affords a greater sense of presence. Users are most effective when they feel more presence in the VE [11]. The large images viewed in our study, however, failed to induce similar experience. The present study utilized the same stimulus for the small and large condition. It is suspected that substantial difference might occur to the image when presented on difference display size. A comparison and closer examination of the images presented on small and



large display revealed a difference in image clarity and sharpness. When viewed on a large display the image was noticeably grainy and less clear compared to a much clearer image presented on the small display. This might account for the lower performance of the large display participants compared to the small display participants. For VE, further loss of image details might have occurred during the process of transferring the original image to VHS tape, thus might explain the lower performance of our VE participants compared to the real image participants. It has been shown that low resolution has an adverse effect on distance judgments [17]. It should be noted, however, that their result suggested only the lowest level resolution (52 x 35 pixels) produced significantly worst estimates. They also suggested that estimation errors were not monotonic function of resolution. A study comparing of participants' performance using various levels of resolution of the images is thus further required.

Direct comparison among distance types reveals that participants yielded more accurate results when estimating vertical length compared to horizontal and transverse distance. This is further supported by the post-test questionnaire result where participants found vertical distance easier to estimate. This result is consistent with the findings of Henry & Furness (1993), who found subjects' performance were almost veridical on vertical distance compared to horizontal distance [5]. This result is expected, as people are generally more familiar with their own height as a scale to other objects. This is further supported by our posttest questionnaire results that revealed subjects do actually use their height to base their estimations from. Very accurate performance in [5] might be due to the difference in the type of stimulus used. Their subjects estimated height of rooms in a museum while our subjects estimated vertical distance of various of objects in an outdoor setting. Interior spaces usually have standard heights and the fact that their subjects come from the architectural background account for the almost perfect estimations in their study. Our study showed that transverse distance give the worst performance. Similar findings by Loomis et al showed that more estimation errors were made on transverse distance than on lateral plane and this error is magnified when distance is increased [18]. For transverse distance, our participants reported less than half of actual value. This inaccuracy is more pronounced for larger distances. A similar observation by Wittner & Kline was reported for egocentric distance estimation. They found distance perception in VE to be less than half (47% of actual distance).

Waller (1999) reported that providing the ability to explore the virtual space would produce accurate result [3]. Our present study, however, revealed larger error in distance estimation especially for transverse distance where performance was on the average less than half of actual distance. In his study, Waller allowed subjects free exploration of the VE. On contrary, subjects in our study have no controlled of their viewpoints in the VE. The restricted movement in our experiment might have accounted for the attenuation in distance estimation. Other studies have indicated that active exploration of the VE produced better result compared to only passive viewing of the VE [19]. Accordingly, a free exploration of the VE would have yielded more accurate

results. The absence of the controlled motion parallax cues, a very effective cue for depth might have added to a lower estimation especially for transverse distance. Limited movement in the environment and low image resolution might explain the less impact of other dynamic cues such as optic flows, edge rate and motion perspective on subjects' estimation. It should be noted that VE used in our experiment, an outdoor setting, is more complex compared to the simple cubic room VE utilized by Waller. Additionally, subjects in our studies were asked to estimate distances among various objects at various locations in the scene. Comparatively, subjects' task in Waller study, which involved distance estimation between two cubes with corrective feedback given, is relatively easier. Besides, he reported error-corrective feedback has the strongest effect on accuracy in addition to the geometric field of view factor.

## 7- Conclusion

Our present study reported that distances were generally underestimated in both the real and virtual environment. The differences in distance perception between video images of the real and VE, within the constraint of the present experiment were small. On average, vertical distance were perceived more accurately compared to horizontal and transverse distance. Transverse distance was perceived less than half of the actual distance. More compression of distance was observed for larger distance. Generally, distances perceived in images presented on a small display produced less estimation error when compared to presentation a larger display. Although, our study reported better performance on small display compared to large display, contrary to the results of previous studies, further investigations are still needed to explain these differences.

Findings from our current works would have significant implications for applications that require spatial judgment tasks. Applications such as reconstruction of accident or crime scenes, where the actual real world scene may no longer exist or have been altered, a virtual model could be used as a substitute. For applications in which the VE model is used to represent its real world counterpart such as crime reconstruction, users performance might not as intended. The choice of display size and the type of image used to view the real or VE might have an effect on the observer's perceptual judgment performance.

## 8- Future Works

Findings from our study have resulted in more questions to be answered, thus entail the need for further research. Further works are needed to investigate the followings:

- The effect of image resolution on distance perception in real and virtual image viewed on different display size
- The effect of user-controlled navigation verses non-user controlled navigation on distance perception in VE

Current works are underway to investigate these effects. Other future studies would also include investigation into the effect of physiological cues and visual cues especially textures and other pictorial cues that are present in the stimulus.

