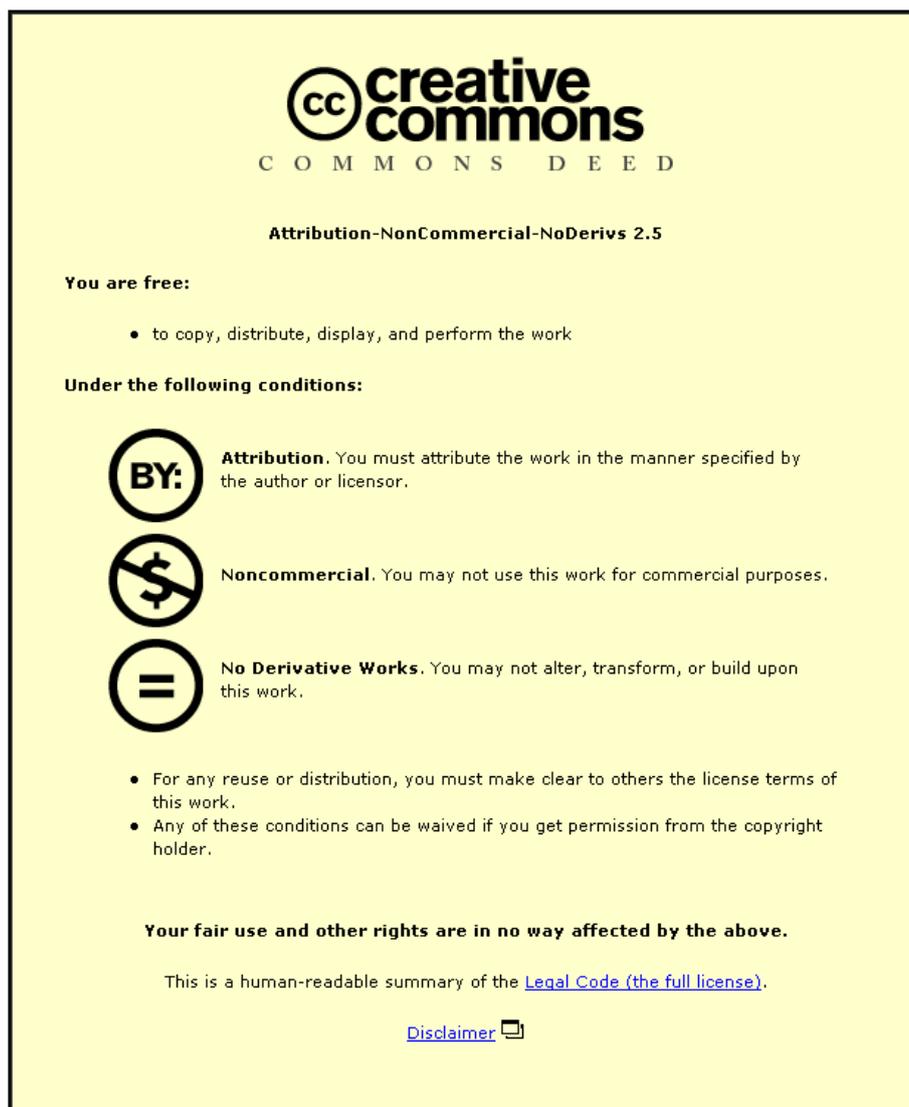


This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (<https://dspace.lboro.ac.uk/>) under the following Creative Commons Licence conditions.



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

**MEASUREMENT OF RANGE OF MOTION OF
HUMAN FINGER JOINTS, USING A COMPUTER
VISION SYSTEM**

By

Abdusalam Ben-Naser

A Doctoral Thesis

Submitted in partial fulfilment of the requirements
for the award of Doctor of Philosophy
of Loughborough University

September 2011

© By Abdusalam Ben-Naser, 2011

Acknowledgements

This research project would not have been possible without the support of many people. First and foremost I offer my sincerest gratitude to my supervisors Dr. David Kerr and Dr. Kaddour Bouazza-Marouf who have supported me throughout my thesis with their guidance and knowledge.

Also I would like to take the opportunity to thank all people at Wolfson School and Loughborough University who created a wonderful environment at this university and made Loughborough one of the top universities in the UK.

To:

*Whom images do not need enhancement,
who scarified their lives to liberate Misrata*

Abstract

Assessment of finger range of motion (ROM) is often required for monitoring the effectiveness of rehabilitative treatments and for evaluating patients' functional impairment. There are several devices which are used to measure this motion, such as wire tracing, tracing onto paper and mechanical and electronic goniometry. These devices are quite cheap, excluding electronic goniometry; however the drawbacks of these devices are their lack of accuracy and the time-consuming nature of the measurement process.

The work described in this thesis considers the design, implementation and validation of a new medical measurement system utilized in the evaluation of the range of motion of the human finger joints instead of the current measurement tools.

The proposed system is a non-contact measurement device based on computer vision technology and has many advantages over the existing measurement devices. In terms of accuracy, better results are achieved by this system, it can be operated by semi-skilled person, and is time saving for the evaluator.

The computer vision system in this study consists of CCD cameras to capture the images, a frame-grabber to change the analogue signal from the cameras to digital signals which can be manipulated by a computer, Ultra Violet light (UV) to illuminate the measurement space, software to process the images and perform the required computation, a darkened enclosure to accommodate the cameras and UV light and to shield the working area from any undesirable ambient light.

Two calibration techniques were used to calibrate the cameras, Direct Linear Transformation and Tsai. A calibration piece that suits this application was designed and manufactured. A steel hand model was used to measure the fingers joint angles. The average error from measuring the finger angles using

this system was around 1 degree compared with 5 degrees for the existing used techniques.

Key words

Direct Linear Transformation, Tsai camera calibration, 3D computer vision, photogrammetry, Camera calibration, camera parameters, non-contact measurement system, fingers range of motion.

Contents

1	Introduction	1
1.1	Problem identification	1
1.2	The thesis motivation and objective	3
1.2.1	Potential advantages of the proposed system	4
1.3	Thesis methodology	5
1.4	Thesis Outline	6
2	Literature review	8
2.1	Anatomy and the motion of the fingers	8
2.2	Existing finger joints measurements devices	9
2.2.1	Visual estimation	10
2.2.2	Tracing onto paper	11
2.2.3	Wire tracing	11
2.2.4	Composite finger flexion or pulp to palm measure	11
2.2.5	Universal Goniometry	12
2.2.6	Inclinometer or gravity-dependent goniometer	14
2.2.7	Electrogoniometer	15
2.2.8	HandMaster	15
2.2.9	Glove technology	16
2.3	Three-dimensional Computer vision system	18
2.3.1	What is a Three-dimensional computer vision system?	18
2.3.2	Why use a computer vision system?	19
2.3.3	Factors that affect the performance of the computer vision system	19
2.4	Some applications of computer vision systems	20
2.5	Elements of a computer vision system	25
2.5.1	Image acquisition	26
2.5.2	Illumination	27
2.5.3	Image formation and focusing	28
2.5.4	Image detection	29
2.5.5	Solid state camera and output signal	29
2.5.6	Image processing techniques	32
2.5.7	Thresholding	34
2.6	Conclusion	35
3	Camera calibration and reconstruction of 3D coordinates	37
3.1	Summary	37
3.2	Introduction	37
3.3	What is camera calibration?	38
3.4	Why We Need Calibration	39
3.5	Calibration Techniques	40
3.5.1	Standard and classical calibration	40
3.5.2	Camera self-calibration	41
3.6	Calibration objects	42
3.7	Camera parameters	44
3.7.1	External parameters (Extrinsic)	44
3.7.2	Internal parameters (Intrinsic)	44
3.8	Tsai's calibration method	44

3.8.1	Image distortion.....	47
3.8.2	The horizontal Uncertainty Factor and image centre.....	48
3.8.3	Tsai camera model.....	49
3.9	Reconstruction of 3-D world coordinates using Tsai method.....	49
3.10	Direct Linear Transformation method.....	53
3.10.1	Configuration of calibration points.....	54
3.11	Conclusion.....	57
4	System design and measurement process.....	58
4.1	Introduction.....	58
4.2	Components of the computer vision system for this study.....	58
4.2.1	Cameras.....	58
4.2.2	Lens.....	58
4.2.3	Frame Grabber.....	59
4.2.4	Illumination.....	59
4.2.5	Marker and markerless computer vision system.....	60
4.2.6	Ultraviolet gel.....	62
4.2.7	Calibration piece.....	64
4.2.8	The computer and the software.....	72
4.3	Design of the hand model.....	72
4.4	Distribution of the cameras around the measurement volume.....	73
4.5	The test rig.....	76
4.6	The procedure for measuring the angles of finger joints.....	78
4.6.1	Calibration process.....	78
4.6.2	Measurement of the angles of finger joints.....	83
4.6.3	Correspondence problem.....	85
4.6.4	Automatic matching of the measurement points.....	86
4.7	Calculation angle between two vectors in space.....	96
4.8	Summary.....	99
5	Experimental work and results.....	100
5.1	Summary.....	100
5.2	Experimental work.....	100
5.2.1	Calculating the FOV angle.....	102
5.2.2	Calculating the horizontal and vertical distances in which the calibration piece can be shifted.....	103
5.3	Measuring the 3D world coordinates for the calibration piece by CMM.....	105
5.4	Calibrating the cameras using DLT.....	106
5.5	Reconstruct the 3D coordinates of the calibration points using DLT.....	109
5.6	Determining the effect of the error in X, Y, Z coordinates on the value of finger joint angles.....	115
5.7	Measuring all 4 fingers of the hand.....	124
5.8	Measuring fingers joints based on the Tsai technique.....	141
5.9	Measurements of a real hand.....	145
5.10	Conclusion.....	148
6	Measuring the full range of finger motion.....	150
6.1	Extension and flexion range of the fingers.....	153
6.2	Design of the new system.....	154

6.3	How to select the relative measurement points for the front cameras	160
	Figure joints	162
6.4	Conclusion	166
7	Discussion, conclusion and further work	168
7.1	Sources of error	174
7.2	Further work	175
	References	180
	Appendix A	187
	Appendix B	217

1 Introduction

1.1 Problem identification

The range of motion (ROM) at a joint in a limb is considered by many clinicians as a measurable and definable entity and is therefore often used as an assessment measure. Joint angle measurements are also used as an indicator for change and as an outcome measure to evaluate the result of medical, surgical and other therapeutic remedies (Ellis, *et al.*, 1997).

In most physiotherapy departments and orthopaedic clinics, medical staff often assesses the range of motion of a joint using traditional methods of measurement including visual estimation, goniometry, composite finger flexion to distal palmar crease (see Ellis, B., *et al.*, 2002 for this method) and wire tracing or more sophisticated techniques such as a goniometric glove and electronic devices (Williams, *et al.*, 2000 and Sturman, *et al.*, 1994).

Reliable and valid measurement of the active range of motion of hand joints is one of the most important factors in evaluating the outcome from injury, treatment, or disease of the hand (Bainbridge, 2000). It is normally required to evaluate three types of finger joints which are metacarpophalangeal (MCP), proximal interphalangeal (PIP) and distal interphalangeal joints (DIP), see Figure 1-1.

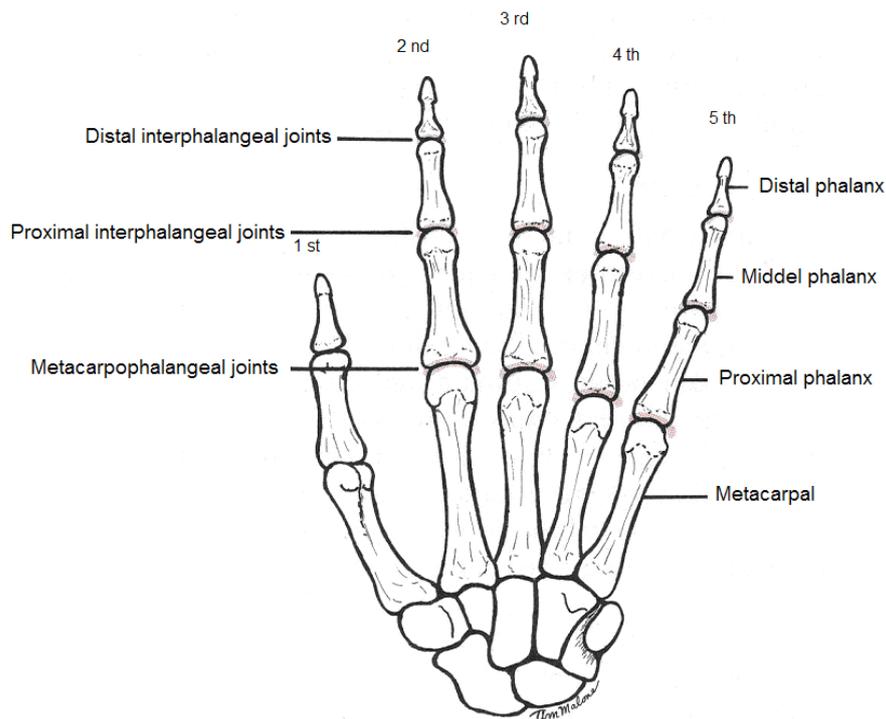


Figure 1-1: An anterior palmar view of the hand (Norkin, *et al.*, 2003)

When the motion of the hand needs to be assessed, the assessor should measure the finger joints, when it bends towards the palm of the hand (flexion motion) and when the finger moved towards the back of the hand (extension motion), see Figure 1-2, and with three joints per digit that means 30 measurements should be obtained for each hand, making a total of 60 readings for both hands. However, due to the difficulties and time expenditure needed to obtain the range of motion, it is uncommon for patients to have their full range of finger motion measured except in legal cases. The doctor or surgeon uses general words such as; the motion is “improved”, “worse” or “same as last time”, as a result of lack of accuracy of the devices used.

The present manual measuring tools produce readings which have an accuracy of no more than $\pm 5^\circ$ and the readings are generally rounded to the nearest 5° or 10° (Bainbridge, 2000).

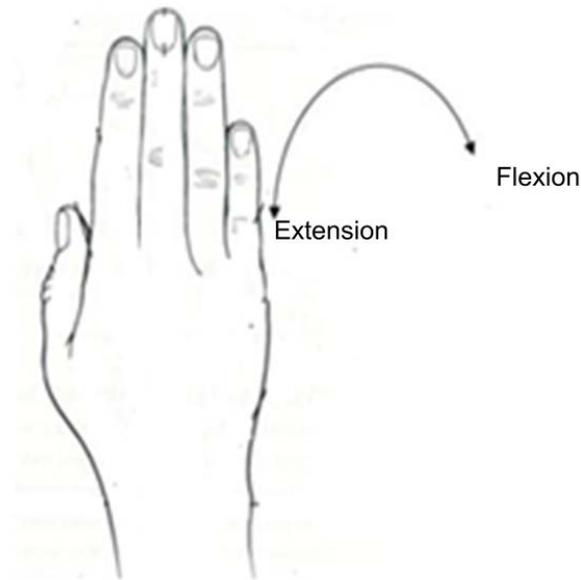


Figure 1-2: Extension and flexion motion of the fingers (Norkin, *et al.*, 2003)

1.2 The thesis motivation and objective

This work was started after consultant surgeon Mr L.C. Bainbridge from Derby Royal Infirmary suggested that if a system that allowed automatic and rapid measurement of the range of motion of the hand could be designed, it would have great value in medical applications. In other words, for the first time doctors would be able serially to assess the improvement of the motion of the hands without difficulties and in a short period of time. Moreover, in legal cases requiring assessments for compensation, the results could be stored in the machine and retrieved if needed as evidence in court (Bainbridge, 2000). They would be also an accurate and repeatable, quantitative method of evaluating treatment regimes.

Therefore, this work is a combination of engineering science represented by 3D computer vision technology and medical science represented by measuring the finger joint angles.

The ultimate scope of the work is to contribute to the knowledge in the field of hand motion assessment. The objective is to design, implement, and validate a computer vision system, which can be utilised in orthopaedic and rehabilitation clinics instead of the existing traditional devices, and help the assessor to have a precise knowledge and assessment of the motion of the patient's fingers after treatment or surgery and to save the assessor's time. Moreover, using this vision system, the data can be graphically displayed and easily recorded for future requirements.

Also, we are aiming to get the process time below 1 minute for a single hand and the angular accuracy down to better than ± 1 degree.

1.2.1 Potential advantages of the proposed system

The proposed system is expected to have the following advantages over existing hand measurement devices.

- As the system will be based on computer technology, the measurement process is expected to take a short time and to produce good accuracy.
- Data will be recorded and retrieved more easily (data on all digits of one or both hands could be integrated).
- In case of legal issues, the outcome of an operation or treatment could be evaluated precisely, so the correct amount of compensation could be determined.
- The proposed system should be user friendly as much as possible in order to be operated by a wide range of medical staff after simple training.
- The system will help physiotherapists to follow up the progress on the screen via a graphical display.
- The system will not include any mechanical or moving parts, so it will not need regular maintenance. However if the cameras are moved from their positions, the system will need to be calibrated again.

- The size of the proposed system will be less than 1m^3 in order to occupy a small volume in the clinic, and could be easily moved around.

1.3 Thesis methodology

In this work the finger joint angles will be measured using computer vision technology and the procedures below will be followed:

- Collecting all the information and equipment for the construction of the computer vision system.
- Designing the calibration piece, according to the information which has been found in the literature and which suits the system and gives the highest accuracy.
- Measuring the direction vectors of the bones in each finger of the hand in 3D space, and then deducing the angles between these fingers. To do this, the finger skin of the patient will be marked at intervals along the line of the bone, and then these marks will be located in 3D space.
- The vector along the skin is assumed to be parallel to the centre of the bone line and that the bones are straight.
- In order to measure all finger angles, several simultaneous views of the hand will be required.
- A plane mirror will be added to the computer vision system to enable capturing images of the fingers at all positions.
- A model of a human hand will be designed and manufactured to use as a test piece for measuring all the finger joints before measuring the real hand.
- A Coordinate Measurement Machine or CMM will be used to measure the 3D world coordinate system (X, Y, Z) for the calibration points and to compare the angles of finger joints obtained from the vision system.

The data from the CMM will be compared with that obtained from the computer vision system.

- Calibrating the cameras will be accomplished by using two popular calibration techniques.
- The image in a computer environment is represented as a matrix, and the techniques used to calibrate the cameras involve matrices solving. As MATLAB is matrices based software package, it will be used to manipulate the images and in writing the code for this study.

1.4 Thesis Outline

A broad range of topics is covered in this thesis. An overview of the chapters is given in this section.

Chapter 1: an introduction to the problem, the motivation for this study and the proposed method and strategy for solving the task.

Chapter 2: gives an idea about the anatomy of the human fingers, and type and range of motion they can achieve. Also it includes a survey covering broadly most previous devices used and under research for evaluating finger joint angles. The advantages and the drawbacks for these tools are also reviewed. The last part of this chapter discusses the elements of the computer vision system and the function of each element.

Chapter 3: discusses how the computer vision system is used to extract 3D information from 2D images. Camera calibration, which is important in the 3D computer vision environment, is also explained. The two most popular camera calibration techniques, DLT and Tsai, are introduced and reviewed.

Chapter 4: describes the components of the computer vision system for this study, and how to design a suitable calibration piece in order to increase the

efficiency of the computer vision system. It also illustrates the procedures that are used to calibrate the cameras and to measure the finger joint angles of a human hand model. The correspondences between the measurement points for all images are also explained.

Chapter 5: describes the mathematical model which is obtained to examine the effect of the shift in X,Y and Z coordinates from the correct hand positions on the measured angles of the finger joints. Also in this part of the thesis; the computer vision system is tested. A hand model is designed and built to measure the angles of the finger joints. The results from the measurement process are discussed. Finally, to examine the system for measuring real fingers, the measurement process is implemented on two joints of a real human hand.

Chapter 6: in chapter five, the measurement process involved assessment of finger joints by placing cameras above the hand. This position of the cameras does not allow images of the measurement points at all flexion positions to be captured. In chapter six, the system is modified so it becomes able to measure the hand movement at all positions.

Chapter 7: summarizes the contribution and draws the conclusions of this research and gives a short outlook into the future.

2 Literature review

2.1 Anatomy and the motion of the fingers

The human finger consists of three joints as illustrated in Figure 1-1, the distal interphalangeal joint, abbreviated as DIP, the proximal interphalangeal joint, abbreviated as PIP and the metacarpophalangeal joint, abbreviated as the MCP. These joints have two types of motion, the primary motion being in the flexion-extension plane, Figure 2-1. Abduction and adduction are limited and occur only at the MCP joints, see Figure 2-2.

The values of the flexion-extension motion vary from one joint to another. Table 2-1 (Norkin, *et al.*, 2003) provides a summary of typical ranges of motion values for the MCP, PIP, and DIP joints of the fingers, for a healthy human (adult) male. Although these values were reported by different sources, it can be concluded that the PIP joints have the greatest amount of flexion followed by the MCP and DIP joints. The MCP joints have the greatest amount of extension.

Table 2-1: Finger motion in degree from selected

Joint	Motion	AAOS (Degrees)	AMA (Degrees)	Hume,M. 1990 (Degrees)	Mallon, W. 1991 (Degrees)
MCP	Flexion	90	90	100	95
	Extension	45	20	0	20
PIP	Flexion	100	100	105	105
	Extension	0	0	0	7
DIP	Flexion	90	70	85	68
	Extension	0	0	0	8

AAOS = American Association of orthopaedic surgery.
AMA = American medical Association.

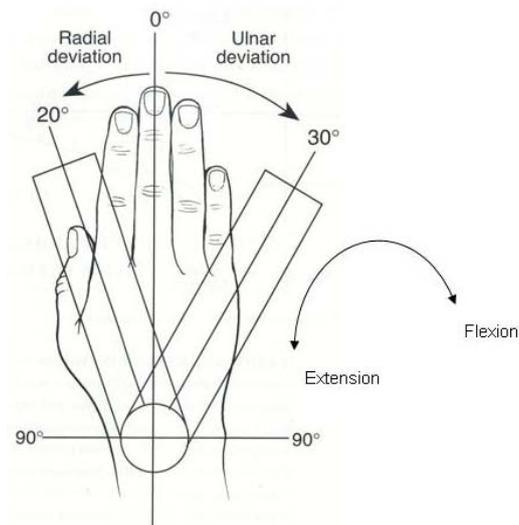


Figure 2-1: Flexion-extension motion (Greene, *et al.*, 1994)

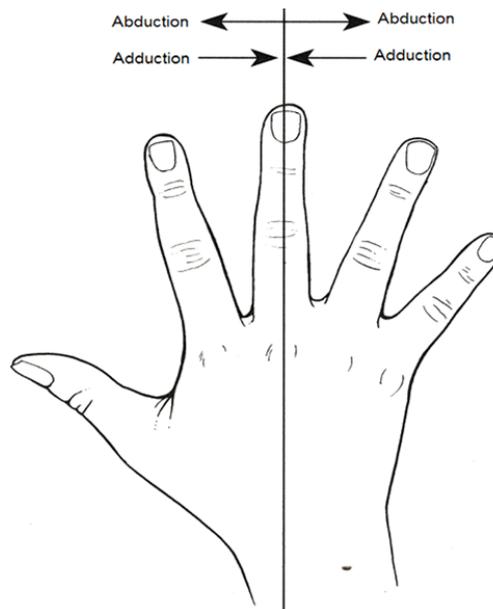


Figure 2-2: Abduction / adduction motion (Greene, *et al.*, 1994)

2.2 Existing finger joints measurements devices

Accuracy of a measurement refers to the difference between the quantity as measured and its true value (Thomas, 2000), and accurate measurement of

active range of motion of finger joints is very important for hand assessment (Macdermid, *et al.*, 2001 and Dipietro, *et al.*, 2003) in case of injuries or treatments.

Precision or repeatability of a measurement is defined as the ability of a device to reproduce the same measurement over and over again under the same conditions (Thomas Y., 2000). A precise assessment allows disease progression and response to therapeutic interventions to be assessed (Rose, *et al.*, 2002). In orthopaedic clinics the assessments are usually done by visual estimation or by means of simple tools, for instance wire tracing, traditional goniometry, or measurement of composite flexion and span (Low, 1976). However, some departments use more sophisticated techniques such as the goniometric glove and other electronic devices (Williams, *et al.*, 2000, Sturman, *et al.*, 1994).

2.2.1 Visual estimation

Visual evaluation was the first method used and is still commonly employed in clinical assessment of joint function (Salter, 1995). It is a quick, easy method, does not require any tools and can be used for all joints. The assessor looks at the joint range and visualises it against an imaginary protractor and the range expected in a healthy joint.

There are potential sources of error when using visual estimation method. For instance, if the evaluator is inconsistent and views the joint from different angles, results may be ambiguous (Simpson, 2002). Rose, *et al.*, 2002, demonstrated that this method is an inaccurate and unreliable technique of obtaining finger range measurements, and should only be used when other methods are unavailable.

2.2.2 Tracing onto paper

Getting the digit joint angles by this method is done with a piece of paper which is placed between the fingers, and the profile of the bent finger is traced onto the paper using a pencil (Ellis, *et al.*, 1997). The angles are then measured from marks made on the paper. This measurement method does not include the MCP joint and is very time consuming.

2.2.3 Wire tracing

Wire tracing is another traditional method in which solder wire is laid along the back of the hand and bent to conform to the angle of each joint. The wire is then lifted off the finger and laid on paper, then the angles are measured by tracing this wire onto a piece of paper (Ellis, *et al.*, 1997), This method is particularly useful for painful hands as there is less need to repeat movements. Wire tracing is a cheap method, and the whole hand can be recorded on one sheet of A4 paper (Simpson, 2002). However it has a low accuracy shows limitations as a reliable assessment tools and is slow (Ellis, *et al.*, 1997).

2.2.4 Composite finger flexion or pulp to palm measure

In this method as the name suggests, a composite measure of finger range of motion, measured as the distance of the finger tip to a fixed point in the palm, is made by using a ruler (Ellis, *et al.*, 2001), see Figure 2-3.

Composite finger flexion is a quick, simple and cheap method and provides a measure which therapists and patients can use as an indicator of progress (MacDermid, *et al.*, 2001). This method measures the distance of the digit from the palm so it reflects the functional ability to form a fist. The disadvantage of the Composite finger flexion method is that, change in range of motion does not

reflect an improvement or deterioration at a specific joint, unlike the measurement carried out using goniometry (Ellis, *et al.*, 2001).

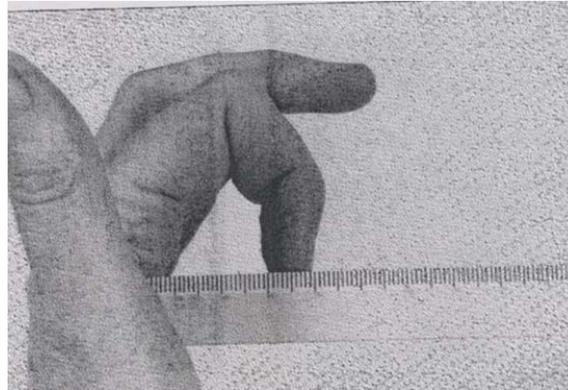


Figure 2-3: Technique for measuring range of motion by using composite finger flexion (Ellis, *et al.*, 2001)

2.2.5 Universal Goniometry

The goniometer is one of the primary measures of hand function and involves measurement of the range of motion of the finger joints with a metal or plastic protractor (Weiss, *et al.*, 1994). The two-arm goniometer is still the most widely used, most economical, and most portable device for evaluation of range of motion. The device consists of two connecting arms; one of these arms remains fixed but the other moves (Lea, *et al.*, 1995). There are many types of goniometer available on the market. Figure 2-4 shows some of them.

A goniometer is a simple measurement device and is a non-invasive method of quantifying the range of motion of joint angles which can be measured in a standing position or in flexion or extension (Jaegger, *et al.*, 2002). However, measurement of total finger flexion by goniometry involves the measurement of flexion at the metacarpophalangeal, proximal inter-phalangeal and distal interphalangeal joints of a finger and adding these together. Total active motion is the total active flexion minus any loss of extension (extension lag) at the three finger joints. Therefore, total active motion measurement requires six separate

measurements for each digit (MacDermid, *et al.*, 2001), which is very time demanding.

Although traditional goniometry may be clinically acceptable, it is a static measure of dynamic function and can only be used to assess one joint at a time and the accuracy is no more than $\pm 5^\circ$. Additionally, the position of proximal joints can influence the range of motion at the measured joint (Williams, *et al.*, 2000). Subject error in finger joint measuring tends to be more complex than large joints, due to the large number of joints in a small space (Weiss, *et al.*, 1994).

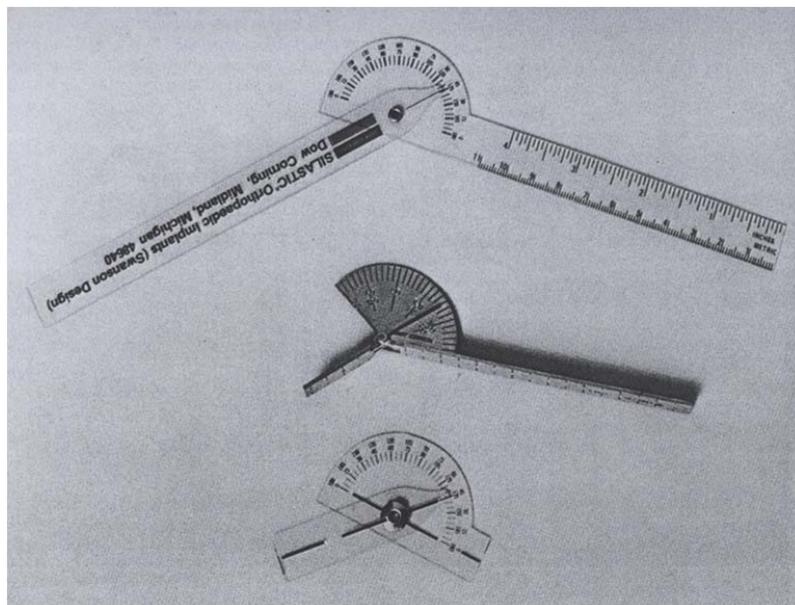


Figure 2-4: Universal goniometer (Cambridge, 1995)

Another limitation of traditional goniometers arises because evaluating ROM simultaneously from all the hand is not easy, as the whole measurement process is tedious and time demanding for the assessor and the patient (Dipietro, *et al.*, 2003). Furthermore, if the fingers are injured the measurement process can cause discomfort to the patient.

The traditional goniometer can be affected by several errors such as inexperience of the assessor (increase in error from not adhering to the standard

measurement technique), and instrument error which arises from using improper size of the instrument, for instance by using a large goniometer to assess a small bones or vice versa (Dipietro, *et al.*, 2003). Also the recording of data is performed manually.

2.2.6 Inclinometer or gravity-dependent goniometer

The inclinometer offers an alternative to the universal goniometer for measuring the motion of the joint (Lea, *et al.*, 1995). The inclinometer uses gravity's effect on pointers and fluid level to measure joint position and motion (Norkin, *et al.*, 2003). There are two major types of inclinometers, mechanical and electronic. The mechanical inclinometers are based upon a weighted pendulum or a fluid level. The pendulum weighted inclinometer has a starting position indicated by the pendulum. In the fluid type the fluid level specifies the horizontal position (Lea, *et al.*, 1995 and Gerhardt, *et al.*, 2002). Figure 2-5 illustrates different examples.

The electronic inclinometer has a specially programmed software and recording system and must be recalibrated horizontally or vertically against known vertical or horizontal system (Lea, *et al.*, 1995).

The main advantages of the inclinometer are that it enables the assessor to identify the same starting position on successive measurements, because gravity does not change (Gerhardt, *et al.*, 2002).

Inclinometer measurement is time consuming for assessing the finger joints as it measures one joint at a time. Moreover they are difficult to use on small joints of the fingers where there is soft tissue deformity or oedema (Clarkson, 2000 and Moore, 1978).

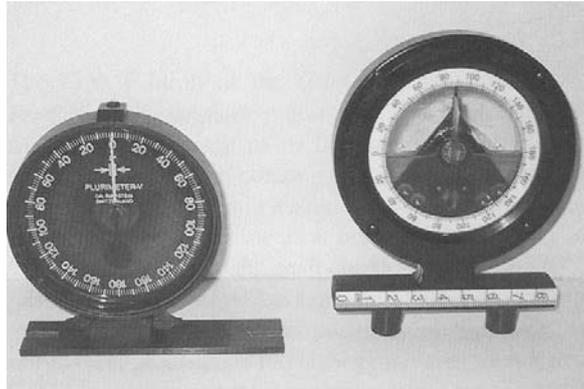


Figure 2-5: Mechanical (on the left) and fluid level inclinometers (Gerhardt, et al., 2002)

2.2.7 Electrogoniometer

The electrogoniometer is an electrical potentiometer that can be used to measure the joint angles. Changes in joint position cause the resistance of the potentiometer to vary, and the voltage output from the potentiometer can be calibrated so that the measured voltage can be read as a joint angle.

Electrogoniometers are expensive and take time to calibrate so they are often used in research rather than in clinical assessment.

2.2.8 HandMaster

The HandMaster is an exoskeleton-like device worn on the fingers and hand. Using sensors (potentiometers) at the finger joints it can accurately measure the three joints of each finger as well as the complex motion of the thumb (Sturman, *et al.*, 1994).

The MCP, PIP and DIP joints for the index finger have been measured by the HandMaster and the goniometer. The results showed that the HandMaster has better reliability and greater sensitive compared with the goniometer (Weiss, *et*

al., 1994). However, this type of measurement devices takes time to be fixed on the hand and is difficult to use if the hand is seriously injured.

2.2.9 Glove technology

Typically, the therapist performs the hand measurements via a simple plastic or metal goniometers (Rand, *et al.*, 1993), which must be placed on each finger joint to evaluate the flexion and extension angles (Dipietro, *et al.*, 2003). Due to the limitations of the mechanical goniometers as mentioned in section 2.2.5 above, Glove technology has been of interest to researchers over the past few years. The gloves measure the range of motion of the finger joints as the hand is flexed and extended (Sturman, *et al.*, 1994).

The measurement process using gloves can be automated, which leads to less time demand for both the therapist and the patient. This in theory could reduce the costs of rehabilitative treatment. Another advantage is that the glove allows dynamic and simultaneous recording of all hand joints which is not achievable with the instruments used at present for the same purpose (Dipietro, *et al.*, 2003).

There are many gloves that have appeared in the literature and in the market place, such as the Sayre glove, DataGlove, Led glove, and the Mattel Power glove, which is based on acoustic tracker technology (Sturman, *et al.*, 1994), and an instrumented glove (Rand, *et al.*, 1993).

The Gonimetric glove shown in Figure 2-6 is composed of a light, flexible grade of Lycra Flexion sensors, placed over the DIP, PIP, and (MCP) joints of the fingers. The accuracy of the glove is within the limits of traditional goniometry. The electronic glove may reduce or remove observer bias noted with traditional measurement devices. In addition, multiple joint angles can be recorded at any one time which will speed up the measurement process, which is an advantage over traditional goniometry (William, *et al.*, 2000). However, gloves cannot be used if the injury to the hand includes swelling, bandages, splints, etc.



Figure 2-6 : The completed goniometric gloves, small and medium size (William, *et al.*, 2000)

Table 2-2: Comparison between existing finger joints measurements devices

	Assessment time	Accuracy	Cost	Repeatability	Resolution	Drawback
Visual estimation	Very short	inaccurate and unreliable	N/A	Low	data not available	Not reliable
Tracing on paper	Time consuming	Low	Cheap	Low	data not available	Does not include the MCP
Wire tracing	Slow	Low	Cheap	data not available	data not available	Not reliable
Universal Goniometry	Time consuming	Low ($\pm 5^\circ$)	Cheap	5 degrees ⁴	data not available	Assess one joint at a time
Goniometric Glove	quick	5.6 degrees ¹	Expensive	± 2.3 degrees ¹	1 degree ¹	Does not suit swelling or injured hands
Inclinometer	Time consuming	± 2 degree ²	Reasonable	data not available	0.01 degree ²	Measures one joint at a time
Electrogoniometer	Time consuming	5-7 degrees ³	Expensive	1 degree ³	data not available	Measures one joint at a time

1) Wise, *et al* (1990).

2) Wikipedia (2013).

3) Thomas, *et al* (2000).

4) Macionis (2013).

All the finger joint measurement techniques explained in section 2.2 above are based on direct contact measurement i.e. the device must be kept in contact with the fingers. Other approaches that are based on non-contact measurement can be used to evaluate the range of motion of the hand, by using optical devices.

There are two common non-contact measurement methods, the first uses markers on the object such as small infrared-reflecting points, flashing infrared LEDs or any other material which glows under some circumstances. Two or more cameras distributed around the subject capture the markers in their fields of view. Software matches the markers from the multi-cameras and Three-dimensional (3D) coordinates for each mark can be obtained using different reconstruction techniques (Sturman, *et al.*, 1994). In this work, a similar non-contact measurement system was designed known as a computer vision system.

The second non-contact measurement approach uses one camera and a mirror, so the camera capture two images of the same object one image of the real object and the second through the mirror. Then the image coordinates of each pair of symmetric points of the two images are used to reconstruct the 3D coordinates of the object (Sturman, *et al.*, 1994 and Zhang, *et al.*, 1998).

2.3 Three-dimensional Computer vision system

2.3.1 What is a Three-dimensional computer vision system?

The Three-dimensional (3D) computer vision system is a system in which a computer understands or extracts features in 3D scenes from visual input (Shirai, 1992, Leta, *et al.*, 2006). The visual input may be a monocular image, a range image or time sequence images.

A computer vision system recovers useful information about a scene from its two-dimensional projections. Since images are a two-dimensional projection of

the three-dimensional world, the information is not directly available and must be recovered (Jain, *et al.*, 1995) by means of software analysis.

A major problem associated with 3D computer vision is that a 3D scene may include many objects which look different depending on the viewer direction, illumination conditions, and geometrical relations among objects (Shirai, 1992).

2.3.2 Why use a computer vision system?

Computer vision technology is important in many industrial applications, and it improves productivity and quality management and provides a competitive advantage to industries that use this technology. These include electronic manufacturing, medical diagnosis, virtual reality systems, glass making, food production, automated assembly, automotive manufacturing, pharmaceutical and medical, container and packaging and so on (Kerr, 2003, Lin, *et al.*, 2003).

With the advent of low-cost computational hardware, computer vision systems have emerged as financially viable devices in automated manufacturing and measurement systems (Lin, *et al.*, 2003, Chen, 2002).

Computer Vision offers accuracy, consistency, and repeatability, in contrast to the subjectivity, fatigue, slowness, and cost associated with human inspection (Leta,*et al.*, 2006). The advantages of using a machine vision system for assessment include a decrease in the time required for measurement as well as greater accuracy of non-contact measurements and better flexibility than the conventional methods (Chen, 2002).

2.3.3 Factors that affect the performance of the computer vision system

As mentioned above the computer vision system has many advantages. However, it requires a lot of effort and cost to develop a reliable 3D measuring

system (Lin,*et al.*, 2006).The essential problem of using computer vision techniques consists in image quality, as image analysis requires the features of interest be well defined. The choice of the most suitable method to pre-processing and threshold the image into its two main components (the object and the background), must be sufficiently robust to obtain images without information loss (Leta, *et al.*, 2006).

Like the human eye, vision systems are affected by the level and quality of illumination. By adjustment of the lighting, the appearance of an object can be changed with the feature of interest clarified or blurred. Therefore the performance of the illumination system can affect the quality of the image and plays an important role in the over all efficiency and accuracy of the system (Brosnan, *et al.*, 2004).

Also, computer vision systems are subjected to human error during the calibration process, as the centre of the calibration points measured by a machine, such as a coordinate measurement machine, may not coincide with the centre measured by the vision system.

Other sources of errors result from electronic devices, for instance the camera analogue signal connected to the computer through the frame grabber. During the sampling process, mismatch between the pixel locations in the camera sensor and those in the sampled image may occur (for more details see(Lenz, *et al.*, 1990 and Horn, 2000)).Moreover cameras often have noticeable geometric distortions caused by the optical system that can affect the accuracy of the computer vision system (Tsai, 1987).

2.4 Some applications of computer vision systems.

3D Computer vision systems have many applications in medical and industrial sectors. These applications include automatic quality control or inventory

management systems, computer-assisted surgery, security surveillance, as well as, 3D human tracking in sport and medical science and 3D medical images (Henderson, 2003).

Ringer, *et al.*, (2000) proposed a tracking technique and 3D reconstruction of human location. The scheme uses a number of video cameras and the aim is to obtain complete and accurate information on the three-dimensional location and motion of the bodies over time, see Figure 2-7.

The system uses markers placed at the joints of the arm(s) or leg(s) being analysed. The location of these markers on each camera's image plane provides the input to the tracking system, see Figure 2-8. It has many applications in medicine, sports analysis and motion capture for animation.

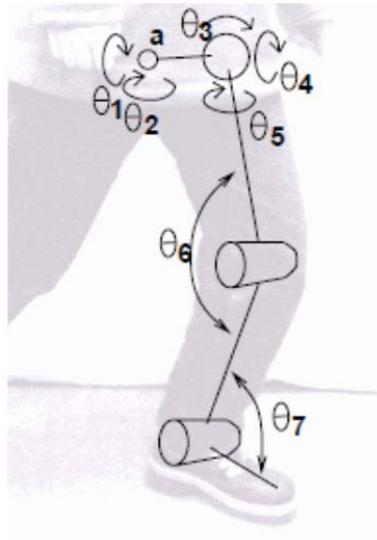


Figure 2-7: Parameters used to describe the position of a leg (Ringer, 2000)

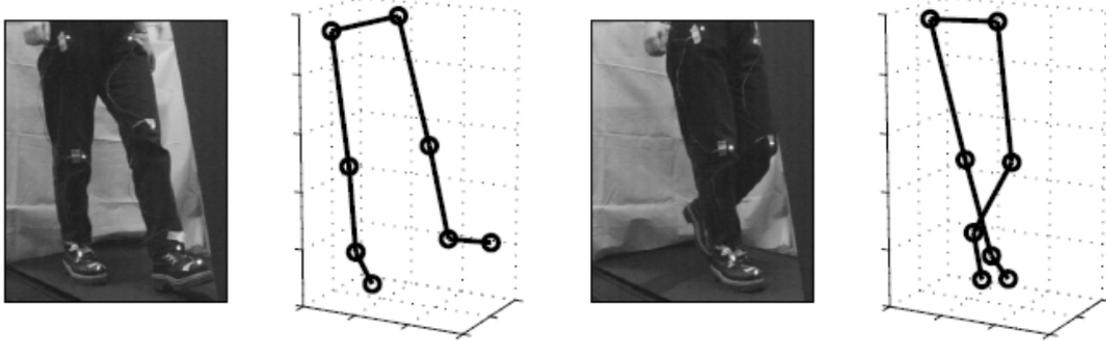


Figure 2-8: Example of the proposed system tracking the legs of a person Walking (Ringer, 2000)

Aguilar, *et al.*, (2005), developed a stereo system to measure free-form surfaces of railway concrete sleepers (Figure 2-9) and calculate track and rail seat dimensional tolerances, to replace Digital measurement of concrete sleeper dimensions using digital verniers and data loggers. The system consists of CCD cameras which calibrated using the calibration object shown in Figure 2-10. During the calibration, the geometry, position and orientation of the cameras are calculated using the Tsai non-coplanar method.



Figure 2-9: Railway concrete sleepers (Aguilar, *et al.*, 2005)

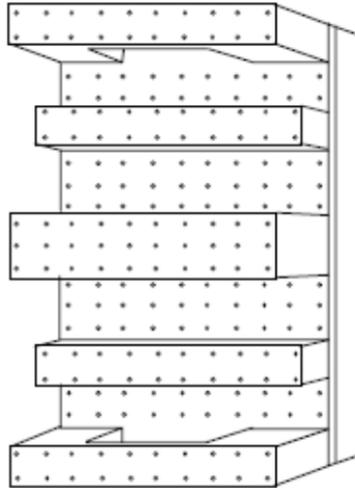


Figure 2-10 : Calibration object used by Aguilar, *et al.*, (2005)

To obtain the 3D coordinates of the measured points the triangulation technique was used. For evaluating the accuracy, the same points are measured with a coordinate measurement machine (CMM) and compared with the obtained results.

Bin, *et al.*, (2010) used Direct Linear Transformation algorithm to reconstruct 3D of femur from biplanar radiography. Firstly the camera parameters were obtained using DLT technique. To produce the 3D information of the femur, they do not use and markers, however this method is based on the principle that the anatomical landmarks can be found in the 2D contours of the two images.

Two X-ray radiographs (antero-posterior (AP) and lateral (LAT)) of the femur were used to reconstruct the 3D of the femur. To process these images, the median filter was applied to filtering out the noise, and then canny edge detection was used to detect the edges, the output image shown in Figure 2-11. Figure 2-12 shows the 3D of the femur.

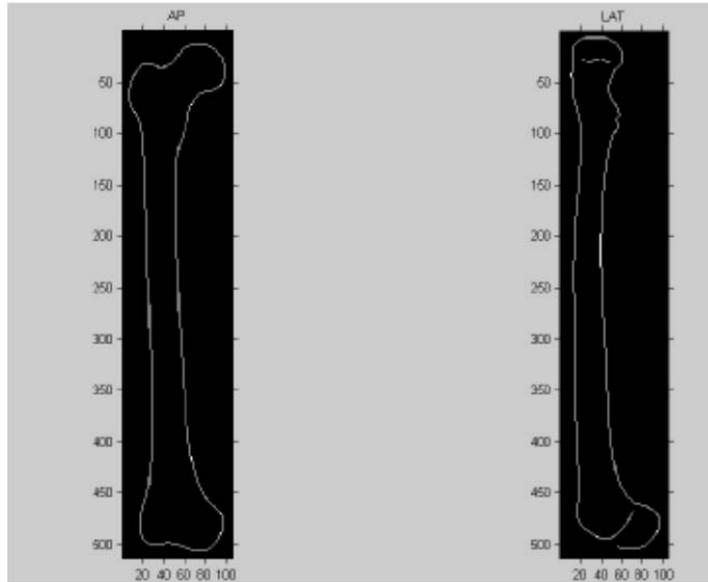


Figure 2-11: The final results of edge detection of two X-ray images of the femur (Bin, *et al.*, 2010)

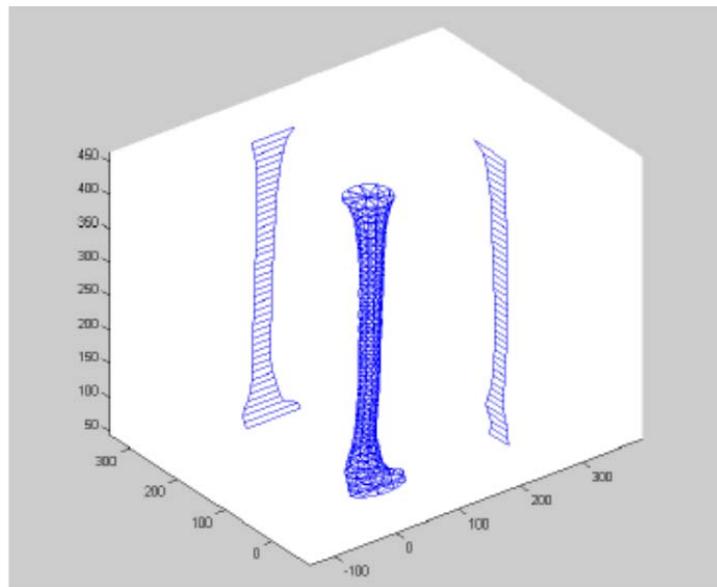


Figure 2-12: 3D femur (Bin, *et al.*, 2010)

Instead of using multiple cameras, and to reconstruct 3D information, Zhang, *et al.*, (1998) utilized a single camera and a plane mirror. Thus the object and its image in the mirror will form a bilaterally symmetric structure. Then the image coordinates of each pair of symmetric points are used to reconstruct the 3D

coordinates of the object. For finding the intrinsic and extrinsic camera parameters, Tsai's algorithm was adopted

Figure 2-13 (a) shows the image of a cardboard box in front of a plane mirror and Figure 2-13 (b) shows the shape of the box determined by the six 6 vertices of the box that are visible both in and outside the mirror.

Due to limitation of the geometry, only a fraction of an object and its corresponding image in a plane mirror can be simultaneously visible in a single view. This fraction is usually less than 1/2. In order to reconstruct a large part of the object, two or more mirrors may be placed at different orientations and locations, so a large part of the object can be reconstructed.

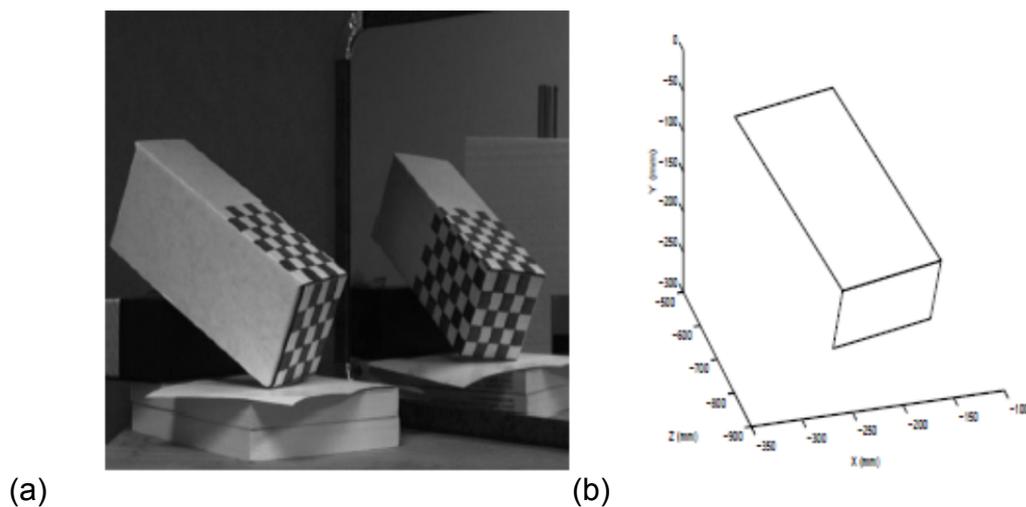


Figure 2-13: The cardboard box and the house model and their recovered structures (Zhang, *et al.*, 1998)

2.5 Elements of a computer vision system

A computer vision system (Figure 2-14) consists of all the elements necessary to obtain a digital representation of a visual image, to modify the data, and to present the data to the external world. In an industrial environment, the system may seem complex due to all the associated manufacturing process equipment

used. However the complexity is reduced when the vision system is divided into three main components (Galbiati, 1990).

- Image acquisition
- Processing
- Output or display

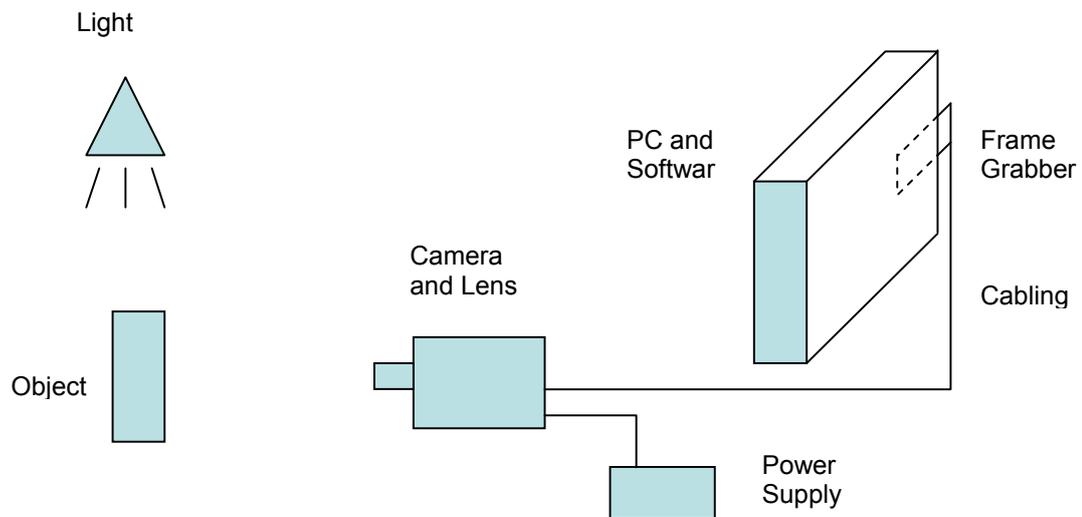


Figure 2-14: Computer vision system

2.5.1 Image acquisition

Image acquisition converts the image of a physical object into a set of digitized data which can be used by the processing unit of the system. The acquisition function can be regarded as consisting of four phases.

