Loughborough University
Department of Mechanical Engineering

INTRAOPERATIVE REGISTRATION FOR ROBOTIC-ASSISTED ORTHOPAEDIC SURGERY:

A DIGITAL X-RAY PHOTOGRAMMETRY BASED TECHNIQUE

By

Ian Browbank, B.Eng. M.Sc.

A Doctoral Thesis

Submitted in Partial Fulfilment of the Requirements for the Reward of
Doctor of Philosophy of Loughborough University.

May 1998

ABSTRACT

Changes in orthopaedic practice have led to an increased reliance upon fluoroscopic image-guidance during fracture fixation (osteosynthesis) procedures. The resulting complexity of surgical techniques, and concerns over X-ray radiation exposure levels to orthopaedic surgeons, has prompted an investigation of *robotic-assisted orthopaedic surgery*, with a view to improving the precision, repeatability and radiation safety of existing fluoroscopically-guided bone drilling procedures.

A prerequisite for such an application is to establish the patient's location with respect to the robotic manipulator. Owing to the inherent nature of osteosynthesis procedures, this *intraoperative registration* process can only be performed by quantifying the existing fluoroscopic examination process. A *digital X-ray photogrammetry* based registration technique, which uses a radiolucent robot-mounted X-ray calibration frame, has therefore been investigated.

When this calibration frame is simultaneously imaged with the patient, discrete registration markers, corresponding to radiopaque fiducials embedded in the frame, are superimposed into the standard intraoperative X-ray images. Digitising these images, using a PC-based frame grabber card, has allowed semi-automatic image analysis routines to be implemented. Applying correction-calibration software, which provides on-line compensation for image distortion effects, then allows the imaged part of the patient's skeleton to be located, with respect to the calibration frame's coordinate system, thus establishing intraoperative registration.

In vitro laboratory-based trials of this registration technique indicate that reconstruction errors are in the sub-millimetre range. As such, this new approach represents a low cost non-invasive registration option, which fully adheres to "operating room compatibility" criteria, and is applicable to a wide range of osteosynthesis procedures. Significantly, by extending the technique to include quantification of the trajectory planning process, it has also been possible to demonstrate major improvements over existing surgical techniques.

TABLE OF CONTENTS

Page

Statement of Originality i
Abstract ii
Table of Contents iii
List of Figures vi
Acknowledgements ix

1. Introduction.

1.1 FIELD OF INTEREST 1
1.2 ORTHOPAEDIC SURGERY 2
1.2.1 Osteosynthesis 4
1.2.2 Radiation Exposure to the Orthopaedic Surgeon 6
1.3 ROBOTIC-ASSISTED ORTHOPAEDIC SURGERY 8
1.3.1 The MEDROSA Research Project 12
1.4 OUTLINE OF THESIS 15
1.4.1 Definition of Research Problem 16
1.4.2 Proposed Research Methodology 18
1.5 THESIS STRUCTURE 20
1.6 SUMMARY OF CHAPTER 1 22

2. Photogrammetry.

2.1 INTRODUCTION 24
2.2 CAMERA-BASED PHOTOGRAMMETRY 26
2.2.1 The Collinearity Condition 28
2.2.2 Stereo-Photogrammetry 30
2.2.3 Implicit Photogrammetric Reconstruction 33
2.2.3.1 Mathematical Modelling of the Image Formation Process 34
2.2.3.2 Minimising Reconstruction Errors 39
2.3 X-RAY PHOTOGRAMMETRY 45
2.3.1 Radiography Equipment 47
2.3.2 Roentgen Stereophotogrammetry 49
2.3.2.1 Roentgen Stereophotogrammetric Analysis 52
2.3.2.2 Subsequent RSA Research 58
2.4 DIGITAL X-RAY PHOTOGRAMMETRY 60
2.4.1 Fluoroscopy Equipment 62
2.4.2 Distortion Correction of X-Ray Image Intensifier Images 67
2.4.2.1 Global Polynomial Correction Functions 71
2.4.2.2 Local Affine Correction Functions 73
2.4.2.3 The Mathematical Modelling Approach 75
2.4.3 Geomagnetic Distortion 76
2.4.3.1 Geomagnetic Distortion Correction 79

2.4.4 Quantitative Three-Dimensional Angiography 80
2.4.4.1 Biplane Coronary Angiography 81
2.4.4.2 Multi-View Methods 83
2.5 SUMMARY OF CHAPTER 2 84

3. Intraoperative Registration.

3.1 OVERVIEW OF THE REGISTRATION PROBLEM 87
3.2 INTRAOPERATIVE SENSOR TECHNOLOGY 91
3.3 ORTHOPAEDIC APPLICATIONS 99
3.3.1 Arthroplasty 102
3.3.2 Spinal Fusion Surgery 109
3.3.3 Osteotomies of the Hip 113
3.3.4 ACL Reconstruction 116
3.4 INTRAOPERATIVE QUANTITATIVE X-RAY IMAGING 117
3.4.1 Minimal Access Surgery Applications 118
3.4.2 Femoral Osteosynthesis Applications 126
3.5 SUMMARY OF CHAPTER 3 134

4. Camera-Based Feasibility Study.

4.1 INTRODUCTION 138
4.2 INTRAOPERATIVE REGISTRATION SCHEME 140
4.3 PRELIMINARY DLT INVESTIGATION 145
4.3.1 DLT Camera Calibration Mathematics 145
4.3.2 Design of a Suitable Calibration Object 149
4.3.3 Software Development 152
4.4 INVESTIGATION OF THE TWO-PLANE CALIBRATION METHOD 156
4.4.1 Two-Plane Camera Calibration Mathematics 156
4.4.2 Calibration Frame Design 162
4.4.3 Software Development 170
4.4.3.1 Preliminary Two-Plane Investigation 172
4.4.3.2 Robot-Compatible Study 174
4.5 SUMMARY OF CHAPTER 4 188

5. X-Ray Based Feasibility Study.

5.1 INTRODUCTION 190
5.2 INITIAL DEMONSTRATION OF FEASIBILITY 191
5.3 PRELIMINARY SOFTWARE DEVELOPMENT 198
5.4 SECOND GENERATION X-RAY COMPATIBLE SOFTWARE 205
5.4.1 Image Analysis Software 206
5.4.2 Investigation of the Image Distortion Problem 213
5.5 SUMMARY OF CHAPTER 5 219

6. Intramedullary Nailing of Femoral Shaft Fractures.

6.1 FEMORAL ANATOMY 221
6.2 FEMORAL SHAFT FRACTURES 224
6.3 LOCKED INTRAMEDULLARY NAILING 226
6.4 CLOSED INTRAMEDULLARY NAILING TECHNIQUE 230
6.4.1 The Distal Locking Problem 236
6.5 ROBOTIC-ASSISTED DISTAL LOCKING 249
6.5.1 "Perfect Circles" Based Approach 250
6.5.2 Enhanced Trajectory Planning 253
6.6 SUMMARY OF CHAPTER 6 259

7. Internal Fixation of Proximal Femoral Fractures.

7.1 PROXIMAL FEMORAL FRACTURES 260
7.2 TREATMENT OF PROXIMAL FEMORAL FRACTURES 265
7.3 INSERTION OF THE SLIDING COMPRESSION HIP SCREW 268
7.3.1 Guide Wire Insertion Problems 273
7.4 ROBOTIC-ASSISTED GUIDE WIRE INSERTIONS 275
7.4.1 Initial L-Frame Investigation 279
7.4.2 Follow-Up L-Frame Investigation 284
7.5 SUMMARY OF CHAPTER 7 290

8. Conclusions.

8.1 INTRODUCTION 291
8.2 EVALUATION OF THE X-RAY PHOTOGRAMMETRY APPROACH 292
8.2.1 The Image Analysis Problem 293
8.2.2 The Image Distortion Problem 295
8.2.3 The Intraoperative Registration Problem 298
8.2.4 Trajectory Planning 300
8.2.5 Compatibility with the Operating Room Environment 302
8.3 RECOMMENDATIONS 305

9. References. 308

Appendix 1 Ionising Radiation Exposure To The Orthopaedic Surgeon A1
Appendix 2 The MEDROSA Project A31
Appendix 3 Intraoperative Registration for Neurosurgery Applications A56
Appendix 4 Registration Graphs A80
Appendix 5 Solution of the Linear Least Squares Problem A86

LIST OF FIGURES

		Page
1.1	Robotic-Assisted Orthopaedic Surgery	18
2.1	The Collinearity Condition	29
2.2	Model of the Stereo Imaging Process	30
2.3	The Epipolar Constraint	32
2.4	The Pinhole Camera Model	35
2.5	The Two-Plane Model	37
2.6	Model of the X-ray Image Formation Process	46
2.7	The Major Components of a Radiography System	47
2.8	Biplane X-Ray Photogrammetry	51
2.9	Stereo X-Ray Photogrammetry	52
2.10	Reconstruction of X-ray Source Location	57
2.11	X-Ray Photogrammetric Reconstruction	57
2.12	Components of a Modern Image Intensifier Tube	63
2.13	Schematic Representation of the Operation of an Image Intensifier Tube	64
2.14	Schematic of a Fluoroscopy System	65
2.15	Mobile X-ray Image Intensifier (Siemens Siremobil 2)	66
2.16	Pincushion Distortion Caused by an X-Ray Image Intensifier	68
2.17	X-Ray Image Intensifier with Fiber-Optic Entrance Window	69
2.18	Superimposed True and Distorted Control Point Cells	74
2.19	Geomagnetic Distortion of the Image Intensifier Image	78
3.1	Co-Registration of Surgical Data Sources	88
3.2	A Typical Anthropomorphic Digitising Arm Design	93
3.3	The Infra-Red Optical Digitiser	95
3.4	A Real-Time Intraoperative Sonic Localisation System	97
3.5	Robotic- and Computer-Assisted Orthopaedic Surgery	100
3.6	ROBODOC System Procedural Flow	104
3.7	Registration Graph for the ISS ROBODOC System	104
3.8	Total Knee Replacement: Resection Sequence for the Femoral Component	106
3.9	Ideal Trajectories for Transpedicular Screw Insertion	109
3.10	Optotrak-Based Computer-Assisted Spine Surgery	111
3.11	The Individual Templates Registration Technique	115
3.12	Intraoperative Registration for Minimally Invasive Renal Procedures	120
3.13	Registration Graph for the Grenoble Robotic Spine Surgery System	123
3.14	Schematic of the OrthoSista™ Robot	127
3.15	The CAOS System Architecture	130
3.16	Registration Graph for the CAOS System	132
4.1	The MEDROSA System's Intraoperative Registration Scheme	141
4.2	Kinematic Configuration & Kinematic Model of the MEDROSA Manipulator	142
4.3	Registration Graph For The MEDROSA System	144
4.4	The DLT Camera Calibration Equations Expressed in Matrix Form [A].x = b	147
4.5	The Calibration 'Cube' Used During Initial Camera Trials	150
4.6	Perspective Distorted Views as Observed by the Two Cameras	151
4.7	The 'Steps' Calibration Object	151
4.8	'Steps' Calibration Object Shown in One of its Several Possible Configurations Within the Calibration Cube	152
4.9	Subdivision of Calibration Plane into Triangular Regions	158
4.10	3-D Reconstruction from the Intersection of Two Lines-of-Sight	159
4.11	Convergent Calibration Frame	164
4.12	Biplane Calibration Frame	164
4.13	The Robot-Mounted Calibration Frame	165
4.14	Schematic of the Robot-Mounted Calibration Frame	165
4.15	Ideal Superimposed Fiducial Pattern	167

4.16	Effect of Angular Orientation Upon Fiducial Measurements	170
4.17	Bar Chart Showing Distinct Fiducial Measurement Ranges	171
4.18	A Typical Histogram for a 'Black' Fiducial on a 'White' Background	177
4.19	Calculation of Predicted Fiducial Marker Locations	178
4.20	The Plan View 'Ball-Bearing' Plane Fiducial Numbering System	179
4.21	Calculation of 'Cross' Plane Grid Centre and Grid Spacing	180
4.22	Estimation of the Side View Camera's Effective Focal Point	182
4.23	Distance From a Point to a 3-D Line	183
4.24	Overlaying the Reconstructed Plan View Line-of-Sight in the Side View Image	187
5.1	Prototype Biplane X-Ray Calibration Frame For Distal Interlocking	192
5.2	Frame Positioning For Biplane Fluoroscopic Examination	192
5.3	Prototype Convergent X-Ray Calibration Frame For Hip Fracturs	194
5.4	Frame Positioning for Convergent Fluoroscopic Examination	194
5.5	Typical X-ray Image of The Biplane Frame/Femur Combination	195
5.6	Mechanical Constraints Upon Arciform C-arm Movement	196
5.7	The Unconstrained C-arm Imaging Geometry	197
5.8	The "Hip Box" Biplane Calibration Frame	195
5.9	Ideal Hip Box X-ray Calibration Image	200
5.10	Automatic Determination of Range and Search Limit Values	201
5.11	Perspex Test Object Inside Hip Box Calibration Frame	203
5.12	Predicted Gray-Level Distribution for an Inverted X-Ray Image of a Sphere	207
5.13	Fiducial Neighbourhood Scanning Order	212
5.14	Image Distortion Calibration Plate	215
5.15	A Typical PA/Lateral Subtraction Image	215
6.1	Anatomy of the Femur	222
6.2	Simple Femoral Shaft Fractures	225
6.3	Complex Femoral Shaft Fractures	225
6.4	The AO/ASIF Universal Femoral Nail	229
6.5	Locked Intramedullary Nailing of a Femoral Shaft Fracture	231
6.6	Patient in Lateral Decubitus Position on Fracture Table	234
6.7	Patient in Supine Position on Fracture Table	234
6.8	The "Perfect Circles" Fluoroscopic Image	238
6.9	Nail-Mounted Distal Targeting Devices	240
6.10	The AO Distal Aiming Device	241
6.11	Distal Interlocking Technique Using the AO Distal Aiming Device	241
6.12	The Grosse-Kempf Fluoroscope-Mounted Targeting Device	242
6.13	The Goodall Laser Guidance System for Distal Interlocking	243
6.14	The Freehand Aiming Method for Distal Interlocking	245
6.15	Direct Exposure of the Surgeon's Hand to X-ray Radiation	238
6.16	Removal of Surgeon's Hands from the Primary X-ray Beam	247
6.17	Fluoroscopic Verification of Transfixion Screw Insertion	248
6.18	The Trajectory Planning Scenario (Assuming Perfect Circles)	251
6.19	Ideal Calibration Images For Distal Interlocking	252
6.20	Reconstruction of The Nail Hole's Centre	254
6.21	Rotation About The Long Axis of the Nail	255
6.22	Test Rig Used To Simulate Nail Rotation	248
6.23	Graphs of Characteristic Hole Measurement Vs. Rotation Angle	256
6.24	The Angle-Based "Characteristic" Measurement	258
7.1	Proximal Femoral Fracture Types	261
7.2	Distribution of Femoral Neck Fracture Cases by Age & Sex (England 1985)	262
7.3	Diagrammatic Representation of Changes in Bone Mass with Ageing	264
7.4	The AO Dynamic Hip Screw	267
7.5	The AO DHS Components	267
7.6	C-arm Position for Fractures of the Proximal Femur	269
7.7	Surgical Insertion of the Sliding Compression Hip Screw	271

7.8 Fixation Failure as a Function of Lag Screw Position 274
7.9 The C-Frame Based Registration Scenario 277
7.10 "L-Frame: Positioning in Relation to Patient Anatomy 278
7.11 The "L-Frame" X-ray Calibration Frame 280
7.12 The Three L-Frame Calibration Plate Configurations 281
7.13 Perspex Test Object Used in Conjunction With The L-Frame 280
7.14 The Three Test Object Locations in Relation to the L-Frame 282
7.15 The Original L-Frame Reconstruction Scenario 285
7.16 The Influence of Plate Separation Upon Measurement Errors 286
7.17 Computer-Assisted Planning of the Guide-Wire Trajectory 287

8.1 X-Ray Photogrammetry Based Intraoperative Registration 303

LIST OF TABLES

Page

2.1 Reconstruction Errors Reported by Champleboux *et al.* 38
2.2 Details of Selected Biomechanical Applications of Photogrammetry 41
2.3 Details of Selected 3-D Machine Vision Applications 41
2.4 The Effect of Control Point Numbers on DLT Reconstruction Errors 42
2.5 Calibration Accuracy Using the Two-Plane Method 43

4.1 Typical Control Point Data for Software Evaluation 154
4.2 Typical Reconstructed Control Point Data (DLT) 154
4.3 Typical Test Point Data for Software Evaluation 155
4.4 Typical Reconstructed Test Point Data (DLT) 155
4.5 Typical Reconstructed Control Point Data (Global Interpolation) 173
4.6 Typical Reconstructed Test Point Data (Global Interpolation) 173
4.7 Typical Reconstructed Test Point Data (Local Interpolation) 174
4.8 Typical Reconstructed Test Point Data (Global Interpolation) 188

5.1 Typical Global Reconstruction Data (Upper Test Object Location) 204
5.2 Typical Global Reconstruction Data (Middle Test Object Location) 204
5.3 Typical Global Reconstruction Data (Lower Test Object Location) 204
5.4 Typical Localised Reconstruction Data (Upper Test Object Location) 218
5.5 Typical Localised Reconstruction Data (Middle Test Object Location) 218
5.6 Typical Localised Reconstruction Data (Lower Test Object Location) 218
5.7 Average "Mismatch" Values 217

7.1 Typical Follow-Up Data From Proximal Femoral Fracture Studies 263
7.2 Typical Global Reconstruction Data (L25M Configuration) 283
7.3 Typical Global Reconstruction Data (L50M Configuration) 283
7.4 Typical Global Reconstruction Data (L100M Configuration) 283
7.5 Typical Localised Reconstruction Data (L25M Configuration) 289
7.6 Typical Localised Reconstruction Data (L50M Configuration) 289
7.7 Typical Localised Reconstruction Data (L100M Configuration) 289
7.8 Average "Mismatch" Values 289

Acknowledgements.

The MEDROSA Project is supported by the Wishbone Trust of the British Orthopaedic Association. The author is grateful for their financial support.

I would like to take this opportunity to acknowledge the contribution that has been made to this thesis by the originators of the MEDROSA Project:

Prof. J.R. Hewit (Department of APEME, University of Dundee).
Mr. R.G. Shedden (consultant orthopaedic surgeon, Mayday University Hospital).
Dr. K. Bouazza-Marouf (Dept. Mech. Eng., Loughborough University).

In particular, the support provided by my research supervisor, Dr. K. Bouazza-Marouf, has gone far beyond the normal requirements of this role, and is much appreciated.

I would also like to thank the members of the academic and technical staff, of the Department of Mechanical Engineering at Loughborough University, who have contributed to the research documented in this thesis.

The following individuals and their hospitals have also been very generous with their time and facilities:

Dr. Carlos Heras-Palou (Glenfield General Hospital, Leicester).
Dr. Stewart Calder (Leicester Royal Infirmary).
Dr. R.J. Minns (Dryburn Hospital, Durham).

However, most of all, I wish to thank my sister, parents, and grandparents for their unconditional love and support.

1. INTRODUCTION.

1.1 Field of Interest.

The aim of this multi-disciplinary thesis is to develop engineering-based solutions to problems which are currently experienced during certain orthopaedic surgery procedures. Significantly, this research was instigated at the request of a consultant orthopaedic surgeon, with a view to finding out if the introduction of new technologies into the operating room could lead to enhanced surgical performance. Owing to the inherent nature of the orthopaedic surgery procedures under investigation, a vision-guided robotic system has therefore been proposed as a potential solution. The current thesis is specifically concerned with the vision-related aspects of this robotic application.

When the working environment of a robotic manipulator is unstructured, some form of sensory input must be provided to guide the execution of the manipulator's task. In the case of the proposed orthopaedic surgery application, the manipulator's task is to accurately position a surgical tool (or a tool guide) in relation to the patient's skeletal anatomy. Since direct access to the patient's skeleton is obviously limited to the extent of the surgical incision, the patient's location with respect to the robotic manipulator must be determined indirectly using X-ray images. This *registration* or *calibration* process therefore requires the extraction of three-dimensional measurements from two-dimensional X-ray images, and as such, falls within the scope of a field of research which is known as *X-ray photogrammetry.*

However, unlike the vast majority of X-ray photogrammetry research, the current thesis involves an *intraoperative* application. That is to say, the three-dimensional measurements must be performed in the operating room at the time of surgery. In addition to the need to comply with surgical sterility, it is therefore essential that the 3-D measurement process does not delay the surgical procedure. As a consequence, the use of *machine vision* or *computer vision* (i.e. digital imaging technology) techniques has been proposed as a means of providing an intraoperative X-ray

photogrammetry capability. This approach initially requires the intraoperative images to be digitised. Semi-automatic *image analysis* algorithms are then applied to rapidly extract calibration data from these digital images. The required triangulation-based 3-D reconstruction process can then be performed using appropriate photogrammetry algorithms.

Given these objectives, the aim of this chapter is to provide an overview of both the orthopaedic and the engineering aspects of the thesis. Accordingly, Section 1.2 starts with a generalised discussion of orthopaedic surgery, and then introduces the problems associated with the specific branch of the orthopaedic speciality (*osteosynthesis*) under investigation. Similarly, Section 1.3 introduces the concepts and terminology associated with *robotic-assisted surgery*, building up to a brief outline of the research project to which this current thesis contributes. Having introduced this background material, a more detailed analysis of the problems which must be addressed by the thesis is then presented in Section 1.4. The organisation of the actual thesis document, in relation to the chapters which follow, is then discussed in Section 1.5. Finally, Section 1.6 provides a summary of the main points raised by the chapter.

1.2 Orthopaedic Surgery.

According to Stedman's Medical Dictionary **[1]**, *orthopaedic surgery* is "the branch of surgery that embraces the treatment of acute and chronic disorders of the musculoskeletal system, including injuries, diseases, dysfunction and deformities in the extremities and spine". A large proportion of an orthopaedic surgeon's workload is therefore concerned with the treatment of fractured bones and their related injuries. As further explained in Section 1.2.1, the management of these fractures increasingly involves the surgical insertion of mechanical fixation devices. The aim of this thesis is to derive engineering-based solutions to a number of problems currently associated with these fracture fixation procedures. Accordingly, this section gives a brief outline of these problems from the orthopaedic perspective, and begins by discussing the nature of orthopaedic surgery as a whole.

Following Röentgen's discovery of X-ray radiation in 1895, diagnostic radiography (i.e. medical X-ray imaging) soon became an essential part of the practice of orthopaedic surgery. In order to compensate for the fact that an X-ray image is a two-dimensional representation of the patient's internal anatomy, and as such does not provide depth information, it is standard practice to acquire images from two different perspectives. For many years, orthopaedic procedures were therefore planned purely on the basis of a pair of preoperative radiographic images. The actual surgical procedure was then performed in a "hands on" manner, via an incision which was large enough to allow direct physical access to the surgical site. Verification of the outcome of the surgical procedure would then be obtained by a postoperative radiographic examination.

In recent years, the development of real-time (instantaneous) intraoperative imaging techniques has allowed a major change in the way many orthopaedic procedures are actually performed. The introduction of arthroscopy (endoscopic examination of a joint) means that procedures involving the knee, shoulder, or ankle, can now be performed using "keyhole" or "minimally invasive" surgery techniques. Similarly, the use of intraoperative fluoroscopy (a real-time radiography technique) has led to the adoption of "closed" surgery techniques. This type of procedure relies upon fluoroscopic visualisation of the surgical site rather that direct physical access, and can therefore be performed via a much smaller incision. This reduced invasiveness means that the patient experiences less surgical trauma (i.e. devitalisation of soft tissue, blood loss, risk of infection, etc.), thus speeding up the postoperative recovery process.

However, in spite of these radiological advances, the nature of orthopaedic surgery has largely remained unchanged. Orthopaedic procedures are still performed, using hand-held surgical instruments and air-powered tools (e.g. drills and oscillating saws), in a freehand manner that is analogous to the "wet carpentry" of bone **[2]**. This subjective "low tech" approach has a number of important implications with respect to surgical results. Owing to the high levels of skill required, the outcome of an orthopaedic procedure is dependent upon the ability and experience of the individual surgeon, and is therefore highly variable. In order to acquire the necessary level of expertise,

trainee orthopaedic surgeons must also follow a long learning curve, during which they are required to operate on living patients.

It can therefore be appreciated that in relation to accuracy and repeatability, existing orthopaedic surgery techniques leave considerable scope for improvement. Since the postoperative complication rates associated with many orthopaedic procedures are directly related to surgical accuracy, any measure which can enhance the precision of a surgical procedure may reduce the need for corrective revision surgery. As a consequence, in addition to the obvious health-related benefits experienced by the patient, improved surgical techniques can also lead to significant financial savings due to the reduced burden that is placed upon orthopaedic resources.

1.2.1 Osteosynthesis.

One of the primary functions of the skeleton is to provide a rigid framework for the muscles which control human motion. When a bone fracture occurs, the resulting loss of skeletal continuity allows the muscles attached to the broken bone fragments to contract. These muscle forces tend to pull the bone fragments away from their natural alignment, and as such, in the absence of medical treatment, the natural bone healing process may result in severe malunion of the fracture (i.e. the bone fragments unite in an abnormal alignment). Alternatively, in cases where the blood supply to the fracture site is compromised, or the ends of the broken bone fragments are not in physical contact, non-union (i.e. failure to unite) of the fracture will occur. In addition to the complications associated with injuries to the surrounding soft tissues, bone fractures can therefore lead to permanent deformity or disability if left untreated.

In order to overcome these potential problems, the orthopaedic treatment of bone fractures involves a three stage process: *reduction*, *stabilisation*, and *bone healing*. The term reduction refers to the restoration under anaesthesia of the broken bone fragments to their anatomically correct positions. In practice, this process involves either a manipulative (closed reduction) or a surgical (open reduction) procedure. Having obtained axial and rotational alignment of the broken bone fragments, a suitable stabilisation technique is then applied to ensure that reduction is maintained

throughout the natural bone healing process (i.e. no relative motion occurs between the bone fragments). Over a period of months, the gap between the fracture surfaces is then progressively filled in by newly formed bone, eventually leading to solid union of the fracture site in the anatomically correct alignment.

Traditionally, stabilisation of the fracture site was achieved by applying either skeletal traction, or a plaster of Paris cast. However, as these conservative treatment options require prolonged immobilisation of the injured body part, complications such as muscle wastage and joint stiffness are inevitable. Given that an extended period of hospitalisation may be required, treatment can also prove to be very expensive. As a consequence, operative treatments are increasingly becoming the preferred method of fracture stabilisation. In particular, *internal fixation* (*osteosynthesis*) techniques are currently very popular, owing to the fact that they allow early postoperative mobilisation of the muscles and joints adjacent to the fracture, thus avoiding many of the problems associated with non-operative treatments. Osteosynthesis procedures achieve excellent stabilisation of the fracture site by surgically attaching mechanical fixation devices (e.g. screws, pins, nails, rods, plates, etc.) to the broken bone fragments. These fixation devices then remain inside the patient for the duration of the bone healing process, and ideally are eventually removed following radiographic confirmation of solid union.

Internal fixation devices are now routinely used to stabilise fractures of the long bones in both the upper (humerus, radius, or ulna) and lower (femur, tibia, or fibula) limbs. The scope of osteosynthesis also includes many spinal, pelvic, and shoulder (scapula and clavicle) fractures. However, in order to obtain the clinical benefits offered by this form of fracture stabilisation, a technically demanding surgical procedure may have to be performed. Many internal fixation devices require the insertion of a metallic component inside the broken bone fragments. The insertion of this component will therefore typically involve a series of "blind" (i.e. without direct visualisation) drilling and reaming processes. The recent changes in orthopaedic practice towards "closed" surgical techniques, also mean that direct physical access to the fracture site may not be available. As a consequence, an orthopaedic surgeon's

only means of monitoring his/her progress during an internal fixation procedure is frequently provided by intraoperative X-ray imaging. Internal fixation procedures are therefore characterised by a heavy reliance upon intraoperative fluoroscopy. As further discussed in Section 1.2.2, this increased use of fluoroscopy has led to growing concerns that the orthopaedic surgeon may now be exposed to unacceptable levels of X-ray radiation.

The use of two-dimensional X-ray images to guide a three-dimensional surgical process also constitutes a depth perception problem. The orthopaedic surgeon is therefore required to mentally correlate the 2-D information provided by the X-ray images with the patient's actual (3-D) anatomy. Since two orthogonal, and hence very dissimilar fluoroscopic views of the patient's anatomy are generally acquired, this process requires high levels of concentration and can prove to be quite stressful. Given that an individual's hand-eye coordination and visuospatial ability (i.e. the ability to comprehend and conceptualise visual representations and spatial relationships when performing a task) are inherent skills, they do not necessarily improve with experience. Many orthopaedic surgeons therefore encounter great difficulties with the trial-and-error techniques used to insert internal fixation devices under fluoroscopic guidance. Consequently, several attempts may be required in order to achieve satisfactory positioning of the fixation device. As each subsequent attempt prolongs the duration of the surgical procedure, the additional anaesthetic received by the patient obviously has a detrimental effect upon postoperative recovery. Irradiation of both the patient and the surgeon is also increased due to the need to acquire additional fluoroscopic images.

1.2.2 Radiation Exposure to the Orthopaedic Surgeon.

Diagnostic radiography directly exposes the patient to a potentially hazardous form of radiation (i.e. X-ray radiation). However, as radiographic examinations tend to be infrequent, and the patient benefits directly from the resulting diagnostic information, it is widely believed that the risks associated with patient irradiation are "acceptable". Unfortunately, the medical staff performing the radiographic examination may also be indirectly exposed to a small dose of X-ray radiation. Given that the detrimental

effects associated with irradiation are known to be cumulative, this repeated occupational exposure to X-ray radiation represents a serious problem.

In the radiology suite, the occupational exposure problem is avoided by placing the control panel for the X-ray imaging equipment behind a lead-lined enclosure. However, owing to the limited amount of space that is available in the operating room, the use of protective screens is not a viable option during orthopaedic procedures involving intraoperative fluoroscopy. Both the surgeon and the members of the surgical team are therefore required to wear radiation attenuating lead aprons. Since exposure to radiation obeys an inverse square law, with respect to distance from the radiation source, the members of the surgical team are able to significantly limit the amount of radiation they receive by maximising their distance from the fluoroscopy equipment at the time of image acquisition. Regrettably, during some internal fixation procedures this is a luxury which is not afforded to the surgeon.

An internal fixation procedure will only be attempted following satisfactory reduction of the fracture. Ideally, closed reduction will be achieved by manipulating the bone fragments under fluoroscopic guidance. In order to avoid loss of reduction, it may be necessary for the surgeon to hold one of the patient's limbs while a fluoroscopic image is acquired. When surgically preparing the hole for an internal fixation device, fluoroscopy is also frequently used to monitor the introduction of a surgical tool into the bone. The surgeon may therefore be required to hold the tool in position, or support the weight of the tool, while a fluoroscopic image is acquired. This level of proximity to both the X-ray beam and the patient makes irradiation of the orthopaedic surgeon inevitable. In addition to the realistic possibility of direct exposure, resulting for example from a hand unintentionally straying into the X-ray beam, the surgeon is indirectly exposed to the scattered radiation which emanates from the patient's body. The parts of the surgeon's body which are not protected by the lead apron, most notably the head/neck region and the hands, therefore receive a quantifiable dose of X-ray radiation.

As discussed at length in Appendix 1, the growing concern that exposure to X-ray radiation now represents a potential occupational health hazard for the orthopaedic surgeon has prompted a number of studies aimed at quantifying surgeon irradiation. Reassuringly, these dosimetric studies tend to suggest that the cumulative radiation dose received by an orthopaedic surgeon, when working at normal patient workloads, is well within the dose limits currently recommended by radiation protection standards. However, the results of these studies also indicate that fluoroscopic screening times and irradiation levels can vary considerably. When performing certain internal fixation procedures, inexperienced orthopaedic surgeons also have a tendency to overuse fluoroscopy. In addition to significantly increasing irradiation of the patient, these prolonged screening times can lead to excessive irradiation of an individual surgeon.

Irrespective of the results derived from dosimetric studies, irradiation of the orthopaedic surgeon remains a controversial issue, due to the uncertainties which currently surround the harmful effects of long-term exposure to low-dose radiation. Existing knowledge of radiation related health risks is mainly derived from studies of populations exposed to high-dose radiation (e.g. the survivors of the Hiroshima and Nagasaki nuclear explosions). In order to set maximum permissible dose limits for occupationally exposed personnel, extrapolation of data relating to high-dose exposure has therefore been necessary. Over the past fifty years, re-evaluation of this high-dose data, and the extrapolation models used to derive risk estimates, has led to several successive reductions in the recommended dose limits. A number of studies have therefore concluded that it is wise to work on the assumption that there is no "safe" dose of radiation.

1.3 Robotic-Assisted Orthopaedic Surgery.

When viewed from an engineering perspective, most orthopaedic surgery procedures appear to involve a sequence of bone "machining" processes (i.e. drilling, sawing, milling, tapping, etc.). Continuing this analogy, the problems outlined in the previous sections would therefore seem to require accurate, repeatable "machining" processes to be performed in a hazardous environment (i.e. in the presence of X-ray radiation). To

an engineer, such an application immediately suggests the use of some form of robotic manipulator. However, the introduction of such a device into the operating room is hindered by two major problems; the need to ensure the safety of both the patient and the surgical team, and the need to maintain surgical sterility. A field of research which is generally referred to as *robotic-assisted surgery* has therefore been established, with the aim of overcoming these problems.

As the name suggests, the ultimate goal of robotic-*assisted* orthopaedic surgery is not to replace the surgeon. Instead, the computer-controlled robotic manipulator is used to complement the capabilities of the orthopaedic surgeon, with a view to obtaining better results than either could achieve alone. The surgeon therefore performs most of the surgical procedure in the usual manner, and then uses the robotic manipulator to assist one or more of the bone "machining" processes. This robotic-assistance may be either invasive or non-invasive. In the non-invasive scenario, the robotic manipulator is used to align a tool guide with a pre-defined trajectory. The manipulator then maintains this alignment while the surgeon inserts a surgical tool into the tool guide, and then manually performs the bone "machining" process. In the invasive scenario, the bone "machining" process is automated by mounting an air-powered surgical tool on the manipulator's end-effector. Alignment with the desired surgical trajectory is therefore followed by robotic "machining" of the bone.

In order to derive positioning/machining instructions for the manipulator, robotic-assisted surgery requires quantification of the surgical procedure. Preoperative CT or MRI scanning is therefore normally used to define a quantitative surgical plan. However, as CT and MRI are not intraoperative imaging modalities, implementation of this surgical plan requires that the *intraoperative registration* problem be solved. In other words, the preoperative surgical plan must be made available in the operating room, and must be redefined in relation to the manipulator's coordinate system. One possible solution to this problem is to introduce calibration markers which are common to both preoperative imaging and the actual surgical procedure. Alternatively, a variety of intraoperative position-sensing techniques can be employed

to determine the locations of anatomical landmarks, which are clearly visible in the preoperative images.

The majority of the safety issues raised by robotic-assisted surgery stem from the manipulator's ability to move automatically under computer control. An obvious solution to this problem is to use a *passive* (motorless) rather than an *active* (motorised) manipulator. Since the intraoperative registration process establishes a direct link between preoperative planning and surgical execution, the position of the manipulator's end-effector can be displayed in real-time with respect to the preoperative MRI or CT images. Alignment with a desired surgical trajectory can therefore be achieved by manually moving the passive manipulator until the displayed end-effector position is coincident with target points in the MRI/CT images. The passive manipulator can then be locked in position to act as a tool guide during manual completion of the procedure.

For the purposes of this thesis, surgical procedures which involve the use of passive guidance and position-sensing systems will be referred to as *computer-assisted surgery* (CAS) applications. As indicated by the example provided in the previous paragraph, CAS systems incur lower levels of risk than their robotic-assisted surgery counterparts. However, this obvious advantage is offset by the fact that computer-assisted surgery is essentially a freehand form of surgery. In spite of the fact that quantitative information is provided to guide the surgical procedure, the precision of computer-assisted surgery is therefore still operator (i.e. surgeon) dependent. Although there are a few notable exceptions, the need to continually reposition a surgical tool/tool guide by hand, also means that the majority of CAS systems tend to be limited to procedures which involve simple two-dimensional trajectories. In order to maximise the benefits obtained by introducing new technologies into the operating room, the safety issues associated with robotic-assisted surgery must therefore be addressed.

The use of the term "*robotic*-assisted surgery" often leads to misconceptions which are not conducive to clinical acceptance. To a surgeon or patient, the term "robot"

immediately conjures up a mental image of the large industrial robots found on car plant assembly lines. However, these six degrees-of-freedom, anthropomorphic robotic arms are primarily designed to satisfy performance-related (i.e. versatility and repeatability) rather than safety criteria. As a consequence, industrial robots must undergo extensive modifications in order to safely achieve the high levels of human/robot interaction required by surgical applications. Bearing in mind that surgical sterility must also be maintained, the use of custom-built manipulators is therefore a more viable option for robotic-assisted surgery.

The kinematic configurations of these special-purpose manipulators provide only those degrees-of-freedom that are actually required by the particular surgical application under investigation. Special-purpose manipulators therefore tend to have a reduced number of degrees-of-freedom (typically four or five), thus simplifying the control process. By mechanically constraining these motions, it is possible to ensure that the patient remains outside the working volume of a "non-invasive" manipulator. In the case of "invasive" manipulators, fail-safe braking mechanisms can be employed to immobilise any degrees-of-freedom which are not required during the invasive stages of the procedure. The safety of the surgical team can also be ensured by operating the manipulator at reduced speeds. In the unlikely event of a malfunction, the surgeon therefore has sufficient reaction time to bring the manipulator to a stop using a dead man's handle device.

Assuming that the relevant safety and sterility issues can be addressed, robotic-assisted orthopaedic surgery offers many potential benefits. The majority of these benefits are obviously associated with the excellent precision (accuracy and repeatability) provided by a robotic manipulator. Ideally, the use of a manipulator to perform/guide bone "machining" processes will lead to significant improvements upon the results obtained by conventional surgical techniques. As a consequence of the increased level of automation, the influence of surgeon variables (i.e. skill, experience, etc.) upon surgical results should also be diminished, thus contributing to the desired standardisation of orthopaedic tasks.

Additional benefits include the fact that the quantitative image data used to define positioning/machining instructions for the manipulator, essentially eliminates the depth perception problem associated with 2-D radiographs. The manipulator's role as a surgical tool/tool guide positioning device, also leads to the possibility of "remote" surgery. That is to say, the orthopaedic surgeon can distance himself from both the patient and the X-ray beam at the time of fluoroscopic image acquisition. Theoretically, surgeon irradiation can therefore be reduced to the background radiation levels encountered in the operating room. Assuming that the enhanced surgical precision offered by robotic-assistance leads to "right-first-time" procedures, irradiation of the patient will also be substantially reduced as a result of fewer fluoroscopic images being required.

Finally, robotic-assisted orthopaedic surgery also has a number of financial implications. Although the initial capital investment required to develop a custom-built system is obviously substantial, and minor running/maintenance costs are also incurred, in the long-term robotic-assistance will ideally lead to more cost-effective surgical procedures. In addition to the previously mentioned lowering of postoperative complication rates, standardisation of performance will also potentially lead to reductions in the duration of surgical procedures. The number of operating room personnel required to perform a particular procedure may also be reduced, with the manipulator replacing a surgical assistant. Changes in surgical practice, including the introduction of new surgical procedures, may also become possible. As a result, the cost of the actual surgical procedure and postoperative patient rehabilitation, may both be significantly reduced.

1.3.1 The MEDROSA Research Project.

The research work documented in this thesis has been conducted as part of a robotic-assisted orthopaedic surgery project, which is known by the acronym "MEDROSA" (MEchatronic Design of a Robot for Orthopaedic Surgery Assistance). As implied by this project title, the MEDROSA research group has adopted the *mechatronic* design philosophy as a means of satisfying the unique requirements of robotic-assisted orthopaedic surgery. Mechatronics is a specialised branch of engineering which

involves the "synergistic integration of mechanical engineering with electronics and intelligent computer control in the design and manufacture of products and processes" **[3]**. Ideally, this approach leads to the development of "smart" or "intelligent" products which offer enhanced performance capabilities.

With the aim of placing the current thesis into context, a detailed discussion of the MEDROSA research project as a whole is provided in Appendix 2. In addition to providing details of the project's background, this appendix also outlines the aspects of the MEDROSA Project that do not fall within the scope of the current thesis. Further details can also be found in a number of peer-reviewed publications **[4-9]**. However, in relation to the current chapter, the most important aspects of the MEDROSA Project can be summarised as follows. Adoption of the mechatronic design philosophy has led to the development of a prototype robotic-assisted surgery system which incorporates three major components:

- A custom-built robotic manipulator.
- A PC-based robot controller and electronics cabinet.
- A PC-based machine vision sub-system which interfaces to existing intraoperative X-ray imaging (fluoroscopy) equipment.

The ultimate goal of this system is to improve the precision and safety (with respect to X-ray radiation) of the fluoroscopically guided drilling processes performed during osteosynthesis procedures. The machine vision system is therefore used to plan drilling trajectories, which are then implemented by the robotic manipulator. Given that osteosynthesis procedures do not routinely involve the use of preoperative MRI or CT scanning, the quantitative image data required to calculate the drilling trajectories is actually extracted from intraoperative fluoroscopic images. These drilling trajectories are then downloaded to the robot controller, which calculates the two translations and two rotations that must be performed by the manipulator, in order to align a drill/drill guide with the desired trajectory. Computer controlled alignment of the manipulator is then followed by either a "non-invasive" drilling process, performed by the surgeon with the assistance of the manipulator positioned drill guide,

or an automatic "invasive" drilling process, performed by the manipulator's "intelligent" drilling unit end-effector.

In order to focus development of the prototype MEDROSA system along the right lines, two specific osteosynthesis procedures have initially been targeted for investigation. Both of these procedures involve the internal fixation of femoral (thigh bone) fractures:

- Locked Intramedullary Nailing of Femoral Shaft Fractures:
  As described in Chapter 6, the most effective means of stabilising a femoral shaft fracture is to insert a stainless steel nail into the central marrow cavity of the femur. Depending upon the severity of the fracture, a process referred to as *distal interlocking* may then be performed to rigidly attach the nail to the femur. This interlocking process requires the insertion of two screws through the nail/femur. However, since the nail is inside the femur, the holes for these screws can only be drilled under fluoroscopic guidance. Thus, in addition to being the most technically demanding stage of the nailing procedure, distal interlocking also leads to higher levels of surgeon irradiation than any other osteosynthesis procedure (as demonstrated by the dosimetric studies discussed in Appendix 1).

- Sliding Compression Hip Screw Fixation of Hip Fractures:
  As described in Chapter 7, the internal fixation of hip fractures is the most commonly performed orthopaedic surgery procedure. At the present time, the preferred fixation device for this type of fracture is the sliding compression hip screw. In order to allow accurate positioning of this device within the femur, a pilot hole is initially drilled using a *guide wire*. The cannulated (hollow) component of the hip screw is then inserted over this guide wire. However, in practice, the fluoroscopically controlled guide wire insertion process is severely complicated by the depth perception problem. Surgeons can therefore require several attempts before a satisfactory guide wire placement is achieved.

Since the problems incurred during these femoral procedures are representative of osteosynthesis as a whole, it has been possible to address the long-term goal of

producing a generic (i.e. applicable to a wide range of osteosynthesis procedures) cost-effective system.

1.4 Outline of Thesis.

Essentially, this thesis deals with the machine vision related aspects of the MEDROSA Project. The thesis therefore represents an investigation of the surgical planning requirements for robotic-assisted osteosynthesis procedures, and is primarily concerned with the quantification of surgical tasks under operating room conditions. As with any robotic-assisted surgery application, this quantification process initially requires the position of the patient to be accurately determined in relation to the robotic manipulator (i.e. a registration/calibration process). Assuming that this spatial relationship can be maintained, drilling trajectories must then be defined with respect to a coordinate system which is compatible with the robotic manipulator. The transfer of this drilling trajectory data to the robot controller then represents the end of the surgical planning process, and hence the scope of this thesis.

As discussed in Appendix 2, past experience has shown that the development process of a robotic-assisted surgery system can take up to ten years. Given that the research work documented in this thesis coincides with the start of the MEDROSA Project, and has therefore been conducted in parallel to the development of the other system components, *in vivo* (i.e. on living human patients) clinical testing of the MEDROSA system has not been possible during the time scale of the thesis. Research has therefore initially been limited to an *in vitro* feasibility study under controlled laboratory conditions. Towards the end of the research period, a limited number of preclinical trials have also been performed under simulated operating room conditions.

The implications of the problems faced by this thesis are discussed from the engineering (machine vision) perspective in Section 1.4.1. A brief outline of the methodology which has been adopted in order to address these problems, then follows in Section 1.4.2.

1.4.1 Definition of Research Problem.

In accordance with the overall goals of the MEDROSA research project, the following criteria have been incorporated into the initial specification of the machine vision system:

- The safety of the patient and theatre staff must be ensured at all times.
- Existing medical equipment must be used (without major modifications).
- Minimum deviation from existing orthopaedic surgery and fluoroscopic practices.
- Full compatibility with the operating room environment.
- Any additional costs or lengthening of the operation time must be fully justifiable in terms of clinical benefits gained.

These criteria obviously stem from the desire to produce a clinically acceptable cost-effective system. However, in relation to the current thesis, they also have the following implications:

- Image Non-linearity:

  As previously noted, osteosynthesis procedures do not involve the use of preoperative computerised imaging (i.e. CT or MRI scanning). In order to define drilling trajectories that can be implemented by the manipulator, quantitative data must therefore be extracted from intraoperative fluoroscopic images. Unfortunately, the fluoroscopy units which are currently used during orthopaedic procedures are designed to provide diagnostic (i.e. qualitative) images, and exhibit non-linear imaging characteristics that are highly undesirable for a metrology application. As a result, quantitative data cannot be directly extracted from a fluoroscopic image.

- Limited Data Processing Time:

  The use of intraoperative images to provide drilling trajectory data also limits the amount of time that is available for data processing. Ideally, the MEDROSA system should neither delay nor prolong the surgical procedure, and as such, the relevant information must be rapidly extracted from the fluoroscopic images. The level of user interaction required to achieve this "image analysis" task must therefore be kept to a minimum.

- An Unconstrained Imaging Geometry.

  Intraoperative fluoroscopy involves the use of a mobile X-ray system which is commonly referred to as a *C-arm unit*. When subjectively positioning this X-ray system to acquire a fluoroscopic image of the patient's anatomy, the only criterion which must be addressed is the acquisition of a satisfactory field-of-view. Accordingly, the imaging geometry of the C-arm unit is totally unconstrained in terms of range/magnification, angle of incidence and repeatability. The resulting lack of prior knowledge, and limited control over the imaging geometry, significantly complicates the machine vision task.

- Choice of X-ray Views.

  During femoral fracture procedures, the intraoperative fluoroscopic examination process routinely involves the acquisition of two orthogonal views: a back-to-front (*posteroanterior* or *PA*) view and a side-to-side (*lateral*) view. Orthopaedic surgeons are therefore accustomed to working with these views. The introduction of differing or additional viewing angles would consequently be met with a certain degree of reluctance, and would also have to be justifiable in terms of the additional radiation exposure to the patient.

- Compatibility with Operating Room Equipment:

  Figure 1.1 depicts the major components required to perform robotic-assisted orthopaedic surgery. Surgical planning requires accurate knowledge of the spatial relationship between the patient and the manipulator. Given that this information is to be indirectly obtained using X-ray images, the registration technique must therefore be fully compatible with the robotic manipulator, the mobile C-arm unit, the patient's anatomy, and the operating table.

- Compatibility with the Operating Room Environment:

  Figure 1.1 is obviously an over-simplification of the actual operating room scenario. In reality, the available workspace is further limited by the presence of additional equipment (e.g. anaesthesia equipment, instrument trays, etc.) and the operating room personnel (e.g. the surgeon(s), anaesthetist, radiographer, scrub nurse, etc.). By machine vision standards, the modern operating room is therefore a

highly unstructured working environment. Additional problems are also caused by the absolute need to prevent contamination/maintain surgical sterility.