## 9- Acknowledgement

The authors would like to thank Kip Shansi and Andre Schappo for their assistance with the experiment, Waad Yuosof for reviewing the statistics figures, and all those either directly or indirectly involved in this study.

## 10- Bibliography

- [1] Waller D., Hunt E., Knapp D.. Measuring Spatial Knowledge In A Virtual Environment: Distances And Angles. In The 39<sup>th</sup> Annual Meeting Of The Psychonomic Society, Dallas -Texas, 1998.
- [2] Golledge R.G. Cognition of Physical and Built Environments. In Garling G and Evans GW (Eds.), *Environment, Cognition, and Action : An Integrated Approach*, 35-62, NY :Oxford Press,1991
- [3] Waller, D. Factors Affecting The Perception Of Interobject Distances In Virtual Environments. *Presence*, 8(6) : 657-670, 1999.
- [4] Witmer, B.G., Kline P.B. Judging Perceived And Traversed Distance In Virtual Environment. *Presence*, 7(2):144-167, 1998.
- [5] Henry D., Furness T. Spatial Perception In Virtual Environments: Evaluating An Architectural Application. In *IEEE Virtual Reality Annual International Symposium*, Vol. 2, Seattle-Washington, 1993.
- [6] Willemson P., Gooch A.A. An experimental Comparison of Perceived Egocentric Distance in Real, Image-Based, and Traditional Virtual Environments using Direct Walking Tasks. In Tech Report UUCS-02-009, School of Computing, University , Utah, 2002.
- [7] Roscoe, S.N. Judgments Of Size And Distance With Imaging Displays. In *Human Factors*, 26(2) :617-629,1984.
- [8] Ruddle R.A., Stephen P.J., Jones D.M. Navigating Large-Scale Virtual Environments: What Differences Occur Between Head-Mounted and Desk-Top Displays? In *Presence*, 8(2) :157-168. 1999
- [9] Heineken D.E., Schulte F.P. Acquiring Distance Knowledge In Virtual Environments. In RTO HFM Workshop On "What Is Essential For Virtual Reality Systems To Meet Military Human Performance Goals?", The Hague-Netherland, 2000.
- [10] Patrick E., Cosgrove D., Slavkovic A., Rode J.A., Verratti T, Chiselko G. Using a Large Projection Screen as an Alternative to Head-Mounted Displays for Virtual Environments. In *CHI 2000*, The Hague-Amsterdam, 2000.
- [11] Tan D.S., Gergle D., Scupelli P.G., Paush R. With Similar Visual Angles, Larger Displays Improve Spatial Performance. In *CHI 2003*, Ft Lauderdale-Florida, 2003
- [12] Awang-Rambli D.R., Kalawsky R.S. The Effect Of Display And Image Type On Inter-Object Distance Estimation in Virtual and Real Environments. In *SIGCHI-NZ Symposium on Computer-Human Interaction*, Hamilton-New Zealand, 2002
- [13] Knapp J.M. The Visual Perception of Egocentric Distance in Virtual Environments. Doctoral Theses Dissertation. University of California, Santa Barbara, California, USA. 1999.
- [14] Henry, D. Spatial Perception in Virtual Environments: Evaluating an Architectural Applications. Master Thesis. University of Washington, USA, 1992
- [15] Yoon, J., Byun, E., Chung N.S. Comparison Of Space Perception Between A Real Environment And A Virtual Environment. In *The IEA 2000/HFES 2000 Congress*, The Hague-Netherland, 2000.
- [16] Egglestons R.G., Janson W. P., Aldrich K.A. Virtual Reality System Effects On Size-Distance Judgments In A Virtual Environment. In *IEEE Annual Virtual reality International Symposium*, Santa Clara-California., 1996.
- [17] Jää-Aro K-M, Kjelldahl L . Effects Of Image Resolution On Depth Perception In Stereo And Non-Stereo Images. In *Stereoscopic Displays and Applications VIII*, SPIE Proceedings, 3012 :319-326, 1997.
- [18] Loomis, J.M., Da Silva, J.A., Philbeck, J. W. ,Fukushima, S. S. Visual Perception Of Location And Distance. In *Current Directions In Psychological Science*, 5(3) :72-77,1996.
- [19] Peruch P. Acquisition of Spatial Knowledge Through the Visual Exploration of Simulated Environments. In *Ecological Psychology* , 7(1) :1-20,1995.