- Illumination
- Image formation and focusing
- Image detection
- Camera and output signal

2.5.2 Illumination

One of the complications that have troubled the computer vision system design is the variability of an object's appearance from one image to another. With slight changes in lighting conditions and viewpoint often come large changes in the object's appearance (Belhumeur, *et al.*, 1998).

How the object appears in a computer vision system can be significantly affected by the illumination. By carefully choosing type, position and direction of the illumination light source, the contrast between the features in the object and the background can be enhanced, thereby simplifying the automatic computer vision task. Light always makes the difference between the success and failure in a vision system (Titus, 2001).

A lighting system must deliver as much light as evenly as possible to the units in the camera's field of view (Masi, 1998). However, uneven illumination, in principle, sometimes can enable the vision system to distinguish between or recognize different objects in the images.

Many industrial computer vision applications in the past have used visible light since the source was available and the application frequently was the automation of a manual inspection task. There are three types of visible lamps most frequently used in computer vision applications: fluorescent, tungsten-halogen, and LEDs. However, the use of illumination outside the visible range, such as ultraviolet, x-ray, and infrared is increasing because of the need to achieve special tasks which are not possible with visible light.

Environmental illumination can affect all lighting methods by changing the total level of illumination on the object which appears as noise in the data. The affect of environmental illumination can be reduced by the use of light shields and barriers which reduce or prevent the amount of stray radiation entering the lens, also by using of automatic thresholding in software.

2.5.3 Image formation and focusing

The image of the scene is focused on the sensing element with a lens, in a way similar to that in a photographic camera. The difference between the photographic camera and the computer vision system is that a camera uses film, and a computer vision system uses a sensor instead, to capture the image. The sensor converts the image to an electrical signal.

The computer vision camera is usually specified separately from the vision system as the capability and feature requirements are dependent on the application. In addition the camera lens specification must be given as it is the element which adapts the camera to the specific task. There are four important parameters connected with the optical lens of the vision system.

1. Magnification
2. Focal length
3. Depth of field
4. Flange back length (Lens mounting)

Magnification (m) is a measure of the relative size of the object to the size of the image formed on the sensor of the camera. The magnification value will be less than 1 in the case of conventional industrial applications, since the dimension of the detector is smaller than the object being viewed. In the case of microscopic application the value of m will be greater than 1 (Burger, *et al.*, 2009 and Bueche, 1977).

The focal length of a lens is the distance along the optical axis from the lens to the focus (or focal point). The f-number (focal ratio) expresses the diameter of the diaphragm aperture in terms of the effective focal length of the lens. The greater the f-number, the less light per unit area reaches the focal plane of the camera sensor (Demant, *et al.*, 1999).

The depth of field is a function of aperture size, magnification, and size of the sensor elements. A smaller aperture provides more depth of field but admits less light; a larger aperture admits more light but reduces the depth of field. The smaller the pixel areas of the sensor the smaller the depth of field (Nelkon, *et al.*, 1979 and Demant, *et al.*, 1999).

The flange distance, also known as lens mount, is the distance from the back of the lens flange to the sensor plane. There are two types of flange back lengths:

C-mount- distance from the back of the lens flange to the sensor is 17.526 mm

CS-mount- distance from the back of the lens flange to the sensor is 12.526 mm (Demant, *et al.*, 1999).

2.5.4 Image detection

In order to acquire a digital image a device should be sensitive to the energy radiation that is reflected or scattered by the object that is being imaged.

The lens focuses the image of the physical object onto the sensor. The sensor element produces an electrical signal representing the visible image. The digitizer then converts the output of the sensor into a digital signal which can be understood and processed by a computer. The amount of electrical output from the sensor is proportional to the light intensity, and the digitizer changes this analogue signal into digital form (Galbiati, 1990).

2.5.5 Solid state camera and output signal

Solid state cameras such as charge coupled device (CCD), charge injected device (CID), or multiplexed photosensor are the most widely used devices in computer vision systems, because they are relatively cheap, reliable and rugged. The camera works as transducer and all image processing happens after the

camera video signal is sampled and digitized to obtain a matrix of numbers in the frame buffer (Wittels, *et al.*, 1989).

The basic concept behind the solid state camera is that a separate electrical signal is produced for each pixel element in the sensor. The sensor can be arranged in either a linear or in a rectangular array. The output of the sensor is a series of voltage pulses representing the light intensity at the pixel location. Solid state cameras are not subject to blooming and flare, virtually no geometric distortion, drift or lag and they are light, rugged and consume little power. Solid state sensors comprise CCD, CID, CMOS (complementary metal oxide semiconductor), and CPD (charge priming device) (Galbiati, 1990).

The charge couple device or CCD is a sensor based on semiconductor technology. CCDs have become the sensor of choice because they do not suffer from geometric distortion and their response to incoming light does not weaken with long exposure times.

A CCD consists of an M by N rectangular grid of photosensors which are sensitive to light intensity. Each photosensor can be considered as a very small rectangular black box which converts light energy to voltage. The physical area of the array is typically 6.4 mm × 4.8 mm (for a half inch format sensor).

When the light falls on a CCD, the photosensors accumulate an amount of electric charge proportional to the illumination time and the intensity of the incident illumination. The output of the CCD array is a continuous electric signal, The signal is then sent to an electronic device called a frame-grabber, where it is digitized into a 2D rectangular array and stored in a memory buffer (the final digital storage area for the image shown by a computer display) (Efford, 2002 and Sonka, *et al.*, 1999).

The output of most analogue image sensors is a continuous voltage waveform in which the amplitude and spatial behaviour are related to the physical

phenomenon being sensed. In order to produce a digital image, the continuously sensed data needs to be converted into digital form. This consists of two processes: sampling and quantization.

Sampling and quantization are the processes that convert the continuous image irradiance as it is projected by an optical system onto the image plane into a matrix of digital numbers that can be stored and processed by a computer (Jähne, 2002).

The result of sampling and quantization of the image function $f(x,y)$ is a two-dimensional array of points. A point on the dimension grid is called a pixel or pel. Both words are an abbreviation of “picture element”. The position of the pixel is given by the common notation for a matrix. The first index, m , denotes the position of the row, the second, n , the position of the column, as shown in Figure 2-15.

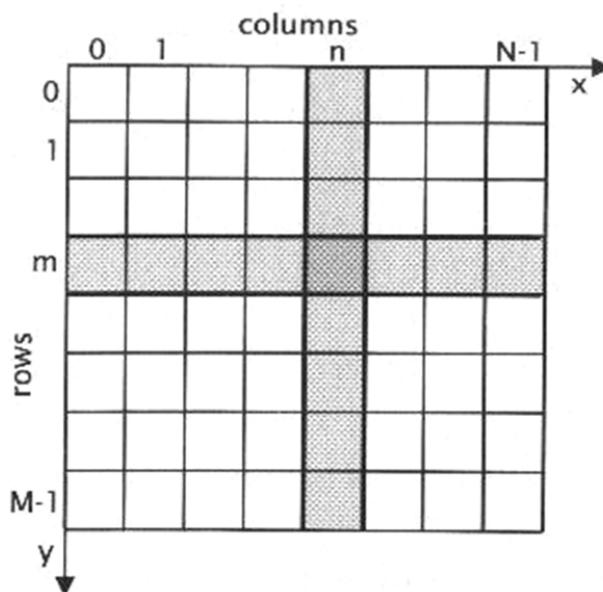


Figure 2-15: Digital image represented by a matrix

If the size of the digital image is M by N pixels, i.e. is represented by an $M \times N$ matrix, the index n runs from 0 to $N - 1$ and the index m runs from 0 to $M - 1$. The vertical axis (y axis) runs from top to bottom not vice versa as it is common in

graphs. The horizontal axis (x axis) runs from left to right. N coordinate values along the first row of the image are (0, 1) and so on until the coordinate values (0, N-1), and the same for the columns (Jähne, 2002).

The digital image can be written in matrix form as:

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, M-1) \\ f(1,0) & \dots & \dots & f(1, M-1) \\ \dots & \dots & \dots & \dots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1, M-1) \end{bmatrix} \quad (2-1)$$

Dense sampling produces a high resolution image in which there are many pixels, each pixel represents the contribution of a very small part of the scene. By contrast coarse sampling produces a low resolution image in which there are few pixels, each pixel represents the contribution of a relatively large part of the scene to the image (Efford, 2002). Image resolution is defined as the smallest number of discernible line pairs per unit distance; for instance, 2540 line pairs per inch (Gonzalez, R, *et al*, 2002).

2.5.6 Image processing techniques

Image processing techniques involve two steps, image enhancement and image segmentation.

a) Image enhancement:

Random and systematic noise can appear in the image data because of the physical limitations of the hardware used for image acquisition. In addition, the image may contain data features which are not of interest or which mask the items of interest. So the image should be enhanced before applying any segmentation technique.

Image enhancement is the process of manipulating an image so that the output image is more suitable than the input image for a certain application. The

enhancement methods are problem oriented. For example a technique that suits the enhancement of X-Ray images which was taken under restricted conditions may not be the best approach for enhancing an image that taken under natural illumination. So choosing the proper enhancement technique depends on the nature of the image under investigation.

Image enhancement is one of the most important issues in image processing and A lot of techniques have been developed.

One of the simplest and most useful image enhancement operations involve the adjustment of brightness and contrast in the image. The usual reason for manipulating these quantities is the need to compensate for the difficulties in image acquisition. Without the aid of image processing, the image may need to be re-acquired several times, adjusting the exposure each time until satisfactory results are obtained (Galbiati, 1990).

b) Image segmentation:

Image segmentation is very essential and critical in many image, video, and computer vision applications. It is used to partition the object of interest from the background, which ideally to identify which part of the data array makes up the objects of interest in the real world. Segmentation supports tasks such as measurement, visualisation, registration and reconstruction, each task of them has specific needs. For example, the demand for accuracy is much higher for measurement than for visualisation. Many segmentation methods have been developed such as edge detection, thresholding, shape detection, nearest neighbour classification and so on.

2.5.7 Thresholding

Thresholding has been a popular tool used in image segmentation for separating objects from background (Tsai, 1995), especially in those images where pixels belonging to the object are substantially different from the grey levels of the pixels belonging to the background.

In the thresholding process, the pixels having a grey scale value at or below a given threshold value are given a zero value (black) and all above the threshold are set at one (white) (Gonzalez, *et al.*, 2002), (Galbiati, 1990).

Many thresholding techniques are available in the literature, (Lin, 2003), and most of these methods can be classified into two categories: global and local thresholding. A global technique finds a single threshold value for the entire image, whereas the local ones use information obtained from a certain size of neighbourhood, or a certain reference domain within the image (Sue, *et al.*, 2003).

a) Global thresholding:

If the background grey level is reasonably constant throughout the image, and if the objects all have approximately equal contrast above the background, then a fixed global threshold (T) will usually work well.

A thresholded image $g(x, y) = 1$ if $f(x, y) > T$ and
 $g(x, y) = 0$ if $f(x, y) \leq T$

Where: $g(x, y)$ is the thresholded image
 $f(x, y)$ is the grey level of point (x, y)

b) Local thresholding:

In many cases, the background grey level is not constant, and the contrast of an object differs between the objects in the scene. In such cases, the thresholding value which can be applied in one area of the image might not work well in other areas. In these situations, it is better to use a threshold grey level that is slowly varying throughout the image, according to some local image property (Gonzalez, *et al.*, 2002 and Castleman, 1996).

When the thresholding value T depends only on $f(x, y)$ (the grey level values) the thresholding is called global. If T depends on both $f(x, y)$ and some local property of the point (x, y) for instance the average grey level of a neighbourhood centred on (x, y) , this type of threshold is called local. If T depends on the spatial coordinates x and y , the threshold is called dynamic or adaptive.

2.6 Conclusions

The human finger consists of three joints DIP, PIP, and MCP. These joints have two types of motion; the primary motion is in the flexion-extension plane. Abduction and adduction are limited and occur only at the MCP joints.

There are several existing tools for measuring the angles of the finger joints, they are considered as contact measuring devices, which sometimes can be difficult to use especially if the hand is suffering from serious injuries or damage.

Some of these existing tools are quite cheap, however the time demand for the assessor and the patient, and the lack of accuracy are the major drawback of these devices. Also most of the existing devices assess one joint at a time, so it takes longer to measure all of the fingers joints.

A computer vision system can be used to assess the finger joints instead of the current tools. Such a system is non-contact and harmless to the patient. As the

proposed system is based on computer technology, the measurement process is expected to be achieved in a few seconds. If the accuracy of the system achieves ± 1 degree, compared with the measurements obtained from the CMM, then it will compete with the existing tools, such as Universal Goniometry and visual estimation. Moreover this system is a non-contact measurement tool which has no effect on human hand in case if it is injured or swollen, unlike wire tracing, Inclinator, and hand gloves which are difficult to use if the hand is seriously injured.

However, to develop a reliable 3D computer vision system takes a lot of effort and cost as there are many factors that influence the overall efficiency and accuracy of the system.

3 Camera calibration and reconstruction of 3D coordinates

3.1 Summary

This chapter explains how the computer vision system is used to extract 3D information from 2D images, and the two most popular techniques which have been used to calibrate the cameras in order to deduce the camera parameters.

The first part illustrates the four steps method, which was introduced by Tsai (Tsai, 1987), and how this method can be implemented to calculate the camera parameters. The mathematical procedures required to reconstruct the 3D world coordinate from the image system are then shown.

The second part shows an alternative calibration technique which was introduced by Abdel-Aziz and Karara (Abdel-Aziz, 1971) and is known as the Direct Linear Transformation or DLT. This method is easier than the four steps method because it is based only on solving a linear equation. Also, this part demonstrates how 3D information can be obtained.

3.2 Introduction

Camera calibration is a crucial problem and an important step towards computational computer vision, i.e. for extracting 3D information from 2D images (Salvi, *et al.*, 2002 and Robert, 1996). Accurate camera calibration and orientation procedures are a necessary prerequisite for the extraction of precise and reliable 3D metric information from images.

In principle, to use cameras to extract 3D information from a scene, there are two techniques: stereovision or by camera calibration. Stereovision is a technique for recovering the 3D structure of a scene from two different viewpoints. From a pair

of images and by using this technique the 3D coordinates of a physical 3D point can be computed by triangulation (Orteu, 2009), for more detail see (Orteu, 2009, Efford, 2000, and Gonzalez, *et al.*, 2002). As soon as the accuracy requirement increases, then this method is no longer sufficient (Würz-Wessel, 2003).

The camera calibration approach is a mathematical model which can be used to deduce 3D world coordinates of an object from two or more 2D images. Depending on the required accuracy, the calibration algorithm starts from solving a linear equation and goes on to use a more sophisticated and complex model.

There are different requirements for camera calibration. In some robotics applications for instance, the calibration procedure should be fast and automatic, however in metrology applications the accuracy is a more important factor (Heikkila, 2000).

3.3 What is camera calibration?

Camera calibration is a necessary step in 3D computer vision in order to extract metric information from 2D images (Zhengyou, 2000). The purpose of this calibration is to create the relationship between 3D world coordinates and their corresponding 2D image coordinates. Once this relationship is established, 3D information can be inferred from 2D information, and vice versa (Luong, *et al.*, 1992).

One common method for extracting 3D information from intensity images is to acquire a pair of images using two cameras placed apart from each other. Two corresponding points in these two images result in 3D information. As an alternative, two or more images, taken from a moving camera, can also be used to compute 3D information (Zehang, *et al.*, 2001).

Camera calibration in the context of three-dimensional computer vision is the process of determining the internal camera geometric and optical characteristics (intrinsic parameters) and the 3D position and orientation of the camera frame relative to a certain world coordinate system (extrinsic parameters), which represent the first step in 3D computer vision measurements. Generally, the overall performance of the computer vision system strongly relies on the accuracy of the camera calibration (Heikkila, 1997).

3.4 Why We Need Calibration

One of the important aims of 3D computer vision is to find the position of objects in real space. We measure everything in real space by establishing a reference frame, which is called the world reference frame. An object in an image is measured in terms of pixel coordinates, which are in the image reference frame.

Because we know the distance between pixels in an image, and do not know the distance between these pixels in the real world, we must create some equations to link the world reference frame and image reference frame in order to find the relation between the coordinates of points in 3D space and the coordinates of the points in the image.

The difficulty is we cannot find the relation between those two reference frames directly. To link these frames together, we need another reference frame called the camera reference frame. The basic idea is to find an equation linking the camera reference frame with the image reference frame, and another equation linking the world reference frame with the camera reference frame. Solving this system results in the relation we are interested in (Zhengyou, 2000).

The problem of camera calibration is to compute the camera extrinsic and intrinsic parameters. Extrinsic parameters are the parameters that define the location and orientation of the camera reference frame with respect to a known

world reference frame. Intrinsic parameters are the parameters necessary to link the pixel coordinates of image points with the corresponding points in the camera reference frame (Zehang, *et al.*, 2001).

A direct measurement of the extrinsic and the intrinsic parameters of a camera positioned in 3D space is technically impossible or not feasible. Because of this, camera parameters are calculated indirectly using suitable calibration techniques (Klette, *et al.*, 1998). Calibration of a camera is the procedure of indirectly measuring these parameters in practice.

3.5 Calibration Techniques

Several calibration techniques for the determination of extrinsic and intrinsic parameters have been published such as Abduel-Azez and Karar (1971), Tsai (1987), Heikkila& Silven (1997), and Zhang (2000). To review them is outside the scope of this work (for different camera calibration approaches see Tsai, R.1986, Salvi, *et al.*, 2002 and Würz-Wessel, 2003). These are all based on the pinhole camera model. They generally include measurements which enable almost direct calculation of camera position and orientation, as well as internal camera parameters (Dapena, *et al.*, 1982).

The calibration techniques can be roughly classify into two categories: standard calibration and self-calibration (Maybank, *et al.*, 1992 and Zhang, 2000).

3.5.1 Standard and classical calibration

The standard calibration method, see Figure 3-1, uses a calibration object, whose geometry in 3D world coordinates is known with very good accuracy. These 3D world coordinates can be obtained by using a modern machine, such as a Coordinate Measurement Machine (CMM), which has an accuracy of 0.7 μm . Then the calibration object is put in the field of view of the cameras in order

to find the relation between world coordinates and image coordinates from which the parameters of the cameras can be extracted (Ji, *et al.*, 2004).

There are many standard calibration techniques available from the literature and the most popular two are the Direct Linear Transformation and Tsai camera calibration. The standard calibration methods are unsuitable in many applications because the camera must be directed toward a calibration pattern. So another calibration technique was developed known as camera self-calibration (Maybank, *et al.*, 1992 and Zhang, 2000).

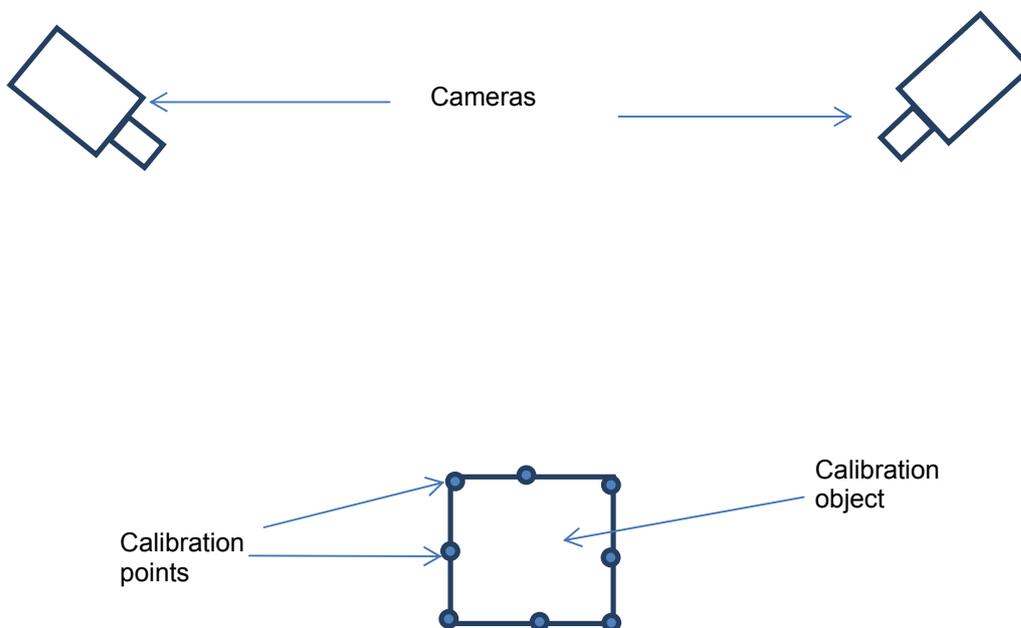


Figure 3-1: Standard calibration method

3.5.2 Camera self-calibration

The self-calibration technique does not use any calibration object, and performs calibration using the matched 2D points in different views (Dai, *et al.*, 2001). Camera self-calibration is especially important in circumstances where reference objects cannot easily be placed in the environment (e.g., for the remote control robot). Therefore, the camera extrinsic parameters, i.e. the position and

orientation of the camera relative to the world coordinate system, are not involved in self-calibration methods. Camera self-calibration is not the interest of this work; more details about this technique can be found in (Maybank, *et al.*, 1992 and Sang, 1996).

3.6 Calibration objects

Choosing the size and the shape of the calibration object depends on the purpose of the application. There are pre-manufactured calibration objects which can be used to calibrate the cameras such as the ones shown in Figure 3-2, 3-3, and 3-4. The 3D world coordinates for these calibration objects are already measured so there is no need to use a CMM or other device to find the 3D world coordinates.

However, the size and distribution of these pre-manufactured calibration objects do not always satisfy all computer vision system applications. In other words, some computer vision systems required a specially designed and manufactured calibration object, as in the case of this work.

The geometry of the objects that are used for calibration needs to be known very accurately (Allard, *et al.*, (1995) stated that Doebelin (1975) recommends use of a calibration standard 10 times as accurate as the accuracy required of the device being calibrated). The 3D coordinates of the calibration object can be obtained with some specific devices such as theodolites (Robert, 1996) or a Coordinate Measurement Machine (CMM).

According to the nature of the measurement process, the calibration points can be light emitting diodes (LEDs) that usually emit infrared light, or light reflecting points that reflect ambient or projected light (Sturman, *et al.*, 1994).

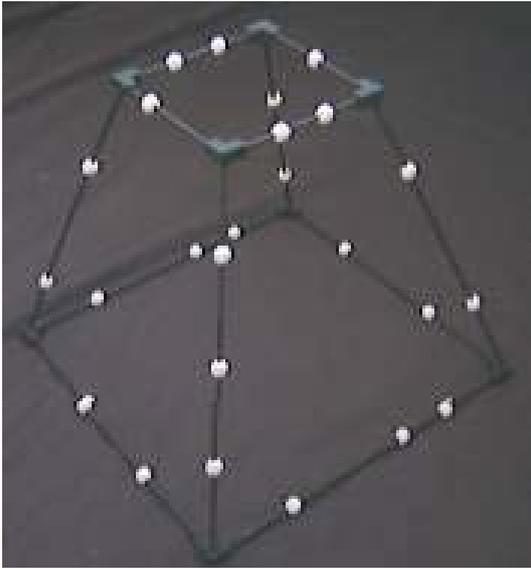


Figure 3-2: framed calibration piece. The calibration points are the white balls (Klette, *et al*,1998)

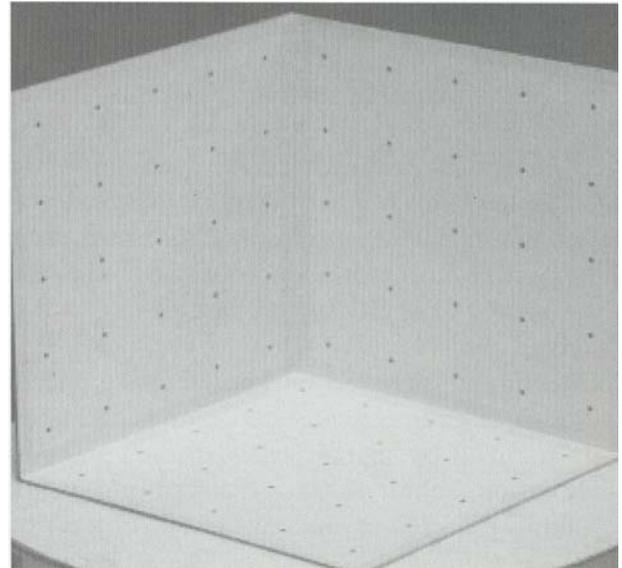


Figure 3-3: open cube calibration piece. The calibration points are the black spots (Klette, *et al.*, 1998)

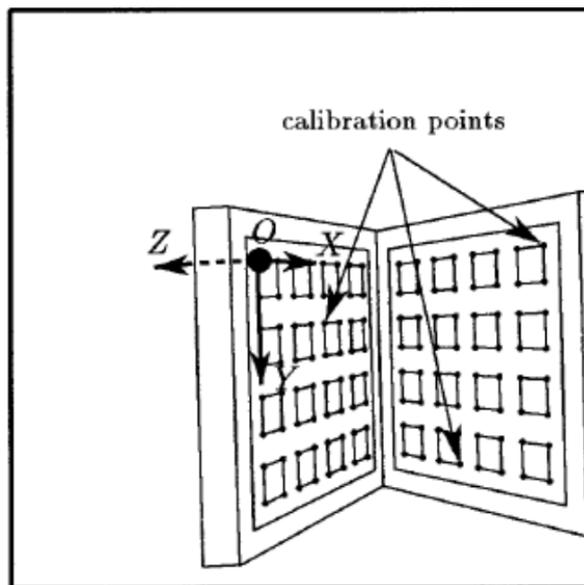


Figure 3-4: Two perpendicular plates, The calibration points are corners of the small squares region(Robert, 1996)

3.7 Camera parameters

Computer vision algorithms for reconstruction of the 3D structure of a scene or for computing the position of objects in space need equations linking the coordinates of points in 3D space with the coordinates of their corresponding image points. There are two types of camera parameters:

3.7.1 External parameters (Extrinsic)

The camera reference frame is often unknown, and a common problem is how to determine the location and orientation of the camera with respect to a known world reference frame, using the image information. The extrinsic parameters are the rotation and translation matrices that define location and orientation of the camera reference frame with respect to the world frame (Shah, *et al.*, 1996).

3.7.2 Internal parameters (Intrinsic)

The intrinsic parameters are the parameters that link the image coordinates (pixels) with the camera coordinates. These parameters include:

- Focal length of lens.
- Lens distortion coefficient, radial and tangential.
- Image centre coordinates.
- Uncertainty scale factor (because of converting from analogue to digital in the frame-grabber (Tsai, 1987).

3.8 Tsai's calibration method

Many techniques for geometric calibration of CCD cameras are available from the computer vision literature. The widely used method proposed by R.Y. Tsai (1986), which is based on a two-step technique (Salvi, *et al.*, 2002) will be described here.

Tsai's technique is complicated and requires an estimate of as many parameters as possible using linear least-squares fitting methods. These estimated parameter values are used only as starting values for the final optimization. In a subsequent step, the rest of the parameters are obtained using a nonlinear optimization method.

The Tsai approach also includes the determination coefficients of radial lens distortion k_1 and k_2 , and of the scaling factor s_x . The method requires at least seven non-coplanar, accurately detected calibration points, which are given in any arbitrary order, but with known geometric configuration.

The Tsai method starts with capturing an image of the calibration piece. Then the image coordinates of the calibration points are found. The correspondences between the world coordinate system of calibration points and their images coordinates are obtained. These data represent the basic data on which the calibration is based.

The Tsai model is based on a pinhole camera; a pinhole camera is a simple camera without a lens and with a single small aperture as shown in

Figure 3-5. Tsai model has eleven parameters in total, five of these are internal (also called intrinsic or interior) parameters:

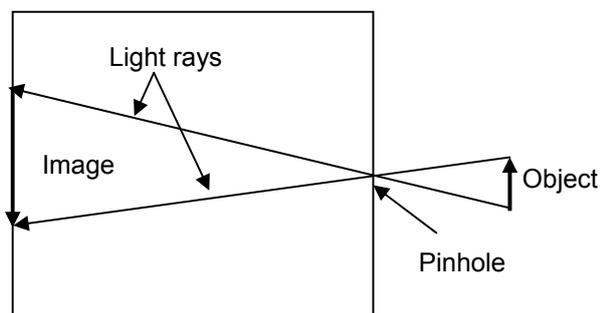


Figure 3-5: Sketch of pinhole camera

- f The distance between the lens centre and image plane (effective focal length of the pin-hole camera),
- k_1 1st order radial lens distortion coefficient,
- C_x, C_y The coordinates of the principle point (i.e. the intersection point of the optical axis with the image plane)
- s_x scale factor to account for any uncertainty in the frame-grabber's re-sampling of the horizontal scanline.

And six are external (also called extrinsic or exterior) parameters:

- R_x, R_y, R_z - rotation angles for the transformation between the world and camera coordinate frames,
- T_x, T_y, T_z - translational components for the transformation between the world and camera coordinate frames.

In addition to the eleven variable camera parameters, Tsai's model has six fixed intrinsic camera constants (which can be read from the manufacturers' data sheets) as below:

- N_{cx} Number of sensor elements in camera's x direction,
- N_{fx} Number of pixels in frame grabber's x direction (in pixels),
- d_x X dimension of camera's sensor element (mm),
- d_y Y dimension of camera's sensor element (mm),
- d'_x Effective X dimension of pixel in frame grabber (mm/pixel), and
- d'_y Effective Y dimension of pixel in frame grabber (mm/pixel).

See Figure 3-6.

3.8.1 Image distortion

Projection in an ideal imaging system is governed by the pin-hole model. Real optical systems suffer from a number of inevitable geometric distortions. In optical systems made of spherical surfaces, with centres along the optical axis, geometric distortion occurs in the radial direction. Figure 3-7 shows the influence of this distortion on the acquired image. When both vertical and horizontal image lines bend in toward the centre of the image, the distortion is called positive distortion (pincushion). Barrel distortion (negative distortion) causes the outlines of an image to curve outward.

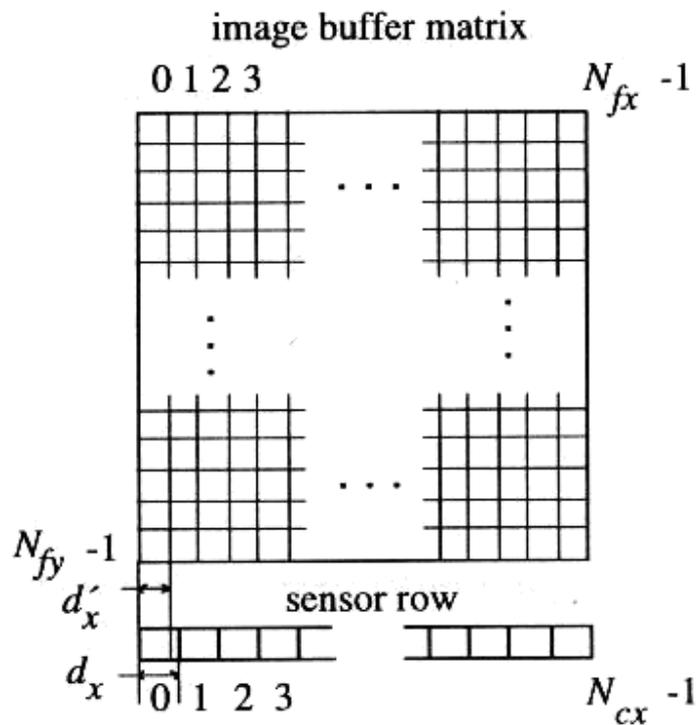


Figure 3-6: Intrinsic camera constants



a) Barrel distortion



b) Pincushion distortion

Figure 3-7: Deformation of the ideal image by radial lens distortion (Montabone, 2010)

The effects of radial lens distortion can be described mathematically. However, an ideal modelling of the lens distortion leads to an infinite number of distortion coefficients. In practice only one or two coefficients are sufficient (Klette, *et al.*, 1998).

3.8.2 The horizontal Uncertainty Factor and image centre

Due to a variety of factors, such as slight hardware timing errors between the image acquisition hardware and camera scanning hardware, or the imprecision of the timing of TV scanning itself, an additional uncertainty parameter has to be introduced. Even a one-percent difference can cause three- to five-pixels error for a full resolution frame. Therefore, an unknown parameter s_x is added to accommodate this uncertainty, and to include it in the list of unknown parameters to be calibrated.

However, since the image is scanned line by line, obviously the distance between adjacent pixels in the y direction is just the same as the centre to centre distance between adjacent CCD sensor elements in the Y direction (Tsai, 1987).

3.8.3 Tsai camera model

Figure 3-8 shows the geometry of the Tsai camera model. The 3D coordinates of the object point P in the 3D world coordinate system are $P(x_w, y_w, z_w)$ and $P(x, y, z)$ are the 3D coordinates of the object point P in the 3D camera coordinate system, which is centred at the optical centre (point O). The z axis is the same as the optical axis. The image coordinate system (X, Y) has the origin at point O_i . f is the distance between the front image plane and the optical centre. (X_u, Y_u) is the image coordinate of $P(x, y, z)$ if a perfect pinhole camera model is used. (X_d, Y_d) is the actual image coordinate which differs from (X_u, Y_u) due to lens distortion. Since the unit of the coordinate system used in the computer (X_f, Y_f) is pixel, additional parameters need to be specified (and calibrated) that relate the image coordinate in the front image plane to the computer image coordinate in the frame memory (X_f, Y_f) .

The overall transformation from world coordinate system (x_w, y_w, z_w) to the computer image coordinate system (X_f, Y_f) is shown in Figure 3-9.

3.9 Reconstruction of 3-D world coordinates using Tsai method

The Tsai equations for camera parameters (see Appendix (A) for more details) are:

$$s_x^{-1} d_x 'X + s_x^{-1} d_x 'X k_1 r^2 = f \frac{r_1 x_w + r_2 y_w + r_3 z_w + T_x}{r_7 x_w + r_8 y_w + r_9 z_w + T_z} \quad (3-1)$$

$$d_y 'Y + d_y 'Y k_1 r^2 = f \frac{r_4 x_w + r_5 y_w + r_6 z_w + T_y}{r_7 x_w + r_8 y_w + r_9 z_w + T_z} \quad (3-2)$$

From equation (3-1) we have:

$$A r_7 x_w + A r_8 y_w + A r_9 z_w + A T_z = f r_1 x_w + f r_2 y_w + f r_3 z_w + f T_x$$

$$(A r_7 - f r_1) x_w + (A r_8 - f r_2) y_w + (A r_9 - f r_3) z_w = f T_x - A T_z \quad (3-3)$$

Where:

$$A = s_x^{-1} d_x ' X + s_x^{-1} d_x ' X k_1 r^2$$

And from equation (3-2) we have

$$B r_7 x_w + B r_8 y_w + B r_9 z_w + B T_z = f r_4 x_w + f r_5 y_w + f r_6 z_w + f T_y$$

$$(B r_7 - f r_4) x_w + (B r_8 - f r_5) y_w + (B r_9 - f r_6) z_w = f T_y - B T_z \quad (3-4)$$

Where:

$$B = d_y Y + d_y Y k_1 r^2$$

By rearranging equations (3-3,3-4) and when the number of cameras is two or more, the following reconstruction matrix is used to find the world coordinates (x_w, y_w, z_w) .

$$\begin{bmatrix} A_1 r_7 - f r_1 & A_1 r_8 - f r_2 & A_1 r_9 - f r_3 \\ B r_7 - f r_4 & B r_8 - f r_5 & B r_9 - f r_6 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ A_m r_7 - f r_1 & A_m r_8 - f r_2 & A_m r_9 - f r_3 \\ B_m r_7 - f r_4 & B_m r_8 - f r_5 & B_m r_9 - f r_6 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \begin{bmatrix} f T_x - A_1 T_z \\ f T_y - B_1 T_z \\ \vdots \\ \vdots \\ f T_x - A_m T_z \\ f T_y - B_m T_z \end{bmatrix} \quad (3-5)$$

Where:

m is the number of camera used in the system.

A detailed explanation of the four steps model and how it can be implemented to obtain the camera parameters and reconstruct the 3-D coordinates can be found in Appendix (A).

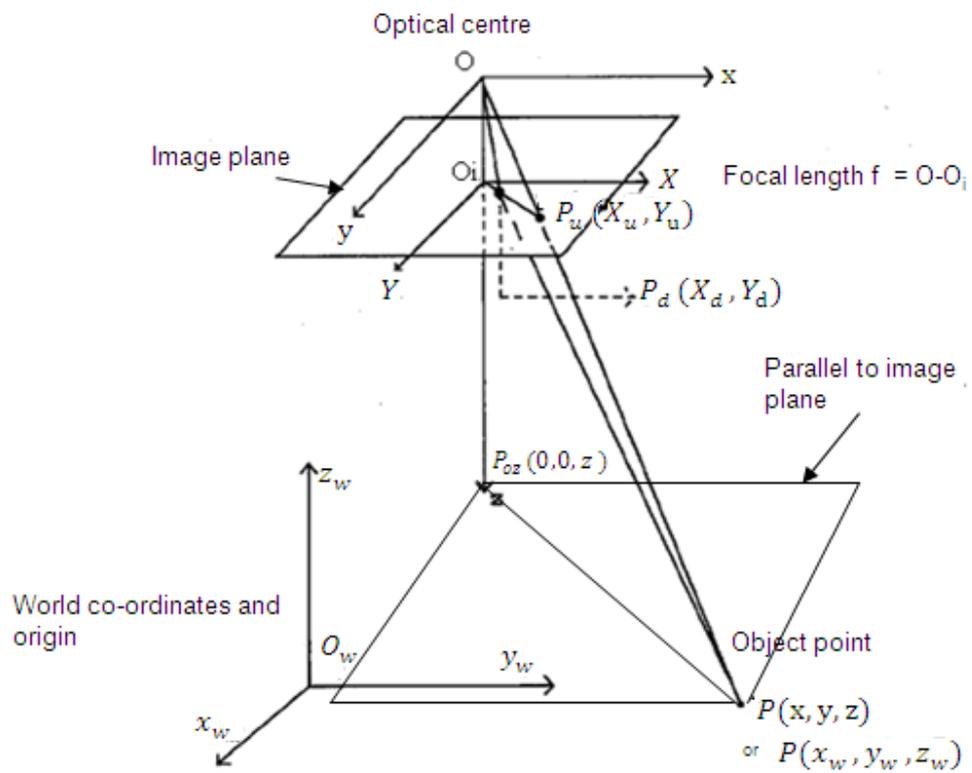


Figure 3-8: Tsai camera model with perspective and radial lens distortion (Tsai, 1987)

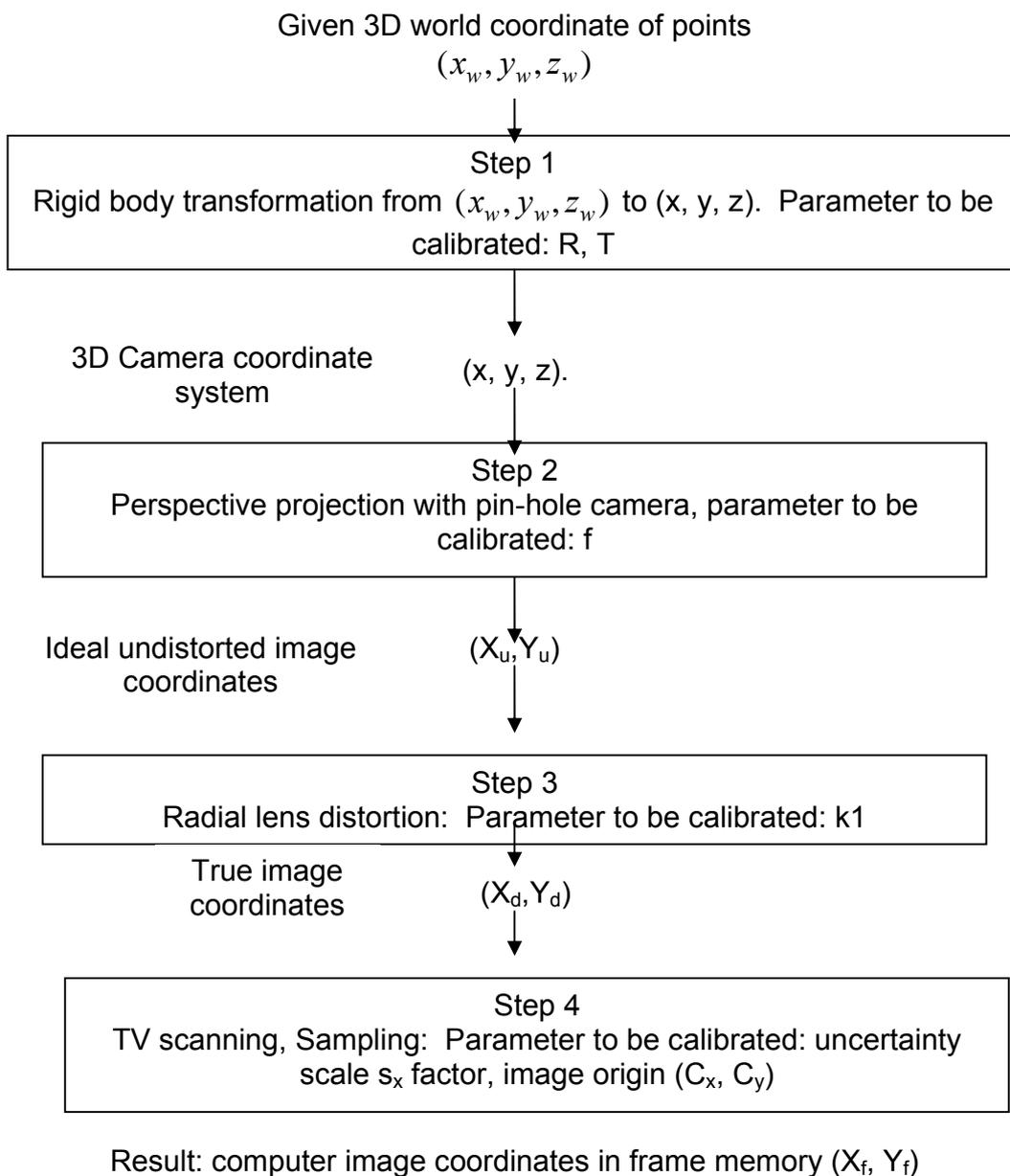


Figure 3-9: Four steps transformation from 3D world coordinate system to computer image coordinate.

3.10 Direct Linear Transformation method

The Direct Linear Transformation (DLT) is one of the most popular techniques used for reconstruction of 3D coordinates from two or more 2D images (Hinrichs, 1995, Challis, *et al.*, 1994). The Direct Linear Transformation method was first proposed by Abdel-Aziz and Karara, 1971; Marzan and Karara, 1975 and it was improved by Hatze, 1988, and Gazzani, 1993. It allows the determination of the 3D coordinates of a point from two or more 2D views of this point. These methods are commonly used in kinematic analysis of human and animal movement. This is because of the accuracy of the results obtained and the great flexibility in camera set-up (Pourcelot, *et al.*, 2000).

Through a calibration procedure, the standard DLT approach determines the DLT parameters. As long as there are at least six control points, the least squares method can be used to determine the DLT parameters. If there are less than six control points, the 11 DLT parameters will be undetermined.

The DLT parameters are solved using a least squares method since the DLT forms an overdetermined system of linear equations. The transformation depends on the position and orientation of the camera, and is characterized by eleven or more parameters. The method is commonly used to obtain 3D coordinates from two or more cameras. The DLT equations are:

$$u = \frac{L_1X + L_2Y + L_3Z + L_4}{L_9X + L_{10}Y + L_{11}Z + 1} \quad (3-6)$$

$$v = \frac{L_5X + L_6Y + L_7Z + L_8}{L_9X + L_{10}Y + L_{11}Z + 1} \quad (3-7)$$

Where:

(X, Y, Z) = the real world coordinate axes for a point in space.

(u, v) = the image-plane coordinate for the same point.

(L_1, L_{11}) = camera parameters.

If we have six or more calibration points with known (X,Y,Z) coordinates, and corresponding image coordinates (u,v) , the camera parameters (L_1, L_{11}) can be obtained as we have twelve or more equations (two for each calibration point) with eleven unknowns. This is done for each camera.

Given the parameters for two or more cameras, which are placed apart from each other, and their image point coordinate data, the unknown world coordinate (X,Y,Z) of any other object located inside the control volume can be calculated because there are four or more equations for the three unknowns (X,Y,Z) .

The major drawbacks of DLT technique are that the control points must be distributed evenly within the measurement space, and location of the control points in space has to be known (Challis, *et al.*, 1992).

The mathematics behind the DLT and how it is used to calculate world coordinates for a given object are explained in full in Appendix (A.2).

3.10.1 Configuration of calibration points

Distribution of the calibration points has an effect on the accuracy obtained from these techniques. Challis, *et al.*, (1992) examined five different configurations within one calibration structure to find out what effect the distribution of the calibration points had on the accuracy. The five configurations are shown in Figure 3-10, Figure 3-11, Figure 3-12, Figure 3-13, and Figure 3-14.

The results show that configurations 1 and 4 gave the least accuracy, with the remaining configurations producing similar results. Frame 1 which has only 8 calibration points gave results comparable with frame 5, even though frame 5 had over four times the number of the calibration points distributed throughout the perimeter of the calibration space. These results illustrate that it is more important to distribute the points around the space in which the measurement is

to take place than to have them inside the space, as this configuration is covering all the measuring volume. Chen, *et al.*, 1994, found that the best results were generally obtained when the calibration points were evenly distributed in the whole calibration region. Leroux, *et al.*, 1991 concluded that increasing the number of calibration points leads to an insignificant increase of the system accuracy. The accuracy of measurement process can be increased by having the control points covering all the measurement volume, in other words all the measurement points, at any hand position and location, have an equivalent calibration points but this is not feasible, so in this work the calibration piece was designed to have control points surround and inside the measuring volume.

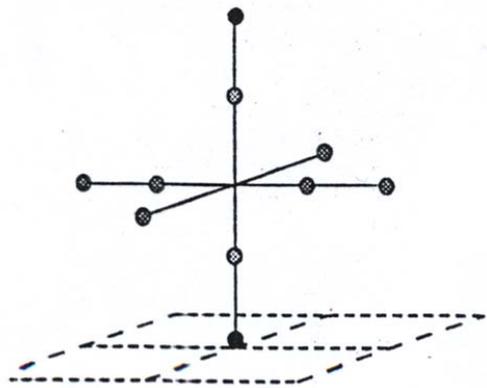


Figure 3-10: Control points configuration 1

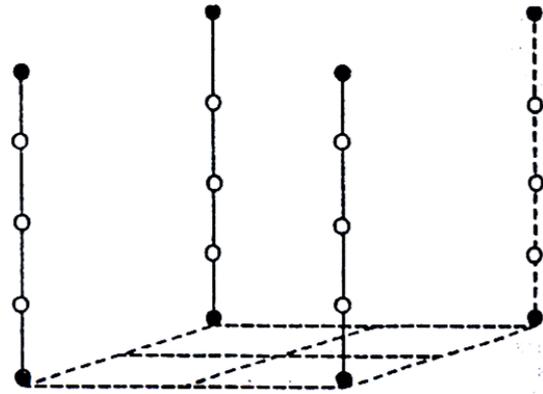


Figure 3-11: Control points configuration 2

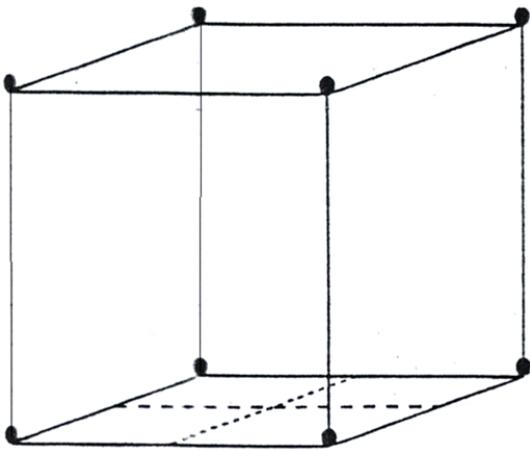


Figure 3-12: Control point configuration 3

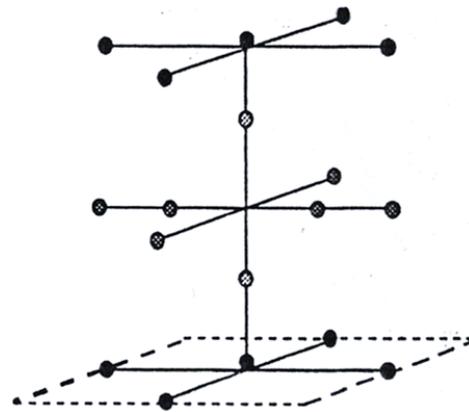


Figure 3-13: Control point configuration 4

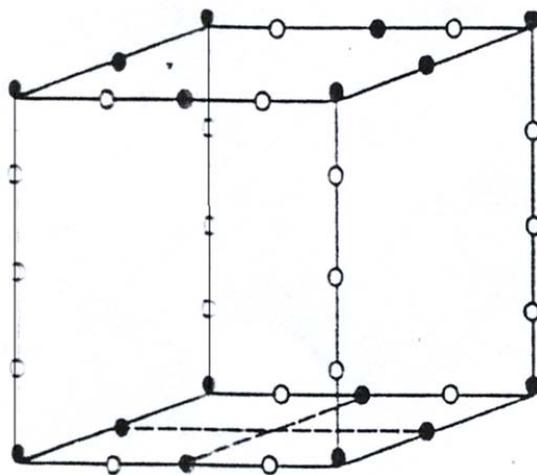


Figure 3-14: Control points configuration 5

3.11 Conclusion

Camera calibration is an important task in a computer vision environment; the cameras need to be calibrated in order to find the parameters that link the 3-D world coordinate system with the image coordinates. There are several methods available to achieve this task.

In this chapter two popular methods have been discussed, one is called the four steps transformation method, also known as the Tsai method, where a non-linear system of equations needs to be solved. The second technique is known as the Direct Linear Transformation or DLT technique. The latter method is easier because it only involves linear equations. In both methods, the location of the cameras need not be known.

In this work DLT and Tsai were used because they are the most widely-used techniques employed for image-based motion analysis. The two techniques are compared back-to-back with regard to performance in chapter 5.

To achieve better accuracy from the measurement system the 3D location of the calibration points must be precisely obtained, and should be evenly distributed across the control region.

4 System design and measurement process

4.1 Introduction

This chapter describes the components of the computer vision system for this study, and also discusses the criteria that are required to design the calibration piece for this system. It also illustrates the human hand model which was used to carry out the measurement of the angles between the finger joints. The procedure for calibrating the cameras and measuring the angles of the human hand model are also explained in this part.