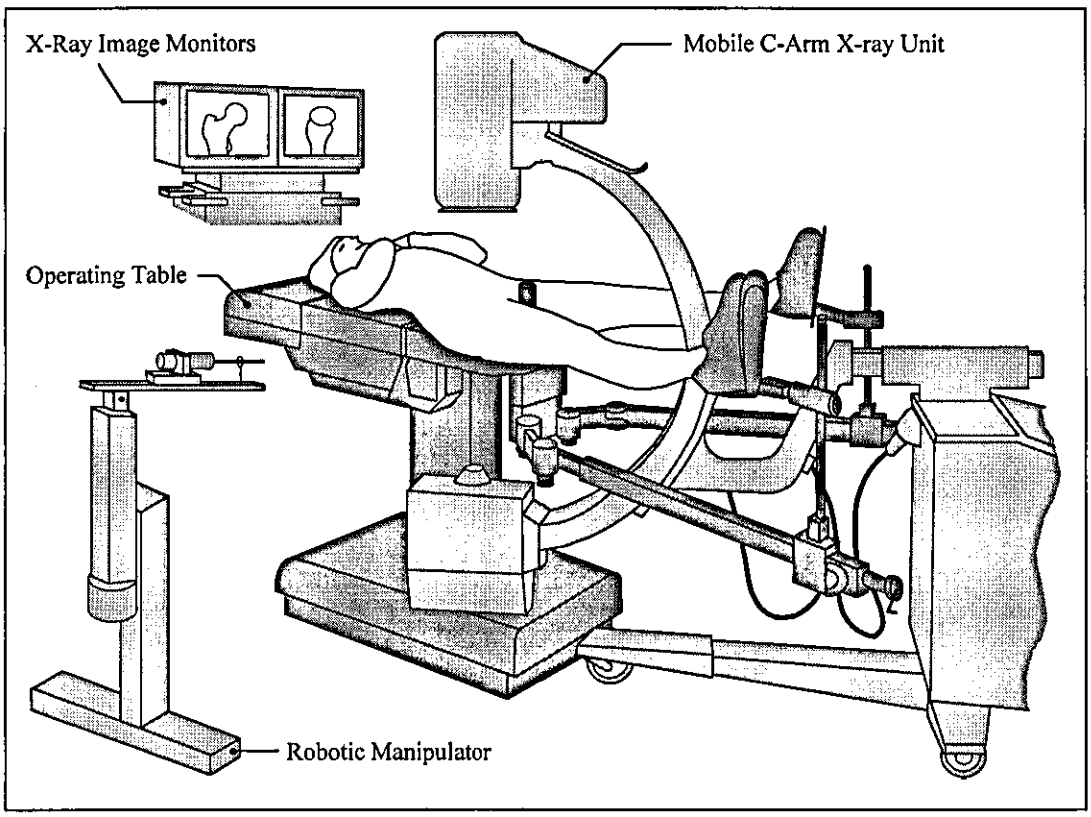


Fig. 1.1 : Robotic-Assisted Orthopaedic Surgery.

In view of these "given" constraints, the problem under investigation may initially be defined as the development of a clinically acceptable, low cost, intraoperative registration technique, which allows the rapid extraction of accurate, quantitative, three-dimensional data from non-linear X-ray images.

1.4.2 Proposed Research Methodology.

A clinically acceptable solution to the issues raised by this thesis can only be provided by quantifying the intraoperative fluoroscopic examination process. Three-dimensional trajectory data must therefore be extracted from a pair of orthogonal (PA and lateral) non-linear fluoroscopic images. Owing to the depth perception problem, this task is currently performed by the orthopaedic surgeon, who is required to mentally integrate the two-dimensional fluoroscopic images into a three-dimensional "model" of the patient's anatomy. However, the reconstruction of three-dimensional

measurements from two or more views of an object, is actually the basis of most three-dimensional machine vision techniques. If the fluoroscopic images can be captured in a digital format, semi-automatic image analysis and calibration algorithms can therefore be applied to rapidly extract the required trajectory data.

Mobile C-arm units display live (continuous) fluoroscopic images by outputting a standard analogue video signal to a monitor. In theory, digital fluoroscopic images can therefore be acquired by using the analogue-to-digital converter on a PC-based frame grabber card to sample this video signal. Accordingly, establishing a video link between a C-arm unit and a PC-based frame grabber card, should allow digital X-ray images to be stored in the card's image memory, or on the PC's hard disc. However, since fluoroscopy units have non-linear imaging characteristics, the acquired images are not a true representation of the patient's anatomy. Prior to extracting measurements from the stored images, it is therefore necessary to apply some form of compensation for this image non-linearity. Ideally, an *internal calibration* process will be performed preoperatively (i.e. off-line), leading to the derivation of correction factors which can subsequently be applied to the intraoperative images.

In machine vision terms, intraoperative registration of the patient with the robotic manipulator involves an *external calibration* process. Given the unconstrained nature of the C-arm fluoroscopy unit's imaging geometry, the analytical calibration techniques favoured by most machine vision applications are unsuitable for intraoperative use. It is therefore proposed that the empirical calibration techniques developed for *X-ray photogrammetry* applications, will be adapted to meet the requirements of digital intraoperative imaging. X-ray photogrammetry techniques are able to compensate for unconstrained imaging geometries by introducing artificial calibration markers into the X-ray images. In practice, this is achieved by positioning a calibration frame within the field-of-view of the X-ray equipment. In this way, metallic markers embedded in the calibration frame are superimposed into the X-ray images of the patient's anatomy.

It is believed that the intraoperative registration problem can be solved by mounting such a calibration frame on the robotic manipulator, and then positioning it around the patient at the time of image acquisition. Having calibrated the C-arm unit and the robotic manipulator, trajectory planning will then require the interactive indication (by the surgeon) of an entry point and a target point in the PA image. Selection of the corresponding points in the lateral image will then allow a triangulation-based reconstruction process to be performed, thereby calculating the actual (3-D) coordinates of the entry point and the target point. This pair of 3-D points can then be used to define the drilling trajectory as a 3-D vector/line in space.

With this proposed scenario in mind, the research work documented in this thesis has involved two distinct phases. Initially, a *feasibility study* was undertaken to evaluate the hardware (i.e. calibration objects) and software requirements of quantitative three-dimensional digital X-ray imaging. During this period of research, unrestricted access to a C-arm fluoroscopy unit was not available. Having established that the video signal from a C-arm unit could be acquired using a PC-based frame grabber card, system development was therefore facilitated by using CCD (charge-coupled device) cameras to simulate the C-arm unit. Having demonstrated the feasibility of the current thesis using this safe convenient development route, a C-arm unit was then purchased to allow the second phase of research to take place. The purpose of this *follow-up study* was to ensure that the benefits offered by the machine vision system could be transferred into the operating room. Accordingly, the techniques developed during the feasibility study were adapted to meet the requirements of the two femoral osteosynthesis procedures under investigation by the MEDROSA Project.

1.5 Thesis Structure.

In order to provide a logical flow of information, this thesis has been written in five sequential sections, some of which are supported by appendices. Section 1 (Chapter 1) introduces the orthopaedic background of the thesis, establishing a need for the introduction of new technologies into the operating room. The most quantifiable aspect of this need, namely irradiation of the orthopaedic surgeon, is addressed in some detail in Appendix 1. A generalised discussion of robotic-assisted orthopaedic

surgery is then followed by a brief introduction to the MEDROSA Project, of which the current thesis is an integral part. Ideally, this section should be read in conjunction with Appendix 2, which gives a more detailed account of the MEDROSA Project. The scope, objectives, and proposed methodology of the thesis are then outlined.

Section 2 (Chapters 2 and 3) serves the dual purpose of providing relevant theory and terminology, as well as a critical review of related research work. In view of the methodology adopted to satisfy the requirements of this thesis, Chapter 2 is devoted to a discussion of photogrammetry techniques. Since CCD cameras have initially been used to allow development of the machine vision system, the fundamentals of photogrammetry are initially presented in terms of their original camera-based applications. Brief discussions of the relevant radiography equipment are then followed by reviews of conventional (film-based) and digital X-ray photogrammetry. The implications of this existing knowledge in relation to the current thesis are then discussed.

By contrast, Chapter 3 is devoted to the intraoperative registration problem. A generalised overview of this problem is therefore followed by a discussion of intraoperative sensor technology. A critical review of the registration strategies adopted by a number of computer/robotic-assisted orthopaedic surgery applications is then provided. A similar treatment of neurosurgery applications is also supplied elsewhere, in Appendix 3. The review then ends with a critique of applications which involve the use of intraoperative fluoroscopy to establish registration. The implications of the chapter's findings are then discussed in relation to the methodology adopted by the current thesis.

Section 3 (Chapters 4 and 5) is concerned with the initial feasibility study. Chapter 4 therefore covers the CCD camera-based research work, and begins by outlining the intraoperative registration strategy of the MEDROSA system. The methodology used to develop the software modules and calibration hardware required to implement a generic 3-D machine vision system is then outlined. Consequently, Chapter 4 deals with the practical aspects of camera calibration, mathematical data reduction

techniques, and semi-automatic image analysis. The preliminary stages of the X-ray based research work are then discussed in Chapter 5. This chapter also includes an investigation of the internal calibration (image distortion) problem.

Section 4 (Chapters 6 and 7) covers the follow-up studies (i.e. pre-clinical trials) involving the femoral osteosynthesis procedures under investigation by the MEDROSA Project: internal fixation of femoral shaft and hip fractures. Chapter 6 therefore begins with a brief introduction to femoral anatomy, and the associated medical terminology, which is applicable to both chapters. Each chapter then presents a brief introduction to the nature of the relevant fracture type, and a discussion of the available treatment options. Details of the surgical procedure which is performed to insert the fixation device under investigation are then provided. The problems associated with the evaluated stages of each procedure, distal interlocking in the case of Chapter 6 and guide wire insertion in the case of Chapter 7, are then discussed. An outline of the solution proposed by this thesis is then followed by a discussion of the progress that has been made to date.

The fifth and final section (Chapter 8) evaluates the outcome of the body of research in relation to the original objectives of the thesis. A range of recommendations are then made, regarding a number of the issues raised by the thesis.

1.6 Summary of Chapter 1.

Changes in orthopaedic practice have led to an increased reliance upon fluoroscopic image-guidance during internal fixation (osteosynthesis) procedures. The resulting complexity of surgical techniques, and concerns over X-ray radiation exposure levels to orthopaedic surgeons, has prompted an investigation of robotic-assisted orthopaedic surgery, with a view to improving the precision, repeatability and radiation safety of existing surgical procedures. A prerequisite for such an application is accurate knowledge of the patient's location with respect to the robotic manipulator. Accordingly, a quantitative 3-D X-ray imaging system has been proposed as a means of performing this intraoperative registration process.

This PC-based vision system will acquire digital X-ray images by establishing a video link with the standard piece of intraoperative X-ray imaging equipment (i.e. the mobile C-arm unit). Operating room compatible calibration techniques, based upon the X-ray photogrammetry methodology, will then be applied, in order to compensate for the non-linear imaging properties of the C-arm unit, and to establish robot-to-patient registration. Having established registration, the vision system will allow the surgeon to interactively define drilling trajectories, which can then be accurately implemented by a robotic manipulator. Given the widespread use of intraoperative fluoroscopy, such a metrology "tool" will be highly beneficial to a large number of orthopaedic surgery procedures.

2. PHOTOGRAMMETRY

2.1 Introduction.

Photogrammetry is defined by the American Society of Photogrammetry as being "the art, science, and technology of obtaining reliable information about physical objects and the environment through the processes of recording, measuring, and interpreting photographic images and patterns of recorded radiant electromagnetic energy and other phenomena" **[10]**. In essence, photogrammetry therefore encompasses all reconstruction techniques which attempt to recreate an object's three-dimensional measurements using two or more two-dimensional images of the object.

As implied by the terminology, photogrammetric reconstruction techniques were originally developed for photographic (camera-based) applications, such as topography and aerial reconnaissance. These early techniques, which are generally referred to as "classical" or "metric" photogrammetry, achieve very high levels of accuracy through the use of specialised (metric) camera equipment. However, the unavoidable delay incurred by the film development process, means that classical photogrammetry cannot be applied to real-time applications. The need to manually extract image coordinates from a pair of photographs, using a device known as a stereo-comparator, further delays the reconstruction process. Severe limitations are also placed upon the imaging geometries that can be used during image acquisition. Classical photogrammetry is therefore generally considered to be an expensive, time-consuming and inflexible metrology method, and as such, is rarely used for non-topographic or close-range applications.

In recent years, developments in the area of computer vision have led to a renewed interest in photogrammetry. PC-based frame grabber cards and CCD video cameras are now readily available at affordable prices. When used in conjunction with digital image processing algorithms, this hardware allows the acquisition and analysis of digital images in near real-time. The film development and manual "digitisation" problems of classical photogrammetry can therefore be avoided, by using off-the-shelf

equipment to automate the reconstruction process. However, as the spatial resolution of current digital imaging systems is poor in comparison to that of photographic (analogue) film, the time gains achievable through the use of a PC-based photogrammetry system must be offset against a loss of reconstruction accuracy. The use of non-metric cameras, also necessitates an additional camera calibration stage in order to indirectly estimate the parameters of the imaging geometry, which are no longer constrained. Nevertheless, for engineering applications which require only moderate accuracy, digital photogrammetry is now a viable proposition.

Camera-based photogrammetry has also found several medical applications in recent years, most notably in the field of biomechanics. Due to substantial similarities between the image formation processes of photographic and X-ray images, photogrammetric techniques can also be applied to diagnostic radiography. However, the problems which are inherent to classical (camera-based) photogrammetry, have also acted as a deterrent to the routine clinical use of X-ray (film-based) photogrammetry. It was therefore not until the early 1970s that X-ray photogrammetry received serious consideration as a preoperative or postoperative diagnostic technique. Contemporary investigations are now beginning to consider the intraoperative use of photogrammetric techniques, which has been made possible by recent developments in digital radiography.

The aim of the remaining sections of this chapter is to summarise and analyse previous photogrammetric research. Section 2.2 covers aspects of camera-based (or visible light) photogrammetry. Classical photogrammetry theory, which forms the basis of both non-metric and X-ray photogrammetry, is initially outlined in order to establish the relevant photogrammetric concepts and terminology. A critique of modern (non-metric) photogrammetric techniques is also presented in order to demonstrate the extent of existing knowledge. A similar treatment of radiographic theory and previous film-based X-ray photogrammetry research follows in Section 2.3. The review of photogrammetric research is then completed by a discussion of image intensifier based "digital" X-ray photogrammetry in Section 2.4. An in-depth discussion of the geometrical image distortion problem, caused by image intensifier imperfections and

the influence of the geomagnetic field, is presented along with a critique of quantitative three-dimensional angiography research. The implications of the chapter's findings with respect to the objectives of the current thesis, are then summarised in Section 2.5.

2.2 Camera-Based Photogrammetry.

A photographic image is formed by the projection of a three-dimensional scene onto a two-dimensional film plane. All depth information is therefore lost during the image formation process, and as a consequence, an image point does not uniquely define the location of the corresponding real-world point. However, by acquiring two (or more) images of the same object from slightly different points of view, it is possible to extract three-dimensional measurements from photographic images using a triangulation-based technique known as *photogrammetric reconstruction.*

Classical photogrammetric reconstruction is based upon an analytical technique that requires an accurate knowledge of the external and internal camera parameters at the time of image acquisition. The external camera parameters define the pose (i.e. position and orientation) of the camera body in terms of the usual six degrees-of-freedom. The internal camera parameters relate to the actual optical system of the camera, and include the focal length and the position of the perspective centre. Once all of the internal and external camera parameters are known, geometric and trigonometric relationships can be applied to determine the spatial position of an object from its corresponding two-dimensional image coordinates. These relationships are based upon the assumption that light travels in straight lines or "rays". The mathematical interpretation of this assumption is known as the *collinearity condition*, and is outlined in Section 2.2.1.

In order to ensure that the internal and external camera parameters can be directly obtained, specialised photographic equipment must be used. Image acquisition for classical photogrammetry applications is therefore performed by a pair of metric cameras which are rigidly attached to one another. Mechanically constraining the imaging geometry in this way, ensures that the external parameters can be accurately

determined. Quantification of the internal parameters is made possible by a series of fiducial marks which are superimposed onto the photographs by the metric cameras. Assuming that the camera parameters are known, the only other piece of information which is required to reconstruct a three-dimensional object point, is the image coordinates of the corresponding points in the two photographs. Traditionally, the difference in position of these two image points, referred to as the *disparity*, was manually measured using a stereo comparator. An analytical method, which is outlined in Section 2.2.2, is then used to derive the object's location from the disparity measurement.

By contrast, knowledge of the actual camera parameters is not absolutely necessary for photogrammetry applications involving non-metric cameras. The use of separately constructed non-metric cameras, which do not superimpose fiducial marks onto the image, necessitates an additional *implicit camera calibration* process before triangulation-based reconstruction can take place. Strictly speaking, the term camera calibration refers to the experimental determination of the internal and external camera parameters. However, unless a given application specifically requires knowledge of the individual camera parameters, a model of the image formation process will normally be adopted that relates image and real-world points using a rigid-body transformation. Implicit "camera calibration" therefore involves the indirect determination of a series of model coefficients, that have no physical significance in relation to the camera parameters.

Implicit camera calibration, the main subject of Section 2.2.3, requires a set of image points whose real-world coordinates are known. In practice, corresponding real-world/image point-pairs are obtained by imaging a calibration object whose dimensions can be accurately determined. The image coordinates of easily identifiable markers on the calibration object, are then measured either manually or automatically using a digital image processing system. A least-squares data reduction technique is subsequently used to derive the coefficients of the rigid-body transformation that relates image and real-world coordinates. Once all of a system's cameras have been

calibrated in this way, providing that the imaging geometry remains identical to that used during calibration, photogrammetric reconstruction can be performed.

2.2.1 The Collinearity Condition.

The relationship between points in three-dimensional object space and their corresponding points in an image can be derived from the collinearity condition. This condition states that an object point and its corresponding image lie on a straight line, which also passes through the perspective centre of the camera, and may be expressed mathematically in terms of the two coordinate systems shown in Figure 2.1, where the axes:

- (X, Y, Z) define an arbitrarily allocated object or world coordinate system (WCS).
- (x, y, z) define a camera or image coordinate system (ICS), with axes (x & y) lying in and thus defining the image plane.

Using these coordinate systems, it is possible to define the following vectors:

- An Image Space Vector:
  A vector from the perspective centre to the image point in terms of image coordinates. $= \begin{bmatrix} x_i - x_p \\ y_i - y_p \\ 0 - c \end{bmatrix}$

  Where: x_i & y_i = image coordinates of a point.
  x_p & y_p = image coordinates of the principle point.
  c = a camera constant (i.e. the principal distance of the camera).

- An Object Space Vector:
  A vector from the perspective centre to the object point in terms of the object coordinate system. $= \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix}$

  Where: X, Y, Z = object/world space coordinates of the point.
  X_0, Y_0, Z_0 = object/world space coordinates of the perspective centre.

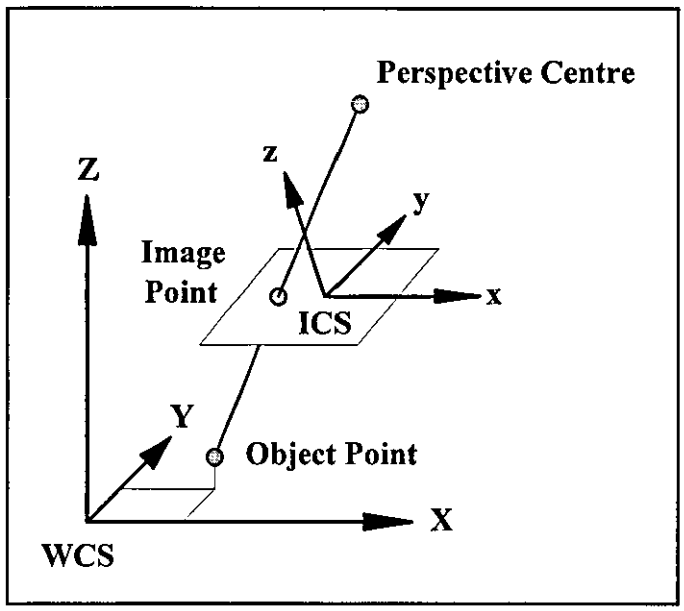


Fig. 2.1 : The Collinearity Condition.
(Adapted from Beyer **[11]**)

Since the collinearity condition states that these two vectors are collinear, they may be related by a multiplication of the object space vector by an orthogonal rotation matrix and scale factor:

$$\begin{bmatrix} x_i - x_p \\ y_i - y_p \\ -c \end{bmatrix} = \lambda . [M] \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix} \qquad .....(2.1)$$

Where:

λ = A linear scale factor.

[M] = A (3 x 3) transformation matrix from the image to world coordinate system.

Given that the scale factor is of little interest, it is common practice to eliminate it by dividing the first two rows of the matrix equation (2.1) by the third. As such, the collinearity condition is more generally stated by the two equations:

$$x_i - x_p = -c.\frac{m_{11}(X-X_0)+m_{12}(Y-Y_0)+m_{13}(Z-Z_0)}{m_{31}(X-X_0)+m_{32}(Y-Y_0)+m_{33}(Z-Z_0)} \qquad .....(2.2)$$

$$y_i - y_p = -c.\frac{m_{21}(X-X_0)+m_{22}(Y-Y_0)+m_{23}(Z-Z_0)}{m_{31}(X-X_0)+m_{32}(Y-Y_0)+m_{33}(Z-Z_0)} \qquad .....(2.3)$$

Where: m_{ij} = the elements of transformation matrix [M].

Equations (2.2) and (2.3) are the fundamental basis of most metric and implicit photogrammetry techniques.

2.2.2 Stereo-Photogrammetry.

Classical photogrammetry involves the use of two identical cameras, which are rigidly attached to one another in order to achieve the imaging geometry depicted in Figure 2.2. This set-up is generally referred to as the *stereo-photogrammetric* camera arrangement, and requires that the film plates of the two cameras lie within the same plane. As a consequence, the optical axes of the two cameras are parallel, and are separated by the distance between the centres of the two lenses, which is known as the *baseline* (B). The objective of such a system is to find the coordinates (X', Y', Z') of a real-world point (A), using only measurements of the corresponding image points $a'_1(x_1, y_1)$ and $a'_2(x_2, y_2)$.

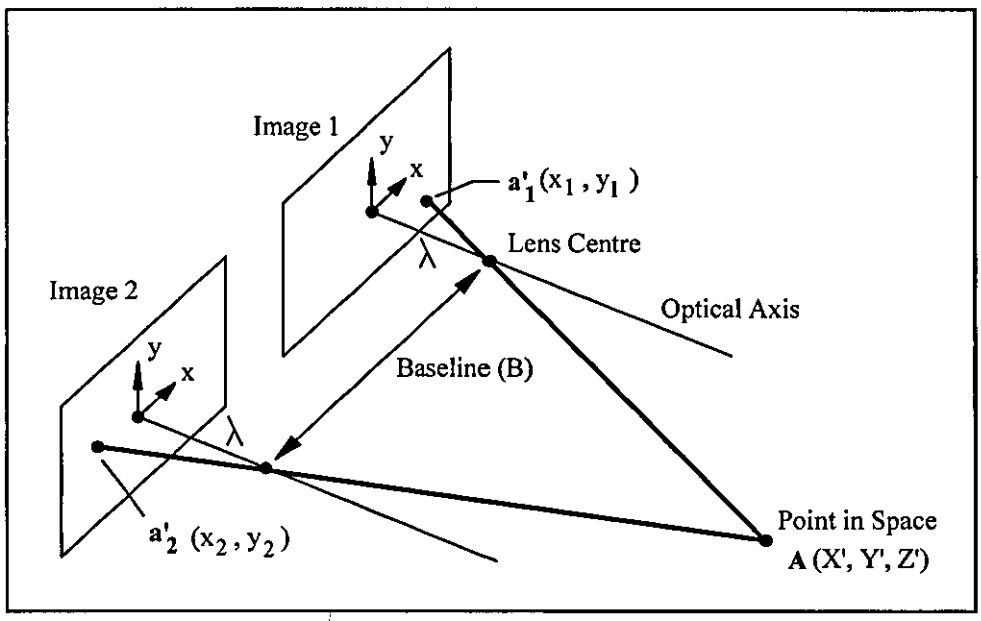


Fig. 2.2 : Model of the Stereo Imaging Process
(Adapted from Fu *et al.* **[12]**)

By convention, the image coordinate system (x, y, z) of each camera, is defined such that the x-y plane is coincident with the image plane and the z-axis lies along the optical axis. Similarly, the WCS origin (X, Y, Z) will generally be allocated as being

either at the midpoint of the baseline, or coincident with the image coordinate system of one of the cameras. For the purposes of this discussion, the WCS origin is assumed to be coincident with the origin of the ICS of image one. By applying the collinearity condition to the imaging geometry depicted in Figure 2.2, it can therefore be shown that the following geometrical relationships exist **[12]**:

$$Z' = \lambda - \frac{\lambda . B}{x_2 - x_1} \qquad .....(2.4)$$

$$X' = \frac{B . x_1}{x_2 - x_1} \qquad .....(2.5)$$

$$Y' = \frac{B . y_1}{x_2 - x_1} \qquad .....(2.6)$$

Where:

- λ is the focal length of the camera lens.
- B is the stereo baseline.
- $(x_2 - x_1)$ is referred to as the disparity (or parallax) and is the difference between the image coordinates of point (A) as recorded by the two cameras.

Assuming that the use of metric cameras allows the parameters baseline (B) and focal length (λ) to be directly obtained, the only unknowns required to calculate the world coordinates of point A(X', Y', Z') from equations (2.4 - 2.6), are therefore the image coordinates of $a'_1(x_1, y_1)$ and $a'_2(x_2, y_2)$. Classical stereo-photogrammetric reconstruction therefore involves two main stages: image acquisition and image data extraction.

Manually measuring the coordinates of corresponding image points from a pair of photographs is a slow, laborious task. Automation of the image extraction processes has therefore been the subject of much attention. Digital image analysis techniques (correlation, edge detection etc.) have been developed which are able to locate and measure an object of interest's position within an image. Applying these techniques to a stereo-pair of images, in order to determine the image coordinates corresponding to a

real-world point, is generally referred to as *image matching* or solution of the *correspondence problem.*

Once the object of interest has been located in the first image, knowledge of the geometrical relationship between the two cameras can be used to simplify the search for the corresponding point in the second image. This is achieved by taking advantage of the *coplanarity condition*, which states that an object point (A), its corresponding image pair (a'_1 & a'_2), as well as the lens centres of the two cameras (L_1 & L_2), must lie in a common plane (refer to Figure 2.3). Any plane that contains the stereo baseline (i.e. the line through L_1 & L_2) is generally referred to as an *epipolar plane*. The intersection of the epipolar plane with the second camera's image plane, defines an *epipolar line* along which the coplanarity condition dictates image point (a'_2) must lie. Analysis of the second image can therefore be constrained to a search along the epipolar line, thus significantly simplifying and speeding up the image matching process.

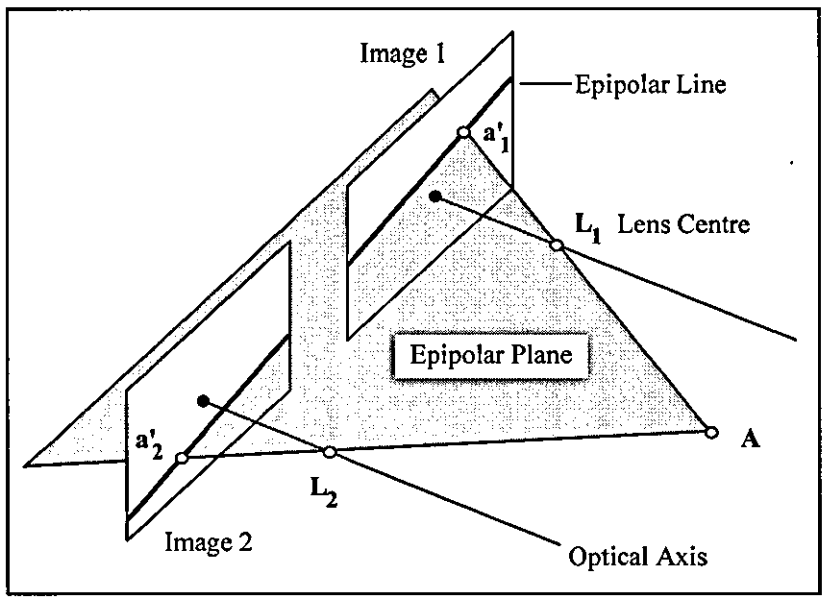


Fig. 2.3 : The Epipolar Constraint

Stereo-photogrammetry relies upon ambient lighting, and as such is considered to be a *passive* imaging technique. Shadows and reflections within the joint field-of-view of the two cameras, can therefore make automatic image matching difficult. One method of overcoming these problems, is to illuminate the imaged scene with controlled

lighting, thereby creating artificial image features which can be easily identified and matched. This approach is known as *active imaging* or *structured light*, and requires an accurate knowledge of the spatial relationship between the cameras and the additional light source.

Several authors have reported the use of structured light in conjunction with a stereoscopic imaging geometry. Robinson *et al.* (Trent Polytechnic, Nottingham) **[13]** describe a robotic vision system that sequentially projects a vertical He-Ne laser line into the field-of-view. Similarly, Gregory *et al.* (University of Bath) **[14]** use a commercial slide projector to provide a reference grid pattern, which aids the three-dimensional reconstruction of facial features. Jarvis **[15]** also gives an account of an alternative approach, which allows "stereo" imaging to be performed using only a single camera and a laser scanner.

2.2.3 Implicit Photogrammetric Reconstruction.

Although the mathematical theory behind classical stereo-photogrammetry is relatively simple, in reality, the required camera alignment can only be achieved through the use of highly specialised camera equipment. The provision of a more cost-effective photogrammetry system, based upon off-the-shelf (non-metric) cameras, therefore requires a two-step empirical approach to the photogrammetric reconstruction process:

- Step 1 : Implicit Camera Calibration.

  Using a collinearity condition based model of the imaging process, an image-to-world homogeneous transformation is defined for each of the imaging geometry's cameras. The coefficients of these models are then computed from the coordinates of a number of three-dimensional points and their corresponding image points. In practice these reference or *control points* are provided by imaging either a 3-D calibration frame, or a 2-D calibration plane which can be accurately moved within the common field-of-view of the cameras.

- Step 2 : Triangulation-Based Reconstruction.

  As the image-to-world transformation coefficients, established for each camera by the calibration process, are only valid for the specific imaging geometry used during calibration, the object to be measured must be placed in the volume previously occupied by the calibration object. These image-to-world transformations allow the *line-of-sight* (i.e. a vector) along which any imaged object point must lie, to be calculated from its corresponding image point coordinates. This process is known as solution of the *back-projection problem.* Applying back-projection to the image coordinates of an object point (P), as imaged by camera one, yields a line-of-sight that passes through point (P). Likewise, applying the same process to the image coordinates of point (P) as imaged by camera two, yields a second line-of-sight which also passes through point (P). The spatial intersection of these two lines-of-sight must therefore triangulate the true three-dimensional position of point (P).

The mathematical aspects of implicit camera calibration are discussed in sub-section 2.2.3.1. Sub-section 2.2.3.2 then outlines factors that can significantly influence reconstruction errors.

2.2.3.1 Mathematical Modelling of the Image Formation Process.

The majority of implicit camera calibration applications adopt the *pinhole* or *central projection* model depicted in Figure 2.4. This model assumes an infinitesimally small aperture located at the optical centre of the lens. All light rays from in front of the camera therefore converge at the pinhole (*centre of projection*) and are projected onto the image plane at the back of the camera. In order to avoid dealing with an inverted image, the image plane is considered to lie in front of the camera at a distance of one focal length from the pinhole. An image point is therefore the point where the line-of-sight passing through the object point and the centre of projection intersects the image plane.

Using homogeneous coordinates, the following theoretical world-to-image (*projection problem*) transformation can be derived from the pinhole model **[17]**:

$$V' = [P].[R].[T].V \qquad .....(2.7)$$

Where V' = Homogeneous image coordinate vector [4 x 1].

V = Homogeneous world coordinate vector [X, Y, Z, 1].

$[R]$ = Rotation matrix [4 x 4].

$[T]$ = Translation matrix [4 x 4].

$[P]$ = Perspective transformation matrix [4 x 4].

This mapping transforms an object point (expressed in homogeneous world coordinates) into the corresponding image point (expressed in homogeneous image coordinates).

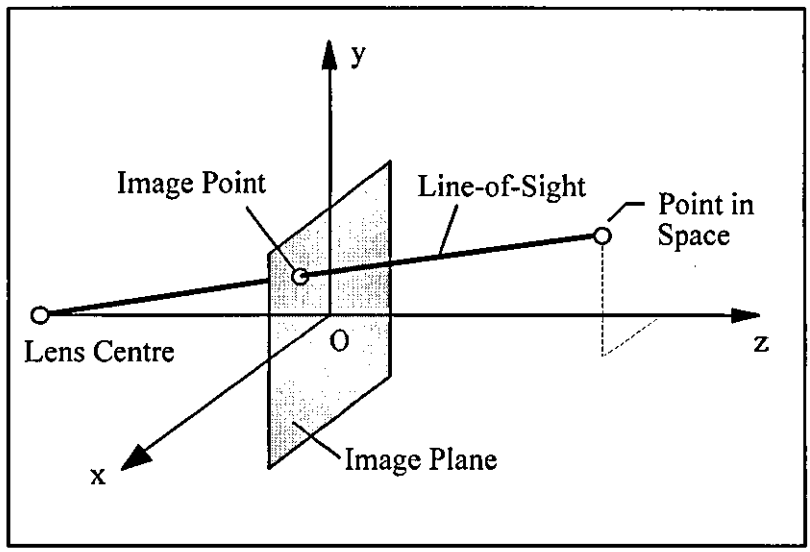


Fig. 2.4 : The Pinhole Camera Model.
(Adapted from Martins *et al.* [16])

The popularity of the pinhole model stems from the fact that this world-to-image transformation is easily invertable. A solution of the back-projection problem (image-to-world) is therefore obtainable from equation (2.7) by inverting the three matrices [R], [T], & [P]:

$$V = [T]^{-1}.[R]^{-1}.[P]^{-1}.V' \qquad .....(2.8)$$

In practice, knowledge of the actual translations and rotations described by matrices [T] and [R] may not be important. As a consequence, implicit camera calibration

frequently adopts a "black box" approach to the world-to-image transformation, by defining a single perspective projection transformation matrix [A] comprised from the matrix product [P][R][T]. Equation (2.7) therefore becomes:

$$V' = [A].V \qquad .....(2.9)$$

Essentially, equation (2.9) is identical to the matrix form of the collinearity condition as expressed by equation (2.1). The most commonly used method of solving this type of collinearity condition system of equations is the Direct Linear Transformation (DLT). The DLT was developed in 1971 by Abdel-Aziz *et al.* (University of Illinois, USA) **[18]**, as a means of simplifying the stereo-comparator technique of measuring photographic images. However, due to its mathematical simplicity, the technique has also become popular for digital photogrammetry applications. Several refinements of Abdel-Aziz's original form of the DLT have therefore been reported **[19, 20]**.

The DLT establishes a linear relationship between image (x, y) and world (X, Y, Z) coordinates in terms of eleven unknown parameters (L_i). This is achieved by simplifying the two collinearity condition equations (2.2 & 2.3) into a more computationally efficient form:

$$x = \frac{L_1.X + L_2.Y + L_3.Z + L_4}{L_9.X + L_{10}.Y + L_{11}.Z + 1} \qquad .....(2.10)$$

$$y = \frac{L_5.X + L_6.Y + L_7.Z + L_8}{L_9.X + L_{10}.Y + L_{11}.Z + 1} \qquad .....(2.11)$$

Due to the inherent linearity of equations (2.10 & 2.11) a direct solution for the unknown DLT parameters may be obtained from six or more real-world/image point correspondences.

Unfortunately, in reality no camera lens produces a perfectly linear world-to-image transformation. A certain amount of image distortion is therefore inevitable. If the modelling errors incurred by deviations from the collinearity condition are not

significant, with respect to the precision required by the photogrammetric application, image distortion may be ignored. However, for high accuracy applications some form of compensation for lens distortion must be provided by the imaging model. The DLT approach can be adapted to take lens distortion into consideration, by adding extra parameters to the DLT equations (2.10 & 2.11) **[21]**. A two-stage method, which uses a non-linear iterative technique to refine the results of an initial linear (DLT-based) calibration process, has also been investigated by a number of research groups **[22, 23]**. However, as these modified techniques can no longer be expressed by a series of linear equations, many of the advantages offered by a direct solution are lost. A number of alternatives to the pinhole camera model have therefore been proposed with a view to providing a simplified solution to the lens distortion problem.

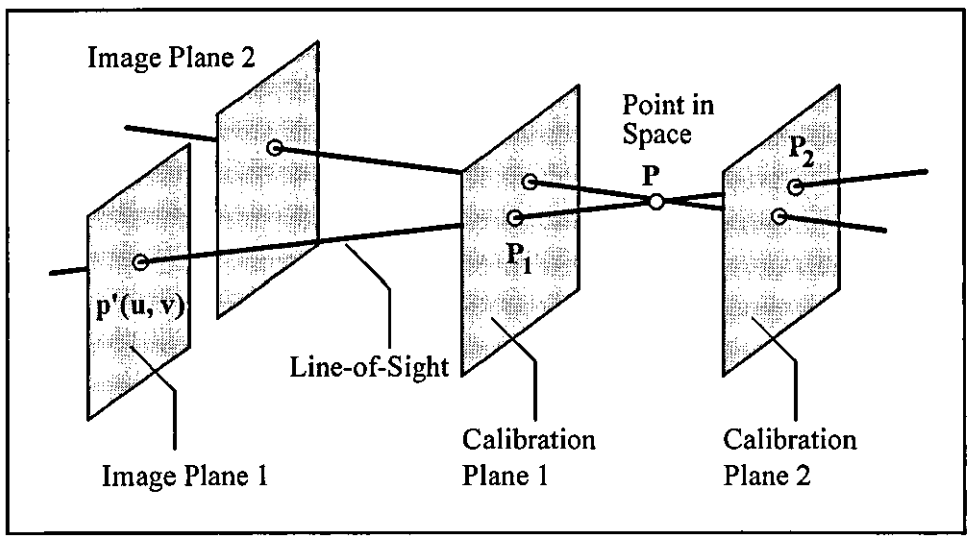


Fig. 2.5 : The Two-Plane Model.
(Adapted from Haigron *et al.* **[24]**)

One of the more promising alternatives to the pinhole model, is the "two-plane" model initially proposed by Martins *et al.* (University of Rhode Island, USA) **[P56]**. Unlike the pinhole model, the two-plane model does not assume a unique lens centre, and as such all lines-of-sight are not forced to go through the same point. The technique is therefore more versatile with respect to lens distortion. In practice, the two-plane method works by deriving independent interpolation functions for two parallel calibration planes placed in front of the camera (refer to Figure 2.5).

Once this calibration process has taken place, an image point p'(u, v) corresponding to a real-world point (P), can be used to calculate the point at which the line-of-sight crosses each of the calibration planes. These two real-world crossing points (P_1 & P_2) can then be used to define the line-of-sight along which real-world point (P) must lie. The intersection of this line-of-sight with the one passing through point (P) and the second camera, will therefore uniquely define the position of point (P).

Martins *et al.* report that accuracies comparable with pinhole-based techniques can be obtained using either linear or quadratic global interpolation functions. However, both of these methods were outperformed by the use of local interpolation techniques (linear spline fitting), which make no assumptions about the uniformity of lens distortion across the image. Champleboux *et al.* (TIMB-IMAG, Faculté de Médecine de Grenoble, France) **[25]** also report improved results, in comparison to the standard transformation matrix approach, through the use of the two-plane model in conjunction with localised interpolation techniques. However, they achieved even better results using an N-plane calibration technique (i.e. more than two calibration planes) and B-spline interpolation functions (refer to Table 2.1). A multi-plane approach has also been advocated by Haigron *et al.* (Université de Rennes, France) **[24]**, who obtained similar results to Champleboux *et al.* using thin plate spline interpolation functions.

Table 2.1 : Reconstruction Errors Reported by Champleboux *et al.* [25]

Method:	Max. Error (mm)	Mean Error (mm)
Perspective Transformation	2.19	0.63
Two-Plane Linear (Global)	3.25	0.90
Two-Plane Quadratic (Global)	1.63	0.39
Two-Plane Local	0.67	0.20
N-Plane B-Spline	0.46	0.15

The original two-plane model developed by Martins *et al.* was only applicable to the back-projection problem (image-to-world). However, Gremban *et al.* (Carnegie Mellon University, Pittsburgh, PA, USA) **[26]** and Wei *et al.* (Chinese Academy of Sciences, Beijing, China) **[27, 28]** have both subsequently extended the two-plane

model to also provide a solution for the projection problem (world-to-image transformation). The two-plane model therefore now offers a complete solution to the camera calibration problem.

2.2.3.2 Minimising Reconstruction Errors.

Although many factors can influence the accuracy of photogrammetric reconstruction, it can be shown that the major ones are:

- The imaging device.
- The imaging geometry.
- The calibration process.
- The accuracy of the calibration data.

Reconstruction accuracy is heavily dependent upon the imaging characteristics of a photogrammetry system's camera equipment. For photographic applications, lens distortion and film unflatness are widely acknowledged as being the main contributors to reconstruction errors. Metric cameras overcome these problems through the use of high quality lens systems and film flattening devices. By contrast, non-metric cameras rely upon specialised data reduction techniques to compensate for the resulting image distortion. Significant errors can therefore occur if the analytical model does not accurately represent reality.

Film unflatness is obviously not a problem for video cameras. However, the sensing elements of *vidicon* cameras, a generic term for video cameras based upon traditional television type technology (i.e. thermionic tubes), also display non-linear imaging properties. A number of authors, including Martins *et al.* **[16]**, have therefore obtained poor reconstruction accuracy when using the linear pinhole model to approximate vidicon cameras. As a result, solid state (CCD) video cameras, which have almost spatially linear imaging properties, are more popular for digital photogrammetry applications. The main errors associated with CCD-based digital applications are lens distortion, which remains the main source for concern, and the sampling or quantisation error caused by digitising the video signal.

The volume which can be calibrated/measured by a photogrammetry system is dictated by the common field-of-view (FOV) of the two cameras. An increase in the FOV, which may be obtained by either decreasing the focal length of the camera lens or increasing the camera-to-object distance, leads to a corresponding decrease in image magnification (i.e. the ratio of image to real-world dimensions). As a consequence, the error associated with the image measurement process, corresponds to a larger error in real-world terms at increased FOVs. Accordingly, the overall reconstruction error tends to increase as the camera-to-object distance (range) or FOV increases. Differences in the imaging geometries used by photogrammetry systems, therefore make direct comparisons ineffective when based purely upon mean, absolute, or root mean square (RMS) errors (expressed in millimetres):

- Table 2.2 summarises a number of biomechanical (human movement analysis) applications of photogrammetry. The measurement volumes for these applications are dictated by human anatomy, and range from measurements of small body parts (e.g. fingers, etc.) to full gait analysis over several strides.

- Table 2.3 summarises a number of studies which typify engineering or scientific applications of photogrammetry. This type of study usually requires sub-millimetre accuracy, and as such, is restricted to smaller measurement volumes than biomechanical applications.

In order to aid comparison, reconstruction errors are frequently expressed as a ratio of either the camera-to-object distance or the field-of-view. Usually, the RMS reconstruction error to camera-to-object distance ratio will be stated in either ratio (i.e. 1:4000) or percentage form (i.e. 0.025%).

When evaluating published results, it is important to observe whether quoted figures are derived from the reconstruction of control points or test points. The reconstruction of control points merely gives an indication of the accuracy of the mathematical calibration technique employed (i.e. the modelling errors). A true measure of system performance can only be obtained through the provision of test points which were not used during the calibration process. In practice, reconstruction errors for these independent test points are always larger than those of control points.

Table 2.2 : Details of Selected Biomechanical Applications of Photogrammetry

Author	Imaging System	Calibration Object	Reconstruction Error
Fioretti *et al.* (1985) [29] Universitá di Ancona (Italy)	Film Cameras (Range = 0.45 m)	Framework: Wire Grid 0.2 m x 0.2 m x 0.09 m	Better than 0.1 % RMS ≈ 0.3 mm
Wood *et al.* (1986) [30] Univ. of Western Australia	Film Cameras (Range = 12.00 m)	Wedge Shape Frame: 3.5 m x 2.5 m x 1.5 m	RMS ≈ 10 mm
Hatze (1988) [19] University of Vienna (Austria)	Film Cameras (Range = 4.45 m)	Frame: 0.4 m x 2.0 m x 2.0 m	1:2860 (wrt FOV) RMS ≈ 0.73 mm
Challis *et al.* (1992) [31] Loughborough University	Film Cameras (Range = 2.55 m)	Rectangular Frame: 1.0 m x 0.6 m x 1.0 m	RMS ≈ 2.5 mm
Bing Yu *et al.* (1993) [32] University of Iowa (USA)	VHS Camcorder (Range = 40.00 m)	Spherical Object: 24.4 m x 2.4 m x 2.4 m	RMS ≈ 20.0 mm

Table 2.3 : Details of Selected 3-D Machine Vision Applications

Author	Imaging System	Calibration Object	Reconstruction Error
Martins *et al.* (1981) [16] University of Rhode Island	CCD Cameras (Range = 25.00 in)	Robot moves white target piece	Average Error = 0.004 inch
Tsai (1986)[22] IBM, New York (USA)	CCD Cameras (Range = 4.50 in)	Plate: 16 black squares 1" x 1" x 0.5"	1:8000 wrt depth 1:4000 wrt FOV
Beyer (1989) [11] Swiss Federal Institute of Tech.	CCD Cameras (Range = 2.40 m)	Plate: white targets 0.5 m x 0.5 m x 0.5 m	RMS ≈ 0.2 mm
Sid-Ahmed *et al.* (1989) [23] Univ. of Windsor, Ontario	CCD Cameras (Range = 0.60 m)	Plate: 40 mm x 40 mm black disks/white card	Average Absolute Error ≈ 0.06 mm
Champleboux *et al.* (1992) Grenoble (France) [25]	Vidicon Cameras (Range = 1.8 m)	Plate with 265 holes: 0.3 m x 0.3 m x 0.3 m	Mean Error = 0.15 mm

Research has shown that reconstruction quality is heavily dependent upon the number and distribution of the control points around the measurement volume **[19, 30, 31, 33]**. The minimum number of control points which can be used, is dictated by the number of unknown parameters in the mathematical model adopted. However, in order to reduce the effects of measurement errors, oversampling is advisable for practical applications. A linear least-squares minimisation technique is therefore normally used to obtain a "best fit" solution based upon more than the minimum permissible number of control points. The optimum number of control points for a particular application depends upon the type of interpolation function being used and the accuracy requirements of the metrology task.

The original (linear) form of the Direct Linear Transformation, which has eleven unknown coefficients, requires a minimum of six control points (each world/image correspondence produces two equations). Further increases in the number of control points, should theoretically lead to improved accuracy. However, studies such as those summarised in Table 2.4 have shown that in practice, additional control points produce only minor improvements in the overall reconstruction accuracy. As these improvements are obtained at the expense of increased computational effort, in terms of both image analysis and data reduction, a point of diminishing returns is eventually reached.

Table 2.4 : The Effect of Control Point Numbers on DLT Reconstruction Errors.

	7CP	11CP	30CP
X	4.0	4.3	4.7
Y	6.7	7.2	6.4
Z	6.5	6.7	6.0
Mean	10.2	10.7	9.9
Abs.	5.7	6.1	5.7

Wood *et al* **[30]**

	7CP	11CP	30CP
X	6.2	5.5	5.3
Y	4.4	4.1	4.2
Z	5.2	4.9	4.8
Mean	5.3	4.8	4.7
Abs.	9.2	8.4	8.3

Hatze **[19]**

	8CP	20CP	36CP
X	2.3	2.3	2.4
Y	2.0	1.9	2.0
Z	2.2	2.2	2.2
Mean	2.2	2.1	2.2
Abs.	3.8	3.7	3.8

Challis *et al.* **[31]**

For this reason, applications involving the linear form of the DLT generally make use of only slightly more than the theoretical minimum number of control points. Gremban *et al.* **[26]** also found that the "density of the calibration grid makes little or no difference to the accuracy of the result", when applying global linear interpolation to the two-plane calibration technique (refer to Table 2.5).

Table 2.5 : Calibration Accuracy Using the Two-Plane Method:
Comparison of results obtained using global and localised interpolation functions.
(Adapted from Gremban *et al.* **[26]**)

Grid Size	Error (mm)	
	Global	Local
3 x 3	0.388	0.296
5 x 7	0.366	0.166
7 x 10	0.350	0.169
15 x 20	0.366	0.147
Range ≈ 0.53 m	1:1400	1:3500

However, when non-linear or localised interpolation functions are used, significant improvements in reconstruction accuracy can be obtained by increasing the number of control points. Gremban *et al.* **[26]** found that when a local linear spline interpolation technique was applied to the two-plane method, the reconstruction error could be halved by increasing the number of control points from nine to three hundred (refer to Table 2.5). Gremban et al's results are also in agreement with a number of other studies **[16, 24, 25]**, which suggest that localised interpolation functions yield better results than global linear interpolation. This is due to the fact that global interpolation techniques assume that image distortion is uniform or symmetrical over the whole image. However, if this is not the case, global functions will over/under correct distortions at a given point in the image.

Wood *et al.* **[30]** demonstrated that extrapolation-based reconstruction of target points lying outside the calibrated volume, incurred a 50-100% increase in measurement errors for the DLT approach. Similar findings have subsequently been reported by Hatze **[19]** and Challis *et al.* **[31]**. The two-plane method has also been shown to produce poor reconstruction accuracy when the target point is not located between the two calibration grids **[26]**. Extrapolation-based errors should therefore be avoided at all costs, by ensuring that the control point distribution completely surrounds the object to be measured. A symmetrical control point distribution should also be used in order to avoid bias in the reconstruction errors.