# The Effect Of Display And Image Type On Inter-Object Distance Estimation in Virtual and Real Environments

Dayang R. Awang-Rambli

Computer Science Dept,  
Loughborough University,

Loughborough, Leicestershire, LE11 3TU, UK,

Tel: +4490) 01509 223047, Fax +44(0) 01509  
223940

E-mail: [D.R.Awang-Rambli@lboro.ac.uk](mailto:D.R.Awang-Rambli@lboro.ac.uk)

Professor Roy S. Kalawsky

Advanced VR Research Centre,

Loughborough University,

Loughborough, Leicestershire, LE11 3TU, UK,

Tel: +4490) 01509 223047, Fax +44(0) 01509  
223940

E-mail: [R.S.Kalawsky@lboro.ac.uk](mailto:R.S.Kalawsky@lboro.ac.uk)

## ABSTRACT

This paper reports on a study to examine the effect of display type (desktop display verses projected display) on inter-object distance estimation in real and virtual environment (VE). Non-stereo images of real and virtual environments were used as stimulus. Participants were asked to estimate two distances: transverse distance (objects lying in the sagittal plane – in depth) and lateral distance (objects on the same horizontal line). Our result shows that distances were generally underestimated. For transverse distance, no significant difference was found for real and virtual images on both type of display. On average, lateral distance estimations yielded more accurate results for virtual image. Participants' performances were better on projected display compared to desktop display on both lateral and transverse distance. A significant effect of display on distance was revealed for lateral distance.

## Keywords

virtual environment, distance perception, inter-object distance estimation, visual cues, display.

## INTRODUCTION

Virtual Environment technologies presently have been attracting profound interest from a variety of fields. The term virtual environment refers to a simulated experience in a three-dimensional computer-generated synthetic environment where users can move around and interact within it [6][5][2]. Providing simulated environments of the physical world, in real time, makes it a potentially attractive/important tool for a wide range of areas such as training, prototyping, architecture, tele-operations of robots, medicine, visualization of complex data sets, architecture and regional planning. VE allows designers, clients, and decision makers in the area of prototyping and architecture, an early preview of the planned 3D space through simulated environments, and thus, allow cost and time saving decisions to be made prior to the delivery of actual physical structure or product. In applications, such as flight training and fire fighters training, or surgery training, VE provide simulated environment of places or situation, which are

rare, remote or dangerous [16][13]. Thus, trainees can practice in the safe VE.

However, in some respects current VE technologies are still inadequate. Several studies have indicated that VE allows users to perceive space differently from the real environment [4][19][9][13][7][16]. In order for VE technologies to be effectively applied to various fields of applications, particularly, those that use VE to represent its real world counterpart; it ought to allow users to perceive the virtual and real environments similarly. Users must be able to apply knowledge acquired in the VE to the real environment. As such, it is necessary for research be directed toward finding answers to these basic questions: how to make a user perceive a VE similar to a real world? To what extent that experience gained in the VE can be used to represent the real world? How is the knowledge acquired in VE transferable to the real world? Similar questions have been the focus of several researchers [4][16][21] and these questions serve to motivate the current and future works in our research.

A VE enables an immersed user to experience a different environment through exploration of 3D virtual space, thus understanding of spatial knowledge plays an important role in determining objects' and participants' sizes, distances and orientation within the environment [18][16][3]. Spatial awareness refers to a person understanding the 3D spatial environment. It involves knowledge of location and orientation of objects and of the participants themselves within the 3D space. In the real world, human perceptual understanding of the 3D space is mainly derived from visual cues for depth and distance [1]. Similarly, within the virtual environments, these cues are used to obtain spatial characteristics of virtual 3D space. One focus of our research is to study the effects of various visual cues on distance estimation, in a goal to generate a simulated 3D environment accurately or closely represents its real world space.

For the immersive experience in the VE, head mounted displays (HMD) have been used to provide users with a seemingly realistic experience compared to desktop monitor [Ruddle et al (1999) cited in [11]]. However, studies have shown that distance perceptions to objects in VE viewed through HMD are constantly underestimated when compared to the real world, and it is unknown why these differences occurs [17]. It was suggested that the display device used might affect distance judgment in VE [17, see also [12]].

This paper reports an initial study to investigate one aspect of spatial knowledge - distance estimation. The influence of display type and image type on distance estimation was investigated. Subjects were required to make distance estimation between objects presented to them in the form of pictures of a real world scene and of a virtual world scene. Non-stereo images were used in this first experiment, as it was desirable to remove stereo-acuity problems that may be experienced by certain users. Another reason is that many available VE displays are non-stereo. However, stereo images will be used in later experiments. Two types of distances compared in this study were: transverse distance and lateral distance. This study is part of a series of experiments in our research to investigate similarity of users' perception of the VE as compared to the real world.