4.2 Components of the computer vision system for this study

4.2.1 Cameras

Pulnix TM-500 analogue monochrome cameras were used to capture images for the calibration and the test piece. These are a low cost ½ inch format CCD camera and give high resolution images. The cameras were connected to the host computer through the frame grabber. More information about the features and the specifications of these cameras are available in Appendix (A.3).

4.2.2 Lens

The working volume in this project should be as small as possible, so as not to occupy a large volume inside the clinic and should be easily moveable if needed. As a working limit, a maximum enclosed volume of not more than 1m³ was decided on. The focal length of the lens is thus restricted by this working volume and the dimensions of the object under investigation, i.e. the calibration piece and a human hand in this application. A lens with 12mm focal length has been used as its field of view can cover the calibration object within the proposed working volume. Reducing the focal length will increase the FOV

and make the job easier, but will reduce the spatial resolution of the image. For a 12 mm focal length lens, the maximum FOV at 1m is 533mm. At a working distance of 750mm, this falls to 400mm, which is more than sufficient to cover the entire hand.

4.2.3 Frame Grabber

The frame grabber, or video capture card, interfaces the camera to the host computer. The frame grabber type DT3155 from Data Translation was used in order to take the image data provided by the Pulnix TM 500 camera in analogue TV form and convert it to digital information so it can be processed by the host PC.

The DT3155 is a programmable, monochrome frame grabber board for the PCI bus. It provides digital video synchronization for reduced pixel jitter which gives high-accuracy data sampling. The DT3155 is suitable for both image analysis and machine vision applications. It also has 4 camera inputs and a MATLAB driver for the image acquisition toolbox.

The DT3155 accepts video signals in many different monochrome formats and digitizes the image. The board either stores the digitized data to the host computer's system memory or transfers the digitized data to the computer's display controller to display images in real time. The board transfers image data to the host computer using PCI burst transfers.

Appendix (A.4) shows the key features of this card and how it is connected to the system.

4.2.4 Illumination

The calibration points, which are distributed throughout the calibration piece in

order to calibrate the cameras and the measurement points, which in turn are used to measure the angles of the fingers joints, consist of fluorescent gel. When these gel points are exposed to UV light, they will fluoresce and give a good contrast against the darker background. For this reason, two ultraviolet 8W lamps were used to illuminate the working area, so that the cameras can see the glowing gel. The dimensions of the lamp are 320 mm × 82 mm × 40 mm (L×D×H as illustrated in Figure 4-1).



Figure 4-1: UV lamp used in this work

4.2.5 Marker and marker-less computer vision system

Some 3D stereo vision systems use markers to create a correspondence between images that captured by cameras placed at different locations, and others systems are marker-less.

There are several markers that have been used in stereo vision systems, such as passive and active markers. Passive markers are not luminescent themselves, but they are covered by reflective materials, which are activated by the arrays of infrared light emitting diodes surrounding the position sensor. While active markers can emit infrared light themselves as they use electronic circuit and batteries (Zhou, *et al.*, 2008).

These type of markers cannot be used in this work because of the size and the way they operate. Figure 4-2 shows examples of passive markers.



Figure 4-2: Reflective markers used in vision systems (Zhou, *et al.*, 2008)

In case of marker-less vision systems, and to reconstruct the 3D information of the object, the images of the object under investigation must have some features which can be used to create matching between the images, such as edges, lines, and so on. The disadvantage of marker-less systems is that some objects do not have clear features that can be used to create correspondence between the images. Marker-less stereo vision system has many applications such as human motion tracking; examples of this application can be found in Caillette, *et al.*, (2004) and Azad, *et al.*, (2008).

One of the main tasks in 3D computer vision system is how to process the 2D digital images, in order to segment the markers or the features which are used to reconstruct the 3D information of the object. From the author's work in digital image processing, enhancement and segmenting digital images to extract features or markers is not always a simple task, particularly if the system has to be done automatically as in case of system proposed in this work. And in many cases, the enhancement techniques do not produce the desired images, so some mathematical estimation, or other techniques, should be used to clarify the objects from the background.

So, to ease the task of enhancing and segmenting the digital image, in order to automatically process the image, then the markers should be easily distinguished from the background. One way to achieve that is using ultraviolet (UV) gel which has been used by Hemmings (2002) and produced a high contrast image. In this work the same gel was used for the reasons mentioned in the following section.

4.2.6 Ultraviolet gel

UV hair gel was used to create the calibration points for the calibration piece and as measurement points for the hand model because of the following advantages:

The images created using the hair gel markers can be easily processed because of the high contrast between the objects and the background. Figure 4-3 shows an image of real hand with UV gel.

The markers can be applied directly to the hand without using any adhesives.

The markers can be removed easily by using soap and water.

The size of the markers can be made very small to increase the accuracy of the measurement process.

The gel has no adverse effect on the skin, i.e. is harmless to patients.

The cost of the gel is negligible, and it is commercially available from high- street stores.



Figure 4-3: The gel points spread on a real human hand

Figure 4-4: shows how the gel points spread on the hand model during the evaluation process.

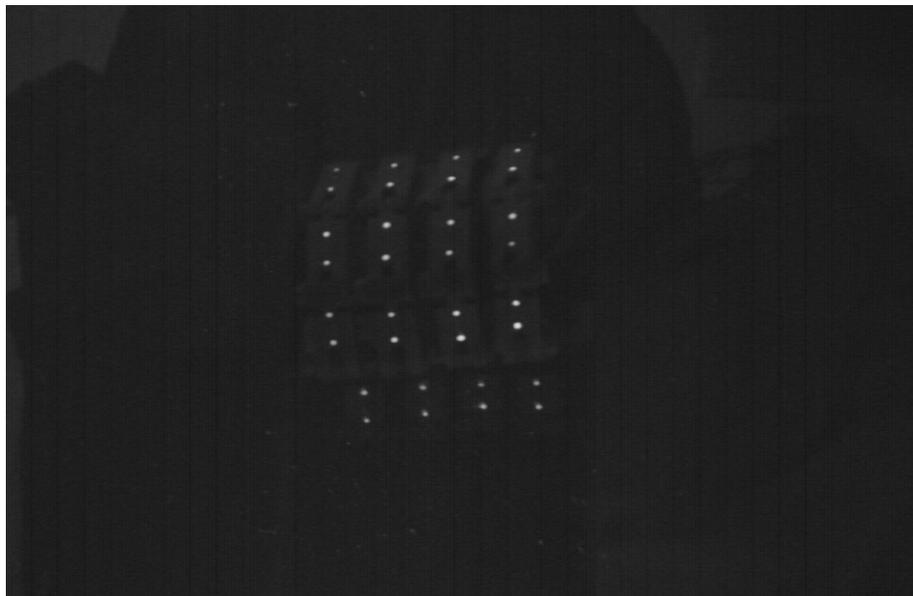


Figure 4-4: UV gel points (white spots) appear on the hand model during the measurement process

4.2.7 Calibration piece

Design of the calibration piece

A proper design of the calibration piece is one of the important issues in 3D computer vision measurements. In order to increase the accuracy of the computer vision system, the calibration piece should have the following characteristics:

The calibration piece should be designed so that it covers all the space occupied by the object to be measured, to avoid extrapolation-based reconstruction of the points located outside the control volume, as Wood, G. *et al*, 1986 and Challis, J. *et al*, 1992 illustrated that the error in measurements increases if extrapolation is used for the DLT technique. In this work, the control volume should enclose the human hand (human fingers model) in both flexion (finger joints are relaxed) and extension (finger joints are bent) positions. Also, all the calibration points should be located within the FOV of the cameras.

1. The control points should be evenly distributed within the measurement volume (Challis, J. *et al*, 1992).
2. To calibrate the cameras by implementing the DLT technique, the number of calibration points should be at least six, and seven for the Tsai approach (so the equations can be solved). Further increases in the number of calibration points should theoretically improve the accuracy of the measurements. However Hatze, 1988 and Challis, *et al*, 1992 found that increasing the number of control points leads only to a minor improvement in the reconstruction accuracy, taking into account the increased cost and effort of the measurement process in terms of image capturing and manipulating.

Calculating the size of the calibration piece

The size of the calibration piece should cover the hand at extension, flexion, and Ulnar and Radial deviation (Figure 2-1 illustrates flexion and extension, and Ulnar and Radial motion of the wrist). At the same time the calibration piece should be covered by the FOV of the cameras within the available working volume, which has been adopted for this work. As the hand occupies a volume in space, then the calibration piece should have a volumetric shape. To determine the suitable size of the calibration piece, the length, width and the movement of a typical human hand should be established first.

Gregory, (2002) stated that, Garrett, (1971) found the typical length of the hand (the distance from the distal wrist crease to the tip of the long finger with the hand extended) for males was 190 mm; and the hand breadth (the distance across the back of the hand) for male was 86mm. The dimension of the male hand was considered in this work because it is larger than the female hand, so the design will also cover the female hand. According to the American Medical Association, maximum wrist motion is 30° and 20° for Ulnar and Radial deviation respectively and 60° for both flexion and extension (Norkin, *et al.*, 2003).

Calculating the length of the calibration piece

Figure 4-5 shows a schematic of the Ulnar and Radial deviation of the hand. In this figure we define the following dimensions:

- W → The distance across the back of the hand
- L → The distance from wrist to the tip of the long finger (middle finger)
- Ø → The Ulnar deviation angle

The calculation of the length of the calibration piece is based on the Ulnar deviation angle (Ø) and not on the Radial deviation angle, because Ø is greater

than the Radial deviation angle, and leads to a design of the calibration piece that covers the hand at all positions, see Figure 4-5.

$$\text{Since } D = L \times \cos \emptyset \quad (4-1)$$

$$H = W \times \sin \emptyset \quad (4-2)$$

And $L = 190$ mm, $W = 86$ mm, and $\emptyset = 30^\circ$ (see section 4.2.7)

Then, the total length of the calibration piece, L_c should be

$$L_c = D + H = 190 \times \cos (30) + 86 \times \sin (30) = 198.5 \text{ mm}$$

Calculating the width of the calibration piece

From Figure 4-5 we have:

$$d = L \times \sin \emptyset \quad (4-3)$$

$$h = L \times \sin \theta \quad (4-4)$$

Where θ is Radial deviation angle, at maximum Radial deviation $\theta = 20^\circ$

Then, the total width of the calibration piece should be

$$W_c = h + W + d = 190 \times \sin (30) + 86 + 190 \times \sin (20) \quad (\text{See Figure 4-5})$$

$$W_c = 222 \text{ mm}$$

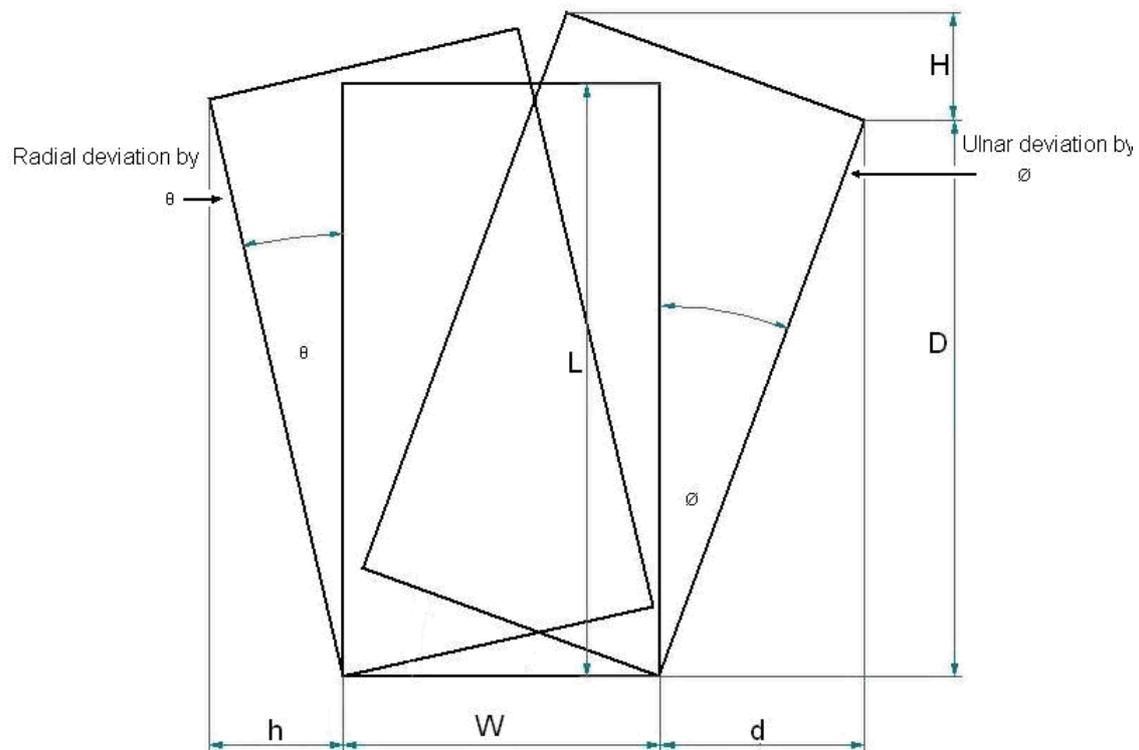


Figure 4-5: The length and width of the calibration piece

Calculating the height of the calibration piece

The height of the calibration piece should correspond to the hand at full extension position. From Figure 4-6 the height of the calibration piece is:

$$H_c = L \times \sin \Phi \quad (4-5)$$

Where: Φ is the extension angle of the wrist.

Given that the maximum extension angle is 60° as reported in Norkin, *et al.*,(2003), then the total height of the calibration piece, when the hand is at maximum extension angle, should be

$$H_c = 190 \times \sin (60) = 164.5 \text{ mm} \quad (4-6)$$

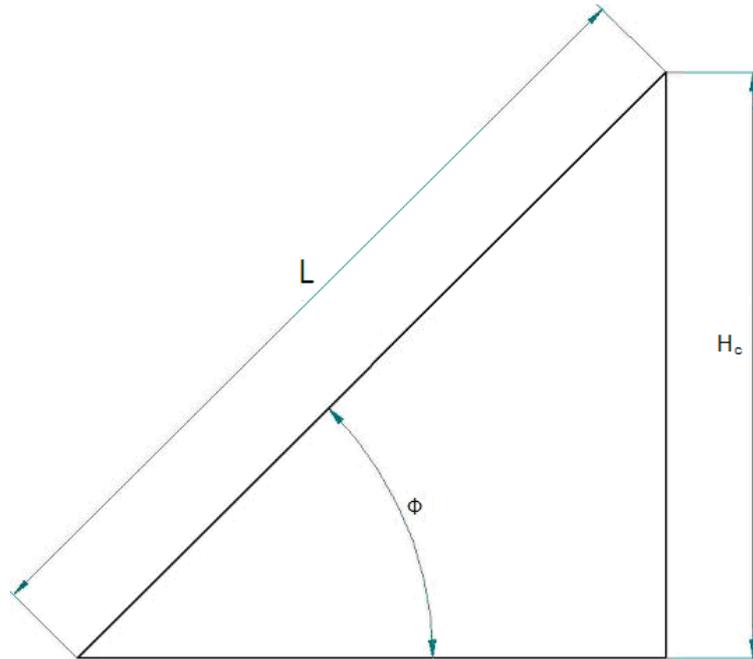


Figure 4-6 : The height of the calibration piece

If the calibration piece is designed so that it includes a volume encompassing the hand at maximum range of motion of the Ulnar and Radial deviation, and the distance from the distal wrist crease to the tip of the long finger with the hand extended, that means the dimensions of the calibration piece as calculated above are $198.5 \times 222 \times 164.5$ ($L_c \times W_c \times H_c$) mm. These dimensions need to be kept as small as possible for the reasons stated previously, and so can be reduced as follows:

1. The effective length of the finger (the actual distance between the first and last measurement points on the finger which is less than the total length of the finger can be used in the calculation instead of the total length, so the length can be reduced by around 30 mm, because the measurement points will not be placed either at the distal wrist crease or at the tip of the middle finger, for more details see Figure 4-7.

The patient's hand can be placed on a support so that the Ulnar and Radial deviation of the hand will be restricted to smaller angles. Reduction from steps 1 and 2 lead to a reduction in the length and the width of the calibration piece.

The height of the calibration piece is a function of the hand length and the extension angle of the wrist (Φ). Reducing this angle by using a hand support, and using the effective length of the hand, leads to a reduction in the height of the calibration piece.

Given the effective hand length of 160 mm, hand breadth of 86 mm, Ulnar deviation angle $\theta = 20^\circ$, Radial deviation angle $\theta = 13^\circ$ and extension angle of wrist $\Phi = 45^\circ$, the dimensions of the calibration piece thus reduce to approximately $180 \times 180 \times 115 \text{ mm}$ ($L_c \times W_c \times H_c$).

Based on the discussion above; the calibration piece was designed and built, with the dimensions stated. It consisted of 6 steel pins, of 6mm diameter, fixed into a wooden base. The base had a square shape of 200 mm in length. Six steel bars of 2 mm diameter were inserted through the pins. Twenty seven points were distributed throughout the calibration piece (9 on the base, 9 in the middle and 9 on the top), making the distance between each pair of points 90 mm. The calibration points were placed with a pen and were approximately circular in shape, with average diameters of around 2 mm. Figure 4-8 (a, b) shows a layout of the calibration piece and the distribution of the calibration points.

The effect of temperature on the steel pins and bars that compose the calibration piece was not considered, as all the experimental work has done under the same indoor temperature, i.e. 20 degrees. In case, and if the temperature changed, and to avoid any error from expansion or contraction of the steel bars and pins, the 3D world coordinates of the measurement points should be re measured by the CMM before calibrating the cameras. Then the new data file that contains the

calibration points replaces the old data file of the same points. And this is very simple task.

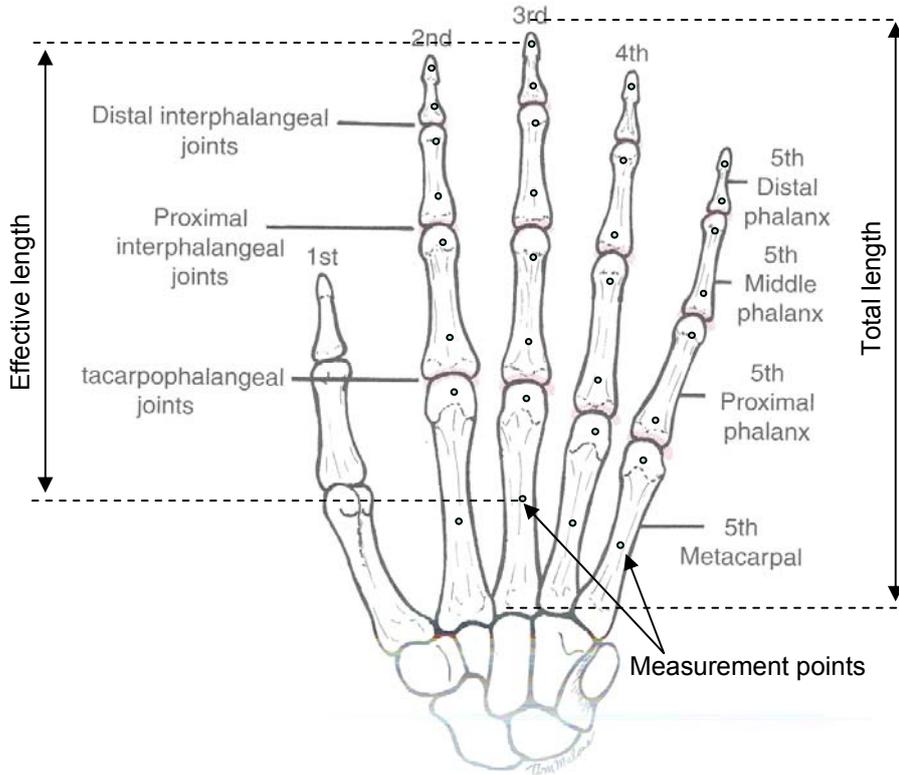


Figure 4-7: Total and effective length of the human hand (Norkin, *et al.*, 2003)

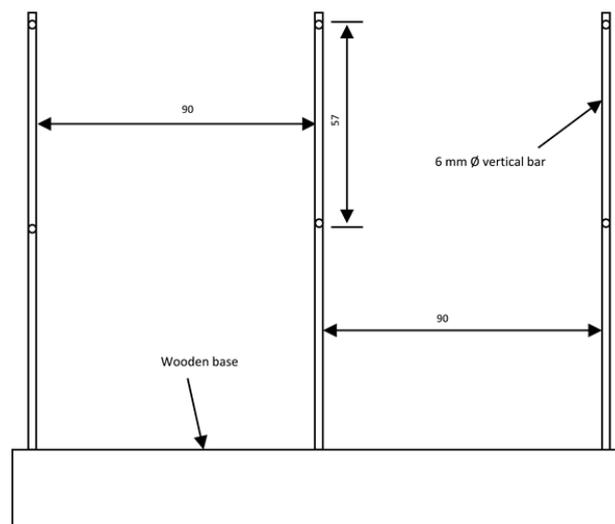


Figure 4-8 (a): Front view of the calibration piece

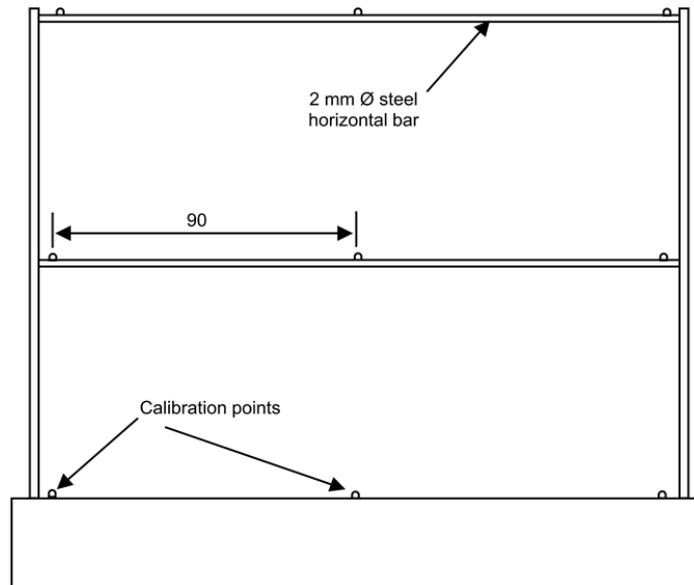


Figure 4-8 (b): Side view of the calibration piece

Because of the absorbing property of the black colour when exposed to white light, the calibration piece was painted with black matt paint, in order to produce a high contrast monochrome image with almost white calibration points on a dark background. Figure 4-9 shows a picture of the calibration piece.



Figure 4-9: Picture of one of the calibration pieces used in this work

4.2.8 The computer and the software

The computer is an important element of a computer vision system. In general, the faster the computer, the less time the vision system will need to process the images. A computer with a Pentium 1.4 MHZ processor and 512 MB memory was used in this design.

Two software programs were utilized to manipulate the data. Image Pro Plus was firstly used then later on MATLAB was used instead. MATLAB is matrix based software and as the digital images are represented by matrices, this makes MATLAB powerful for dealing with these images. It is also well designed for solving equations simultaneously in matrix form; i.e. the DLT equations. It has many built-in functions which makes the code easier to write and faster to execute. Also it is compatible with the DT3155 framegrabber.

4.3 Design of the hand model

A model of the human hand was designed and manufactured from steel. The hand was painted matt black. The same remarks can be made about the benefit of the black paint as were made for the calibration piece above. The model only included four fingers, the thumb is not considered because it lies on a different plane. The bones for each finger can be bent and extended up to an angle similar to the real hand. This model was used to carry out all the angle measurement procedures similar to the measurement of real human finger joints. The lengths of the bones for the hand model are shown in Table 4-1. These dimensions were chosen to be similar to human hand.

Table 4-1: Dimensions of finger bones used in hand model (Parish, 1968, Wagner, 1988, and Gregory, 2002)

Bone's name	Length (mm)
Metacarpus	50
Proximal Phalanx	40
Middle Phalanx	30
Distal Phalanx	20



Figure 4-10: Human hand model

In order to measure the angles between each pair of bones, two points of UV gel were placed on each bone, so a line between each pair of points represents a vector in space. Then the angles of the joint between each adjacent pair of bones are calculated from the intersection of two vectors in space. Figure 4-10 above shows a picture for the human hand model.

4.4 Distribution of the cameras around the measurement volume

The cameras should be distributed so that the maximum accuracy for the calibration parameters can be achieved. However there are some factors which restrict this requirement, for instance:

The size of the working space should be reasonable, so that it does not occupy a large volume. The volume of the measurement area should not exceed 1m^3 (the smaller the size, if possible, the more practical the system as it will occupy a small volume when it is placed in the clinic and can be shifted or replaced easily).

One of the aims of this work is to reconstruct 3D world coordinates from two or more images, that means the calibration points and the measurement points should be captured by at least two cameras. So when the cameras are placed this factor should be considered.

The nature of this work requires that the cameras be placed at positions so that images for the human hand that is under assessment can be obtained with either the hand at flexion (relax) or extension location.

The focal length of the lenses also has to be considered; lenses with a shorter focal length will increase the field of view, but at the same time produce images with low spatial resolution. To cover the control volume in this work with a maximum focal length, 12mm focal length lenses were used.

The cameras were placed initially in positions where they could acquire images of the hand model when all the measurement points could be seen by all cameras directly, i.e. the situation when the hand is at the full extension position was not included (this situation will be considered in chapter six).

The cameras' location and the calibration piece can be represented by a cylinder which has diameter D and height h . The three cameras were placed at 120° apart from each other on the upper circumference of this cylinder and the calibration piece was placed at the base point of the cylinder. Figures 4-11 and 4-12 are sketches following the above geometry, showing the distribution of the cameras around the calibration piece.

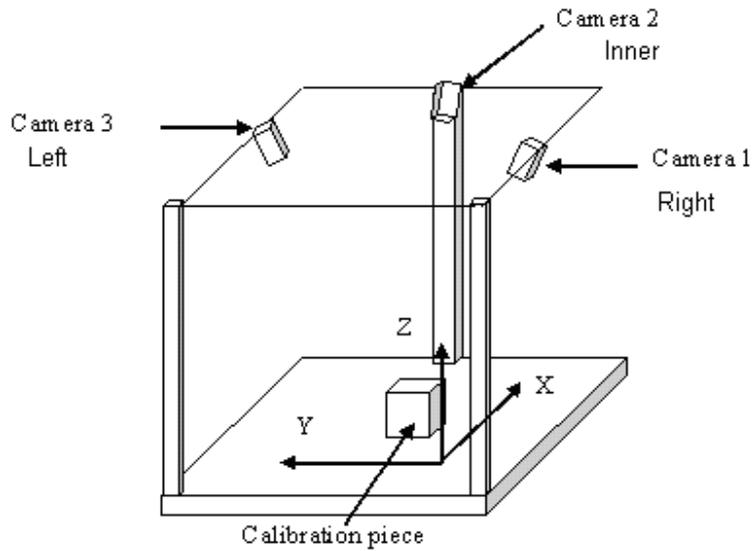


Figure 4-11: Locations of the cameras and the calibration piece (cameras Z coordinates are longer than X and Y coordinates)

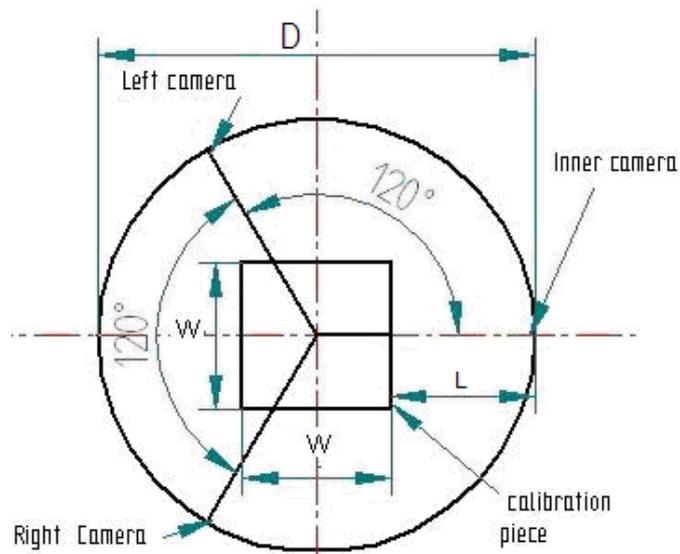


Figure 4-12: Cameras location with respect to the calibration piece, top view

Figure 4-13 shows the front view of one camera and the calibration piece. From this figure we have:

h - The camera height in mm (from the base).

L - The horizontal distance from the calibration piece to the camera in mm.

W - The side length of the calibration piece in mm.

$$\theta = \tan^{-1} (L/h) \quad (4-7)$$

$$\phi \text{ (Angle of FOV)} = 2 \tan^{-1} (\text{width of CCD}/2 \times f) \quad (4-8)$$

Δx is the horizontal distance in which the calibration piece can be shifted.

$$\Delta x = X - (X - \Delta x)H_c = L \times \sin \Phi \quad (4-9)$$

Where:

$$X = \tan (\phi + \theta) \times h - W - L \quad (4-10)$$

$$\text{And } (X - \Delta x) = \tan(\phi + \theta) \times l \quad (4-11)$$

Δz is the vertical distance in which the calibration piece can be shifted.

$$\Delta z = \Delta x / \tan(\phi + \theta) \quad (4-12)$$

The numerical values for all the above parameters (D, W, h, θ , etc) will be illustrated in chapter five.

4.5 The test rig

The rig consists of four parts:

- 1- An aluminium support frame to which the cameras are fixed. The upper part of the frame has a square shape with side length of 650mm. This part is placed on three stands, each 640mm in height. All the cameras are mounted on the upper part, so that they cover the volume under investigation.
- 2- Two lamp holders to hold the UV lamps at a position where a maximum fluorescent of the gel points can be achieved whilst keeping the lamps outside the FOV of the cameras.
- 3- The base where the calibration piece or hand is placed; the calibration piece should be at a fixed position by using two 200mm in length angle steel bars fitted on the base so that they form a right angle between each other, the height of the bars being 20mm (see Figure 4-14). Thus, the

correspondence between the points in 3D world coordinates and the image coordinates can be established automatically.

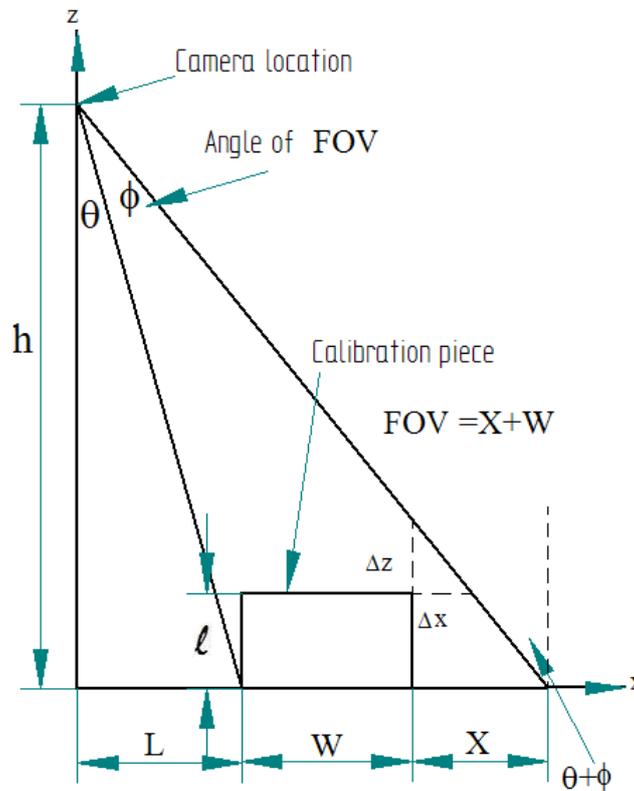


Figure 4-13: cameras location with respect to the calibration piece, front view, left camera

- 4- A cubical frame, with length 75 mm to enclose all the three parts above. The frame was covered with black sheets to exclude day light from all sides excluding the bottom and the front sides. A black curtain was placed over the front side to give access for the measurement process.

Figure 4-14 illustrates all the components of the rig, cameras and the UV lamps for the computer vision system in this project.

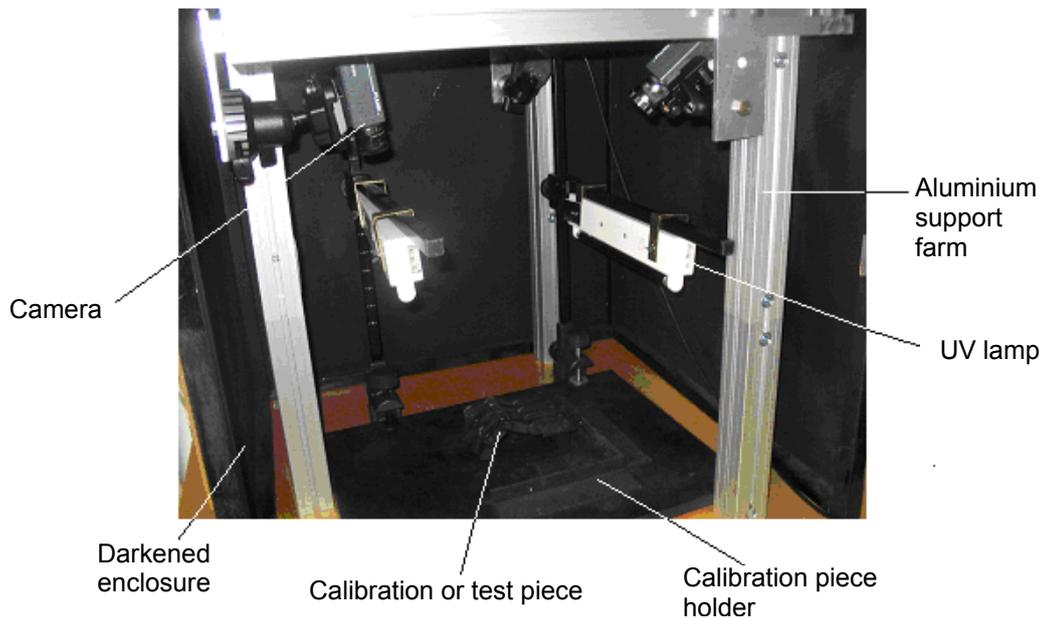


Figure 4-14: The rig used for this study

4.6 The procedure for measuring the angles of finger joints

The measurement of the finger joint angles consists of two steps, the first is to obtain camera parameters from the calibration process and the second is to measure the finger joint angles.

4.6.1 Calibration process

To calibrate the cameras the following procedures were implemented:

1) The calibration piece was placed at a fixed position (in order to automatically create a matching between the calibrations points from different images) inside the calibration box. Then the function `image_capture` (written using MATLAB) is run to order the cameras to capture a sequence of images, two or three

depending on the number of the cameras used. The images are saved at a certain location or processed directly (they have been saved for reference purposes). Notice that, all the software functions mentioned in section 4-6 were written using MATLAB. Some of these functions are shown in full in Appendix (B).

2) A thresholding function `threshold_value` was written to threshold the monochrome images. This function separates the objects (calibration points in this case) from the background. The thresholding process produces binary images, with all the pixels values either 1 for white or 0 for black.

The histogram shown in Figure 4-15-b illustrates that grey levels in image of the calibration piece are normally distributed, so to find the thresholding value, the standard deviation and the mean of the grey level image are first found, and then the following formula used to compute the thresholding value:

$$T = (3 * STD + M) * 1.1 \quad (4-13)$$

Where: STD is the standard deviation of the global image grey level.

M is the global mean grey level of the image.

This empirically-derived formula gives a good result for both the calibration piece and the test piece under the circumstances of the experimental work. Figure 4-15 (a, c) are grey level and binary image of the calibration piece.

After thresholding the images, and getting the binary versions, the `calb_piece_image_processing` function was implemented, the output from which is a three column matrix, where the first column includes the object number, the second and third containing the x_{im} and y_{im} image coordinates for the calibration points. The number of rows of this matrix depends on the number of calibration

points under investigation. The purpose of putting data in a matrix form is to benefit from the power of MATLAB for handling matrices.

The image coordinates of the calibration and measurement points were found by using a built in Matlab function called `regionprops`. This function calculates the centre of mass (centroid) rather than the centre, which calculated by CMM. There was an idea to write a Matlab function that calculate the centre on the calibration and control points, but there is no guarantee the result will have significant improvement, taking into account the shape of calibration and measurement points is approximately a circle with around 2 mm diameter, and the centroid and the centre of a circle are the same.

3) From the step (2) above we have 3 data files produced from processing the 3 digital images of the calibration piece that captured by the three cameras. Each file contains a matrix with 27 rows and 3 columns. Number of rows represents the calibration points and the first column represents the calibration point number and the second and third columns represent the x and y image coordinate of the calibration points (an example of this format can be seen in Table 5-2).

Nevertheless, and because of the cameras located at different location, as a result the data on each file does not have the same order, in other words, calibration point located in row one in the data file produced by camera one will not be at the same row for the data file of camera two, the same for camera three.

To calculate the camera parameters, the three data files of the three images of the calibration piece must be first rearranged in order so that all files have the same order for each calibration point. I.e. calibration point located at row number one for the first camera will have the row location for cameras two and three, and the same for the rest of the calibration points.

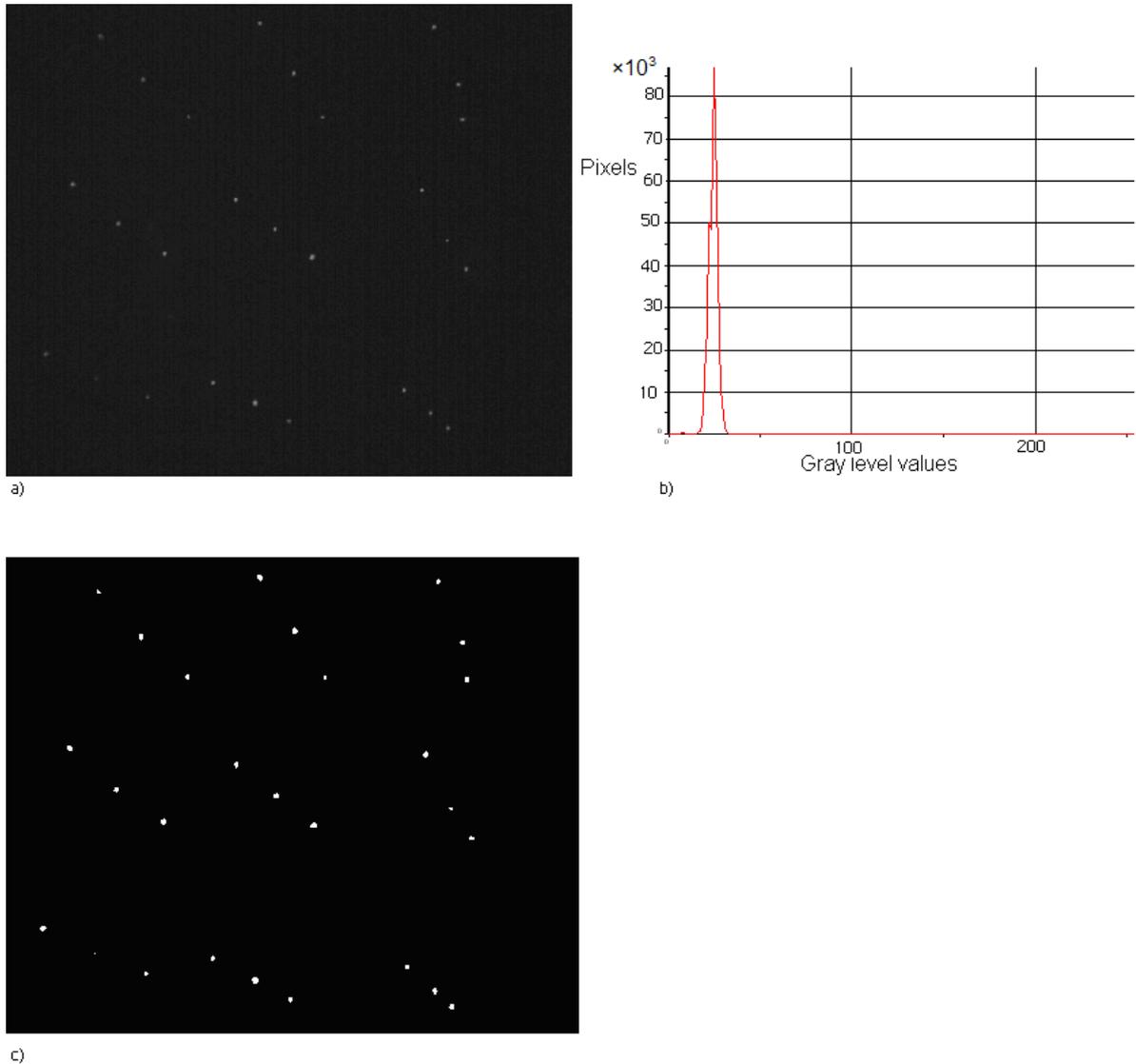


Figure 4-15:

- a) A grey level image for the calibration piece (white spots represent the fluorescent gel calibration points)
- b) An image histogram of the calibration piece.
- c) An image for the calibration piece after thresholding at level T.

4) Having that, the calibration piece located at a fixed place with respect to the cameras, not as in the case of the finger joints, so we know in advance where the location of calibration points captured by each camera. For example the calibration points captured by camera number one, for instance, located at row number one will be located at row number 27, and the same for the rest of the calibration points. See Figure 4-16 for more clarification.

5) The fourth step of calibrating the cameras is to run the function `auto_match` which automatically order the calibration points from the three images. For example if the calibration point number one located at the first row of the data file of image number one, then the same points will locate at first row of the images number two and three, and so on. Also this function creates matching between the 3D world coordinates (x_w, y_w, z_w) and the image coordinates (x_{im}, y_{im}) of the calibration points.

Notice that, the function `auto_match` can be written by any programming language such as C, however the code of this function was written by Matlab, as Matlab contains built in functions which make manipulating matrix much easier, given that the data files of the three images represent matrices.

Once the correspondence is established, then the data are now ready to be processed by the function `camera_parameters`, which computes the parameters for each camera.

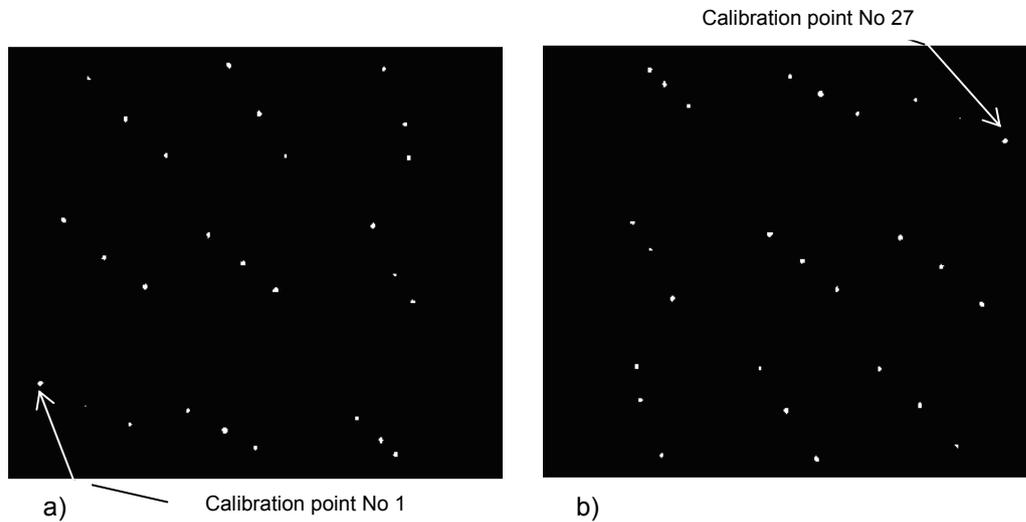


Figure 4-16:a) Image of calibration points captured by camera No (1) shows the location of the calibration point No 1, b) image of calibration points captured by camera No (2) shows the change of location of the calibration point No 1 to point No 27.

5) To find the cameras' parameters, using the DLT (or another technique) a camera_parameters function has been written, the inputs of which are the 3D world coordinates for the calibration points and the x, y image coordinates for the same points. Output from this function is a matrix with two or more columns, depending on the number of cameras, and each column represents the twelve parameters for each camera in the case of using the DLT method.

4.6.2 Measurement of the angles of finger joints

Once the parameters for all the cameras have been found, the next step is to measure the angles between each pair of finger bones. To measure these angles, firstly one model finger, which was designed previously by Hemmings, M.2002 (see Figure 4-17), was utilized to carry out the measurement process. The results from these measurements are shown in chapter five.

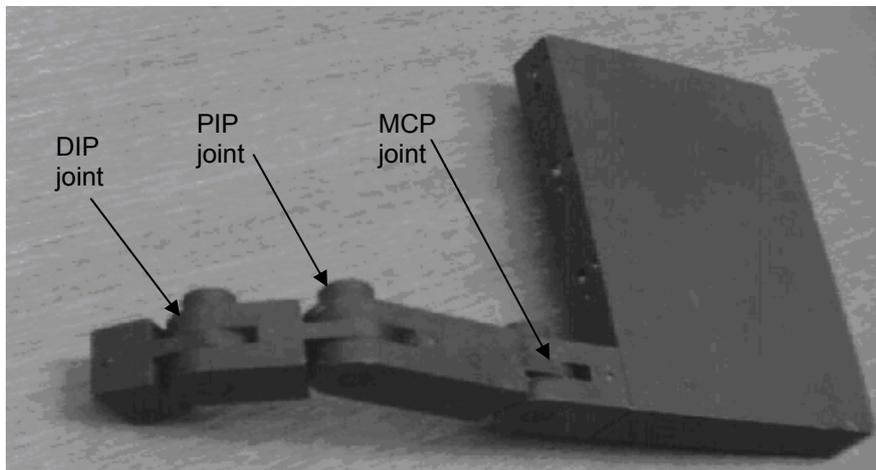


Figure 4-17: One finger test piece

Measuring the angles for one finger is straightforward, because it has only eight measurement points. However measuring all 4 fingers of the whole hand with 32 measurement points is quite complicated, because it requires sorting out and handling the correspondence matter for the 32 measurement points. To measure the whole hand, the following technique was adopted:

Firstly the hand model, with 32 UV gel points distributed on the fingers so each bone has two points, was placed inside the control volume, and procedures 1 and 2 in section (4.6.1) were used to capture and threshold the images of the hand. Then the function `hand_image_processing` was implemented. The results are two or more matrices depending on the number of cameras the system contains. Each matrix has three columns, object number, and x and y image coordinates for the measured points, and 32 rows characterising the number of the measured points. Figure 4-18 shows three thresholded images of the hand taken from three cameras at different positions, (see Figure 4-11 for the relative positions of these cameras).

Reconstructing the 3D coordinates for an object in space, using the DLT or other similar technique such as Tsai camera calibration which will be implemented in this work as well, requires two or more images for the same object captured by cameras located at different positions so that the mathematical equations for the technique can be solved.

Figures (4-19, 4-20 and 4-21) show three images of measurement points of the hand model after applying thresholding and inverting the images. Notice that these images were not taken from one experiment, and are intended to illustrate that the location of the measurement points can be at various positions, i.e. sometimes they are close to each other and sometimes not.

The coordinates of the measurement points are different from one image to another because of the location of the cameras. For instance, the image coordinates (x,y) of the pair of points that are located on the Metacarpal bone, for the image captured by the right camera, are not the same as those captured by the left or inner camera, see Figures 4-19, 4-20 and 4-21.

4.6.3 Correspondence problem

The primary problems to be solved in 3D computer vision are calibration, correspondence, and reconstruction. Calibration and reconstruction have already been described. The correspondence problem involves extracting features such as points, lines and contours in the images and then establishing their correspondences between images. The images can be taken from a different point of view, at different times, and with objects in the scene in general motion relative to the camera(s). Both the process of feature extraction and of feature matching is often computationally expensive and noise sensitive (Lee, *et al.*, 1993).

In general, there are two basic ways to find the correspondences between two or more images which are area-based and feature-based. The area-based stereo (ABS) method creates a description for each image pixel location, usually by producing a measure of the local intensity profile of the area surrounding the pixel and compares this measure to the candidate target pixels in the other image (Goulermas, *et al.*, 2001).

The area-based techniques have a disadvantage in that they use intensity values at each pixel directly, and are thus sensitive to distortions as a result of changes in viewing position as well as changes in absolute intensity, contrast, and illumination (Dhond, *et al.*, 1989).

Feature-based stereo (FBS) techniques use symbolic features derived from intensity images rather than image intensities themselves. Hence, these systems are more stable towards changes in contrast and ambient lighting. The features used most commonly are edge points, line segments and corners.

Also, feature-based methods allow for simple comparisons between attributes of the features being matched, and are hence faster than correlation-based area matching methods (Dhond, *et al.*, 1989).

It is not possible in general to say whether ABS is “better” than FBS or vice versa. It is nevertheless the nature of the images under investigation that push towards choosing the proper matching approach or to develop a new method to achieve the best effective correspondence technique.

In this work, and to create an effective and robust a matching between all images, a new technique was developed based on the following features that can be detected from the captured images (see images Figure 4.17).

The images contain a certain number of measurement points.

All the images contain the same number of measurement points.

Measurement points have an approximately circular shape.

The size of the measurement points is located within a certain interval.

The measurement points are separated from each other.

Each finger has not more than 8 measurement points.

The measurement points have a good contrast with respect to the background.

Four measurement points (MPs) are needed to produce one finger joint angle.

The calibration and measurement points have an average diameter of 6 pixels and 1 pixel equal 0.38 mm at the optical magnification used. This relation can be found by placing a known length inside the calibration volume, and find out how many pixels per one unit measurement.

The image features mentioned above were used to develop a new technique to mach all the MPs that are captured by all the cameras. To make the system robust and to speed up the measurement process, the computer vision system should handle the correspondence matter automatically as explained in the following section.

4.6.4 Automatic matching of the measurement points

A critical problem in the design of this computer vision system is how it can automatically create correspondence between each measurement point in one

image with the rest of the other images, so that the system can automatically reconstruct the 3D world space of the measurement points, allowing the finger joint angles to then be obtained.