The final major influence upon system accuracy is the quality of the measured calibration data. The theoretical limit upon a photogrammetry system's accuracy is the precision to which the control point and image coordinates are measured. The extraction of image coordinates must therefore be optimised. Calibration objects must also provide control points whose location can be determined to a level of accuracy which is greater than that required by the photogrammetry application. In practice, this is achieved by surveying the locations of control point markers on a three-dimensional calibration object, or alternatively, by providing a two-dimensional calibration plate whose location can be either accurately positioned or measured. As indicated by the cross-section of studies outlined in Tables 2.2 and 2.3, several calibration object design variants have previously been investigated.

Owing to the large calibration volumes involved, biomechanical applications tend to opt for a rigid three-dimensional framework onto which control point markers are either painted or attached (refer to Table 2.2). For this type of frame-based application, control point locations are usually surveyed using an independent optical technique; Wood *et al.* **[30]** employed a metric camera, while Challis *et al.* **[31]** used a laser-based surveying system. In addition, Hatze **[19]** and Challis *et al.* **[31]** both report the extraction of image coordinates from enlargements of film negatives/positives, using a combination of a projector and a high resolution (0.025 mm is cited by Challis *et al.*) digitising tablet.

The five machine vision applications summarised in Table 2.3, all apply automatic image analysis techniques to extract the image coordinates of control point markers. In order to aid this process, a simple high contrast calibration image is required. The design of control point markers (shape, size etc.) has also been shown to play an important role in the accuracy of image analysis **[34, 35]**. Two-dimensional calibration objects, which incorporate black and white control point patterns, are therefore frequently used for this type of application. Alternatively, the method adopted by Champleboux *et al.* **[25]**, back-lighting a calibration plate containing an accurately drilled grid of holes, can be used to achieve the same effect. One of the many digital image analysis techniques discussed in the reviews by West *et al.* **[36]**

and Tian *et al.* [37], can then be used to extract the image coordinates of control point markers to sub-pixel accuracy.

2.3 X-Ray Photogrammetry.

Radiography is essentially a specialised form of photography in which X-rays are used instead of visible light. The X-ray image formation process is analogous to the back lighting photography technique, in which a silhouette or shadowgraph of an object is obtained by placing the object between the light source and the camera. However, as X-rays cannot be focused by lenses, the X-ray source must be as small as possible in order to ensure that the radiographic image has sharp edges. If the following assumptions are made, radiography can therefore be considered to be a special case of central projection, in which all rays pass through a perspective centre (i.e. the focus of the X-ray source):

- The X-ray source can be approximated to a perfect point source.
- X-rays travel in straight lines and are not deflected when passing through an object.

Mathematical interpretation of this imaging geometry is generally based upon the model in Figure 2.6, which depicts a divergent beam or "cone" of X-rays emanating from a point source, before passing through an object to the corresponding point on an image plane. As the image plane and perspective centre lie on different sides of the object, a magnified image is always produced. The extent of this magnification depends upon the ratio of the source-to-object and source-to-image distances, and as such, perspective distortion occurs with object points nearer the source experiencing more magnification than points on the opposite (image plane) side of the object.

The absence of lens distortion means than any of the linear calibration techniques originally developed for camera-based photogrammetry, including the DLT method, can be adapted to radiography applications. Alternatively, a simple vector geometry approach, which is comparable to the two-plane method, can be used to provide a mathematical solution to the central projection problem. X-ray photogrammetry therefore offers the physician an accurate, non-contacting method of measuring and

recording relative displacements within the musculoskeletal system. However, as X-rays are a form of ionising radiation, every image acquired during an investigation incurs an associated radiation dose to the patient. As a consequence, unlike camera-based photogrammetry, X-ray photogrammetry applications very rarely involve more than two exposures per investigation.

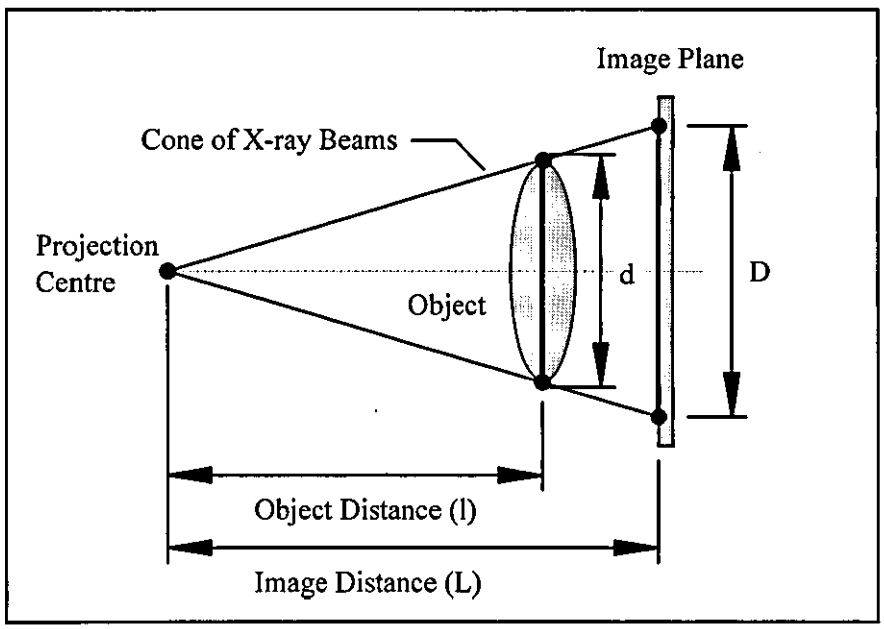


Fig. 2.6 : Model of the X-ray Image Formation Process:

$$\text{Image Magnification} = \frac{D}{d} = \frac{L}{l}$$

X-ray radiation was discovered in 1895 by the German born physicist Dr. Wilhelm Conrad Roentgen. The term *Roentgen stereophotogrammetry* is therefore frequently used to distinguish X-ray photogrammetry applications from their visible light counterparts. In the context of this section, Roentgen stereophotogrammetry is used to denote photogrammetry applications based upon radiography. That is to say, applications in which the X-ray images are recorded on photographic film. Digital X-ray imaging systems introduce additional distortions into the image formation process, and are therefore worthy of the separate discussion provided in Section 2.4.

Section 2.3.1 introduces the components of a typical radiography system. A more detailed discussion of the radiation physics associated with radiography is presented elsewhere, in Section A1.3.1 of Appendix 1. A critique of Roentgen

stereophotogrammetry research is then provided in Section 2.3.2. Clinicians have been performing three-dimensional reconstructions based upon radiographic images for almost one hundred years. However, it was not until the introduction of the Roentgen Stereophotogrammetry Analysis (RSA) method [38] in the early 1970's, that X-ray photogrammetry became a clinically accepted diagnostic technique. Section 2.3.2 is therefore restricted to a discussion of RSA and the subsequent developments in the area of Roentgen stereophotogrammetry.

2.3.1 Radiography Equipment.

Figure 2.6 is obviously an over-simplification of the true diagnostic radiography scenario. In reality, the acquisition of high quality X-ray images, obtained with the minimum possible amount of patient irradiation, requires the radiography system components depicted in Figure 2.7.

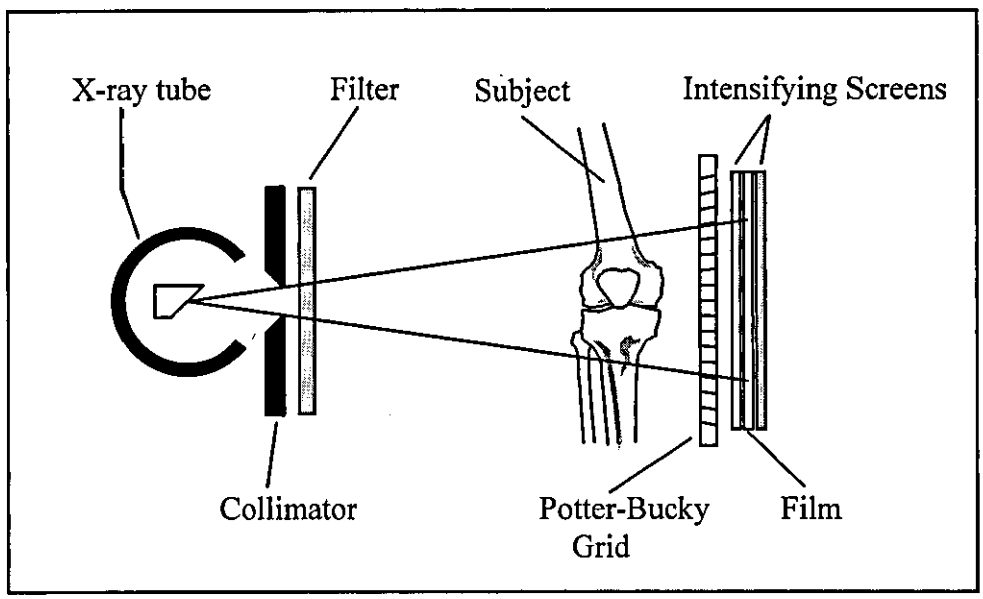


Fig. 2.7 : The Major Components of a Radiography System
(Adapted from [39])

X-rays are produced in a device known as an X-ray tube, which in turn is powered by a high kilovoltage generator (not shown in Figure 2.7). By controlling the X-ray tube voltage and current, a continuous spectrum of X-ray radiation may be obtained, whose penetrative properties are matched to the requirements of the particular radiographic examination. Collimation and filtration of the X-ray beam emerging from the tube is then performed, in order to reduce patient irradiation. Collimation involves the use of

lead shutters to limit the field of exposure, thereby ensuring that only the parts of the patient's anatomy that are to be imaged are irradiated. Filtration on the other hand, is necessary to remove the low energy component of the X-ray spectrum, which is unable to penetrate the patient, and does not therefore contribute to the imaging process. In practice this is usually achieved using an aluminium exit window in the X-ray tube housing.

As the human body is heterogeneous, the *primary* X-ray beam which leaves the X-ray tube housing, is subjected to a series of differential attenuation processes as it passes through the patient's anatomy. The amount of attenuation is related to the thickness, atomic number and density of the tissues in the path of the X-ray beam. For medical applications, the important practical consequence of this relationship is the fact that the attenuation coefficient of bone is significantly higher than that of soft tissue **[40]**. The attenuated or *secondary* X-ray beam, which emerges from the patient, therefore carries information about internal anatomical structures in the form of a distribution of X-ray radiation intensities.

Attenuation processes within the patient can cause X-ray beams to be deflected from their original trajectories. Unfortunately, anatomical information is only provided by X-ray beams which pass straight through the patient (i.e. beams which obey the collinearity condition). If this deflected or *scattered radiation* is allowed to reach the image plane, information which is not directly related to anatomy (i.e. noise) will therefore be introduced into the resulting image. Consequently, a device known as a Potter-Bucky grid must be located in front of the image plane, in order to improve image contrast. These grids consist of lead slats (separated by low-atomic number material) oriented in such a way that only X-rays travelling directly from the X-ray tube can be transmitted **[40]**.

As X-ray energy is not directly visible, all forms of radiology involve some form of image transduction **[41]**. Radiography applications capture the image using a high resolution photographic emulsion. However, as photographic film is relatively insensitive to X-rays, the film must be sandwiched between two fluorescent

(intensifying) screens in a light-proof box (film cassette). The film therefore effectively records the image formed at the surface of the intensifying screens, where X-ray photons are converted into visible light.

2.3.2 Roentgen Stereophotogrammetry.

Diagnostic radiography is a qualitative imaging technique aimed at providing non-invasive investigations of the musculoskeletal system. A standard radiographic examination involves the acquisition of an orthogonal image pair of the body site of interest. This pair of images then allows the physician to build up a three-dimensional "impression" of the patient's internal anatomy. However, the extraction of quantitative three-dimensional data from a conventional pair of radiographs is complicated by several factors:

- The (3-D to 2-D) perspective transformation.
- The image distortion problem (i.e. perspective distortion and film unflatness).
- The poor quality of X-ray images (i.e. superimposed structures **[42, 43]**).
- The lack of easily identifiable anatomical landmarks in the human body.

However, some of these problems can be overcome by the introduction of artificial calibration features into the X-ray images. Photogrammetric techniques can then be used to reconstruct the three-dimensional locations of imaged object points. For radiography-based applications, solution of the photogrammetric reconstruction problem requires that the following criteria are satisfied **[44]**:

- The orientation of the two film planes must be known in terms of a common coordinate system.
- The location of the two X-ray (point) sources must also be known in terms of the same coordinate system.
- Finally, the projected object points on each of the X-ray films must be determinable in terms of the common coordinate system.

Given that standard X-ray equipment does not allow accurate, repeatable positioning of the X-ray source with respect to the film plane, a purely analytical approach based upon known dimensions is not possible for X-ray photogrammetry. An initial

calibration process is therefore required to reconstruct the imaging geometry of the radiography system, using only the information contained in the radiographs themselves.

The simplest method of calibrating an X-ray photogrammetry system, is to radiograph a calibration object whose control points establish the required world coordinate system. If the imaging geometry of the X-ray system can be rigidly maintained, the calibration object and the patient can be imaged separately. However, as this is rarely the case, a calibration object design is required that allows control point markers to be superimposed into the diagnostic X-ray images, without obscuring the anatomical features of interest. Fortunately, the design of calibration objects for Roentgen stereophotogrammetry is much simpler than the corresponding problem for visible light applications.

The desired calibration/registration image can be obtained by surrounding the anatomical region of interest within a Perspex framework, into which a regular array of metallic control point markers have been inlaid. Perspex having a relatively low density (1190 kg/m^3) provides an X-ray image of hardly recognisable contrast, and is therefore said to be *radiolucent* or *radiotransparent*. It can also be shown that refraction of the X-ray beam as it passes through a Perspex plate is insignificant (typically 0.1 μm) **[38, 45]**. Accordingly, it is possible to construct a calibration framework which is not visible in the X-ray image, and does not create any image distortion. Metallic or *radiopaque* control point markers mounted into this framework at known locations, will however be superimposed into the diagnostic radiograph when the patient and frame are imaged together.

Compliance with the standard radiographic examination process requires a photogrammetric set-up similar to that shown in Figure 2.8. In order to allow the acquisition of calibrated front-to-back (*anteroposterior* or AP) and side-to-side (*lateral*) projections, the patient must be surrounded by a rectangular calibration frame consisting of two pairs of parallel plates. As an orthogonal pair of film cassette holders are incorporated into the calibration frame design, this arrangement is

generally referred to as *biplane X-ray photogrammetry*. Ideally, two X-ray tube/film cassette systems should be used to acquire the orthogonal radiographs by simultaneous exposure. If this is not possible, some form of patient positioning (immobilisation) device may have to be used to ensure that the patient does not move between image acquisitions.

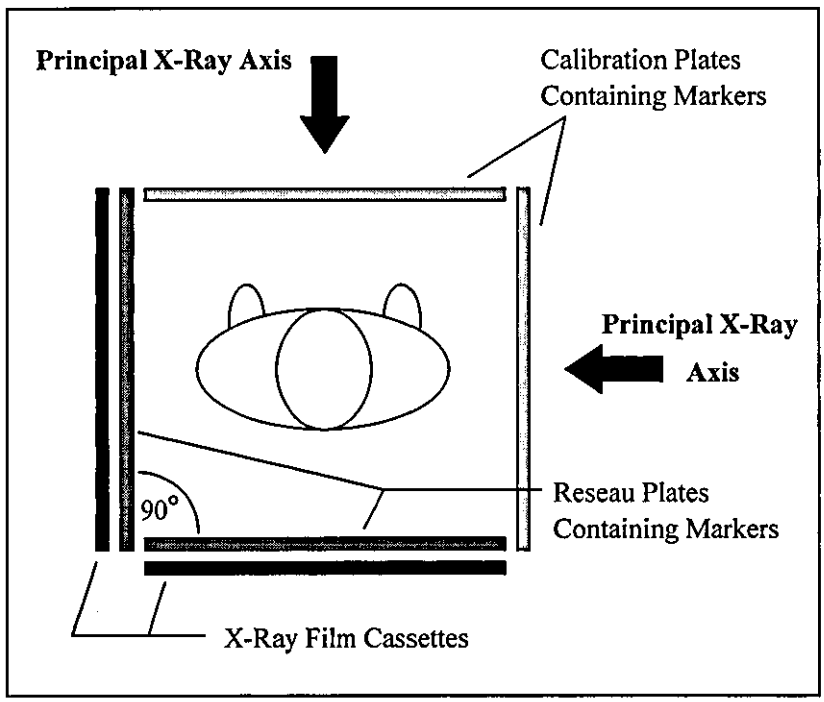


Fig. 2.8 : Biplane X-ray Photogrammetry
(Adapted from Lippert *et al.* [44])

Surrounding the anatomical region of interest with a rectangular calibration frame is not always the most practical option. For this reason, the alternative photogrammetry configuration shown in Figure 2.9 is frequently used instead of biplane photogrammetry. This technique employs oblique radiographic views (i.e. off-angle from AP or lateral) of the patient, resulting in a convergent imaging geometry requiring only a single image plane. The term *stereo X-ray photogrammetry* is therefore generally used to describe this arrangement. As only a single pair of parallel calibration plates are required, large volumes of the patient, such as the spine/torso and pelvis/hip areas, can be accommodated by this technique. However, as the standard diagnostic views are no longer being used, the stereo arrangement can require the acquisition of additional radiographs.

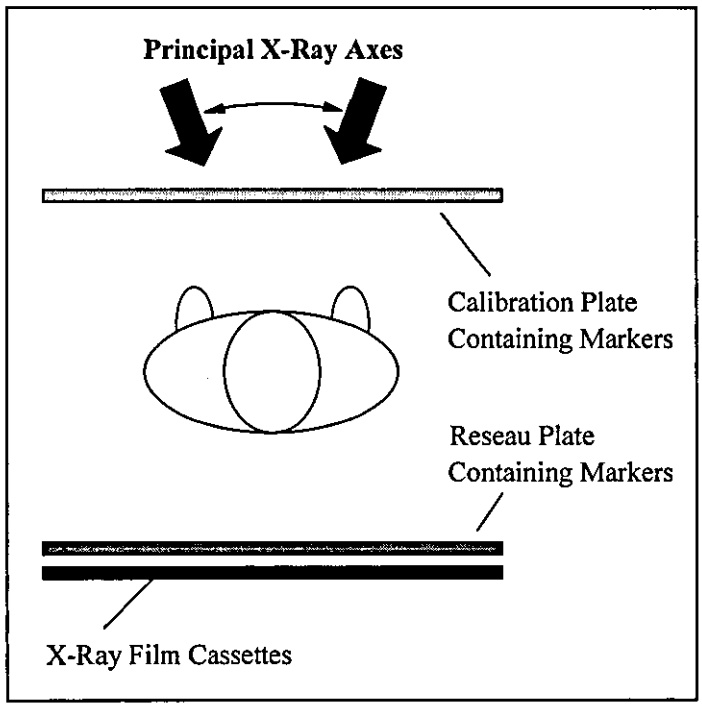


Fig. 2.9 : Stereo X-Ray Photogrammetry.
(Adapted from Lippert *et al.* **[P39]**)

The world coordinate system defined by the calibration frame, allows photogrammetric calibration and reconstruction to be performed using any of the optical photogrammetry techniques (i.e. DLT, two-plane method, etc.). However, as a result of the popularity of the Roentgen stereophotogrammetric analysis technique, which is discussed in the following section (2.3.2.1), most radiography-based photogrammetry applications adopt a straightforward geometrical solution to the central projection problem.

2.3.2.1 Roentgen Stereophotogrammetric Analysis.

The term *Roentgen stereophotogrammetric analysis* (RSA) is used to describe a specialised form of X-ray photogrammetry, which achieves excellent levels of precision (better than 0.2 mm) by implanting radiopaque markers into the radiographed bone structure. When imaged, these markers provide distinct measurement points, thereby avoiding the image matching problems associated with anatomical landmarks. Determining the positions of at least three of these implanted markers, allows a bone to be modelled as a rigid body. RSA therefore offers an

accurate method of quantifying the rotation of one body part in relation to another (i.e. skeletal kinematics).

The RSA technique was developed by Göran Selvik (Department of Anatomy, University of Lund, Sweden) between 1972-74 **[38]**. The original aims of Selvik's research was to develop a clinically usable Roentgen stereophotogrammetry method which would allow:

- High accuracy determination of radiopaque marker positions in an object.
- Determination of the movements of skeletal segments in terms of rotations and translations.

Existing knowledge of X-ray photogrammetry, kinematic analysis, and implantation techniques, was therefore used to develop a computerised system that would allow any body part to be investigated using standard radiography equipment **[46]**.

Selvik's RSA investigation technique can be broken down into four main parts **[47]**:

- Implantation of tantalum markers.
- The radiographic examination.
- Measurement of the radiographs.
- Photogrammetric calibration and reconstruction.

The insertion of spherical tantalum markers (diameter 0.5, 0.8, or 1.0 mm) into the bone is performed intraoperatively, or percutaneously under local anaesthesia, using a specially designed spring-release implantation instrument. Tantalum has excellent corrosion resistance and biocompatability properties, and as such, is ideally suited to permanent implantation into the human body. This metallic element's relatively high atomic number (Z = 73), also ensures that the spherical implants are easily identifiable in radiographic images. Although rigid body analysis theoretically requires only three markers per bone, in practice, five or more well spaced markers are actually implanted. The additional markers improve reconstruction accuracy and also provide a degree of redundancy.

Selvik's original account of the RSA technique describes two test (or calibration) "cage" designs that allow simultaneous calibration and patient examination **[38]**. Test cage "1A" supports the stereo X-ray photogrammetry configuration, while test cage "2" provides the four calibration planes required for biplane X-ray photogrammetry. Both of these cages were constructed from Plexiglas plates, with nine tantalum control point markers (0.5 mm diameter pins or balls) being used per plate. A third test cage ("1B") design, aimed at examinations of larger objects, was also described. This test cage was constructed from glass plates, onto which tantalum control point markers were glued, and facilitated a special case of stereo X-ray photogrammetry in which the cage and patient are imaged separately.

Routine clinical use has obviously led to refinements of the original RSA radiographic examination process. According to some of the more recent review articles **[46 - 48]**, the biplane calibration configuration still remains the method of choice for examinations of the extremities (i.e. the skull, knee, ankle etc.). However, several variants upon the stereo calibration configuration have subsequently been developed for large anatomical volumes which cannot easily be examined inside a test cage (i.e. hip, spine etc.). Selvik **[46]** and Kärrholm **[47]** both refer to the use of a modified stereo calibration cage situated below the examination table. Although strictly speaking, the use of this calibration set-up means that the patient is not actually situated within the volume defined by the test cage (i.e. the "calibrated" volume), due to the linear nature of the X-ray image formation process, RSA reconstructions based upon extrapolated data should not incur large errors. The second frequently quoted stereo technique involves radiographing the patient and the test cage separately. To improve the accuracy of this method, an additional reference plate containing tantalum control points markers is placed in front of the film cassette. The cage and the reference plate are radiographed, and thereafter the patient and the plate **[47]**.

Once the radiographic examination process has taken place, the film coordinates of the control points and implanted markers must be extracted. This process is performed manually, using a measuring table design that incorporates a television camera with a magnification factor of about twenty **[46]**. As the radiographic image itself is already magnified, measurements can be extracted to a precision of 5-25 microns. These

measured image coordinates, along with the known/independently determined real-world control point coordinates, then form the input data for the photogrammetric reconstruction process. RSA adopts a three-stage geometrical solution to the reconstruction of implanted marker positions:

- Determination of the Projective Transformation.

  As previously mentioned, a pre-requisite of Roentgen stereophotogrammetry is the ability to transform measured film coordinates into the world coordinate system established by the test cage. The required image-to-world projection transformation is obtained by ensuring that the "base" plate of the test cage is placed as near as possible to the film cassette. The optical photogrammetry term *reseau* plate is frequently used to distinguish the plate nearest the film cassette. Alternatively, Selvik uses the term *fiducial* markers to describe the control points in this plate **[38]**. By arbitrarily designating one of these fiducial markers to be the origin of the frame/world coordinate system, the required linear relationship between the frame and film coordinate systems may be obtained from four or more corresponding film/frame coordinate point-pairs.

- Determination of the X-ray Source Positions.

  The technique used to reconstruct the two X-ray source locations is summarised in Figure 2.10. More specifically, this figure shows the estimation of the left X-ray source location (S_L), of a stereo X-ray photogrammetry set-up, from the spatial intersection of a bundle of four reconstructed X-rays. However, the same technique is equally applicable to the biplane arrangement, and the process can theoretically be achieved using only two reconstructed rays. Nevertheless with reference to Figure 2.10, the first stage of the process is the selection of two or more control point markers ($P_{1\ calib}$ to $P_{4\ calib}$) of known world locations in the upper calibration plate. The coordinates of the corresponding points in the film plane ($P_{1\ film}$ to $P_{4\ film}$) are then measured, and transformed into reseau plate points (P_{1R} to P_{4R}) using the projective transformation. Having established four real world calibration/reseau plate point-pairs ($P_{1\ calib}$ & P_{1R} etc.), the three-dimensional X-ray beams passing through these point-pairs can be reconstructed. According to the collinearity condition all of these reconstructed X-ray beams must

pass through the X-ray source. A least-squares multi-line intersection method can therefore be used to estimate the location of (S_L).

- Determination of Object Coordinates.

  Once calibration of the RSA imaging geometry is complete, photogrammetric reconstruction of the implanted marker locations is performed using the two-line spatial intersection technique summarised in Figure 2.11. A target point (T) is imaged at film plane locations $T_{R\ film}$ and $T_{L\ film}$. The measured film coordinates of these points are then transformed into the corresponding reseau plate points ($T_{R\ reseau}$ and $T_{L\ reseau}$) using the relevant projective transformations. As the X-ray source locations (S_L & S_R) have previously been estimated, the two X-rays passing through (S_L & $T_{L\ reseau}$) and (S_R & $T_{R\ reseau}$) can be reconstructed. In practice, measurement errors and the finite size of the X-ray source, ensure that the two reconstructed X-rays do not in fact intersect. The point of nearest approach between the two reconstructed X-rays must therefore be used to estimate the target point's location.

Once the locations of the individual implanted markers are known, kinematic analysis can take place. The reported accuracy for this process using the RSA technique (10-250 μm & 0.03-0.6°) is an order of magnitude better than that of conventional radiography (1-5 mm & 1-6°) **[47]**. As a consequence, RSA is currently the most accurate radiography-based method available. In addition, extensive clinical experience has shown RSA to be both reliable and safe; during the period 1973-1990 Selvik reports that approximately 2000 examinations were conducted at Lund without complications **[46]**. Several independent research sites have therefore developed Roentgen stereophotogrammetry systems based upon Selvik's original description of RSA.

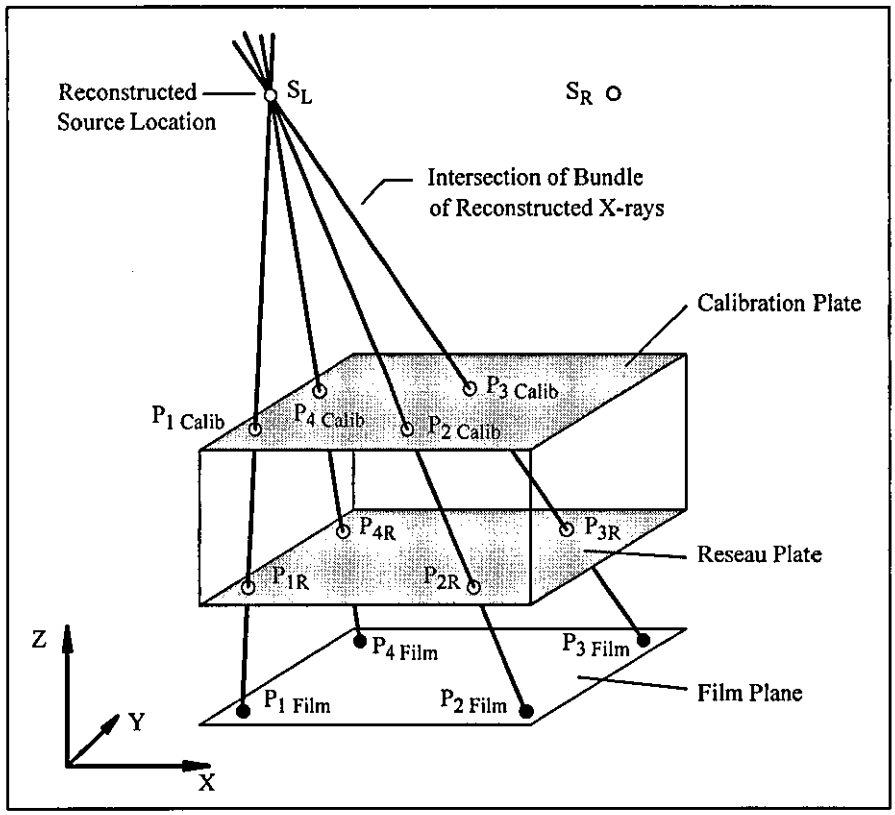


Fig. 2.10 : Reconstruction of X-ray Source Location

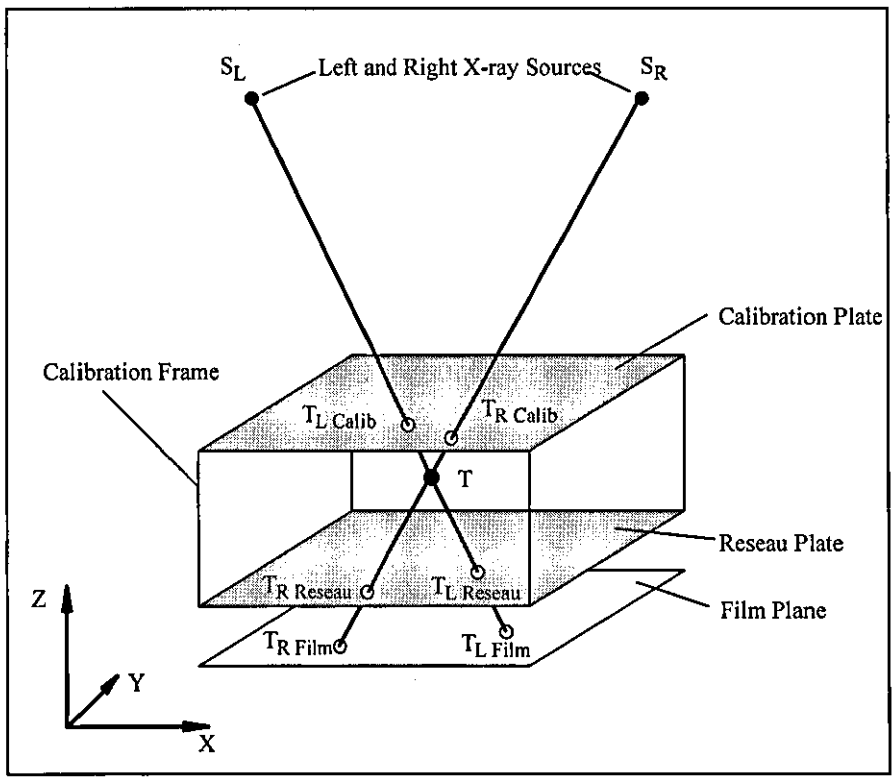


Fig. 2.11 : X-Ray Photogrammetric Reconstruction.
(Adapted from Fraser *et al.* [49])

2.3.2.2 Subsequent RSA Research.

RSA provides an accurate method of studying joint kinematics and stability *in vivo*. Functional anatomical investigations of the knee **[50]**, ankle **[51]**, wrist **[52]** and shoulder **[53]** joints, are therefore frequently reported in medical journals. However, in addition to these conventional kinematic studies of the musculoskeletal system, RSA has also found a number of applications which are specific to X-ray techniques **[39]**. Unlike conventional radiography, due to the use of permanently implanted reference markers, RSA allows an accurate series of postoperative follow-up examinations to be performed over a period of years. RSA can therefore be used to study skeletal growth **[54]** and bone healing **[55]** processes, and as such, is a valuable tool in orthopaedics, orthodontics and craniofacial surgery. The most successful orthopaedic applications of RSA are undoubtedly the study of prosthetic fixation, and to a lesser extent studies of the spine. The remainder of this section therefore concentrates upon these two application areas.

- **RSA Studies of Prosthetic Fixation:**

RSA evaluation of prosthesis fixation requires the implantation of radiopaque markers at the time of surgery. In addition to the bones on either side of the joint, radiopaque markers are also inserted into the plastic components of the prosthetic implant. Easily identifiable landmarks on the profiles of metallic prosthesis components, can also be used to derive the implant's location. As the metallic implant may obscure some of the implanted bone markers in the radiographic images, it is important to provide a degree of redundancy in the marker distribution. The initial postoperative RSA examination is conducted a few days after the arthroplasty procedure. This "reference" examination is performed without any load being applied to the artificial joint, and provides an initial estimate of the implant's location with respect to the surrounding bone segments. Subsequent examinations, conducted at regular intervals for a period of up to two years, are performed with full weight bearing on the joint. Comparison of the results from these follow-up examinations with those of the reference examination, allow determination of the migration of the implant (rigid body) with respect to the bone (rigid body).

The 1992 review paper by Ryd **[48]**, gives a good overview of the total hip replacement (THR) studies conducted by Selvik *et al.* at Lund. This research work was instigated in 1976 by Baldursson *et al.* **[56]**, who applied RSA to the measurement of acetabular component migration. Using variations upon the stereo X-ray photogrammetry configuration, with a 30-40° angle between the two X-ray beams, RSA is reported to yield a reconstruction accuracy of better than 0.2 mm for THR applications **[48]**. Subsequent research by Mjöberg *et al.* **[57]**, aimed at investigating methods of improving implant/bone attachment, has established a relationship between prosthetic loosening and the heat damage which occurs within bone during the polymerisation of bone cement.

During the late 1970s, stereo X-ray photogrammetry systems were also developed by Baumrind *et al.* (University of California at San Francisco, USA) **[58]** and Veress *et al.* (University of Washington, Seattle, USA) **[44, 49]**, in order to detect prosthesis loosening. More recently, an RSA system has also been established by Turner-Smith *et al.* at the Nuffield Orthopaedic Centre in Oxford, with the aim of measuring the migration of femoral and acetabular THR components **[43, 59]**.

RSA has also been used to evaluate prosthetic fixation after total knee arthroplasty (TKA). Selvik reports that during the 1980s, over four hundred RSA studies of TKA were performed at Lund **[46]**. In the majority of these investigations tantalum markers were inserted at the time of surgery, into both the tibia and the plastic tibial component. The artificial knee joint was subsequently radiographed using a biplane X-ray photogrammetry configuration. A series of follow-up RSA examinations were then undertaken over a period of years, in order to study migration of the tibial component. More recent studies have evaluated new cementless tibial prosthesis designs, which aim to increase bone ingrowth into the porous implant surfaces by improving initial fixation and hence minimising micromotions **[60]**.

- **RSA Studies of the Spine:**

As most of the thirty-two vertebrae comprising the spine provide anatomical features which are easily recognisable over a wide range of projections (i.e. radiographic landmarks), the spine is an ideal candidate for both conventional X-ray

photogrammetry and RSA. Many of the early X-ray photogrammetry investigations were targeted at the diagnostic interpretation of scoliosis (abnormal curvature of the spine) **[61 - 65]**. Using three or more anatomical landmarks per vertebra, the relative locations or kinematic motions of a number of vertebral rigid bodies can be determined. By performing a series of follow-up examinations, it is therefore possible to monitor the progress of spinal abnormalities in response to surgical or conservative corrective treatments. Similar RSA-based evaluations of skeletal kinematics have also been reported by Lee *et al.* (University of Iowa, Iowa City, IA, USA) **[66]** and Olsson *et al.* (University of Lund, Sweden) **[67]**.

The other major application area for RSA of the spine, is in the treatment of low-back pain. As this condition is often attributed to instability between vertebral segments, a spinal fusion procedure is frequently performed as a pain relief measure. This type of surgical procedure achieves stabilisation of two of more vertebrae, through the use of an external mechanical fixator. In order to evaluate the effectiveness of such fixations, Olsson *et al.* (University of Lund, Sweden) **[68]** have used RSA to investigate the mobility of vertebral segments after spinal fusion.

2.4 Digital X-Ray Photogrammetry.

In spite of the obvious benefits offered by Roentgen photogrammetry, clinical acceptance of the technique has been somewhat limited. Due to the need to initially develop and then manually digitise the film using specialist equipment, film-based photogrammetry methods such as RSA can be very time-consuming (up to two hours) and expensive. These inherent limitations also preclude the use of Roentgen photogrammetry techniques from intraoperative applications. In order to make the progression from the current research applications into widespread routine clinical usage, improvements must therefore be made with respect to ease of use and the time taken to perform Roentgen photogrammetric investigations. Such improvements can potentially be achieved through the use of digital X-ray imaging.

The simplest method of producing a digital X-ray image is to use either a video camera/frame grabber combination or a laser film scanner, to digitise a conventional radiograph. However, as this approach still incurs the delay caused by the film development process, more direct methods of obtaining a digital radiograph are preferred. Several digital X-ray imaging systems, which are based upon "film-replacement" transducer technology, have therefore been developed **[69]**. One of the more promising digital systems, generally referred to as photostimuable phosphor computed radiography (PPCR), allows the direct acquisition of a digital X-ray image using standard radiography equipment **[70 - 72]**. However, as state-of-the-art techniques such as PPCR are just beginning to find their way into hospitals, the most widely used method of directly obtaining digital X-ray images is to digitise the video signal from a fluoroscopy unit. The remainder of this section, therefore, concentrates upon photogrammetry applications which are based upon digital fluoroscopy, and as such, the concepts and terminology associated with fluoroscopic equipment are initially introduced in Section 2.4.1.

Photogrammetric reconstruction using fluoroscopy equipment is a far less attractive proposition than conventional Roentgen photogrammetry. Due to the use of television equipment in the imaging process, the spatial resolution of a fluoroscopic image is poor, in comparison to that obtained by film-based radiography. This imposes an obvious limit upon the reconstruction accuracy which can be obtained. In addition to the perspective distortion resulting from central projection, which is experienced by all forms of radiography, fluoroscopic images are also further distorted during the image transduction process. Fluoroscopy systems use image intensifier tubes to produce dynamic, real-time X-ray images. As discussed in Sections 2.4.2 and 2.4.3, these image intensifier tubes have inherently non-linear imaging properties, and are also susceptible to the influence of strong magnetic fields. The extraction of accurate three-dimensional data from two or more fluoroscopic images of an object, therefore requires additional "image dewarping" or internal calibration procedures to be performed, in order to correct the resulting geometric image distortion.

Owing to these limitations, fluoroscopy-based photogrammetry has mainly been restricted to real-time radiography procedures. The most notable of these application areas is angiography, which involves real-time visualisation of blood vessels after the injection of a radiopaque contrast medium. Accordingly, Section 2.4.4 presents a critique of previous quantitative three-dimensional angiography research.

2.4.1. Fluoroscopy Equipment.

The generic term *fluoroscopy* is used to describe radiographic techniques which allow the visualisation of dynamic X-ray images. Traditional fluoroscopy employs a simple fluorescent screen to produce instantaneous X-ray images. These screens consist of a plastic base coated with a thin layer of fluorescent material, which in turn is mounted onto a lead-glass plate **[41]**. As the lead-glass provides adequate radiation protection, the optical image can be viewed directly through the glass plate. Unfortunately, the low brightness of the resulting image means that it can only be viewed in a darkened room. Modern fluoroscopy systems therefore employ a device known as an X-ray image intensifier tube in order to produce an image whose brightness is substantially higher than that of a simple fluorescent screen.

As shown in Figure 2.12, the modern X-ray image intensifier is an electron-optical system housed within a vacuum tube. The main components of this system are an intensifying screen at either end of the tube, a photocathode, and a series of focusing electrodes. These components achieve the required brightness gain by converting incident X-ray radiation into electrons, which are then accelerated by a potential difference, prior to being converted into visible light. The overall effect of these conversion processes it to increase the number of light quanta produced for a given amount of X-ray radiation (as compared with a simple fluorescent screen). The brightness gain produced by the tube is further increased, due to the fact that the output image is much smaller than the input image, thereby concentrating the light quanta into a smaller area. An output image which is approximately 5000 times brighter than that of a conventional fluorescent screen, is therefore produced as a result of the electron acceleration and image reduction processes **[40]**.

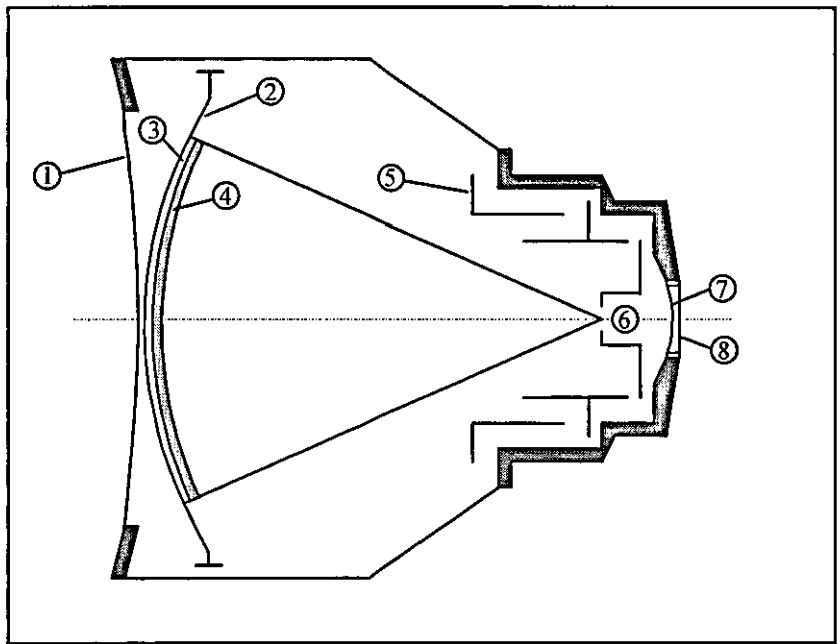


Fig. 2.12 : Components of a Modern Image Intensifier Tube
(Adapted from Garrett *et al.* [41])

A more detailed breakdown of the conversion processes, within an image intensifier tube, is provided by the schematic shown in Figure 2.13. Looking initially at the input side of the tube, X-rays entering the tube housing initially pass through a high transmission input window. Upon reaching the fluorescent input phosphor these incident X-rays are absorbed, resulting in the production of light quanta (i.e. photons). The photons then hit the photocathode which is in contact with the input phosphor, where they cause photoelectrons to be emitted. These photoelectrons are accelerated by an applied electrostatic field, and are focused by a "crossover" electron-optical system towards the anode aperture. After passing through this "electron focus", the high kinetic energy photoelectrons strike the output where they produce the final visible light image.

The minified fluoroscopic image which is formed at the output window of the image intensifier tube is too small to be of any clinical use; the diameter of the output screen is typically 2.5 cm [73]. Depending upon the imaging requirements of the diagnostic

application, the output image from an image intensifier is therefore either recorded on film (still or ciné), or viewed with the aid of television type technology. The latter option uses a lens system to focus the output image onto the photoconductive surface of either a television camera (Vidicon) tube or a CCD camera. The amplified output from the camera is then transmitted as a video signal and can be viewed in real-time on the screen of a conventional television monitor. A permanent record of the fluoroscopic examination can be obtained on video tape, or alternatively, by digitally sampling and storing the video signal.

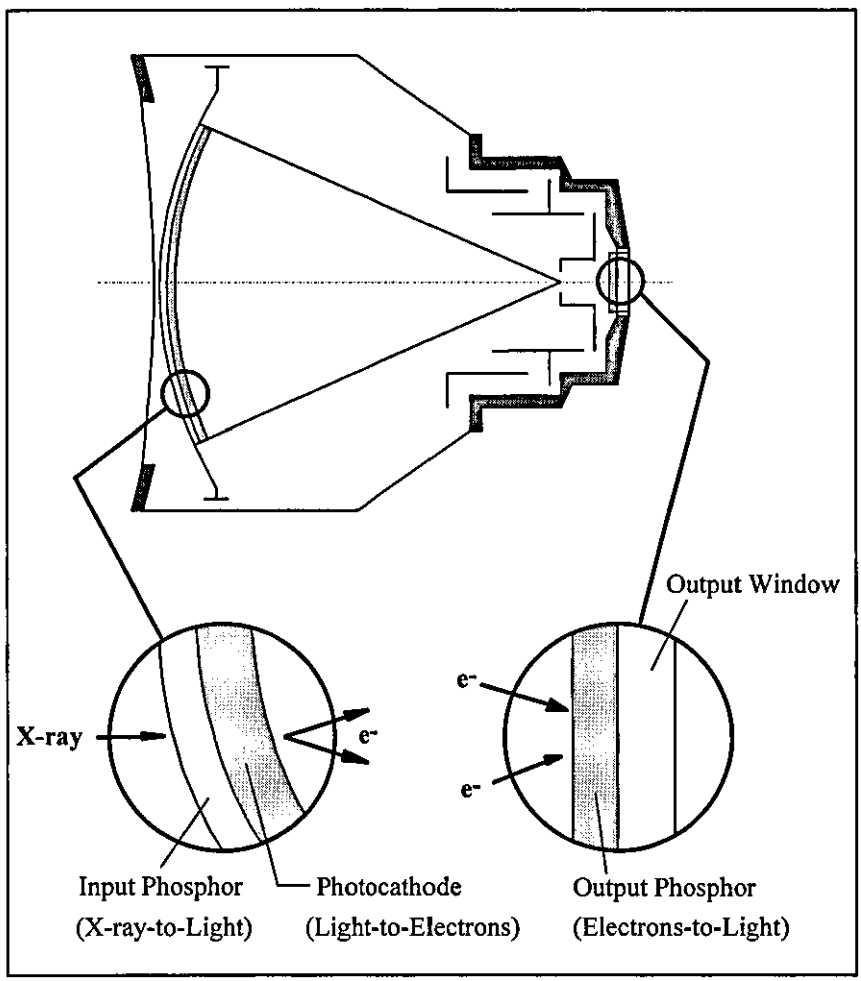


Fig. 2.13 : Schematic Representation of the Operation of an Image Intensifier Tube.
(Adapted from Vosburgh *et al.* [74])

The image intensifier/television imaging system is usually incorporated into a modern fluoroscopy system in the manner depicted schematically in Figure 2.14. In order to minimise patient irradiation, a technique known as *automatic dose-rate* (ADR) control

can be used to drive the X-ray generator. ADR uses exposure control circuits to adjust the X-ray tube settings (kV & mA), in response to changes in the intensity of the light leaving the image intensifier; as measured by a light sensor. Since the light intensity leaving the image intensifier is related to the transmission of X-ray radiation through the patient, ADR provides an effective means of obtaining satisfactory fluoroscopic images using minimal X-ray radiation.

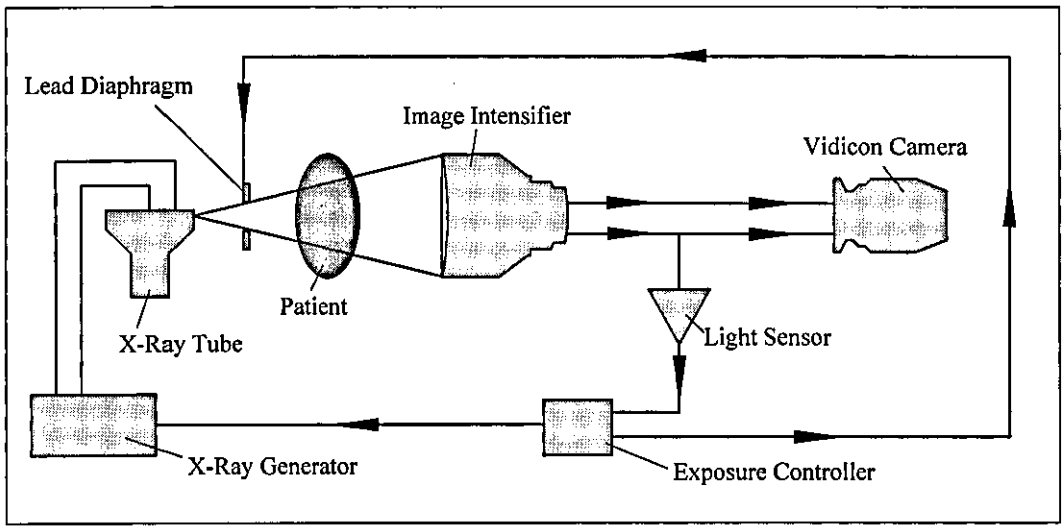


Fig. 2.14 : Schematic of a Fluoroscopy System
(Adapted from Garrett *et al.* **[41]**)

Fluoroscopy allows processes that involve either movement or flow (i.e. coronary angiography, investigations of the gastrointestinal tract etc.) to be studied. However, real-time fluoroscopic visualisation or *screening* obviously requires that the X-ray tube runs continuously. Although the high-gain of the image intensifier tube means that the doses associated with simple fluoroscopy procedures are relatively low, large accumulated doses can still result from prolonged screening times. A technique referred to as *pulsed mode fluoroscopy* is therefore frequently used to limit the radiation output of modern fluoroscopy equipment. Pulsed mode fluoroscopy employs very short exposure times to produce a persistent image, which is updated incrementally approximately every tenth of a second. While the resulting images can sometimes be visually unsatisfying, pulsed mode fluoroscopy can lead to dose reductions of up to 70% **[75]**.