EXPERIMENT

The overall aim of this initial study was to focus on users' perception of distance in real world versus a virtual world model: is distance measured in a virtual world similar to a real world measurement? Factors such as motion parallax (resulting from head movements) and stereoscopic cues were eliminated in this study by the use of non-stereo images. Specifically, effects of the image type (real world picture and virtual world picture) and display type (desktop and projector display) on the subjects' inter-object distances were investigated. In this study, participants were asked to estimate two distances: transverse distance and lateral distance (see Figure 1).

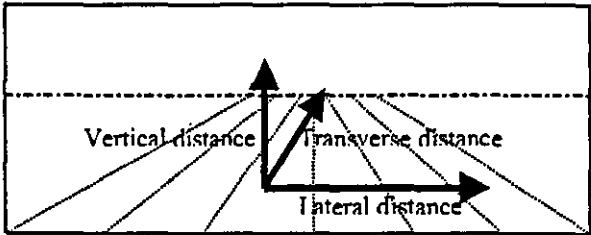


Figure 1 Transverse and lateral distance.

METHODOLOGY

Participants

Forty participants (six females and thirty-four males), comprising of staff, students and faculty members from Loughborough University took part in the study. The ages

of the participants range from 15 to 51 years with an average age of 30.

Materials/Apparatus  
Pictures/Images

A photograph of a location on campus was taken and placed on a Microsoft PowerPoint slide (in full-screen mode) for the real picture condition. A virtual model of this scene was created using REALAX software on a Windows NT machine. Appropriate textures (trees, grass, road, sky) were taken from the real picture to match the virtual model as close as possible to the real picture. The viewpoint in the virtual model is set to 1.5m above the ground, at the same point where the picture is taken in the real world. A snapshot of the virtual model is taken and placed on a Microsoft PowerPoint slide to represent the virtual picture condition.

Display Types (Desktop Pc and LCD Projector)

The images were displayed using a Windows NT machine with a 17" monitor display for the desktop condition. An LCD projector was connected to a Windows NT machine and was used to project the pictures (real and virtual) to a large white paper (135 x 95 cm) on the wall in the projected display condition.

Procedures and Experiment Setup

The aim of the experiment is to observe participants' estimation of distances between objects in a real world picture and virtual world picture presented under the following two conditions: on a desktop display and on a projected display. The experiment involves a 2 x 2 factorial between- subject design. The two independent variables (IV) are display type and the image type. The two levels of the display type IV are desktop and projected display. The two levels of image type IV are real world picture and virtual model picture. The dependent variable is the estimated distance between objects.

The participants were divided into four groups of ten participants each. Presentation of the pictures for each group is summarized in Table 1.

| Group 1                               | Group 2                                | Group 3                                  | Group 4                                    |
|---------------------------------------|----------------------------------------|------------------------------------------|--------------------------------------------|
| Real picture<br>(desktop<br>display ) | Real picture<br>(projected<br>display) | Virtual picture<br>(desktop<br>display ) | Virtual picture<br>(projected<br>display ) |

Table 1. Presentation of conditions taken for each group

Participants were given instructions verbally and as well as written on the computer and projected display. To reduce differences of a meter length concept among participants, a meter long tape was shown to them prior to the start of the experiment. Participants were required to estimate two distances: X (transverse distance- distance between two lampposts) and Y (lateral distance – distance between a

lamppost and a hedge). Estimations were to be made in meter unit (see [19]). To avoid participants changing their mind very often (refer to [14]), each participants were given 15 seconds only for each estimation and then was required to write down the estimation on a data sheet. Each participant then completed a short post-test questionnaire afterwards.

## RESULTS

Initial examination of the data revealed outliers in the real picture/desktop condition. The values of distance X and Y in two cases of the data occurred far from the middle of distribution (i.e. more than 75<sup>th</sup> percentile in a box plot). These outliers were removed prior to further analysis of the data. It was noted that the observed results produced a slightly skewed distribution. However, this data set showed consistency after a direct comparison of the means was made. The means for both images on desktop shows little difference. A similar observation was true for both images on projected display.