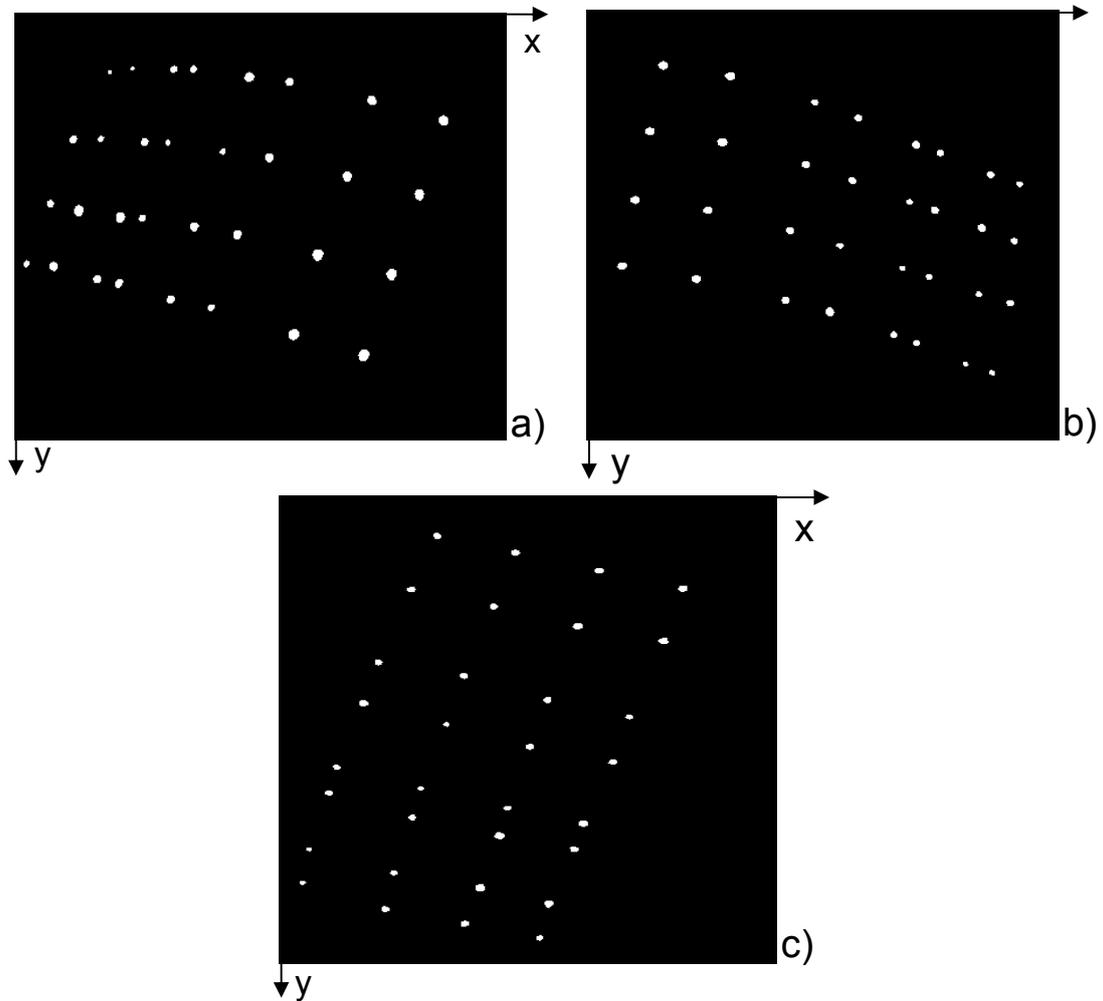


Figure 4-18: Thresholded images of the hand taken from three different cameras.
a) Right camera (see Fig 4-11)
b) Left camera (see Fig 4-11)
c) Inner camera (see Fig 4-11)

The technique, which is explained below, starts with determining the measurement points for each finger and for each bone in that finger for all the images, three in this case. Then the system matches each point in an image to

its equivalent location in the other images (see Figures 4-18, 4-19 and 4-20), The following sections explains how this task was achieved.

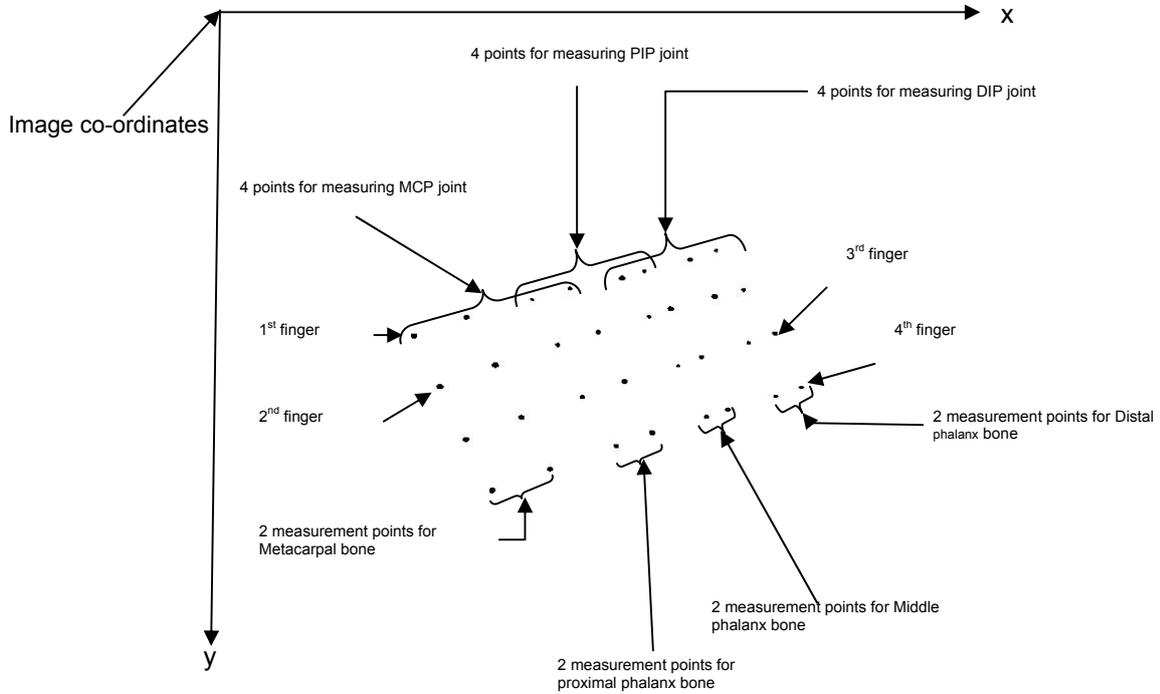


Figure 4-19 : Image of the hand's measurement points captured by the right camera

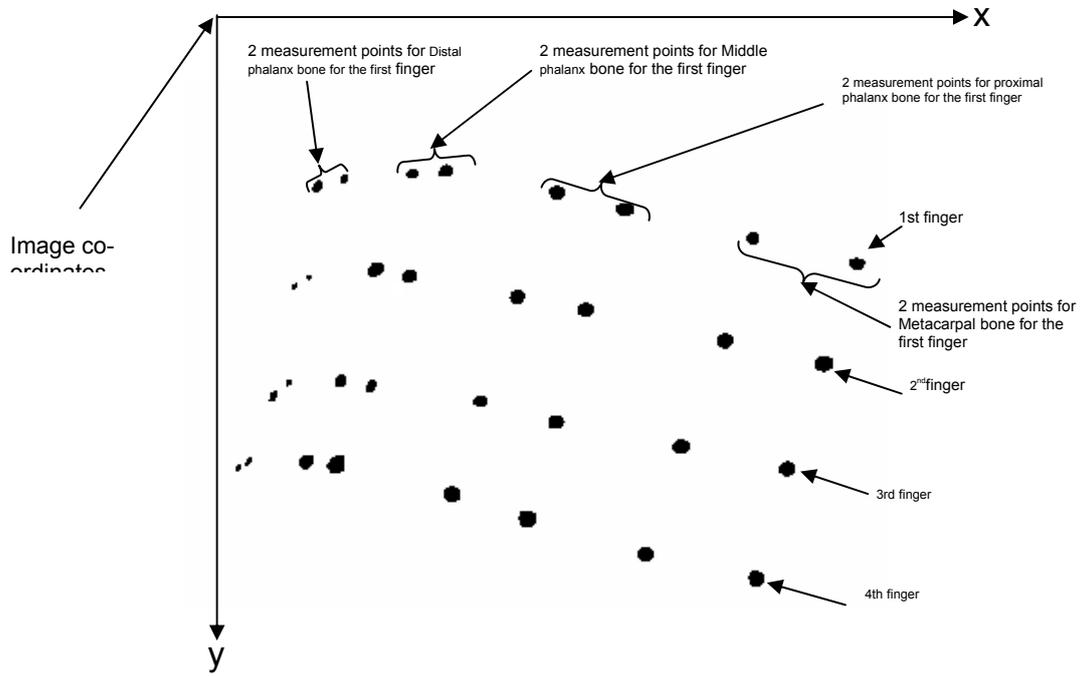


Figure 4-20: Image of the hand's measurement points captured by the left camera

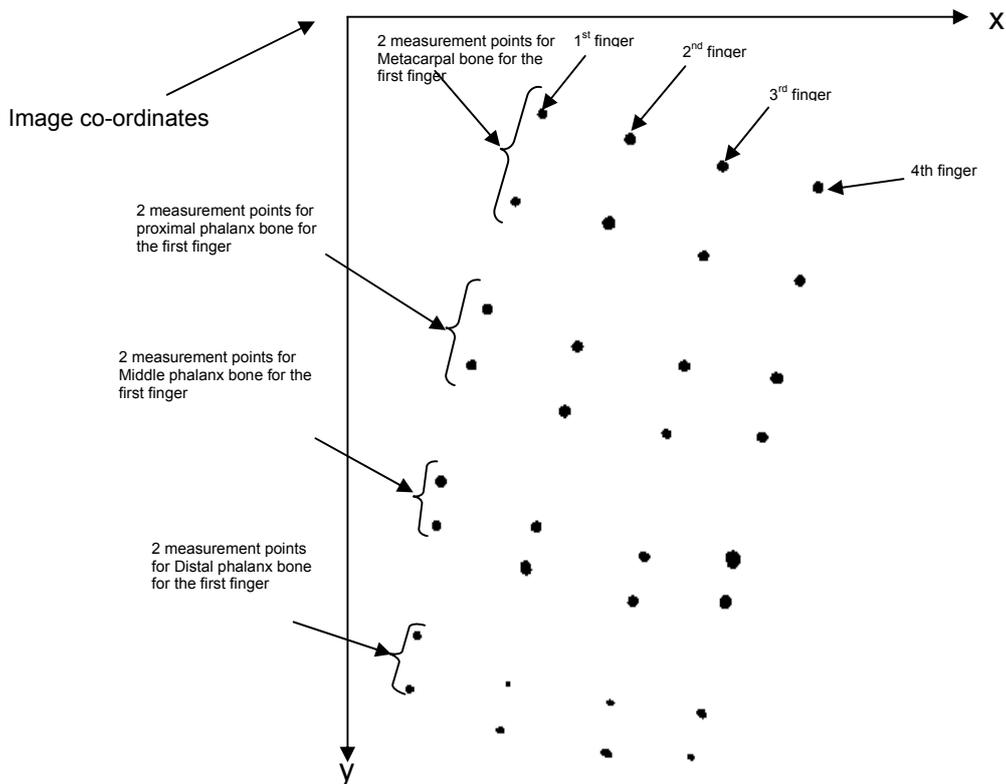


Figure 4-21: Image of the hand's measurement points captured by the Inner camera

A) Arranging the measurement points from the inner camera

After processing the images of the hand model by using the `hand_image_processing` function as explained in section (4.6.2), the outputs from this function are three matrices. Each matrix has 32 rows that characterize the number of measurement points, and three columns that represent object number, and x and y image coordinates for the same points. The six steps below explain how to sort out the measurement points for the inner camera and locate each point with its corresponding finger.

1) Determine the first measurement point in the image. Figure 4-22 shows the image of the measurement points of the hand, captured by the inner camera. The first measurement point, in this case point E in Figure 4-22, was determined

based on the image y-coordinates. In other words, the point with minimum y-coordinates. So, point E is the first point for the index finger.

Mathematically the first point is found from sorting the matrix data for the inner camera obtained from section (4.6.2) above, in ascending order based on the y-coordinates, then selecting the first row from the matrix.

2) Calculate the distance d between the first point obtained from step 1 above and the rest of the measurement points, in terms of pixels, using the formula,

$$d_n = \sqrt{(x_1 - x_n)^2 + (y_1 - y_n)^2} \quad (4-14)$$

For all $n = 2 \dots \dots \dots M$

Where $M =$ total number of measurement points.

3) Find the 3 closest measurement points to the first point (point E in figure 4-21). Figure 4-23 shows points G, F and H which are the three closest points to point E.

4) Calculate the horizontal difference (Δx) between the first point (E) and points G, F and H, see Figure 4-23. Once Δx has been found, choose the two points that have the smallest Δx values, points G and H in Figure 4-23. Two points were chosen because the point which has minimum Δx is not always the next measurement point of the finger under investigation.

5) Compute the absolute difference between Δx_G and Δx_H i.e find $\text{abs}(\Delta x_G - \Delta x_H)$. If the difference ≤ 5 pixels (this value was determined from the experimental work), then the next point will be the point which has minimum y coordinates, otherwise the next point is the point which has minimum Δx .

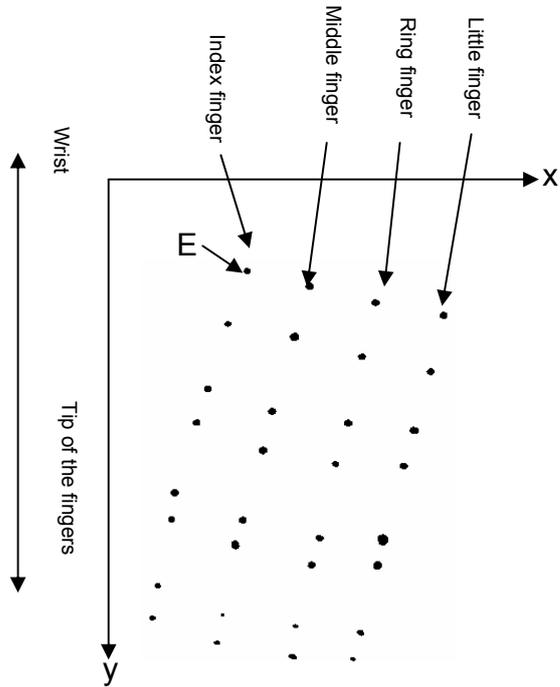


Figure 4-22: point (E) is the first point in the

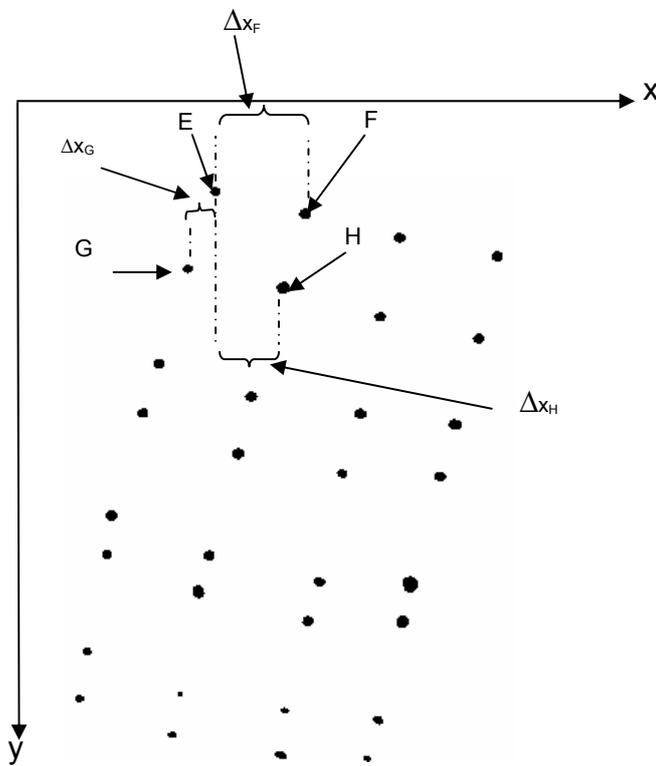


Figure 4-23: Closest 3 points to E and Δx

6) Once the second measurement point for the index finger has been determined, point G in this case (see Figure 4-24), then the steps 1 to 5 above are repeated to determine the third measurement point for the same finger (point K). Notice that in step 1, the points F,I and J in Figure 4-24 have y-coordinates greater than the y-coordinate for the first point (point G in this case), so these points are not included in the code because they are not related to the finger under investigation (index finger in this case, see Figure 4-24).

7) After determining the 8 measurement points for the Index finger, the same procedures 1 to 6 are then applied to establish the measurement points for the Middle and Ring fingers.

8) Once the measurement points for the first three fingers have been determined (Index, Middle and Ring fingers), the remaining 8 points must belong to the last finger (Little finger). Sorting them out requires only ranking them based on their y-coordinates, see Figure 4-25.

9) So far, the eight measurement points for each finger have been established, then the fingers are arranged, based on their first points on the Metacarpal bone, according to x-coordinates that match their location in the hand, see Figure 4-26.

10) The output from step 9 above is a matrix containing 32 rows representing all the measurement points of the hand and 3 columns representing point number and x and y image coordinates for the measurement points. The first eight rows belong to the Index finger starting from the first point on the Metacarpal bone and ending on the last point on the Distal phalanx bone. The same applies for the Middle, Ring, and Little fingers.

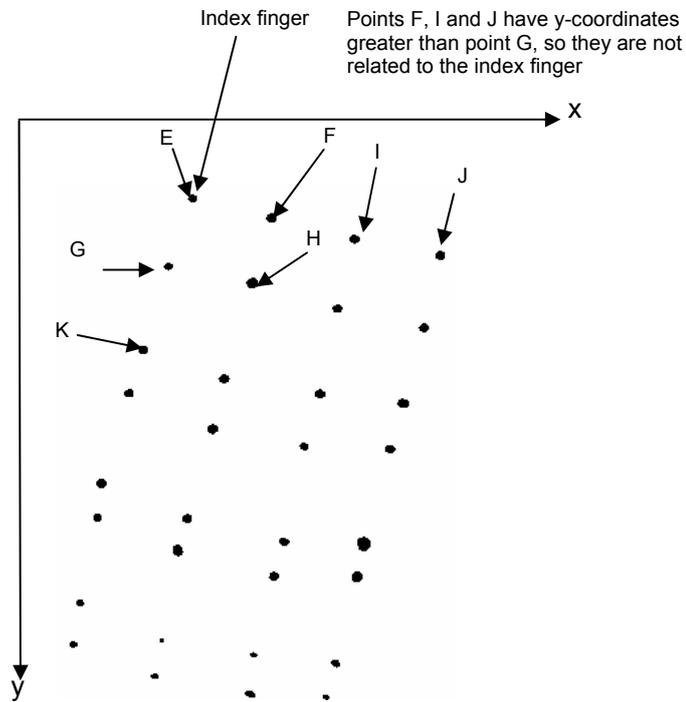


Figure 4-24: Points F,I,J and E are not included when determining the measurement points of the Index finger

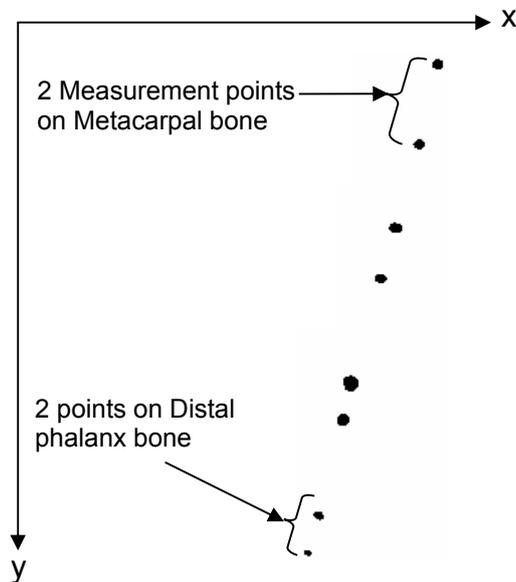


Figure 4-25: The remaining No of points in the image = 8, need to be sorted according to their y locations

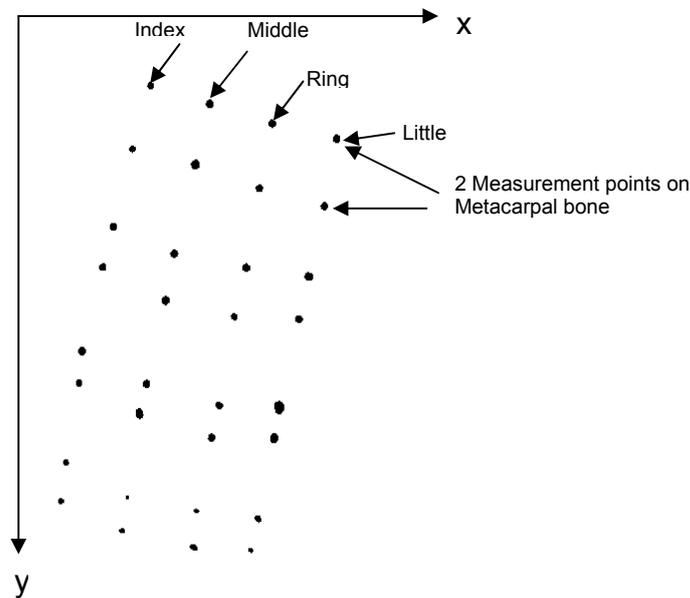


Figure 4-26: Fingers are arranged, based on their first point on Metacarpal bone, according to x-coordinates

B) Arranging the measurement points from the left and right cameras

For arranging the points that are captured by the left and right cameras, a similar technique to that applied for the inner camera, with some changes, has been used. For instance in step (1) above, the matrix data for the inner camera were sorted in ascending order based on y-coordinates to determine the first point. But for the left and right cameras the data are sorted in ascending order based on x-coordinates because the fingers are extended in the x direction, unlike the Inner camera where the fingers are extended in the y direction, see Figure 4-19 and Figure 4-20 in order to see the difference between the direction of the measurement points.

The technique adopted above is easy to implement and can be applied at different camera positions, taking into account minimal changes in the code(see some samples of the codes in Appendix (B)).

Up to this point, the hand was placed at positions in which the measuring points could be seen directly from all the cameras. Later, an explanation is given of a method which can be used in the case where the measurement points cannot be seen directly by all cameras. Also so far, the number of measurement points has been 32. However the points can be reduced to 26 and still allow the measurements to be made as described in chapter (7). This reduction of points will make the measurement process easier and faster.

Once each finger is matched with its own measurement points and each point is linked to its corresponding point in the other images, the next step is to reconstruct the 3D coordinates for each point using the DLT method (Appendix A.2).

4.7 Calculation angle between two vectors in space

Having the 3D coordinates (x_w, y_w, z_w) for the points, the angles of the finger joints can be computed by finding the angles between two adjacent vectors in space (see Figure 4-27) using the following equation:

$$A \cdot B = |A||B|\cos(\theta) \quad (4-15)$$

$$\cos(\theta) = \frac{A \cdot B}{|A||B|} \quad (4-16)$$

Where A and B are two vectors in space and θ is the angle between them.

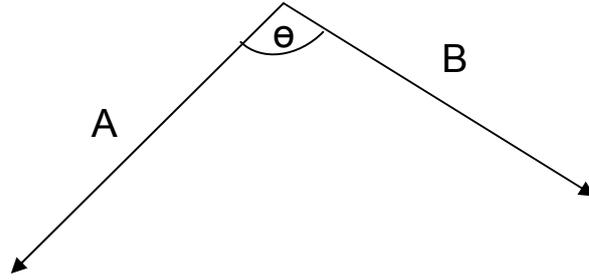


Figure 4-27: Angle between two vectors in space

Figure 4-28 summarizes the process for the camera calibration and measurement of the angles between the fingers bones. Each one of these processes contains several MATLAB functions (see Appendix (A)).

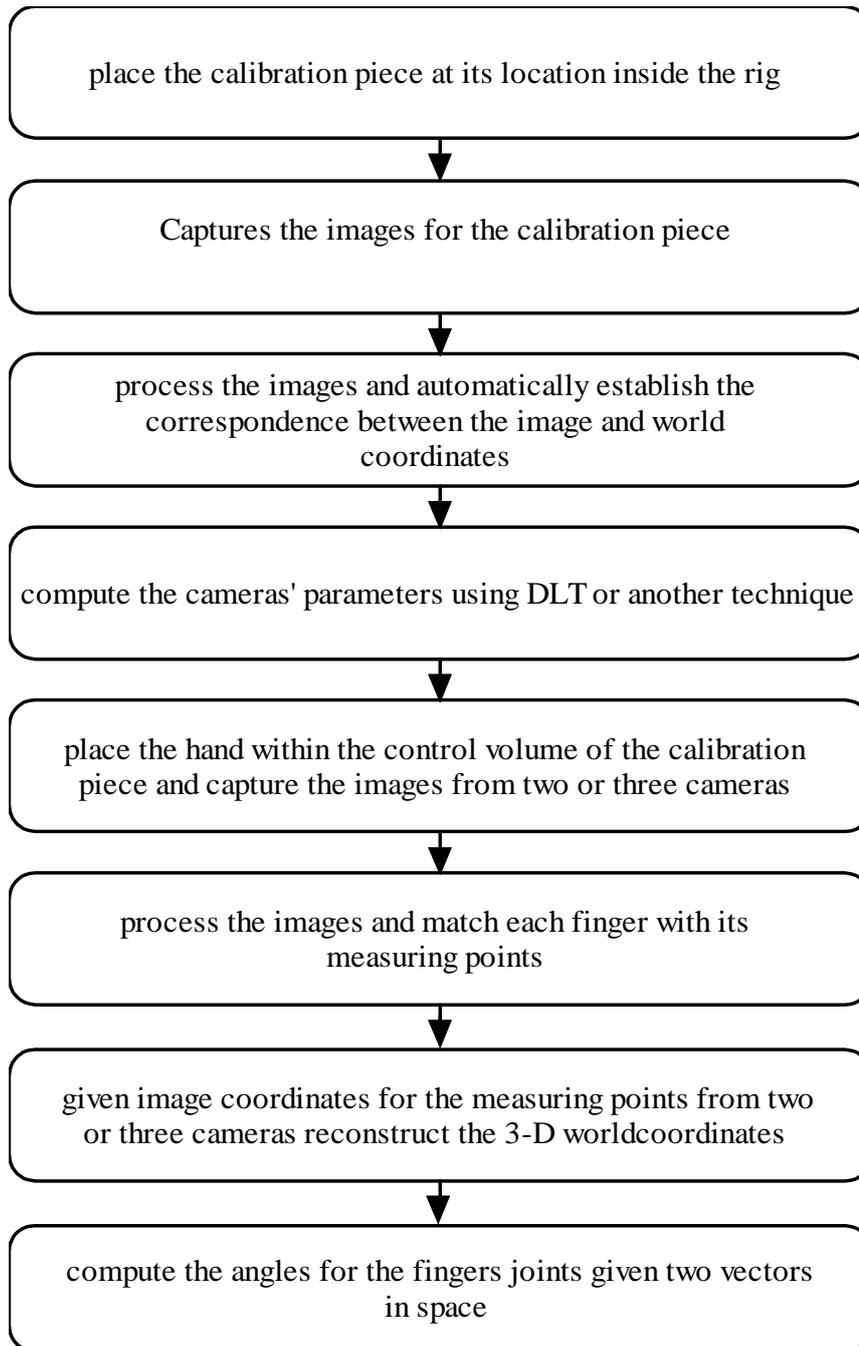


Figure 4-28: the procedures that have been used to measure the angles for the fingers joints

4.8 Summary

Any computer vision system generally consists of camera, lens, frame grabber, illumination, software and the processing unit. Some differences can be found among vision systems, for instance if a digital camera is used, there is no need to use a frame grabber. Also some systems involve only natural illumination and some use special types of light such as UV, the light that has been utilised in this work.

Choosing the proper calibration piece for a certain computer vision task is an important matter, so that a system with good accuracy can be designed. The calibration piece for this work was designed so that it occupied the measurement volume best suited to the average male hand at all positions, and the control points were evenly distributed within the measuring volume as much as possible. The calibration piece was painted with black matt paint to increase the contrast between it and the UV gel control points.

A model of the human hand was designed and manufactured from steel and painted matt black. 32 measurement points produced from UV gel were placed on the fingers, 2 points for each finger bone, to form a vector in space. The calibration and measurement points have an average diameter of 6 pixels (1 pixel = 0.38 mm at the optical magnification used).

The last part of this chapter explained how the cameras were calibrated, and the technique adopted to obtain correspondence between the measurement points so the finger angle can be measured using the hand model.

5 Experimental work and results

5.1 Summary

This chapter covers the first part of the experimental work for this study. Firstly the errors in X,Y and Z coordinates of the measured control points were examined. Then to see the effect of this error on the finger joint angles, a mathematical model was used. Then the angles for one finger were evaluated.

The hand model with 12 joints representing all the fingers was then used to carry out a full measurement process. To check the robustness of the computer vision system, the hand was placed at different locations and the finger joint angles were measured based on the DLT and the Tsai techniques separately. Finally to examine the system for measuring real fingers, the measurement process was implemented on two joints of a real human hand.

5.2 Experimental work

During the first stages of this study, the experimental work was carried out by using a single “finger”, which was designed and built previously by Hemmings (Hemmings, 2002), as shown in Figure 4-17. This simulated finger has three “joints”, i.e., the MCP, PIP, and DIP joints.

To evaluate the finger joint angles, firstly the cameras have to be calibrated. The calibration piece shown in Figure 5-1 was used, its dimensions being shown in Appendix 5. It has a pyramid shape with a square base and made from wood.

There are many shapes of the calibration objects available from the literature such as the ones mentioned in section 3.6. However, the pyramid shape was chosen, because the nature of this shape makes all the calibration points are

seen by the three cameras. As well as this shape is easy to manufacture from wood.



Figure 5-1: Pyramid calibration piece

Notice that, this calibration piece was designed at the beginning of this work and to cover only one finger. However this pyramid shape did not work when we want to measure the four fingers, as increasing the size of the pyramid, in order to locate the fingers inside the calibration volume, will lead to increase the size of the working enclosure which is not comply with this computer vision system. As a result, the new calibration piece which described in section 4.2.6 was introduced.

Three cameras were distributed symmetrically around the calibration piece, the diameter D of the cylinder was 440 mm, and the height h of each camera, from the base, was 520 mm, see Figure 5-2 for more details.

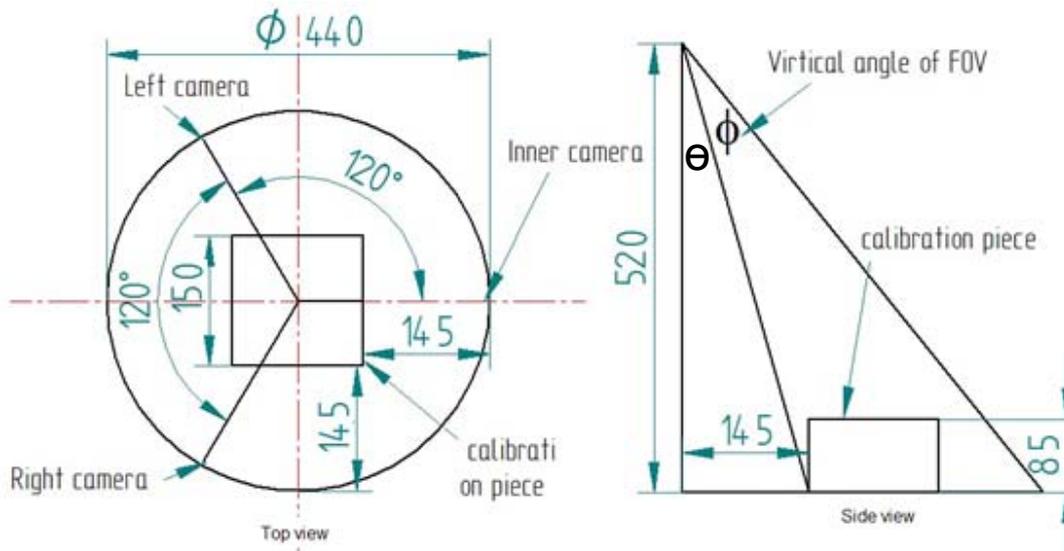


Figure 5-2: Distribution of the cameras within the working volume

The size of the cameras and camera holders restricted and reduced the effective working volume (dimensions of the camera are 45mm x 39mm x 92mm (W X H X L)) and the length of the camera holder is 80mm). If the size of the cameras and holders are reduced, this will give more space and lead to an increase in the FOV of the system.

The FOV of the cameras must cover the calibration piece, under the available working volume and the given focal length of the lens. The following determines the FOV of the camera that is required for this computer vision system.

5.2.1 Calculating the FOV angle

Figure 5-3 shows the general field of view angle of a camera, which can be calculated as below:

$$\tan\left(\frac{\phi}{2}\right) = \frac{\text{sensor width}}{2f} \quad (5-1)$$

$$\therefore \phi = 2 \times \tan^{-1}(\text{sensor width} / 2 \times f) \quad (5-2)$$

Where:

f is the focal length of the lens in mm.

ϕ is the angle of field of view in degrees.

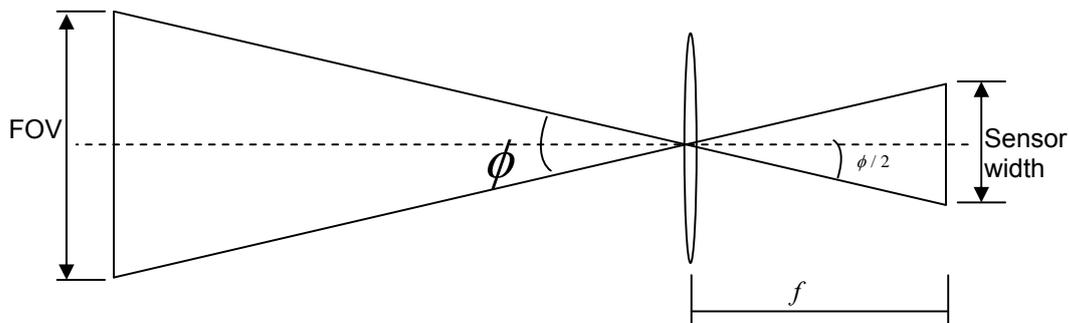


Figure 5-3: Field of view angle (ϕ)

Cameras with a $\frac{1}{2}$ " CCD sensor format have been used, i.e. the height of the sensor is 4.8 mm. The angle of the maximum field of view (FOV) of the camera with the 12 mm lens is:

$$\phi = 2 \times \tan^{-1}(4.8 / 2 \times 12) = 22.6 \text{ degrees} \quad (5-3)$$

5.2.2 Calculating the horizontal and vertical distances in which the calibration piece can be shifted

Placing the calibration piece, shown in Figure 5-1, at the extreme edge of the FOV as shown in Figure 5-4, allows the calibration piece to be moved horizontally by Δx and vertically by Δz . Firstly Δz and Δx will be determined and then the effect of the position of the calibration piece on the calibration accuracy will be tested.

Based on Figure 5-4 we have:

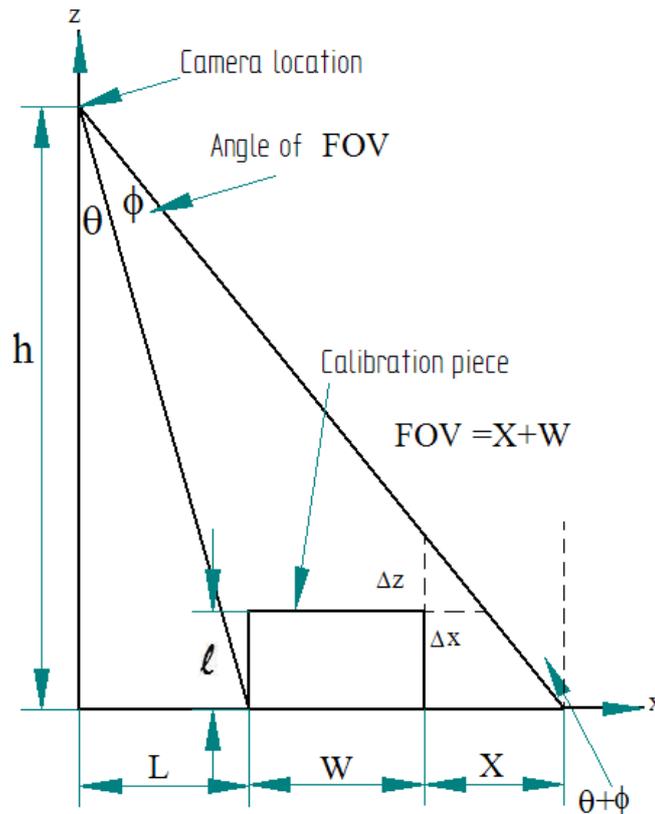
Δx the horizontal distance that the calibration piece can be moved within the FOV.

Δz the vertical distance that the calibration piece can be moved within the FOV.

- h Camera height = 520 mm (from the base).
- L The horizontal distance from the calibration piece to the camera = 145 mm.
- W The side length of the calibration piece = 150 mm.
- l The height of the calibration piece = 85 mm

$$\theta = \tan^{-1}(L/h) \tag{5-4}$$

$$\theta = \tan^{-1}(14.5/52) = 15.6 \text{ degrees}$$



$$h = 520 \text{ mm}, W = 50 \text{ mm}, l = 85 \text{ mm}, L = 145 \text{ mm}$$

$$X = (\tan(\phi + \theta) \times h) - W - L$$

Figure 5-4: Cameras location with respect to the calibration piece, front view, left camera

Given angle of FOV $\phi = 22.6^\circ$ from equation (5-3), the horizontal distance in which the calibration piece can be shifted (Δx), and still located in the FOV of the cameras, can be found as:

$$\Delta x = X - (X - \Delta x) \quad (5-5)$$

Where:

$$X = (\tan(\phi + \theta) \times h) - W - L$$

$$X = (\tan(22.6 + 15.6) \times 520) - 150 - 145$$

$$X = 114.2 \text{ mm}$$

And $(X - \Delta x) = \tan(\phi + \theta) \times l$

$$(X - \Delta x) = \tan(22.6 + 15.6) \times 85 = 66.9 \text{ mm}$$

From equation (5-5)

$$\Delta x = 114.2 - 66.9 = 47.3 \text{ mm}$$

And Δz , the vertical distance in which the calibration piece can be shifted, and still located in the FOV of the cameras, is:

$$\Delta z = \Delta x / \tan(\phi + \theta) \quad (5-6)$$

$$\Delta z = 47.3 / \tan(22.6 + 15.6) = 60.1 \text{ mm}$$

5.3 Measuring the 3D world coordinates for the calibration piece by CMM

The 3D world coordinates for the calibration points of the calibration piece shown in Figure 5-1 were measured using a CMM machine; to do so firstly the origin was chosen to be at the bottom left corner of the calibration piece. To measure the 3D world coordinates for each calibration point, then the probe of the machine was placed on the centre of the calibration point. When the probe touches the calibration point, the machine beeps to indicate the measurement have done and the data saved. All the calibration points were measured by the same method and data file which contains the 3D world coordinates for the calibration points was obtained in order to use it to calibrate the cameras within the vision system. The measurements are shown in Table 5-1.

Table 5-1: Calibration points 3D world coordinates from CM machine

Point No	X (mm)	Y(mm)	Z (mm)
1	2.67	2.14	17.16
2	147.68	2.80	17.21
3	18.36	18.35	34.55
4	133.61	18.07	34.54
5	35.85	33.19	51.94
6	120.12	33.83	52.03
7	50.78	49.95	69.33
8	104.16	48.97	69.31
9	65.19	63.47	86.47
10	89.43	63.99	86.43
11	64.78	87.88	86.51
12	90.03	88.05	86.44
13	51.45	103.66	69.47
14	105.58	102.03	69.44
15	36.15	117.40	52.12
16	120.24	118.17	51.94
17	19.01	132.57	34.53
18	132.94	133.24	34.55
19	3.47	146.79	17.20
20	147.32	147.29	17.11

5.4 Calibrating the cameras using DLT

The calibration piece shown in Figure 5-1 was then placed inside the FOV of the three cameras as shown in Figure 5-2. Three images (one from each camera) were captured and processed to extract the calibration points from the background and to find the x, y image coordinates for the same points. To do this task; the function `Image_capture` (see Appendix B) was written using Matlab. This function instructs the vision system to capture the three images for the calibration piece.

Also another function called `image_processing` (see Appendix B) was written. The task of this function is to automatically process the three images and segment the calibration points from the background. To find the x and y image coordinates for each calibration point, a built-in Matlab function called

regionprops was implemented (see Matlab help for this function). Figure 5-5 shows an image of the calibration piece before and after thresholding.

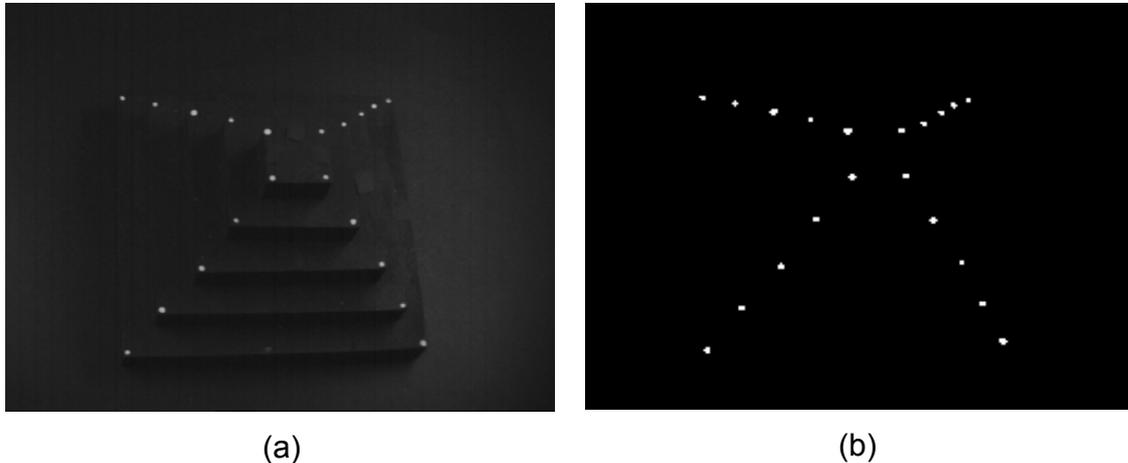


Figure 5-5: Two images for the calibration piece, a) Grey level image b) binary image after thresholding

The output of the second function (image_processing) is in three data files, comprising the object number and x and y coordinates for the calibration points. Table 5-2 shows the output from the image_processing function for one camera.

After implementing the function that automatically matches the calibration points from all images (see section 4.6.1 point 5 for more details), the twelve camera parameters were calculated using the DLT technique (see Appendix A.2 for DLT techniques and camera parameters).

Table 5-3 shows the twelve parameters produced by the DLT technique for the three cameras. This table illustrates that the parameters cannot indicate how much the accuracy is, which is different from the Tsai technique, wherein each parameter has a physical meaning and should be located in a certain interval.

Table 5-2 : x and y coordinates for the calibration points captured by the Right camera

Calibration point No	X coordinate (pixel)	y coordinate (pixel)
1	164.4	134.4
2	171.4	492.4
3	210.0	143.0
4	219.7	432.8
5	264.3	155.1
6	275.0	373.9
7	316.2	165.2
8	323.1	306.6
9	367.1	181.6
10	374.0	246.9
11	442.5	181.2
12	448.5	245.3
13	473.7	171.1
14	487.0	308.1
15	498.3	156.7
16	516.0	146.0
17	526.7	368.5
18	536.0	138.1
19	556.1	426.3
20	584.8	479.7

Table 5-3: Camera parameters for the three cameras by using DLT

Parameter No	Camera 1	Camera 2	Camera 3
1	1.0000	-2.3364	0.4381
2	2.0000	-0.0427	-2.0600
3	3.0000	-0.4654	-1.2168
4	4.0000	134.6732	680.4959
5	5.0000	0.0821	-2.1341
6	6.0000	2.0058	0.0540
7	7.0000	-1.0049	-0.9732
8	8.0000	132.9118	111.6558
9	9.0000	0.0000	0.0004
10	10.0000	-0.0004	0.0005
11	11.0000	-0.0014	-0.0013
12	12.0000	0.0000	0.0000

5.5 Reconstruct the 3D coordinates of the calibration points using DLT

Having the camera parameters for all cameras and the x and y co-ordinates for the calibration points for the three images, the 3D world coordinates (X, Y, Z) for the control points were reconstructed using the DLT reconstruction equations explained in section 3.10 and Appendix A.2.1. Table 5-4 shows the 3D coordinates for the calibration points obtained from the vision system.

Table 5-4 : 3D coordinates for the calibration points obtained from the vision system(The calibration piece placed as shown in figure 5-7)

Points	X (mm)	Y (mm)	Z (mm)
1	2.60	2.41	16.91
2	147.43	3.22	17.75
3	18.36	18.26	34.28
4	133.83	18.14	34.55
5	35.92	32.97	51.83
6	120.43	33.77	52.01
7	50.46	49.97	69.50
8	104.17	49.11	69.41
9	65.33	63.64	86.47
10	89.22	64.32	86.46
11	65.03	87.67	86.55
12	89.71	88.24	86.53
13	51.57	103.43	69.59
14	105.41	101.91	69.67
15	36.30	117.46	52.10
16	120.29	118.27	52.15
17	18.86	132.76	34.41
18	133.04	133.33	34.73
19	3.36	146.64	17.31
20	147.06	147.09	17.62

To see the effect of the location of the calibration piece on the system accuracy, the calibration piece was placed at various positions within the FOV of the three cameras and within the limited movement that was obtained in sections 5-2.1 and 5.2.2.

The data from the vision system was then compared with the data obtained from the coordinate measuring machine. The results are shown below:

First position:

The centre of the calibration piece placed at the base point of the cylinder, see Figure 5-6.

L = 145mm,

h = 520 mm (from the base to the camera),

To compare the performance of the CVS and the CMM the Root Mean Square error and the Standard Deviation were calculated as described below.

Root Mean Square error (RMS) can be calculated using the formula:

$$RMS\ error = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2}{n}} \tag{5-7}$$

Where:

x is the difference between the readings obtained from the two systems, and

n is number of samples.

As the whole data are used in the analysis, so the standard deviation population was found by the following formula:

$$STD = \sqrt{\sum_{i=0}^n (x_i - m)^2} \tag{5-8}$$

Where:

m is the mean.

n is number of samples.

The Root Mean Square (RMS) error and the standard deviation between the CVS obtained from first position shown in Table 5-5.

Table 5-5: RMS error and Standard deviation between the CVS obtained from first position

	X (mm)	Y(mm)	Z(mm)
RMS error	0.18	0.15	0.15
Standard Deviation	0.10	0.08	0.10

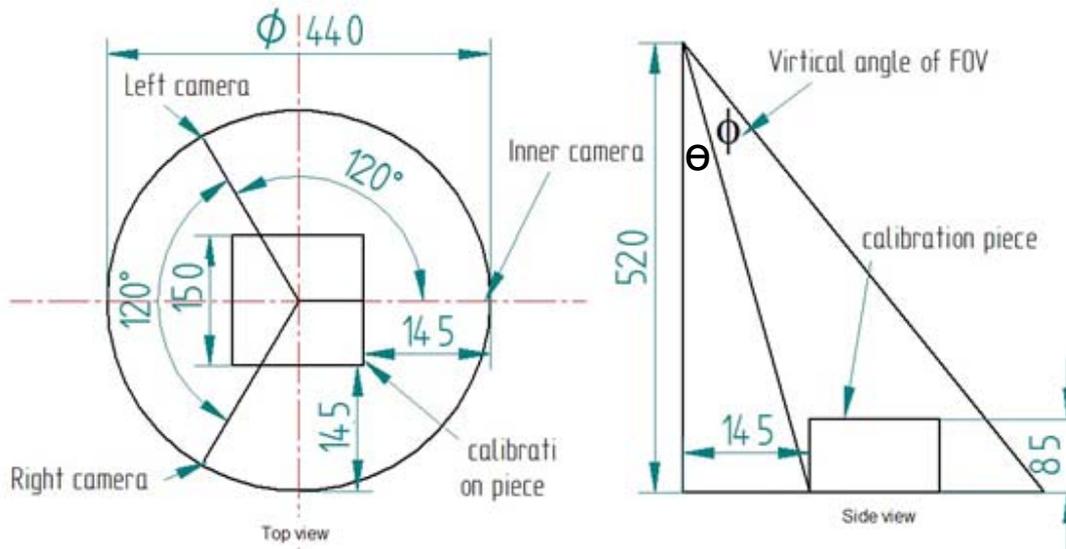


Figure 5-6: First position

Second position:

In this test the vertical distance from the base was increased by 70 mm. (notice that 70 mm > Δz (see step 1 above), but the pyramid shape of the calibration piece gives more space in the vertical direction), i.e. h = 450 mm, and L = 145 mm, see Figure 5-7. The RMS error and the standard deviation between the CVS and CMM are shown in Table 5-6.

Table 5-6: RMS error and Standard deviation between the CVS obtained from second position

	X (mm)	Y(mm)	Z(mm)
RMS error	0.17	0.16	0.23
Standard Deviation	0.11	0.08	0.16

Third position:

L =145,

Vertical distance was reduced by 35mm, i.e. h = 485 mm, see Figure 5-8.

The RMS error and the standard deviation between the CVS and CMM are shown in

Table 5-7: RMS error and Standard deviation between the CVS obtained from third position

	X (mm)	Y (mm)	Z (mm)
RMS error	0.17	0.15	0.19
Standard Deviation	0.09	0.08	0.14

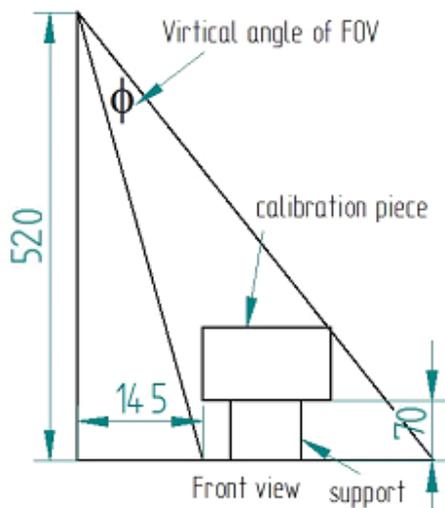


Figure 5-7: Second position, side view (top view same as Figure (5-6))

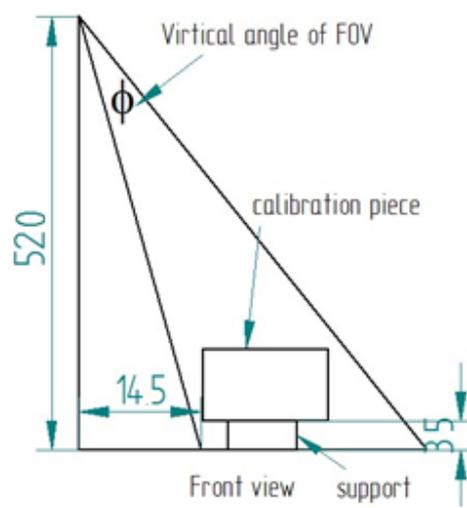


Figure 5-8: Third position, side view (top view same as Figure (5-6))

Forth position:

H = 520mm, L =145mm. The calibration piece was rotated by 15°, see Figure 5-9.

The RMS error and the standard deviation between the CVS and CMM are shown in Table 5-8.

Table 5-8; RMS error and Standard deviation between the CVS obtained from fourth position

	X (mm)	Y(mm)	Z(mm)
RMS error	0.16	0.18	0.22
Standard Deviation	0.10	0.10	0.17

Fifth position:

H = 520mm, L = 145mm and the calibration piece was shifted forward by 30mm, Figure 5-10.

The RMS error and the standard deviation between the CVS and CMM are shown in Table 5-9.

Table 5-9: RMS error and Standard deviation between the CVS obtained from fifth position

	X (mm)	Y (mm)	Z (mm)
RMS error	0.16	0.14	0.17
Standard Deviation	0.09	0.08	0.09

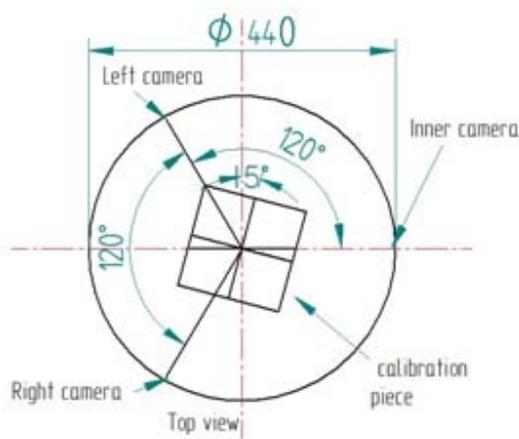


Figure 5-9: Forth position, Top view (side view same as Figure (5-6))

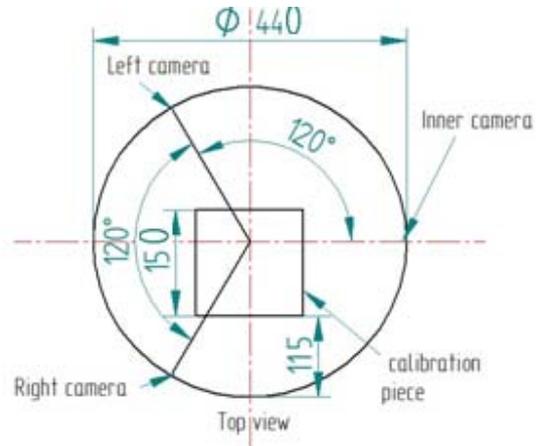


Figure 5-10: Fifth position, Top view (side view same as Figure (5-6))

Sixth position:

H = 520mm, L 145mm and the calibration piece was shifted backward by 30mm as depicted in Figure 5-11.