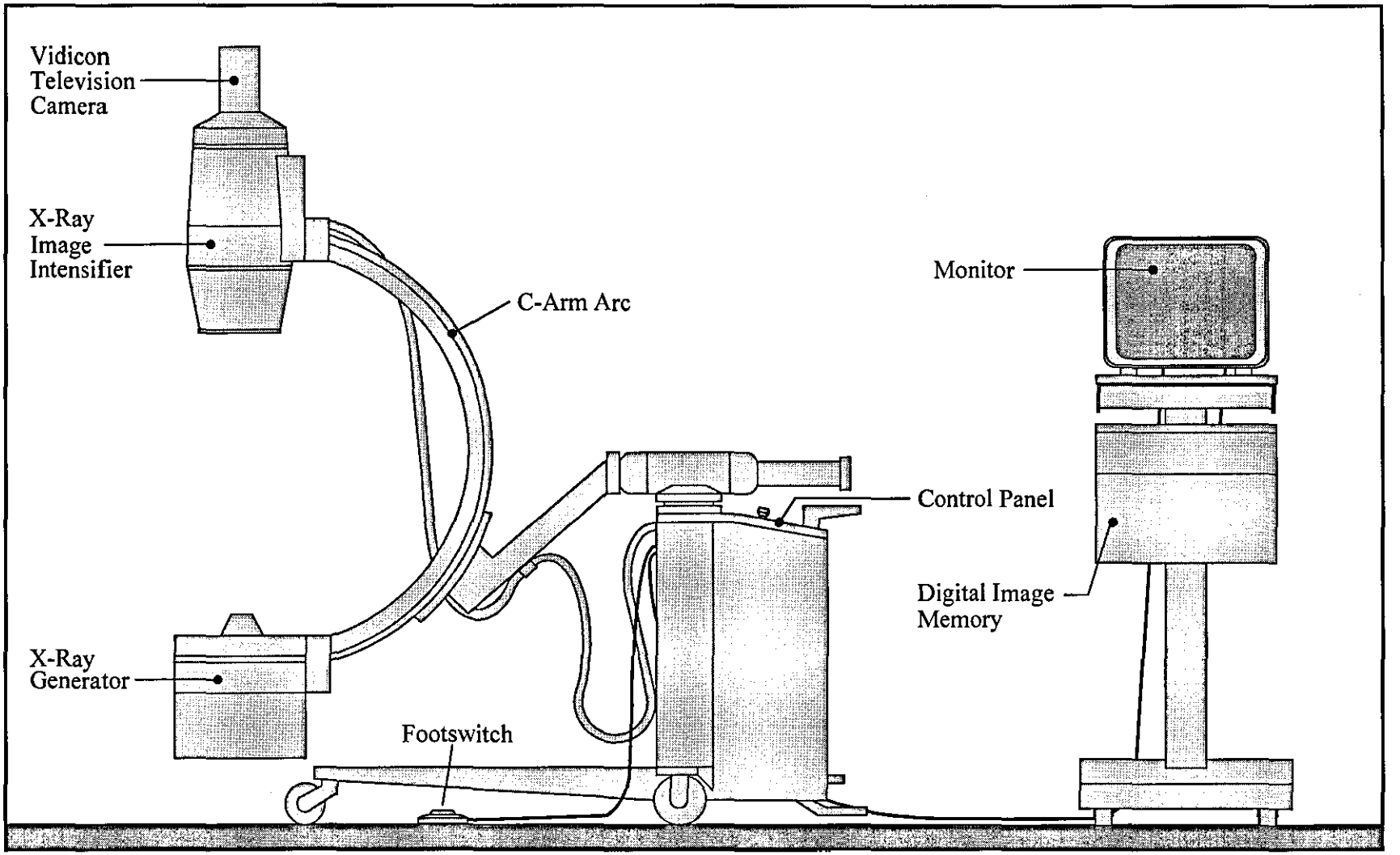


Fig. 2.15 : Mobile X-Ray Image Intensifier (Siemens Siremobil 2)

Fluoroscopy is also frequently used as the primary radiographic imaging technique during orthopaedic surgery procedures. Intraoperative fluoroscopy requires an X-ray imaging system that can be easily manoeuvred, facilitates a wide range of projection angles, and does not restrict the surgeon's access to the patient. In order to meet these requirements, several of the major manufacturers of medical imaging equipment have developed mobile image intensifier systems specifically for use in the operating room. As depicted in Figure 2.15, these systems employ a counterbalanced C-arm geometry mounted on a wheeled base unit. X-ray tube and image intensifier/TV camera housings are supported a fixed distance apart on opposite ends of the C-arm arc. The adoption of this arrangement, in conjunction with five positioning degrees-of-freedom, allows an infinite number of projection angles to be obtained once the mobile unit has been wheeled into position.

Mobile X-ray image intensifiers are operated from a control panel in the base unit. In order to reduce irradiation of the theatre staff, a footswitch can also be used to remotely control fluoroscopic exposures. In modern C-arm models, such as the Philips BV29 (Philips Medical Systems, Netherlands) and Siemens Siremobil 2000 (Siemens Medical Engineering Group, Erlangen, Germany), a CCD camera is used in combination with a digital image processor/image store, thus ensuring that the image intensifier/TV system is fully digitised. The results of continuous, pulsed, or snapshot (single-pulse) fluoroscopy examinations can therefore be enhanced prior to display on a dual-monitor mobile viewing station. Typically, one monitor is used to view the live fluoroscopic image, while the other displays a stored image. However, these features are not always provided by some of the older systems which can still be found in routine clinical use. Single monitor models, such as the Siemens Siremobil 2 depicted in Figure 2.15, are based upon Vidicon television camera technology, and have a very limited image memory capability.

2.4.2 Distortion Correction of X-Ray Image Intensifier Images.

The electron-optical system of an image intensifier tube employs an electrostatic lens design, in which the photocathode and anode are both curved spherically and have a common centre-point **[76]**. As a result, the electron "image" formed at the

photocathode is produced by central projection from the X-ray source onto a curved surface. These electrons are then accelerated and focused towards the output phosphor which, in order to be compatible with either a television or a film camera, is usually flat. The image produced by an image intensifier tube, therefore represents a projection onto a curved photocathode surface, followed by a mapping onto a flat output window. The outcome of this image formation process is a fluoroscopic image, which exhibits noticeable spatial distortion and associated intensity variations.

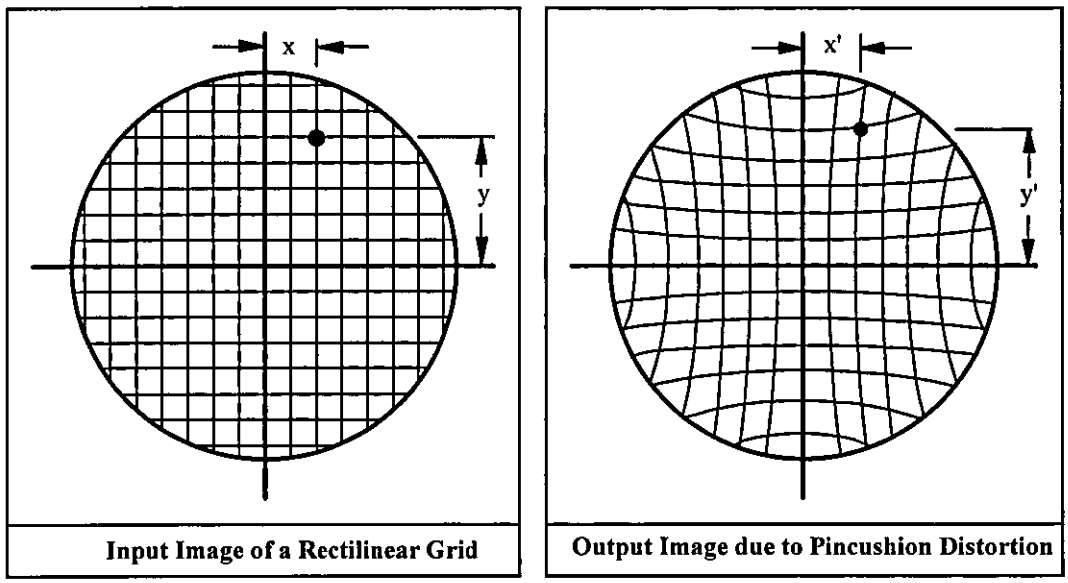


Fig. 2.16 : Pincushion Distortion Caused by an X-ray Image Intensifier
(Adapted from Rudin [77])

Imperfections in the image intensifier's electron-optical system, the optical lens/television camera system, or even a subsequent digitisation process, can all contribute to fluoroscopic image distortion. However, the spatial distortion which is observed in the displayed image is mainly attributable to the projection of an X-ray radiation distribution onto a curved photocathode surface. Image distortion due to photocathode curvature is generally referred to as *pincushion distortion.* This term is derived from the appearance of the image produced when a regular grid pattern is imaged by a fluoroscopy system, as shown in Figure 2.16.

In addition to spatial distortion, a fluoroscopic image also suffers from brightness non-uniformity, with the centre of the image being brighter than the periphery. This image degradation phenomenon is referred to as *vignetting,* and occurs as a direct

consequence of pincushion distortion. The increased magnification at the edges of the image means that a given area element (i.e. a grid square) is projected onto a larger area at the periphery of the image. The increased size of the area element leads to a corresponding decrease in intensity, since the X-ray beam is spread over a larger area.

Pincushion distortion and vignetting are of no great consequence during the majority of diagnostic fluoroscopy procedures. A standard fluoroscopic examination is a qualitative procedure aimed at establishing the presence, as opposed to the exact three-dimensional location, of anatomical abnormalities. However, for applications requiring the extraction of quantitative information, such as photogrammetry, the fact that the relationship between the world coordinates (x, y) and the corresponding image coordinates (x', y') of an imaged point is non-linear, creates a major problem.

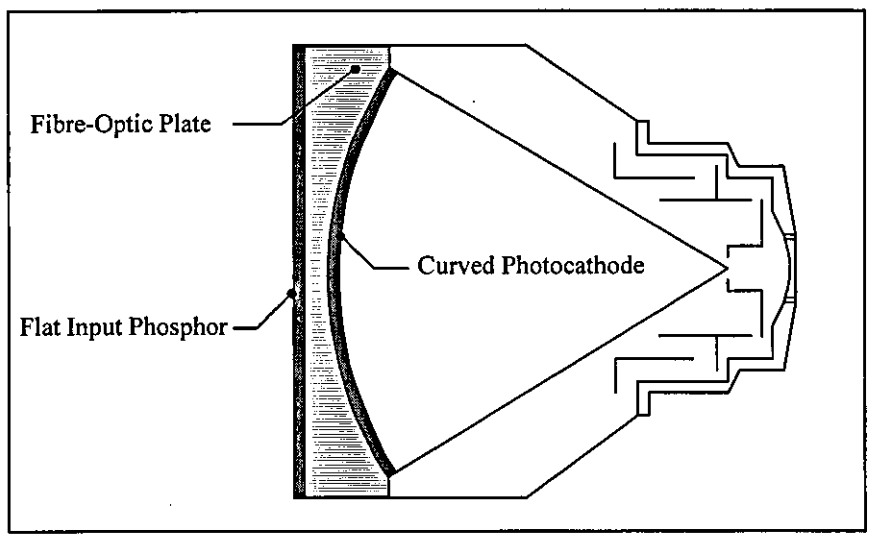


Fig. 2.17 : X-ray Image Intensifier with Fibre-Optic Entrance Window

One possible method of minimising the effects of pincushion distortion is hardware modifications based upon fibre-optic technology. As shown in Figure 2.17, a fibre-optic plate can be used to optically couple the curved photocathode surface to a flat input phosphor [78]. Although this arrangement mimics the flat image plane used in conventional film-based radiography, in practice, producing a fibre-optic plate which covers the entire input window is not a financially viable option. However, as the output window of the image intensifier tube is much smaller than the input window, correction of pincushion distortion at the output side of the image intensifier tube is a

more economical proposition. Consequently, a number of the more modern image intensifier systems, most notably models manufactured by Philips Medical Systems (BV25, BV29, etc.), employ a fibre-optic output window.

When using older fluoroscopy equipment to perform digital X-ray photogrammetry, additional software calibration techniques must be applied, in order to compensate for the effects of pincushion distortion and vignetting. The process of correcting a digital image's spatial (pixel coordinates) and grey scale (pixel values) integrity, is frequently referred to as *image dewarping*. Most image dewarping techniques work by establishing a relationship between the displayed (distorted) output image and a hypothetical (undistorted) input image. The adoption of this approach leads to an image coordinate correction scheme, which when applied, produces corrected image coordinates which correspond to those that would have been obtained by a film-based photogrammetry application. After application of the distortion correction algorithm, photogrammetric reconstruction therefore reverts to solution of the central projection problem, thus allowing the techniques developed for roentgen photogrammetry to be applied.

In order to establish a black box relationship between an imaging system's input and output, corresponding input/output image coordinate point-pairs are required. Unfortunately, in the case of the image intensifier/digitised video system, only data pertaining to the distorted output image is available. The undistorted input "image" data is therefore usually created by acquiring a reference image of a calibration target containing a repeating grid pattern. It is standard practice for this calibration target to be placed as near as possible to the image intensifier's input window, thus ensuring the sharpest possible image, and more importantly, simulating the image plane of a conventional film-based radiography system.

The use of a calibration target makes it possible to quantify the effects of pincushion distortion. A non-linear transformation that maps the measured image point coordinates of the distorted digital image, into the corresponding known control point locations of the calibration target, can therefore be estimated. A distortion correction

algorithm can then be implemented as either a global interpolation function (refer to Section 2.4.2.1), or a series of localised interpolation functions (refer to Section 2.4.2.2). Alternatively, as further discussed in Section 2.4.2.3, a purely geometrical relationship can be derived in order to relate output (distorted) and input (undistorted) image coordinates.

In addition to the numerical correction of pixel coordinates, it is also possible to reconstruct and display a true image in which the pixel values have also been corrected. Compensating for the effects of vignetting typically involves a grey level interpolation technique. For each pixel location in the corrected image, the corresponding location in the distorted digital image is calculated. The pixel value at this location in the digital image is then transferred to the corrected image. However, the calculated point in the distorted image will generally not lie precisely on an acquired pixel location (i.e. the result will not be an integer). The new pixel value must therefore be computed using a bilinear interpolation technique, which samples the grey scale values of the pixels in the immediate neighbourhood of the calculated point. Reconstruction of a "true" image in this manner, is frequently used as a means of demonstrating the effectiveness of a distortion correction scheme. Ideally, comparison of the "before" and "after" images of a calibration object's grid will show a pincushion pattern and a regularly spaced grid pattern.

2.4.2.1 Global Polynomial Correction Functions.

If it is assumed that image distortion is uniform (i.e. symmetrical) over the field-of-view, a global polynomial function can be used to estimate the desired relationship between output (distorted) and input (undistorted) image coordinates. The coefficients of such a relationship can be determined by imaging a standard calibration grid that provides a series of known control points throughout the entire field-of-view.

Image dewarping is frequently performed using a single parameter polynomial, which is based upon classical optical theory for image aberrations. The technique developed by Casperson *et al.* **[79]** (University of California, Los Angeles) is a prime example of this optics-based approach. Making the assumption that only pincushion distortion

was present, Casperson *et al.* reduced the image dewarping problem to the following relationship between measured (distorted) image points (x', y') and undistorted coordinates (x, y):

$$x' = x + D.x.(x^2 + y^2) \quad .....(2.12)$$

$$y' = y + D.y.(x^2 + y^2) \quad .....(2.13)$$

Where: D = An empirically determined distortion parameter.

In recent years, it has become more common to express this relationship as:

$$r' = r + D.r^3 \quad .....(2.14)$$

Where: r' = The distorted radial distance from the centre of the image intensifier.
r = The true radial distance from the centre of the image intensifier.

A subsequent study by Rudin *et al.* **[77]** (State University of New York at Buffalo), claims that Casperson et al's method can be outperformed by a two parameter (odd-power) polynomial technique, which takes into account both photocathode curvature and electron focusing effects:

$$r' = r + D_t.r^3 + E_t.r^5 \quad .....(2.15)$$

Where: D_t & E_t = Empirical distortion parameters.

More complex polynomial correction schemes have also been favoured by Fujita *et al.* **[80]** and Guggenheim *et al.* **[81]**, both of whom applied a fifth order polynomial. The digital correction unit developed by Haaker *et al.* **[82]** (Philips GmbH Forschungslaboratorium, Hamburg, Germany) also employs a high order polynomial correction scheme. This system is initially calibrated in a pre-processing step, with a third order polynomial model being applied to the usual calibration grid image:

$$x_i = f_0 + f_1.y_s + f_2.x_s + f_3.y_s^2 + f_4y_s.x_s + f_5.x_s^2 + f_6.y_s^3 + f_7.y_s^2.x_s + f_8.y_s.x_s^2 + f_9.x_s^3$$

$$y_i = g_0 + g_1.y_s + g_2.x_s + g_3.y_s^2 + g_4.y_s.x_s + g_5.x_s^2 + g_6.y_s^3 + g_7.y_s^2.x_s + g_8.y_s.x_s^2 + g_9.x_s^3$$

.....(2.16) & (2.17)

Where: (x_j, y_j) = Distorted input image coordinates.

(x_s, y_s) = Undistorted output image coordinates.

The coefficients of this model, determined by least-squares fitting, are then programmed into an address memory, which acts as a look up table during the subsequent correction stage. Using this hardware system, it is reported that a digital image (512 x 512 x 8 bit) can be corrected and displayed within 100 milliseconds, thus allowing real-time analysis.

2.4.2.2 Local Affine Correction Functions.

In reality, image distortion is often found to be non-uniform across the field-of-view. As a consequence, the application of a globally derived polynomial correction scheme can occasionally lead to unacceptably high errors. One possible solution to this problem is the modified polynomial method proposed by Baltzopoulos (Department of Movement Science, University of Liverpool) **[83]**. This technique introduces non-negative weighting functions for each calibration point, thus ensuring that only control points within an application specific threshold distance of the target point are used by the polynomial correction scheme. However, a more popular alternative to the global polynomial method is provided by the use of a mosaic of localised correction functions.

The localised correction technique was originally developed by Butler *et al.* **[84]** (Oregon State University) for industrial machine vision applications. The aim of this research was to develop a distortion correction (internal calibration) scheme that would allow easier factory floor metrology using off-the-shelf imaging equipment. This goal was achieved by adopting a *correction-calibration approach* involving separate distortion correction and external calibration processes. A correction function is initially applied to make small adjustments to the measured (distorted) pixel coordinates (d). Photogrammetric reconstruction can then be performed by applying a simple linear pixel-to-world transformation to the resulting corrected pixel coordinates (c).

As the coefficients of the correction function can be derived off-line, and the post-correction pixel-to-world transformation becomes linear, the actual in situ factory floor calibration procedure is greatly simplified. In practice the correction function consists of a number of locally affine (i.e. a polynomial of degree n = 1) functions pieced together, to produce a continuous "global" function. The control points are then used to sub-divide the image space into a regular pattern of cells, with each cell being split into two triangles called a *simplex.* The simplex pattern is then 'superimposed' on the distorted image, allowing corresponding control points and distorted points to be matched, as shown in Figure 2.18. A six parameter linear transformation function, which maps the three distorted points associated with a simplex onto their corrected positions, can then be calculated. Once all such simplex transformations are known, any uncorrected image point (d) can be mapped into a corrected point (c), following the identification of the particular simplex it lies within.

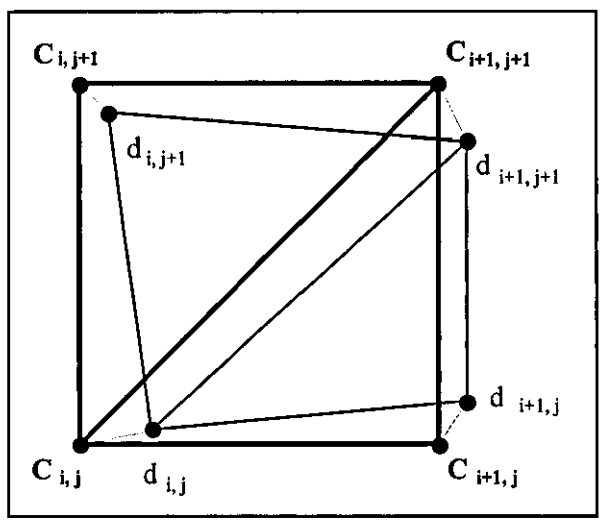


Fig. 2.18 : Superimposed True (C_{ij}) and Distorted (d_{ij}) Control Point Cells.
(Adapted from Butler *et al.* [84])

The local affine function technique has subsequently been applied by Boone *et al.* [85] (Thomas Jefferson University, Philadelphia, PA) as part of a step-by-step correction scheme, aimed at correcting the individual errors arising during the propagation of the image signal through the image intensifier/TV digitiser system. By imaging a rectilinear grid pattern, a best-fit undistorted grid was calculated, based upon the mean separation of a row of control points at the centre of the image. Nearest-neighbour

comparisons were then used to match the best-fit grid points to their corresponding points in the distorted image. Localised affine transformation functions could then be calculated by partitioning the image into a series of triangular regions, each being bounded by three control points.

Reimann *et al.* **[86]** (Henry Ford Hospital, Detroit, MI) have also developed a piecewise affine transformation approach, with a view to hardware-based automatic correction of X-ray image intensifier distortion (both pincushion and vignetting). Once again, a reference image of a calibration grid was broken up into triangular regions, in order to derive the transformation between control point image and real-world coordinates. However, the form of the affine function adopted differed slightly from that of Boone *et al.* **[85]** and Butler *et al.* **[84]**, with Reimann *et al.* choosing to fit the three control points to a unique plane.

Although the localised approach undoubtedly provides a more satisfactory solution to the image intensifier distortion problem, the improvements which can be gained, with respect to the results of global polynomial methods, are achieved at the expense of substantial increases in computational effort. Six transformation parameters must be stored for each triangular sub-region of the image. Additional extrapolation functions are also required for areas near the edges of the image which are not bounded by three control points. Before correction can take place, it is also necessary to identify which triangular region corresponds to the point whose coordinates are to be corrected.

2.4.2.3 The Mathematical Modelling Approach.

In addition to empirically determined (global and local) correction schemes, several analytical expressions have also been proposed as solutions for the image intensifier distortion problem. These geometrical models usually describe distortion due solely to projection of the X-ray image onto the curved input phosphor (i.e. pincushion distortion), and therefore make the assumption that the other sources of distortion can be safely neglected. Further assumptions must also be made with respect to the shape of the photocathode, which is typically modelled as a sphere. Nevertheless, the analytical approach remains a popular choice for correcting the distortion associated

with gantry-based fluoroscopy and angiography systems. The imaging geometry of these systems is highly constrained, with a fixed examination table and a constant source-to-image (SID) distance usually being employed. The angular settings of the gantry are also frequently instrumented, thus allowing accurate determination of the angles required by the mathematical model of the imaging process.

The previously discussed paper by Rudin *et al.* **[77]** provides a derivation of an analytical expression for pincushion distortion, caused by a spherical photocathode of radius (Ro) for a given source-to-image distance (s). However, Rudin *et al.* found that this purely geometrical treatment of image distortion performed poorly, in comparison to their two parameter empirical approximation technique. It was therefore concluded that the photocathode curvature could not be approximated by a sphere, and that distortion caused by the electron-optical system should not be neglected. Similar geometrical models have also been reported by Pietka *et al.* **[87]**, Ning *et al.* **[88]** and Chakraborty **[89]**.

2.4.3 Geomagnetic Distortion.

The majority of the distortion correction schemes outlined in the previous sections (2.4.2.1 to 2.4.2.3), employ a calibration grid attached to the image intensifier input screen, in order to quantify image distortion. The correction parameters calculated from this single reference image, are usually assumed to be valid for all of the possible imaging geometries of the fluoroscopy system. For this assumption to be true, the image intensifier tube would have to produce an identical image of the attached grid at every spatial orientation of the gantry or C-arm. Unfortunately, in reality it can easily be shown, using digital image subtraction techniques, that significant differences exist between the images produced at the various spatial orientations. These frequently overlooked discrepancies are attributable to the influence of the geomagnetic field upon the electrons inside the image intensifier tube.

The influence of a magnetic field (B-field) upon a charged particle, such as an electron, is described by the Lorentz force law **[90]**, which states that as the electron travels through the magnetic field it experiences an induced force, which acts at right

angles to both its velocity and the B-field. In the case of the image intensifier tube, this induced force deflects electrons from their true trajectories, and therefore represents an additional source of image distortion. As depicted in Figure 2.19, the geometrical form of this image distortion depends upon the orientation of the image intensifier tube with respect to the earth's magnetic field **[91]**. If the image intensifier tube is perpendicular to the geomagnetic field, the induced Lorentz force causes shifting (displacement) of the image in one direction. By contrast, for the parallel orientation case, the induced Lorentz force causes rotation of the image.

As the manufacturers of image intensifier tubes are obviously aware of the geomagnetic distortion problem, several techniques have been investigated as a means of suppressing the influence of the earth's magnetic field. The most commonly employed method is to surround the image intensifier tube with a cylindrical metallic shield. This "magnetic" shield focuses the geomagnetic field towards itself, thus weakening the field inside the image intensifier tube. However, as the image intensifier entrance and exit windows cannot be shielded, a residual geomagnetic field remains within the tube. As a consequence, although image rotation and shifting is still present, due to the magnetic shielding, these effects are noticeably reduced at the edges of the image. Geomagnetic distortion in modern image intensifier tubes is therefore characterised by a combination of S-shaped warping and image shifting.

The most important practical implication of geomagnetic distortion is the fact that, unlike pincushion distortion, it is dependent upon the spatial orientation of the image intensifier tube relative to the geomagnetic field **[92]**. Ideally, distortion correction schemes must therefore also take into consideration the spatial orientation of the fluoroscopy equipment. As discussed in the following section (2.4.3.1), in the case of fixed gantry systems this requirement can be met by off-line calibration. By obtaining distortion measurements at several different gantry angulations and orientations, a database of correction coefficients can be created. During a quantitative fluoroscopic or angiographic examination, the relevant correction coefficients can then be looked up and applied.

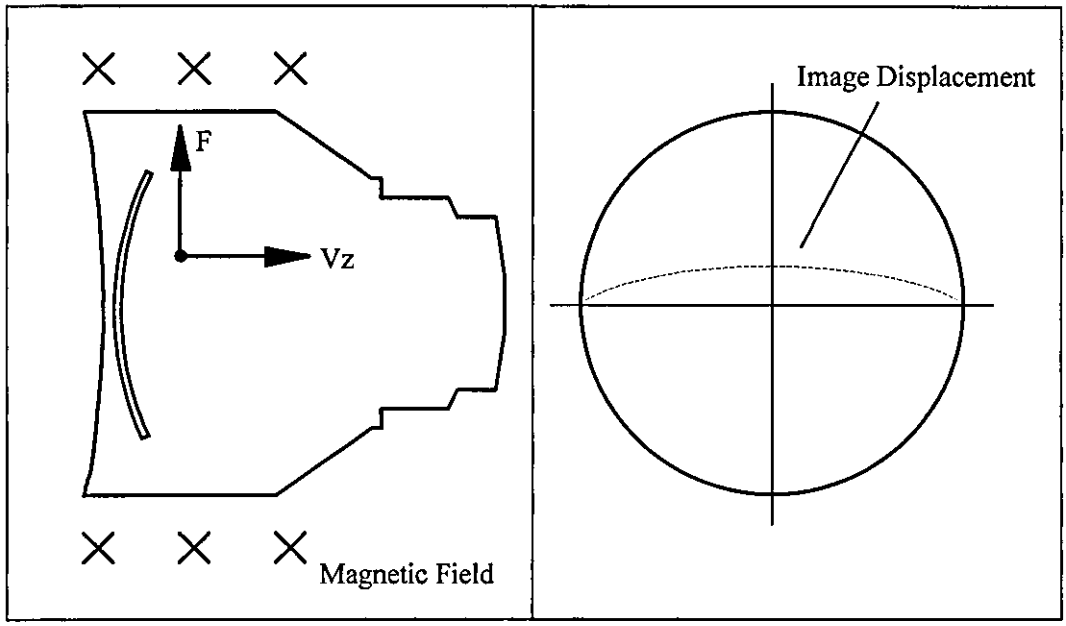


Image Intensifier Oriented Perpendicular (Transverse) to the Magnetic Field: The magnetic field acts upon the longitudinal component of an electron's velocity (Vz) resulting in a Lorentz Force (F) at right-angles to both the field and velocity. As a result the output image is shifted in one direction. (Due to magnetic shielding of the image intensifier, image shifting is reduced at the edges on the image.)

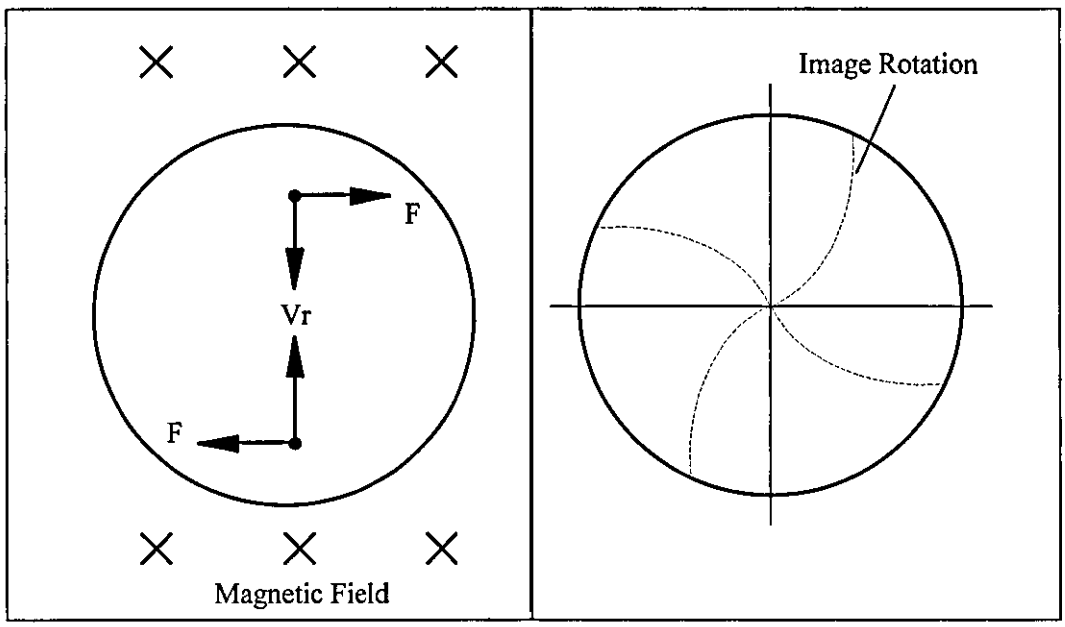


Image Intensifier Oriented Parallel (Longitudinal) to the Magnetic Field: The magnetic field acts upon the radial component of an electron's velocity (Vr) resulting in a Lorentz Force (F) at right-angles to both the field and velocity. As a result the output image is rotated. (Due to magnetic shielding of the image intensifier, image rotation is decreased at the edges on the image resulting in the characteristic S-shaped distortion.)

Fig. 2.19 : Geomagnetic Distortion of the Image Intensifier Image

However, off-line calibration is not a viable option for mobile fluoroscopy equipment, such as the C-arm unit, whose orientation with respect to the earth's magnetic field is completely unconstrained. In such situations, all image distortion (geomagnetic, pincushion, etc.) must therefore be corrected in situ, using only the information available in the diagnostic images themselves.

2.4.3.1 Geomagnetic Distortion Correction.

The limited number of published studies which actually take geomagnetic distortion and the spatial orientation of the X-ray equipment into consideration, all involve gantry-based quantitative angiography applications. In the case of Koppe *et al.* **[93]**, Beier *et al.* **[94]** and Gronenschild **[92]**, a combined pincushion and geomagnetic distortion correction scheme could therefore be implemented, using a standard calibration object and a global polynomial function to provide correction for each spatial orientation of the gantry. Van Der Zwet *et al.* **[95]** (University Hospital Leiden, The Netherlands) also describe a combined pincushion and geomagnetic correction scheme, which applies a localised bilinear transformation to the calculated correction vectors, associated with the four control points surrounding the point of interest.

Two stage techniques have also been reported in which geomagnetic and pincushion distortion are corrected separately. In addition to the geometrical treatment of pincushion distortion discussed in Section 2.4.2.3 **[88]**, Ning *et al.* also apply a global dewarping method to correct for S-shaped geomagnetic distortion **[96]**. Through personal correspondence with the University of Kiel, Germany, it has been established that Onnasch *et al.* also split the correction of image distortion into two distinct stages. Geomagnetic distortion is initially corrected with the aid of four small bronze washers, which are permanently attached to the centre of the image intensifier's input screen **[97]**. Comparison of the measured (distorted) image coordinates of these washers, with their known locations, allows an eight parameter global dewarping transformation to be calculated. Pincushion distortion is then corrected using a single parameter relationship similar to that of Casperson *et al.* **[79]**.

2.4.4 Quantitative Three-Dimensional Angiography.

The most successful applications of digital radiography are undoubtedly in the field of angiography. Since the human vascular structure cannot be imaged directly using X-rays, angiography requires the introduction of a contrast material to render the vasculature radiopaque. However, as the opacified vasculature is superimposed on the patient's skeleton in the subsequent image, the results of a conventional angiographic examination are often far from ideal. The traditional solution to this problem is a time consuming photographic subtraction technique, aimed at removing the superimposed anatomy from the angiographic images. Angiography was therefore considered to be an ideal candidate for digital image processing, both in terms of the enhancement of conventional angiograms, and the automation of the image subtraction process.

Angiographic image subtraction is now frequently performed using the well established clinical technique of digital subtraction angiography (DSA). DSA is a two-stage imaging technique, which initially involves the acquisition of a radiograph of the background anatomy, prior to the introduction of the contrast medium. This initial image is generally referred to as the *mask image*. The contrast medium is then introduced into the appropriate vein, thus allowing a series of opacified images to be obtained. Automatic removal of the background anatomy is then achieved by digitally subtracting the mask image from each of the opacified images. In this way, the resulting *difference images* contain the vascular structures of interest, without the superimposed bone and soft tissue density. If required, additional image enhancement can then be applied to improve the appearance of the DSA images.

The use of digital image processing in conjunction with image intensification, means that both dynamic angiography and DSA can now be performed on-line. The use of X-ray photogrammetry techniques to perform three-dimensional reconstructions of vascular structures has therefore become a viable prospect, and as such, is currently being investigated at several research sites world-wide. For the purposes of this review, these applications have been grouped according to the number and orientation of the angiographic views employed. As the majority of the research effort is focused towards coronary angiography, sub-section 2.4.4.1 deals with digital X-ray

photogrammetry systems that are based upon conventional biplane coronary angiography. Sub-section 2.4.4.2 then goes on to discuss multi-view systems, which mimic the tomographic techniques associated with CT scanning.

2.4.4.1 Biplane Coronary Angiography.

Conventional coronary angiography involves the simultaneous acquisition of two orthogonal projections of the opacified cardiovascular system. The resulting images are then used to subjectively diagnose congenital and acquired heart defects. However, by applying photogrammetric techniques to the angiographic examination process, it is also possible to accurately reconstruct the coronary vessel structure from the two angiographic projections. This quantitative data allows the progression of degenerative changes in the cardiovascular system to be followed much more precisely, and can therefore significantly aid treatment planning and the verification of surgical results.

One of the most common forms of heart disease in industrial countries is coronary arteriosclerosis (hardening of the arteries). The most effective means of evaluating the severity of the stenosis (narrowing of the arteries) caused by this type of condition, is to measure the blood flow through the diseased vessels. Since blood flow is usually defined as the volume of fluid passing through a selected arterial cross-section divided by the corresponding time interval, angiographic determination of coronary blood flow requires accurate measurement of arterial lengths and diameters. Biplane photogrammetric techniques aimed at performing these vessel measurements have been developed by Guggenheim *et al.* **[98]**, Coatrieux *et al.* **[99]**, Oswald *et al.* **[100]**, and Prause *et al.* **[97]**.

The film-based technique developed by Guggenheim *et al.* **[98]** (Cardiology Centre, University Hospital, Geneva, Switzerland) applies the conventional photogrammetric reconstruction approach to digitised coronary cineangiograms. Following completion of the standard angiographic examination, the patient is removed and a calibration object is placed in the location previously occupied by the patient's heart. This calibration object, which consists of a cube incorporating fifteen steel balls, is then

radiographed with the same imaging geometry used to examine the patient. Measurement of the image coordinates of the steel balls then allows a projection transformation and the location of the X-ray source to be determined for each of the two views. Once this calibration process has been performed, reconstruction of the vascular structure can begin **[101]**.

Coatrieux *et al.* **[99]** (University of Rennes, France) also describe the use of a Perspex calibration cube and semi-automatic vessel detection algorithms, to reconstruct coronary vasculature from biplane DSA images. Oswald *et al.* **[100]** (German Heart Institute, Berlin, Germany) on the other hand, have developed a biplane examination technique that avoids the use of a calibration object. Their method is an enhancement of the biplane radiography scheme proposed by Metz *et al.* **[102]**, which assumes that the image coordinates of eight or more object points of unknown three-dimensional location can be accurately determined in both views. The two source-to-image plane distances, as well as the origins of the image coordinate systems, must also be determined without the use of a calibration object.

Observing that it is very unusual for eight or more easily identifiable anatomical landmarks to be present in angiographic images, Oswald *et al.* **[103]** have further reduced the required number of landmarks to five, by making use of additional known imaging geometry parameters. Following interactive marking of these landmarks, which may be either anatomical features such as vessel bifurcations (i.e. forking points where vessels split into two branches) or artificial cardiomarkers mounted on a catheter, an iterative process is used to refine system calibration and reconstruction errors. Semi-automatic detection algorithms are then used to provide two-dimensional vessel centreline and edge data, prior to photogrammetric reconstruction. However, it is worth noting that Oswald *et al.* **[103]** state that interactive reconstruction of the full coronary vessel structure can take up to two hours. Ideally, the level of operator interaction should therefore be further reduced through increased automation of the image analysis process.

Onnasch *et al.* **[97]** (Clinic for Paediatric Cardiology, University of Kiel, Germany), also appear to perform biplane ventricular volume determination without the use of a three-dimensional calibration object. Their reconstruction technique is based upon a rigorous image pairing method, which brings the two images into perfect geometric alignment, so that each pair of corresponding rows in the images match a common cross-section. The three-dimensional ventricle shape can then be reconstructed by piling up the resulting cross-sectional slices. In practice, achieving this level of image alignment requires a complicated series of image pre-processing steps **[104]**. The spatial rotations required to bring the two images into perfect alignment are then calculated.

2.4.4.2 Multi-View Methods.

Clinical acceptance of quantitative biplane angiography is currently hindered by the required levels of operator intervention, which can lead to prolonged reconstruction times. However, this limitation can potentially be overcome through the development of multi-view angiography techniques. By mounting an angiography system on a motorised gantry, it is possible to acquire a series of images of the opacified vessel structure as the imaging chain rotates around the patient. By calibrating each of these images, using a suitably designed calibration object, the three-dimensional vessel structure can subsequently be reconstructed. Since the number of views involved makes interactive calibration techniques impracticable, multi-view techniques are characterised by high levels of automation. As a result of this reduced reliance upon user interaction, near real-time vessel reconstructions become possible.

Koppe *et al.* **[93]** (Philips GmbH Forschungslaboratorien, Hamburg, Germany) have developed an image calibration strategy that allows accurate three-dimensional vessel reconstruction using the Philips INTEGRIS V3000 imaging system to perform "Rotational Angiography". Image intensifier distortion (pincushion and geomagnetic) is initially corrected, with the aid of a calibration plate attached to the image intensifier input screen. A second calibration plate (rigidly coupled in front of the first plate) is then used to derive the location of the X-ray source. The geometrical projection parameters are then determined by imaging a custom-built polyhedron

calibration object, (a calibration cube is unsuitable for multi-view imaging). After this two-stage calibration procedure is complete, the patient acquisitions are performed using the same imaging geometry. It is stated that this approach has been successfully applied *in vivo* during human patient trials.

Rougée *et al.* **[105]** (General Electric Medical Systems Europe, Buc, France) also describe a three-dimensional "Computerised Angiography (CA)" technique which uses a prototype imaging system called a "Morphometer". Two perpendicular imaging chains are mounted on an existing CT gantry, in such a way that a complete biplane cardiovascular digital imaging system can effectively be rotated around the patient. Automatic calibration of the various viewing angles is achieved through the use of a dedicated calibration object **[106]**. This Lucite cylinder incorporates a helical control point distribution, which is claimed to simplify automatic image analysis (labelling) and calibration. In order to facilitate DSA imaging, mask images can be acquired while the gantry rotates in one direction, and opacified images are then acquired while the gantry rotates in the opposite direction. A modified biplane angiographic imaging system has also been developed by Ning *et al.* **[96]** (University of Rochester Medical Centre, Rochester, NY) to provide a similar "Tomographic Angiography" capability.

2.5 Summary of Chapter 2.

The term photogrammetry is used to describe image-based metrology techniques that allow the three-dimensional reconstruction of an object from two or more disparate images of the object. Photogrammetric reconstruction techniques were originally developed for specialised photographic applications, and have subsequently found more widespread use, due to recent developments in the area of machine vision. However, given the objectives of the current thesis, the most important aspect of photogrammetry is the fact that in addition to visible light images, the technique is equally applicable to X-ray images. Thus having initially established the relevant theory and terminology in terms of camera-based photogrammetry, this chapter has concentrated upon the application of photogrammetric techniques to diagnostic radiography.

Film-based X-ray photogrammetry techniques have successfully been applied to a wide range of orthopaedic applications. Many of these photogrammetric techniques, most notably Roentgen stereophotogrammetric analysis (RSA), have demonstrated sub-millimetre levels of precision *in vivo*, thereby establishing the feasibility of the current thesis. Unfortunately, owing to the delay caused by the film development process, these conventional radiography approaches are limited to preoperative and postoperative examinations. Clinical acceptance of film-based X-ray photogrammetry is also hindered, by the time-consuming process of manually measuring the image coordinates of the calibration markers imaged on the radiographs.

Fortunately, it would appear that the intraoperative registration capability required by the current thesis, can be achieved by implementing a digital (fluoroscopy-based) X-ray photogrammetry system. However, in reviewing the relevant published literature, it is noticeable that the non-linear imaging properties and the reduced spatial resolution of the image intensifier tube, have previously led to a reluctance to implement fluoroscopy-based photogrammetry. Fluoroscopic photogrammetry has therefore mainly been restricted to quantitative angiographic examinations.

Upon close examination, it is evident that previous quantitative angiography research has avoided many of the more complicated issues associated with the orthopaedic surgery applications envisaged by the current thesis. Image acquisition during angiographic examinations is usually performed using a fixed-base, gantry-mounted fluoroscopy installation. This type of constrained imaging geometry allows off-line compensation for both pincushion and geomagnetic image distortion, and can also lead to a simplified external calibration process. In addition, the gantry system frequently provides a rigidly maintainable imaging geometry, thus allowing the patient and the calibration object to be imaged separately.

The development of a C-arm based photogrammetry system therefore requires solutions to be found for a number of unresolved issues. The initial problems which have to be overcome, relate to the design of a suitable calibration frame. Given that

C-arm units cannot be accurately re-positioned, the patient and the calibration frame must be imaged together. As a consequence, calibration features have to be introduced into the diagnostic images in an *operating room compatible* manner.

Assuming that a satisfactory means of introducing calibration features into the diagnostic images can be developed, the next significant problem requiring investigation is the image analysis process. In order to ensure that the intraoperative registration process does not prolong the duration of the surgical procedure, user interaction must be kept to a minimum. Semi-automatic image analysis routines must therefore be developed with the aim of achieving rapid data processing.

Once the control point image data has been obtained, the photogrammetric system must then be calibrated. Due to the unconstrained nature of the imaging geometry of the C-arm unit, it is very difficult to predict the effects of geomagnetic distortion under practical conditions. Image dewarping, to compensate for the effects of pincushion and geomagnetic distortion, must therefore be performed in situ, using only the calibration information contained within the diagnostic images. A suitable external calibration/trajectory planning scheme is then required to perform the photogrammetric reconstruction process.

In conclusion, the published literature relating to X-ray photogrammetry, strongly suggests that the objectives of the current thesis can be achieved using existing fluoroscopy equipment, in conjunction with digital image analysis and calibration techniques. However, due to the unique nature of the intraoperative applications under investigation, practical solutions to a number of previously avoided problems must be developed, in order to ensure clinical acceptance of the proposed photogrammetry system.

3. INTRAOPERATIVE REGISTRATION.

3.1 Overview of the Registration Problem.

In the majority of robotic-assisted and computer-assisted surgery (CAS) applications, the patient is initially examined using a tomographic imaging modality; such as computed tomography (CT) or magnetic resonance imaging (MRI). The resulting cross-sectional images are then combined to reconstruct a three-dimensional model of the patient's anatomy. Using a graphical workstation to interact with this model, the surgeon is then able to pre-plan an optimum series of surgical trajectories. However, this pre-surgical plan is defined with respect to the coordinate system associated with the tomographic imaging system, as opposed to the real-world situation encountered by the surgeon in the operating room. Although in theory this problem can be avoided by using intraoperative MRI or CT scanning, for a number of practical reasons, these techniques are currently limited to a few state-of-the-art systems. Most applications therefore require a *registration* process to be undertaken in order to make the preoperative plan "available" in the operating room.

Depending upon the complexity of the surgical procedure, the registration process will involve some/all of the stages shown schematically in Figure 3.1. As preoperative surgical planning can require the use of more than one imaging modality, *cross-modality registration* may become necessary, in order to develop a three-dimensional model of patient anatomy that incorporates multi-modality information. For instance, neurosurgical procedures frequently employ a combination of anatomical (i.e. CT, MRI, etc.) and functional (i.e. SPECT, PET, etc.) imaging modalities in order to gain a better understanding of the location and nature of brain abnormalities. In such cases, cross-modality registration typically involves attaching a series of fiducial markers to the patient's head. When imaged along with the patient, these fiducial markers provide easily identifiable image points which are common to the various imaging modalities. Spatial relationships can therefore be established between pairs of imaging modalities, thereby allowing corresponding points in anatomical and functional images to be identified.

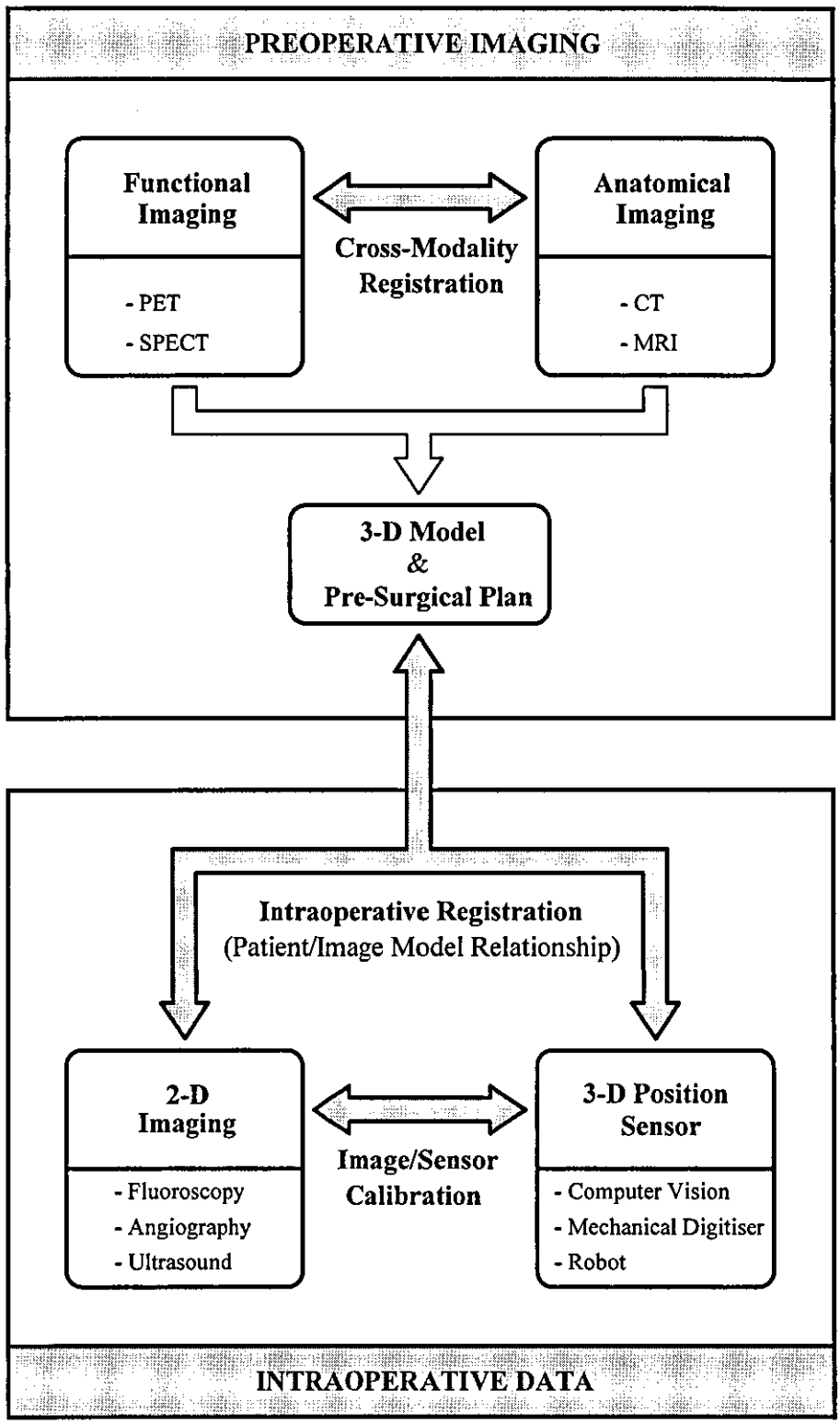


Fig. 3.1 : Co-Registration of Surgical Data Sources.