Accuracy is measured in terms of how close is the estimated distance to the actual distance. For the purpose of this study, the following formula computes the percentage of (under/over) estimation from the actual distance. Underestimation is shown by the negative values, whilst positive values indicate an overestimation of distance.

$$\% \text{ of Estimation to actual distance} = \left( \frac{\text{Estimated distance} - \text{Actual distance}}{\text{Actual distance}} \right) * 100$$

For distance X, nearly all participants underestimated the distance. Estimation varies between conditions for distance Y, where distances were underestimated and overestimated. General ANOVA/MANOVA was used to analyze these data further. Image (real versus virtual) and display (desktop versus projected display) were the independent variables, and estimated distances were used as the dependent variable. Two sets of analyses were performed: one on distance X and the other on distance Y.

### Distance X: Transverse distance

A 2 (desktop versus projected display) x 2 (real versus virtual picture) ANOVA/MANOVA revealed a significant effect of display on distance estimation ( $F = 5.212213$ ,  $p\text{-level} = .028802$ ). Significant level was initially set at 0.05. Projected display participants made more accurate estimation compare to the desktop participants ( $M_{\text{desktop}} = 8.66$ ;  $M_{\text{projected display}} = 13.34$ ) in both images (real and virtual picture). Underestimation by desktop participants is notably large (-61.28%), whereas, on the projected display, average % estimation to actual distance is -0.42%.

A plot of means (2-way interaction -on display (desktop versus projected display) and image (real versus virtual

picture)) for distance X (Figure 2) reveals no significant interaction ( $p < 0.9522$ ). The results showed that all participants underestimated distance. On the average, no significant difference was observed on the estimates for real and virtual picture on both display condition ( $M_{\text{real desktop}} = 8.63$ ,  $n = 8$ ;  $M_{\text{virtual desktop}} = 8.7$ ,  $N = 10$ ;  $M_{\text{real projected display}} = 13.19$ ,  $n = 8$ ;  $M_{\text{virtual projected display}} = 13.5$ ,  $N = 10$ ).

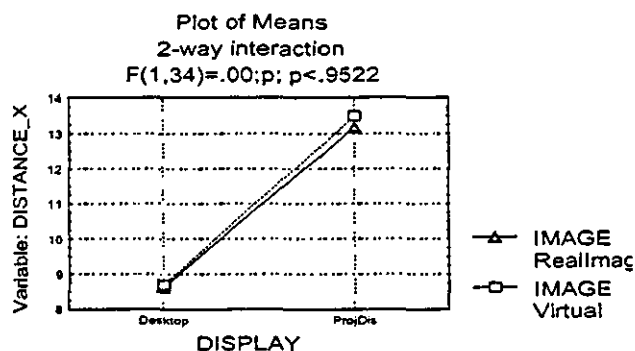


Figure 2 Plot of means - 2-way interaction for distance X - display (desktop versus projected display, real versus virtual picture). Actual distance is 22.8m.

### Distance Perception in Virtual and Real Environment for Distance X

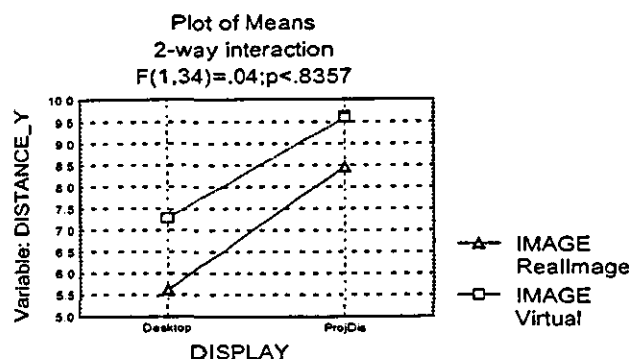
Estimations made in virtual and real environments are generally underestimated. Overall no significant difference was observed between virtual and real images, estimation was averaged at half of the actual distance (% estimation of actual = -50.15; % estimation of actual = -50.45). A direct comparison made between virtual and real environments estimations on desktop reveals that no significant difference is observed between both images ( $M_{\text{real desktop}} = 8.63$ ,  $M_{\text{virtual desktop}} = 8.7$ ). A similar observation was noted on projected display condition ( $M_{\text{real projected display}} = 13.19$ ,  $M_{\text{virtual projected display}} = 13.5$ ). VE and real environment participants performed generally better on a projected display than on a desktop.

### Distance Y: Lateral distance

A 2 (desktop versus projected display) x 2 (real versus virtual picture) ANOVA observe no significant effect or interaction, although the effect of the independent variable display approaches significant ( $F = 4.059417$ ,  $p = 0.051889$ ). Participants on the average tend to underestimate distances in both images, but to a much greater extent in the real image ( $M_{\text{real}} = 7.05$ ;  $M_{\text{virtual}} = 8.45$ ). Overall, estimates are more accurate for the projected display conditions, average magnitude estimation to actual is 3.49% compared to -22.35% for desktop participants ( $M_{\text{projected display}} = 9.03$ ;  $M_{\text{desktop}} = 6.46$ ).