The RMS error and the standard deviation between the CV system and CMM are shown in Table 5-10.

Table 5-10: RMS error and Standard deviation between the CVS obtained from sixth position

	X (mm)	Y(mm)	Z(mm)
RMS error	0.19	0.18	0.27
Standard Deviation	0.10	0.10	0.20

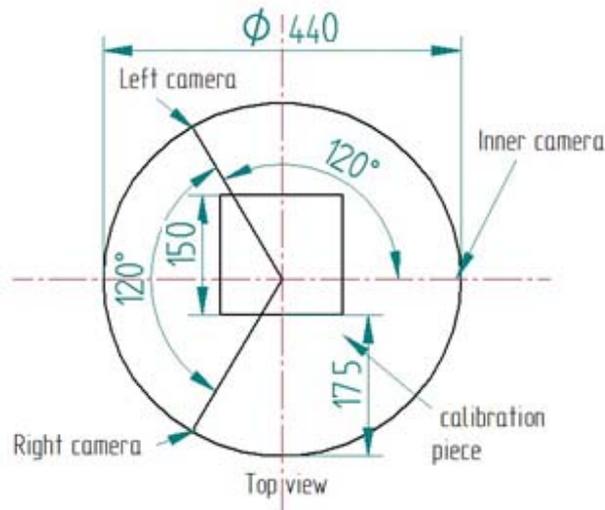


Figure 5-11: Sixth position, Top view(side view same as Figure (5-6))

From the all six results above, the absolute mean errors between the CMM and the CVS in X, Y and Z are summarized in Table 5-11.

Table 5-11: The absolute mean errors between the CMM and the CVS

	X(mm)	Y(mm)	Z(mm)
RMSE, first position	0.18	0.15	0.15
RMSE, second position	0.17	0.16	0.23
RMSE, third position	0.17	0.15	0.19
RMSE, fourth position	0.16	0.18	0.22
RMSE, fifth position	0.16	0.14	0.17
RMSE, sixth position	0.19	0.18	0.27

Mean error	0.17	0.16	0.20
St Dev	0.01	0.02	0.04

From table 5-5 above, Z coordinate shows slightly higher mean error and St Dev (but not significant) than the X and Y coordinates. This is probably because the Z coordinates of the cameras are larger than their X and Y coordinates, due to the physical constraints of the apparatus.

In general the standard deviation for the six tests in X, Y, and Z are very low. In other words it can be concluded that, variations in placing the calibration piece over a small range of locations within the FOV of the cameras as calculated in sections 5.2.1 and 5.2.2, does not significantly affect the error between the CVS and the CMM. Moreover this relaxes the positioning constraints of the calibration piece, making it easier for a non-expert to set up the calibration procedure.

5.6 Determining the effect of the error in X, Y, Z coordinates on the value of finger joint angles

The ultimate aim of this work is to measure the angles for the finger joints, each joint being defined by the intersection of two vectors in space. To indicate how much the error in terms of X, Y and Z co-ordinates will affect the angles of interest, the following formulae have been derived.

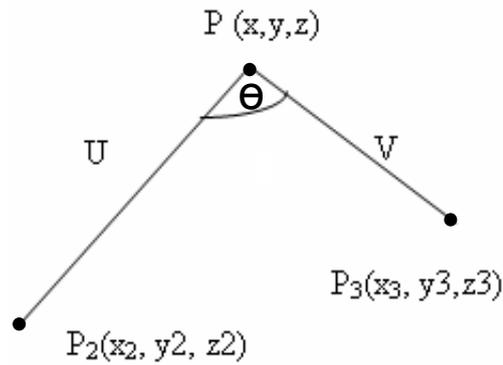


Figure 5-12: Angle between two vectors in space

Let: $P(x, y, z), P_2(x_2, y_2, z_2), P_3(x_3, y_3, z_3)$ in Figure 5-12 be three points in space, so the two vectors between these points are:

$$PP_2 = [x_2 - x, y_2 - y, z_2 - z] = U \text{ and}$$

$$PP_3 = [x_3 - x, y_3 - y, z_3 - z] = V$$

and let theta (θ) be the angle between the two vectors, so that

$$\begin{aligned} \cos(\theta) &= \frac{U \cdot V}{|U||V|} = \\ &= \frac{[x_2 - x, y_2 - y, z_2 - z] \cdot [x_3 - x, y_3 - y, z_3 - z]}{\sqrt{(x_2 - x)^2 + (y_2 - y)^2 + (z_2 - z)^2} \sqrt{(x_3 - x)^2 + (y_3 - y)^2 + (z_3 - z)^2}} = \\ &= \frac{[x^2 - (x_2 + x_3)x + y^2 - (y_2 + y_3)y + z^2 - (z_2 + z_3)z + x_2x_3 + y_2y_3 + z_2z_3]}{\sqrt{x^2 - 2x_2x + x_2^2 + y^2 - 2y_2y + y_2^2 + z^2 - 2z_2z + z_2^2} \sqrt{x^2 - 2x_3x + x_3^2 + y^2 - 2y_3y + y_3^2 + z^2 - 2z_3z + z_3^2}} \end{aligned} \quad (5-9)$$

Squaring each side of the equation above, we get

$$\begin{aligned} \cos^2(\theta) &= \\ &= \frac{[x^2 - (x_2 + x_3)x + y^2 - (y_2 + y_3)y + z^2 - (z_2 + z_3)z + x_2x_3 + y_2y_3 + z_2z_3]^2}{[x^2 - 2x_2x + x_2^2 + y^2 - 2y_2y + y_2^2 + z^2 - 2z_2z + z_2^2][x^2 - 2x_3x + x_3^2 + y^2 - 2y_3y + y_3^2 + z^2 - 2z_3z + z_3^2]} \end{aligned} \quad (5-10)$$

Now let N=

$$[x^2 - (x_2 + x_3)x + y^2 - (y_2 + y_3)y + z^2 - (z_2 + z_3)z + x_2x_3 + y_2y_3 + z_2z_3]^2 \quad (5-11)$$

And let M =

$$[x^2 - 2x_2x + x_2^2 + y^2 - 2y_2y + y_2^2 + z^2 - 2z_2z + z_2^2][x^2 - 2x_3x + x_3^2 + y^2 - 2y_3y + y_3^2 + z^2 - 2z_3z + z_3^2] \quad (5-12)$$

And let

$$A = [x^2 - 2x_2x + x_2^2 + y^2 - 2y_2y + y_2^2 + z^2 - 2z_2z + z_2^2] \quad (5-13)$$

$$B = [x^2 - 2x_3x + x_3^2 + y^2 - 2y_3y + y_3^2 + z^2 - 2z_3z + z_3^2] \quad (5-14)$$

Hence

$$M = A \cdot B, \text{ so } \cos^2(\theta) = \frac{N}{M} \quad (5-15)$$

To find the change of theta with respect to x, y and z for point P we find $\frac{\partial \theta}{\partial x}, \frac{\partial \theta}{\partial y}$

and $\frac{\partial \theta}{\partial z}$.

From equation (5-11), taking partial derivatives of θ with respect to x gives

$$(-2 \cos \theta \sin \theta) \frac{\partial \theta}{\partial x} = \frac{M \cdot \frac{\partial N}{\partial x} - N \frac{\partial M}{\partial x}}{M^2} \quad (5-16)$$

From equation (5-9)

$$\frac{\partial N}{\partial x} = 2\sqrt{N}(2x - (x_2 + x_3)) \quad (5-17)$$

and from equation (5-10)

$$\begin{aligned} \frac{\partial M}{\partial x} &= A \frac{\partial B}{\partial x} + B \frac{\partial A}{\partial x} \\ &= A \cdot (2x - 2x_3) + B \cdot (2x - 2x_2) \end{aligned} \quad (5-18)$$

By substituting equations (5-15, 5-16) in (5-14) we obtain

$$\frac{\partial \theta}{\partial x} = \left(\frac{-1}{\cos \theta \sin \theta} \right) \left[\frac{2M\sqrt{N}(2x - (x_2 + x_3)) - N(A(2x - 2x_3) + B(2x - 2x_2))}{M^2} \right] \quad (5-19)$$

$$\text{Given } \sin^2 \theta + \cos^2 \theta = 1 \rightarrow \sin \theta = \sqrt{1 - \cos^2 \theta} \rightarrow \sin \theta = \sqrt{1 - \frac{N}{M}} \quad (5-20)$$

By putting the values of $\cos \theta$ and $\sin \theta$ from equations (5-15 and 5-20) in equation (5-19), we obtain

$$\frac{\partial \theta}{\partial x} = \left(\frac{-1}{\sqrt{\frac{N}{M}} \sqrt{1 - \frac{N}{M}}} \right) \left[\frac{2M\sqrt{N}(2x - (x_2 + x_3)) - N(A(2x - 2x_3) + B(2x - 2x_2))}{M^2} \right] \quad (5-21)$$

Similarly

$$\frac{\partial \theta}{\partial y} = \left(\frac{-1}{\sqrt{\frac{N}{M}} \sqrt{1 - \frac{N}{M}}} \right) \left[\frac{2M\sqrt{N}(2y - (y_2 + y_3)) - N(A(2y - 2y_3) + B(2y - 2y_2))}{M^2} \right] \quad (5-22)$$

$$\frac{\partial \theta}{\partial z} = \left(\frac{-1}{\sqrt{\frac{N}{M}} \sqrt{1 - \frac{N}{M}}} \right) \left[\frac{2M\sqrt{N}(2z - (z_2 + z_3)) - N(A(2z - 2z_3) + B(2z - 2z_2))}{M^2} \right] \quad (5-23)$$

Similarly for the Point P_2 we have

$$\frac{\partial \theta}{\partial x_2} = \left(\frac{-1}{\sqrt{\frac{N}{M}} \sqrt{1 - \frac{N}{M}}} \right) \left[\frac{2M\sqrt{N}(x_3 - x) - N(B(2x_2 - 2x))}{M^2} \right] \quad (5-24)$$

$$\frac{\partial \theta}{\partial y_2} = \left(\frac{-1}{\sqrt{\frac{N}{M}} \sqrt{1 - \frac{N}{M}}} \right) \left[\frac{2M\sqrt{N}(y_3 - y) - N(B(2y_2 - 2y))}{M^2} \right] \quad (5-25)$$

$$\frac{\partial \theta}{\partial z_2} = \left(\frac{-1}{\sqrt{\frac{N}{M}} \sqrt{1 - \frac{N}{M}}} \right) \left[\frac{2M\sqrt{N}(z_3 - z) - N(B(2z_2 - 2z))}{M^2} \right] \quad (5-26)$$

And for the Point P_3 we have

$$\frac{\partial \theta}{\partial x_3} = \left(\frac{-1}{\sqrt{\frac{N}{M}} \sqrt{1 - \frac{N}{M}}} \right) \left[\frac{2M\sqrt{N}(x_2 - x) - N(A(2x_3 - 2x))}{M^2} \right] \quad (5-27)$$

$$\frac{\partial \theta}{\partial y_3} = \left(\frac{-1}{\sqrt{\frac{N}{M}} \sqrt{1 - \frac{N}{M}}} \right) \left[\frac{2M\sqrt{N}(y_2 - y) - N(A(2y_3 - 2y))}{M^2} \right] \quad (5-28)$$

$$\frac{\partial \theta}{\partial z_3} = \left(\frac{-1}{\sqrt{\frac{N}{M}} \sqrt{1 - \frac{N}{M}}} \right) \left[\frac{2M\sqrt{N}(z_2 - z) - N(A(2z_3 - 2z))}{M^2} \right] \quad (5-29)$$

A change in the angle between two vectors in space with respect to the 3D world co-ordinates depends on the location of these vectors in space. For instance, if the two vectors are located in a plane parallel to the Y-Z plane, that means $\frac{\partial \theta}{\partial x} = 0$, $\frac{\partial \theta}{\partial y} \neq 0$ and $\frac{\partial \theta}{\partial z} \neq 0$. The change of θ with respect to the X, Y, Z coordinates is directly proportional to the value of θ .

Given that the fingers are placed parallel to the y-axis as depicted in Figure 5-13 and that the two vectors U and V that form the finger joint angle θ are as shown in Figure 5-14, the minimum change of θ will be with respect to the X coordinate.

Let the values of P, P_2, P_3 be (57, 50, 77), (40, 50, 60) and (60, 40, 65) mm respectively; these three points are located within the control volume of the calibration piece (Figure 5-1). Then the magnitude of the vectors $PP_2 = 24$ mm and $PP_3 = 15.9$ mm, which are similar to the distance between pairs of measurement points on the fingers of the hand model.

To find out the rate of change of theta with respect to x, y and z at the given points P, P_2, P_3 , we substitute these 3D coordinates into equations (5-19, 5-20, 5-21).

MATLAB code was written to find $\frac{\partial \theta}{\partial x}$, $\frac{\partial \theta}{\partial y}$ and $\frac{\partial \theta}{\partial z}$, the values were calculated as

follows:

$$\frac{\partial \theta}{\partial x} = -1.84 \text{ degree/mm}$$

$$\frac{\partial \theta}{\partial y} = -0.64 \text{ degree/mm}$$

$$\frac{\partial \theta}{\partial z} = -2.88 \text{ degree/mm}$$

Given that 1 pixel = 0.38 mm, from the vision system, we have 1 mm = 2.63 pixel,

then $\frac{\partial \theta}{\partial x} = -0.70 \text{ degree/pixel}$, $\frac{\partial \theta}{\partial y} = -0.24 \text{ degree/pixel}$ and $\frac{\partial \theta}{\partial z} = -1.0$

degree/pixel

Similarly

$$\frac{\partial \theta}{\partial x_2} = -1.22 \text{ degree/mm} = -0.46 \text{ degree/pixel}$$

$$\frac{\partial \theta}{\partial y_2} = 1.63 \text{ degree/mm} = 0.62 \text{ degree/pixel}$$

$$\frac{\partial \theta}{\partial z_2} = 1.22 \text{ degree/mm} = 0.46 \text{ degree/pixel}$$

$$\frac{\partial \theta}{\partial x_3} = 3.07 \text{ degree/mm} = 1.16 \text{ degree/pixel}$$

$$\frac{\partial \theta}{\partial y_3} = -0.98 \text{ degree/mm} = -0.37 \text{ degree/pixel}$$

$$\frac{\partial \theta}{\partial z_3} = 1.59 \text{ degree/mm} = 0.6 \text{ degree/pixel}$$

The total changes of θ with respect to x, y and z coordinates are:

$$\frac{\partial \theta}{\partial x} + \frac{\partial \theta}{\partial x_2} + \frac{\partial \theta}{\partial x_3} = 0.01 \text{ degree/mm}$$

$$\frac{\partial \theta}{\partial y} + \frac{\partial \theta}{\partial y_2} + \frac{\partial \theta}{\partial y_3} = 0.01 \text{ degree/mm}$$

$$\frac{\partial \theta}{\partial z} + \frac{\partial \theta}{\partial z_2} + \frac{\partial \theta}{\partial z_3} = -0.07 \text{ degree/mm}$$

This means if the errors in X, Y, Z in mm for instance are 0.156, 0.128 and 0.112 respectively (these values are the average absolute error between CMM and the vision system for the calibration piece shown in Figure 5-2), then θ will change by 0.0015° , 0.0012° and 0.0078° for X, Y, and Z. From these results it can be seen that the error between the readings obtained from the CMM and from the vision system has an insignificant effect on the deduced angle between the two vectors in space. We can therefore conclude at this stage that the level of joint angle measurement accuracy, for all the fingers measured by this computer vision system is very encouraging.

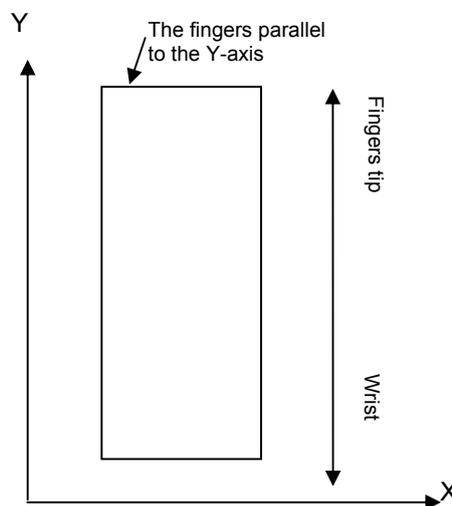


Figure 5-13: Position of the fingers with respect to world co-ordinates (Top view)

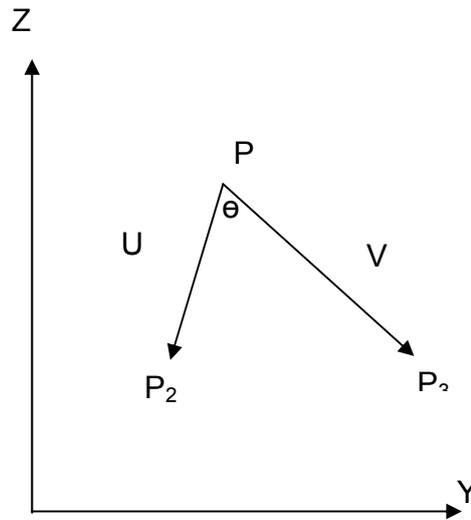


Figure 5-14: Two vectors in space represent finger angle joint (Side view)

Up to now the finger joint measurements have been carried out on one model finger which was designed previously by Hemmings (2002) as shown in Figure 4-16. This finger contains three angles, each angle formed from the intersection of two vectors in space, so 8 points are needed to measure these angles.

After the cameras had been calibrated, the three angles were measured by the CMM machine. To measure these angles by the CMM, firstly the 3D world coordinates for the 8 points were found as explained in section 5.3. To measure any joint angle, the four points which represent the two vectors that form this angle are selected, then the CMM software calculate the angle as described in section 4.7.

The measurement process by the CMM was repeated three times to find out how much human error can affect the evaluation of the angles, see Table 5-12.

Table 5-12: Finger angles measurement from CMM

Finger joint	First measurement (degree)	Second measurement (degree)	Third measurement (degree)	Max (degree)	Min (degree)	Range (Max-Min) (degree)
MCP joint	20.19	20.49	20.10	20.49	20.10	0.39
PIP joint	17.03	16.73	17.28	17.28	16.73	0.55
DIP joint	28.07	28.10	27.92	28.10	27.92	0.18

From Table 5-12 above, there is some discrepancy between the angles obtained from the CMM (mainly due to human error); the highest difference between the maximum and the minimum reading was 0.55°. Thus we have a worst case error of ± 0.27 degrees in the CMM readings (half the maximum range).

The finger was then placed inside the control volume, and the computer vision measurement process was carried out to reconstruct the 3D world coordinates of the measurement points as discussed in section 5.5. Then joint finger angles were calculated based on the formula shown in section 4-7. Table 5-13 displays the results from the computer vision system.

Table 5-13 : Finger angles measurement for the CVS

Finger joint	Angle (degrees)
MCP joint	20.6
PIP joint	16.82
DIP joint	28.9

Comparing, the tables 5-6 and 5-7 above, the PIP joint angle obtained from the CVS lies within the range of the measurements that were obtained from the CMM.

Thus, the MCP joint angle measured by the CVS can be considered as within the range of the measurement obtained by the CMM machine, as the difference between the two measurements is about 0.1 degrees.

However, in the case of the DIP angle, the angle measured by the CV system is outside the range of the CMM machine measurement by 0.8 degrees, which is quite high compared with the PIP and MCP angles. For reasons that may be behind this error, see Figure 5-17 and the discussion related to that figure. Nevertheless, this error is still below the specification target of this work which is 1 degree.

The absolute error between the three measurements from the CMM and readings from the CV system are shown in Table 5-14.

Table 5-14 : Absolute error between the three measurements from the CMM and readings from the CV system

Finger joint	Absolute error, first measurement (degree)	Absolute error, second measurement (degree)	Absolute error, third measurement (degree)	Average error (degree)
MCP joint	0.46	0.16	0.54	0.38
PIP joint	0.186	0.12	0.43	0.24
DIP joint	0.82	0.79	0.97	0.86
Mean error	0.49			

From Table 5-14 the overall mean error between the CMM and the CVS when one finger joints were measured is 0.5°, which is lower than our target.

5.7 Measuring all 4 fingers of the hand

So far, the angles for only one finger have been evaluated. To measure the joint angles for all four fingers, a new calibration piece was designed, as the size of

the one shown in Figure 5-1 was not sufficient to enclose all four fingers. The new calibration piece was designed according to the factors discussed in section (4.2.6) and is shown in Figure 4-9.

In this test the FOV of the cameras was increased by increasing the height of the three cameras to 800 mm (see Figure 5-15) in order to cover the calibration piece.

The new calibration piece has 27 calibration points evenly distributed, the 3D world coordinates for these points were measured by CMM and then it obtained from the Computer Vision System (CVS) the data is shown in Table 5-15.

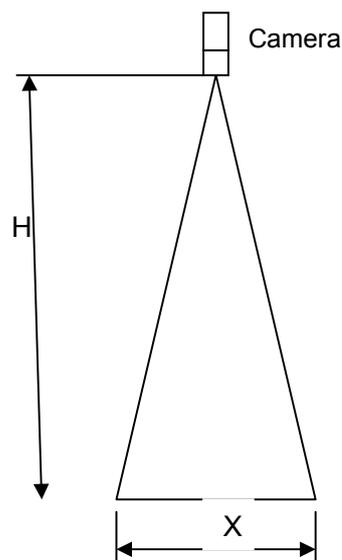


Figure 5-15: Increasing the height of the camera (H) to 800 mm leads to increase the field of view (X)

Once the 3D coordinates of the calibration piece had been obtained, it was placed at a fixed position inside the working volume by using two steel strips with L cross section as location stops. Placing the calibration piece at a fixed position helps to establish an automatic correspondence between the calibration points captured by different cameras. The cameras' parameters were then obtained by using the DLT technique.

The 3D coordinates for the calibration points were reconstructed and compared with the ones from the CMM. Table 5-16 displays the mean and standard deviation of the error between CVS and CMM for the calibration points in mm.

Table 5-15 : 3D world coordinates of the calibration points for the calibration piece shown in Figure 4-9 (measured by CVS and CMM)

Calibration point	Measured by CVS			Measured by CMM		
	X(mm)	Y(mm)	Z(mm)	X(mm)	Y(mm)	Z(mm)
1	-21.53	19.90	22.68	-22.26	19.54	24.59
2	-25.13	11.28	80.26	-24.52	11.43	79.75
3	-21.19	12.84	137.38	-20.53	13.18	137.03
4	-21.37	101.13	22.49	-21.16	100.99	22.78
5	-19.70	100.36	80.91	-19.58	100.41	81.00
6	-18.15	100.03	138.90	-18.37	99.91	138.85
7	-22.81	184.01	24.29	-22.43	184.01	23.91
8	-23.15	188.09	82.01	-23.28	187.22	83.05
9	-23.23	187.59	138.11	-23.86	186.17	140.06
10	-99.81	11.02	23.38	-99.29	11.66	22.29
11	-100.91	10.56	81.19	-100.66	10.98	80.71
12	-95.22	11.90	139.36	-95.03	12.00	139.78
13	-97.82	98.17	22.03	-97.58	98.56	21.67
14	-99.26	100.23	81.32	-99.31	100.71	81.16
15	-102.25	99.62	140.60	-102.60	99.88	140.46
16	-100.40	187.32	23.58	-100.65	187.39	23.11
17	-100.11	189.84	81.84	-100.59	189.02	82.67
18	-97.43	188.62	139.00	-97.97	186.93	140.64
19	-178.26	13.66	24.94	-177.92	13.86	24.04
20	-178.73	11.57	81.76	-178.80	11.57	81.43
21	-177.37	12.23	138.17	-177.74	12.19	138.49
22	-182.51	100.46	23.49	-182.56	100.60	23.53
23	-181.93	99.51	81.68	-182.24	99.62	81.56
24	-181.69	99.36	139.83	-182.39	99.59	139.11
25	-177.37	185.03	25.00	-177.96	184.98	25.08
26	-177.75	188.60	82.46	-178.44	188.19	83.05
27	-176.36	187.95	138.90	-177.11	186.97	140.09

Table 5-16: Absolute error between CMM and CVS (see Table 5-15)

Calibration point	X(mm)	Y(mm)	Z(mm)
1	0.73	0.36	1.91
2	0.61	0.15	0.51
3	0.66	0.34	0.35
4	0.21	0.14	0.29
5	0.12	0.05	0.09
6	0.22	0.13	0.05
7	0.38	0.00	0.38
8	0.13	0.87	1.03
9	0.63	1.42	1.95
10	0.53	0.64	1.09
11	0.25	0.42	0.49
12	0.19	0.09	0.42
13	0.24	0.39	0.36
14	0.05	0.49	0.16
15	0.35	0.27	0.13
16	0.25	0.07	0.47
17	0.48	0.82	0.82
18	0.54	1.70	1.64
19	0.34	0.21	0.90
20	0.07	0.01	0.33
21	0.37	0.04	0.32
22	0.05	0.15	0.04
23	0.31	0.10	0.12
24	0.70	0.23	0.72
25	0.59	0.05	0.08
26	0.69	0.41	0.59
27	0.75	0.99	1.19
Mean error	0.39	0.39	0.61
STDEV	0.22	0.42	0.54

Looking to Table 5-16 above, the mean error in X direction is 0.39 mm and the standard deviation is 0.22 mm which means the data is very close to the mean. For the error in Y direction the mean error is same as X direction; however the standard deviation is higher than X direction. Looking to the data in Y direction it is clear that most of the data is very small, i.e. very close to the mean except two readings; 1.7 mm and 1.42 mm (see Table 5-16). There is no explanation for these high errors unless; they could be a human error while using CMM (see Table 5-14) or/and errors from electronic devices that composes the CVS.

In the case of the error from the Z direction, the mean and standard deviation is higher than both X and Y directions, the same reason behind Y direction can be mentioned here. Nevertheless and in general the data in Z coordinate is slightly higher than the X and Y coordinates. This is could be the same reason that mentioned about the data shown in Table 5-11. Reducing the Z coordinates of the cameras will lead to a reduction in the FOV, which is not acceptable. Also increasing the X and Y coordinates of the cameras will lead to an increase in the working volume which is undesirable as well.

The hand model depicted in Figure 4-10 was then placed into the control volume, so the finger joint angles could be evaluated. In order to test the robustness of the system, two categories of angles were measured. In the first category, the range of joints angles varied between 21° and 43° (the finger joints can be bent up to around 100°). However these angles have been used in order to enable the cameras to see the measurement points directly). In the second category, the angles varied between 2° and 8°. Also, for each category the hand model was placed at different positions within the control volume. Then the angles measured by the CMM and the CVS were compared, and to avoid compound angles the four measurement points were placed so their top projection shapes straight line, as much as possible. The results from the two categories were as below:

A) First category (joints angles varied between 21° and 43°)

The values of pre-set joint angles measured by the CMM (Table 5-17) were:

Table 5-17: Joint angles measured by CMM

	MCP joint(°)	PIP joint(°)	DIP joint(°)
First finger	21.35	23.39	33.16
Second finger	23.47	30.72	33.62
Third finger	22.89	20.32	28.65
Fourth finger	27.71	38.29	41.27

The hand model with above set angles was placed at different positions in the control volume and the measurement process was carried out as illustrated below:

First position

The hand was placed at the centre of the control volume, see Figure 5-16. The absolute errors between the CMM and the CVS are shown in Table 5-18.

Table 5-18: Absolute errors between the CMM and the CVS (first position)

	MCP joint(°)	PIP joint(°)	DIP joint(°)
First finger	0.23	0.14	0.11
Second finger	0.04	0.92	0.04
Third finger	0.25	0.05	0.06
Fourth finger	0.61	0.47	0.59
Mean	0.29		
STDEV	0.28		

From Table 5-18, the mean error is 0.29° and the standard deviation is 0.28°. This is a very low error as most of the results are much less than one degree except one angle which 0.92 degree.

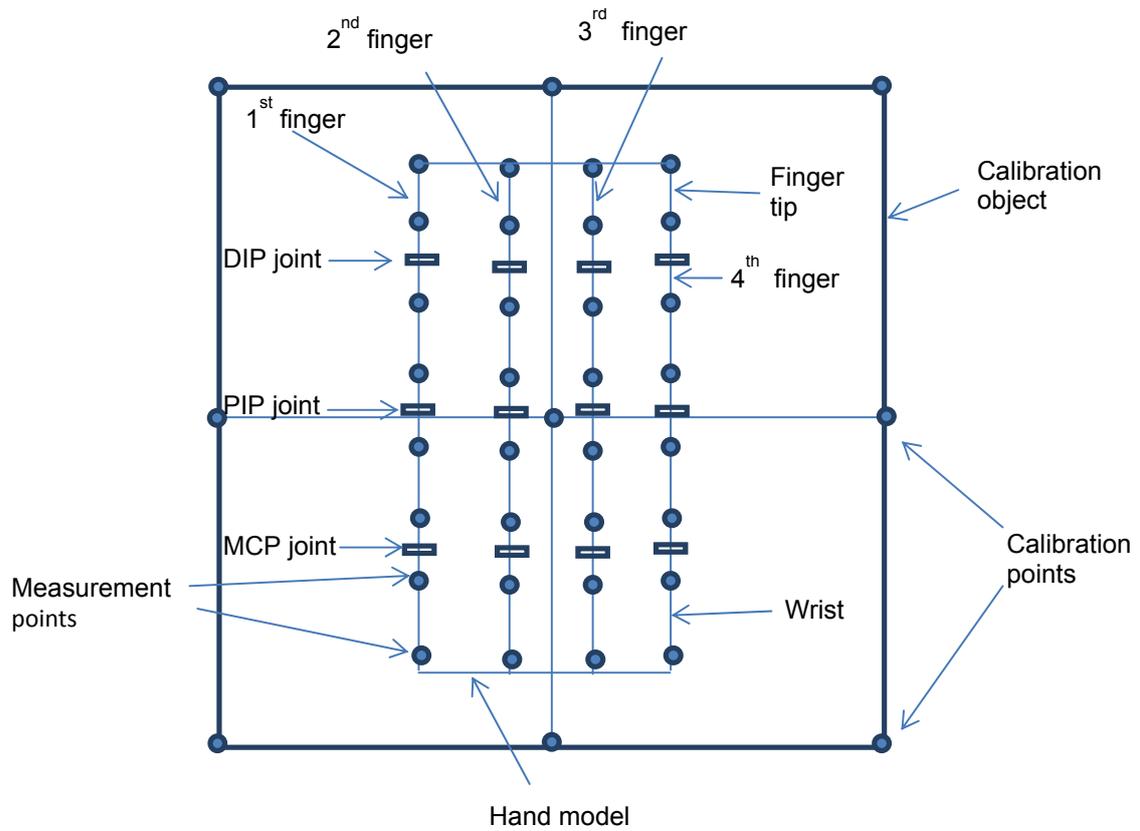


Figure 5-16: First position, hand model placed at the centre of the calibration piece.

Second position

The hand model was shifted to the right by 30 mm from the first position, see Figure 5-17. The absolute errors between the CMM and the CVS are shown in Table 5-19:

Table 5-19: Absolute errors between the CMM and the CVS (second position)

	MCP joint(°)	PIP joint(°)	DIP joint(°)
First finger	0.08	0.65	0.23
second finger	0.15	0.47	0.15
Third finger	0.40	0.71	1.18
Fourth finger	0.01	1.24	0.61
Mean	0.49		
STDEV	0.39		

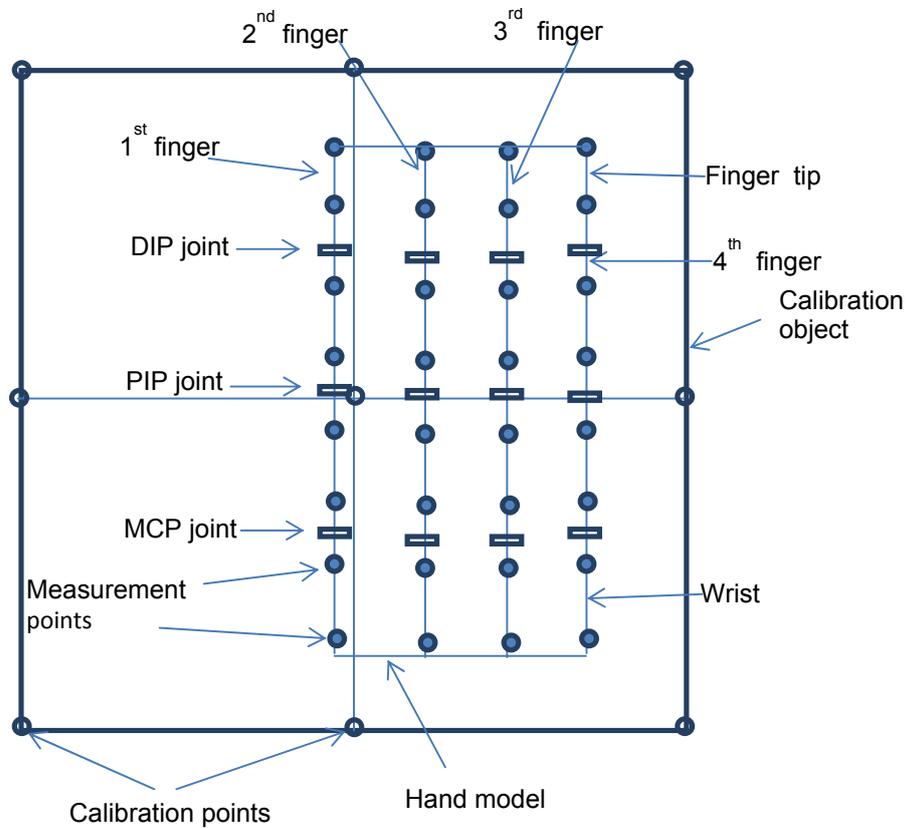


Figure 5-17: Second position, the hand model shifted from the first position to the right by 30 mm

Third position

The hand model was shifted to the left by 30 mm from the first position, see Figure 5-18. The absolute errors between the CMM and the CVS are shown in Table 5-20.

Table 5-20: Absolute errors between the CMM and the CVS (third position)

	MCP joint(°)	PIP joint(°)	DIP joint(°)
First finger	0.35	0.88	0.15
Second finger	0.20	0.02	0.47
Third finger	0.47	0.29	0.43
Fourth finger	0.61	2.35	1.80
Mean	0.67		
STDEV	0.67		

Comparing Table 5-19 and Table 5-20, the errors in MCP joints for the two measurements are very close to each other and they are much less than 1 degree. However the PIP and DIP joints of the fourth finger obtained from the third position is significantly higher than the errors at the same joints obtained from the second position. This should not be the case, as moving the hand model by 30 mm to the right or to the left from the first position, should not have significant difference between the two positions because in both cases the hand model located within the control volume which has evenly distributed calibration points.

Moreover, at the second position the fourth finger located at the same position with respect to the calibration points as the first finger in the third position as shown in Figure 5-17 and Figure 5-18. So the errors in these two fingers should be similar. However the errors increased from 0.65 degree to 2.35 degree for the PIP joint and from 0.23 degree to 1.80 degree for the DIP joint. Having that the two tests were done under the same conditions and the same camera parameters were used, the difference only was in the location of the hand model

which should not have a significant effect in the results. So these random errors could be from the electronic devices.

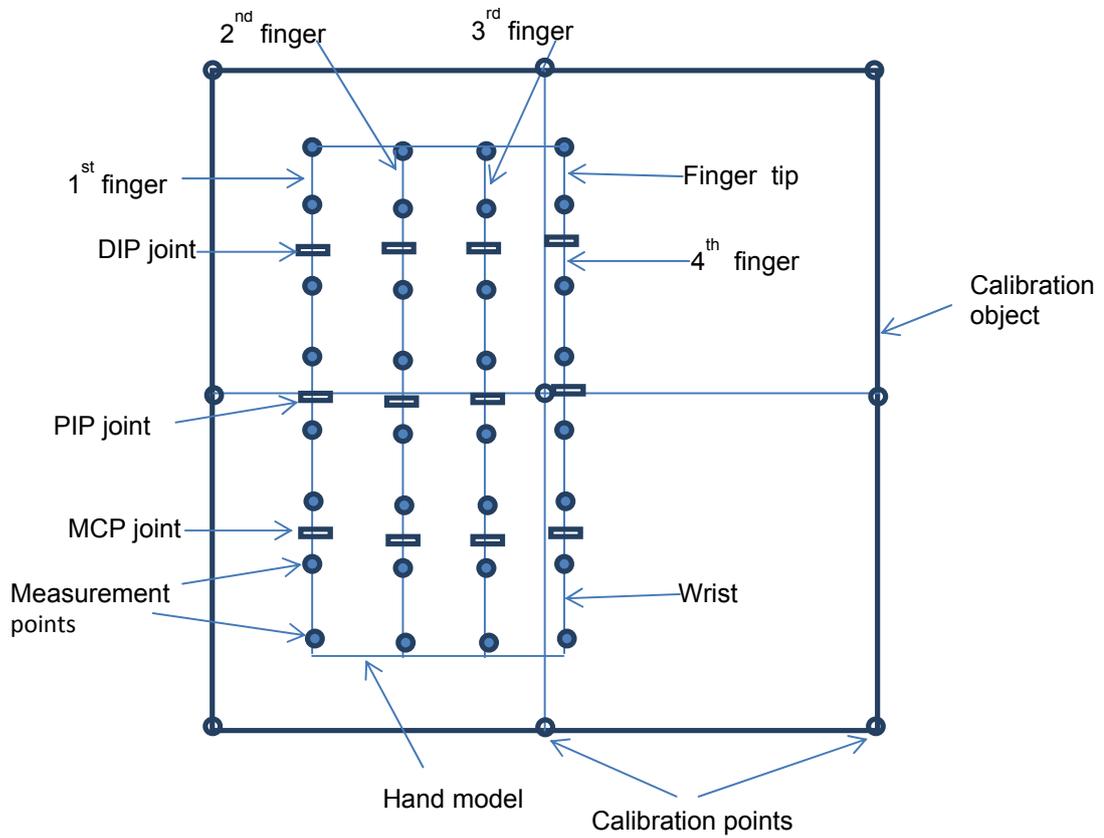


Figure 5-18: Third position, the hand model shifted from the first position to the left by 30 mm

Forth position

The hand model was rotated to the right by 15°, see Figure 5-19. The average errors between the two readings are shown in Table 5-21.

Table 5-21: Absolute errors between the CMM and the CVS (fourth position)

	MCP joint(°)	PIP joint(°)	DIP joint(°)
First finger	0.11	0.65	1.03
Second finger	0.04	0.08	0.23
Third finger	0.10	0.04	0.03
Fourth finger	0.96	0.44	0.32
Mean	0.34		
STDEV	0.35		

From Table 5-21 the mean and the standard deviation are similar to the first position. The table shows twelve errors, out of 12, which are much less than one degree and two errors are almost on degree each.

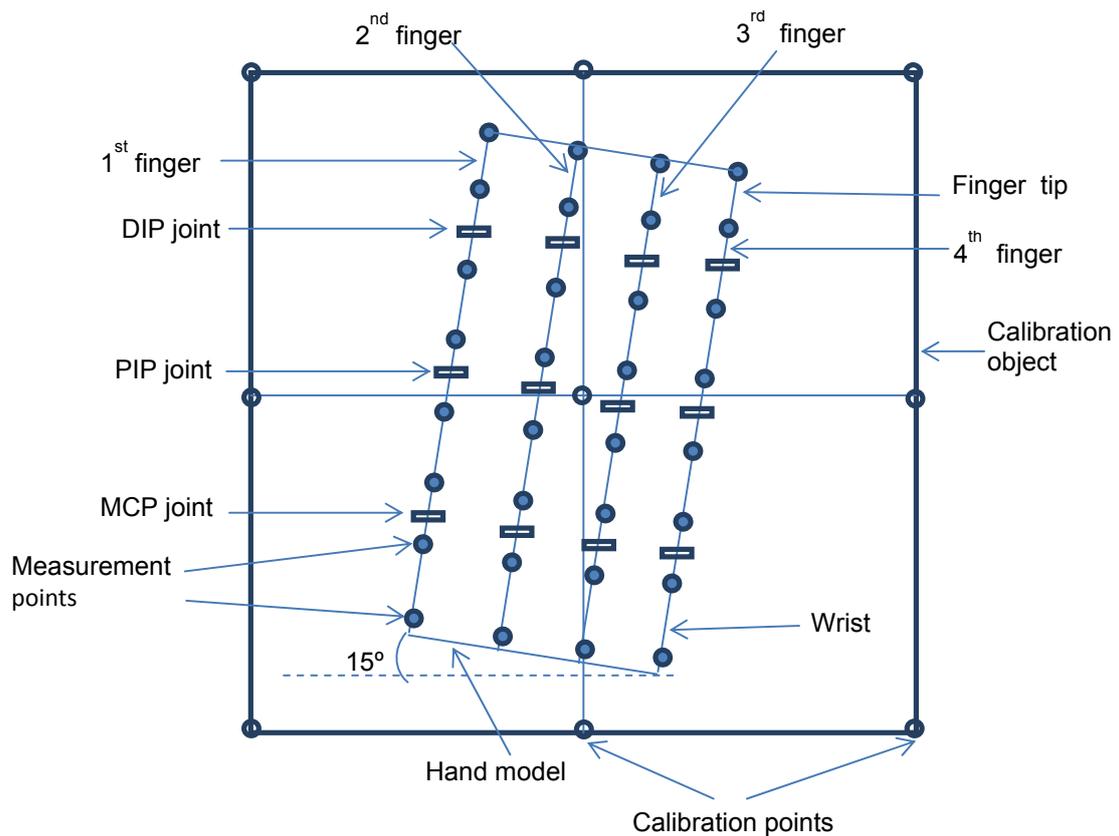


Figure 5-19: Fourth position, the hand model was rotated by 15 degrees to the right from first position

Fifth position

The hand model was rotated to the left by 15°, see Figure 5-20. The absolute errors between the CMM and the CVS are shown in Table 5-22.

Table 5-22: Absolute errors between the CMM and the CVS (fifth position)

	MCP joint(°)	PIP joint(°)	DIP joint(°)
First finger	0.02	0.28	0.19
Second finger	0.11	0.01	0.70
Third finger	0.60	0.83	0.20
Fourth finger	0.45	1.11	0.74
Mean	0.44		
STDEV	0.34		

From Table 5-22 the mean and standard deviation is very low. And all the errors are much less than the target of this work, except the PIP joint for the fourth finger which is 1.11 degree. Nevertheless it is still very acceptable as it is not significantly higher than one degree.

By comparing Table 5-21 and Table 5-22 it can be concluded that, the rotating the handle model to the right and to the left produced errors much less than the target error.

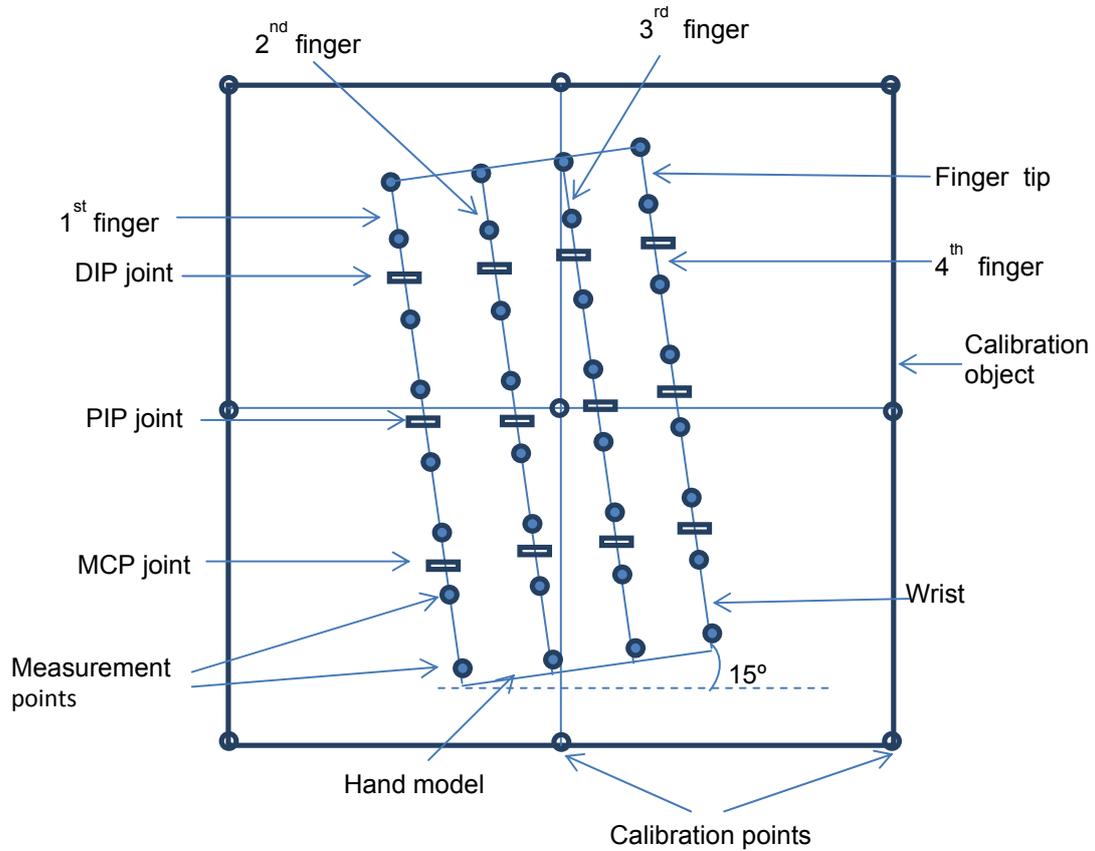


Figure 5-20: Fifth position, the hand model was rotated by 15 degrees to the left from first position.

Sixth position

The hand model was same as first position and lifted upward by 50 mm. The absolute errors between the CMM and the CVS are shown in Table 5-23.

Table 5-23: Absolute errors between the CMM and the CVS (sixth position)

	MCP joint(°)	PIP joint(°)	DIP joint(°)
First finger	0.45	0.46	0.10
second finger	0.32	0.92	1.18
Third finger	0.17	1.15	1.14
Fourth finger	0.61	1.93	1.42
Mean	0.82		
STDEV	0.53		

From Table 5-23 above the mean error and the standard deviation are higher than the first position and they should be similar as the first position. This result as a consequence of the high errors that produced by the PIP and DIP joints of the fourth finger. The reason of these errors could be the same one that disrobed about the result shown in Table 5-20.

Seventh position

The hand model was same as third position and lifted upward by 50 mm. The absolute errors between the CMM and the CVS are shown in Table 5-24.

Table 5-24: Absolute errors between the CMM and the CVS (seventh position)

	MCP joint(°)	PIP joint(°)	DIP joint(°)
First finger	1.08	0.60	0.82
Second finger	0.07	0.32	0.73
Third finger	0.30	0.06	0.87
Fourth finger	0.34	0.30	2.24
Mean	0.64		
STDEV	0.57		

All the errors shown in Table 5-24 are very low, i.e. less than the target; however the error from the DIP joint for the fourth finger is very high.

Eighth position

The hand model was as second position and lifted upward by around 50 mm. The absolute errors between the CMM and the CVS are shown in Table 5-25.

Table 5-25: Absolute errors between the CMM and the CVS (eighth position)

	MCP joint(°)	PIP joint(°)	DIP joint(°)
First finger	0.40	1.26	1.34
Second finger	0.18	0.21	0.19
Third finger	0.15	0.52	0.68
Fourth finger	0.20	1.54	3.62
Mean	0.86		
STDEV	0.96		

From the eight data tables above which obtained by placing the hand model at different locations within the control volume, it is obvious that most of the finger joint angles measured by the CVS have an error of less than one degree compared with the CMM machine. However, there are some tables shown one or two errors, out of 12 readings, which are more than one degree. These errors should not be existed as placing the hand model at different locations within the calibration piece should not significantly affect the result. Given that the calibration points were evenly distributed and all the tested were used the same cameras parameters. So the electronic devices could be behind these high errors.

B) Second category: new finger angles were set and the angles varied between 2° and 8°

The values of pre-set joint angles measured by the CMM are shown in Table 5-26.

Table 5-26: Joint angles measured by CMM

	MCP joint(°)	PIP joint(°)	DIP joint(°)
First finger	2.39	2.91	5.59
second finger	3.29	0.72	7.36
Third finger	3.22	2.36	1.49
Fourth finger	4.78	1.11	0.44

In the second category the hand model was placed at different positions similar to category A. The mean error from category B is higher than category A. And after investigation it was found that for all hand locations the largest error always emerged on two joints only; the DIP joint for the third finger gave an error of around 4° and the DIP joint for the fourth finger gave an error of around 2°. The rest of the joints produced a very small error comparable with the error for the category (A). The reason for this large error could be human error from the CMM measurements, and because the DIP joint angle is produced from the intersection of the two smallest vectors in the finger. The first vector is formed from a pair of measurement points for the Distal phalanx bone and the second vector from a pair of measurement points for the Middle phalanx bone. It is suggested that the error for the joint angle (from the CMM) produced from short vectors would be expected to be larger than the one produced from long vectors, see Figure 5-21. Another cause of this error might be mismatching between the points used to calculate the vector in space by the CMM machine and the points from the vision system. In the vision system the centroid of the points is used to find the vectors. If the measuring probe of the CMM does not coincide with the position of the vision system, an additional error between the two readings will emerge (see Figure 5-22).

The height of the UV gel (Z) shown in Figure 5-22 was not considered because all the measurement process were carried out after the UV gel had dried, and in this case the Z height is insignificant.

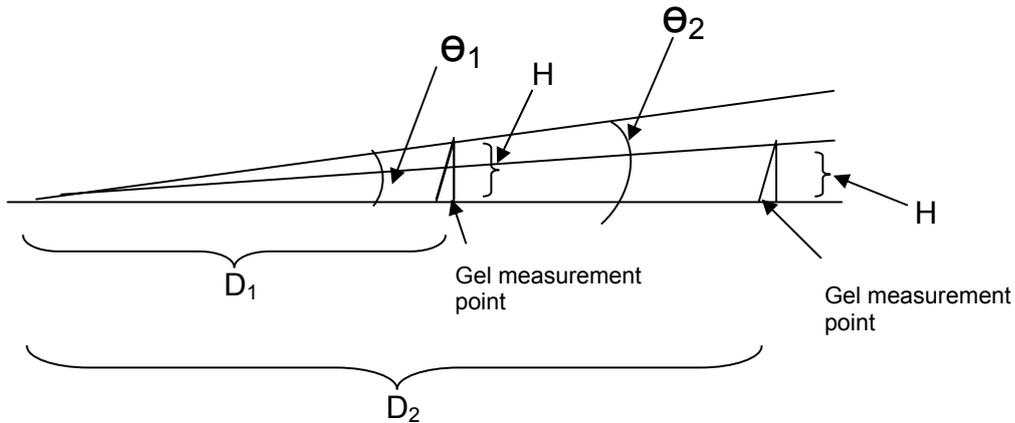


Figure 5-21: Relationship between error in angle and the distance $\theta_2 > \theta_1$ despite the height of the gel measurement points (H) is the same, this is because of $D_2 > D_1$

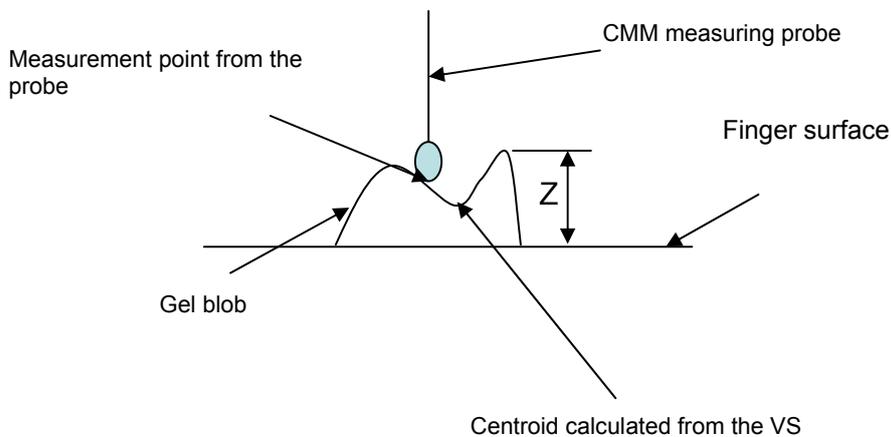


Figure 5-22: Mismatching between the two measuring devices (CMM and VS)

To test the repeatability of the system, i.e. to see the degree of consistency between successive measures of the same joints while, all parameters are fixed. The finger joints angles were vary from less than one degree to more than 39 degrees in order to make sure the CVS works fine in the case of small and large angles as well. The fingers joints were measured 4 times in sequence by using the computer vision system, the results are shown in Table 5-27. From this table,

there are difference between some reading despite all of them are used the same cameras parameters and the hand model was placed at fixed position, but only difference between these tests is the captured images, so the reason for these differences are the image coordinates for the measurement points which is not related to the calibration technique but related to the electronic devices.