In order to accurately reproduce the planned surgical procedure in the operating room, the *intraoperative registration* problem must first be solved. Essentially, intraoperative registration aims to establish a spatial relationship between the reference system associated with the preoperative patient model, and an arbitrarily defined operating room (world) coordinate system. In practice, the operating room coordinate system for a CAS application is normally associated with some form of intraoperative position-sensor. For "open" surgery procedures in which there is direct physical access to the surgical site, this intraoperative sensor may be an encoded electro-mechanical digitising arm or a triangulation-based quantitative imaging system. When using these technologies, the intraoperative registration problem is solved by digitising/measuring the three-dimensional location of a series of points which can also be easily identified in the preoperative images. These points can be either naturally occurring anatomical landmarks, or alternatively, artificial fiducials which are rigidly attached to the patient's anatomy. The required transformation between patient model and operating room coordinates is then derived from this corresponding point-pair data.

For "closed" surgery procedures, visualisation of the surgical site can only be achieved through the use of a two-dimensional intraoperative imaging technique (i.e. fluoroscopy, ultrasound, etc.). Intraoperative registration therefore initially requires a spatial relationship to be established between the frames of reference associated with the 3-D preoperative images and the 2-D intraoperative images. Once again this registration process usually involves an image matching method which makes use of artificial fiducials or anatomical landmarks. A second registration process, which will be referred to as *image/sensor calibration* for the purposes of this chapter, must then be performed in order to complete the intraoperative registration process. The aim of image/sensor calibration is to establish a transformation between the frames of reference associated with the intraoperative imaging system and the operating room coordinate system. This can be achieved by either measuring the actual position and orientation (pose) of the imaging system using a 3-D position-sensor, or alternatively, by imaging a calibration object whose pose can then be accurately determined by a 3-D position-sensor.

Once intraoperative registration is complete the surgical plan can then be accurately reproduced in the operating room. The majority of CAS systems provide the surgeon with a guiding system to aid his execution of the surgical plan. Ordinarily, this involves tracking the position of a surgical tool in real-time using the 3-D intraoperative position-sensor. As the transformation between operating room and preoperative image coordinates has been established by the registration process, a graphical workstation can then be employed to display the measured tool position on the preoperative images, thus establishing the required link between surgical planning and execution.

By contrast, in robotic-assisted surgery applications the operating room coordinate system is usually defined by a frame of reference attached to the base of the electro-mechanical manipulator. Kinematic analysis and calibration of the manipulator's configuration, therefore allows the pose of the end-effector to be determined with respect to the intraoperative coordinate system. As such, during "open" surgery the intraoperative registration problem can be solved by using the manipulator as a 3-D position-sensor while performing a tactile search for artificial fiducials or anatomical landmarks. Alternatively, by acquiring an intraoperative image of a calibration object which is attached to the manipulator's end-effector, image/"sensor" calibration can be performed in order to establish intraoperative registration for "closed" surgery procedures.

Owing to the nature of the robotic-assisted orthopaedic surgery procedures proposed by the current thesis, an in-depth discussion of cross-modality registration will not be presented in this chapter. The remainder of the chapter therefore concentrates upon the intraoperative aspects of the registration process. Section 3.2 introduces the various forms of sensor technology that can be applied to solve the intraoperative registration problem. A critique of the intraoperative registration strategies which have previously been used during computer-assisted and robotic-assisted orthopaedic procedures is then presented in Section 3.3. A similar treatment of intraoperative

registration strategies developed for neurosurgery applications is also provided elsewhere, in Appendix 3.

The reason for including this review of neurosurgery applications, is that neurosurgical procedures generally require higher levels of precision than their orthopaedic counterparts. As a result, techniques developed for neurosurgery represent the state-of-the-art in intraoperative registration. In addition, much of the pioneering CAS and robotic-assisted surgery research was targeted at neurosurgery, due to the fact that it is the surgical speciality which can potentially benefit the most from the introduction of these new technologies into the operating room.

Finally, a critique of surgical applications which involve the use of quantitative intraoperative X-ray imaging is presented in Section 3.4. The research work covered in this section is of direct relevance to the current thesis, and includes two independent studies which are also specifically targeted at femoral osteosynthesis procedures. The implications of this closely related research, as well as the findings of the chapter as a whole, are therefore discussed with respect to the objectives of the current thesis in Section 3.5.

3.2 Intraoperative Sensor Technology.

Intraoperative registration strategies usually require the acquisition of accurate 3-D coordinate data in the operating room. Software registration techniques are then applied to match this intraoperative data with the corresponding preoperative image data, thereby allowing the required patient model/operating room coordinate transformation to be established. As a result, intraoperative 3-D position-sensing techniques form the cornerstone of most algorithm-based registration methods. When applied to CAS and robotic-assisted surgery applications, position-sensing techniques are invariably referred to as *3-D localisers*, *digitisers* or *navigation systems*. However, in spite of this terminology, the individual components of these systems are often based upon existing medical equipment or well established computer vision technology. The basic principles underlying the operation of these various sensing

techniques are therefore presented in this section, with a view to clarifying subsequent descriptions of the intraoperative registration process.

Intraoperative 3-D position-sensing systems tend to employ one of the sensing techniques listed below:

- Passive (mechanical) digitising arms.
- Passive imaging.
- Active imaging.
- Sonic localisation.
- Electromagnetic localisation.
- Localised ultrasound imaging.

However, since the photogrammetric reconstruction process has already been outlined in Chapter 2, and the nature of the current thesis requires that registration strategies employing quantitative X-ray imaging are discussed separately (in Section 3.4), passive imaging systems are not covered in this section. Similarly, owing to the "hybrid" nature of localised ultrasound imaging, which achieves "3-D" ultrasound imaging by using a mechanical, optical, or electromagnetic localiser to track the location of a standard ultrasound imaging probe, a detailed analysis of this technique is also unwarranted. The following discussion is therefore limited to the four remaining sensing techniques, and begins by looking at passive digitising arms.

Mechanical position-sensing arms usually adopt an anthropomorphic, six degrees-of-freedom kinematic configuration, similar to the one depicted in Figure 3.2. This articulated arrangement allows the surgeon to manually bring the tip of the measuring probe into direct physical contact with a point whose 3-D location is to be measured (e.g. an anatomical landmark). If the operating room coordinate system is defined as being coincident with a frame of reference attached to the base of the arm, the task of intraoperative position-sensing then reverts to solution of the forward kinematics problem. Intraoperative position-sensing can therefore be achieved in real-time by instrumenting each of the arm's joints with a potentiometer or an optical encoder, and interfacing these rotary sensors with a personal computer.

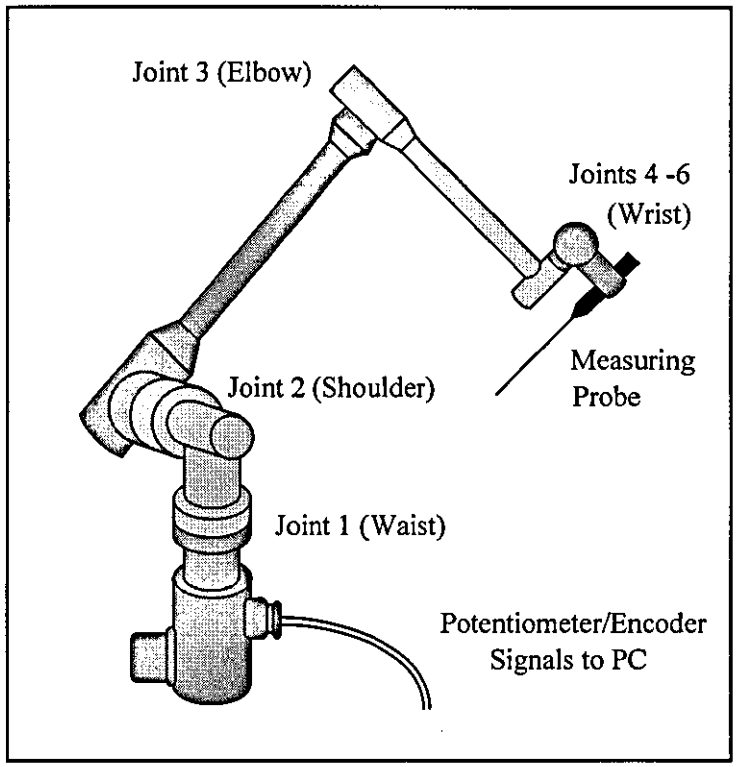


Fig. 3.2 : A Typical Anthropomorphic Digitising Arm Design

As the probe's tip position is calculated from measured joint angles and a series of "known" lengths, the measurement accuracy of a digitising arm is dependent upon the precision of its optical encoders/potentiometers, as well as the accuracy of the arm calibration process. Since the absolute measurement error is also related to the size of the working volume, digitising arms are predominantly applied to neurosurgery and ear, nose and throat (ENT) procedures which involve small working volumes (typically 0.4 m x 0.4 m x 0.4 m, or less). For applications of this magnitude, most digitising arms are capable of providing measurements to sub-millimetre accuracy.

Commercially available arm-based systems, such as the Viewing Wand (ISG Technologies, Ontario, Canada), are currently being applied to a wide range of surgical procedures **[107 - 110]**. Successful clinical trials have also been conducted using custom-built prototype digitising arm systems **[111 - 114]**. The use of mechanical or electromagnetic braking systems has also been investigated, as a means of allowing the arm to rigidly maintain a desired position **[112, 115, 116]**. The arm can then be used as either a tool-holder or as a support for an imaging device. It is

also worth noting that Lavallée *et al.* **[117]** (TIMC-IMAG, Faculté de Médecine de Grenoble, France) have investigated an alternative to the standard (algorithmic) registration approach, which involves mounting a video camera onto an articulated passive arm. This video-based registration technique works by interactively aligning an on-line video image of the patient's anatomy, with a simulated image derived from preoperative CT-image data.

In practice, manoeuvring a digitising arm in the restricted operating room environment can prove to be very awkward. As a result, more research effort has gone into the development of "armless" (non-contacting) position-sensing methods. The advantages offered by active imaging systems were previously outlined in Section 2.2.2. Structured light can significantly simplify and speed up the image analysis process, and as such, has found widespread acceptance as an intraoperative position-sensing method. Accordingly, several studies have documented the use of laser scanners and projector-based structured light systems during neurosurgery **[118 - 120]**, orthopaedic **[121]**, and radiotherapy **[122]** applications. However, structured light is by no means the most popular form of active imaging technique employed in CAS and robotic-assisted surgery. That accolade belongs to a group of intraoperative sensors which henceforth will be referred to as *infra-red optical digitisers*.

The principle of operation of an infra-red optical digitiser, requires that a group of active targets are physically attached to the object whose three-dimensional location is to be reconstructed. A multi-camera system is then used to triangulate the three-dimensional positions of these active targets, thus allowing the object to be modelled as a rigid body. For the vast majority of applications, the active targets are infra-red light emitting diodes (IR-LEDs), which are attached to either a surgical tool or part of the patient's anatomy. As these diodes obviously require a source of power, the introduction of electrical wiring into the space around the patient is an unavoidable inconvenience when using infra-red optical digitisers. However, the target identification problem can be significantly simplified by sequentially pulsing the IR-LEDs in synchronisation with the cameras.

Most infra-red optical digitisers employ the three-camera imaging geometry shown schematically in Figure 3.3. Typically, each of these cameras consists of a cylindrical lens placed in front of a high resolution (4096 element) line-scan CCD sensor. This lens arrangement focuses the light from an IR-LED onto a single element of the sensor array, and as a result, each camera effectively defines a three-dimensional plane containing the diode. By mounting the two end cameras horizontally and the central camera vertically, the diode's location can therefore be triangulated from the intersection of three planes.

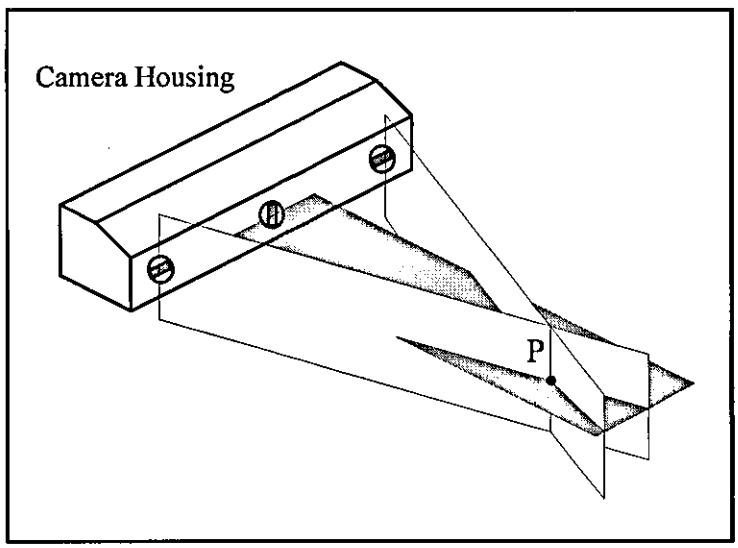


Fig. 3.3 : The Infra-Red Optical Digitiser.
Triangulation of an infra-red LED's location (P) from the intersection of three planes.
(Adapted from Galloway *et al.* **[123]**)

Infra-red optical digitisers that incorporate line-scan CCD cameras are arguably the most accurate form of intraoperative position-sensors. Although measurement accuracy depends upon the size of the working volume, sub-millimetre precision is not unusual in volumes of up to 1 m^3. The high data rates associated with line-scan cameras also ensure that PC-based systems are capable of simultaneously tracking several rigid bodies in real-time. Commercially available systems have therefore been adapted to a wide variety of surgical procedures. The most noteworthy of these systems is Optotrak 3020 (Northern Digital Inc., Ontario, Canada), which has been successfully applied to orthopaedic surgery **[124 - 128]**, neurosurgery **[129, 130]**, craniofacial surgery **[131]** and radiotherapy **[132]**.

However, it must also be appreciated that a rigid body's location can only be reconstructed if a direct line-of-sight exists between the CCD cameras and the IR-LEDs. In addition, the need to physically attach active markers to both the patient and the surgical tools, necessitates the inconvenience of a dry sterilisation technique (typically ethylene oxide), and restricts the use of infra-red optical digitisers to open surgical procedures.

The current batch of sonic localisation systems are based upon modified ultrasonic range-finding technology. Ultrasonic range-finders work by measuring the time taken for an acoustic pulse to travel from a sound emitter to a microphone. Knowledge of the speed of sound in air then allows the required distance between the sound emitter and the microphone to be calculated. However, by adapting this technique to include the use of three or more microphones, whose relative locations are accurately known, it is also possible to reconstruct the three-dimensional location of the sound emitter. Using a simple synchronisation technique, the distance from the sound emitter to each of the microphones is initially established. As each of these distances effectively constrains the sound emitter's location to the surface of a sphere, whose centre is at the relevant microphone location and whose radius is equal to the emitter-to-microphone distance, the location of the sound emitter can be derived from the intersection of three spheres. Intraoperative position-sensing can therefore be performed by sequentially determining the location of a group of sound emitters attached to a measuring probe.

Figure 3.4 depicts the use of a typical sonic localisation system during a neurosurgical procedure. Assuming that a preoperative CT or MRI examination has already taken place, intraoperative registration requires the locations of previously imaged fiducial markers to be established with respect to the operating room coordinate system. This is achieved by bringing the tip of a hand-held probe/wand into contact with each of the fiducial markers. Once in position, ultrasonic emitters attached to the wand sequentially emit a series of pulses, which are then detected by a synchronised array of microphones. By reconstructing the emitter positions in this way, the known geometry of the probe allows its tip location and hence the fiducial marker location to

be determined. Intraoperative registration can therefore be achieved, thus allowing subsequent probe tip locations to be measured and displayed, with respect to the preoperative images on the system's graphical workstation.

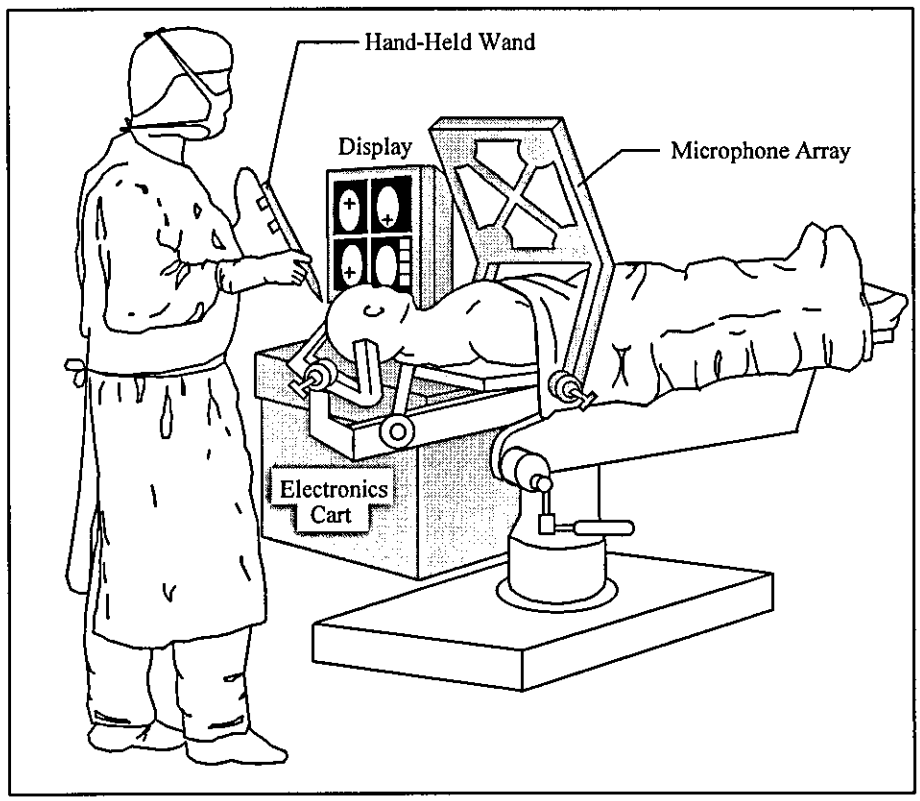


Fig. 3.4 : A Real-Time Intraoperative Sonic Localisation System
(Adapted from Williams *et al.* [133])

As sonic localisation is a time-of-flight based metrology technique, the accuracy of a "measured" distance is only as good as that of the speed of sound value from which it is calculated. Unfortunately, the speed of sound in air is both temperature and humidity dependent. In order to compensate for these environmental influences, real-time determination of the speed of sound is therefore a prerequisite for accurate sonic localisation. This requirement is usually achieved by incorporating reference sound emitters into the microphone array. Immediately prior to each measurement process, these reference emitters fire a sequence of acoustic pulses at the microphone array. Knowledge of the reference emitter-to-microphone distances then allows the required speed of sound value to be calculated from the measured transition times of the reference pulses.

In addition, problems also arise from the need to maintain an unobstructed "line-of-sight" between the sound emitters and the microphone array. Great care must be taken to ensure that any reflected acoustic pulses (i.e. echoes) or background noises are filtered out during signal processing. A Doppler type effect can also be experienced as a result of air turbulence in the operating room. In order to minimise these distortion effects, it is advisable to reduce the distance between the hand-held probe and the microphone array; hence the table-mounted microphone array depicted in Figure 3.4. However, due to the cumulative effect of these potential error sources, the absolute measurement accuracy of sonic localisation systems is generally quite poor (typically ±1 mm or worse). Intraoperative utilisation of sonic localisers has therefore been limited to a small number of neurosurgical and ENT procedures **[112, 134, 135]**.

Electromagnetic localisation techniques work by using a fixed position magnetic source (e.g. an electromagnet) to establish a low-frequency magnetic field in the vicinity of the measurement volume. Given that the characteristics (strength and orientation) of this magnetic field are known as a function of distance from the magnetic source, the relative spatial position and orientation of an object can be determined by measuring the magnetic field at the object's location. Accordingly, if the origin of the operating room coordinate system is taken as being coincident with the magnetic source, intraoperative position-sensing can be performed by attaching magnetic field sensors to a surgical instrument or part of the patient's anatomy. Real-time tracking of these objects is then facilitated by monitoring the modulation of the magnetic field as they move.

Unfortunately, the fact that electromagnetic localisation is based upon measurement of a reference magnetic field, means that it is also highly sensitive to external factors that can lead to distortion of the magnetic field. As such, electromagnetic noise produced by standard operating room equipment can cause considerable problems. Great care must also be taken to ensure that, when placed between the magnetic source and sensor, surgical instruments made from metallic alloys do not lead to significant levels of interference. Alternative instrument materials, and hence sterilisation techniques,

may therefore have to be adopted in order to accommodate the use of electromagnetic localisation.

In addition to problems caused by incompatibility with the operating room environment, the widespread use of electromagnetic technology for intraoperative position-sensing is also hindered by the limited accuracy of commercially available systems. The manufacturers of popular systems such as 3Space Isotrack (Polhemus Inc., Colchester, VT) and Regulus (Stereotactic Medical Systems, Rochester, MN), typically quote measurement accuracies of the order of ±3 mm. Consequently, electromagnetic navigation systems have tended to be aimed at neurosurgery **[136, 137]** and ENT **[138]** procedures that require only moderate accuracy. However, the technique has also been investigated for use during spinal fusion surgery **[139]**, and has even been applied to a couple of robotic applications **[140, 141]**.

3.3 Orthopaedic Applications.

In many respects, orthopaedic procedures are ideal candidates for the introduction of computer-assisted surgery (CAS) techniques. The orthopaedic speciality deals with problems of the musculoskeletal system, and as such, surgical procedures generally involve interactions with inherently rigid structures (i.e. bones). Given that these structures can be immobilised and will not distort during surgery, registration strategies based upon rigid coordinate transformations can be applied with a certain degree of confidence. The "frameless stereotaxy" approach, discussed in Appendix 3, can therefore be employed to preoperatively plan and execute orthopaedic procedures. Computed tomography (CT), which gives excellent differentiation between bone and soft tissues, is normally used to generate the three-dimensional models required for preoperative planning. This plan is then implemented using a combination of intraoperative position-sensing and guidance systems. However, orthopaedic CAS applications also benefit from the fact that intraoperative feedback can be obtained using a wide variety of imaging techniques (e.g. fluoroscopy, arthroscopy, ultrasound, etc.).

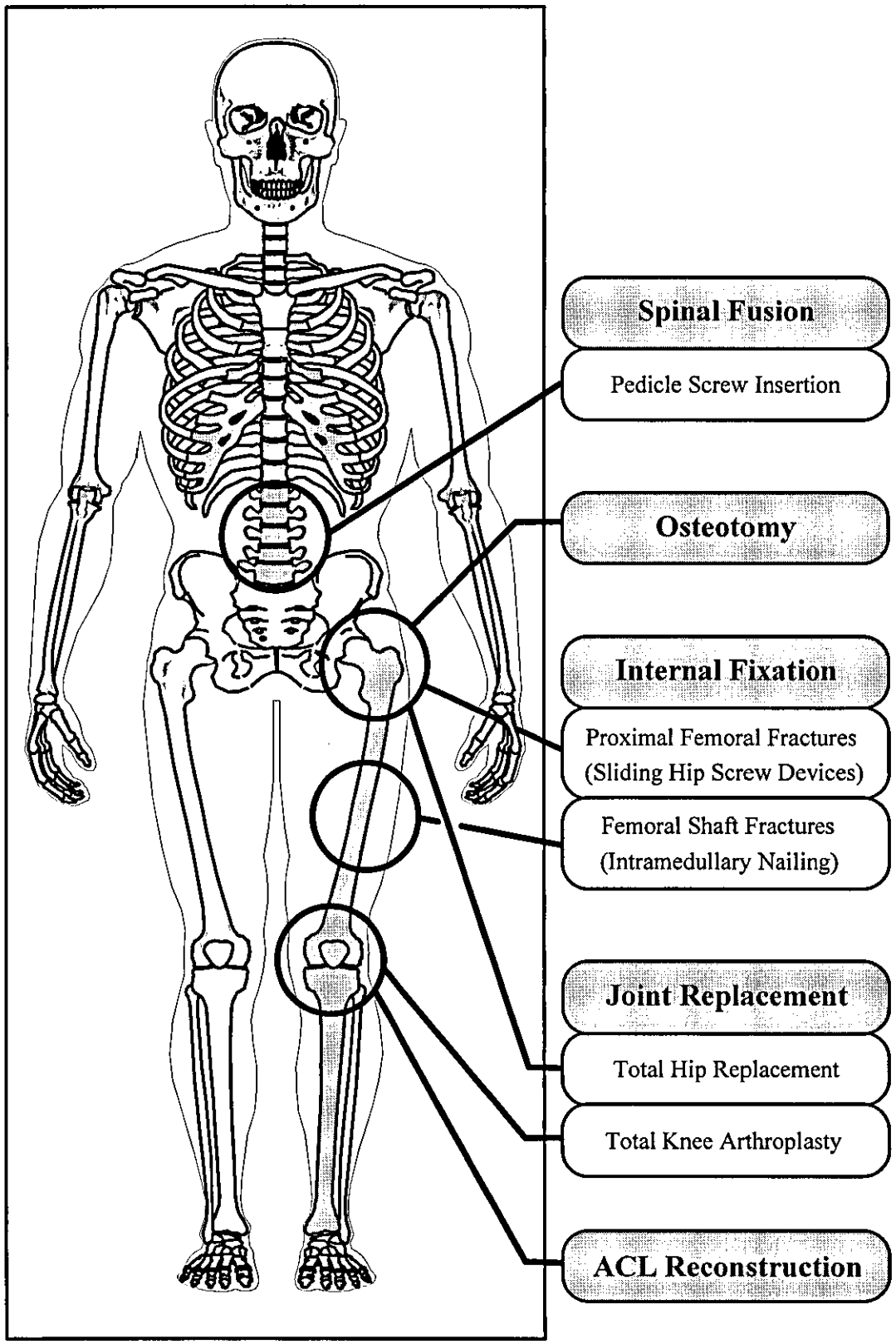


Fig. 3.5 : Robotic- and Computer-Assisted Orthopaedic Surgery:
Procedures currently under investigation world-wide.

Several research groups have also investigated robotic-assisted surgery as a means of improving the accuracy of orthopaedic procedures. Many of these applications are non-invasive, with a robotic manipulator merely being used to position tool guides. However, an alternative CAD/CAM approach can also be adopted, whereby the preoperative plan is converted into a series of robot-compatible commands. By attaching an air-powered surgical tool to the robot's end-effector, and establishing intraoperative registration, the robotic manipulator can therefore be used to automate the bone "machining" processes currently undertaken by the surgeon. In comparison to neurosurgical applications discussed in Appendix 3, the accuracy requirements for robotic-assisted orthopaedic surgery are far more relaxed. Bone is a "forgiving" or "adaptive" living tissue, which is able to compensate for small machining errors after a period of time. With a few notable exceptions, the potential damage that can be caused by tool-positioning errors of a few millimetres is also greatly reduced.

For the purposes of this critique, robotic and CAS applications involving orthopaedic procedures have been grouped according to the five categories illustrated in Figure 3.5. The first of these categories is arthroplasty or joint replacement surgery. The aim of this type of procedure is to replace the natural surfaces of a damaged or worn joint with prosthetic components. In practice, a complex sequence of bone "machining" processes must be undertaken to ensure optimal positioning of the prosthetic components. Accordingly, Section 3.3.1 discusses the registration strategies of a number of robotic-assisted surgery systems, which are targeted at the total hip replacement (THR) and total knee replacement (TKR) procedures.

By contrast, the second category of procedures has benefited from the frameless stereotaxy techniques originally developed for neurosurgical applications. Spinal fusion surgery is a form of osteosynthesis that requires the insertion of screws into the patient's vertebrae. Unfortunately, due to the proximity of the spinal cord, protrusion of these screws from the narrow vertebral pedicles can cause substantial damage. Section 3.3.2 therefore discusses a number of CAS strategies aimed at ensuring the accurate placement of pedicle screws. The next category covers another group of procedures that involve complex bone machining processes. Osteotomy typically

involves the removal of a portion of bone, as part of the corrective treatment of malaligned fractures or osteoarthritis. In addition to robotic-assisted osteotomy, Section 3.3.3 also details an alternative registration strategy which involves the use of custom-made templates.

The two remaining categories cover procedures which do not involve preoperative planning. However, the internal fixation of long bone fractures is the subject of the current thesis, and is therefore discussed separately in Section 3.4.2. Consequently, the final section of this review of orthopaedic applications (Section 3.3.4) covers a research project aimed at optimising reconstruction of the anterior cruciate ligament (ACL).

3.3.1 Arthroplasty.

The early attempts at joint replacement surgery achieved fixation of the prosthetic components by using "bone cement" as a grouting material. However, it soon became clear that these "cemented" prosthetic joints had a limited life span. During the inevitable revision arthroplasty procedures, it was also discovered that the use of bone cement made the removal of prosthetic components very difficult. "Cementless" fixation techniques have subsequently been introduced as a possible solution to this problem. Cementless fixation relies upon bone ingrowth into porous surfaces on the prosthetic components, and as such, requires a close fit to be established between the bone and the implant. Unfortunately, the required levels of precision are very difficult to achieve using hand-held surgical instruments. Gaps will therefore generally be observed between the bone and the implant, thus reducing the scope for bone ingrowth. As a result, micromotion/mechanical loosening problems can occur, ultimately leading to failure of the prosthetic joint.

In an attempt to reduce the failure rates associated with cementless prostheses, a number of research groups have investigated robotic-assistance, with a view to improving the initial bone/implant interface. The most advanced of these projects has led to the development of the ROBODOC™ Surgical Assistant System (Integrated Surgical Systems, Inc., Sacramento, CA), which is initially targeted at cementless total

hip replacement (THR) procedures **[142]**. To be more specific, the ROBODOC system uses a customised five-axis Sankyo Seiki industrial robot to prepare the cavity into which the femoral component of the THR device is to be inserted. Having undergone a long development process (discussed elsewhere in Section A2.6 of Appendix 2), in November 1992 ROBODOC became the first robot to actively participate in an invasive orthopaedic procedure, and is currently undergoing large-scale multi-centre clinical trials **[143]**.

The ROBODOC system achieves robotic milling of the femoral cavity using the sequence of events summarised in Figure 3.6. The patient initially undergoes a minor surgical procedure to percutaneously insert three titanium calibration pins into the involved femur. A CT examination is then undertaken with these pins firmly in position. The system's ORTHODOC™ Preoperative Planning Workstation is then used to generate a 3-D graphical model of the patient's femur from the CT image data. The surgeon can then plan the surgical procedure by superimposing prosthesis models, derived from the manufacturer's specification, onto the femur model. A CAD/CAM package is then used to transform the preoperative plan into a set of tool path instructions for the robot's cutter **[144]**.

The preliminary stages of the surgical procedure, including femoral neck osteotomy and insertion of the prosthetic acetabular cup, are performed in the usual manner by the surgeon. The three calibration pins are then exposed, and the femur is immobilised using a fixation device attached to the robot's base. Intraoperative registration is then established by determining the locations of the three calibration pins with respect to the robot's coordinate system. This is achieved by inserting a ball probe into the robot's end-effector, and manually guiding the robot to the vicinity of the first calibration pin. The end-effector's force sensor then allows the robot to physically locate the centre of the pin using a tactile search strategy. The same technique is then used to "teach" the robot the locations of the other two calibration pins. The ball probe is then replaced with a high-speed rotary cutter, and robotic milling of the femoral cavity is performed using the tool path instructions derived from the preoperative plan.

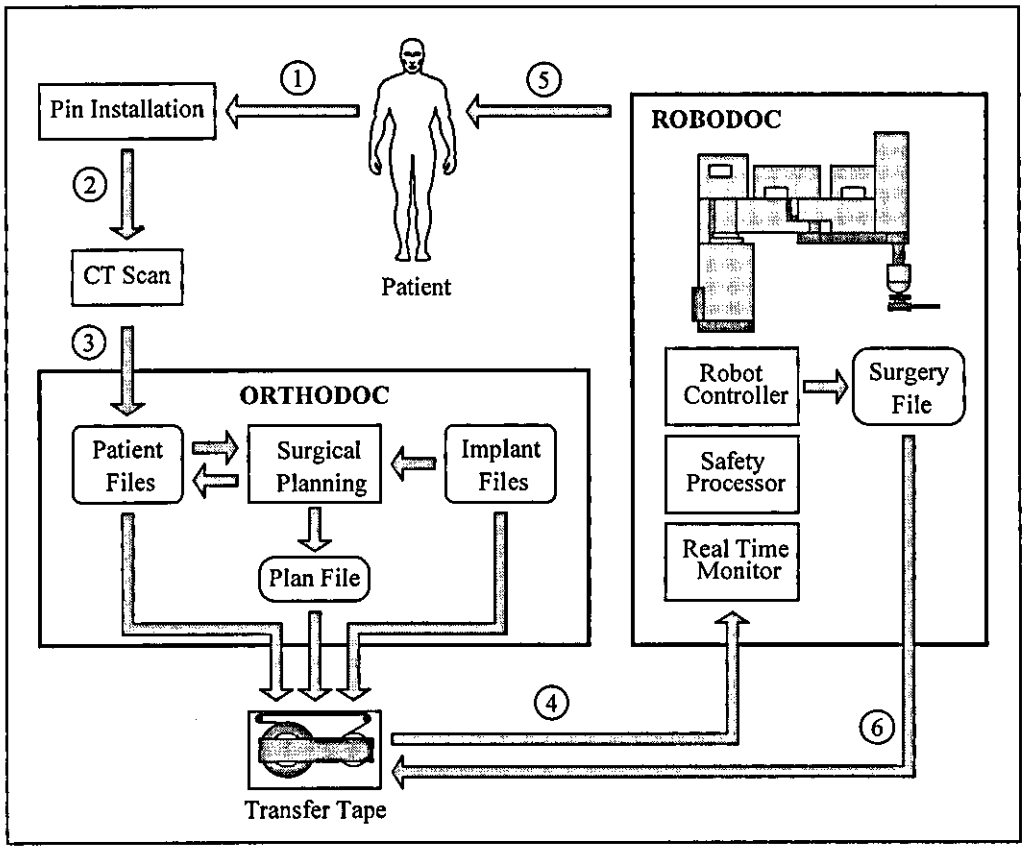


Fig. 3.6 : ROBODOC System Procedural Flow
(Adapted from Kazanzides *et al.* **[142]**)

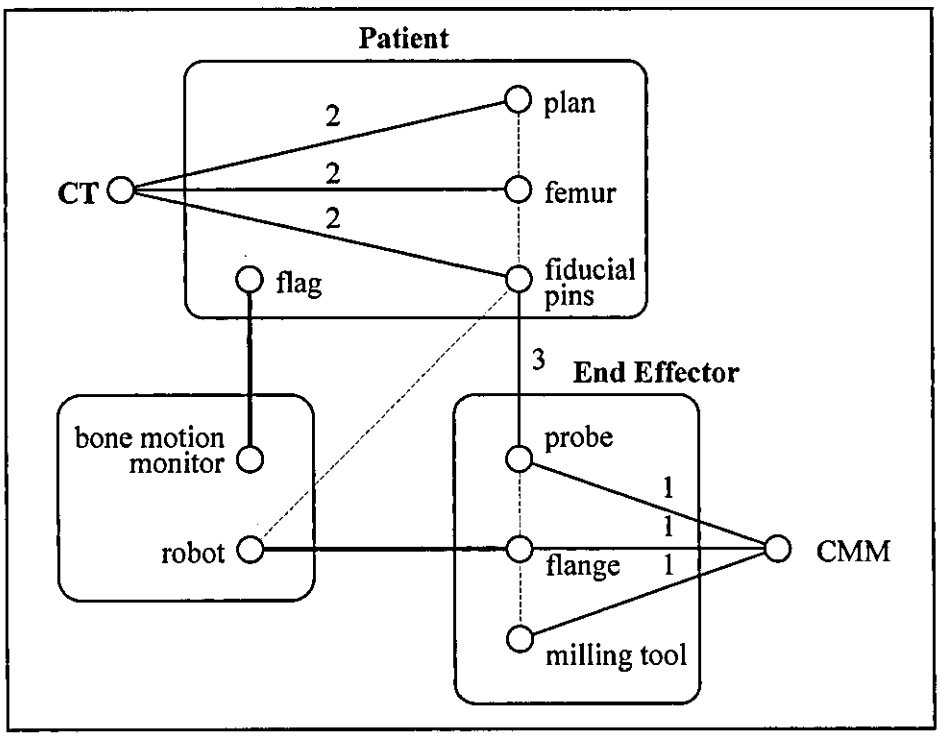


Fig. 3.7 : Registration Graph for the ISS ROBODOC System.
(Adapted from Lea *et al.* **[145]**)

From the preceding description, it can be appreciated that the ROBODOC system employs a point-based registration strategy, which relies upon artificial fiducials (i.e. three titanium calibration pins). Accordingly, registration is established by determining the location of these discrete fiducial points with respect to both the preoperative and intraoperative coordinate systems. Upon closer examination, the registration process therefore involves the sequence of measurements depicted in the Figure 3.7 *registration graph* (A full explanation of this diagrammatic technique is provided elsewhere, in Appendix 4). The use of titanium fiducials minimises artefacts during image acquisition, thus allowing the pin positions to be determined with respect to the CT-scanner coordinate system. The preoperative plan is then defined with respect to the same coordinate system.

As previously mentioned, intraoperative position-sensing is performed by bringing a ball probe, which is held by the robot's end-effector, into physical contact with each of the calibration pins. At the time of contact with a calibration pin, the position of the ball probe must therefore be known with respect to the robot coordinate system (i.e. the [robot : fiducial pins] link). As indicated by Figure 3.7, this transformation is established indirectly by a combination of measurement processes (i.e. the subgraph [robot : flange : probe : fiducial pins]). The probe's fixed relationship with the end-effector is initially determined off-line using a coordinate measuring machine (CMM). The intraoperative location of the end-effector, with respect to the robot coordinate system, is then calculated by applying forward kinematics to the measured joint angles. Assuming that the fixed relationship between the surgical tool and the end-effector has been established off-line using the CMM, registration is therefore established by the subgraph [plan : femur : fiducial pins : robot : flange : tool].

However, as highlighted by the induced [robot : fiducial pins] link on the registration graph (i.e. the dashed line), any relative motion between the femur and the robot will invalidate registration. The application of the fixation device prior to the tactile pin finding process, is therefore mandatory for successful registration of the ROBODOC system. As part of the system's safety protocol, a bone motion monitor is also used as a backup. If this sensor detects a displacement of the femur of more than 2 mm, the

milling process is safely interrupted, and the surgeon is instructed by the system to re-establish registration by redoing the tactile search for the pins **[142]**. The need to surgically implant calibration pins into the patient's femur is also perceived to be a shortcoming of the ROBODOC registration strategy. Less invasive surface-based or "pinless" registration techniques are therefore being developed to reduce the trauma experienced by the patient.

The application of surface-based registration techniques to cementless THR procedures is also being investigated by Digioia *et al.* at Carnegie Mellon University (Pittsburgh, PA, USA) **[118]**. In addition to being one of the participating groups in the ROBODOC multi-centre clinical trials, Digioia *et al.* are also investigating robotic milling of the acetabulum as a potential solution to problems associated with the implantation of acetabular THR components **[146]**. A robotic testbed, which is similar to the one used during the early stages of the ROBODOC project, has therefore been commissioned. A number of intraoperative position-sensing techniques have also been evaluated, with a view to developing a "hip navigation" system that allows optimal placement of the acetabular implant. The current implementation of this system employs the Optotrak (Northern Digital Inc., Ontario, Canada) infra-red optical digitiser **[147]**.

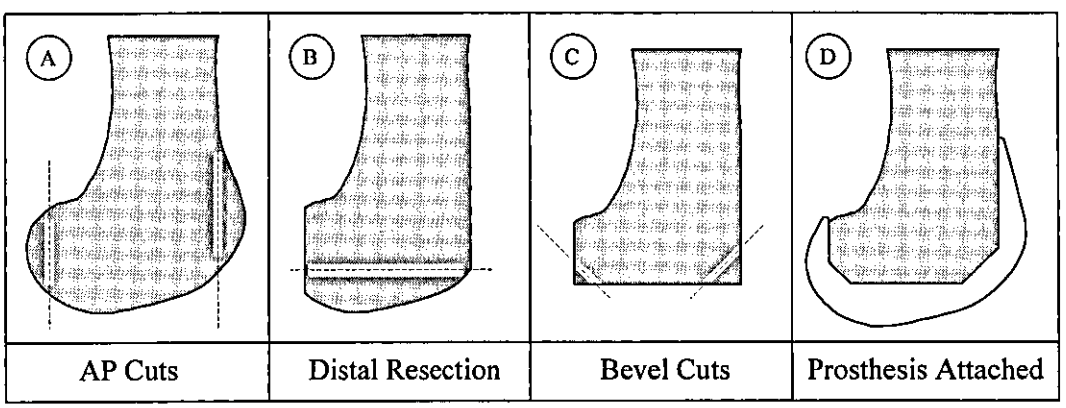


Fig. 3.8 : Total Knee Replacement.
Resection sequence for the femoral component.

Cadaver trials have also been conducted by research groups investigating the potential benefits of robotic-assistance during total knee replacement (TKR) procedures. As

illustrated by Figure 3.8, TKR surgery involves a complex sequence of resections prior to insertion of the femoral prosthetic component. A single resection is also required to insert the tibial prosthetic component. These cuts are currently performed with the aid of consecutive jig-based systems, and as such, a given jig positioning error will effect all of the subsequent cuts. The contact area between a cementless implant and the femur/tibia may therefore be less than ideal for bone ingrowth. However, obtaining an adequate bone/implant interface is not the only criterion governing the positioning of the prosthetic components. In order to restore normal function to the knee joint, the femoral and tibial components must also be accurately aligned with one another. Unfortunately, as only the knee joint is exposed, in practice judging the correct alignment of the entire leg proves to be a difficult task.

Matsen *et al.* **[148]** at the University of Washington (Seattle, WA, USA), have investigated the use of a PUMA 260 robot to position saw and drill guides during femoral arthroplasty procedures. Instead of the usual preoperative CT examination, the registration strategy employed by this group involves the use of 3-D femoral templates. Following immobilisation of the femur, the template which corresponds to the chosen femoral component is mounted on the robot's end-effector. The robot is then switched into passive mode, thus allowing the surgeon to manually bring the template into physical contact with the articular surface of the femur. Having "taught" the robotic system the fixed location of the femur, the corresponding sequence of resections can be calculated. In response to instructions from the surgeon, the robotic system will then position tool guides at the desired locations.

A more advanced, non-invasive robotic TKR system is currently being developed by Kienzle *et al.* **[149]** at Northwestern University (Evanston, IL, USA). The registration strategy adopted by this system is virtually identical to that of ROBODOC. The patient undergoes a preoperative CT examination of the entire leg, with three fiducial markers "attached" to both the femur and the tibia. Preoperative planning is then performed using a graphical model of the patient's leg derived from the CT data. In this way, correct alignment of the femoral and tibial prosthetic components is assured. Following immobilisation of the hip, knee and ankle, intraoperative registration is

established by "teaching" a modified industrial robot the location of the three femoral fiducials. The robot then positions a drill guide, which is used by the surgeon to drill a set of holes in the distal femur. These location holes are then used to accurately attach a cutting block to the femur, thus allowing manual resection of the articular surfaces. The robot is then "taught" the location of the tibial fiducials, and the alignment process is repeated to perform the tibial resection. In future, it is intended that the robot will also align cutting guides.

However, the Northwestern TKR system does differ from ROBODOC in one respect: a combination of artificial and anatomical fiducials is used to reduce the invasiveness of the registration procedure. To avoid preoperative insertion of landmark pins into parts of the leg that are not exposed during surgery, a variety of fiducial marker types are employed. Thus, in addition to a pair of titanium screws implanted in the proximal (knee) end of the tibia, an external landmark pin is also held in place by a fibreglass cast at the patient's ankle. For the same reason, femoral registration is achieved using two titanium screws at the distal (knee) end of the femur, and an "inferred" anatomical landmark defined by the centre of the femoral head. This point is identified with respect to the preoperative coordinate system by selecting a series of imaged points on the surface of the femoral head, and calculating the corresponding centre point. Intraoperative determination of the centre of the femoral head, requires the surgeon to manually flex and abduct the entire leg while the robot tracks the corresponding locations of the knee. The required point is then calculated by fitting a sphere to the recorded knee positions.

A number of research groups are also known to be developing invasive TKR systems. The robotic system developed by Fadda *et al.* **[150]** (Istituti Orthopedici Rizzoli, Bologna, Italy) employs a point-based registration strategy that relies upon implanted titanium screws. A non-invasive registration technique involving a laser scanner system is also being investigated by this group **[151]**. Davies *et al.* **[152]** (Imperial College, London) on the other hand, are investigating the use of force control to actively constrain the motions of a custom-built semi-active robot. This system requires the surgeon to manually move a cutter mounted on the end of a robot arm. By

varying the stiffness of the arm the surgeon is effectively "guided" to make the pre-planned sequence of cuts. It is also reported that the ROBODOC system is currently being adapted to perform the TKR procedure **[143]**. However, it remains to be seen whether improved surgical precision, obtained through the use of robotic-assistance, will actually prolong the life span of cementless prosthetic joints.

3.3.2 Spinal Fusion Surgery.

The surgical treatment of spinal injuries (e.g. fractures, dislocations, etc.) and degenerative conditions increasingly involves stabilisation of the effected vertebrae using an internal fixation device. In such cases, two or more vertebrae are "fused" together by using a series of screws to rigidly attach them to a fixation device (i.e. a metal plate or rod) which acts as a splint. In order to maximise the purchase obtained by these screws, they are frequently inserted into the pedicles of the vertebral bodies in the manner shown in Figure 3.9.

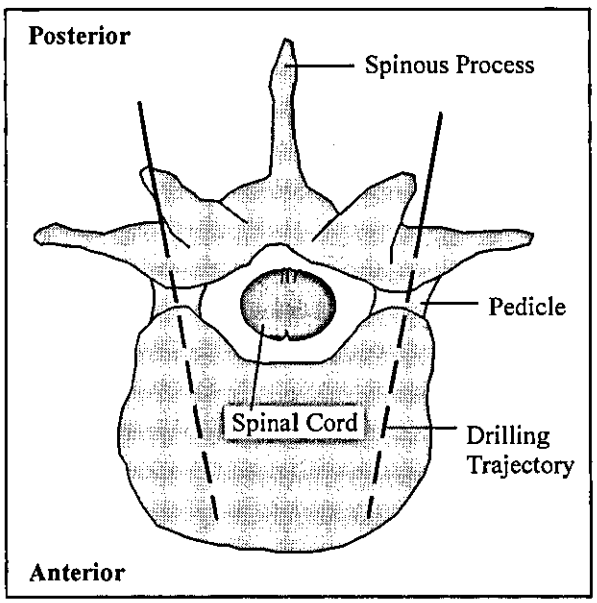


Fig. 3.9 : Ideal Trajectories for Transpedicular Screw Insertion

A posterior surgical approach is therefore required, with a large midline incision being performed to gain direct access to the posterior aspect of the vertebrae **[153]**. The entry points for the transpedicular screws are then identified with the aid of a number of visible anatomical landmarks on the vertebrae. Image intensification is then used to determine the direction of the pedicle axis, thus allowing a Kirschner wire to be

inserted through the centre of the pedicle. If required, the diameter of the screw-hole can then be enlarged using a cannulated surgical drill.

Owing to the narrow diameter of the pedicle, existing techniques for transpedicular screw insertion incur unacceptably high complication rates **[154, 155]**. Given the close proximity of the spinal cord, misplaced screws can cause substantial neurovascular damage. Loss of spinal stabilisation can also occur as a result of misplaced screws working loose. Frameless stereotaxy techniques are therefore currently being adapted to improve localisation of the pedicle axis. Since the vertebra provides several easily identifiable anatomical landmarks, reasonable accuracy can be obtained without the use of artificial fiducials. The combination of preoperative CT imaging and intraoperative position-sensing also eliminates the need for image intensification, thereby reducing irradiation of the surgical staff. However, as prolonged immobilisation of the spine is not possible, registration strategies have to take into account any vertebral motion that occurs during surgery. Continuous tracking of the patient's spine is therefore a prerequisite of accurate computer-assisted spinal fusion surgery.