For distance Y, a two-way interaction (2 (desktop versus projected display) x 2 (real versus virtual picture)) plot of means (see Figure 3) reveals that participants underestimated distance in both real and virtual image for

the desktop condition, though estimates made for virtual picture on the average are more accurate than in real picture ( $M_{\text{real desktop}} = 5.63$ ,  $M_{\text{virtual desktop}} = 7.30$ ). However, when viewed on a projected display, real image participants produce very accurate estimation compared to virtual participants ( $M_{\text{real projected display}} = 8.47$ ;  $M_{\text{virtual projected display}} = 9.6$ ).



**Figure 3** Plot of means-two-way interaction-for distance Y (desktop versus projected display, real versus virtual picture). Actual distance is 8.73m

#### Distance Estimation in Real and Virtual Environment for Distance Y

Distances are generally underestimated for both real and virtual environment. Participants in the VE yielded more accurate estimates, averaging -3.21% underestimation of actual compared to -17.40% for real image participants. For images presented on desktop, participants tend to produce better estimates for virtual image (-16.38% of actual) compared to real image (-35.42% of actual). However, real image participants perform much better than virtual image participants for the projected display conditions (Real image: -2.97% of actual; Virtual image: 9.96% of actual).

#### Posttest Questionnaires Results

In the posttest questionnaires, participants were asked to describe how they made their estimation and to indicate which distance (X or Y) is easier to estimate. Generally, most participants reported estimations based on objects features and locations in the pictures such as the hedges, roads and lampposts. For the virtual environment, some participants commented that they tried to visualize the scene as a real one based on everyday experience. Nearly everyone commented that distance Y is easier to estimate because no perspective was involved here.

#### ANALYSIS

An important observation made in this experiment is the effect of display on distance estimation: projected display yields more accurate estimation in both types of distances, X and Y. Both images (virtual and real) produce more accurate estimations on a projected display compared to desktop.

No significant difference was found for real and virtual picture viewed on both displays for distance X. The slightly better estimations of virtual image on projected display may be due to other factors not investigated in this study. Surprisingly, for distance Y, distance estimations were more accurate for a virtual picture viewed on a desktop; error made was less than half of real picture estimation. But when viewed on a projected display, a real image produces more accurate results compared to a virtual image.

Overall, estimations made in real and virtual picture in distance X shows not much difference. (Real picture, of estimate = -50.15,  $M = 10.91$ ; Virtual, of estimate = -50.45,  $M = 11.1$ ). For distance Y, virtual image yields more accurate estimation compared to real image (Real picture, of estimate = -17.4,  $M = 8.66$ ; Virtual, of estimate = -3.21,  $M = 13.34$ ).

#### DISCUSSION

Consistent with the findings of [19][9][4] distances are generally underestimated in the real and VE for both distance X and Y. This inaccuracy is expected as the stimulus used were pictures. Lumsden indicated that inter-object distance distortion occurs when viewing a photograph of three-dimensional scene, our results show similar occurrence for computer-displayed images [10]. For distance X, the present study reveals no significant difference on distance perceived in real and virtual images, whilst VE participants' performance were more superior compared to the real world for distance Y. Witmer & Kline, however, found that egocentric distances are underestimated more in VE than in real world are perceived less than half of the actual distance [19]. They reported 72% of true distance for real world performance and 47% of true distance for VE. Our participants, however, yield more accurate result on distance Y (average % underestimate is 3.21 for virtual and 17.40 for real), while for distance X both real and virtual participants underestimated approximately 50% of the actual. More visual cues (familiar objects and perspective cues) available in our stimuli (images) might account for this difference. In the real world estimations, however, estimates on average ranges between 87-91% of actual distance [Wright (1995), cited in [19]]

Waller and Yoon et al indicate that people can perceive distance in virtual world similar to the real world [15][21]. Corroborating these findings, our current study, though an underestimation, on the average (for distance X) reveals similar observations. The virtual model utilized textures from the real world picture, makes it closely resemble real picture. No significant difference between real and virtual environment in our results might be attributed to this resemblance. A similar but comparable experiment conducted by [20] reported that with regards to relative perception of horizontal and vertical extents, a snapshot of a VR scene on a desktop is similar to a picture. This might



further explain our result since most participants reported in the posttest questionnaires used visual cues such as the height of the lamppost, hedges and trees.