From this table the difference is ranged from 0.27 degree to 1.11 degree and can appear at any finger joint randomly. This error and in addition to the error from the calibration technique leads to increase the error to more than one degree in one or two joints as seen in Table 5-24.

Table 5-27: Results from examine the CVS repeatability

		Finger joints angles in degrees						
	finger joints	1st test	2nd test	3rd test	4th test	Min	Max	Max-Min
1 st finger	MCP	17.31	17.43	16.85	17.63	16.85	17.63	0.78
2 nd finger	MCP	15.52	15.69	15.37	15.9	15.37	15.9	0.53
3 rd finger	MCP	22.45	22.51	22.28	22.91	22.28	22.91	0.63
4 th finger	MCP	29.43	29.15	29.13	29.58	29.13	29.58	0.45
1 st finger	PIP	2.41	2.1	2.44	3.15	2.1	3.15	1.05
2 nd finger	PIP	1.75	1.41	1.72	0.64	0.64	1.75	1.11
3 rd finger	PIP	19	18.8	19.11	19.63	18.8	19.63	0.83
4 th finger	PIP	39.7	39.86	39.02	39.38	39.02	39.86	0.84
1 st finger	DIP	5.21	5.98	5.26	4.92	4.92	5.98	1.06
2 nd finger	DIP	21.52	21.5	21.65	20.69	20.69	21.65	0.96
3 rd finger	DIP	5.69	5.73	5.46	5.56	5.46	5.73	0.27
4 th finger	DIP	1.37	1.33	1.83	1.45	1.33	1.83	0.5

5.8 Measuring fingers joints based on the Tsai technique

All the results above were carried out using the DLT technique. The Tsai technique was then used as another calibration technique to estimate the camera parameters. The Tsai technique was used because the literature says it should be superior to the DLT and confirmation of this was required.

Having the camera parameters, the hand model was placed at five different positions within the control volume and the finger joints angles were measured using the computer vision system, and then compared with the angle values obtained from the CMM. The results are shown in Table 5-28. The mean error, standard deviation and the root mean square error for all the results is less than one degree.

Also from this table there are some results have errors less than or around 1 degree, as seen in the first and fifth tests, while the third test contains two errors, out of twelve readings, which are slightly higher than one degree. For the second and fourth test, each contains two errors ranged between 1.52 degree and 1.87 degree. The result from using Tsai camera calibration is similar to the ones obtained from using DLT technique as shown in Table 5-22 to Table 5-25.

Table 5-28: Absolute error between CMM and TSAI

	Finger joints	Absolute error between CMM and TSAI				
		1 st test	2 nd test	3 rd test	4 th test	5 th test
1 st finger	MCP	0.05	0.09	0.27	0.20	0.23
2 nd finger	MCP	0.71	0.85	1.34	0.94	1.25
3 rd finger	MCP	0.06	0.17	0.42	0.21	0.07
4 th finger	MCP	0.84	0.66	0.27	0.02	0.36
1 st finger	PIP	0.13	0.86	0.50	0.88	0.02
2 nd finger	PIP	1.18	1.67	0.54	1.30	0.89
3 rd finger	PIP	0.16	0.10	1.21	1.55	0.93
4 th finger	PIP	0.96	1.87	1.42	1.52	0.93
1 st finger	DIP	0.47	0.01	0.41	0.95	0.50
2 nd finger	DIP	0.20	0.60	0.35	0.92	0.38
3 rd finger	DIP	0.19	0.56	1.25	1.34	0.57
4 th finger	DIP	0.50	0.20	0.62	0.41	0.64
Mean		0.64				
STDEV		0.48				
RMSE		0.80				

To find which camera calibration method gives better result, i.e. DLT or Tsai, the finger joint angles were measured by using the DLT and Tsai under the same

conditions. To do so, firstly the cameras were calibrated using DLT technique as explained in section 5.4. Then the calibration piece was used to calibrate the cameras by Tsai method as explained in detail in Appendix A.1.

At this stage we have the camera parameter obtained by both techniques stored in the vision system. Then the hand model was placed inside the calibrated volume to capture the images of the hand.

In order to automatically match all the measurement points of the hand that captured by the cameras, the technique explained in section 4.6.4 was implemented.

Then the 3D world coordinates of the measurement points were reconstructed using DLT technique as explained in section 5.5. And to calculate the fingers joints angles the formula shown in section 4-7 was used.

With the same images which used for DLT, the 3D world coordinates of the measurement points were reconstructed using Tsai technique as explained in Appendix A.1.2, then the fingers joints angles were calculated as above.

Table 5-29 and Figure 5-23 compare the results from using both DLT and Tsai techniques.

From Table 5-29 and Figure 5-23 it can be concluded that the Tsai method gave a better result than the DLT as expected. Also from this table all the results obtained by Tsai method is less than one degree except one reading which is slightly higher than one degree (1.25). Comparing the error at DIP joints for the third finger, it is clear that error from DLT is much higher than Tsai, but this does not necessary mean the high error obtained by the DLT is correct, as this could be from the electronic devices rather than the calibration technique.

Despite the fact that the Tsai method is more computationally expensive and complex as it needs to solve non-linear system of equations. However, recent advances in computer software and hardware makes this complexity much easier to solve at relatively low cost. Nevertheless, in the case where a high degree of accuracy is not required, then it is much better to calibrate the cameras by the DLT to avoid the mathematical complexity.

Table 5-29: Comparison between DLT and Tsai

	Finger joints	DLT ,CMM (°)	TSAI ,CMM (°)
1 st finger	MCP	0.40	0.23
2 nd finger	MCP	1.07	1.25
3 rd finger	MCP	0.07	0.07
4 th finger	MCP	0.17	0.36
1 st finger	PIP	0.14	0.02
2 nd finger	PIP	0.86	0.89
3 rd finger	PIP	0.91	0.93
4 th finger	PIP	1.76	0.93
1 st finger	DIP	0.16	0.50
2 nd finger	DIP	1.10	0.38
3 rd finger	DIP	1.99	0.57
4 th finger	DIP	0.31	0.64
Mean		0.75	0.56
STDEV		0.65	0.38
MSE		0.94	0.45
RMSE		0.97	0.67

MSE = mean squared error

RMSE = root mean squared error

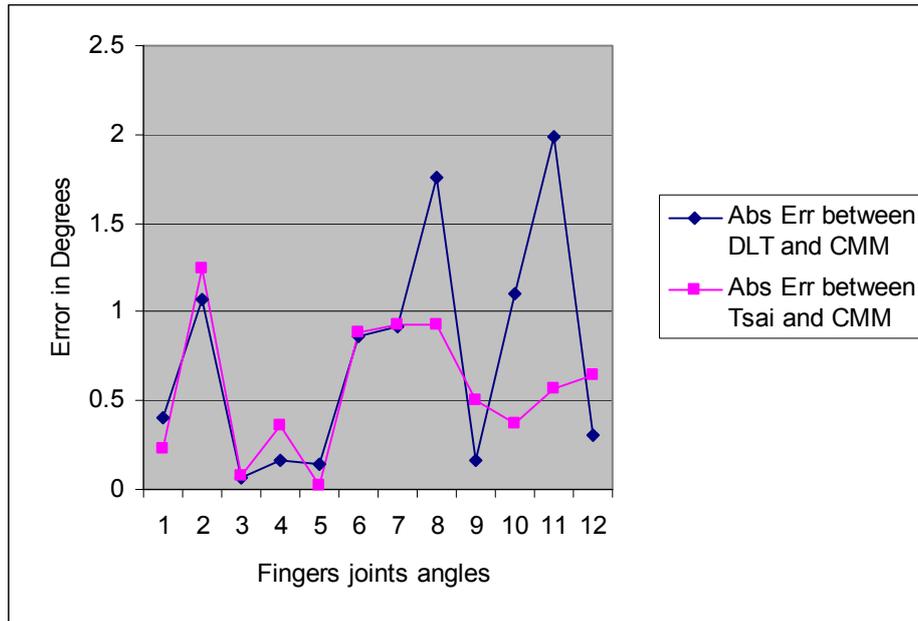


Figure 5-23: Comparison between DLT and Tsai based on CMM

5.9 Measurements of a real hand

To measure the finger joint angles of a real hand, 8 gel points were placed on a real hand and two images were captured by two cameras, while the hand holding a tennis ball for two reasons. The first is to bend the fingers and the second to ensure that the fingers did not move between measurements by CVS and protractor.

The images of the real hand that obtained from the vision system are shown in Figure 5-24 and Figure 5-25. These images were processed and the MCP joint angles for the index and middle fingers were evaluated.

Then the hand moved from the vision system, and while it is holding the tennis ball, the same angles were measured by the protractor shown in Figure 5-26. The results from the two methods are illustrated in Table 5-30. To ensure that the finger did not move between measurements by CVS and

protractor, the hand was measured holding a tennis ball during all measurement processes.

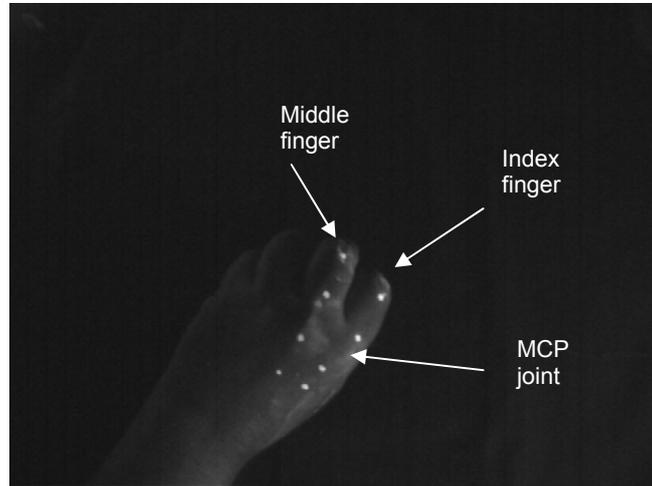


Figure 5-24: Real hand grey level image with UV gel captured by the left camera

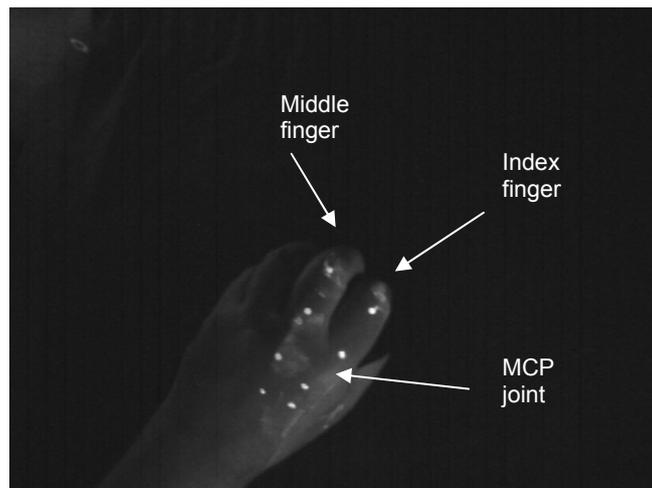


Figure 5-25: Real hand grey level image with UV gel captured by the right camera

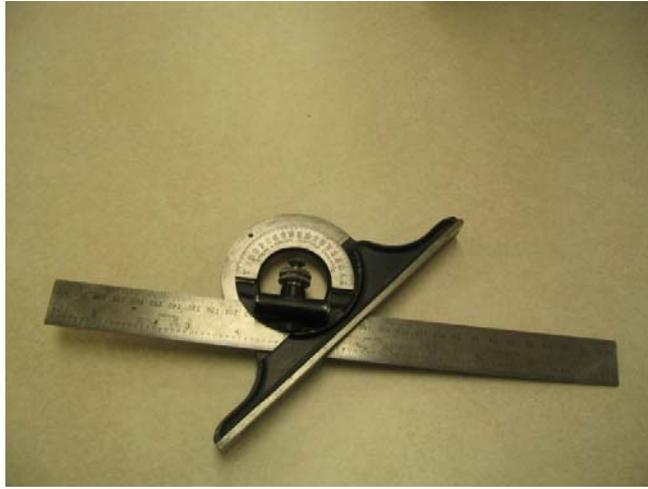


Figure 5-26: Protractor which used to measure the MCP joints

Table 5-30: Results from using CVS and a protractor

	Results from CVS in degrees	Results from protractor in degrees
MCP joint for the Middle finger	43.5	48
MCP joint for the Index finger	52	57

From Table 5-30 above it is obvious that the difference between the reading obtained by the CVS and the protractor was around 5 degrees. This result is similar to the one found in literature where the measurement process carried out by the traditional goniometry (Bainbridge, 2000 and Williams, *et al.*, 2000). From this test it is clear that the protractor is not a reliable one for measuring finger angles; also in general the error from using a protractor or any goniometry can be up to ± 5 degrees.

5.10 Conclusion

The computer vision system in this study has proved to be a reliable and accurate tool for evaluation of the range of motion of the fingers. The results from using a steel hand model showed that the proposed system is reliable and produced encouraged results.

Firstly, the FOV of the cameras is calculated in order to guarantee that it covers the calibration and test piece. The results show the location of the cameras under the available working volume and the given focal length of the lens is sufficient to cover the calibration and test piece. The FOV of the cameras can be increased if the size of the cameras and cameras holders are replaced by smaller ones.

The 12 hand finger joints were measured by using the CMM, then the hand model was placed at several locations inside the control volume and the finger joints angles were measured by the CVS and the two results were compared. The comparison showed that the error between the CMM and the CVS was very small for most of the joints (less than one degree). However, some results show one or two errors out of twelve are more than one degree. After investigation it was found that these large errors were not related to the CVS but as a result of human error while measuring the angles by the CMM or from the electronic devices that compose the CVS, the later can be avoided by replacing the analogue cameras by digital ones.

Two calibration approaches were used to calibrate the cameras, DLT and Tsai. A comparison between the DLT and the Tsai techniques showed that the latter produced better results when the finger joints angles were evaluated.

The experimental work showed that moving the calibration piece within the allowable FOV had an insignificant effect on the accuracy of the system.

A real hand was measured and it was illustrated that using the UV gel as measurement points gave good contrast with the human skin. However, the error between the computer vision system and the protractor used to measure the finger joint angles was quite high (around 5 degrees). This error is because of the unreliability of the protractor for assessing the finger joints.

6 Measuring the full range of finger motion

So far, this work has involved assessment of fingers joints by placing cameras above the hand. This position of the cameras does not allow images of the measurement points at all flexion positions to be captured, see Figure 6-1. To cover all the hand at different values of flexion and extension gestures, the four ideas below have been considered.

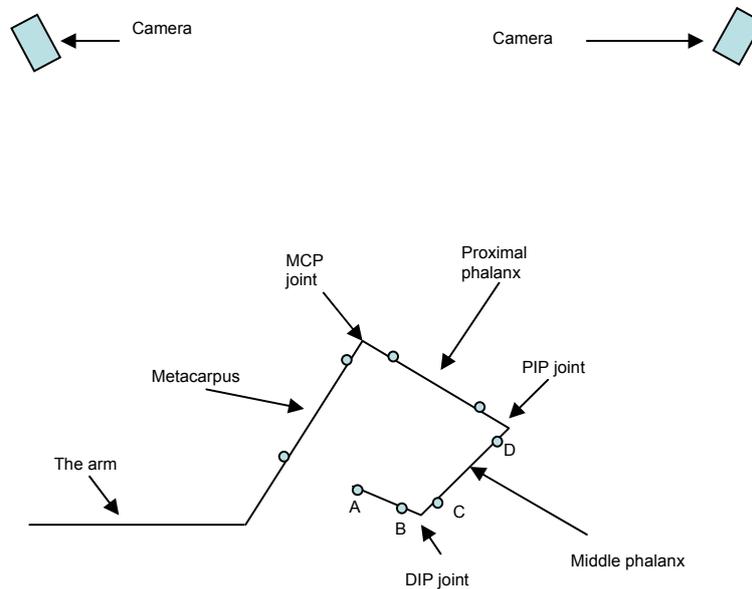


Figure 6-1: Finger at maximum flexion position, at this position the top cameras cannot see the UV gel measurement points A, B, C, and D

1. Placing cameras underneath the hand as well as above it, so images for the hand can be captured while they are in the flexed position. This option was discarded because it would lead to a significant increase in the size of the working volume (to double the existing size) and would add to the cost. Keeping the size of the working volume as small as possible is a major factor in this study, so that the vision system occupies a small volume in the clinic.

2. The second option was to place the hand so that its back is facing the cameras, and then to capture the measurement points which are available in view. The points which can't be captured from this position can be imaged by turning the hand through 180 degrees. However, using this method will increase the assessment time, which is not desirable and may lead to more patient discomfort.
3. The third idea was to place a concave mirror underneath the hand, so that the cameras placed above it can see all the measurement points on the fingers even if the hand is at the fully flexed position. From the properties of concave mirrors, if the measurement points are located between the centre of curvature and the focal point, the image will be located beyond the centre of curvature. In this case the virtual image can be placed inside the control volume. However, because of the geometrical distortion from using this type of mirror, this idea was discounted.
4. The fourth idea, which was adopted in this work, was to use a plane mirror or a 50:50 Beam Splitter (BS). If the cameras cannot see the measurement points directly, then they will be seen through the BS or mirror. The BS or mirror is placed underneath the hand and parallel to the X, Y world plane, so the location of the virtual points will differ from their real ones in Z coordinates only. In either case (BS or plane mirror), it is important to be able to distinguish between real and virtual measurement points.

Beam Splitter

In the case of the beam splitter (BS), the discrimination between the virtual measurement points and the real ones can be achieved based on their brightness values (grey level). The virtual points that are reflected from the BS will have brightness values 50% or so less than the real measurement points. After

calculating the Z world coordinates for the virtual points by using either the DLT or Tsai method, the actual Z coordinates for the virtual points can be calculated as shown below.

$$Z_{actual} = 2 \times (Z_{BS} - Z_{virtual}) + Z_{virtual} \quad (6-1)$$

$$= 2 \times Z_{BS} - Z_{virtual} \quad (6-2)$$

Where:

Z_{actual} - The actual Z coordinate for the real point, see Figure 6-2.

Z_{BS} - The Z coordinate of the beam splitter (known from the design).

$Z_{virtual}$ - The Z coordinate of the virtual point calculated by the CVS.

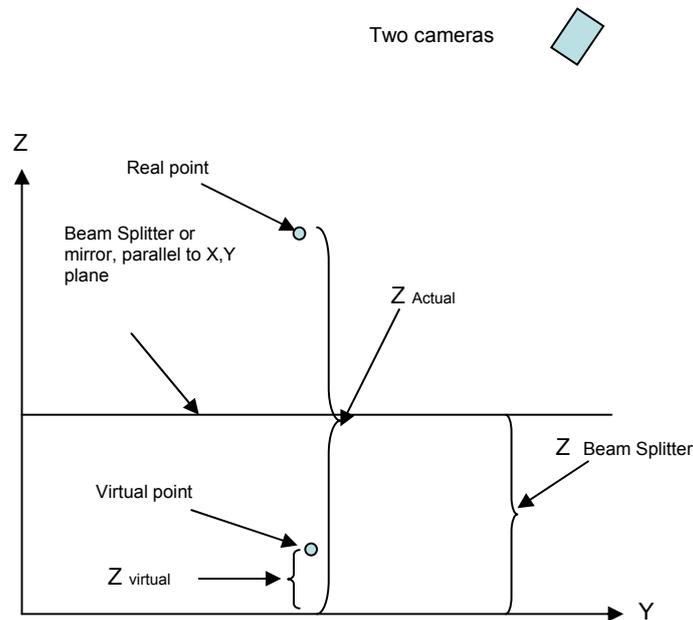


Figure 6-2: Location of real and virtual points

Plane mirror

In the case of using a plane mirror, both real and virtual points have equal brightness. So after reconstruction of the 3D world coordinates, the virtual points

have Z coordinates less than the Z coordinate for the plane of the mirror. If any two points have similar X,Y coordinates, and one of them is virtual and the other is real, this means that the cameras see the virtual and the real point at the same time. In this case, the real one should be chosen and the virtual one eliminated.

6.1 Extension and flexion range of the fingers

Table 6-1 shows the typical expected finger motion in degrees from four literature sources. From this table, the maximum flexion angle of the DIP joints according to the AAOS (American Association of Orthopaedic Surgery) is 90°, which cannot be achieved unless a force is applied to the Distal phalanx bone toward the flexion direction of the DIP joint. Since this is not the case when measuring the finger joint angles in this study, this source has been discarded.

Table 6-1: Maximum finger motion in degrees from selected sources

Joint	Motion	AAOS	AMA	Hume, M. 1990	Mallon, W. 1991
MCP	Flexion	90	90	100	95
	Extension	45	20	0	20
PIP	Flexion	100	100	105	105
	Extension	0	0	0	7
DIP	Flexion	90	70	85	68
	Extension	0	0	0	8

AAOS = American Association of orthopaedic surgery.

AMA = American medical Association.

Similarly, the maximum extension values of the PIP and DIP joints measured by Mallon (1991) were 7° and 8° respectively. These angles also can't be achieved unless a force is applied toward the extension direction. The maximum extension angle for the MCP joint obtained by Hume (1990) is 0°. By observation of typical hand motion, it is obvious that this joint angle should be > 0°.

Based on the discussion above, the maximum extension and flexion movement of the fingers used by the AMA (American medical Association) has been chosen to be utilized in the new design process. Figure 6-3 shows a sketch of the maximum finger flexion angles as obtained by the AMA.

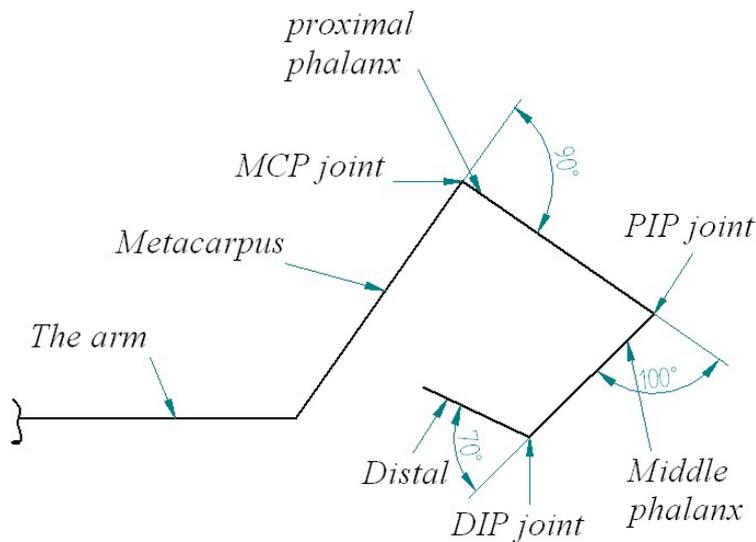


Figure 6-3: Maximum finger flexion motion according to AMM (side view)

6.2 Design of the new system

In order that the measurement points can be captured at all finger positions, the new system will use four cameras, two of them to capture the measurement points of the Middle Phalanx and the Distal, and the other two to deal with the Metacarpus and Proximal Phalanx or all the finger joints, depending on the position of the hand.

The BS is made from a coated glass plate of thickness 3 mm and with dimensions 186×76 mm. It must be placed so that it is parallel to the X, Y plane of the world coordinate system so that it does not affect the X,Y location of the

measurement points and so that the points seen through the BS will be shifted only in the Z direction.

Figure 6-4 shows the Beam Splitter placed on a specially designed support, at a height of 32 mm in the Z direction in order to insure that the virtual points are located inside the control volume of the calibration piece. Any error that might appear from the extrapolation can thus be avoided, see Figure 6-5.

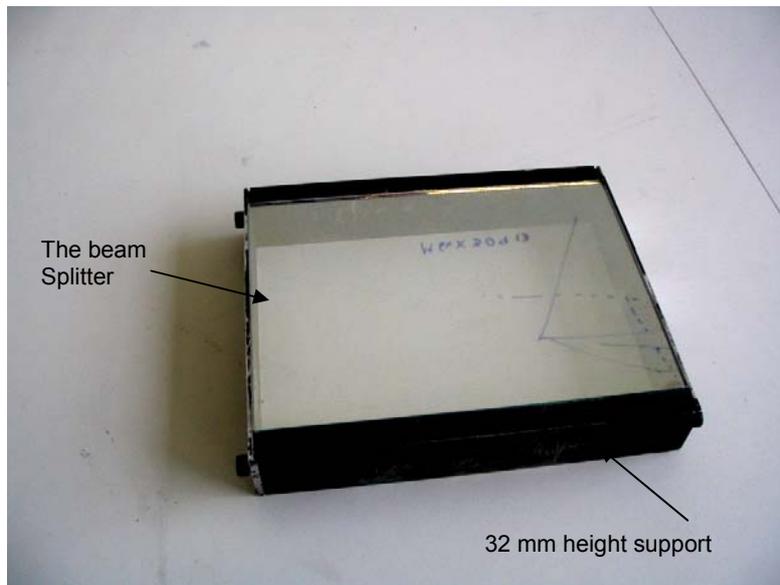


Figure 6-4: The Beam Splitter used in this work

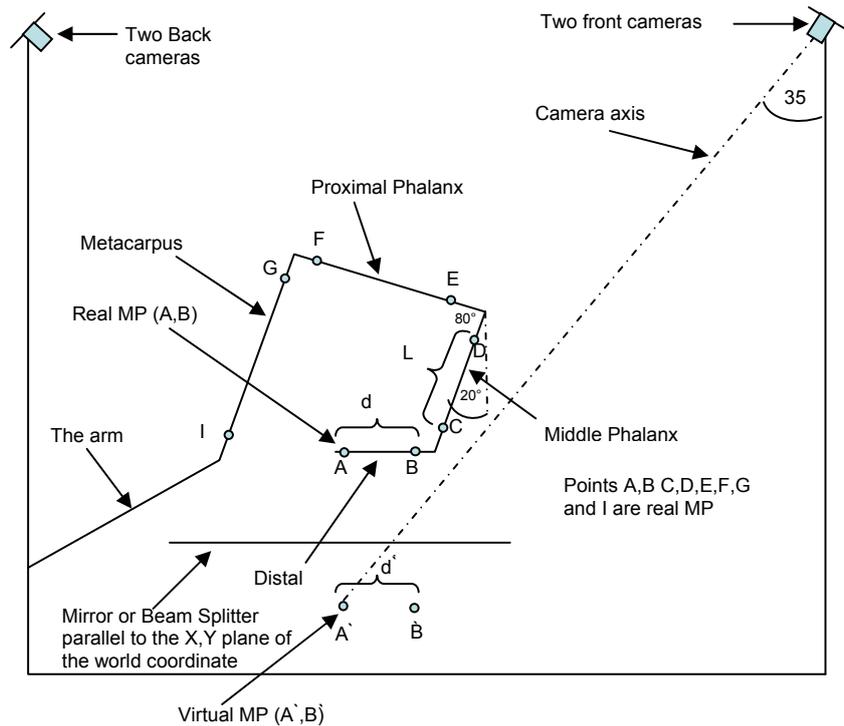


Figure 6-5: Illustrates the position of the hand with respect to the cameras

Movement of the DIP joint (see Figure 6-3) leads naturally to movement of the PIP joint and vice versa. The linear relationship between the two joints can be expressed by equation (6-3), Lee, *et al.*, (1993).

$$\theta_{DIP} = \frac{2}{3} \theta_{PIP} \quad (6-3)$$

Applying equation (6-3) to this work will help to determine the number of measurement points at different positions of the DIP joint and the PIP joint. However this formula can't be applied in the case of injured fingers, where there is a possibility that the movement of the DIP and PIP joints may not follow the natural movement of the fingers as indicated by equation (6-3).

In order to reduce the number of virtual points that are involved in the measurement process, the hand is placed at a position such that the virtual measurement points for the distal bone reflected by the Beam Splitter are used

and not any other virtual measurement points. Before carrying out the measurement process we need to determine the following:

1. The distance between the two measurement points of the Middle Phalanx (points C and D in Figure 6-5 in terms of pixels, when the hand is at maximum flexion. In this case only the real measurement points (MPs) of the Middle Phalanx bone are used to measure the angle. The points C and D in the image should be spaced apart from each other at a distance where they can be recognized.
2. As the finger moves towards the extension position, the front cameras will see the real MPs of the distal bone. We need to determine the distance between the two measurement points of the distal bone in terms of pixels at the position when the front cameras start seeing the real measurement points of the distal bone.

Given the above we can conclude that:

- (i)- The front cameras axes and Middle phalanx make angles of 35° and 20° with the vertical axis respectively.
- (ii)- The distance between the two MPs of the Middle Phalanx must be 15 mm (± 3 mm) (This distance is for the measurement points of the Middle Phalanx bone of the little finger and is larger for the Middle Phalanx of the rest of the fingers).
- (iii)- According to the system configuration, $1\text{mm} \equiv 2.3$ pixels.

To show how the measurement points for the Middle phalanx bone are projected on the camera sensor, the measurement points for the Middle phalanx bone and the camera sensor shown in Figure 6-5 can be transferred and redrawn as illustrated in Figure 6-6.

Figure 6-6 illustrates the distance between the measurement points for the Middle phalanx (L), the projection of L on the camera sensor in terms of pixels (X), and angle between the vector L and the camera sensor (θ).

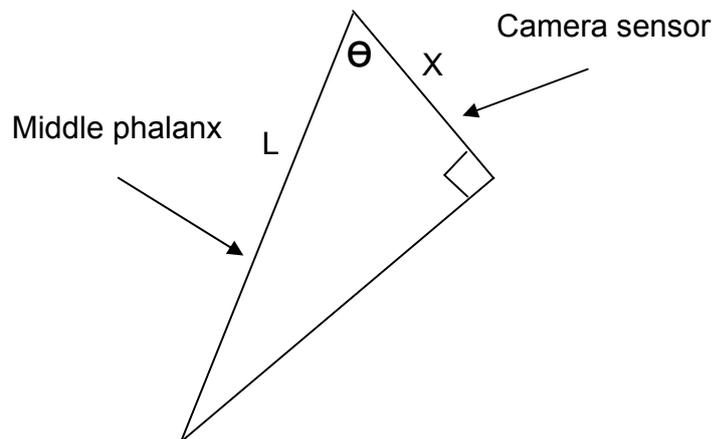


Figure 6-6: The relation between the MP for the Middle phalanx (L) and its projection on the camera sensor X (see Figure 6-5 for more details)

The value of θ from Figure 6-5 above is:

$$\theta = 20^\circ + 55^\circ = 75^\circ \quad (6-4)$$

Then the distance X can be calculated from the following formula:

$$X = \cos(75) \times 15 \times 2.3 = 9 \text{ Pixels} \quad (6-5)$$

From the equation(6-5) above, the distance between the MPs of the Middle Phalanx is 9 pixels, when the hand is at maximum extension and this distance will be increased as the hand moves towards the flexion position.

As mentioned above, when the hand is at maximum flexion, the virtual measurement points of the distal phalanx will be seen by the front cameras only. Also the measurement points for the distal phalanx and the camera sensor shown in Figure 6-5 can be transferred and redrawn as illustrated in Figure 6-7.

Figure 6-7 shows the relationship between the virtual measurement points of the distal phalanx and the camera sensor. From this figure we have:

d^* is the distance between the virtual MPs of the Distal phalanx for the little finger, 10mm.

C is the projection of d^* on the camera sensor C .

α is the angle between the vector D and the camera sensor (35 degrees) and the scaling factor is 2.3 pixels/mm as before. Then the length of C in terms of pixels can be calculated as:

$$C = \cos(\alpha) \times d^* \times 2.33 = 16.4 \text{ pixels} \quad (6-6)$$

Distance C increases as the hand moves towards its extension position and it reaches the maximum value when d^* is parallel to the camera sensor (i.e. when the Distal phalanx is parallel to the axis of the camera). When the Distal phalanx angle with respect to the vertical axis is equal to or less than the angle of the camera axis, then the camera starts “seeing” the real measurement points of the Distal phalanx as well as virtual ones, until the Distal phalanx becomes parallel to the vertical axis, when at this stage the cameras will “see” two real measurement points and only one virtual point. When the Distal phalanx is extended further, only the real points will be seen by the front cameras.

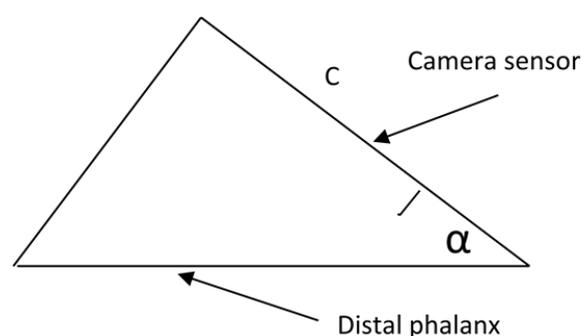


Figure 6-7: The relationship between the virtual M.P of the Distal phalanx(d^*), and their projection on the camera sensor (C) (see Figure 6-5 for more detail)

6.3 How to select the relative measurement points for the front cameras

When the fingers are at maximum flexion, the two front cameras (see Figure 6-5) will “see” 8 MPs, 2 for the Proximal phalanx (real), and 4 for the Middle phalanx (2 real and 2 virtual). The last 2 measurement points for the Distal phalanx are virtual. Figure 6-8 illustrates a sketch of the finger when it is at maximum flexion.

From Figure 6-8 if all 4 virtual points are seen (points E, F, G and I); this means that the finger is at maximum flexion. The 2 virtual measurement points which have maximum X coordinates (points G and I) can thus be removed because these points belong to the middle Phalanx, and the front cameras can see the real measurement points of this bone (points C and D). Also the two measurement points which have minimum X coordinates (points A and B) can be removed, because the back cameras will deal with them instead.

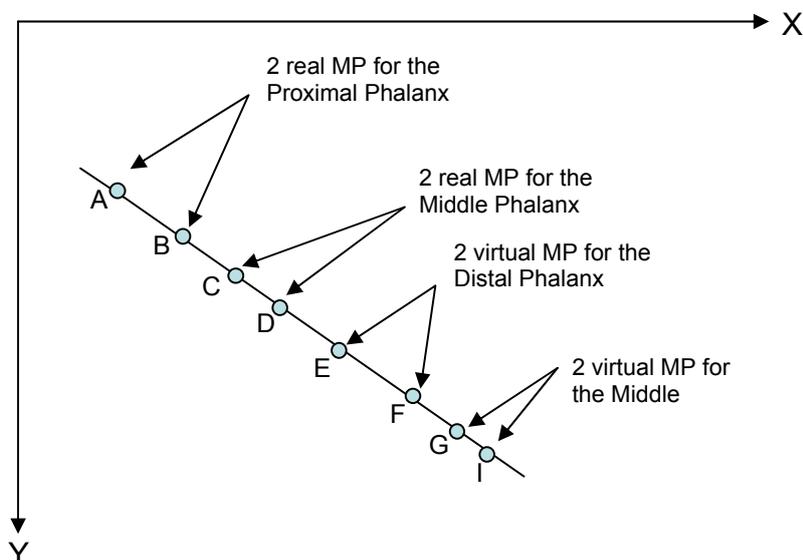


Figure 6-8: Finger at maximum flexion

As the finger moves towards the extension position, the number of measurement points will be reduced from 5 to 0. If the number of measurement points seen by

the front cameras is less than 2, then the points captured by the back cameras will be used for the calculation rather than those from the front cameras.

The two back cameras will process the real measurement points, either 4 or 8 depending on the number of points seen by the front cameras.

The correspondence matter has already been considered in Chapter five to obtain matching between the measurements points of the fingers, and the same technique was applied here. Thus, the 3D reconstruction of the points can now be achieved.

The hand model was placed inside the control volume and the fingers were bent to approximately the maximum flexion position so that some of the measurement points were reflected from the Beam Splitter and the camera could only see the virtual images of these measurement points.

The finger joint angles were then measured at different hand positions using the new system. Results from the measurement process carried out are shown in Table 6-2 and Figure 6-9 illustrates some of the results obtained by using the DLT and Tsai techniques.

Table 6-2: The absolute error between DLT and CMM, Tsai and CMM

Figure joints	CMM (degrees)	DLT (degrees)	Tsai (degrees)	ABS Err DLT & CMM (degrees)	ABS Err TSAI & CMM (degrees)
1 st finger MCP	32.16	33.42	33.27	1.25	1.11
2 nd finger MCP	30.45	32.13	31.92	1.67	1.46
3 rd finger MCP	29.76	31.15	31.04	1.40	1.29
4 th finger MCP	28.18	30.65	30.41	2.47	2.23
1 st finger PIP	35.49	35.07	35.04	0.43	0.46
2 nd finger PIP	41.19	39.61	39.82	1.59	1.37
3 rd finger PIP	49.49	49.78	50.21	0.29	0.72
4 th finger PIP	45.82	43.79	44.78	2.03	1.04
1 st finger DIP	85.22	83.86	83.66	1.36	1.56
2 nd finger DIP	85.30	88.43	87.96	3.13	2.66
3 rd finger DIP	73.63	74.18	73.48	0.55	0.16
4 th finger DIP	84.85	85.81	84.50	0.96	0.35
Mean				1.43	1.20
Max				3.13	2.66
Min				0.29	0.16
STDEV				0.84	0.74

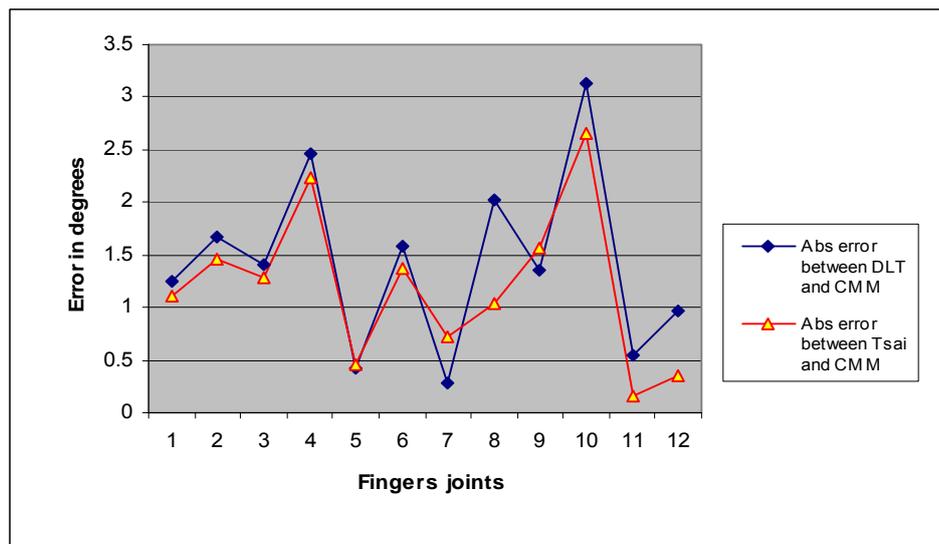


Figure 6-9: Absolute error between DLT and CMM, Tsai and CMM

Joints numbers 1 to 4 are MCP joints and each joint produced from 4 real points, Joints numbers 5 to 8 are PIP joints and each joint produced from 4 real points, Joints numbers 9 to 12 are DIP joints and each joint produced from two real and two virtual.

Upon inspection of the data, it was found that there were large errors which had not existed before. To find what had caused these errors, the hand model was placed at a fixed location and the finger joint angles were measured six times,

using the same camera parameters for all the 6 tests. It was found that there was significant error between the readings despite all the parameters being the same. Ideally there should be no significant error between the 6 tests. The results in Figure 6-10 below shows there are significant errors obtained from the finger joint angles 6, 8 and 10.

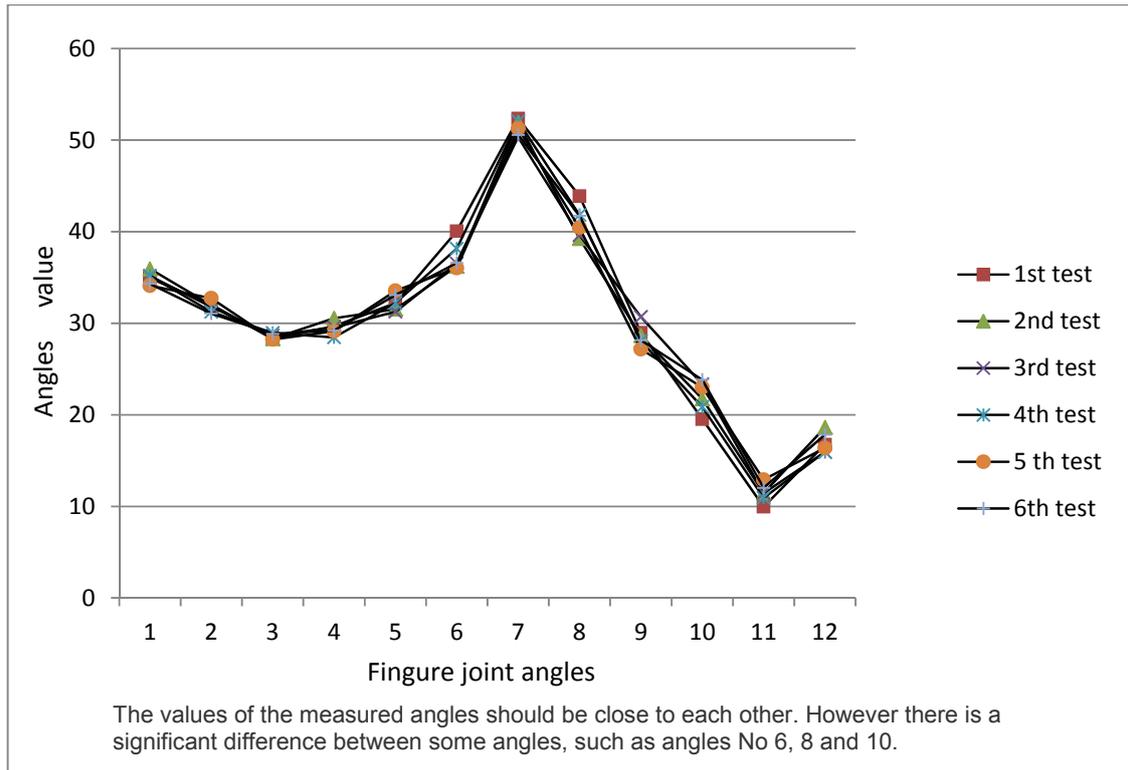


Figure 6-10: Errors between six measurements made under the same parameters

Now let us investigate further, in order to understand how these errors may have arisen.

Table 6-3 shows two values of a finger joint that was captured twice using the same camera parameters and under the same conditions. The comparison shows 3.2 degrees difference between the two readings. Having that all the parameters are the same for the two tests except each test has its own images,

so the difference in joint angles measurement should be from the image coordinates.

Table 6-3: Result of measuring the same joint twice under the same conditions

	First reading	Second reading	Difference between readings
Finger joint angle	84.4	87.6	3.2

Table 6-4 shows the X, Y and Z coordinates that were used to calculate the finger joint angles show in Table 6-3.

Table 6-4: 3D world coordinates for the points that used to calculate the angle shown in Table 6-3

	X1 mm	Y1mm	Z1 mm	X2 mm	Y2 mm	Z2 mm	X2-X1	Y2-Y1	Z2-Z1
1 st point	-258.78	291.28	107.69	-258.78	291.23	107.85	0.00	-0.05	0.15
2 nd point	-259.48	283.35	100.78	-259.48	283.34	100.83	0.00	-0.01	0.05
3 rd point	-258.73	280.21	81.89	-258.83	279.89	81.19	-0.10	-0.32	-0.70
4 th point	-259.54	288.08	70.95	-259.52	288.18	71.12	+0.03	+0.10	+0.17

The table above (the last column) illustrates that there is a difference in Z coordinate between the first and second test. For instance, in the second test, the 3rd measurement point shifted downwards by 0.7 mm and 4th point shifted upward by 0.17 mm (See Figure 6-11). Also the 1st and 2nd points moved upwards by 0.15 mm and 0.05 mm. According to the finger position with respect to the 3D world coordinates, see Figure 6-11, the error in Z direction has the maximum influence on the change in the angle between the two vectors (Vectors A and B in Figure 6-11).

From Figure 6-11, it is clear that the shifts in Z coordinates in the second measurement process led to increase the angle between vector A and vector B. This explains the reason behind the difference between the same joint angle that measured twice under the same conditions (see Table 6-3).

Having that the two measurements were done under the same conditions, and the only difference was the image coordinates, so the error in the angle between the two measurements was produced from the shift in image pixels.

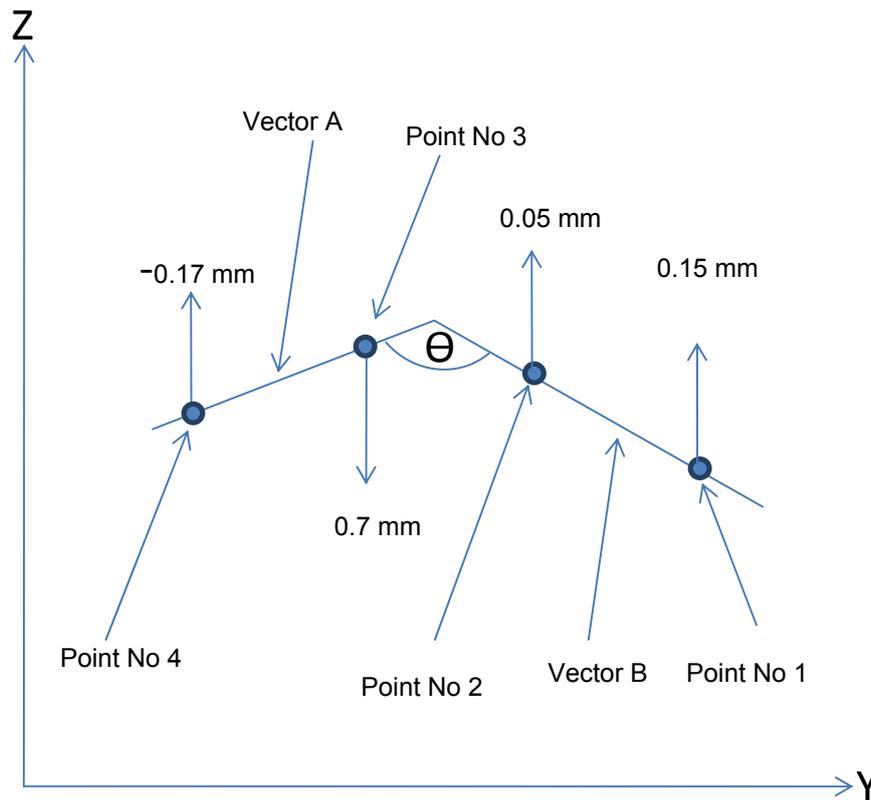


Figure 6-11. The error between the two measurements led to an increase in the Value of Θ for the second measurement.

Further investigation into the x, y pixel location of the measurement points used to calculate the finger joint angle, shown in see Table 6-3 above, demonstrated that there was up to 0.5 pixel shift between the set of images for the first and second measurement processes. Also, it was found that if the pixel shift is not significant (around 0.1 pixel), then there is no significant difference of the finger joint angles between the readings.

Also, by comparing results shown in Table 6-2 and Figure 6-10 with Table 5-21, Table 5-24 and Table 5-29 it can be concluded that the system have developed more errors which are not existed before.

From the data sheet of the frame-grabber, the shift of the pixel should not be more than 0.15 pixels. After contacting the frame-gabber manufacturer (Data Translation) it was concluded that the frame-grabber has a fault which was not existed during the measurement process shown in chapter 5.

6.4 Conclusion

To assess the finger angle joints using the computer vision system, the system should cover the hand at all flexion and extension motions. To achieve that several ideas have been discussed, each one has advantages and drawbacks. The one which was applied in this prototype system is based on the property of reflection of objects from a plane mirror or a Beam Splitter. In case of a plane mirror the brightness of the real and the virtual object is the same, however the virtual objects reflected from the Beam Splitter have brightness less than real ones, depending on the coating property of the Beam Splitter. The disadvantage of using a Beam Splitter is that the illumination needs more consideration in order to increase the efficiency of the Beam Splitter.

The Beam Splitter was added to the computer vision system, and all the design theory behind placing the Beam Splitter with respect to the hand was discussed so that the system can cover the hand at all extension and flexion motions.

The new system was then used to evaluate the hand. Several measurement processes were carried out and the system appeared to operate satisfactorily; however its accuracy was slightly reduced compared with the results obtained before introducing the Beam Splitter (as described in Chapter five).

The angles with large error were checked further and it was found that there was a large shift in pixel location which could be associated with the measured angular error. The pixel shifting was 0.50 pixel (some measurements produced a pixel shift around one pixel), however according to the data sheet for the frame grabber used in this work, the pixel shifting should be limited to between ± 0.15 pixels. As the pixel coordinates are used to reconstruct the 3D coordinates of the measurement points, they affect the 3D location of the measurement points and as a result the angles between two vectors in space will be affected (finger joint angles). To avoid such problems in the future, digital cameras should be used instead of the analogue ones or a better frame-grabber should be used.

7 Discussion, conclusion and further work

Accurate measurement of finger joints is an important procedure in hand therapy in order to examine the function of the hand after treatment or surgery. The evaluation of finger ROM involves assessment of the metacarpophalangeal (MCP), proximal interphalangeal (PIP) and distal interphalangeal joints (DIP).

There are several existing tools for measuring the angles of the finger joints that are quite cheap and easy to use, such as visual estimation, wire tracing, and traditional goniometry. However the time demanded for the assessor and the patient, and the lack of accuracy are the major drawbacks of these devices. Also they only allow assessment of one joint at a time. Other devices are more expensive and sophisticated; for instance the goniometric glove and electrogoniometer. However these devices cannot be used if the finger is injured.

To replace the existing devices for measuring finger motion, a new measurement system based on computer vision technology was designed and built. The proposed system is a non-contact measurement device and has many advantages over the existing measurement devices. In terms of accuracy an error of less than 1 degree has been achieved by this system; which is very low compared to the existing devices. For instance, traditional goniometry has an error of 5 degrees (Bainbridge, 2000 and Williams, *et al.*, 2000). Also the new system is better than the goniometric glove and electrogoniometer which are currently still the subject of research and have an error similar to the traditional goniometer (Williams, N. *et al*, 2000).