The two CAS projects that have reached clinical trials, both employ the Optotrak 3020 (Northern Digital Inc.) system to perform intraoperative position-sensing and continuous tracking of the patient's vertebrae. The Orthopaedic Surgery Planning System (OSPS) developed by Nolte *et al.* **[124]** (University of Bern, Switzerland) is a modified version of an existing neurosurgery system. The original clinical trials of this "pedicle navigator" were performed in 1994, and have since been followed up by a number of larger trials **[156, 157]**. A similar system developed by Lavallée *et al.* **[125]** (TIMC, Faculté de Médecine de Grenoble, France) is also reported to have undergone clinical trials in 1994. However, this system represents only a fraction of the research conducted into computer-assisted spine surgery at Grenoble. In addition to the video-based registration method previously mentioned in Section 3.2, Lavallée *et al.* **[158]** have also investigated localised ultrasound and X-ray based registration techniques.

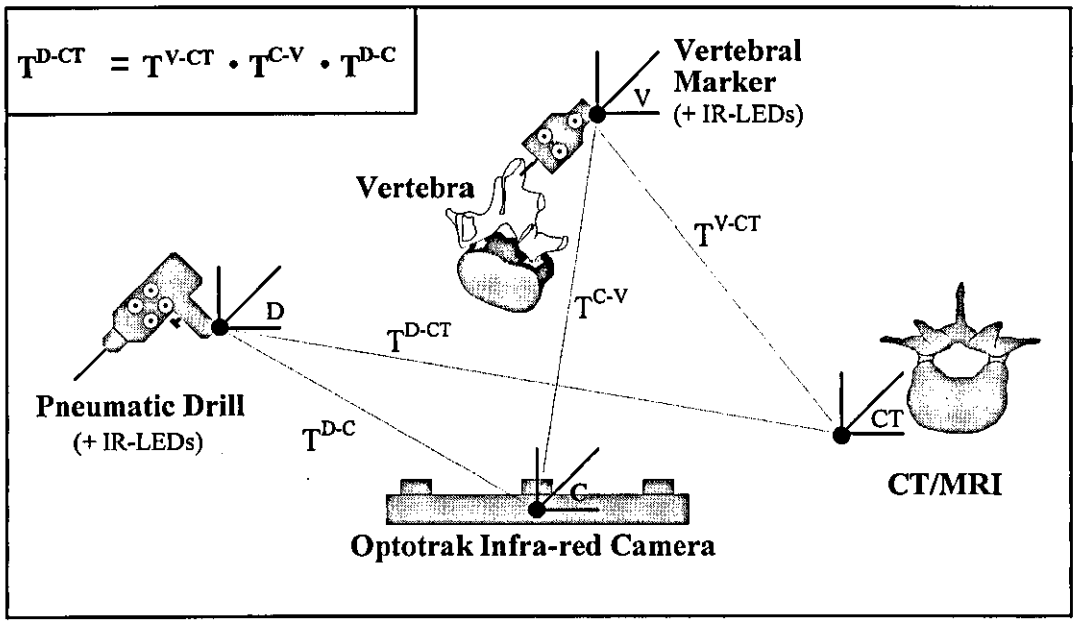


Fig. 3.10 : Optotrak-Based Computer-Assisted Spine Surgery.
(Adapted from Nolte *et al.* [159])

The Bern and Grenoble Optotrak-based systems both employ the components shown schematically in Figure 3.10 (a registration graph which is also applicable to both systems is supplied elsewhere in Appendix 4). Preoperative planning is performed by interactively marking the desired screw positions on CT/MRI images of the vertebrae. These screw positions are defined in terms of entry points and target points, and thus allow surgical trajectories to be established. Following the standard surgical exposure of the spine, a vertebral marker is firmly attached to the spinous process of the vertebra undergoing surgery. When tracked by the Optotrak system, the infra-red LEDs mounted on this vertebral marker define a dynamic operating room coordinate system. An intraoperative position-sensing capability that compensates for vertebral motions can therefore be achieved by simultaneously tracking the vertebral marker, and a digitising probe which is also equipped with IR-LEDs. The Bern group initially use this digitising probe to locate a series of anatomical landmarks (typically 4-8 points) on the vertebra, and then implement a point-based registration technique. If required, a further 30-60 random surface points are then digitised to allow refinement by surface-based registration. By contrast, the Grenoble group digitise approximately fifty random surface points, and directly apply a surface-based registration technique.

The Bern and Grenoble systems also employ similar intraoperative guidance techniques. The Optotrak system is used to track the position of a surgical drill equipped with IR-LEDs. Since registration has been established, the coordinate transformations summarised in Figure 3.10 allow the drill's position to be displayed in real-time with respect to the preoperative images. A graphical user interface can therefore be used to aid alignment of the drill bit with the desired trajectory/pedicle axis. In practice, this alignment process is achieved by displaying a CT/MRI image that is orthogonal to the desired trajectory. In this image the pre-planned trajectory appears as a single point, and is highlighted by an appropriate symbol (e.g. a coloured circle or cross). The position of the tip of the drill bit is then superimposed in real-time onto this image. By following his progress on the graphical user interface, the surgeon is therefore able to bring the tip of the drill bit into contact with the desired entry point. In order to achieve alignment with the pedicle axis, the drill must then be pivoted about this entry point. The position of a second point on the drill bit is therefore superimposed in real-time onto the CT/MRI image, thus allowing perfect alignment to be seen to be established when the three displayed points are coincident.

A number of studies have also reported preliminary results for systems that do not (as yet) take vertebral motions into consideration. An attempt by Roberts *et al.* **[160]** (Dartmouth-Hitchcock Medical Centre, NH) to apply a sonic localisation technique to extracranial neurosurgery procedures involving the lumbar spine, proved to be unsuccessful. Since relative motion between the skin covering the lower back and the spine is far greater than the corresponding motion between the scalp and the skull, the skin markings/temporarily attached fiducials registration technique proved to be totally unsuitable for spine surgery applications. Measurement errors of over 6 mm were therefore encountered when implementing a microscope-based image fusion system. The frameless stereotaxy system evaluated by Amiot *et al.* **[139]** (Université de Montréal, Canada), used an electromagnetic localisation technique to track the location of a Kirschner wire "pointing tool". When using a point-based registration strategy in conjunction with preoperative CT scanning, laboratory trials of this system indicated a total error of 4.5 mm ± 1.0 mm.

Glossop *et al.* **[110]** (University of Toronto, Canada) are currently evaluating the application of the Viewing Wand (ISG Technologies) neurosurgical navigation system to spine surgery. This system's passive digitising arm has been used in conjunction with preoperative CT scanning to perform *in vitro* tests on a section of human cadaver spine. Point-based registration using seven vertebral landmarks led to an average measurement error of 3.66 mm, which was further reduced to 2.8 mm by performing an additional surface-based refinement involving 30-35 random surface points. The feasibility of incorporating real-time vertebral motion tracking into the registration strategy is also being investigated **[126]**.

Glossop *et al.* **[110]** have also proposed an alternative motion compensation technique, involving the attachment of a second articulated arm to the vertebra. The resulting "two-arm" system would be very similar to one that is currently being developed by Besant *et al.* **[161]** (Imperial College, London).

3.3.3 Osteotomies Of The Hip.

The generic term *osteotomy* is used to describe a surgical procedure that involves the cutting of bone. In the context of orthopaedic surgery, such a procedure may be necessary as part of the surgical treatment of several hip related conditions:

- "Non-anatomic" reduction of hip fractures.
- Correction of improper fracture healing (i.e. nonunion or malunion).
- Treatment of degenerative joint disease (e.g. chronic osteoarthritis).
- Correction of congenital deformity.
- Correction of asymmetrical bone growth (e.g. femoral shortening/lengthening).

The majority of hip osteotomies are currently planned using an orthogonal pair of preoperative X-ray images. From this limited two-dimensional information, the surgeon has to determine the location of the resection plane(s) that will provide the required modification of the hip joint or realignment of the femur. In addition, it is also necessary to determine which type of internal fixation device will provide adequate stabilisation of the newly created bone fragments. Execution of the surgical plan is then performed in a free-hand manner using a hand-held pneumatic saw. As

such, the surgeon's only form of guidance during a three-dimensional bone cutting procedure, comes from two-dimensional X-ray images of the corresponding patient anatomy. The outcome of an osteotomy procedure is therefore directly related to the experience and ability of the surgeon.

In an attempt to improve the treatment of osteoarthritic joints, Lord *et al.* **[162]** (Massachusetts Institute of Technology, Cambridge, MA, USA) have developed a computerised surgery simulation tool aimed at intertrochanteric osteotomy procedures. This system uses the results of a CT or MRI examination to reconstruct three-dimensional surface mesh models of the patient's hip. By interacting with these models, the surgeon is therefore able to plan the desired resection planes with the benefit of enhanced visualisation of the patient's anatomy. An "expert" system then evaluates the proposed surgical plan with respect to several criteria (e.g. femoral head contact area, cartilage coverage, joint kinematics, etc.), and gives an indication of the predicted outcome of the procedure. In this way it is possible to optimise the preoperative planning of the resection planes.

A similar computerised simulation system has also been developed by Moctezuma *et al.* **[163]** (Technical University of Munich, Germany). In addition to preoperatively planning the osteotomy procedure using CT derived models, this research group also intends to use a robotic arm to accurately reproduce the planned resections. A prototype robotic-assisted osteotomy system has therefore been commissioned. This system employs a clean room PUMA 560 robot to position a custom-built sawing device **[164]**, and has already been used to perform a series of tests on human cadaver femurs. It is reported that intraoperative registration of the system will be established "with the aid of an X-ray based sensor" and three metal pins implanted in the femur **[163]**.

However, the approach which has achieved the most success with hip osteotomies, is the "individual templates" concept advocated by Radermacher *et al.* **[165]** (Aachen University of Technology, Germany). This technique initially requires a 3-D surface model on the femur/pelvis to be reconstructed from preoperative CT image data. A

CAD/CAM package then uses this model to generate a machining sequence which, when downloaded to a NC milling machine, allows a 3-D template or "negative" of a relevant section of the model to be machined into a plastic block. Assuming that the CT image data is accurate, this template should also be an exact negative of the actual bone surface. Intraoperative registration can therefore be established by bringing the template into contact with the corresponding bone segment, and performing "hardware-based" surface matching to find the unique orientation in which the template is in full contact with the bone. As illustrated by Figure 3.11, by incorporating tool guides into the individual templates, preoperative surgical plans can then be accurately reproduced without resorting to the intraoperative use of a robot.

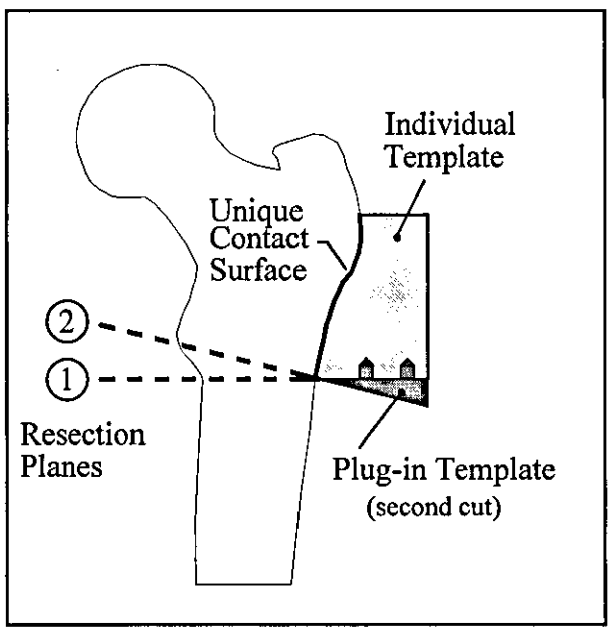


Fig. 3.11 : The Individual Templates Technique.
The templates shown are for intertrochanteric femoral osteotomy.
(Adapted from Radermacher *et al.* [165])

The first surgical procedure involving individual templates was performed in 1992 by Radermacher *et al.* **[166]**. During this triple hip osteotomy the required resections were guided by three independent, sterilised templates. Verification of the surgical results was then obtained using intraoperative X-ray imaging. The success of this procedure, has led to the individual templates technique subsequently being applied to a wide variety of orthopaedic applications (e.g. spine surgery, TKR, THR, etc.).

However, the need to establish physical contact between the bone and the template does limit the scope of the technique to open surgery procedures.

3.3.4 ACL Reconstruction.

The following discussion of ligament reconstruction is included, to highlight the fact that the bone machining applications covered in Sections 3.3.1-3.3.3, represent only a small fraction of the scope of orthopaedic surgery. In dealing with the musculoskeletal system, orthopaedic surgeons are also required to perform a wide variety of procedures involving non-rigid structures (e.g. muscles, tendons, cartilage, nerves etc.). Due to the registration problems created by the intraoperative deformation of "soft" tissues, most robot and computer-assisted surgery applications have avoided these more demanding procedures. However, in cases where the surgical outcome depends upon the accurate positioning of a non-rigid structure in relation to adjacent bone structures, soft tissue procedures can obviously be performed using one of the established "skeletal" registration techniques. The surgical reconstruction of ruptured anterior cruciate ligaments (ACLs) is a prime example of this type of procedure.

The function of the ACL, which connects the tibia and the femur, is to prevent dislocation of the knee by restricting certain motions of the joint. In order to restore normal knee joint kinematics, a torn ACL must therefore be replaced with a ligament graft. The long-term success of this reconstruction procedure depends upon the positioning of the graft; ideally when inserted the graft should maintain a constant length during flexion-extension of the knee joint. Consequently, the existing surgical procedure requires an accurately positioned tunnel to be drilled in both the femur and the tibia. The graft is then securely fixed into position inside these tunnels. However, as this procedure is currently performed under arthroscopy (i.e. endoscopic visualisation of the knee joint), accurate placement of the tunnels is a non-trivial task.

Accordingly, Dessenne *et al.* **[128]** (TIMC-IMAG, Grenoble, France) have developed an Optotrak-based approach which optimises placement of the ACL graft. In order to allow this process to be achieved without preoperative imaging or intraoperative

X-rays, reference markers equipped with IR-LEDs are firmly attached to both the femur and the tibia. The Optotrak system is then used to record the relative locations of these two reference markers at a variety (typically 20-50) of knee flexion angles. The surgeon then uses an IR-LED equipped pointing device to digitise random surface points (typically 20-100) in the vicinity of the proposed tibial and femoral attachment sites. Both sets of data are then used to automatically calculate attachment sites which ensure that the graft will maintain an almost constant length over the full range of joint motion.

By simultaneously tracking the locations of the two reference markers, and a surgical drill equipped with IR-LEDs, it is possible to display the position of the drill in real-time with respect to surface models derived from the digitised points. A graphical guidance technique can therefore be used to ensure that the surgeon aligns the drill with the optimum attachment sites, thus allowing accurate drilling of the tibial and femoral tunnels. It is reported that this technique has been validated on twelve patients **[128]**. However, only long-term postoperative follow-up studies will reveal if the system truly represents an improvement upon existing surgical procedures.

3.4 Intraoperative Quantitative X-ray Imaging.

For many years, the application of photogrammetric reconstruction techniques to intraoperative fluoroscopy was deterred by the non-linear imaging properties of the image intensifier tube. However, it has now become possible to overcome this hurdle using a PC-based digital image processing system and image dewarping algorithms. A limited number of research groups have therefore begun to investigate the potential benefits offered by intraoperative quantitative X-ray imaging. In the majority of cases, this research is motivated by the current trend towards *minimal access surgery* (MAS). As the name implies, the aim of this high-tech form of surgery is to minimise the extent of the surgical incision, thus reducing the surgical trauma experienced by the patient. One method of achieving this goal, is to perform surgical procedures under fluoroscopic guidance, rather than direct physical access to the surgical site.

Accordingly, quantitative fluoroscopic imaging has the potential to allow existing "open" surgery procedures to be accurately performed using MAS techniques.

For the purposes of the following critique, research projects employing X-ray based intraoperative registration strategies have been divided into two groups. Section 3.4.1 covers three applications which are conceptually similar to those proposed by the current thesis, in that they all involve fluoroscopically-guided drilling or guide wire insertion procedures. Section 3.4.2 then goes on to discuss projects which are actually investigating the same surgical procedures (i.e. internal fixation of femoral fractures) as the current thesis.

3.4.1 Minimal Access Surgery Applications.

The three research projects discussed in this section, have all developed quantitative fluoroscopic guidance techniques with a view to performing surgical procedures percutaneously (i.e. through "unbroken" skin). Potamianos *et al.* **[167]** describe a surgical navigation system aimed at improving existing percutaneous needle insertion procedures. As these procedures do not involve preoperative MRI or CT scanning, this system is a rare example of a computer-assisted surgery application in which surgical planning is performed purely on the basis of intraoperative imaging. However, the feature which really makes this application noteworthy, is the fact that it deals with deformable or "soft" tissues. By contrast, Santos-Munné *et al.* **[168]** and Lavallée *et al.* **[158]** are both investigating the percutaneous placement of pedicle screws. As previously discussed in Section 3.3.2, existing surgical techniques to perform these spinal fusion procedures require direct physical access to the patient's vertebrae. The successful outcome of either one of these projects would therefore allow a major change in current surgical practice.

The computer-assisted surgery system described by Potamianos *et al.* **[169]** (Imperial College, London) is initially targeted at percutaneous nephrolithotomy (PCNL). The aim of this MAS procedure is to restore normal renal function by removing renal calculi (i.e. "kidney stones") from the patient's kidney. In practice, this is currently achieved by inserting a narrow cannula into the kidney. MAS instruments are then

introduced down this cannula to grasp and remove the stone(s) under endoscopic vision. However, prior to the insertion of the cannula, a percutaneous track must initially be established to gain access to the target site in the kidney. Under fluoroscopic guidance, a needle is therefore inserted into the kidney via the patient's lower back. A guide wire can then be introduced through this hollow needle, thus allowing subsequent dilation of the track by a consecutive series of dilators.

Unfortunately, owing to the inadequacies of the mobile C-arm unit, the initial "access" stage of the PCNL procedure is currently performed in the radiology suite, rather than the operating room. A radiologist is therefore required to perform the complex three-dimensional task of percutaneous needle insertion, under two-dimensional image guidance. The urologist then has the option of either completing the surgical procedure in situ, or transferring the patient to the operating room. As this scenario obviously represents an inefficient use of medical resources, Potamianos *et al.* **[167]** have developed a 3-D navigation and guidance system, with the aim of allowing the entire PCNL procedure to be performed in the operating room. This system consists of a custom-built passive robotic manipulator (PRM) interfaced to a mobile C-arm unit via a personal computer. Registration of the system, allows the position of a needle held in the PRM's end-effector to be superimposed on 2-D fluoroscopic images of the patient. The existing trial-and-error based needle insertion technique, is therefore replaced by a precise image-guided manipulator alignment process.

Given that preoperative imaging is not required for this application, strictly speaking, intraoperative "registration" of the system actually involves an *image/sensor calibration* process. Nevertheless, for reasons of convenience, the term "registration" is used throughout the following discussion to imply this image/sensor calibration process. In order to superimpose the needle's position onto the intraoperative fluoroscopic images, registration must be performed for every C-arm configuration used to acquire a fluoroscopic image. Noting that their mobile C-arm unit could not be accurately repositioned, Potamianos *et al.* **[167]** concluded that registration would have to be established by imaging the patient and a calibration object together. However, due to the wide variety of patient torso sizes and the limited clearance

between the image intensifier and the patient's lower back, it was also concluded that the use of an intraoperative calibration cage was not practicable during kidney surgery procedures. Three registration strategies have therefore been evaluated as alternatives to the calibration cage technique, with Figure 3.12 depicting the preferred technique.

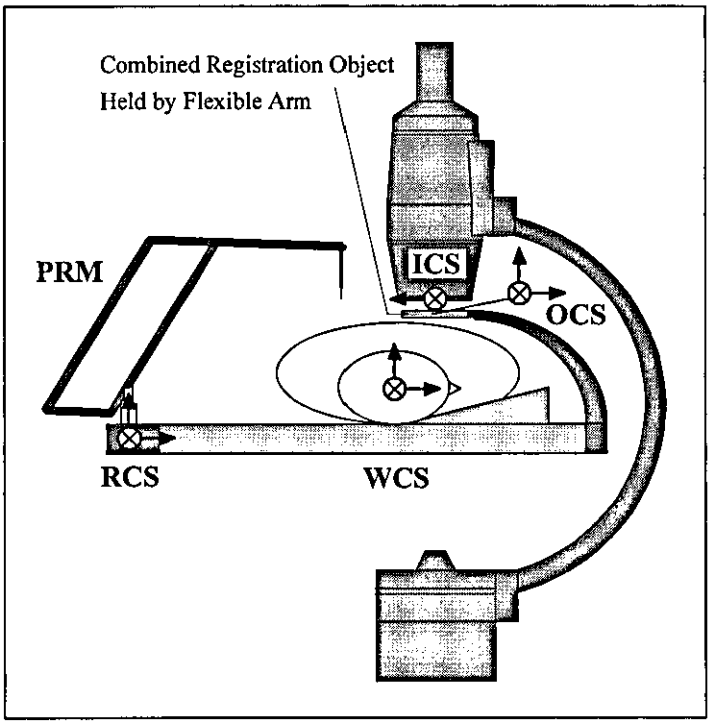


Fig. 3.12 : Intraoperative Registration for Minimally Invasive Renal Procedures.
(Adapted from Potamianos *et al.* [167])

For each of these scenarios, the origin of the operating room or *world coordinate system* (WCS), is defined as being a point inside the patient, which is either on or slightly above the operating table. By rigidly mounting the PRM on the operating table, a fixed relationship is established between the WCS and the *robot coordinate system* (RCS). Since the surgical tool (i.e. the needle) is held by the PRM's end-effector, applying forward kinematics to the PRM's potentiometer readings allows the tool's location to be established with respect to the RCS, and hence the WCS. The registration process therefore requires a transformation to be established between the WCS and a two-dimensional *image coordinate system* (ICS). This ICS is defined by a planar calibration object, which is attached in front of the image intensifier's input window during a preoperative pincushion distortion compensation procedure **[170]**.

Under the registration scenario depicted in Figure 3.12, a radiolucent calibration object is mounted on a flexible arm **[167]**. A series of radiopaque fiducials imbedded in this calibration object, define an (intermediate) object coordinate system (OCS). By acquiring a fluoroscopic image of the patient with this object in the field-of-view, an OCS to ICS transformation can therefore be established for that particular configuration of the C-arm unit. As the calibration object also contains contact registration points, whose coordinates are known with respect to the OCS, the corresponding WCS to OCS transformation can then be established by using the PRM to digitise the locations of these registration points.

However, the only practicable means of introducing a calibration object into the field-of-view, is to position it in the narrow space between the patient and the image intensifier (as shown in Figure 3.12). The calibration object therefore consists of two Perspex plates separated by a distance of only 30 mm, into which a limited number (typically six or more) of radiopaque markers are embedded near the periphery. As a consequence, the patient is actually situated outside the calibrated volume (i.e. the volume between the two Perspex plates). Photogrammetric reconstruction of a target point inside the patient, from two (or more) "calibrated images", must therefore be performed on an extrapolation basis. As previously discussed in Chapter 2, this extrapolation scenario generally leads to larger reconstruction errors when used in conjunction with a non-linear imaging system. Nevertheless, in spite of this acknowledged limitation, *in vitro* trials of this registration scenario suggest that a surgical tool's position can still be displayed with sub-millimetre accuracy on intraoperative images **[167]**.

Returning now to the percutaneous placement of pedicle screws, the robotic systems developed by Santos-Munné *et al.* **[168]** and Lavallée *et al.* **[158]** initially appear to have much in common. In both cases, a CT examination is performed to allow preoperative planning of the surgical procedure. Intraoperative registration is then established by acquiring fluoroscopic images of the patient, while the robotic arm holds a radiolucent calibration object within the field-of-view. Having established registration in this way, the robotic arm is able to align a tool guide with the desired

drilling trajectory. Using this tool guide to define his entry point and orientation, the surgeon can then percutaneously drill the screw-hole. However, upon close examination, the registration strategies adopted by these two systems can be seen to vary considerably. These differences are in accordance with the differing motivations of the two research groups.

In the case of Santos-Munné *et al.* **[168]** (Northwestern University, Evanston, IL, USA), the ultimate aim of the project is to obtain the levels of precision associated with stereotactic procedures, while at the same time remaining as close as possible to current clinical practices. Surgical planning is therefore performed without reconstructing a 3-D model of the vertebra(e) from CT image data. Instead, the system applies a well-established three stage trajectory planning process, which is based upon orthogonal 2-D images of the vertebra(e). The first of these images is derived from a preoperative CT examination, the results of which are used to create a 2-D transverse plane or "slice" through the involved vertebra(e). The system then acquires intraoperative anteroposterior (AP) and sagittal images of the vertebrae, by using a PC-based frame grabber to digitise the video output signal from a mobile C-arm unit. The surgeon then interactively indicates the desired entry points on the AP image, and the desired trajectory angles on the sagittal and transverse images.

Intraoperative registration of the Northwestern system is established by ensuring that the PUMA 560 robot holds a radiolucent end-effector within the field-of-view of both the AP and sagittal images. This dual-purpose end-effector incorporates eight radiopaque fiducials (steel balls imbedded in the corners), as well as a built-in drill guide. Following correction for pincushion distortion, these fiducial points allow a transformation between robot and image coordinates to be obtained for each view. The pre-planned surgical trajectories can therefore be reconstructed in relation to the robot coordinate system, thus allowing accurate alignment of the drill guide. However, from the description given by Santos-Munné *et al.* **[168]**, it is unclear as to whether or not this image calibration process takes respiration-induced vertebral motions into consideration.

The system described by Lavallée *et al.* **[158]** has been developed as part of the computer-assisted medical interventions (CAMI) project at TIMC-IMAG, Faculté de Médecine de Grenoble, France. In accordance with the general methodology advocated by this project **[171]**, intraoperative registration is established using an anatomy-based registration strategy which relies upon matching anatomical structures that are segmented from the preoperative and intraoperative images **[172]**. The preoperative data for this matching process is obtained by reconstructing a 3-D surface model of the patient's vertebra(e) from a set of CT images. The corresponding intraoperative data is then provided by segmenting 2-D surface contours from fluoroscopic images of the patient's vertebrae. The application of a least-squares minimisation algorithm to these data sets then allows a 3-D surface to 2-D projection matching process to be performed.

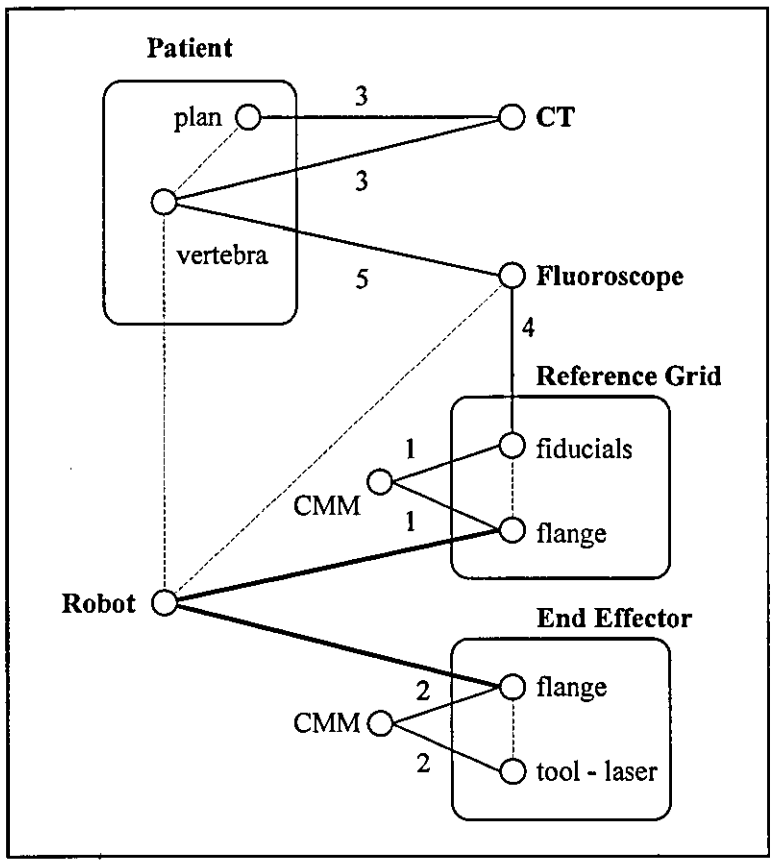


Fig. 3.13 : Registration Graph for the Grenoble Robotic Spine Surgery System.
(Adapted from Lea *et al.* **[145]**)

In practice, implementation of this anatomy-based registration strategy requires the sequence of measurement processes depicted in Figure 3.13. Following a preoperative

CT examination of the patient's vertebrae, the system's DEC 5000 workstation is initially used to perform 3-D surface segmentation. During this semi-automatic process the contours of the vertebra(e) are segmented from each CT slice. The desired 3-D surface model (S) is then generated by interpolating between these segmented vertebral contours. In order to speed up the intraoperative surface matching process, a pre-processing step is also applied to calculate an octree spline based distance map for this 3-D surface model **[172]**. Surgical planning is then performed by interactively indicating target points and straight line trajectories in relation to the 3-D surface model.

It is reported that intraoperative imaging actually involves the use of two X-ray devices **[173]**. As highlighted by the differing event tags (i.e. events 4 & 5) in Figure 3.13, this allows image/robot calibration to be performed independently of patient imaging. The N-Planes B-Spline (NPBS) image calibration technique **[25]** is therefore implemented immediately prior to the acquisition of the two diagnostic X-ray images. As previously explained in Section 2.2.3.1, the NPBS technique is a variation upon the two-plane image calibration method. Accordingly, the robotic arm is used to position a Perspex calibration plate, which incorporates radiopaque fiducial markers, within the field-of-view of the fluoroscopy unit **[174]**. An image of the calibration plate is then acquired, and the process is repeated by moving the plate to a second location within the field-of-view. The calibration plate is then removed to allow the diagnostic X-ray images to be acquired. Segmentation of the 2-D vertebral contours from the diagnostic X-ray images is then performed by interactive designation. However, automation of this segmentation process is also currently under investigation **[174]**.

Having segmented surface data from both the preoperative and the intraoperative images, registration is established by performing a 3-D/2-D matching process. Using interpolation functions derived by the NPBS calibration technique, a line-of-sight vector is initially reconstructed for each of the contour pixels segmented from the intraoperative X-ray images. In theory, when transformed from operating room to preoperative coordinates, these line-of-sight vectors should be tangents to the 3-D surface model (S). Registration is therefore established by defining a signed

Euclidean distance function (d), which estimates the distance from the a line-of-sight vector to the surface (S) **[175]**. Iterative least-squares minimisation of this distance function is then performed by applying the Levenberg-Marquardt algorithm to an initial estimate of the 3-D/2-D transformation parameters. Percutaneous insertion of the pedicle screws can then be performed using a laser-guided tool-alignment technique **[173]**.

Clinical trials of this CAMI spine surgery system have yet to be undertaken. However, the IGOR neurosurgery system discussed in Appendix 3 is also a CAMI application, and as such, the 3-D/2-D matching algorithm has already been successfully demonstrated during 600 surgical procedures **[176]**. *In vitro* laboratory trials of the spine surgery system also indicate that sub-millimetre accuracy is potentially obtainable **[158]**. However, as these tests were performed on an isolated phantom vertebra, the problem of respiration induced vertebral motions was not taken into consideration.

As indicated by the induced [Robot : Vertebra] link in Figure 3.13, any relative motion between the robot and the vertebra invalidates the registration strategy of this system. Real-time tracking of the patient's vertebra will therefore be necessary during clinical trials. It has been proposed that this problem could be solved by using the two X-ray devices to track the vertebral contours **[158]**. However, the merits of this technique are debatable on the grounds of patient irradiation. The Optotrak-based solution of a rigidly attached vertebral marker may therefore have to be adopted, thus further increasing the overall cost of the system.

It is also interesting to note that having separated the image/robot calibration and patient imaging processes, any relative motion between the X-ray devices and the robot can also invalidate registration (as indicated by the induced [Robot : Fluoroscope] link in Figure 3.13). However, for practical reasons two CCD cameras were used to simulate the intraoperative X-ray imaging devices during the laboratory trials. It is therefore unclear as to whether any measures have been taken to ensure that a fixed relationship is maintained between the robot and the two X-ray devices.

3.4.2 Femoral Osteosynthesis Applications.

As previously discussed in Chapter 1, osteosynthesis procedures are ideal candidates for robotic/computer assistance. Predictably, research groups working in this application area have initially targeted the most frequently performed (i.e. repair of femoral neck fractures), and the most technically demanding (i.e. locked intramedullary nailing of femoral shaft fractures) internal fixation procedures. As a consequence, in addition to the MEDROSA Project, three unrelated research groups are also known to be currently investigating femoral osteosynthesis procedures. The first of these applications **[177]** is believed to involve an adapted version of the Optotrak-based spine surgery system developed by Nolte *et al.* (University of Bern, Switzerland). Regrettably, owing to the limited amount of published information relating to this system, an in-depth analysis of its registration strategy cannot be provided in this section. The remainder of the section therefore concentrates upon the research of Finlay *et al.* **[178]** (Armstrong Projects Ltd.) and Phillips *et al.* **[179]** (The University of Hull). However, it is also worth noting that locked intramedullary nailing of femoral shaft fractures is also widely quoted as being a potential future application of the ROBODOC™ system **[180]**.

The MEDROSA project was originally intended to be a joint venture between the Department of Mechanical Engineering at Loughborough University, and Armstrong Projects Limited (Beaconsfield, Hertfordshire, UK). However, after an initial period of collaboration, this partnership failed to materialise. As a result, both groups have subsequently gone on to develop their own robotic-assisted femoral osteosynthesis systems. Given the common origins of these "independently" developed systems, it is inevitable that they are conceptually very similar. Accordingly, the OrthoSista™ system developed by Armstrong Projects Limited also employs an active robotic manipulator to perform the initial X-ray guided drilling stages (i.e. guide wire insertion) of surgical procedures to repair proximal femoral fractures.

In order to gain clinical acceptance, the OrthoSista™ system is initially intended to function as a non-invasive surgical assistant. As such, the custom-built OrthoSista

robot (shown schematically in Figure 3.14) has been designed to align a drill guide with the desired guide wire trajectory. The four degrees-of-freedom required to perform this task are provided by mounting the drill guide support on a sliding cross-member, which in turn is mounted between a double cartesian (i.e. two parallel XY frames) articulation. Each of these cartesian articulations, which consist of a pair of stepper motor driven lead screws, can be independently raised and lowered in a vertical direction. Similarly, the two lead screw carriages can also be independently moved in the horizontal direction. A combination of these four linear motions is therefore used to provide the two translational and two rotational degrees-of-freedom required for drill guide alignment. It is also claimed that this configuration offers exceptional rigidity **[178]**.

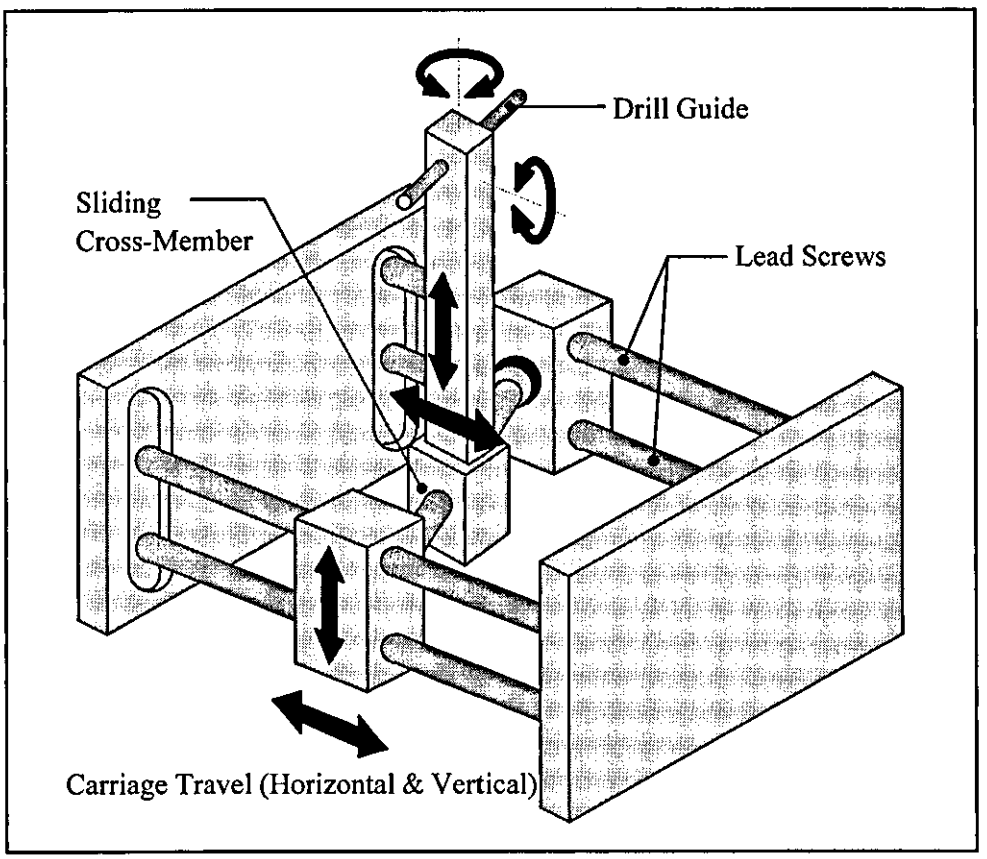


Fig. 3.14 : Schematic of the OrthoSista™ Robot
(Adapted from **[181]**)

Intraoperative "registration" (i.e. image/sensor calibration) of the OrthoSista™ system is to be established by using the guide wire as a temporary calibration object **[178]**. As a result, the guide wire must be located within the field-of-view of the

intraoperative X-ray images. It is therefore proposed that having achieved closed reduction of the proximal femoral fracture, the surgeon will insert the guide wire into the robot-mounted drill guide, and proceed to percutaneously introduce it into the patient's leg until contact is made with the femur. The standard PA and lateral views of the proximal femur will then be acquired using a mobile C-arm unit. These fluoroscopic images will then be digitised and displayed by the system's frame grabber, thus allowing the surgeon to interactively indicate both the desired guide wire trajectory, and the locations of three fiducial markers (i.e. notches on the guide wire). The guide wire will then be removed from the drill guide to allow computer-controlled alignment with the desired guide wire trajectory. The surgeon can then use the robot-mounted drill guide to aid percutaneous insertion of the guide wire into the femoral neck/head.

Surprisingly, Finlay *et al.* **[178]** state that this guide wire based registration strategy was chosen in preference to a previously investigated technique, which involved the use of a robot-mounted calibration frame. By way of explanation, it is reported that while testing of the prototype system, it was discovered that the calibration frame could impede positioning of the mobile C-arm unit. Given that an interactive fiducial segmentation technique was employed, the frame-based approach also proved to be time-consuming. However, the favoured guide wire approach also has a number of inherent limitations. Ideally, the fiducial markers used to establish registration should be evenly distributed throughout the entire field-of-view. In order to compensate for measurement errors, it is also common practice to perform oversampling (i.e. to use more fiducials than are actually required). Unfortunately, neither of these conditions is satisfied when using the guide wire as a calibration object. The description provided by Finlay *et al.* **[178]** also tends to suggest that the guide wire is in contact with the lateral aspect of the proximal femur during image acquisition. Assuming that the standard PA view is acquired, the guide wire will therefore appear at the periphery of the image. Consequently, interactive indication of the fiducial markers is performed in the region of the image that is subjected to maximum pincushion distortion.

As acknowledged by Finlay *et al.* **[178]**, the use of a calibration object that provides only three fiducial markers, also imposes a number of constraints upon the intraoperative imaging geometry. Firstly, a unique solution to the photogrammetric reconstruction problem actually requires at least four non-coplanar fiducial markers **[182]**. As a result, the surgeon must also interactively indicate the location of a marker wire attached to the image intensifier. Secondly, when less than six fiducials markers are used, an analytical technique based upon trigonometric functions must be applied in order to solve the photogrammetric reconstruction problem. In the case of Finlay *et al.* this analytical solution assumes that the PA and lateral views are perfectly orthogonal. However, in practice, achieving this ideal imaging geometry is very difficult when using a mobile C-arm unit. Intraoperative deviations from the analytical model may therefore lead to large reconstruction errors. Finally, for the same reason, it is also imperative that the base of the C-arm unit is not moved between the acquisition of the PA and the lateral images.

By contrast, Phillips *et al.* **[179]** (The University of Hull & Hull Royal Infirmary) are investigating femoral osteosynthesis procedures as part of a research project aimed at developing computer-assisted orthopaedic systems (CAOS). The motivation for this research stems from a desire to enhance implementation of the "Basic Orthopaedic Principle". This term is used by Mohsen *et al.* **[183]** to describe orthopaedic procedures which involve "the placement of an object (guide wire, screw, etc.) at a specific site within a region, via a trajectory which is planned from X-ray based 2-D images or other imaging modalities and is governed by 3-D anatomical constraints". Accordingly, an intraoperative navigation and tool guidance system, referred to as the "non-invasive intelligent orthopaedic guide", has been developed to assist the distal interlocking of femoral intramedullary nails. This system has also subsequently been adapted to assist guide wire insertions during the internal fixation of proximal femoral fractures. A number of applications relating to the spine are also currently under investigation.

As depicted in Figure 3.15, the CAOS system employs three distinct components: an "Intelligent Image Intensifier", a "Trajectory Tactician", and an "Intelligent Trajectory

Guide". The aim of the intelligent image intensifier is to improve the existing intraoperative imaging process, by providing quantitative distortion-free X-ray images. A video link is therefore established between the mobile C-arm unit and the system (via a frame grabber), thus allowing a series of image calibration and registration processes to be undertaken. The trajectory tactician (a procedure specific software module) is then used to select the surgical trajectory. In practice, this is achieved by segmenting features of interest from the distortion free-images, and then reconstructing their three-dimensional locations. A set of surgical parameters are then applied in order to optimise selection of the trajectory.

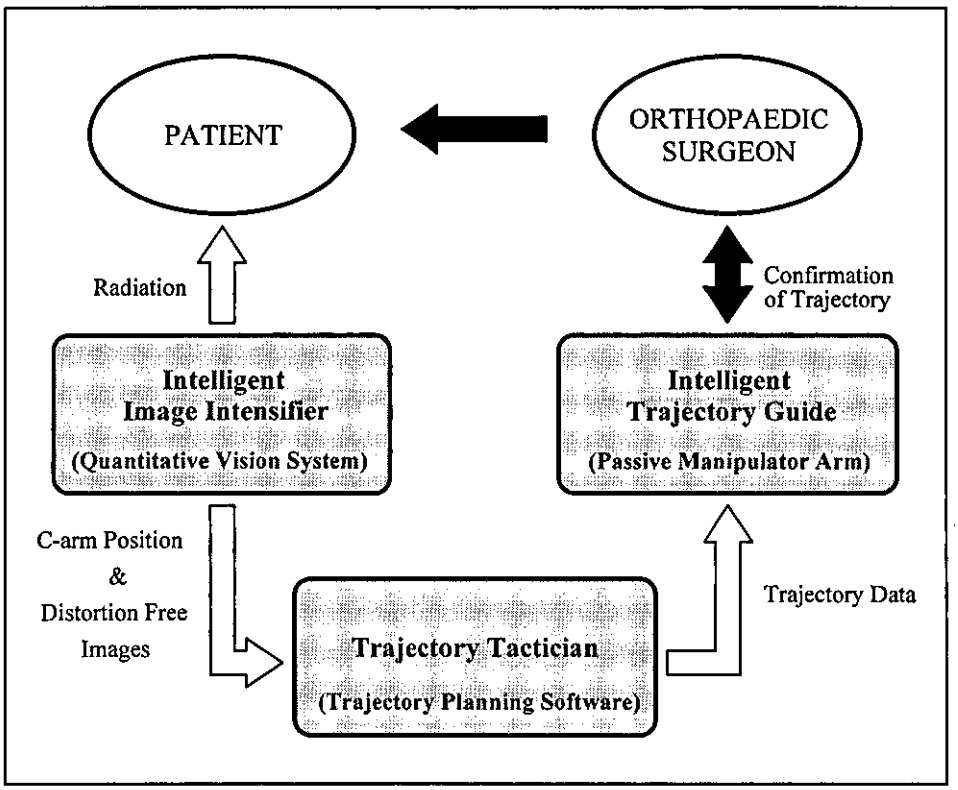


Fig. 3.15 : The CAOS System Architecture.
(Adapted from Viant *et al.* **[184]**)

Having established intraoperative registration, the intelligent trajectory guide is then used to align a drill guide with the optimal trajectory. This surgical guidance system comprises a custom-built, passive manipulator arm, and a real-time graphical display **[179]**. The surgeon is therefore required to manipulate the passive arm until the displayed drill guide location is coincident with the displayed trajectory. An

electromagnetic braking system then allows the arm to be locked in position, prior to manual completion of the drilling process.

The CAOS system employs the Optotrak™ (Northern Digital Inc.) infra-red optical digitiser to register the mobile C-arm unit (intelligent image intensifier), and the passive manipulator arm (intelligent trajectory guide) with the operating room coordinate system. However, prior to establishing intraoperative registration, a number of image calibration procedures must also be performed. The first of these procedures is internal calibration of the image intensifier tube. In order to compensate for image distortion, a radiolucent calibration plate is placed against the input window of the image intensifier housing. This plate contains the usual grid of radiopaque fiducial spheres which, when imaged, allow an image dewarping function to be derived. By applying this dewarping function across an entire image, the CAOS system is able to display "distortion-free" images. However, as previously discussed in Section 2.4.3, when using a mobile image intensifier unit it is very difficult to compensate for the effects of geomagnetic/local magnetic fields. Phillips *et al.* **[179]** therefore acknowledge that their distortion correction scheme does not (as yet) take into consideration magnetic distortion effects.

The aim of the second image calibration procedure is to determine the location of the X-ray source in relation to the image intensifier. For the purposes of this calibration procedure, and all subsequent measurements, the image intensifier's location is defined by the calibration plate that was used to perform distortion correction. This plate represents an ideal image plane or "virtual screen" **[185]**, and allows 2-D image coordinates to be transformed into the operating room coordinate system. As such, the location of the X-ray source is actually determined in relation to this virtual screen, by placing a second calibration plate halfway between the image intensifier and the source. This second radiolucent plate contains both radiopaque fiducial markers and IR-LEDs, thus allowing the Optotrak system to locate the plate and hence the radiopaque fiducials **[179]**. By acquiring an image of the plate, lines-of-sight which pass through each radiopaque fiducial marker and the corresponding imaged point (transformed into virtual screen coordinates) can then be reconstructed. The desired source location is then obtained from the intersection of these lines-of-sight.

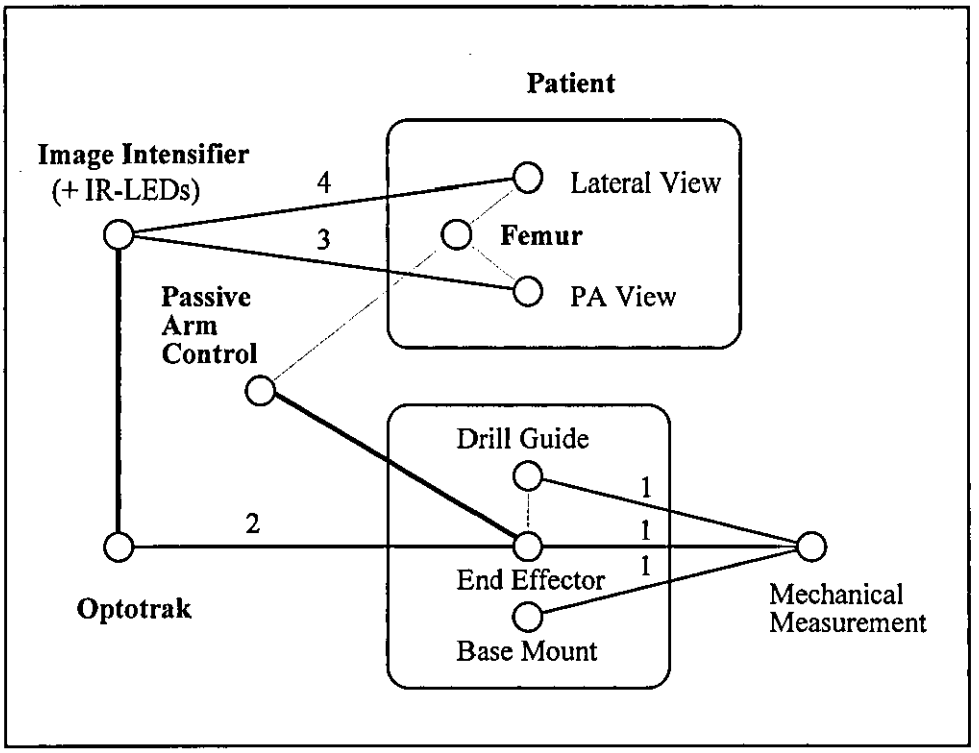


Fig. 3.16 : Registration Graph for the CAOS System.
(Adapted from Phillips *et al.* [179])

The registration strategy of the CAOS system is summarised in Figure 3.16. The measurement processes labelled with event tag number one, perform kinematic calibration of the passive manipulator arm. However, because the passive arm incorporates relative optical encoders, upon power-up, it is also necessary to manually position the end-effector in a home position. The passive arm is then registered with the operating room coordinate system (event 2), by using the Optotrak system to locate a group of IR-LEDs mounted on a specially designed end-effector plate. Similarly, registration of the image intensifier with the operating room coordinate system is also established using a radiolucent plate equipped with IR-LEDs. This plate is attached to the input window of the image intensifier housing, thus allowing an image of its implanted radiopaque fiducial markers to be obtained. However, the plate must be removed in order to acquire the diagnostic PA and lateral images (events 3 & 4) of the femur. A group of IR-LEDs are therefore permanently attached to the side of the image intensifier housing, thus allowing the Optotrak system to maintain registration of the image intensifier.