The present study differs from those of [15] and [21] who reported that inter-object distances in VE are generally overestimated. Our study reported that distances are generally underestimated. When viewing a photograph of a three-dimensional scene that has been magnified, Lumsden suggested that distortions of inter-object distance occur when two or more identical objects are viewed at increasing distances from the observer causing an apparent decrease in the distance between the objects [10]. This might account for the underestimations made in distance X. However, for distance Y, although an underestimation, our participants' performances in VE were unexpectedly very accurate compared to those in real world. This result contrasts those of [7] and [19], whose findings reported a more accurate judgment in real world compared to VE. The virtual model utilized textures from the real world picture, makes it closely resemble the real picture. This might account for a more accurate estimation but this does not explain its better performance over real image. Although shadow is not present in the pictures, other lighting effects (brightness & contrast) in the real image were initially closely matched to the virtual picture. However, a direct comparison of the real and virtual picture revealed that objects in virtual picture are sharper and clearer (more contrast) than in the real picture. This might account for better estimation in virtual picture.

Our present study indicates that generally participants reported larger estimates when viewing images on a projected display than when viewing on a desktop display. However, these larger values tend to correspond to better estimation results, actually, quite accurate results were produced especially for distance Y where error made is only 3.49% of actual. It was proposed that vertical overestimation would increase if a picture were distended such as projecting it onto a larger screen [20]. Even though vertical estimation is not investigated in this study, but most participants reported using objects' height in the scene to base their estimation. This might account for the larger estimate values made when images are viewed on larger screen. In a comparable study, investigating spatial knowledge gained in VE viewed in three conditions (HMD, desktop monitor and large projection screen), it was found that the large projection screen performance was more accurate than the other two conditions [11]. Patrick et. al. suggested that this better performance might be due to the image sizes that are large enough to induce a realistic appearance on the participants, thus better judgment of relative position was perceived [11]. The accurate result for the projected display in our study may also be due to the participant having similar experience. The scene depicted by both pictures were similar but the angles subtended by

both display types differed by a few degrees with the desktop having slightly larger field of view, this might have an affect on perception of distances.

A direct comparison between estimation made in distance X and distance Y reveals that participants produce more accurate estimations when judging objects on the same plane (distance Y). Correspondingly, Loomis et. al. found that more estimation errors were made on transverse than on lateral plane [8]. They also found that the degree of perceptual distortion increases with distance. This might explain the greater distortion in distance X in our study.

## CONCLUSION

Generally, our findings show that most distance judgments are underestimated. Images viewed on a projected display generally produced more accurate estimation compared to when viewed on a desktop display. Significant effect of display on distance was shown.

However, on average our current study reveals no significant difference between objects perceived in real or virtual world for transverse distance estimations. Contrary to most studies, an unexpected outcome is the better performance of VE participants over the real world participants for lateral distance on desktop display. Accepting these results would have great implications on applications such as the reconstruction of accident or crime scene where, a virtual model of the scene would accurately represent the actual scene compared to pictures taken. It should be noted that factors such a motion parallax and stereopsis were not present in this experiment. Other visual cues, which may be present such as linear perspective, relative size, relative height, foreshortening, occlusion, and texture gradient might account for the observed results.

## FUTURE WORK

The results of this work are considered extremely important for any application where spatial judgment is required. Such examples include training tasks where people have to observe and interact with synthetically generated scenarios. Transfer of knowledge gained in virtual environments to real situations may not be as effective as desired. Results from the present study do not clearly elucidate the better performance of larger display over small display. It is also not clear why virtual environment yields more accurate result compared to real world for transverse distance on projected display and lateral distance on desktop display. It is hoped that further research will identify the associated impact of this discrepancy between the real world and the virtual world interactions.

Future works would include investigating the effects of other visual cues such as textures, object heights, other depth cues (linear perspective, relative size and height, foreshortening and occlusion) and the content of the scene

on distance perception. Studies into the effect of various visual cues in vertical estimation as well as for transverse and lateral estimations will also be part of the follow-on experiments.

## ACKNOWLEDGEMENT

The authors would like to thank Mohd Y Nayan, Alan Chamberlain, Kip Shansi, for their assistance with the experiment, Waad Yuosof for reviewing the statistics figures, and all those either directly or indirectly involved in this study.