The system can be operated by a semi-skilled person and all four fingers of one hand can be assessed at the same time. Also the measurement process takes only a few minutes to prepare the patient and then less than 5 seconds to run the computer vision system and obtain the joints angles, so it saves the evaluator's time.

As the new system is a non-contact measurement device, it causes no discomfort for the patient especially if the hand is injured. This facility is not available for the existing measurement tools such as the conventional goniometer and the goniometric glove.

It is possible to display the output from the measurement system process graphically; for instance a 3D simulation of the patient's hand could be reconstructed and displayed on a screen so that the assessor can visually check the progress of the finger range of motion, making it easier to monitor changes in the condition of the hand. Also the data can be electronically saved and retrieve swiftly at any time.

The 3D computer vision system has a lot of advantages as mentioned above, however to design a high precision 3D measuring system requires a lot of effort and cost as there are many factors affecting the efficiency of this system.

For instance, the level and quality of illumination can improve or decrease the image captured by the vision system. Choosing the optimal light distribution, the appearance of an object can be improved and the feature of interest clarified. Once a good quality image is obtained from the system, then the image enhancement techniques can be applied to process and analyse with fewer difficulties, and the features of interest can be extracted without information loss.

From the experimental procedure for this work, the effect of illumination on the captured images was very clear particularly when the Beam Splitter was used. The idea behind introducing the Beam Splitter was to distinguish between the virtual and real measurement points based on the brightness of these points. However, because of the illumination this task was difficult to achieve. As a result, a mirror was used instead of the Beam Splitter to cope with these illumination problems.

Also, the electronic devices which are essential components of the computer vision system, such as camera sensor and frame-grabber, may have a negative contribution to the accuracy of the system. When converting the analogue signal from the camera to a digital signal, to be read by the computer, some spatial mismatch between pixel's locations in the camera sensor and sampled image may occur. Also pixel shifting can have negative effect of on the system as shown in Figure 6.11. Moreover cameras often have noticeable geometric distortions caused by their optical systems.

Before carrying out the measurement processes using the proposed 3D computer vision system, the cameras must be calibrated using a proper calibration piece. At the beginning of this work a pyramid shape calibration piece was used to cover only one finger. However it was not possible to use the same shape to cover the four fingers, as increasing the size of the pyramid, will lead to increase the size of the working enclosure which is not comply with this computer vision system. As a result, the new calibration piece was introduced.

The calibration piece used in this was designed so that it covered the hand at extension, flexion, and Ulnar and Radial deviation. At the same time the calibration piece should be covered by the FOV of the cameras within the available working volume, which has been adopted for this work. As the hand occupies a volume in space, then the calibration piece should have a volumetric shape. The size of the calibration piece was calculated based on the dimension of the male hand because it is larger than the female hand, so the design will also cover the female hand.

Reconstructing the 3D coordinates for an object in space, using the DLT or Tsai technique, requires two or more images of the same object captured by cameras located at different locations. Consequently the (x,y) image co-ordinates for the captured object will differ from one image to another.

To use the DLT or Tsai technique for this work, a match or correspondence between all the measurement points of the finger joints from one image and the rest of the images should be established first. Also, to make the computer vision system efficient and reliable, the correspondence problem should be solved automatically. A new automated technique was developed and implemented to resolve this issue and it works at all hand locations within the calibration volume and at all hand positions.

A Coordinate Measurement Machine (CMM) was used, as a reference, to measure the X, Y, Z world co-ordinates for the measurement points. Then the result was compared with the one obtained from the computer vision system. To see the effect of the difference between the two results on the finger joints angles, a mathematical model was written. The model shows the effect of error on the finger joints angles is insignificant.

At the first stage of this work the measurement process did not cover the hand at all positions, i.e. the cameras did not capture the measurement points when the hand was at maximum flexion. The measurement process was carried out using both the DLT and the Tsai techniques.

The hand model was placed at several locations inside the control volume in order to measure the joints angles at different flexion and extension angles and at different hand locations. Then finger joints angles were measured by the CVS. The results show the errors between the CMM and the CVS are less than one degree at most angles; however there are one or two errors out of twelve are larger than one degree. After investigation it was found that these large errors were not related to the CVS but as a result of human error while measuring the angles by the CMM or from the electronic devices that compose the CVS such as frame grabber. The Table 7-1 below compares the results from the two techniques based on the CMM control data.

Table 7-1: Comparison between DLT and Tsai

	Std. Deviation (°)	Mean (°)	MSE (°)	RMSE (°)
DLT	0.62	0.60	0.75	0.86
Tsai	0.48	0.64	0.64	0.80

The Table 7-1 above shows that the Tsai method gave better results than the DLT, despite being more complicated than the DLT and given that its equations required optimization methods to be solved.

At the second stage of this work, to cover all the hand at different values of flexion and extension of the fingers, a Beam Splitter was added to the system to enable the cameras to see the measurement points. A plane mirror could have been used instead of the Beam Splitter. The difference between the plane mirror and the Beam Splitter is only the method of separating the virtual points from the real ones. Also the location of the light needs more care if the Beam Splitter is used so that the brightness of the real points is considerably higher than virtual ones (unlike using the plane mirror). From the experimental results, using a plane mirror for this work is better than the Beam Splitter because of the illumination matter.

The measurement process was carried out on a real human hand, in order to show that the UV gel works well and gives a good contrast between the skin and the measurement points. The finger joint angles obtained from the computer vision system were compared with the ones measured by a manual protractor. The results showed there was a large difference between the two measurement modalities. We cannot rely on the result from the protractor because it is not designed for this purpose; even if a proper manual goniometer was used we cannot use it to evaluate the computer vision system because of its large documented error.

Figure 7-1 below summarizes the finger joint angles measurement process.

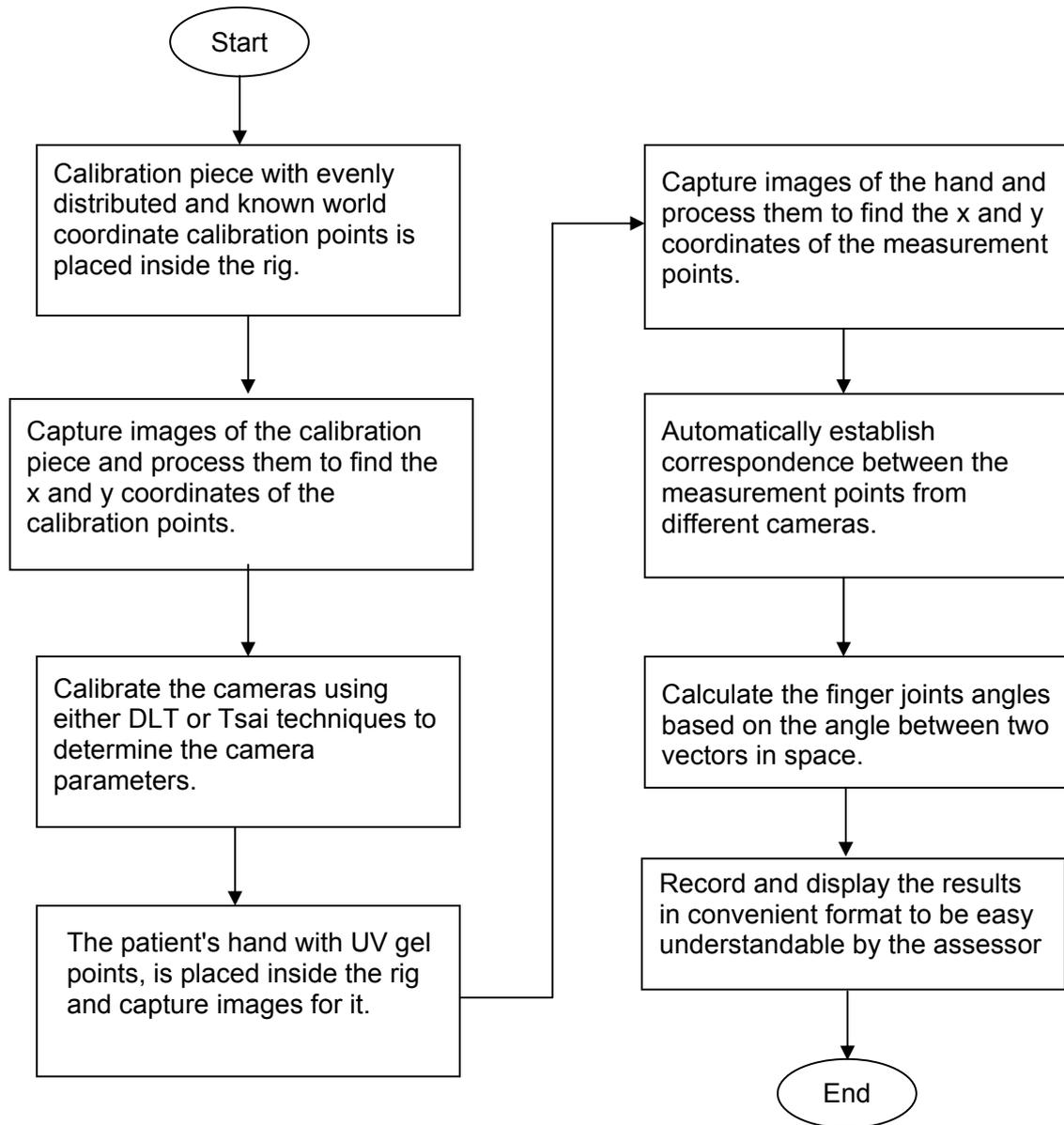


Figure 7-1: The procedures of the measurement process

7.1 Sources of error

There are several errors which can decrease the accuracy of the computer vision system, for instance:

a) The calibration points and the model hand finger joint angles were measured with a high accuracy Coordinate Measurement Machine; however human error can reduce the accuracy of the measurement process. To examine the human error the measurement process was repeated three times under the same conditions. The results showed the error to be around 0.35 degrees. This repeatability error affects the results from the computer vision system as the reading from the Coordinate Measurement Machine was used as a reference datum. In order to decrease this error, a robust statistical approach should be used when using the CMM to measure the calibration and the measurement points. For example, several measurements of the same point should be taken, and the mean or median value used.

b) When the calibration or the measurement points were obtained from the Coordinate Measurement Machine, the measurement probe of the machine, with a spherical shape, was placed roughly at the centre of the points, and when the images of the points were processed to obtain the x,y image coordinates for either the calibration or measurement points, the calculation of the x and y coordinates were based on the centroid, not on the centre. This lead to a mismatch between the points that the CMM used to calculate coordinates and the ones from the image system (although the centre and the centroid of a circle are the same, still less than half a pixel difference can contribute to the overall error).

c) Even if the image system is adjusted to calculate the x and y image coordinates based on the centre of the points, we cannot guarantee that the measurement probe picks up the same centre. This leads to an additional error

between the measurements obtained from the CMM and the ones from the computer vision system. It is very hard to determine the value of this error because of the existence of many parameters such as optical and electronic devices, but from the experimental work 0.3 shifting in pixel location can increase the error significantly.

d) The location of the calibration or the measurement points with respect to the camera's position can lead to a difference in the centroid of the same point, which adds an error to the system. If it is experimentally possible, all the cameras should see the same centroid of all the calibration and measurement piece are the same. For instance the original shape of the blob should ideally be a regular (convex) shape, if possible an ellipse or circle.

e) Shifting in the pixel location caused by the capture card, when it converts the signal from analogue to digital, can increase the error; this error can be avoided by using better capture card in terms of pixel shifting or by replacing the analogue cameras with digital ones.

7.2 Further work

The total number of measurement points used to measure the whole hand was 32. However, this number can be reduced without significant loss of useful information. For example, instead of 8 points for the Metacarpus bones (back of the hand), these can be reduced to 2 points only for the following reasons

- The Metacarpus bones are converging and placing the measurement points (the UV gel) needs extra care, not like the other finger bones where they are separated.
- In some people the shape of the back of the hand and the blood vessels makes placing the gel quite difficult.

- The change in angle value between the Metacarpus bones and the Proximal Phalanx bones depends on the movement of the Proximal Phalanx bones and not the Metacarpus bones.

When using a reduced number of points for the Metacarpus bone, the following procedures must be adopted:

- Place two points on one finger, for instance the middle finger.
- These two points represent a vector in space which can be used to measure the angle between this vector and all the other Proximal Phalanx bones.
- If it is desired to evaluate the progress that the patient has achieved after a certain period, we can compare the results from a specific time to another time.
- If it is desired to measure the angles and compare them with the standard joint angle (i.e. for a medically healthy person), it is necessary to determine the range of angles between the vector on the metacarpal bone for the middle finger and all the other Proximal Phalanx bones for healthy people by using the CVS or another available measurement method.

The size of the cameras and camera holders restricted and reduced the actual working volume, so they should be replaced with smaller ones. This will give more space and lead to an increase in the FOV of the system.

Finally the computer vision system explained in this thesis can replace the existing finger joint measurement devices and to be commercialized the following work needs to be done.

1) The user interface needs to be designed and the range of motion of the hand should be graphically displayed in 2D or 3D view so that the assessor can follow up any change of this motion graphically.

2) The computer vision system so far was mainly tested on a black painted hand model; this hand model has different properties from the real hand such as the colour, skin of the hand, and hand vessels and hair. So the system should be tested on real hands to see if the range of error is same as the one from testing the hand model. However the contrast between the skin of the hand and the measurement point was tested and the measurement point can be easily separated from the background (skin).

When the hand model was used to check the accuracy of the computer vision system, the results were compared with those obtained from the Coordinate Measurement Machine, with a theoretical accuracy of $0.7 \mu\text{m}$. However the Coordinate Measurement Machine cannot be used to measure the finger joint angles for a real hand because any object measured by a Coordinate Measurement Machine must be placed at a fixed position and not moved at all, which is very hard to achieve with the real hand.

From literature it has not been possible to find an accurate method to evaluate the angles measured by the computer vision system for the real hand. The traditional goniometer, which is currently used for evaluating the range of motion of the hand, cannot be used to assess the data obtained from the computer vision system because the traditional goniometer has too large error.

One method which could be used to evaluate the data from the computer vision system, for the real hand, is by using 2D photography. In this method one finger would be evaluated at a time and the finger imaged from the side, not from the top as in the computer vision system, so that the projection of the measurement points on the image plane shape the angles between the finger bones, see Figure 7-2 for more detail. The angles between the finger joints would be calculated from finding the angle between two vectors in a 2D plane.

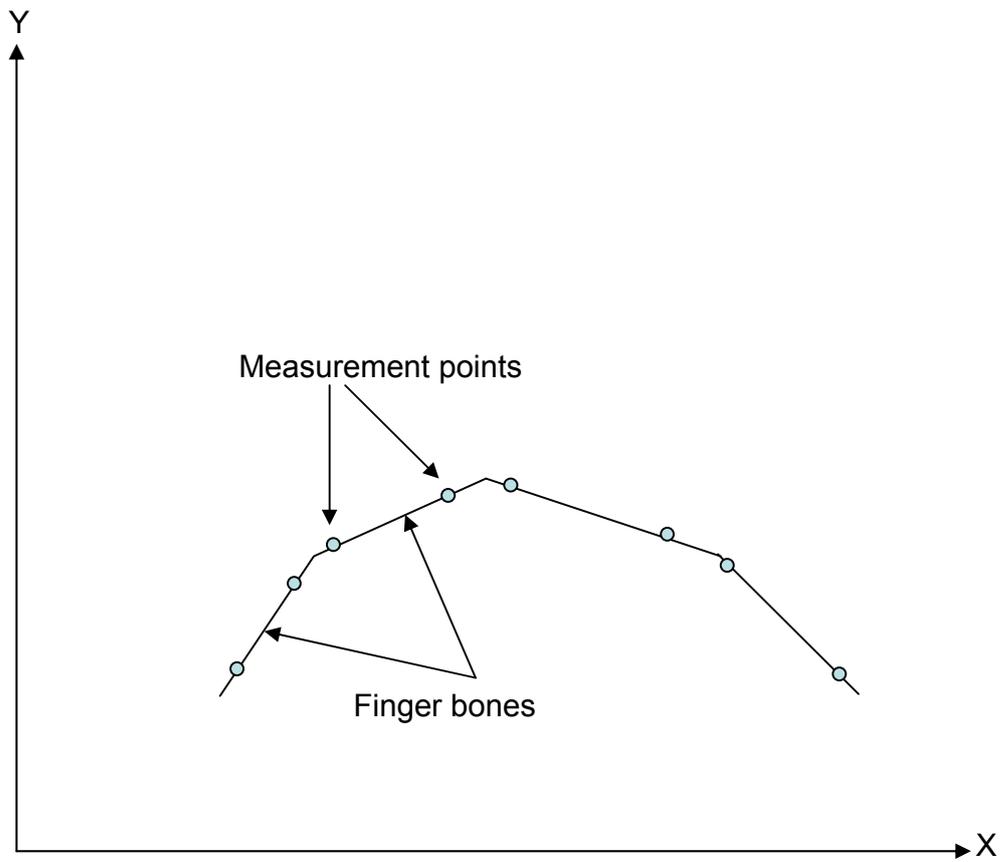


Figure 7-2: Projection of the measurement points for one finger on the X,Y image plane

REFERENCES

- Abdel-Aziz, Y. and Karara, H. (1971). Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. *Proceedings of the Symposium on Close-Range Photogrammetry*. Urbana, Illinois. pp. 1-18.
- Aguilar, J., Lope, M., Torres, F. and Blesa, A. (2005). Development of a stereo vision system for non-contact railway concrete sleepers measurement based in holographic optical elements. *Measurement*, **38**(2), pp. 154-165.
- Allard, P., Stokes, I. and Bianchi, J. (1995). *Three-dimensional analysis of human movement*. Champaign, USA. Human Kinetics.
- Azad, P., Asfour, T. and Dillmann, R. (2008). Robust real-time stereo-based markerless human motion capture. In *IEEE/RAS International Conference on Humanoid Robots*, Daejeon, Korea. pp. 700-707.
- Bainbridge, L. (2000). Automatic Assessment of Hand Range of Motion. Derbyshire Hospital.
- Belhumeur, P., and Kriegman, D. (1998). What Is the Set of Images of an Object Under All Possible Illumination Conditions? *International Journal of Computer Vision*, **28**(3), pp. 245.
- Bin, Z., Shaobin, S., Jinwei, S., Zhiyong, C., and Chunyang, X. (2010). 3D Reconstruction method from biplanar radiography using DLT Algorithm: Application to the Femur, *First International Conference on Pervasive Computing Signal Processing and Applications (PCSPA)*, Harbin, China, pp. 251-254.
- Brosnan, T. and Sun, D. (2004). Improving quality inspection of food products by computer vision—a review. *Journal of Food Engineering*, **61**(1), pp. 3-16.
- Bueche, F. (1977). *Principles of physics*. New York: McGraw-Hill.
- Burger, W. and Burge, M. (2009). Principles of digital image processing fundamental techniques [Homepage of Springer], [Online].
- Caillette, F. and Howard, T. (2004). Real-time markerless human body tracking using colored voxels and 3D blobs, *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 266-267.
- Cambridge, C. (1995). Range of motion measurement of the hand. In: Hunter, J., Mackin, E., and Callahan A., *Rehabilitation of the Hand: Surgery and Therapy*. 4th edition, St Louis: Mosby, pp. 93-107.
- Castleman, R. (1996). *Digital image processing*. Upper Saddle River, New Jersey: Prentice-Hall International.
- Challis, J. and Kerwin, D. (1992). Accuracy assessment and control point configuration when using the DLT for photogrammetry. *Journal of Biomechanics*, **25**(9), pp. 1053-1058.
- Chen, L., Armstrong, C. and Raftopoulos, D. (1994). An investigation on the accuracy of three-dimensional space reconstruction using the direct linear transformation technique. *Journal of Biomechanics*, **27**(4), pp. 493-500.

- Chen, M. (2002). Roundness measurements for discontinuous perimeters via machine visions. *Computers in Industry*, **47**(2), pp. 185-197.
- Chi-Fang, L. and Chih-Yang, L. (1999). A new approach to high precision 3-D measuring system. *Image and Vision Computing*, **17**(11), pp. 805-814.
- Clarkson, H. (2000). *Musculoskeletal Assessment: Joint Range of Motion and Manual Muscle Strength*. 2nd edition. Baltimore: Lippincott Williams & Wilkins.
- Dai, S. and Ji, Q. (2001). A new technique for camera self-calibration, *Proceedings of the International Conference on Robotics and Automation*, Seoul, Korea, pp. 2165-2170, vol.3.
- Dapena, J., Harman, E. and Miller, J. (1982). Three-dimensional cinematography with control object of unknown shape. *Journal of Biomechanics*, **15**(1), pp. 11-19.
- Demant, C., Streicher-Abel, B., and Waszkewitz, P. (1999). *Industrial image processing: visual quality control in manufacturing*. Berlin; New York: Springer.
- Dhond, U., and Aggarwal, J. (1989). Structure from Stereo A Review. *IEEE Transactions on Systems ON SYSTEMS, Man, and Cybernetics*, **19**, pp. 1489-1510.
- Dipietro, L., Sabatini, A. and Dario, P. (2003). Evaluation of an instrumented glove for hand-movement acquisition. *Journal of Rehabilitation Research and Development*, **40**(2), pp. 179-190.
- Doebelin, E. (1990). *Measurement systems: application and design*. New York: McGraw-Hill.
- Efford, N. (2000). *Digital image processing: a practical introduction using Java*. Harrow, England; New York: Addison-Wesley.
- Ellis, B., Bruton A. and Goddard J. (1997). Joint angle measurement: a comparative study of the reliability of goniometry and wire tracing for the hand. *Clinical Rehabilitation*, **11**(4), pp. 314-20.
- Ellis, B. and Bruton, A. (2002). A study to compare the reliability of composite finger flexion with goniometry for measurement of range of motion in the hand. *Clinical rehabilitation*, **16**, pp. 568-570.
- Galbiati, L. (1990). *Machine vision and digital image processing fundamentals*. Englewood Cliffs: Prentice Hall.
- Garrett, J. (1971). The adult human hand: some anthropometric and biomechanical considerations. *Human factors*, **13**(2), pp. 117-131.
- Gazzani, F. (1993). Comparative assessment of two algorithms for calibrating stereo photogrammetric systems. *Journal of Biomechanics*, **26**(12), pp. 1449-1454.
- Gerhardt, J., Cocchiarella, L. and Lea, R. (2002). *The Practical Guide to Range of Motion Assessment*. Chicago: American Medical Association.
- Gonzalez, R. and Woods, R. (2002). *Digital image processing*. Upper Saddle River, N.J.: Prentice Hall.

- Goulermas, J. and Liatsis, P. (2001). Hybrid symbiotic genetic optimisation for robust edge-based stereo correspondence. *Pattern Recognition*, **34**(12), pp. 2477-2496.
- Greene, W., and Heckman, J. (1994). *The clinical measurement of joint motion*. Rosemont: American Academy of Orthopaedic Surgeons.
- Gregory, W. (2002). *Biomechanics and control of torque production during prehension*, PHD thesis, Pennsylvania State University.
- Hatze, H., 1988. High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified DLT-approach. *Journal of Biomechanics*, **21**(7), pp. 533-538.
- Heikkila, J. and Silven, O., (1997). A four-step camera calibration procedure with implicit image correction. *Proceedings 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, pp.1106-1112.
- Heikkila, J. (2000). Geometric camera calibration using circular control points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**(10), pp. 1066-1077.
- Hemmings, M. (2002). Measurement the location of finger joints enabling assessment of manual dexterity and flexion using computer vision. *Final year project (undergraduate) report, Loughborough University, Mechanical Engineering*.
- Henderson, H. (2003). *Encyclopedia of Computer Science and Technology*. New York, NY: Facts On File.
- Hinrichs, R. and Mclean, S. (1995). NLT and extrapolated DLT: 3-D cinematography alternatives for enlarging the volume of calibration. *Journal of Biomechanics*, **28**(10), pp. 1219-1223.
- Horn, B. (2000). *Tsai Camera Calibration Method Revisited*, Massachusetts Institute of Technology.
- Hume, M., Gellman, H., Mckellop, H. and Brumfield J. (1990). Functional range of motion of the joints of the hand. *The Journal of Hand Surgery*, **15**(2), pp. 240-243.
- Jaegger, G., Marcellin-Little D. and Levine, D. (2002). Reliability of goniometry in Labrador Retrievers. *American Journal of Veterinary Research*, **63**(7), pp. 979-86.
- Jähne, B. (2002). *Digital image processing*. New York: Springer.
- Jain, R., Kasturl, R. and Schunck, B. (1995). *Machine vision*. New York: McGraw-Hill.
- Ji, Q. and DAI, S. (2004). Self-Calibration of a rotating camera with a translational offset. *IEEE Transactions on Robotics and Automation*, **20**, pp. 1-14.
- Kerr, D. (2003). Lecturer notice in computer vision and industrial image processing. Loughborough University.
- Klette, R., Schlüns, K., and Koschan, A. (1998). *Computer vision: three-dimensional data from images*. New York; London: Springer.

- Lea, R. and Gerhardt, J. (1995). Range-of-motion measurements. *The Journal of Bone and Joint Surgery. American volume*, **77**(5), pp. 784-798.
- Lee, C., Cooper, D. and Keren, D. (1993). Computing correspondence based on regions and invariants without feature extraction and segmentation, *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1993*, pp. 655-656.
- Lenz, R. and Fritsch, D. (1990). Accuracy of videometry with CCD sensors. *ISPRS Journal of Photogrammetry and Remote Sensing*, **45**(2), pp. 90-110.
- Leroux, M., Allard, P. and Murphy, N. (1991). Accuracy and precision of the direct linear transformation technique (DLT) in very close range photogrammetry with video cameras. *Journal of Biomechanics*, **24**(3-4), pp. 262.
- Leta, F., Feliciano, F., Souza, I. and Cataldo, E. (2006). Discussing accuracy in an automatic measurement system using computer vision techniques. *ABCM Symposium Series in Mechatronic*.
- Lin, K. (2003). Fast image thresholding by finding the zero(s) of the first derivative of between-class variance. *Machine vision and applications*, **13**, pp. 254-262.
- Low, J. (1976). The reliability of joint measurement. *Physiotherapy*, **62**(7), pp. 227-229.
- Luong, Q. and Faugeras, O. (1992). Self-calibration of a camera using multiple images, *Pattern Recognition. Proceedings of the 11th International Conference on Computer Vision and Applications*, Sophia- Antipolis, France, pp. 9-12.
- MacDermid, J., Fox, E., Richards, R. and Roth, J. (2001). Validity of pulp-to-palm distance as a measure of finger flexion. *The Journal of Hand Surgery. Volume*, **26**(5), pp. 432-435.
- Macionis, V. (2013). Reliability of the standard goniometry and diagrammatic recording of finger joint angles: a comparative study with healthy subjects and non-professional raters. *BMC Musculoskeletal Disorders*. Vol. **14**(1):17 (2013)
- Mallon, W., Brown, H. and Nunley, J. (1991). Digital ranges of motion: Normal values in young adults. *The Journal of Hand Surgery*, **16**(5), pp. 882-887.
- Marzan, G. and Karara, H. (1975). A computer program for direct linear transformation solution of collinearity condition and some applications of it. *Proceedings of the Symposium on Close-Range Photogrammetric Systems*, pp. 420-476.
- Masi, C. (1998). Lighting Makes Its Mark on Vision Systems-Light sources give many illuminations options. *Test and Measurement World*, **18**(6), pp. 11-16.
- Maybank, S. and Faugeras, O. (1992). Theory of self-calibration of a moving camera. *International Journal of Computer Vision*, **8**(2), pp. 123-151.
- Montabone, S. (2010). *Beginning digital image processing: using free tools for photographers*. New York, NY: Apress.
- Moore, M. (1978). *Clinical assessment of joint motion*. In: *Basmajian, Therapeutic exercise*. 3rd edition. Baltimore: Williams & Wilkins.
- Nelkon, M. and Parker, P. (1979). *Advanced level physics*. London: Heinemann.

- Norkin, C. and White, D. (2003). *Measurement of Joint Motion: A Guide to Goniometry*. Philadelphia, USA: Davis Company.
- Orteu, J. (2009). 3-D computer vision in experimental mechanics. *Optics and Lasers in Engineering*, **47**(3-4), pp. 282-291.
- Parish, J. Skeletal hand charts in inherited connective tissue disease (1968). *Journal of Medical Genetics*. Vol. 4(4).pp 227-38
- Pourcelot, P., Audigie, F., Degueurce, C., GEIGER, D. and DENOIX, J. (2000). A method to synchronise cameras using the direct linear transformation technique. *Journal of Biomechanics*, **33**(12), pp. 1751.
- Rand, D. and Nicol, A. (1993). An instrumental glove for monitoring MCP joint motion. *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of Engineering in Medicine*, **207**(4), pp. 207-210.
- Ringer, M. and Lasenby, J. (2000). Modelling and tracking articulated motion from multiple camera views. Cambridge, England: Signal Processing Group, Dept. of Engineering, University of Cambridge.
- Robert, L. (1996). Camera Calibration without Feature Extraction. *Computer Vision and Image Understanding*, **63**(2), pp. 314-325.
- Rose, V., Nduka, C., Pereira, J., Pickford, M. and Belcher, H. (2002). Visual estimation of finger angles: do we need goniometers? *The Journal of Hand Surgery: British & European Volume*, **27**(4), pp. 382-384.
- Salter, N. (1955). Methods of measurement of muscle and joint function. *The Journal of Bone and Joint Surgery. British volume*, **37-B**(3), pp. 474-491.
- Salvi, J., Armangué, X. and Batlle, J. (2002). A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, **35**(7), pp. 1617-1635.
- Sang, D. (1996). A self-calibration technique for active vision systems. *IEEE Transactions on Robotics and Automation*, **12**(1), pp. 114-120.
- Shah, S. and Aggarwal, J. (1996). Intrinsic parameter calibration procedure for a (high-distortion) fish-eye lens camera with distortion model and accuracy estimation. *Pattern Recognition*, **29**(11), pp. 1775-1788.
- Shirai, Y. (1992). 3D computer vision and applications. In: 11th *International Conference on Pattern Recognition. International Association for Pattern Recognition*, IEEE Computer Society Press, pp. 236-245.
- Shirai, Y. (1987). *Three-dimensional computer vision*. Berlin; New York: Springer-Verlag.
- Simpson, C. (2002). *Hand assessment guide for therapists*. UK: Salisbury: APS.
- Sonka, M., Hlavac V. and Boyle, R. (1999). *Image processing, analysis, and machine vision*. 2nd edition. London: PWS Publishing.
- Sturman, D. and Zeltzer, D. (1994). A Survey of Glove-based Input. Clumsy intermediary devices constrain our interaction with computers and their applications. Glove-based input

devices let us apply our manual dexterity to the task. *IEEE computer graphics and applications.*, **14**(1), pp. 30.

Sue, W. and Adnan, A. (2003). Automatic thresholding of gray-level using multistage approach. *Proceeding, Seventh International Conference on Document Analysis and Recognition*, 2003, vol.1, pp. 493-497.

Thomas Y. , Robert G. (2000). Comparison between using spectral analysis of electrogoniometer data and observational analysis to quantify repetitive motion and ergonomic changes in cyclical industrial work. *Ergonomics*, 2000, VOL. 43, NO. 1, 106-132

Titus, J. (2001). Features - Machine vision - Lights make machine vision shine - The proper light source can ensure the success of your vision system. *Test & measurement world*, **21**(7), pp. 33.

Tsai, D. (1995). A fast thresholding selection procedure for multimodal and unimodal histograms. *Pattern Recognition Letters*, **16**(6), pp. 653-666.

Tsai, R. (1986). Efficient and accurate camera calibration technique for 3d machine vision. *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 1986, New York, NY, USA, pp. 364-374.

Tsai, R. (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal on Robotics and Automation*. *3*, 323-344.

Weiss, P., August, S., Peters, G. and Sampalis, J. (1994). Using the exos handmaster to measure digital range of motion: reliability and validity. *Medical engineering & physics*, **16**(4), pp. 323-328.

Wikipedia (2013). Inclinator [online]. Available at: <http://en.wikipedia.org/wiki/Inclinator> [Access 15 March 2013].

Williams, N, Penrose, J., Caddy, C., Barnes, E., Hose, D. and Harley, P. (2000). A Goniometric glove for clinical hand assessment: Construction, calibration and validation. *The Journal of Hand Surgery: British & European Volume*, **25**(2), pp. 200-207.

Wagner, C (1988). The pianist's hand: anthropometry and biomechanics *Ergonomics*. Vol. 31 No 1. pp 97-131.

Wise S, Gardner W, Sabelman E, Valainis E, Wong Y, Glass K, Drace J, & Rosen J. (1990). Evaluation of a fiber optic glove for semi-automated goniometric measurements. *Journal of Rehabilitation Research and Development*. 27, 411-24.

Wittels, N., McClellan, J., Cushing, K., Howard, W. and Palmer, A. (1989). *How to select cameras for machine vision*. Cambridge, MA, USA.

Wood, G. and Marshall, R. (1986). The accuracy of DLT extrapolation in three-dimensional film analysis. *Journal of Biomechanics*, **19**(9), pp. 781-785.

Würz-Wessel, A. (2003). *Free formed surface mirrors in computer vision systems*. PHD thesis, Tübingen University, Germany.

Zehang, S. and Christian, R. (2001). Camera calibration. *Intership with UNR and Ford Motor company*, .

Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol.22, **22**(11), pp. 1330.

Zhang, Z. and Tsui, H. (1998). 3D reconstruction from a single view of an object and its image in a plane mirror. *Proceedings of the 14th International Conference on Pattern Recognition 1998*, Washington, DC, USA, vol.2, pp. 1174-1176.

Zhengyou, Z. (2000). A Flexible New Technique for Camera Calibration. *IEEE transactions on pattern analysis and machine intelligence*, **22**(11), pp. 1330-1334.

Zhou, H. and HU, H. (2008). Human motion tracking for rehabilitation—A survey. *Biomedical Signal Processing and Control Biomedical Signal Processing and Control*, **3**(1), pp. 1-18.

Appendix A

A.1 Tsai camera model

The following explains how we can find the camera parameters by the method proposed by Tsai and how the 3-D space co-ordinates for a point can be obtained.

This technique involves four steps of transformation as illustrated below:

Step 1:

Transformation from the object world coordinate system (x_w, y_w, z_w) to the camera 3D coordinate system (x, y, z) shown as step 1 in Figure (7-1), can be described by the following equation.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T \quad (7-1)$$

Where is R is the 3 X 3 rotation matrix defining the camera orientation and T is a translation vector representing the camera position, which can further be defined as

$$R \equiv \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \quad (7-2)$$

And the translation vector as

$$T \equiv \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$

There are six extrinsic parameters: the three angles of rotation yaw θ , pitch ϕ , and tilt ψ , and the three components for the translation vector T . The rotation matrix R can be expressed as function of θ, ϕ and ψ as follows:

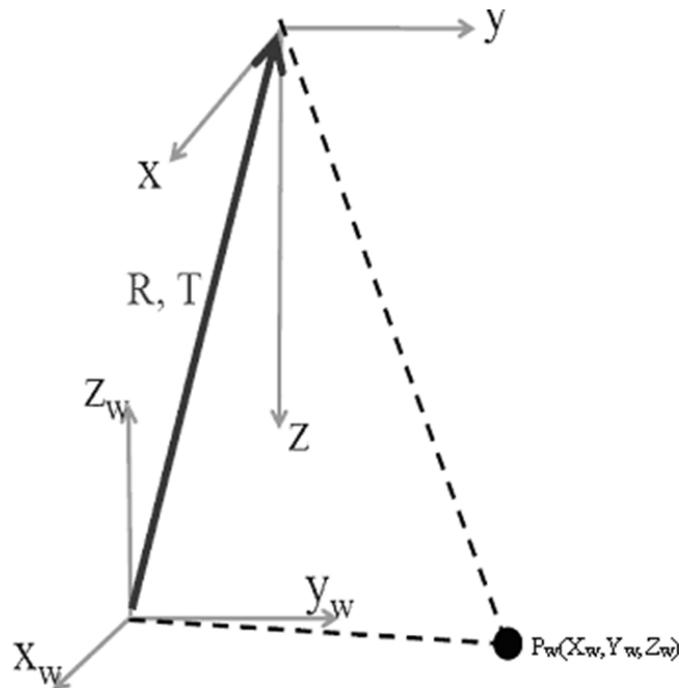


Figure A-1: World-to-camera Transformation

$$R \equiv \begin{bmatrix} \cos \psi \cos \theta & \sin \psi \cos \theta & -\sin \theta \\ -\sin \psi \cos \phi + \cos \psi \sin \theta \cos \phi & \cos \psi \cos \phi + \sin \psi \sin \theta \sin \phi & \cos \theta \sin \phi \\ \sin \psi \sin \phi + \cos \psi \sin \theta \cos \phi & -\cos \psi \sin \phi + \sin \psi \sin \theta \sin \phi & \cos \theta \sin \phi \end{bmatrix} \quad (7-3)$$

The parameters to be calibrated are R and T .

Step 2:

Transformation from 3D camera coordinates (x, y, z) to ideal (undistorted) image coordinates (X_u, Y_u) using perspective projection with pinhole camera geometry Fig (7-2).

$$X_u = f \frac{x}{z} \quad (7-4)$$

$$Y_u = f \frac{y}{z} \quad (7-5)$$

The parameter to be found is the effective focal length f .

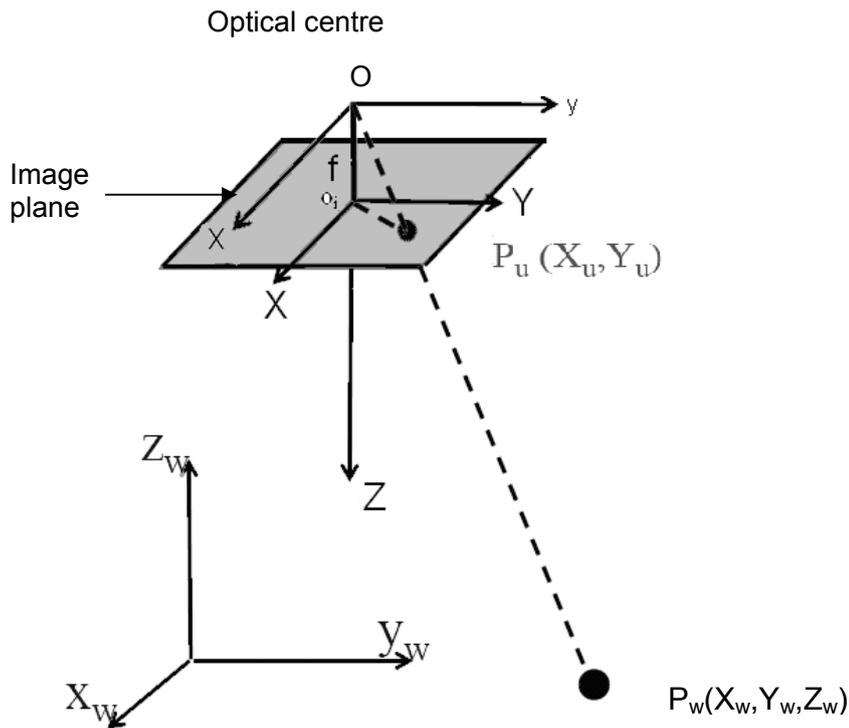


Figure A-2: Ideal transformation from 3D camera coordinates to image Coordinates

Step 3: Transformation from ideal (undistorted) (X_u, Y_u) to distorted image coordinates (X_d, Y_d)

The distorted image coordinates (X_d, Y_d) can be determined from undistorted image coordinates (X_u, Y_u) using the equations

$$D_x = X_u - X_d \quad (7-6)$$

$$D_y = Y_u - Y_d \quad (7-7)$$

Where:

(X_d, Y_d) is the distorted image coordinate on the image plane, and

$$D_x = X_d(k_1 r^2 + k_2 r^4 + \dots)$$

$$k_1, k_2, \dots$$

k_1, k_2 are distortion coefficients

$$D_y = Y_d(k_1 r^2 + k_2 r^4 + \dots)$$

$$r = \sqrt{X_d^2 + Y_d^2}$$

Tsai wrote that from his experience of an industrial machine vision application, only radial distortion needs to be considered, and only one term is needed (Tsai, R.Y. 1987). The positive values of k_1 means that the image has to be stretched towards the corner (pincushion distortion). On the other hand a negative value of k_1 means the inverse situation (barrel distortion).

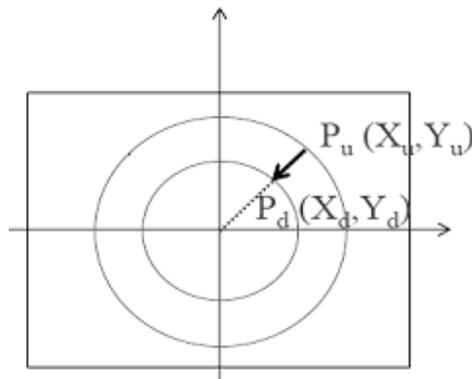


Figure A-3: location of point P before and after applying radial Lens Distortion

Step 4:

Transformation from real image coordinates (X_d, Y_d) to computer image coordinates: (X_f, Y_f)

$$X_f = s_x d'_x{}^{-1} X_d + C_x \quad (7-8)$$

$$Y_f = d_y{}^{-1} Y_d + C_y \quad (7-9)$$

$$d'_x = d_x \frac{N_{cx}}{N_{fx}}$$

Where:

(X_f, Y_f) row and column numbers of the image pixel in computer frame memory.

(C_x, C_y) row and column numbers of the centre of computer frame memory.

d_x Centre to centre distance between adjacent sensor elements in X (scan line) direction (mm).

d_y Centre to centre distance between adjacent sensor elements in Y direction (mm).

N_{cx} Number of sensor elements in camera's x direction.

N_{fx} Number of pixels in frame grabber's x direction .

s_x Is the uncertainty scale factor in vertical direction only.

The parameter to be found is the uncertainty image scale s_x .

A.1.1 How to Implement the Tsai technique

The following procedure illustrates how the Tsai method can be used to calibrate cameras.

A. Computing the distorted image coordinates (X_d, Y_d)

1. capture an image for the calibration points. From the image find out the row and column number of each calibration point and call it (X_{fi}, Y_{fi})

2. find N_{cx}, N_{fx}, d'_x and d_y using the information supplied by the camera manufacturer.

3. find the centre pixel of frame memory (C_x, C_y) .

4. calculate (X_{di}, Y_{di}) using equations (8,9)

$$X_{di} = s_x^{-1} d'_x{}^{-1} (X_{fi} - C_x)$$

$$Y_{di} = d_y (Y_{fi} - C_y)$$

Where:

$i=1, \dots, N$, N is the total number of calibration points.

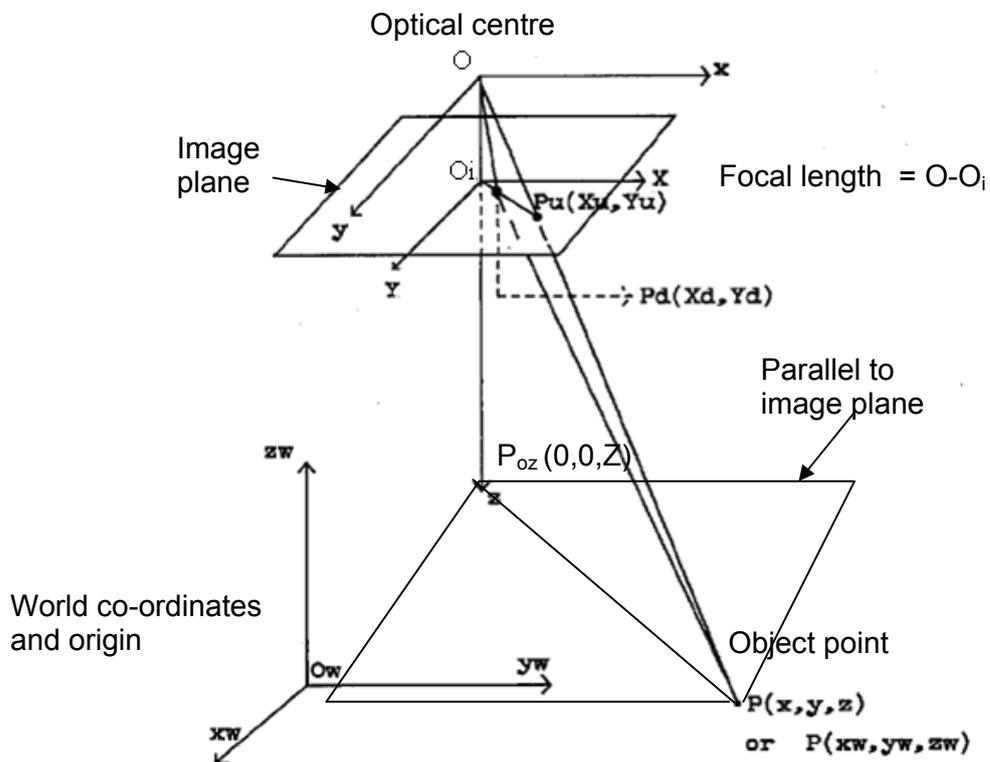


Figure A-4: Tsai camera model

From Fig (7-4) the image plane parallel to the plane passing through P and P_{oz} and $\overline{O_i P_d} \parallel \overline{P_{oz} P}$ and they are the intersection of a plane passing through O, P_{oz} and P . Similarly $\overline{O_i P_d} \parallel \overline{P_{oz} P}$.

Since the cross product of any two parallel vectors is the null vector then

$$\overline{O_i P_d} \times \overline{P_{oz} P} = 0 \quad (7-10)$$

$$\overline{O_i P_d} = (X_d - 0, Y_d - 0) = (X_d, Y_d) \quad (7-11)$$

Similarly

$$\overline{P_{oz} P} = (x - 0, y - 0, z - z) = (x, y) \quad (7-12)$$

By substituting the value of $\overline{O_i P_d}$ and $\overline{P_{oz} P}$ from equations (11, 12) into (10) we get:

$$(X_d, Y_d) \times (x, y) = 0 \quad (7-13)$$

$$(X_d \cdot y - Y_d \cdot x) = 0 \quad (7-14)$$

From equation (1) and (2) we have

$$x = r_1 x_w + r_2 y_w + r_3 z_w + T_x \quad (7-15)$$

$$y = r_4 x_w + r_5 y_w + r_6 z_w + T_y \quad (7-16)$$

$$z = r_7 x_w + r_8 y_w + r_9 z_w + T_z \quad (7-17)$$

Substitute the values of (x, y, z) from equations (15, 16) into (14) we get,

$$Y_d \cdot (r_1 x_w + r_2 y_w + r_3 z_w + T_x) = X_d \cdot (r_4 x_w + r_5 y_w + r_6 z_w + T_y) \quad (7-18)$$

$$Y_d r_1 x_w + Y_d r_2 y_w + Y_d r_3 z_w + Y_d T_x - X_d r_4 x_w - X_d r_5 y_w - X_d r_6 z_w = X_d T_y \quad (7-19)$$

Dividing (19) by T_y^{-1} we get:

$$Y_d T_y^{-1} r_1 x_w + Y_d T_y^{-1} r_2 y_w + Y_d T_y^{-1} r_3 z_w + Y_d T_y^{-1} T_x - X_d T_y^{-1} r_4 x_w - X_d T_y^{-1} r_5 y_w - X_d T_y^{-1} r_6 z_w = X_d \quad (7-20)$$

By rearranging the terms in (20), the following equation is obtained.

$$\begin{bmatrix} Y_{di}x_{wi} & Y_{di}y_{wi} & Y_{di}z_{wi} & Y_{di} & X_{di}x_{wi} & -X_{di}y_{wi} & -X_{di}z_{wi} \end{bmatrix} \begin{bmatrix} T_y^{-1}s_xr_1 \\ T_y^{-1}s_xr_2 \\ T_y^{-1}s_xr_3 \\ T_y^{-1}s_xT_x \\ T_y^{-1}r_4 \\ T_y^{-1}r_5 \\ T_y^{-1}r_6 \end{bmatrix} = X_{di} \quad (7-21)$$

Where:

s_x is the uncertainty scale factor and its initial value is equal to 1

B. compute $T_y^{-1}s_xr_1, T_y^{-1}s_xr_2, T_y^{-1}s_xr_3, T_y^{-1}s_xT_x, T_y^{-1}r_4, T_y^{-1}r_5, T_y^{-1}r_6,$

Using equation (21) with the number of control points i at seven or more, an overdetermined system of linear equations can be solved for the seven unknowns $T_y^{-1}s_xr_1, T_y^{-1}s_xr_2, T_y^{-1}s_xr_2, T_y^{-1}s_xT_x, T_y^{-1}r_4, T_y^{-1}r_5, T_y^{-1}r_6$

C. compute $(r_1, \dots, r_6, T_x, T_y)$ from $T_y^{-1}s_xr_1, T_y^{-1}s_xr_2, T_y^{-1}s_xr_2,$

$$T_y^{-1}s_xT_x, T_y^{-1}r_4, T_y^{-1}r_5, T_y^{-1}r_6$$

1) Finding $|T_y|$ from $T_y^{-1}s_xr_1, T_y^{-1}s_xr_2, T_y^{-1}s_xr_2,$

$$T_y^{-1}s_xT_x, T_y^{-1}r_4, T_y^{-1}r_5, T_y^{-1}r_6$$

Let

$$a_1 = T_y^{-1}s_xr_1 \quad a_2 = T_y^{-1}s_xr_2 \quad a_3 = T_y^{-1}s_xr_3 \quad a_4 = T_y^{-1}s_xT_x$$

$$a_5 = T_y^{-1}r_4 \quad a_6 = T_y^{-1}r_5 \quad a_7 = T_y^{-1}r_6$$

Where:

a_1, a_2, \dots, a_7 were determined from equation (21)

To compute $|T_y|$

Form the orthonormal property of R and the definition of a_5, a_6 and a_7 we have

$$\begin{aligned} (a_5^2 + a_6^2 + a_7^2)^{-1/2} &= ((T_y^{-1}r_4) + (T_y^{-1}r_5) + (T_y^{-1}r_6))^{-1/2} \\ &= |T_y| \cdot (r_4^2 + r_5^2 + r_6^2)^{-1/2} = |T_y| \end{aligned}$$

$$\text{Since } (r_4^2 + r_5^2 + r_6^2) = 1$$

Then

$$|T_y| = (a_5^2 + a_6^2 + a_7^2)^{-1/2}$$

To determine the sign of T_y , we chose a point i in the computer image coordinate whose image position (X_{fi}, Y_{fi}) lies far away from the principal point (C_x, C_y) and its world coordinate is (x_{wi}, y_{wi}, z_{wi}) . First the sign of T_y is assumed to be positive. We then find the following parameters by putting the value of $s_x = 1$.

$$r_1 = (T_y^{-1} s_x r_1) \cdot T_y \quad r_2 = (T_y^{-1} s_x r_2) \cdot T_y$$

$$r_3 = (T_y^{-1} s_x r_3) \cdot T_y \quad r_4 = (T_y^{-1} r_4) \cdot T_y$$

$$r_5 = (T_y^{-1} r_5) \cdot T_y \quad r_6 = (T_y^{-1} r_6) \cdot T_y$$

$$T_x = (T_y^{-1} s_x T_x) \cdot T_y$$

$$x = r_1 x_w + r_2 y_w + r_3 z_w + T_x$$

$$y = r_4 x_w + r_5 y_w + r_6 z_w + T_y$$

where:

$$\begin{aligned} &T_y^{-1} s_x r_1, T_y^{-1} s_x r_2, T_y^{-1} s_x r_3, T_y^{-1} s_x T_x, T_y^{-1} r_4, T_y^{-1} r_5, \\ &T_y^{-1} r_6 \end{aligned}$$

are determined from equation (21)

if $(x$ and $X_{fi})$ have the same sign and $(y$ and $Y_{fi})$ have the same sign, then T_y is positive, else T_y is negative.