The use of the Optotrak system, to continuously monitor the location of the image intensifier, allows the mobile C-arm unit to be freely moved without invalidating registration. As long as the IR-LEDs mounted on the image intensifier housing remain clearly visible to the Optotrak system, the C-arm can therefore be arbitrarily positioned to acquire PA and lateral views of the femur. Having measured the image intensifier's pose, the corresponding X-ray source position can then be calculated using the transformation determined by the calibration process. Knowledge of the source and image intensifier locations, then allows three-dimensional reconstruction of features appearing in both the PA and the lateral images. However, as indicated by the induced [passive arm control : femur] link, registration can be invalidated by patient motion.

Unfortunately, Phillips *et al.* **[185]** have found that maintaining an unobstructed line-of-sight between the Optotrak sensor and the C-arm unit is not always practicable in the cluttered operating room environment. Intraoperative position-sensing techniques that eliminate the need for a clear line-of-sight are therefore currently under investigation, with a view to eventually replacing the Optotrak system. It is reported by Mohsen that one of these techniques "utilises an arrangement of markers (typically ten) placed at a known relationship to each other and within the imaged anatomical region" **[186]**. Nevertheless, the Optotrak-based CAOS system is viewed by Phillips *et al.* as being an effective "concept demonstrator". Preliminary *in vitro* trials of the system have verified that it can successfully insert a guide wire (3.2 mm diameter) through the distal holes (6.8 mm diameter) of an intramedullary nail inside a phantom femur **[185]**. A similar set of trials has also established that a guide wire can be inserted into the head of a phantom femur to an accuracy of ±1.5 mm **[183]**.

The trajectory tactician developed by Phillips *et al.* **[185]** to optimise planning of the distal locking procedure, requires the reconstruction of three geometrical features: the long axis of the intramedullary nail, an angle of rotation about this nail axis, and a point on the nail axis which corresponds to the centre of the distal locking hole. Location of the nail axis is performed by interactively indicating the edges of the nail.

Bisecting these two nail boundaries yields a 2-D projection of the nail axis which, when combined with the known X-ray source location, allows a projection plane that passes through the actual nail axis to be defined. As such, the intersection of projection planes calculated from the PA and lateral images, defines the nail axis in relation to the operating room coordinate system.

The desired angle of rotation is obtained by analysing the oval-shaped projection of the distal locking hole **[186]**. Accordingly, an area of interest containing the hole's profile is interactively selected, and a gradient based edge detection algorithm is applied to segment the boundaries of the hole. A correction scheme is then applied to compensate for perspective distortion caused by the hole not being centred in the middle of the X-ray image. The major axis of this corrected hole is then calculated using a least-squares optimisation technique, and a number of measurements of the hole's shape are extracted. Comparison of these measurements with a mathematical model of the distal locking hole, allows the desired angle of rotation to be determined. The required point on the nail axis is then obtained from the intersection of the nail axis with a line-of-sight, which passes through the X-ray source and the centre of gravity of the projected hole.

Having extracted these geometrical features from the X-ray images, the optimum drilling trajectory is defined by the vector which passes through the centre of the distal hole (i.e. the reconstructed point on the nail axis), and has a direction vector which is both perpendicular to the reconstructed nail axis, and at the estimated rotation angle to the horizontal. However, as a vector defined in this manner obviously results in two possible trajectories, a further optimisation stage is required to eliminate the redundant trajectory.

3.5 Summary of Chapter 3.

In the context of robotic/computer-assisted surgery, the term *registration* is used to describe the process of establishing a frame of reference (coordinate system) which is common to all system components. Given that the ultimate aim is to clinically demonstrate distinct improvements over existing freehand surgical techniques, the vast

majority of robotic/computer-assisted surgery applications involve preoperative MRI or CT imaging of the patient. A process referred to as *intraoperative registration* is therefore performed in order to establish a transformation linking the preoperative and intraoperative coordinate systems. Additional *image/sensor calibration* processes may also be required to co-register an intraoperative imaging system with a position-sensing device, passive guidance system, or robotic manipulator.

The aim of this chapter was to place the current thesis into context, by providing an overview of state-of-the-art intraoperative registration techniques. Accordingly, the chapter has covered both the position-sensing technology and the registration strategies adopted by research groups investigating orthopaedic surgery applications. In relation to the current thesis, the nature of the osteosynthesis applications under investigation precludes the use of many of these sensing techniques. In particular, the fact that the distal interlocking of femoral intramedullary nails is essentially a percutaneous procedure, limits the selection to X-ray or ultrasound imaging based techniques. Given the desire to comply with current surgical practices, and to use existing operating room equipment (i.e. the mobile C-arm unit), a passive X-ray imaging technique (i.e. X-ray photogrammetry) has therefore been selected.

In many respects, the orthopaedic applications reviewed in this chapter can be seen to be lagging behind their neurosurgical counterparts (discussed in Appendix 3). Although frameless stereotaxy techniques are beginning to be applied, many orthopaedic applications still rely upon the use of permanently implanted fiducial markers (i.e. invasive registration strategies). Nevertheless, researchers working in the field of orthopaedics have developed a number of interesting alternatives to the algorithm-based registration approach (e.g. individual templates and video-based registration).

With regards to the current thesis, the most significant aspect of the orthopaedic applications reviewed in this chapter is the fact that some of them employ quantitative X-ray imaging in order to establish registration. In line with the current thesis, these X-ray based applications also tend to involve some form of radiolucent end-effector

design, and digital image acquisition via a frame grabber/video link to a mobile C-arm unit. The existence of these independently developed systems is therefore interpreted as providing a strong indication that the objectives of the current thesis are indeed valid.

Of these X-ray based applications, it would appear that only the research of Lavallée *et al.* **[176]** (TIMC-IMAG, Faculté de Médecine de Grenoble, France) significantly pre-dates the current thesis. It is reported that the Grenoble CAMI projects started in 1985, and that an X-ray based registration technique was first applied in conjunction with the IGOR robotic-assisted neurosurgery system in 1989. However, during the original applications of this system, image acquisition was performed by a pair of conventional radiography (i.e. film-based) systems **[187]**, and as such, it is unclear when the transition to digital fluoroscopic imaging was actually made. When this X-ray based registration technique was subsequently adapted to an orthopaedic surgery application (i.e. spinal fusion surgery), two intraoperative fluoroscopy systems had to be used in order to accommodate the preferred modification of the two-plane image calibration technique. In addition to the obvious financial implications, such an arrangement is also undesirable from the point of view of reducing the available work space in the operating room.

Among the contemporary research projects involving X-ray based registration, three are known to be investigating femoral osteosynthesis applications. Two of these projects involve the use of the Optotrak™ (Northern Digital Inc.) infra-red optical digitiser in conjunction with a mobile C-arm unit **[177, 179]**. However, owing to problems arising from the need to maintain an unobstructed line-of-sight between the Optotrak sensor and the object(s) being tracked, Phillips *et al.* **[179]** (University of Hull, UK) are known to be actively seeking an alternative to this approach. The use of the Optotrak™ system also significantly increases the overall cost of a robotic/computer assisted surgery system. The third study involving a femoral osteosynthesis procedure is a commercial venture by Armstrong Projects Limited **[M178]**. The research and development carried out by this company shares the same origins as the current thesis. However, having initially investigated the robot-mounted

calibration frame technique, registration of the OrthoSista system is currently established by using the guide wire as a calibration object. Although this approach is theoretically valid, in practice it severely limits the flexibility of the image acquisition process, and leaves a large scope for error.

In conclusion, a final point worthy of attention is the fact that when adapting X-ray photogrammetry techniques to surgical applications, most research groups have found it necessary to strike a balance between accuracy and ease of use. This point is typified by the percutaneous needle insertion application investigated by Davies *et al.* **[167]**. For reasons of anatomical and imaging equipment compatibility, this research group employs a narrow radiolucent end-effector, instead of the more accurate calibration cage technique. However, the reconstruction errors associated with this technique are still within the allowable limits for the particular surgical procedure under investigation. An acceptable loss of precision is therefore offset against several ergonomic benefits.

4. CAMERA-BASED FEASIBILITY STUDY.

4.1 Introduction.

As indicated in Chapter 1, the limited amount of research funding available during the early stages of the MEDROSA project, precluded the purchase of a mobile C-arm unit, or the construction of a lead-lined radiography enclosure. It was therefore proposed that the objectives of this thesis would initially be investigated by using CCD cameras to simulate the intraoperative images acquired by the mobile C-arm unit. This research methodology is obviously subject to the following limitations:

- The image distortion introduced by the image intensifier tube cannot be investigated.
- Mounting the CCD cameras on adjustable camera stands leads to optical imaging geometries which are more constrained than their fluoroscopic counterparts.
- Tests performed under laboratory conditions cannot fully address issues such as surgical sterility, or compatibility with operating room equipment.

However, in the long-term, it was believed that the MEDROSA applications were ideally suited to a two stage *correction-calibration* approach. Under such a scenario, an image distortion correction scheme is initially applied, in order to compensate for the non-linear imaging properties of the image intensifier tube. Following this internal calibration stage, completion of the fluoroscopy-based photogrammetric reconstruction process is then achieved by applying a simple (linear) external calibration technique. As outlined in Chapter 2, these external calibration techniques are based upon the central projection model of X-ray image formation, which in turn is derived from the collinearity condition. Significantly, the image formation models used by most camera calibration applications (i.e. the pinhole model, the two-plane model, etc.) are also derived from the collinearity condition. Acknowledging that internal calibration would have to be addressed at some point in the future, a camera-based simulation of the external calibration stage of the proposed X-ray correction-calibration approach was therefore a perfectly valid proposition.

The use of CCD cameras to simulate fluoroscopic imaging processes obviously has a number of safety-related advantages. Since a CCD array/camera lens combination produces an image which is not noticeably distorted, it was also believed that trials of a prototype (camera-based) MEDROSA system would be possible at an earlier stage of the project. In order to perform the proposed camera-based feasibility study, an IV120(c) (J-L Automation Ltd., Tyne and Wear) frame grabber card, and a pair of TM526 CCD cameras (PULNiX Europe Ltd., Basingstoke, Hampshire) fitted with 25 mm 'C' mount lenses, were therefore purchased.

This frame grabber card fits into the PC AT main back plane, and allows the acquisition of digital images with grey level resolutions of 8-bits per pixel, or alternatively 6-bits per pixel with two pseudo colour RGB overlay planes. These digital images can then be displayed to either a high (512 x 512 pixel) or a low (256 x 256 pixel) spatial resolution. In terms of software support, the frame grabber card was supplied with a basic menu-driven image processing package (JLGenias), and a library of Pascal-callable software units (JLGenial), which can be linked into Microsoft Pascal programs. In order to take advantage of these library functions, it was therefore decided that the software required by this thesis would be written in the DOS-based Microsoft Pascal (version 4.0) language.

Following a brief description of the registration strategy of the MEDROSA system (Section 4.2), the remainder of this chapter outlines the methodology which was adopted in order to simulate this X-ray based process. Accordingly, Section 4.3 describes an initial investigation of the practical aspects of implementing a digital version of the Direct Linear Transformation (DLT). With a view to obtaining a more X-ray compatible approach to the photogrammetric reconstruction process, the focus of the research then switched to an investigation of the two-plane calibration technique (Section 4.4). The major findings and implications of the research work documented in this chapter are then summarised in Section 4.5.

4.2 Intraoperative Registration Strategy.

As discussed in Chapter 1, the desire to achieve compatibility with existing surgical and fluoroscopic practices has led to the adoption of an X-ray photogrammetry-based registration strategy. Accordingly, it is proposed that intraoperative registration of the MEDROSA system will involve the system components depicted schematically in Figure 4.1. Under this scenario, the custom-built manipulator is used to hold an anatomically compatible calibration frame in the vicinity of the patient's proximal or distal femur (not shown in Figure 4.1). The radiographer will then adjust the mobile C-arm fluoroscopy unit, until the anatomical features of interest and the calibration frame are both within the field-of-view. Either a posteroanterior (PA) or a lateral fluoroscopic image of the patient/frame combination will then be acquired.

In order to acquire the corresponding (i.e. lateral or PA) orthogonal image of the patient/frame combination, the arc of the C-arm unit must then be rotated through ninety degrees. In practice, the calibration frame must therefore be attached to the manipulator's end-effector, using a bracket which is perpendicular to the end-effector's drill feed axis. The offset created by this bracket allows the arc of the C-arm unit to be rotated from the vertical position shown in Figure 4.1, to a horizontal position (or vice-versa), without making contact with either the calibration frame or the robotic manipulator. Having acquired an orthogonal pair of X-ray images, into which a radiopaque fiducial pattern has been superimposed, photogrammetric reconstruction techniques can then be applied to establish intraoperative registration.

As indicated by Figure 4.1, the MEDROSA system's intraoperative registration strategy is therefore defined by a series of spatial transformations, which link four arbitrarily allocated coordinate systems. The most important of these frames of reference is the *manipulator coordinate system* (MCS), which effectively defines the operating room or real world coordinate system. Accordingly, the ultimate aim of the intraoperative registration strategy is to accurately define drilling trajectories with respect to this MCS. In accordance with most robotic applications, the origin of the MCS has been allocated to a non-moving part of the manipulator's base.

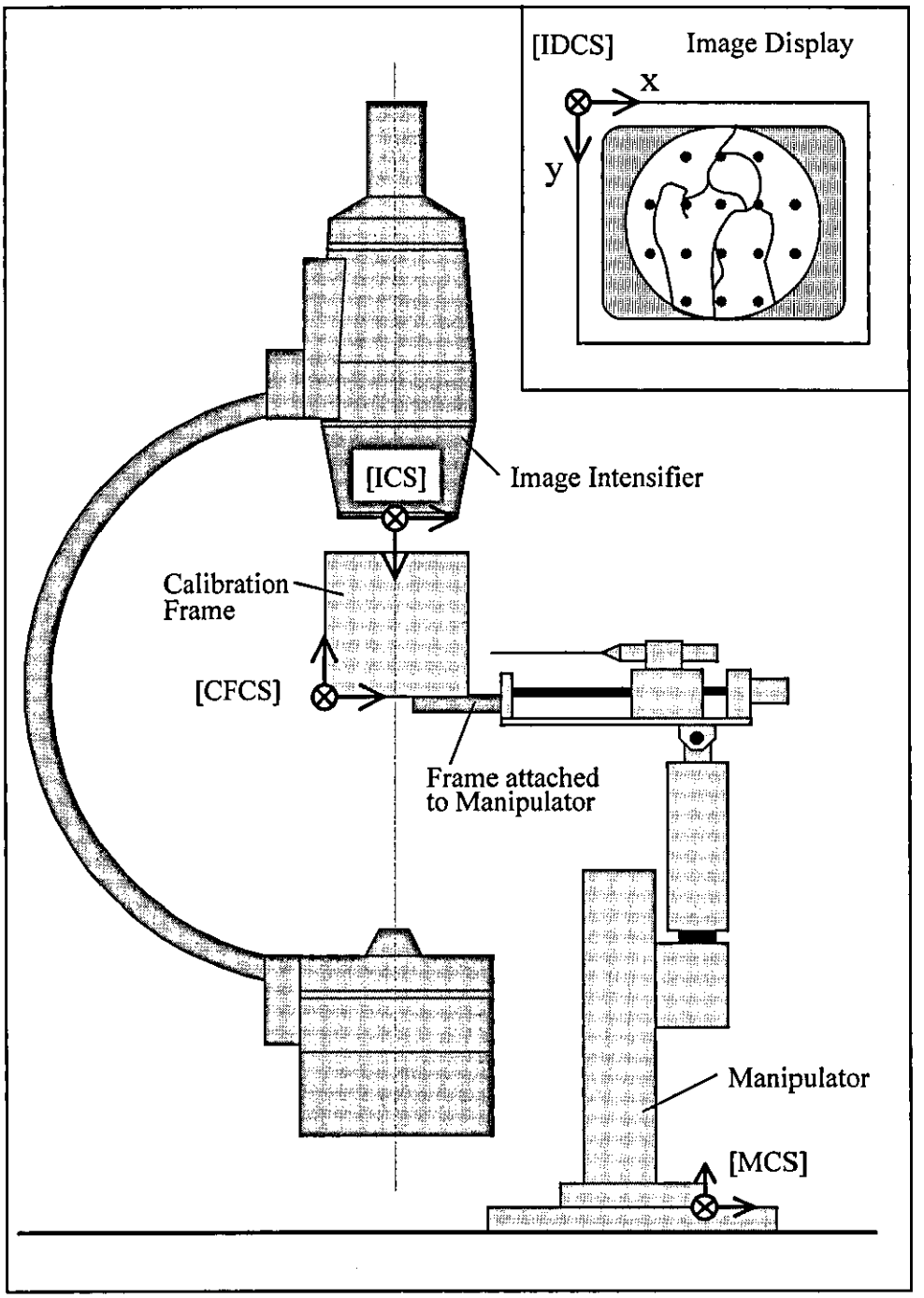


Fig. 4.1 : The MEDROSA System's Intraoperative Registration Scheme.

A kinematic model of the manipulator's configuration (refer to Figure 4.2) has then been developed using the standard Denavit-Hartenberg parameters approach **[189]**. As such, the end-effector's location can be calculated in real-time, with respect to the MCS, by applying forward kinematics to the manipulator's encoder readings and 'known'/calibrated lengths.

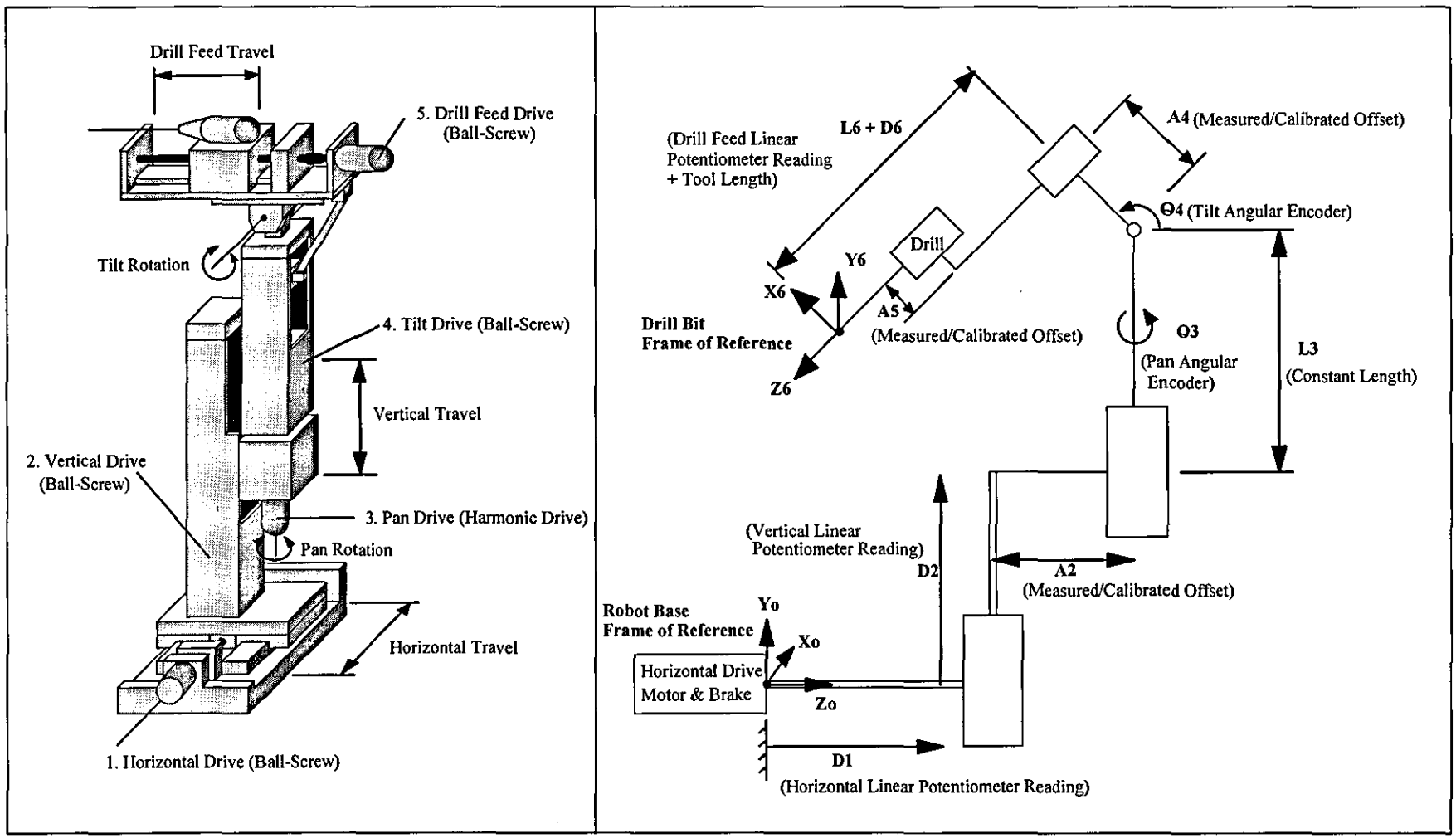


Fig. 4.2 : Kinematic Configuration and Kinematic Model of the MEDROSA Manipulator.

Since the Perspex calibration frame is rigidly attached to the manipulator's end-effector, extending the kinematic model to include both the attachment bracket and calibration frame allows a spatial transformation to be established between the *calibration frame coordinate system* (CFCS) and the MCS. Given that the CFCS is defined by the radiopaque fiducial markers embedded in the calibration frame, and that these fiducial markers are superimposed into the diagnostic fluoroscopic images, the calibration frame therefore fulfils an intermediary role, allowing an image-to-world (or vice-versa) transformation to be determined indirectly.

Having opted to digitise the intraoperative fluoroscopic images, the imaged locations of the fiducial markers are initially measured in terms of pixels, with respect to the *image display coordinate system* (IDCS). By convention, the origin of this two-dimensional IDCS is located at the top left hand corner of the screen. The x-axis therefore runs along the top row of pixels from left-to-right, while the y-axis runs down the first column of pixels from top-to-bottom. Depending upon the image distortion correction scheme adopted, it may subsequently be necessary to convert these pixel values into real world coordinates, with respect to an *image coordinate system* (ICS) located in front of the input window of the image intensifier housing.

As discussed in Chapters 2 and 3, the coordinate system associated with this 'ideal' image plane can be defined by a radiolucent calibration plate containing radiopaque fiducial markers. Should the need arise, a transformation from the ICS to the MCS can be obtained by initially using the images of the calibration frame to reconstruct the location of the X-ray source. Knowledge of the fixed relationship between the X-ray source and the image intensifier housing may then be exploited, to locate the origin of the ICS with respect to the MCS.

Thus in summary, the overall registration strategy of the MEDROSA system involves the measurement processes summarised in Figure 4.3. When represented in this registration graph format, the most obvious feature of the MEDROSA applications is the absence of preoperative CT or MRI scanning. Consequently, surgical planning must be performed solely on the basis of intraoperative fluoroscopic images. As

indicated by the subgraph [robot : end-effector : bracket : fiducials], it is therefore proposed that a combination of preoperative robot calibration (event tags 1 and 2) and intraoperative position-sensing (sustained link [robot : end-effector]) will be used to determine the location of a radiolucent calibration frame, with respect to the operating room coordinate system. The acquisition of PA and lateral fluoroscopic views of the patient's femur, along with the calibration frame (event tags 3 and 4), will then allow the location of the femur to be reconstructed with respect to the calibration frame's coordinate system (subgraph [femur : C-arm : fiducials]), and hence the operating room coordinate system (induced link [femur : robot]).

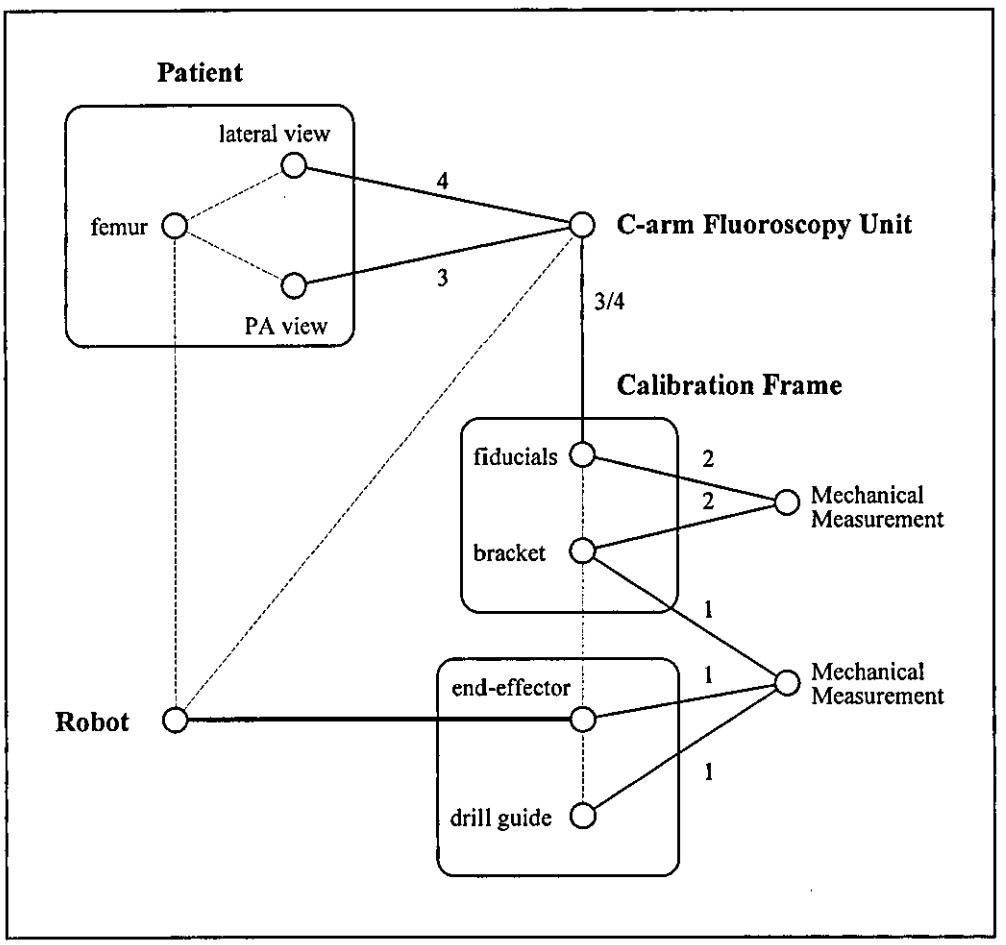


Fig. 4.3 : Registration Graph for the MEDROSA System.
(Manual drilling option).

Having established intraoperative registration the calibration frame will then be removed, allowing the surgical procedure to be performed either manually, using the robot-mounted drill guide (subgraph [femur: robot: end-effector : drill guide]), or

automatically using the end-effector's drilling unit. However, as indicated by the induced [femur : robot] link, registration is only valid as long as there is no relative motion between the patient's leg and the robotic manipulator. Although it is believed that the intraoperative traction applied to the patient's limb, via the fracture table leg supports, will provide adequate immobilisation, this limitation of the proposed registration strategy will have to be evaluated during future clinical trials. If required, additional patient immobilisation techniques will then be applied.

4.3 Preliminary DLT Investigation.

In order to gain familiarity with the practicalities of photogrammetric reconstruction, a laboratory investigation of the Direct Linear Transformation (DLT) was initially conducted. The primary aim of this study was to develop and evaluate software algorithms, which would implement the mathematical techniques required to perform three-dimensional reconstructions from an orthogonal pair of images. These software modules could then be used as building blocks during a more realistic simulation of the proposed X-ray photogrammetry scenario. As a consequence, image processing and analysis techniques were not investigated at this early stage of the research period.

4.3.1 DLT Camera Calibration Mathematics.

As explained in Chapter 2, the DLT establishes a linear relationship between two-dimensional image (x, y) and three-dimensional world (X, Y, Z) coordinates. This relationship is defined by the following equations:

$$x = \frac{L_1.X + L_2.Y + L_3.Z + L_4}{L_9.X + L_{10}.Y + L_{11}.Z + 1} \qquad .....(2.16)$$

$$y = \frac{L_5.X + L_6.Y + L_7.Z + L_8}{L_9.X + L_{10}.Y + L_{11}.Z + 1} \qquad .....(2.17)$$

Which can be rearranged to yield:

$$x = L_1.X + L_2.Y + L_3.Z + L_4 - x.L_9.X - x.L_{10}.Y - x.L_{11}.Z \qquad .....(4.1)$$

$$y = L_5.X + L_6.Y + L_7.Z + L_8 - y.L_9.X - y.L_{10}.Y - y.L_{11}.Z \qquad .....(4.2)$$

As such, when using the DLT approach to perform photogrammetric reconstruction, the eleven unknown transformation parameters (L_i) must initially be determined for each camera. In practice, this camera calibration process is performed by imaging a group of control point markers, whose three-dimensional coordinates are accurately known with respect to the world coordinate system. Given that the measured image (pixel) coordinates of each control point allow a pair of equations of the form (4.1) and (4.2) to be defined, a minimum of six control points must therefore be physically provided in order to calculate the eleven unknown parameters.

When the theoretical minimum of only 'five and a half' control points is used, the most straightforward method of solving the resulting simultaneous equations is to express them in the matrix form shown in Figure 4.4. Such a mathematical problem is usually represented by the shorthand version:

$$\mathbf{[A].x \equiv b} \qquad \text{.....(4.3)}$$

Where:

- The [11 x 1] vector **x** contains the unknown L_i terms.
- The [11 x 1] vector **b** contains the control point image coordinate data.
- The [11 x 11] matrix **[A]** contains the terms which satisfy equations (4.1) and (4.2).

In this form, the required **x** vector coefficients can be calculated using the Gauss Elimination technique. However, in reality, the control point data substituted into equations (4.3) will contain measurement errors. As a consequence, although the calculated **x** vector coefficients represent an exact mathematical solution to the calibration problem at the control point locations, when applied to non-control point locations, the resulting reconstruction accuracy is usually quite poor. The photogrammetric studies reviewed in Chapter 2 therefore strongly recommend the use of *oversampling* (i.e. more than six control points), as a means of compensating for measurement errors in the calibration data. Under this oversampling scenario, equation (4.3) becomes:

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1.X_1 & -x_1.Y_1 & -x_1.Z_1 \\ X_2 & Y_2 & Z_2 & 1 & 0 & 0 & 0 & 0 & -x_2.X_2 & -x_2.Y_2 & -x_2.Z_2 \\ X_3 & Y_3 & Z_3 & 1 & 0 & 0 & 0 & 0 & -x_3.X_3 & -x_3.Y_3 & -x_3.Z_3 \\ X_4 & Y_4 & Z_4 & 1 & 0 & 0 & 0 & 0 & -x_4.X_4 & -x_4.Y_4 & -x_4.Z_4 \\ X_5 & Y_5 & Z_5 & 1 & 0 & 0 & 0 & 0 & -x_5.X_5 & -x_5.Y_5 & -x_5.Z_5 \\ X_6 & Y_6 & Z_6 & 1 & 0 & 0 & 0 & 0 & -x_6.X_6 & -x_6.Y_6 & -x_6.Z_6 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1.X_1 & -y_1.Y_1 & -y_1.Z_1 \\ 0 & 0 & 0 & 0 & X_2 & Y_2 & Z_2 & 1 & -y_2.X_2 & -y_2.Y_2 & -y_2.Z_2 \\ 0 & 0 & 0 & 0 & X_3 & Y_3 & Z_3 & 1 & -y_3.X_3 & -y_3.Y_3 & -y_3.Z_3 \\ 0 & 0 & 0 & 0 & X_4 & Y_4 & Z_4 & 1 & -y_4.X_4 & -y_4.Y_4 & -y_4.Z_4 \\ 0 & 0 & 0 & 0 & X_5 & Y_5 & Z_5 & 1 & -y_5.X_5 & -y_5.Y_5 & -y_5.Z_5 \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \\ L_5 \\ L_6 \\ L_7 \\ L_8 \\ L_9 \\ L_{10} \\ L_{11} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}$$

Fig. 4.4 : The DLT Camera Calibration Equations Expressed in the Matrix Form [A].x = b

Exact solution case using theoretical minimum number of control points (5½).

$$[A].x \cong b \qquad .....(4.4)$$

Where:

- **[A]** is a [m x 11] matrix (m > 11).
- **b** is a [m x 1] vector.

Since the number of simultaneous equations is now greater then the number of unknown coefficients, an exact solution to the calibration problem can no longer be obtained. Equations (4.4) therefore require a 'best fit' solution, which is most readily provided by one of a number of mathematical techniques referred to as the *linear least squares* methods. A detailed discussion of three of these techniques is provided elsewhere in Appendix 5.

Assuming that a linear least squares technique has been applied to calibrate a pair of cameras, the photogrammetric reconstruction process involves the following calculations. Rearranging equations (4.1) and (4.2) we obtain:

$$x - L_4 = (L_1 - x.L_9).X + (L_2 - x.L_{10}).Y + (L_3 - x.L_{11}).Z \qquad .....(4.5)$$

$$y - L_8 = (L_5 - y.L_9).X + (L_6 - y.L_{10}).Y + (L_7 - y.L_{11}).Z \qquad .....(4.6)$$

If the image (pixel) coordinates (x, y) of a target point are known, along with the transformation parameters (L_1 to L_{11}), equations (4.5) & (4.6) are of the form:

$$a_1.X + b_1.Y + c_1.Z + d_1 = 0 \qquad .....(4.7)$$

$$a_2.X + b_2.Y + c_2.Z + d_2 = 0 \qquad .....(4.8)$$

As a result, equations (4.5) and (4.6) each represent a plane in three dimensions containing the imaged target point. Taken as a pair, these two equations (planes) contain sufficient information to define the 'line-of-sight' upon which the target point must lie. Consequently, if two calibrated views are available, the three-dimensional coordinates of the target point can be reconstructed from the intersection of two lines-of-sight. In mathematical terms, the DLT-based photogrammetric reconstruction process can therefore be represented by the matrix equations:

$$\begin{bmatrix} L_1 - x_1.L_9 & L_2 - x_1.L_{10} & L_3 - x_1.L_{11} \\ L_5 - y_1.L_9 & L_6 - y_1.L_{10} & L_7 - y_1.L_{11} \\ M_1 - x_2.M_9 & M_2 - x_2.M_{10} & M_3 - x_2.M_{11} \\ M_5 - y_2.M_9 & M_6 - y_2.M_{10} & M_7 - y_2.M_{11} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x_1 - L_4 \\ y_1 - L_8 \\ x_2 - M_4 \\ y_2 - M_8 \end{bmatrix} \quad \text{.....(4.9)}$$

Where:

- (x_1, y_1) are the image (pixel) coordinates of the target point in image 1.
- (x_2, y_2) are the image (pixel) coordinates of the target point in image 2.
- (L_1 to L_{11}) are the DLT parameters for camera 1.
- (M_1 to M_{11}) are the DLT parameters for camera 2.
- (X, Y, Z) are the required three-dimensional target point coordinates (which are common to both views).

Since the matrix system (4.9) contains four equations and only three unknown parameters (X, Y, Z), it is of the form **[A].x ≅ b**. Having performed two linear least squares calculations to obtain the DLT parameters (L_1 to L_{11}) and (M_1 to M_{11}), the photogrammetric reconstruction process therefore requires an additional linear least squares calculation to solve equations (4.9)

4.3.2 Design of a Suitable Calibration Object.

In addition to the need for six or more control points, the review of photogrammetric research documented in Chapter 2 also highlighted the following requirements:

- The control points must be non-coplanar.
- The control points should surround the measurement volume, thus avoiding extrapolation-based photogrammetric reconstructions.
- The control points should be symmetrically distributed around the measurement volume.

In the long-term, it is proposed that the MEDROSA system will satisfy these requirements by providing a radiolucent (Perspex) calibration frame, into which radiopaque (metallic) fiducial markers have been embedded. However, when using

optical (CCD) cameras, simulating such an arrangement is not a trivial task. The need to provide non-coplanar control points is generally satisfied through the use of two calibration planes. Accordingly, the control point distribution in the calibration plane farthest from the camera, must be clearly visible through the calibration plane nearest to the camera. Unfortunately, Perspex or glass calibration planes can lead to reflection (glare) related problems when used in conjunction with optical cameras.

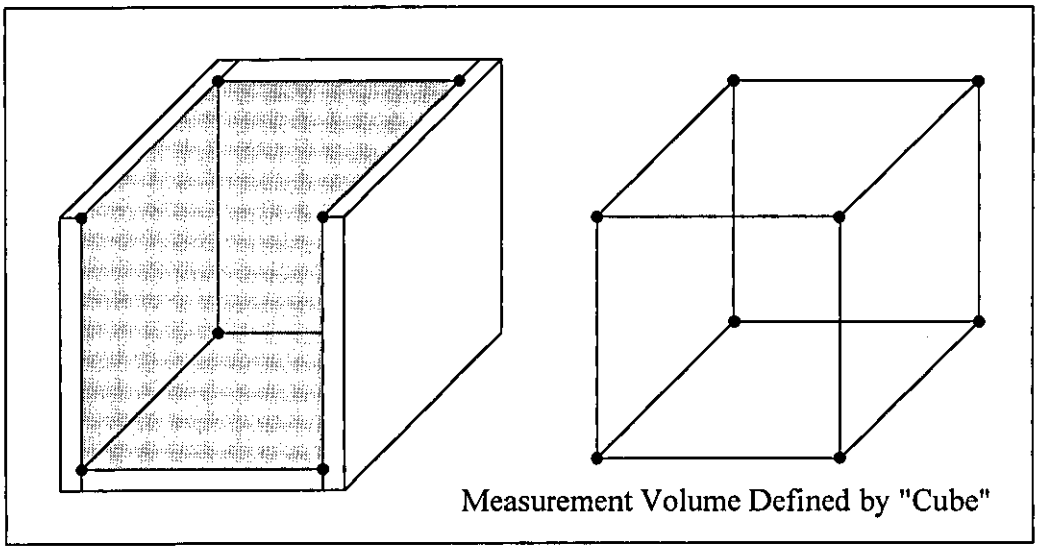


Fig. 4.5 : The Calibration 'Cube' Used During Initial Camera Trials.
The eight internal corners provide easily identifiable control points in both views.

In order to avoid such problems the calibration object depicted in Figure 4.5 was used during the initial DLT investigation. This calibration 'cube' was designed to be compatible with a biplane imaging geometry in which a plan view camera would be mounted one metre above the 'cube', and a side view camera would be mounted one metre in front of the 'cube'. Accordingly, it was manufactured from aluminium plate in such a way that its internal structure defined a 'cube' of 120 mm long sides. Owing to the effects of perspective distortion, the eight corners of this internal 'cube' appeared within the field-of-view of both cameras in the manner shown in Figure 4.6. By arbitrarily allocating one of these internal corners to be the origin of the world coordinate system (X, Y, Z), it was therefore possible to use the calibration 'cube' to provide eight easily identifiable control points.

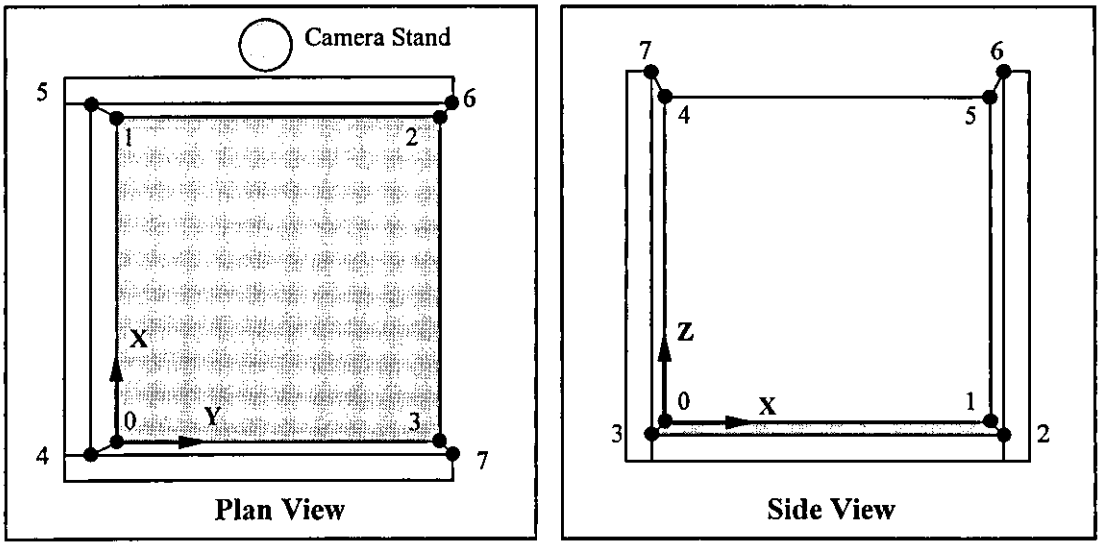


Fig. 4.6 : Perspective Distorted Views as Observed by the Two Cameras.

(A world coordinate system is arbitrarily allocated to an internal corner as shown.)

In addition to the provision of control points, a thorough evaluation of photogrammetric reconstruction accuracy also requires the introduction of independent test (non-control) points within the field-of-view of both cameras. The 'steps' calibration object (120 mm x 120 mm x 60 mm) depicted in Figure 4.7 was therefore designed to be placed inside the calibration 'cube' following completion of the camera calibration process.

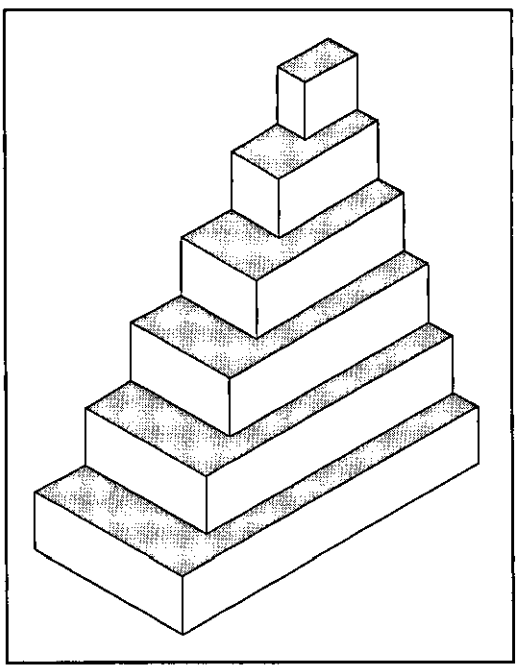


Fig. 4.7 : The 'Steps' Calibration Object.

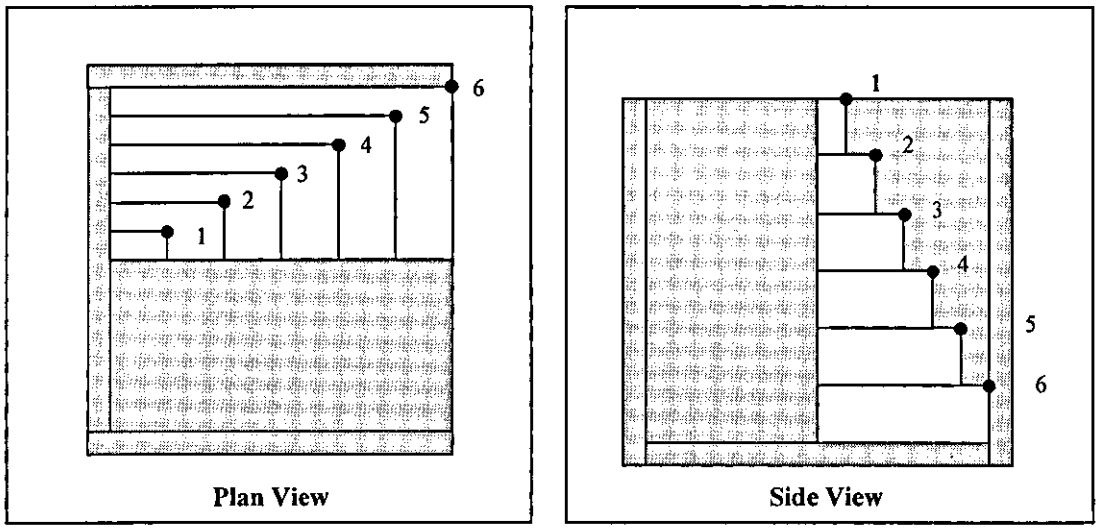


Fig. 4.8 : 'Steps' Calibration Object Shown in One of its Several Possible Configurations within the Calibration Cube.

Once in place, as indicated by Figure 4.8, this layered structure provided six easily recognisable test points of known world coordinates. Through careful design it was also possible to ensure that the 'steps' could be placed inside the 'cube' in several different orientations, thus providing many test point combinations.

4.3.3 Software Development.

Having written a simple software routine to acquire two views (plan and side) of the calibration 'cube', a means of extracting the control point locations from the resulting digital images was required. Additional code was therefore written to allow the interactive indication of the eight internal corners of the 'cube' using a PC-mouse driven cursor. This control point image (pixel) data was then stored in array format, prior to being written to an output file. In this way, it was possible to obtain calibration data of the form shown in Table 4.1.

As outlined in Section 4.3.1, obtaining a solution to the camera calibration (equation 4.4) and photogrammetric reconstruction (equation 4.9) problems requires the use of a linear least squares technique. For comparison purposes, the following programs were therefore developed in order to implement the three linear least squares techniques described in Appendix 5:

- *Pseudo.pas* (a pseudo inverse program developed from first principles).
- *Givens.pas* (a Q-R decomposition program based upon a Givens transformation algorithm supplied by Nash **[190]**).
- *SVD.pas* (a singular value decomposition program based upon an algorithm supplied by Nash **[190]**).

Given that each of these techniques involves a long chain of matrix multiplication processes, substantial errors can occur as a result of the propagation of rounding errors. When the following algorithm checks were implemented, it was therefore discovered that declaring the array elements to floating-point precision (i.e. REAL in MS-Pascal) did in fact lead to poor results:

- Sample Calculations:

  In order to allow a step-by-step verification of the algorithms against calculator obtained values, the programs were initially developed to solve least squares problems involving small matrices. Whenever possible, textbook examples were also used for verification purposes. These algorithms were then adapted to solve the larger systems of matrix equations associated with the camera calibration problem.

- Identity Matrix Checks:

  A matrix [A] multiplied by its inverse $[A]^{-1}$ should yield the identity matrix [I]. Similarly, an orthogonal matrix [Q] multiplied by its transpose $[Q]^{T}$ should also yield the identity matrix [I].

- Residual Checks.

  Having solved the $\mathbf{[A].x \cong b}$ problem, the calculated **x** vector was evaluated by calculating the residual vector $\mathbf{r = [A].x - b}$.

Consequently, a switch to double precision (REAL8 in MS-Pascal) was required to obtain the level of accuracy required by these large matrix computations. This use of REAL8 declarations obviously leads to a substantial increase in the computational effort required to solve linear least squares problems.

Table 4.1 : Typical Control Point Data for Software Evaluation.

	Image Control Point Data (pixels)				World Control Point Data (mm)		
	Plan View		Side View		X	Y	Z
	x	y	x	y			
Control Point 0	128	439	81	467	0	0	0
Control Point 1	120	105	429	473	120	0	0
Control Point 2	461	94	447	490	120	120	0
Control Point 3	468	436	57	480	0	120	0
Control Point 4	104	461	90	128	0	0	120
Control Point 5	97	86	435	133	120	0	120
Control Point 6	478	76	457	104	120	120	120
Control Point 7	486	454	65	97	0	120	120

Having developed these basic linear least squares algorithms, they were then incorporated into programs (*Pseudo8.pas*, *QRDecomp.pas*, and *SVD4.pas*) which allowed the entire camera calibration/photogrammetric reconstruction process to be performed. In practice, this initially involved reading the calibration data from a file, and then calculating the DLT parameters for both cameras (i.e. solving equation 4.3 twice). Target point data was then read in from a second file, thus allowing three-dimensional reconstructions, with respect to the 'cube's' world coordinate system, to be performed by solving equation (4.9).

Table 4.2 : Typical Reconstructed Control Point Data (DLT).

	Reconstructed Coordinates (mm)			Abs. Error (mm)
	X	Y	Z	
Control Point 0	-0.003	0.253	-0.201	0.323
Control Point 1	120.093	-0.046	0.293	0.311
Control Point 2	120.013	119.935	-0.287	0.295
Control Point 3	-0.095	119.858	0.206	0.268
Control Point 4	-0.047	-0.264	119.864	0.301
Control Point 5	119.939	0.061	120.063	0.106
Control Point 6	119.954	120.088	119.914	0.131
Control Point 7	0.145	120.115	120.144	0.234
			Average =	0.246 mm

In order to obtain an indication of the accuracy limitations of the photogrammetric reconstruction process, the first batch of target point data substituted into these programs was the image coordinates of the control points themselves. When reconstructing the control point locations in this way, it was found that the results

provided by the three linear least squares techniques were always in agreement to at least three decimal places. As indicated by Table 4.2, it was also found that the average absolute reconstruction error was of the order of 0.25 mm.