## REFERENCES

1. Crvarich, Gina (1999). An Exploration Of Techniques To Improve Relative Distance Judgments Within An Exocentric Display. Master's Thesis. University of Washington. 1995. Available at <http://www.html.washington.edu/publications/crvarich/1.html>
2. Ellis Stephen R (1992). What Are Virtual Environments? IEEE Computer Graphics & Applications, January 1994, pp17-22.
3. Heineken D.E., Schulte F.P. (2000). Acquiring Distance Knowledge In Virtual Environments. Paper Presented At RTO HFM Workshop On "What Is Essential For Virtual Reality Systems To Meet Military Human Performance Goals?", Hague, Netherland, 13-15 April 2000, published in RTO MP-058, pp17-1-17-5
4. Henry Daniel, Furness Tom (1993). Spatial Perception In Virtual Environments: Evaluating An Architectural Application. IEEE Virtual Reality Annual International Symposium, Vol. 2, Sept 18-22, 1993, Seattle, Washington, pp 33-53.
5. Isdale J (1998). What Is Virtual Reality? Accessed Via WWW <http://www.cse.dmu.ac.uk/~cph/vr/whatisvr.html> (5 Oct 1998).
6. Kalawsky R S (1993). The Science Of Virtual Reality And Virtual Environments. Addison Wesley, 1993.
7. Lampton, D.R., Bliss J.P., Knerr B.W (1994). Object Recognition, Size Estimation, And Distance Estimation In Real-World And Virtual Environments (abstract only). Proceedings Of The Human Factors And Ergonomics Society 38<sup>th</sup> Annual Meeting -1994. Vol. 2, pp 980.
8. Loomis, J. M., Da Silva, J. A., Philbeck, J. W. & Fukusima, S. S. (1996). Visual Perception Of Location And Distance. Current Directions In Psychological Science, 5(3), pp 72-77.
9. Loomis, J M, Knapp, J M (1999). Visual Perception Of Egocentric Distance In Real And Virtual Environments. To Appear In L.J. Hettinger and M.W. Haas (Eds.), Visual and Adaptive Environments. Hillsdale NJ:Erlbaum.
10. Lumsden, E.A. (1980). Problem Of Magnification And Minification: An Explanation Of The Distortion Of Distance, Slant, Shape, And Velocity. In Hagen, M(Ed.) The Perception Of Pictures, pp91-135. New Jersey: Academic Press, pp91-135.
11. Patrick et al. (2000). Using a Large Projection Screen as an Alternative to Head-Mounted Displays for Virtual Environments. CHI 2000, The Hague, Amsterdam, 1-6 April 2000, pp478-485.
12. Roscoe, Stanley N, (1984), "Judgments Of Size And Distance With Imaging Displays", Human Factors, 1984, 26(2), pp 617-629
13. Sinai, M.J., Krebs, W.K., Darken, R.P., Rowland, J. H., And McCarley, J.S. (1999). Egocentric Distance Perception In A Virtual Environment Using A Perceptual Matching Task. In Proceedings Of The 43rd Annual Meeting Human Factors And Ergonomics Society, pp1256-1260.
14. Storms R L, Zyda M J (2000). Interactions in Perceived Quality of Auditory-Visual Displays. Presence, Vo. 9, No. 6, December 2000, pp557-580. MIT 2001.
15. Waller, David (1999). Factors Affecting The Perception Of Interobject Distances In Virtual Environments. Presence, Volume 8, No. 6, December 1999, pp 657
16. Waller D, Hunt E, Knapp D (1998). Measuring Spatial Knowledge In A Virtual Environment: Distances And Angles. Paper Presented At The 39<sup>th</sup> Annual Meeting Of The Psychonomic Society, Dallas TX, 21 Nov. 1998.
17. Willemson P, Gooch AA (2002). An experimental Comparison of Perceived Egocentric Distance in Real, Image-Based, and Traditional Virtual Environments using Direct Walking Tasks. Tech Report UCCS-02-009, School of Computing, University of Utah. Feb 2000
18. Wilson P N, Foreman N, Tlauka M (1997). Transfer Of Spatial Information From A Virtual To Real Environment. Human Factors, 1997, 39(4), pp526-531.
19. Witmer, Bob G, Kline Paul B (1998). Judging Perceived And Traversed Distance In Virtual Environment. Presence, Volume 7, No.2, April 1998, pp144-167.
20. Yang T L, Dixon M W, Proffitt DR (1999). Seeing Big Things: Overestimation Of Heights Is Greater For Real Objects Than Objects In Pictures. Perception, 1999, Volume 28, pp445-467.
21. Yoon, Jungsun, Byun, Eunhee, Chung Nyang S. (2000). Comparison Of Space Perception Between A Real Environment And A Virtual Environment. In Proceedings Of The IEA 2000/HFES 2000 Congress, pp1515-1518.