2) Determine s_x

$$\begin{aligned} & (a_1^2 + a_2^2 + a_3^2)^{-1/2} |T_y| = \\ & ((T_y^{-1} s_x r_1)^2 + (T_y^{-1} s_x r_2)^2 + (T_y^{-1} s_x r_3)^2)^{-1/2} |T_y| \\ & (a_1^2 + a_2^2 + a_3^2)^{-1/2} |T_y| = s_x (r_1^2 + r_2^2 + r_3^2)^{-1/2} \end{aligned} \quad (7-22)$$

From the orthonormal property of the rotation matrix R we have

$$(r_1^2 + r_2^2 + r_3^2) = 1$$

Then

$$s_x = (a_1^2 + a_2^2 + a_3^2)^{-1/2} |T_y| \quad (7-23)$$

3) Computing the rotation matrix $R(r_1, r_2, r_3, \dots, r_9)$

$$\begin{aligned} r_1 &= a_1 \cdot T_y / s_x & r_2 &= a_2 \cdot T_y / s_x & r_3 &= a_3 \cdot T_y / s_x \\ r_4 &= a_5 \cdot T_y & r_5 &= a_6 \cdot T_y & r_6 &= a_7 \cdot T_y \\ T_x &= a_4 \cdot T_y \end{aligned}$$

Where a_1, a_2, \dots, a_7 are the seven variables which computed in (1).

Once r_1, r_2, \dots, r_6 have been found, which are the first two rows of the rotation matrix R , the third row (r_7, r_8, r_9) can be computed as the cross product of the first two rows, using the orthonormal property of R as follows

$$\begin{aligned} r_7 &= r_2 \times r_6 - r_3 \times r_5 \\ r_8 &= r_3 \times r_4 - r_1 \times r_6 \\ r_9 &= r_1 \times r_5 - r_2 \times r_4 \end{aligned}$$

4) Computing the focal length, distortion coefficients, and z position.

By rearranging equations (8, 9) we obtain

$$X_d = (X_f - C_x) s_x^{-1} d'_x$$

$$Y_d = (Y_f - C_y) d_y$$

Substitute X_d and Y_d in equations (6, 7)

$$X_u = (X_f - C_x) s_x^{-1} d'_x + (X_f - C_x) s_x^{-1} d'_x k_1 r^2$$

$$Y_u = (Y_f - C_y) d_y + (Y_f - C_y) d_y k_1 r^2$$

From equations (4, 5) and the values of X_u, Y_u which are obtained above, the computer coordinates are related to the 3D coordinate of the object point in the camera coordinate system according to the following equations.

$$s_x^{-1} d'_x X + s_x^{-1} d'_x X k_1 r^2 = f \frac{x}{z} \quad (7-24)$$

$$d'_y Y + d'_y Y k_1 r^2 = f \frac{y}{z} \quad (7-25)$$

Where:

$$r^2 = (s_x^{-1} d'_x X)^2 + (d'_y Y)^2$$

$$X = (X_f - C_x)$$

$$Y = (Y_f - C_y)$$

Substitute values of (x, y, z) from (15, 16, 17) into (24, 25) we get

$$s_x^{-1} d'_x X + s_x^{-1} d'_x X k_1 r^2 = f \frac{r_1 x_w + r_2 y_w + r_3 z_w + T_x}{r_7 x_w + r_8 y_w + r_9 z_w + T_z} \quad (7-26)$$

And

$$d'_y Y + d'_y Y k_1 r^2 = f \frac{r_4 x_w + r_5 y_w + r_6 z_w + T_y}{r_7 x_w + r_8 y_w + r_9 z_w + T_z} \quad (7-27)$$

5a) Find initial values of f and T_z by putting the value of the lens distortion coefficient equal to zero in (27)

$$d'_y Y = f \frac{r_4 x_w + r_5 y_w + r_6 z_w + T_y}{r_7 x_w + r_8 y_w + r_9 z_w + T_z} \quad (7-28)$$

(28)

Equation (28) can be written as

$$d'_y Y w + d'_y Y T = f y \quad (7-29)$$

Where:

$$y_i = r_4 x_{wi} + r_5 y_{wi} + r_6 z_{wi} + T_y$$

$$w_i = r_7 x_{wi} + r_8 y_{wi} + r_9 z_{wi}$$

From equation (29) we get

$$d_y 'Y_i w_i = y_i f - d_y 'Y T_z \quad (7-30)$$

Using equation (30) with several calibration points, gives an overdetermined system of linear equations with unknown f and T_z . Tsai suggested using the steepest decent method to solve these equations, but here we are using a least squares method as below:

From equation (30):

$$f y_i - d_y T_z Y_i = d_y Y_i w_i$$

Using least squares method

$$Error = E(f, T_z) = \sum_{i=1}^n (f y_i - d_y T_z Y_i - d_y Y_i w_i)^2$$

The least value of the squared error occurs when the partial derivative of the error with respect to f and T_z is equal to zero.

$$\begin{aligned} \frac{\partial E}{\partial f} &= 2 \sum_{i=1}^n (f y_i - d_y T_z Y_i - d_y Y_i w_i) y_i \\ &= 2 \sum_{i=1}^n f y_i^2 - 2 d_y T_z \sum_{i=1}^n Y_i y_i - 2 \sum_{i=1}^n d_y Y_i w_i y_i \\ &= 2 f \sum_{i=1}^n y_i^2 - 2 T_z \sum_{i=1}^n d_y Y_i y_i - 2 \sum_{i=1}^n d_y Y_i w_i y_i = 0 \\ f \sum_{i=1}^n y_i^2 - T_z \sum_{i=1}^n d_y Y_i y_i &= \sum_{i=1}^n d_y Y_i w_i y_i \end{aligned} \quad (7-31)$$

Similarly $\frac{\partial E}{\partial T_z}$ gives

$$f \sum_{i=1}^n y_i Y_i - T_z \sum_{i=1}^n d_y Y_i^2 = \sum_{i=1}^n d_y Y_i^2 w_i \quad (7-32)$$

By rearranging equations (31, 32) we get:

$$\begin{bmatrix} \sum_{i=1}^n y_i^2 & -\sum_{i=1}^n d_y Y_i y_i \\ \sum_{i=1}^n y_i Y_i & -\sum_{i=1}^n d_y Y_i^2 \end{bmatrix} \begin{bmatrix} f \\ T_z \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n d_y Y_i w_i y_i \\ \sum_{i=1}^n d_y Y_i^2 w_i \end{bmatrix} \quad (7-33)$$

Where:

$$y_i = r_4 x_{wi} + r_5 y_{wi} + r_6 z_{wi} + T_y$$

$$w_i = r_7 x_{wi} + r_8 y_{wi} + r_9 z_{wi}$$

By solving equation (33) the approximate values of f and T_z are obtained.

5b) finding the exact values of f, T_z and k_1

With R, T_x and T_y have been determined previously, equation (27) becomes a nonlinear equation with f, T_z and k_1 as unknowns which can be found as explained below.

$$f y_i = d_y Y_i w_i + d_y Y_i k_1 r^2 w_i + d_y T_z Y_i + d_y T_z Y_i k_1 r^2 \quad (7-34)$$

Using least squares method

$$Error = E(f, T_z, k_1) = \sum_{i=1}^n (f y_i - d_y Y_i w_i - d_y Y_i k_1 r^2 w_i - d_y T_z Y_i - d_y T_z Y_i k_1 r^2)^2$$

$$\frac{\partial E}{\partial f}, \frac{\partial E}{\partial T_z}, \frac{\partial E}{\partial k_1}$$

$$\frac{\partial E}{\partial f} = \sum_{i=1}^n 2(f y_i - d_y Y_i w_i - d_y Y_i k_1 r^2 w_i - d_y T_z Y_i - d_y T_z Y_i k_1 r^2) y_i$$

$$\begin{aligned}
&= 2 \sum_{i=1}^n (f y_i^2 - d_y Y_i w_i y_i - d_y Y_i k_1 r^2 w_i y_i - d_y T_z Y_i y_i - d_y T_z Y_i k_1 r^2 y_i) = 0 \\
&y_i^2 - d_y k_1 \sum_{i=1}^n Y_i r^2 w_i y_i - d_y T_z \sum_{i=1}^n Y_i y_i - d_y T_z k_1 \sum_{i=1}^n Y_i r^2 y_i = d_y \sum_{i=1}^n Y_i w_i y_i \\
&f = \frac{d_y \sum_{i=1}^n Y_i w_i y_i - d_y k_1 \sum_{i=1}^n Y_i r^2 w_i y_i + d_y T_z \sum_{i=1}^n Y_i y_i + d_y T_z k_1 \sum_{i=1}^n Y_i r^2 y_i}{\sum_{i=1}^n y_i^2} \quad (7-35)
\end{aligned}$$

And similar from $\frac{\partial E}{\partial T_z}$ we have

$$k_1 = \frac{d_y \sum_{i=1}^n Y_i w_i r^2 [d_y Y_i r^2 (w_i - T_z)]^2 + f \sum_{i=1}^n d_y y_i Y_i r^2 (w_i - T_z) + d_y T_z \sum_{i=1}^n Y_i (d_y Y_i r^2 (w_i - T_z))}{\sum_{i=1}^n [d_y (Y_i r^2 (w_i - T_z))]^2} \quad (7-36)$$

And from $\frac{\partial E}{\partial k_1}$ we have

$$k_1 = \frac{d_y \sum_{i=1}^n Y_i w_i r^2 [d_y Y_i r^2 (w_i - T_z)]^2 + f \sum_{i=1}^n d_y y_i Y_i r^2 (w_i - T_z) + d_y T_z \sum_{i=1}^n Y_i (d_y Y_i r^2 (w_i - T_z))}{\sum_{i=1}^n [d_y (Y_i r^2 (w_i - T_z))]^2} \quad (7-37)$$

Having the initial values of f and T_z from equation (33) and zero as an initial guess for k_1 , and by substituting these initial values into the equations (35,36,37) the exact values of f , T_z and k_1 can be obtained through iterations.

A.1.2 Reconstruction of 3D world coordinates using Tsai technique

From equation (26) we have:

$$\begin{aligned}
&A r_7 x_w + A r_8 y_w + A r_9 z_w + A T_z = f r_1 x_w + f r_2 y_w + f r_3 z_w + f T_x \\
&(A r_7 - f r_1) x_w + (A r_8 - f r_2) y_w + (A r_9 - f r_3) z_w = f T_x - A T_z \quad (7-38)
\end{aligned}$$

Where:

$$A = s_x^{-1} d_x' X + s_x^{-1} d_x' X k_1 r^2$$

And from equation (27) we have

$$\begin{aligned} B r_7 x_w + B r_8 y_w + B r_9 z_w + B T_z &= f r_4 x_w + f r_5 y_w + f r_6 z_w + f T_y \\ (B r_7 - f r_4) x_w + (B r_8 - f r_5) y_w + (B r_9 - f r_6) z_w &= f T_y - B T_z \end{aligned} \quad (7-39)$$

Where:

$$B = d_y Y + d_y Y k_1 r^2$$

By rearranging equations (38,39) and with two or more cameras, the following reconstruction matrix is used to find the world coordinates (x_w, y_w, z_w) .

$$\begin{bmatrix} A_1 r_7 - f r_4 & A_1 r_8 - f r_5 & A_1 r_9 - f r_6 \\ B r_7 - f r_4 & B r_8 - f r_5 & B r_9 - f r_6 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ A_m r_7 - f r_4 & A_m r_8 - f r_5 & A_m r_9 - f r_6 \\ B_m r_7 - f r_4 & B_m r_8 - f r_5 & B_m r_9 - f r_6 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \begin{bmatrix} f T_y - A_1 T_z \\ f T_y - B_1 T_z \\ \vdots \\ \vdots \\ f T_y - A_m T_z \\ f T_y - B_m T_z \end{bmatrix} \quad (7-40)$$

Where:

m is the number of camera used in the system.

A.2 Mathematical formula for DLT

In any camera system the light is passed through a lens onto a point on the image plane. The colinearity condition states that the world coordinates of the object (O), and the image coordinates of an image (I) lie on a line, which passes through the optical centre of the lens (N) Fig (8-6).

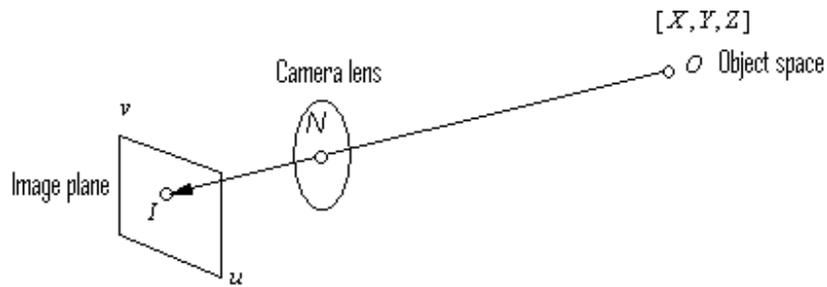


Figure A-5: the colinearity condition

Using the colinearity condition principle, it is possible to project the image back through the lens centre onto a comparator plane; the aim of this projection is to remove inversion which otherwise will be present.

Fig (7-6) shows the object-space reference frame $[X, Y, Z]$ and the new image-plane reference frame $[u, v]$. The space coordinates for the point (O) is $[X, Y, Z]$, and $[u, v]$ is the image-plane coordinates for the image point (I).

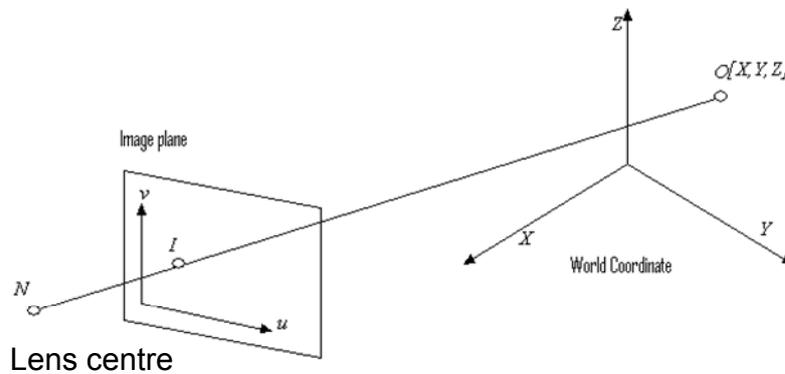


Figure A-6

Since points (O) and (N) are collinear, then it could be assumed that the coordinates for point (N) are (X_0, Y_0, Z_0) and vector A can be represented by $(X - X_0, Y - Y_0, Z - Z_0)$ Fig (7-7).

In order to make the image-plane reference frame 3-dimensional a third axis (W) was added. In this case the value of (W) for any point on the image plane is always 0, therefore the 3-dimensional position for point (I) is (u,v,0). Fig (7-7).

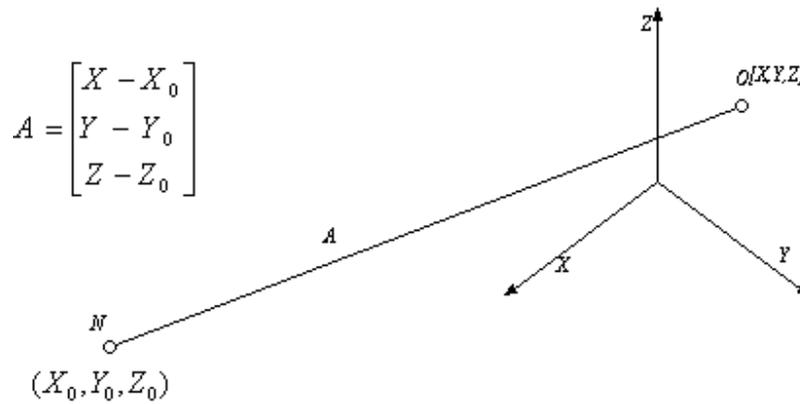


Figure A-7

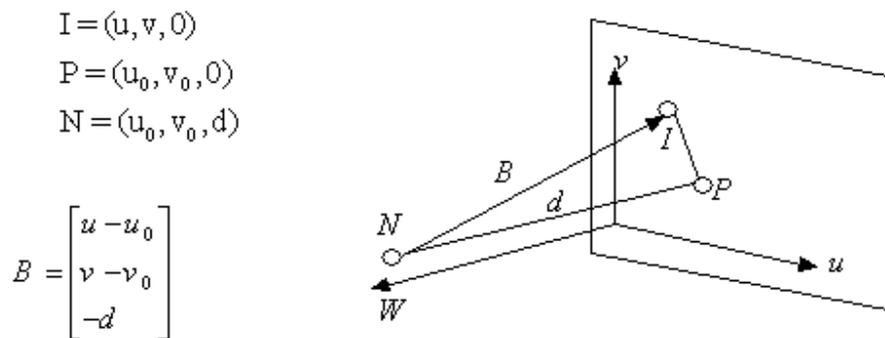


Figure A-8

In Fig (7-8) the point P (the principal point) was added to the image-plane, and a perpendicular line was drawn from N to P and parallel to the axis W . Line (NP) is called the principle axis and the principle distance d is the distance between points P and N .

Given those points O, I, N are collinear, then the vectors A in Fig. (7-7) and B in Fig (7-8) form a single straight line, and the relationship between A and B can be expressed as:

$$B = cA \quad (7-41)$$

Where c = the camera scale factor.

It is obvious that, vectors A and B were expressed in the object-space reference frame and the image-plane reference frame, respectively. In order to describe these vectors in a common reference frame, vector A is transferred to the image-plane reference frame by using the following transformation matrix.

$$T_{I/O} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

$$A^{(I)} = T_{I/O} \cdot A^{(O)} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot A^{(O)} \quad (7-42)$$

Where:

$A^{(I)}$ = vector A described in the image-plane reference frame,

$A^{(O)}$ = vector A described in the object-space reference frame, and

$T_{I/O}$ = the transformation matrix from the object-space reference frame to the image-plane reference frame.

By substituting values of A and B in equation (41) we get:

$$\begin{bmatrix} u - u_0 \\ v - v_0 \\ -d \end{bmatrix} = c \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix} \quad (7-43)$$

Or

$$u - u_0 = c[r_{11}(X - X_0) + r_{12}(Y - Y_0) + r_{13}(Z - Z_0)] \quad (7-44)$$

$$v - v_0 = c[r_{21}(X - X_0) + r_{22}(Y - Y_0) + r_{23}(Z - Z_0)] \quad (7-45)$$

$$-d = c[r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)] \quad (7-46)$$

From equation (46)

$$c = \frac{-d}{[r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)]} \quad (7-47)$$

Substitute (47) for c in (44) and (45) gives:

$$u - u_0 = -d \frac{[r_{11}(X - X_0) + r_{12}(Y - Y_0) + r_{13}(Z - Z_0)]}{[r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)]} \quad (7-48)$$

$$v - v_0 = -d \frac{[r_{21}(X - X_0) + r_{22}(Y - Y_0) + r_{23}(Z - Z_0)]}{[r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)]} \quad (7-49)$$

The image coordinates u, v, u_0 and v_0 in equations (8) and (9) are in the real-world length unit, such as mm. In reality, however, the image-plane system may use different length units, such as pixels. Therefore we need a scale factor conversion in order to use the pixel units as the measurement units.

$$u - u_0 \Rightarrow \lambda_u (u - u_0)$$

$$v - v_0 \Rightarrow \lambda_v (v - v_0)$$

$$u - u_o = \frac{-d [r_{11} (X - X_0) + r_{12} (Y - Y_0) + r_{13} (Z - Z_0)]}{\lambda_u [r_{31} (X - X_0) + r_{32} (Y - Y_0) + r_{33} (Z - Z_0)]} \quad (7-50)$$

$$v - v_o = \frac{-d [r_{21} (X - X_0) + r_{22} (Y - Y_0) + r_{23} (Z - Z_0)]}{\lambda_v [r_{31} (X - X_0) + r_{32} (Y - Y_0) + r_{33} (Z - Z_0)]} \quad (7-51)$$

Where $[u, v]$ = the unit conversion factors for the u and v axis, respectively.

Rearranging equations (47) and (88) gives the two DLT equations originally developed by Abdel-Aziz and Karara in 1971. The basic equations of the 3D DLT are:

$$u = \frac{L_1 X + L_2 Y + L_3 Z + L_4}{L_9 X + L_{10} Y + L_{11} Z + 1} \quad (7-52)$$

$$v = \frac{L_5 X + L_6 Y + L_7 Z + L_8}{L_9 X + L_{10} Y + L_{11} Z + 1} \quad (7-53)$$

Where:

$$[d_u, d_v] \equiv \left[\frac{d}{\lambda_u}, \frac{d}{\lambda_v} \right] \quad (7-54-a)$$

$$D \equiv (X_0 r_{31} + Y_0 r_{32} + Z_0 r_{33}) \quad (71-b)$$

$$L_1 = \frac{u_0 r_{31} - d_u r_{11}}{D} \quad (71-c)$$

$$L_2 = \frac{u_0 r_{32} - d_u r_{12}}{D} \quad (71-d)$$

$$L_3 = \frac{u_0 r_{33} - d_u r_{13}}{D} \quad (71-e)$$

$$L_4 = \frac{(d_u r_{11} - u_0 r_{31}) X_0 + (d_u r_{12} - u_0 r_{32}) Y_0 + (d_u r_{13} - u_0 r_{33}) Z_0}{D} \quad (71-f)$$

$$L_5 = \frac{v_0 r_{31} - d_v r_{21}}{D} \quad (71-g)$$

$$L_6 = \frac{v_0 r_{33} - d_v r_{22}}{D} \quad (71-i)$$

$$L_7 = \frac{v_0 r_{33} - d_v r_{23}}{D} \quad (71-j)$$

$$L_8 = \frac{(d_v r_{21} - v_0 r_{31})X_0 + (d_v r_{22} - v_0 r_{32})Y_0 + (d_v r_{23} - v_0 r_{33})Z_0}{D} \quad (71-k)$$

$$L_9 = \frac{r_{31}}{D} \quad (71-l)$$

$$L_{10} = \frac{r_{32}}{D} \quad (71-m)$$

$$L_{11} = \frac{r_{33}}{D} \quad (71-n)$$

Coefficients ($L_1 - L_{11}$) in equations (52, 53) are the DLT parameters which reflect the relationships between the object reference frame and the image-plane reference frame.

By rearranging the DLT equations (52, 53) the following equations can be produced for the 11 DLT parameters.

$$\begin{bmatrix} X & Y & Z & 1 & 0 & 0 & 0 & 0 & -uX & -uY & -uZ \\ 0 & 0 & 0 & 0 & X & Y & Z & 1 & -vX & -vY & -vZ \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \\ L_5 \\ L_6 \\ L_7 \\ L_8 \\ L_9 \\ L_{10} \\ L_{11} \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix} \quad (7-55)$$

With the number of calibration points larger than eleven, an overdetermined system of linear equations can be established and solved for the unknowns (L_1-L_{11}) using a least squares method.

The formula (56) below illustrates the equation (55) in case of using many (n) calibration points.

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -u_1X_1 & -u_1Y_1 & -u_1Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1X_1 & -v_1Y_1 & -v_1Z_1 \\ \vdots & \vdots \\ \vdots & \vdots \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & -u_nX_n & -u_nY_n & -u_nZ_n \\ 0 & 0 & 0 & 0 & X_n & Y_n & Z_n & 1 & -v_nX_n & -v_nY_n & -v_nZ_n \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ \vdots \\ L_{10} \\ L_{11} \end{bmatrix} = \begin{bmatrix} u_1 \\ v \\ \vdots \\ \vdots \\ u_n \\ v_n \end{bmatrix} \quad (7-56)$$

Where:

n is the number of calibration points.

The aspheric nature of the camera lenses used in the image acquisition system can cause distortion and make flat objects appear slightly curved. So the optical errors from the lens $\Delta u, \Delta v$ may be included in equations (52) and (53).

$$u - \Delta u = \frac{L_1 X + L_2 Y + L_3 Z + L_4}{L_9 X + L_{10} Y + L_{11} Z + 1} \quad (7-57)$$

(57)

$$v - \Delta v = \frac{L_5 X + L_6 Y + L_7 Z + L_8}{L_9 X + L_{10} Y + L_{11} Z + 1} \quad (7-58)$$

Where $(\Delta u, \Delta v)$ = the optical errors. Optical errors can be expressed as

$$\Delta u = \xi (L_{12} r^2 + L_{13} r^4 + L_{14} r^6) + L_{16} \eta \xi + L_{10} (r^2 + 2\xi^2) \quad (7-59)$$

$$\Delta v = \eta (L_{12} r^2 + L_{13} r^4 + L_{14} r^6) + L_{16} \eta \xi + L_{10} (r^2 + 2\xi^2) \quad (7-60)$$

Where:

ξ is the component of this deviation in the u direction

η is the component of this deviation in the v direction

$$\xi = u - u_0 \quad \eta = v - v_0$$

From equation (59) and (60) we can calculate 5 extra camera parameters L_{12} to L_{16} , however in this project the twelfth parameter is only added. As a result the equations (59), (60) were reduced to:

$$\Delta u = \xi L_{12} r^2 \quad (7-61)$$

$$\Delta v = \eta L_{12} r^2 \quad (7-62)$$

Where:

r is the deviation of the point on the image from the image centre P.

$$r^2 = \xi^2 + \eta^2$$

The values of u_0 and v_0 can be calculated after getting the first eleven camera parameters, as following:

From equation (54) we have

$$\begin{aligned}
 &(DL_1)(DL_9) + (DL_2)(DL_{10}) + (DL_3)(DL_{11}) = \\
 &(u_0 r_{31} - d_u r_{11}) r_{31} + (u_0 r_{32} - d_u r_{12}) r_{32} + (u_0 r_{33} - d_u r_{13}) r_{33} = \\
 &u_0 (r_{31}^2 + r_{32}^2 + r_{33}^2) - d_u (r_{11} r_{31} + r_{12} r_{32} + r_{13} r_{33}) = u_0
 \end{aligned} \tag{7-63}$$

Also from equation (54) we have

$$\begin{aligned}
 L_9^2 + L_{10}^2 + L_{11}^2 &= \frac{1}{D^2} [r_{31}^2 + r_{32}^2 + r_{33}^2] = \frac{1}{D^2} \\
 D^2 &= \frac{1}{L_9^2 + L_{10}^2 + L_{11}^2}
 \end{aligned} \tag{7-64}$$

From (63, 64)

$$u_0 = D^2 (L_1 L_9 + L_2 L_{10} + L_3 L_{11}) = \frac{L_1 L_9 + L_2 L_{10} + L_3 L_{11}}{L_9^2 + L_{10}^2 + L_{11}^2}$$

And similarly the value of v_0 can be found

$$\begin{aligned}
 &(DL_5)(DL_9) + (DL_6)(DL_{10}) + (DL_7)(DL_{11}) = v_0 \\
 v_0 &= D^2 (L_5 L_9 + L_6 L_{10} + L_7 L_{11}) = \frac{L_5 L_9 + L_6 L_{10} + L_7 L_{11}}{L_9^2 + L_{10}^2 + L_{11}^2}
 \end{aligned}$$

Where: $(L_1 - L_{11})$ are calculated in equation (55) above

By rearranging equations (52) above we have

$$\begin{aligned}
 L_9 uX + L_{10} uY + L_{11} uZ + u &= L_1 X + L_2 Y + L_3 Z + L_4 + \Delta u R \\
 u &= L_1 X + L_2 Y + L_3 Z + L_4 - L_9 uX - L_{10} uY - L_{11} uZ + \Delta u R
 \end{aligned} \tag{7-65}$$

Where:

$$R = L_9 X + L_{10} Y + L_{11} Z + 1$$

Similarly by rearranging equations (53) above we have

$$v = L_5 X + L_6 Y + L_7 Z + L_8 - L_9 u X - L_{10} u Y - L_{11} u Z + \Delta v R \quad (7-66)$$

Equations (65, 66) can be written in a matrix form as.

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -u_1 X_1 & -u_1 Y_1 & -u_1 Z_1 & \xi r_1^2 R_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1 X_1 & -v_1 Y_1 & -v_1 Z_1 & \eta r_1^2 R_1 \\ \vdots & \vdots \\ \vdots & \vdots \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & -u_n X_n & -u_n Y_n & -u_n Z_n & \xi r_n^2 R_n \\ 0 & 0 & 0 & 0 & X_n & Y_n & Z_n & 1 & -v_n X_n & -v_n Y_n & -v_n Z_n & \eta r_n^2 R_n \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ \vdots \\ L_{10} \\ L_{12} \end{bmatrix} = \begin{bmatrix} u_1 \\ v \\ \vdots \\ \vdots \\ u_n \\ v_n \end{bmatrix}$$

The camera parameters $(L_1 - L_{12})$ are calculated using the formula above.

A.2.1 Reconstruction of world coordinates (X, Y, Z)

From equations (52) above we have

$$\begin{aligned} \alpha L_9 X + \alpha L_{10} Y + \alpha L_{11} Z + \alpha &= L_1 X + L_2 Y + L_3 Z + L_4 \\ (\alpha L_9 - L_1) X + (\alpha L_{10} - L_2) Y + (\alpha L_{11} - L_3) Z &= L_4 - \alpha \end{aligned} \quad (7-67)$$

Where:

$$\alpha = u - \Delta u$$

Similarly From equations (53) above we have

$$(\alpha L_9 - L_5) X + (\alpha L_{10} - L_6) Y + (\alpha L_{11} - L_7) Z = L_8 - \beta \quad (7-68)$$

Where:

$$\beta = v - \Delta v$$

From equations (67, 68) and with the number of cameras at two or more, the following reconstruction matrix is used to find the world coordinates (X, Y, Z)

$$\begin{bmatrix}
 \alpha^{(1)} L_9^{(1)} - L_1^{(1)} & \alpha^{(1)} L_{10}^{(1)} - L_2^{(1)} & \alpha^{(1)} L_{11}^{(1)} - L_3^{(1)} \\
 \beta^{(1)} L_9^{(1)} - L_5^{(1)} & \beta^{(1)} L_{10}^{(1)} - L_6^{(1)} & \alpha^{(1)} L_{11}^{(1)} - L_7^{(1)} \\
 \vdots & \vdots & \vdots \\
 \vdots & \vdots & \vdots \\
 \alpha^{(m)} L_9^{(m)} - L_1^{(m)} & \alpha^{(m)} L_{10}^{(m)} - L_2^{(m)} & \alpha^{(m)} L_{11}^{(m)} - L_3^{(m)} \\
 \beta^{(m)} L_9^{(m)} - L_5^{(m)} & \beta^{(m)} L_{10}^{(m)} - L_6^{(m)} & \alpha^{(m)} L_{11}^{(m)} - L_7^{(m)}
 \end{bmatrix}
 \begin{bmatrix}
 X \\
 Y \\
 Z
 \end{bmatrix}
 =
 \begin{bmatrix}
 L_4^{(1)} - \alpha^{(1)} \\
 L_8^{(1)} - \beta^{(1)} \\
 \vdots \\
 \vdots \\
 L_4^{(m)} - \alpha^{(m)} \\
 L_8^{(m)} - \beta^{(m)}
 \end{bmatrix}
 \quad (7-69)$$

Where:

m is the number of camera used in the system

A.3 Pulnix TM-500 camera

A.3.1 the features

- Miniature size.
- High resolution.
- Shuttering to 1/10,000 sec.
- Auto-shutter option.
- External sync option.
- C and CS Mount.
- Excellentshock and vibration resistance.
- Low cost.

A.3.2 specification of Pulnix TM-500 camera

Imager	1/2" CCD
Pixel	500(H) x 582(V)
Cell size	12.7μm(H) x 8.3μm(V)
Scanning	625 lines, 50 Hz CCIR
Lens mount	C/CS mount
Min. illumination	0.2 lux @ F 1.2
Power required	, 12V DC < 2.4 W
Operating temperature	-10°C to 50°
Size (W X H X L)	45mm x 39mm x 92mm
Weight	235 grams

A.4 DT3155 board

A.4.1 Key features of the board

- Operates on the PCI local bus interface;
- Digitizes 8-bit monochrome video from any one of four 60 Hz or 50 Hz video input channels;
- Synchronizes to any of the video inputs;
- Accepts an external trigger with selectable polarity;
- Provides programmable black and white levels;
- Provides a 256 x 8-bit input look-up table (ILUT);
- Provides a 256 x 8-bit passthru look-up table;
- Provides passthru scaling to 1/4 of the frame size; and
- Provides eight TTL-level digital output signals for general-purpose use. (DT3155 data sheet)
- compatible with Image Pro Plus and MATLAB software



Figure A-9 : The DT3155, monochrome frame grabber for the PCI Bus

Video Input Channels

The DT3155 supports monochrome video input from one of four software-selectable video channels (0 to 3). The channel is software selectable. By default, channel 0 is selected.

In this card there are two video connectors, J1 and J2

Video Input Connector J1

Connector J1 is a 15-pin, male, D-shell connector that accepts all the signals brought out by the DT3155 board through the user-designed cable or EP306 cable. The cable has been designed for this project. Fig (7.10) shows the pin locations for connector J1.

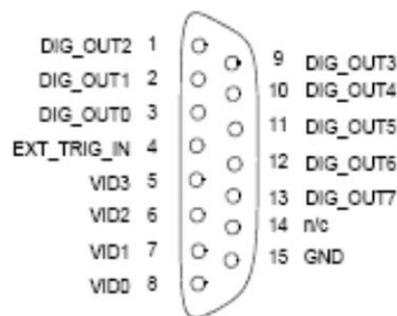


Figure A-10 : Connector J1

Video Coaxial Connector J2

Connector J2 is a female coaxial connector that connects to the video output of the video source using a 75 Ω coaxial cable with a male connector. The single-use BNC input connector, J2, is shared with the VID0 signal (pin 8) on video input connector J1. So both signals should not be matched, otherwise, the two video sources will be shorted together, and could damage the video sources. Fig (7.11) illustrates coaxial connector J2 and Fig (7-12) shows how the cameras connected to the system.

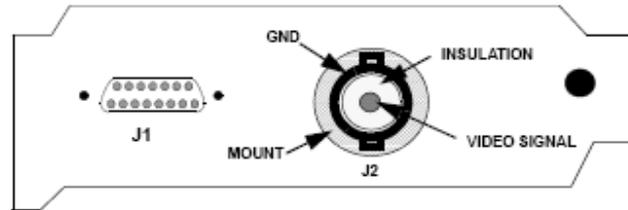


Figure A-11: coaxial connector J2

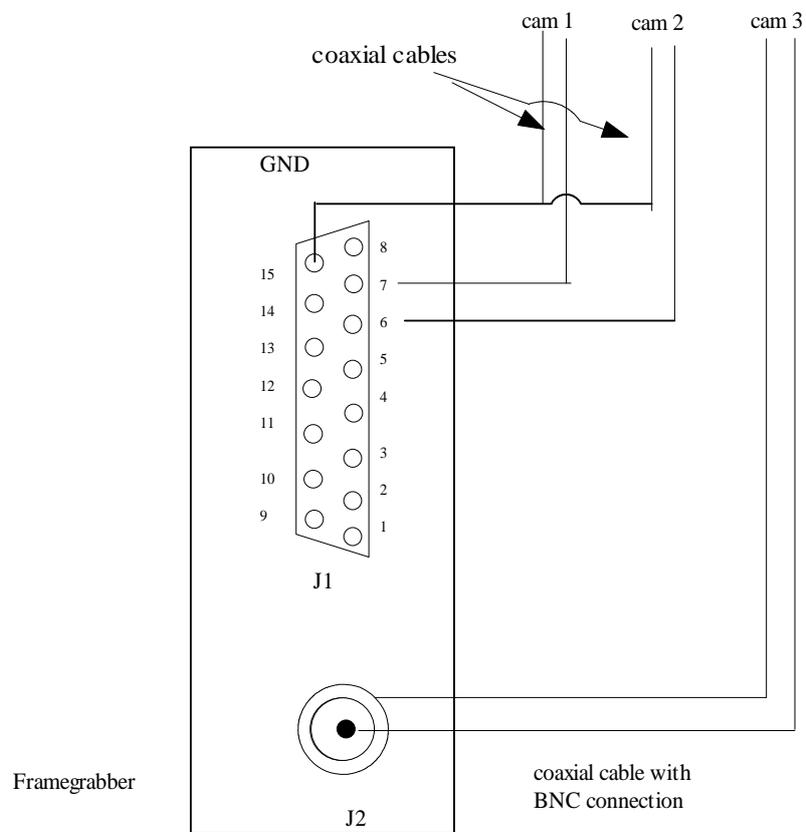


Figure A-12: connecting 3 cameras to the frame grabber

A.5 Dimension of the pyramid calibration piece

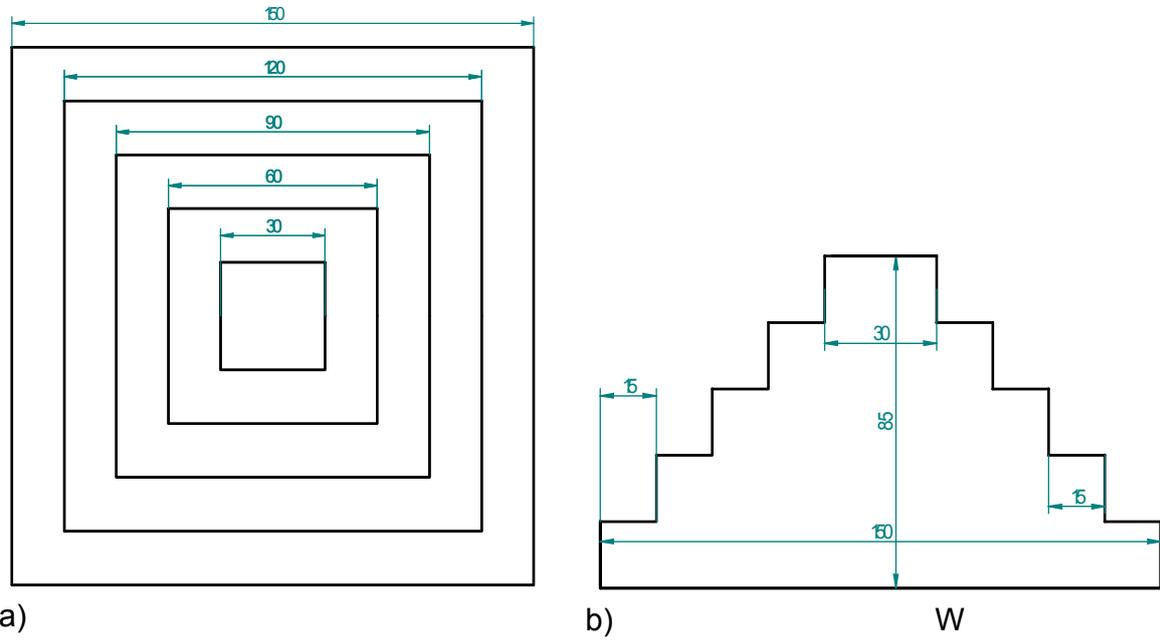


Figure A-13: Pyramid calibration piece a) top view, b) side view (dimension in mm)

Appendix B

```
% function returns the angles for all the fingers
function H = display_the_fingers_angles

clear
% create an empty matrix in order to fill it by the angles for the
%4 fingers.
H=[];
for I = 1 : 4
%-----
----
%if the No of points = 8 then sort the matrix in Ascending
%order based on the X coordinate (the second column).
%the function (points_from_all_cameras) returns 3 matrices, these
%matrices are, points_from_Right_camera,points_from_Inner_camera and
%points_from_Left_camera, each represents the points on the fingers
%in order that match the equivalent points in the other cameras.
%The size of each matrix is 32 rows by 3 columns

[points_from_Right_camera,points_from_Inner_camera,points_from_Left_cam
era] = points_from_all_cameras;

R = size (points_from_Right_camera,1); % no of rows in the matrix.
%points_from_Right_camera.
C = size (points_from_Inner_camera,1); %no of rows in the matrix.
%points_from_Inner_camera.
P = size (points_from_Left_camera,1); %no of rows in the matrix.
%points_from_Left_camera.
%-----
----
%the formula below (1+((R-24)*(I-1)) Produces Numbers 1,9,17,25 for the
%for loop I =1 ,2,3 ,4 respectively,% so we can choose 8 rows from the
%matrix each time and the formula((R-24)*I)Produces Numbers 8 ,16,24,
%32 for I=1 ,2,3 ,4 respectively. so the first for loop chooses the
%matrix (1:8,:),and the second (9:16,:), the third (17:24,:)and the
%last for loop chooses the matrix (25:32,:),so we can choose 8 rows
%from the 32 rows matrix each time, and each matrix represents the
%points for each finger

points_from_Right_camera = points_from_Right_camera(((1+((R-24)*(I-
1))):((R-24)*I)),:);
points_from_Inner_camera = points_from_Inner_camera (((1+((C-24)*(I-
1))):((C-24)*I)),:);
points_from_Left_camera = points_from_Left_camera (((1+((P-24)*(I-
1))):((P-24)*I)),:);

%-----
% function (find_fingers_angles) gives the angles between the finger
% joints.
g =
find_fingers_angles(points_from_Right_camera,points_from_Inner_camera,p
oints_from_Left_camera);
```

```

H =[H;g];
%-----

end% end for (for I = 1 : 4 )

disp('The angles for the fingers are ');
disp(H);

% *****
% *****

% this function is used for ordering the image that captured by inner
% camera
function [finger_1,finger_2,finger_3,finger_4]=
order_points_inner_camera (ICRim)

%reorder the column of the matrix
finger_markers = ICRim(:,[2 1 3 4]);

% delete column No 1, The remaining columns are the points No ,
%X position and Y position

finger_markers = finger_markers(:,2:4);

% this function (new_matrix_min_y)finds out the min value of Y
% coordinats which will be used as a first point in the matrix for
% the image that has been captured by the inner camera
finger_markers = new_matrix_min_y (finger_markers);

%this for loop represents the 4 fingers of the hand.
for I = 1 : 4

% for the 4th finger we do not need to go through the all
% calculation,because we have only the 8 points for the fourth
% finger, so we need only to sort them.
if I < 4

% creat an empty matrix in order to fill it with the out put
A=[];

m = finger_markers;

%7 times for loop, because we want to find 7 points for a
%certain finger based on the first point, i.e. the total No
%of points for each finger are 8
for k = 1 : 7

[N,M] = size(m);

```

```

%if number of rows in matrix m <= 2 , the next point is the
%last point in the matrix
if N >2

%function call: this function (dist_between_points_inner_cam)
%calculates the distance between the first point and all
    %other points
    c = dist_between_points_inner_cam (N,m);

% sortrows is a MATLAB function used to sort the rows of
% the matrix (in this case "c") in ascending order based on
% a specified columns(in this case first column which has the
% distances between the first point and all other points).
    sort_c = sortrows(c,[1]);

    [H,W] = size (sort_c);

%call the function that calculate the min distance(3or rows)
    x = min_distance_inner_cam (H,m,sort_c);

% find the size of the matrix (x)
    [dd,hh]= size (x);

%if No of rows in matrix (x) >=2,then sort out the
%matrix in asending order based on the first row
% which has min delta x, and select the first two
    % rows
if dd >= 2
%-----
%this function used in case of No of rows in the
%matrix x =3 (matrix x consists of 3 columns the
%1st is Delta_x,2nd is Point No, 3rd is X
% coordinates and the 4th is y coordinates). and
% it tests if the 3 points located on the same
%finger or not, in case of the 3 points located
% on the same finger this function choose the two
% points that closer to our reference point, by
%comparing the X coordinates for the 3 points.
%if the deference in X coordinates <= 5 pixels
%that means all the 3 points belong to the same
    %finger, otherwise not.
if dd == 3
        x = select_two_points_inner_camera(dd,x);

end%end for (if dd == 3)

%-----
% sort matrix y in asending order based on the
%firs column
        sort_X = sortrows(x,[1]);

```

```

% select the two points which have min delta x,
%ie the first and second rows; because some
%times the point with min delta x is not the
%next point, so the two points are selected ,
min_2delta_x = sort_X (1:2,:);

% this function (find_next_point_x) finds the next
% point in case of inner camera.
next_point = find_next_point_x (min_2delta_x);

%else for (if dd >= 2). if No of rows in the
% matrix (x)= 1 then this is the next point
else

        next_point = x;
end%end for (if dd >= 2)

% remove the first column which contains the min
%distance, so the matrix has the point No, X
% coordinates and Y coordinates
        next_point = next_point (:,2:4);

% remove the row that contains the second point
% from the matrix c.
        c = c(all(c~= next_point(1,1),2),:);

% remove the first column from the matrix i.e the
% column that has the distance between the points
% in order to continue order to continue finding
%the next point
        c = c(:,2:4);

%Append the row that contains the next point to
%the matrix c
        c = [next_point;c];
        m = c;

else% else for if N >2

        next_point = m(2,:);

end% end for if N >2

        A =[A;next_point];

end% end for loop (k=1:7)

```

```

% add the first row i.e the row that contains the
% first point to the matrix to produce the eight
% points which represent the finger
    A = [finger_markers(1,:);A] ;

% this function (remove_points) remove the 8 points of the
%finger from the whole matrix,
    D = remove_points (finger_markers,A);

%this function (new_matrix_Y)finds the first point of the next
%finger and returns a new matrix,the first row has the data for
%the first point for the next finger.
    finger_markers = new_matrix_Y (D);

else%else for if I < 4

% if the No of points = 8 then sort the matrix in Ascending
% order based on the Y %coordinate (the third column )
    finger_markers = sortrows(finger_markers,[3]);
    A = finger_markers;
end%end for if I < 4

if I == 1
    finger_A = A; % 8 points for the first finger.
elseif I == 2
    finger_B = A ; % 8 points for the second finger.
elseif I == 3
    finger_C = A; % 8 points for the third finger.
elseif I == 4
    finger_D = A ; % 8 points for the fourth finger.
end

end% end for I = 1 : 4

% this function (rank_fingers_x_coord)sort out%the fingers

[finger_1,finger_2,finger_3,finger_4] = rank_fingers_x_coord
(finger_A,finger_B,finger_C,finger_D);

% *****
% *****

%Function to calculate angles between joints given points in space
function angles = angles(ThreeD)

numang = size(ThreeD,1);
numang = (numang/2)-1;

```

```

for i = 1:numang+1;

    v(i,:) = ThreeD(2*i,:) - ThreeD(2*i-1,:);

end

for i = 1:numang;

    ang(i) =
acos(sum(v(i,:).*v(i+1,:))/(sqrt(sum(v(i,:).^2))*sqrt(sum(v(i+1,:).^2))
));

end

angles = (180/pi)*ang;

% *****
% *****
% this function returns the angles between the joints for the 4
fingers.
function angle = find_fingers_angles
(points_from_Right_amera,points_from_Inner_amera,points_from_Left_amera
);

cam1m(:,1) = points_from_Inner_amera(:,2);
cam1m(:,2) = points_from_Inner_amera(:,3);
cam2m(:,1) = points_from_Left_amera(:,2);
cam2m(:,2) = points_from_Left_amera(:,3);
cam3m(:,1) = points_from_Right_amera(:,2);
cam3m(:,2) = points_from_Right_amera(:,3);

TwoD1(:,:,1) = cam1m;
TwoD1(:,:,2) = cam2m;
TwoD1(:,:,3) = cam3m;

    load dozparams.dat
    pars1(:,1) = dozparams(:,1);
    pars1(:,2) = dozparams(:,2);
    pars1(:,3) = dozparams(:,3);

    format short
ThreeD1 = recon(pars1,TwoD1);

angle = angles(ThreeD1);

% *****
% *****
%this function calculates the X and Y coordinate for the 3 images of

```

```

%the calibration piece that captures by Right, Left and the Inner
%cameras. Each image contains 27 points represent the calibration
%points.

function []= calib_piece_images_processing

%Change directory to access the files under investigation.

cd u:\6_pins_calib_piece\1st_test_new_calb_piece

%this for loop determines which image will be processed
for k=1:3

if (k==1)

%imread is a matlab function to read image from graphics file

I=imread('u:\6_pins_calib_piece\1st_test_new_calb_piece\Right_camera_calib_piece.tiff');

elseif (k==2)

    I = imread
('u:\6_pins_calib_piece\1st_test_new_calb_piece\Left_Camera_calib_piece
.tiff'); % to read image from the file
elseif (k==3)
    I =
imread('u:\6_pins_calib_piece\1st_test_new_calb_piece\Inner_Camera_calib_piece.tiff'); % to read image from the file
end

%this function (foreground_image)put all back ground pixels into
%zero and the pixels in foreground remain as their original
%value.

    I = foreground_image(I);

%[labeled,numObjects]= bwlabel(I,n) returns a matrix labeled, of
%the same size as I, containing labels for the connected objects
%in imageI. n can have a value of either 4 or 8,where 4
%specifies 4-connected objects and 8 specifies 8-connected
%objects. and numObjects has number of connected objects found in
%the image I. Label components.

    [labeled,numObjects]= bwlabel(I,8);

    t= numObjects;

    figure; imshow(labeled);
    pixval on

%regionprops is a MATLAB function Measure properties of matrix
%labeled, if properties is the 'basic', then these measurements

```

```

%are : 'Area', 'Centroid', and 'BoundingBox'.Only Area and
% Centroid will be calculated

data = regionprops(labeled,'basic');

% No of calibration points in the image is 27
for j=1:27

%Object_No is No of control points for the calibration piece
Object_No(j)=(j);

%Calculate the area of the points in the imageI. we do no
%need the area just for checking
Area(j) = data(j).Area;

%Calculate the centroid of the control points in X direction
Centre_X(j)=data(j).Centroid(1);
%Calculate the centroid of the control points in Y direction
Centre_Y(j)=data(j).Centroid(2);
end% end for j=1:32

if (k==1)
m =[Object_No',Area',Centre_X',Centre_Y'];

%save the data(Object_No',Area',%Centre_X',Centre_Y') to cnt file
%or to any data file such as xls.

save('u:\6_pins_calib_piece\1st_test_new_calb_piece\Right_camera_calib_
piece.cnt','m','-ascii')
elseif (k==2)
m =[Object_No',Area',Centre_X',Centre_Y'];

save('u:\6_pins_calib_piece\1st_test_new_calb_piece\Left_Camera_calib_p
iece.cnt','m','-ascii')
elseif (k==3)
m =[Object_No',Area',Centre_X',Centre_Y'];

save('u:\6_pins_calib_piece\1st_test_new_calb_piece\Inner_Camera_calib_
piece.cnt','m','-ascii')
end%end if (k==1)

end% end for k=1:3

% *****
% *****

%this function threshold the input image I by using the thresholding
%value 3*the standard deviation of the image I + the mean of the

```

```

%image (I) and return back the binary image.

function A = thresholdvalue(I)

threshold = (3*(Std2(I))+ mean2(I))*1.1; % threshold value

A=(I>=threshold); % A is alogical array

% *****
% *****

% The input of this function (find_next_point_y)is the matrix that
% contains the min 2 delta y. By Finding the deference between the
% delta y. i.e. The deference between the first and second rows of
% the first column. If the deference <= 5 pixels, then the next point
%will be the point which has min X coordinate, otherwise the next point
%is the point which has min delta y. this function is used in case of
%Left and Right Cameras.
function next_point = find_next_point_y (min_2delta_y)

% Find the deference between the delta y, i.e. the deference between
%the first and second rows of the first column.

A = abs(min_2delta_y(1,1)-min_2delta_y(2,1));
%-----
if A <= 5

% this sort out the matrix min_2delta_y, which has two rows
% represent the min delta_y, in ascending order based on
% X_coordinta (the third column),then the point that has
% min X_coordinate is the next point.

min_X_coordinat = sortrows (min_2delta_y,[3]);

next_point = min_X_coordinat;

else

    next_point = min_2delta_y (1,:);

end

```