Table 4.3 : Typical Test Point Data for Software Evaluation.

	Image Test Point Data (pixels)				World Test Point Data (mm)		
	Plan View		Side View		X	Y	Z
	x	y	x	y			
Test Point 1	163	242	291	125	70	20	120
Test Point 2	229	210	319	181	80	40	100
Test Point 3	288	179	349	239	90	60	80
Test Point 4	349	150	381	297	100	80	60
Test Point 5	407	120	414	359	110	100	40
Test Point 6	463	93	448	425	120	120	20

In order to gain a more realistic indication of the measurement errors which would be obtained in practice, test point data sets were then obtained by interactively indicating the six known points provided by the 'steps' calibration object (i.e. non-control points). The results summarised in Table 4.3 represent a typical set of test point data obtained by imaging the 'steps' calibration object in the configuration shown in Figure 4.8. As indicated by the results summarised in Table 4.4, the reconstruction errors associated with non-control points (i.e. points not used during the calibration process) were found to be larger than the corresponding values for control points.

Table 4.4 : Typical Reconstructed Test Point Data (DLT).

	Reconstructed Coordinates (mm)			Abs. Error (mm)
	X	Y	Z	
Test Point 1	69.748	20.094	120.462	0.535
Test Point 2	79.612	40.669	99.986	0.773
Test Point 3	89.651	59.670	79.725	0.553
Test Point 4	99.594	79.965	60.209	0.458
Test Point 5	109.806	99.889	40.256	0.340
Test Point 6	119.611	119.753	19.912	0.469
			Average =	0.521 mm

This trend was also observed during the photogrammetry research reviewed in Chapter 2, and highlights the need to provide an independent means of evaluating

reconstruction accuracy. However, once again, all three linear least squares techniques were in agreement to at least three decimal places.

4.4 Investigation of the Two-Plane Calibration Method.

Having successfully implemented a DLT-based photogrammetric reconstruction scheme, a preliminary investigation of the two-plane camera calibration technique was subsequently undertaken for comparison purposes. Owing to the substantial similarities with the X-ray photogrammetry approach, it soon became clear that the two-plane method was a more promising candidate for further development than the DLT. As a consequence, when a decision was made to undertake a more realistic camera-based simulation of the MEDROSA system's registration strategy, photogrammetric reconstruction was achieved by adapting the two-plane technique to meet the requirements of this particular application. The remainder of this section outlines the methodology adopted during this follow-up study, leading up to the development of a robot-compatible calibration frame and semi-automatic image analysis techniques.

4.4.1 Two-Plane Camera Calibration Mathematics.

As explained in Chapter 2, the two-plane camera calibration method works by deriving independent interpolation functions for two parallel calibration planes. These interpolation functions then provide sufficient information to allow the reconstruction of a line-of-sight (LOS), along which an imaged target point must lie. Assuming that two views of the target point are available, photogrammetric reconstruction is therefore facilitated by the intersection of two such lines-of-sight.

In practice, a wide variety of interpolation functions can obviously be used to establish the required relationship between image and world coordinates. However, the majority of these functions are a variation upon either the global or the local interpolation approach. As the name suggests, global methods attempt to define a single interpolation function which is valid for the entire calibration plane. A 'best fit' solution is therefore obtained by adopting an oversampling/least squares technique.

For the purposes of this thesis, such a technique has been implemented by defining the matrix system:

$$[W] = [A][P] \qquad .....(4.10)$$

Where for the *n* control point case:

- Matrix [W] is a [3 x n] matrix containing world control point data (X_i, Y_i, Z_i).
- Matrix [P] is a [3 x n] matrix containing image (pixel) control point data (x_i, y_i).
- Matrix [A] is a [3 x 3] matrix containing the coefficients which define the required linear interpolation function.

As a consequence, expanding (4.10) yields:

$$\begin{bmatrix} X_1 & X_2 & X_3 & \dots & X_n \\ Y_1 & Y_2 & Y_3 & \dots & Y_n \\ Z_1 & Z_2 & Z_3 & \dots & Z_n \end{bmatrix} = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ y_1 & y_2 & y_3 & \dots & y_n \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \qquad .....(4.11)$$

Given that matrix [A] contains nine unknown coefficients [$A_{i,j}$], and that each control point allows the derivation of three simultaneous equations, a minimum of three control points are therefore required per calibration plane. When oversampling (i.e. n > 3) is practised, estimates for the [$A_{i,j}$] coefficients are obtained by applying the least squares pseudo inverse method to (4.10), thereby yielding:

$$[A] = [W].[P]^T.(P.P^T)^{-1} \qquad .....(4.12)$$

Having performed this camera calibration calculation, the point at which the line-of-sight associated with a target point passes through the calibration plane, can be calculated from:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{bmatrix} \begin{bmatrix} x_{target} \\ y_{target} \\ 1 \end{bmatrix} \qquad .....(4.13)$$

Where (x_{target}, y_{target}) are the measured image (pixel) coordinates of the target point.

By contrast, localised methods define a series of spatial relationships, each of which is only valid for a specific sub-region of the calibration plane. As a consequence, in order to apply the correct interpolation function to an imaged target point, the sub-region in which the point lies must initially be identified. However, the actual interpolation function itself, generally provides an exact rather than a 'best fit' solution. As illustrated by Figure 4.9, for the purposes of this thesis, a localised interpolation scheme has been implemented by using the control point markers to define triangular sub-regions. Applying relationship (4.10) to the image/world coordinate data associated with the three control points at the vertices of a triangle, then allows a localised interpolation matrix $[A_i]$ to be defined for each triangular sub-region.

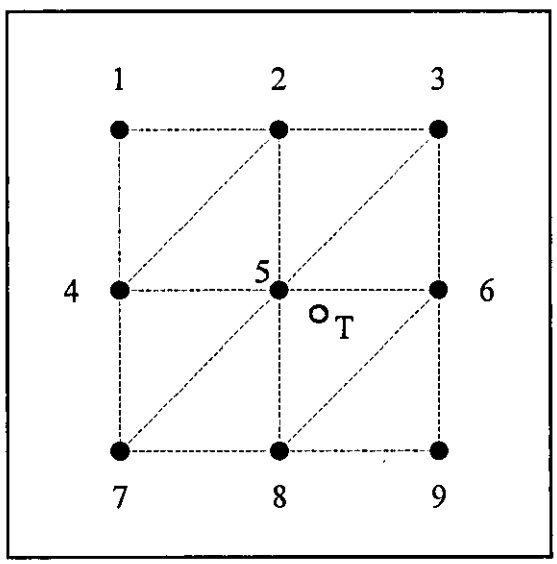


Fig. 4.9: Subdivision of Calibration Plane into Triangular Regions
(Numbered discs represent imaged control point markers)

Previous photogrammetric research tends to suggest that the accuracy of localised interpolation methods improves as the density of the control points is increased. However, increasing the number of control points also increases the number of $[A_i]$ matrices that must be calculated. The storage requirements associated with localised interpolation schemes can therefore be prohibitive. Accordingly, a method which avoids pre-computing the $[A_i]$ matrices has been adopted. Given the image (pixel) coordinates of a target point (T), this method initially identifies the corresponding triangular region, by calculating the three control points that are nearest to point (T).

An exact solution for the required $[A_i]$ matrix is then derived by forming the appropriate $[W_i] = [A_i][P_i]$ system of equations.

For the example supplied by Figure 4.9, this relationship would be:

$$\begin{bmatrix} X_5 & X_6 & X_8 \\ Y_5 & Y_6 & Y_8 \\ Z_5 & Z_6 & Z_8 \end{bmatrix} = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{bmatrix} \begin{bmatrix} x_5 & x_6 & x_8 \\ y_5 & y_6 & y_8 \\ 1 & 1 & 1 \end{bmatrix} \quad .....(4.14)$$

The required $[A_i]$ matrix coefficients are then obtained from:

$$[A_i] = [W_i][P_i]^{-1} \quad .....(4.15)$$

In practice, the inverse matrix $[P_i]^{-1}$ can be calculated by several methods. However, having already used the Crout reduction (refer to Appendix 5) in conjunction with the pseudo inverse method, this familiar technique has once again been used to solve equation (4.15). Having performed this camera calibration process, the intersection of the target point's line-of-sight with the calibration plane is obtained by making the appropriate substitutions into equation (4.13):

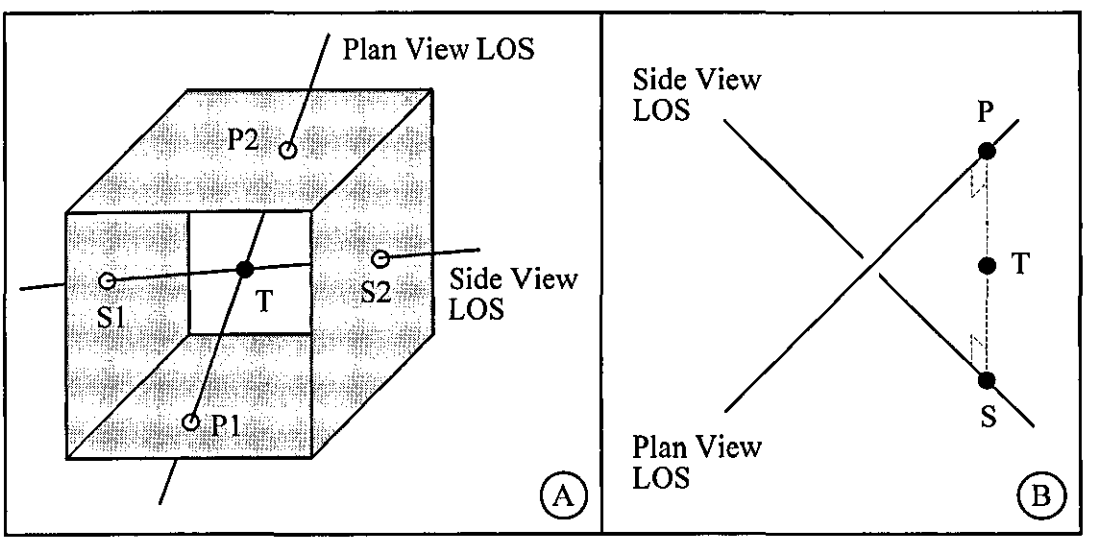


Fig 4.10 : 3-D Reconstruction from the Intersection of Two Lines-of-Sight.

Assuming that either a global or a local interpolation function has been calculated for each calibration plane, the photogrammetric reconstruction process can then be performed by adopting vector analysis techniques. By way of example, such an

approach is outlined here in relation to the biplane imaging geometry shown in Figure 4.10(A). Having calculated interpolation functions for the two plan view (top and bottom) planes, the intersections (P_1 and P_2) of the line-of-sight, passing through a target point (T), with the relevant planes can be calculated from equation (4.13). Similarly, the intersection points (S_1 and S_2) can also be calculated using the appropriate side view interpolation functions.

In vector analysis, a pair of 3-D points provide sufficient information to allow the equation of a 3-D vector (line) to be defined in the form:

$$r = (\text{Point on Line}) + \lambda.(\text{Direction Vector}) \qquad .....(4.16)$$

Where: λ is a constant.

The calculated intersection points P_1, P_2, S_1, and S_2 therefore allow the following line-of-sight vectors to be defined:

$$\text{Plan View LOS} = r_p = \begin{pmatrix} P_{1x} \\ P_{1y} \\ P_{1z} \end{pmatrix} + \lambda_p . \begin{pmatrix} P_{2x} - P_{1x} \\ P_{2y} - P_{1y} \\ P_{2z} - P_{1z} \end{pmatrix} \qquad .....(4.17)$$

$$\text{Side View LOS} = r_s = \begin{pmatrix} S_{1x} \\ S_{1y} \\ S_{1z} \end{pmatrix} + \lambda_s . \begin{pmatrix} S_{2x} - S_{1x} \\ S_{2y} - S_{1y} \\ S_{2z} - S_{1z} \end{pmatrix} \qquad .....(4.18)$$

In practice, owing to measurement errors, it is highly unlikely that these reconstructed lines-of-sight will actually intersect at the target point. As indicated by Figure 4.10(B), the most appropriate means of determining the target point's location is therefore to calculate the points (P) and (S), which correspond to the shortest distance between the two non-intersecting (or skew) lines-of-sight. Having calculated points (P) and (S), the target point location (T) is then estimated by assuming that it lies at the midpoint of the line segment joining (P) and (S). The mathematical technique which is employed to perform these calculations, takes advantage of the fact that at the point of closest approach, the line segment joining (P) and (S) will be perpendicular to both the plan view line-of-sight and the side view line-of-sight.

Hence, defining the line segment $\vec{PS}$ to be:

$$\vec{PS} = r_s - r_p \qquad \text{.....(4.19)}$$

We obtain:

$$\vec{PS} = \begin{bmatrix} (S_{1x} - P_{1x}) + (S_{2x} - S_{1x}).\lambda_s - (P_{2x} - P_{1x}).\lambda_p \\ (S_{1y} - P_{1y}) + (S_{2y} - S_{1y}).\lambda_s - (P_{2y} - P_{1y}).\lambda_p \\ (S_{1z} - P_{1z}) + (S_{2z} - S_{1z}).\lambda_s - (P_{2z} - P_{1z}).\lambda_p \end{bmatrix} \qquad \text{.....(4.20)}$$

Since the line segment $\vec{PS}$ is perpendicular to the direction vector of the plan view line-of-sight we have:

$$\begin{bmatrix} (S_{1x} - P_{1x}) + (S_{2x} - S_{1x}).\lambda_s - (P_{2x} - P_{1x}).\lambda_p \\ (S_{1y} - P_{1y}) + (S_{2y} - S_{1y}).\lambda_s - (P_{2y} - P_{1y}).\lambda_p \\ (S_{1z} - P_{1z}) + (S_{2z} - S_{1z}).\lambda_s - (P_{2z} - P_{1z}).\lambda_p \end{bmatrix} \bullet \begin{bmatrix} (P_{2x} - P_{1x}) \\ (P_{2y} - P_{1y}) \\ (P_{2z} - P_{1z}) \end{bmatrix} = 0 \qquad \text{.....(4.21)}$$

Similarly, since the line segment $\vec{PS}$ is also perpendicular to the direction vector of the side view line-of-sight:

$$\begin{bmatrix} (S_{1x} - P_{1x}) + (S_{2x} - S_{1x}).\lambda_s - (P_{2x} - P_{1x}).\lambda_p \\ (S_{1y} - P_{1y}) + (S_{2y} - S_{1y}).\lambda_s - (P_{2y} - P_{1y}).\lambda_p \\ (S_{1z} - P_{1z}) + (S_{2z} - S_{1z}).\lambda_s - (P_{2z} - P_{1z}).\lambda_p \end{bmatrix} \bullet \begin{bmatrix} (S_{2x} - S_{1x}) \\ (S_{2y} - S_{1y}) \\ (S_{2z} - S_{1z}) \end{bmatrix} = 0 \qquad \text{.....(4.22)}$$

When expanded, equations (4.21) and (4.22) yield a pair of equations of the form:

$$A.\lambda_s + B.\lambda_p = C \qquad \text{.....(4.23)}$$

$$D.\lambda_s + E.\lambda_p = F \qquad \text{.....(4.24)}$$

Solving these equations simultaneously yields the values of λ_s and λ_p at points (S) and (P) respectively. By substituting these λ_s and λ_p values into equations (4.17) and (4.18) the actual three-dimensional coordinates of points (S) and (P) can be calculated. The reconstructed target point (midpoint) is then obtained from:

$$T = \begin{pmatrix} (P_x + S_x)/2 \\ (P_y + S_y)/2 \\ (P_z + S_z)/2 \end{pmatrix} \qquad .....(4.25)$$

Finally, since the length of the line segment $\overrightarrow{PS}$ defines the degree of separation or "mismatch" of the two lines-of-sight, the following calculation can also be performed in order to obtain an indication of calibration accuracy.

$$|PS| = \sqrt{(P_x - S_x)^2 + (P_y - S_y)^2 + (P_z - S_z)^2} \qquad .....(4.26)$$

In general, large $|PS|$ values are interpreted as a sign of poor reconstruction accuracy.

4.4.2 Calibration Frame Design.

A preliminary investigation of the two-plane camera calibration technique was performed with the aid of the calibration 'cube' depicted in Figures 4.5 and 4.6. The eight control points provided by this object allow the following parallel pairs of calibration planes to be defined:

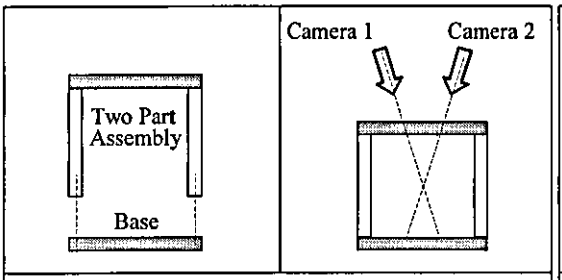
- Plan View: A 'top' plane defined by control points: 4, 5, 6 & 7.
  A 'bottom' plane defined by control points: 0, 1, 2 & 3.
- Side View: A 'front' plane defined by control points: 2, 3, 6 & 7.
  A 'back' plane defined by control points: 0, 1, 4 & 5.

Given that the two-plane method requires a minimum of only three control points per plane, it was therefore possible to develop software algorithms which implement both the global (pseudo inverse) and the localised interpolation schemes outlined in Section 4.4.1. However, in order to follow-up this initial success it became necessary to design a new calibration frame.

In the short term, the main motivation for developing this second calibration frame was a desire to investigate the effect of control point density upon reconstruction accuracy. Since the calibration 'cube' provides only eight control points, its usefulness

in relation to the two-plane method was limited to an initial proof of principle role. Further evaluation of the two-plane technique, and in particular the localised interpolation approach, therefore warranted the design of a calibration frame which provided a densely packed control point distribution. However, the provision of a large number of control points also offers a number of benefits, in relation to the long-term objectives of the MEDROSA applications. The most notable of these benefits is the resulting degree of redundancy in the control point distribution. In other words, by providing more control points than can actually appear within the imaged field-of-view, many of the constraints placed upon the imaging geometry (e.g. perfect alignment of the cameras with the calibration frame) are relaxed, thereby leading to a more flexible calibration process.

An initial evaluation of these imaging requirements led to the design and manufacture of the calibration frames shown schematically in Figures 4.11 and 4.12. With a view to simulating X-ray photogrammetry of the lower thigh/knee region (i.e. the distal interlocking of femoral shaft fractures), both of these frames were designed to encompass a measurement volume consisting of a cube with 180 mm long sides. As a consequence, when imaged from a range of between 750 mm and 1000 mm, the corners of this anatomically compatible measurement volume are outside the field-of-view of the CCD cameras. In order to perform photogrammetric reconstruction, a regular pattern of control points must therefore be distributed throughout the entire field-of-view. In the case of the calibration plane farthest away from the camera, these control points were provided by milling a grid pattern into an aluminium plate. The intersections of these grid lines are then interpreted as control point locations. In order to allow visualisation of this plate, the calibration plane nearest to the camera consists of an aluminium framework supporting a wire mesh. By offsetting these grid/mesh patterns, it was therefore possible to obtain calibration images containing approximately fifty control points per plane.



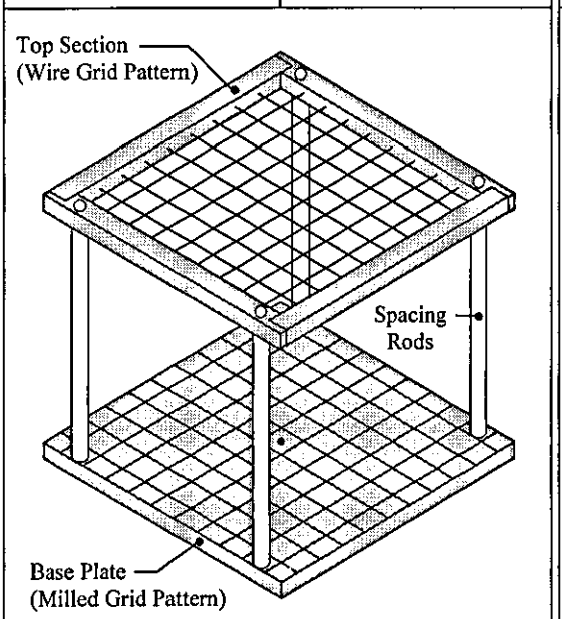
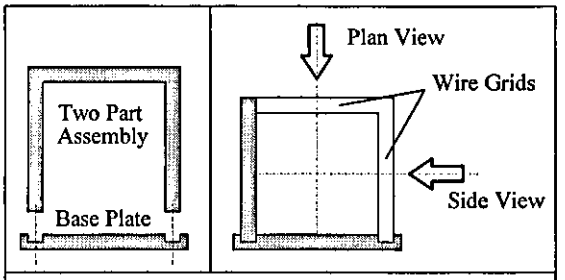


Fig. 4.11 : Convergent Calibration Frame
(198 mm (H) x 210 mm (L) x 210 mm (W))



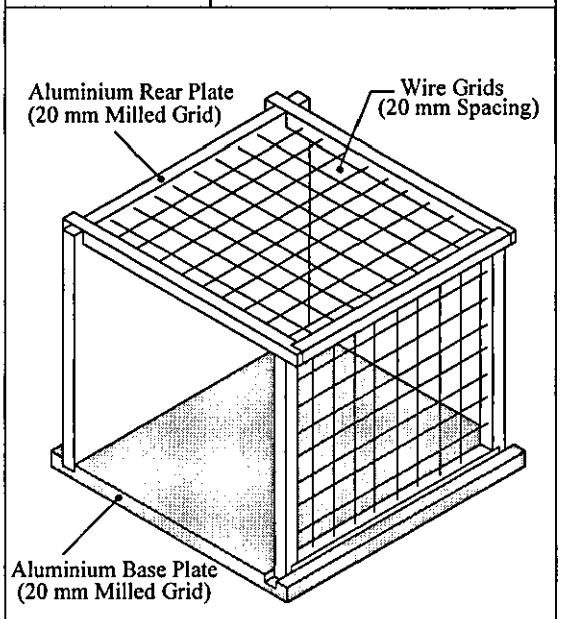


Fig. 4.12 : Biplane Calibration Frame
(195 mm (H) x 222 mm (L) x 228 mm (W))

When the frame depicted in Figure 4.11 was manufactured, convergent imaging geometries were still under investigation as a potential means of quantifying the radiographic examination of the hip joint (refer to Chapter 5 for additional details). Accordingly, this 'convergent' frame consisted of a single pair of calibration planes separated by four spacing rods. Bearing in mind the orthopaedic applications under investigation, this two piece configuration would allow the frame to be assembled around the patient's leg, thus providing a certain degree of anatomical compatibility. In order to acquire the necessary oblique views of the frame, a vertical camera stand was modified to allow a single CCD camera to be positioned at the two convergent viewpoints. By contrast, the frame depicted in Figure 4.12 was designed to support the more conventional biplane imaging geometry, and incorporates the usual four calibration planes (two milled plates and two wire meshes).

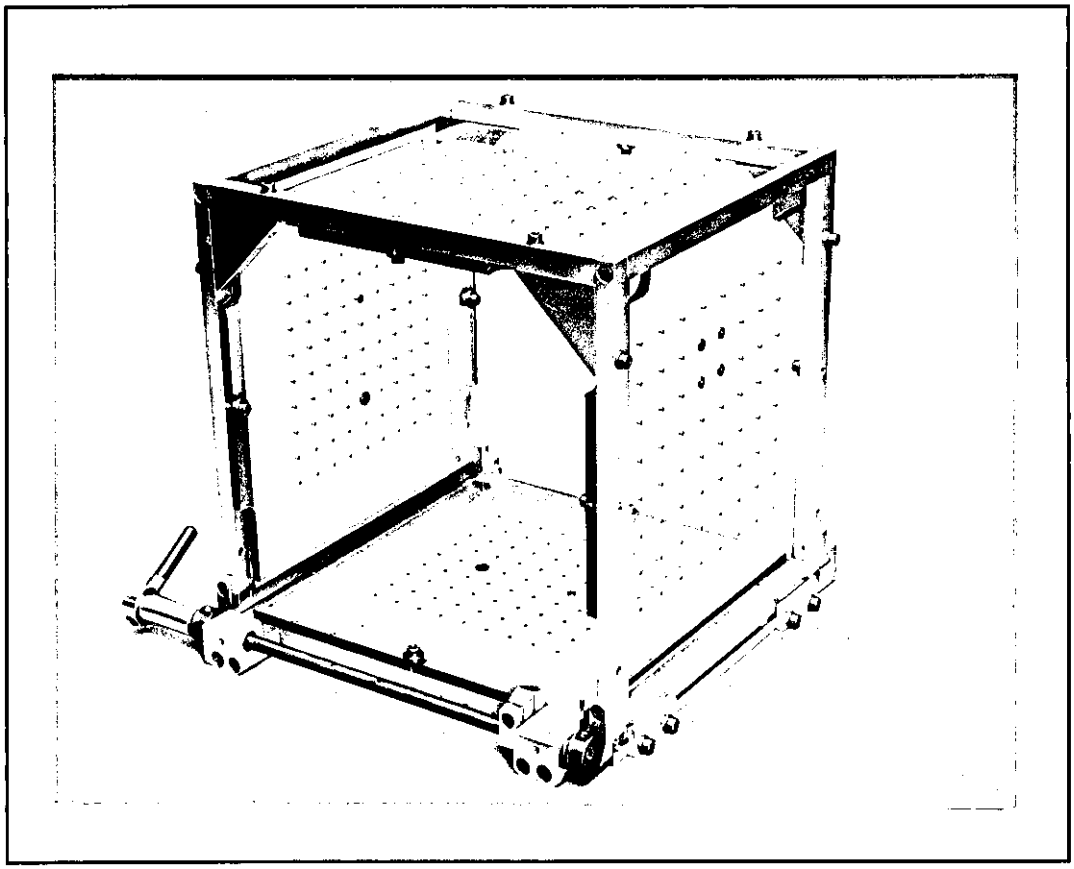


Fig. 4.13 : The Robot-Mounted Calibration Frame

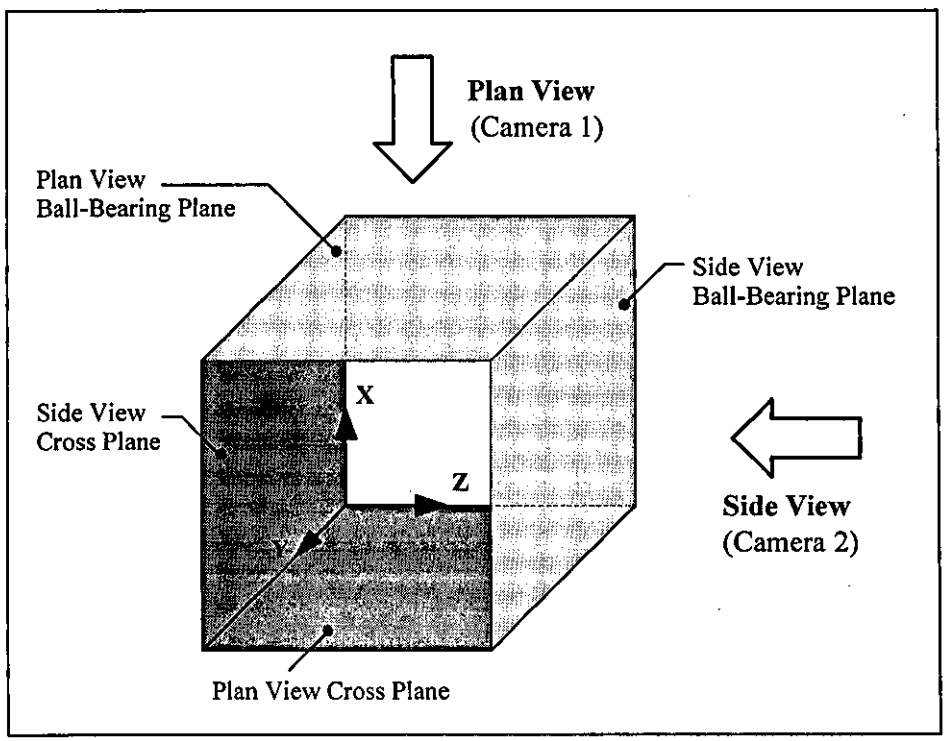


Fig. 4.14 : Schematic of the Robot-Mounted Calibration Frame

Unfortunately, although these 'wire mesh' calibration frames proved to be an effective means of introducing control points into the field-of-view, problems were encountered when attempting to automate the image analysis process. Owing to the large numbers of control points involved, interactive indication of their locations was both time-consuming and tedious. The subjective nature of this measurement process also leads to poor repeatability, as a result of the unavoidable human error. However, owing to the similarities between the two superimposed grid patterns, automatically differentiating between the two calibration planes proved to be impracticable. It was therefore felt that these 'wire mesh' frames made the image analysis process unnecessarily difficult.

Following an intermediate period of software development, the use of the 'wire mesh' frames was therefore abandoned in favour of the robot-compatible calibration frame shown in Figure 4.13. This calibration frame was designed in collaboration with the other members of the MEDROSA research group, with the aim of allowing a full investigation of the registration scheme depicted in Figure 4.1. The base section of the frame can therefore be rigidly attached to the end-effector of the MEDROSA manipulator. A quick-release mechanism then allows the three-sided top section to be attached/detached, thus accommodating the patient's leg. As illustrated by Figure 4.14, this robot-mounted calibration frame supports a biplane imaging geometry, and has been designed to give optimum results at a camera-to-frame (base) range of 750 mm. A slightly larger measurement volume (225.4 mm x 225.4 mm x 220.0 mm) was also provided, with a view to accommodating a wider range of anatomical variations.

Figure 4.13 actually shows the X-ray compatible version of the frame, with metallic fiducial markers embedded into the four Perspex calibration plates. However, these Perspex plates are detachable, allowing camera compatible calibration plates to be inserted in their place. In practice, this was achieved by manufacturing pairs of blank Perspex plates which, when attached to the aluminium calibration frame, could be used to sandwich 'simulated' control point distributions in place. These simulated control point distributions consist of patterns printed on acetate transparencies, in the

case of the calibration plane nearest to the camera, or paper, in the case of the calibration plane farthest from the camera.

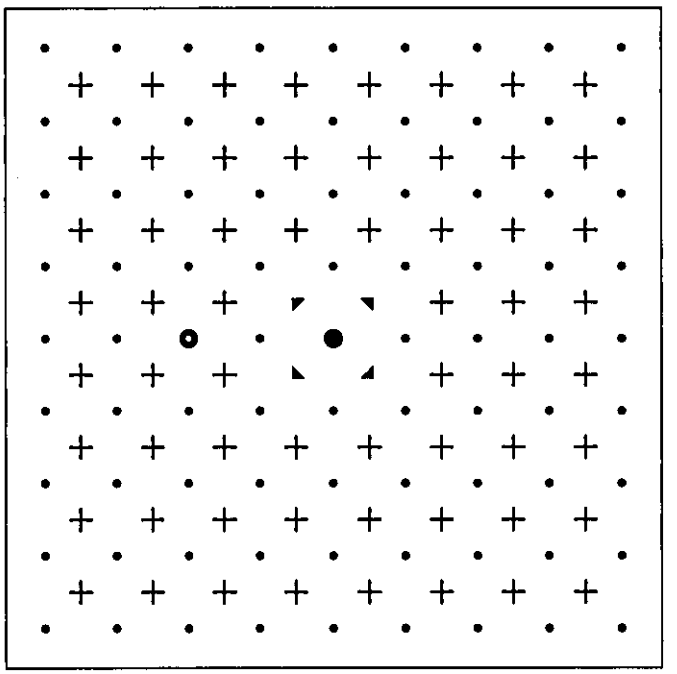


Fig. 4.15 : Ideal Superimposed Fiducial Pattern

Adopting this methodology allowed several control point distributions to be evaluated, with respect to ease of automatic image analysis, without the need to machine several Perspex plates. In the light of the problems encountered with the 'wire mesh' frames, and a preliminary set of X-ray trials (discussed elsewhere in Chapter 5), the following features were eventually incorporated into the chosen control point distributions. The calibration planes closest to the camera consist of a nine-by-nine grid of 2 mm diameter black disks, which simulate the ball-bearings used in the X-ray compatible frames. By contrast the calibration planes farthest from the cameras consist of an eight-by-eight grid of 6 mm wide crosses, which simulate the wire crosses which were initially used during the X-ray trials.

By offsetting these two patterns, and customising the grid spacings (14 mm in the case of the 'ball-bearing' planes, and 20 mm in the case of the 'cross' planes) to take into account the effects of perspective distortion, the image analysis software is therefore ideally required to work on a calibration image of the form shown in Figure 4.15.

However, in practice, the imaged field-of-view does not contain the entire grid patterns.

Given that camera calibration techniques work by establishing image/world point pairs, the image analysis software must be able to uniquely identify or 'label' each of the control points appearing within the field-of-view. The image coordinates of the control points' centroids can then be measured, thus allowing the required point pairs to be established. Owing to the unconstrained nature of the imaging geometries associated with the MEDROSA applications, unique control point identification can only be achieved by introducing a pair of reference or datum markers into the field-of-view. The first of these reference markers indicates the centre of the calibration plane closest to the camera. As indicated by Figure 4.15, this *centre marker* is represented by a black disk (ball-bearing) which has a larger diameter than the other control points in the 'ball-bearing' calibration plane. The in-plane angular orientation of the calibration frame can then be determined by locating the second reference marker. This *orientation marker* is represented by a large hollow disk (metallic washer) in Figure 4.15.

Having located this pair of reference markers, it is possible to estimate the 'expected' locations of the other control points in the 'ball-bearing' plane. Each of these predicted locations can then be automatically scanned in the correct numerical order. Since the out-of-plane angular orientations of the calibration frame are also unconstrained, it cannot be guaranteed that the centres of the 'cross' (far) and 'ball-bearing' (near) calibration planes will appear as coincident points. As a consequence, in order to subsequently scan the 'cross' plane control points in the correct numerical order, reference markers must also be provided to allow the centre of this far plane to be identified. Owing to the fact that the two control point distributions are offset, the most straightforward means of calculating the centre of the 'cross' plane is to place four easily identifiable fiducial markers at the corners of the central grid square. These reference markers are indicated by the four triangles at the centre of Figure 4.15. However, given that the calibration frame is a rigid structure, and that the in-plane orientation has already been established using the 'ball-bearing' control points,

the predicted 'cross' plane control point locations can be calculated without providing an orientation marker in this plane.

Having calculated the expected location of a fiducial marker, this vicinity must then be searched, thus allowing the fiducial marker's actual position to be measured. Ideally, this 'windowing' technique will only locate the fiducial marker of interest. However, owing to the unconstrained imaging geometry and the high control point density, it is inevitable that two or more fiducial markers will occasionally appear within the search region. In order to cope with such an eventuality, evaluation criteria which allow the fiducial marker of interest to be distinguished from the other fiducial markers must therefore be provided. During the camera-based stages of this thesis, these criteria were provided by extracting the following image measurements:

- The perimeter of the fiducial marker.
- The area of the fiducial marker.
- A shape-based 'circularity' function defined by the equation:

$$\text{Circularity} = (4.\pi.\text{Area})/(\text{Perimeter})^2 \qquad .....(4.27)$$

Having adopted these criteria, it was necessary to ensure that the phenomenon displayed in Figure 4.16 was avoided. As indicated by this figure, as a result of the digital sampling process, the measured area of a long narrow fiducial marker can vary considerably with angular orientation. In particular, large fluctuations are observed at the 0°, 90° and 180° cases, which correspond to a perfect alignment of the object with the pixel grid. In order to avoid this type of digitising error, compact (i.e. having a low perimeter to area ratio) isolated fiducial markers should ideally be used **[34]**.

As illustrated by Figure 4.16, circular fiducials offer the best solution to this problem, and also have the added advantage of being invariant to spatial orientation. However, the use of cross-shaped fiducial markers does incur this problem to a certain extent. As a consequence, based upon the experience gained during this camera-based study, cross-shaped fiducials were not used during the later X-ray based studies.

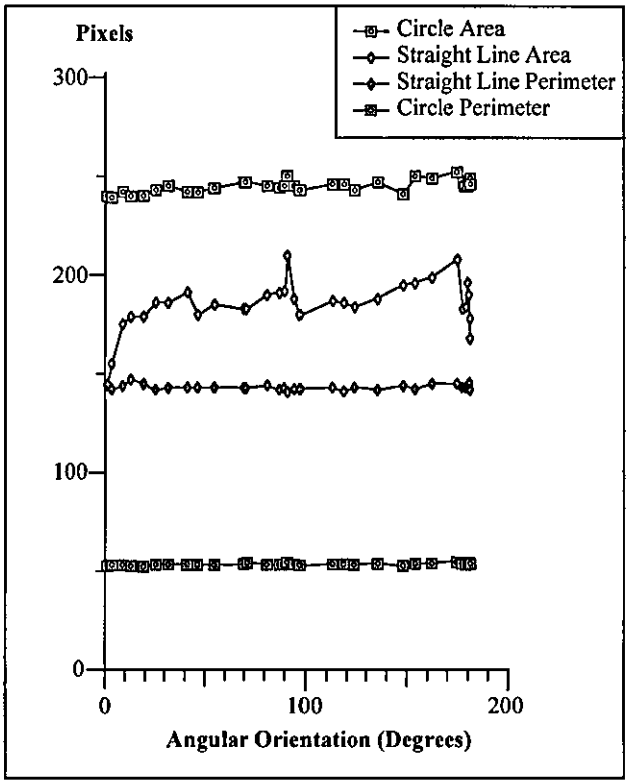


Fig. 4.16 : Effect of Angular Orientation Upon Fiducial Measurements

Nevertheless, as demonstrated by Figure 4.17, by customising the shape and the dimensions of the fiducial markers it was possible to define a unique set of error bands, in relation to the three evaluation criteria, for each type of fiducial marker. Having located an unknown fiducial marker, its identification can therefore be established, thus allowing it to be either accepted as the fiducial of interest or alternatively to be rejected.

4.4.3 Software Development.

The software algorithms described in this section fulfil two main roles: automatic image analysis and photogrammetric reconstruction. In the case of the image analysis algorithms, in order to reduce the duration of the development process, the JLGenial library functions supplied with the IV120(c) frame grabber card have been used whenever possible. As a consequence, the programs adopt the menu-driven user interface supported by these library functions. The digital images acquired by the frame grabber card are also displayed on an independent 'image processing unit' (IPU) monitor, rather than the PC's VDU.

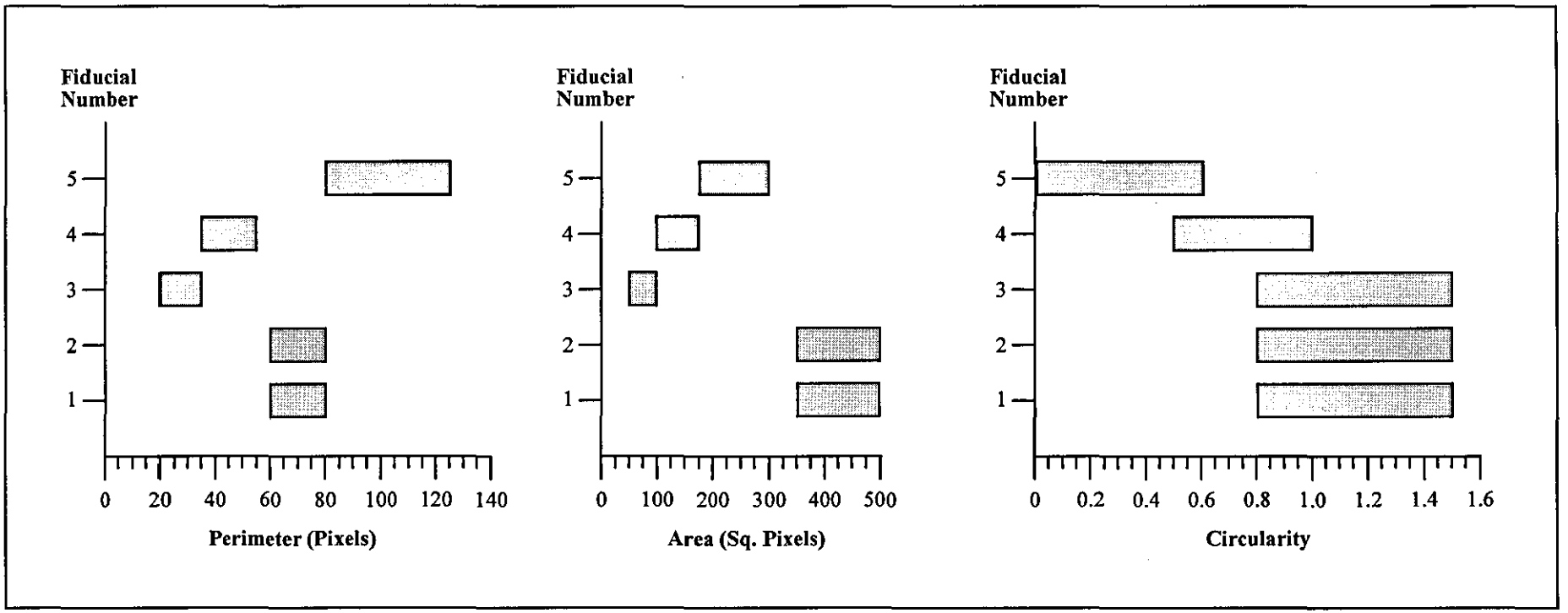


Fig. 4.17 : Bar Charts Showing Distinct Fiducial Measurement Ranges.

Fiducial Number 1 = Ball-bearing plane centre marker (larger ball-bearing).
Fiducial Number 2 = Ball-bearing plane orientation marker (washer).
Fiducial Number 3 = Ball-bearing plane fiducials (smaller ball-bearings).
Fiducial Number 4 = Cross plane centre markers (triangles).
Fiducial Number 5 = Cross plane fiducials (crosses).

4.4.3.1 Preliminary Two-Plane Investigation.

In order to allow direct comparisons with the previously developed DLT algorithms, the same sets of input data were also used to evaluate the first batch of two-plane algorithms. The particular set of results reported in this section therefore corresponds to the control point data supplied in Table 4.1, and the test point data supplied in Table 4.3. As such, it is worth remembering that the image data supplied in these tables was interactively extracted from the relevant calibration images.

With a view to implementing the global interpolation scheme described by equation (4.10), the pseudo inverse algorithms developed for the DLT study were adapted to solve equation (4.12). A program (*twotwo.pas*) was then written to perform this calculation for each of the four calibration planes. The same program then performed the simple matrix multiplication process required to satisfy equation (4.13), before outputting the calculated plane intersection data to a file. A second program (*tworecon.pas*) was then written to read in this intersection data, and perform the three-dimensional reconstruction process described by equations (4.17 to 4.26).

When this pair of programs was used to reconstruct the control point locations themselves, results similar to those supplied in Table 4.5 were obtained. However, while checking these results, it was realised that the use of the calibration 'cube' to implement a simple two-plane scheme, incurred a number of extrapolation cases. As can be appreciated from Figure 4.6, due to the effects of perspective distortion the calibration plane farthest from the camera occupies a smaller region of the image than the calibration plane nearest to the camera. The control points in the near plane therefore lie outside the calibrated region defined by the far plane control points, thereby necessitating an extrapolation-based reconstruction process. Since it is assumed that image distortion effects are minimal when using CCD cameras, the results supplied in Table 4.5 should not be significantly influenced by this potential source of errors. However, when designing the subsequent 'wire mesh' and 'robot-mounted' calibration frames, measures were taken to ensure that the extrapolation case was avoided.

Table 4.5 : Typical Reconstructed Control Point Data (Global Interpolation).

	Reconstructed Coordinates (mm)			Abs. Error (mm)
	X	Y	Z	
Control Point 0	0.235	0.138	0.066	0.280
Control Point 1	119.773	-0.089	-0.108	0.263
Control Point 2	120.278	120.088	-0.211	0.360
Control Point 3	-0.275	119.856	0.254	0.401
Control Point 4	0.247	-0.080	119.923	0.271
Control Point 5	119.744	0.079	120.079	0.279
Control Point 6	120.189	119.911	120.246	0.323
Control Point 7	-0.198	120.081	119.759	0.323
			Average =	0.313

When using the 'steps' calibration object to provide independent test points, it is also possible to incur an extrapolation scenario. In the case of the results summarised in Table 4.6, test points 1 and 6 are subject to such a condition. Nevertheless these results are still comparable to those obtained by the DLT method (refer to Table 4.4). The fifth column of this table supplies the calculated values for the line segment length $|PS|$ described by equation (4.26). Since these values give an indication of the degree of 'mismatch' between the two reconstructed lines-of-sight, their magnitude therefore tends to suggest that the photogrammetric reconstruction process has been performed to an acceptable level of precision.

Table 4.6 : Typical Reconstructed Test Point Data (Global Interpolation).

	Reconstructed Coordinates (mm)			Line Seg. Length (mm)	Abs. Error (mm)
	X	Y	Z		
Test Point 1	69.669	19.874	120.522	0.163	0.631
Test Point 2	79.527	40.245	100.063	0.218	0.537
Test Point 3	89.595	59.188	79.808	0.388	0.927
Test Point 4	99.599	79.555	60.299	0.009	0.670
Test Point 5	109.921	99.673	40.363	0.434	0.494
Test Point 6	119.859	119.851	20.050	0.317	0.211
					Av. = 0.578

In order to evaluate the localised interpolation approach, a program (*local4.pas*) was written to implement the scheme exemplified by equations (4.14) and (4.15). After reading in the appropriate calibration data from a file, this program initially identifies

the relevant triangular image sub-regions by calculating the three closest control points. The corresponding $[A_i]$ interpolation matrices are then obtained by using the Crout reduction to obtain the $[P_i]^{-1}$ matrices. Calculation of the plane intersection points, and the three-dimensional reconstruction process, is then performed using the same approach adopted by the global interpolation programs.

Table 4.7 : Typical Reconstructed Test Point Data (Local Interpolation)

	Reconstructed Coordinates (mm)			Line Seg. Length (mm)	Abs. Error (mm)
	X	Y	Z		
Test Point 1	69.670	19.886	120.506	0.246	0.615
Test Point 2	79.481	40.263	100.068	0.103	0.586
Test Point 3	89.502	59.201	79.838	0.198	0.955
Test Point 4	99.571	79.551	60.252	0.068	0.670
Test Point 5	109.785	99.653	40.387	0.145	0.562
Test Point 6	119.594	119.789	20.186	0.415	0.494
					Av. = 0.647

Given that this localised method provides an exact solution, feeding the image coordinates of the eight control points into the program, as test points, results in a perfect reconstruction of their actual world coordinates. However, as indicated by Table 4.7, when test point locations were reconstructed, the localised two-plane approach was found to be less accurate than both the DLT and the global two-plane methods.

The main reason for this loss of precision is probably the sparsity of the control point distribution provided by the calibration 'cube'. In theory, improved results should be obtained by reducing the grid size from the 120 mm x 120 mm 'grid' established by the 'cube'. Further improvements may also be possible be avoiding the use of extrapolation-based measurements.

4.4.3.2 Robot-Compatible Study.

The outcome of the investigations performed with the 'wire mesh' and 'robot-mounted' calibration frames was the development of a group of software modules which are called by two main programs. The first of these programs (*adapt6.pas*) allows the acquisition and semi-automatic analysis of the required calibration images. The

control point image data is then written in array format to the relevant files. The second program (*twoplane.pas*) then reads in this image data, prior to performing the appropriate camera calibration calculations. Given that the ultimate aim is to define a drilling trajectory which can be implemented by the robotic manipulator, the user is then prompted to indicate the position of a pair of target points in both images. Photogrammetric reconstruction of this pair of points then allows the required drilling trajectory to be defined, with respect to the calibration frame coordinate system, and hence the manipulator coordinate system.

When the image analysis program *adapt6.pas* is run, the user is provided with the following main menu options:

- A)cquire an Image.
- C)alibrate System.

The first of these options allows the user to check the alignment of the two CCD cameras by acquiring a live image of the calibration frame. A graphically overlaid 'gun-sight' type target (i.e. cross-hairs encompassed by a circle) is then provided, to aid alignment of the CCD camera with the centre of the calibration frame. Having verified the alignment of both the plan and side view cameras in this way, the camera calibration process is then instigated by selecting the 'C)alibrate System' menu option. Upon selecting this menu option, the program automatically acquires and displays a live video image from the plan view camera. The user is then prompted to press the left mouse button, thus grabbing a still image of the calibration frame. Ideally, this image will be of the form shown in Figure 4.15.

Having acquired this plan view image of the frame, the first task faced by the image analysis scheme is to locate the centre and orientation fiducial markers of the 'ball-bearing' calibration plane (i.e. the plane nearest to the plan view camera). Although this task could be automated, in practice, it would require the entire image to be scanned. Given that 'global' operations of this type can be very time-consuming, a compromise solution involving interactive indication of these fiducial locations has been adopted. Accordingly, the user is initially prompted to indicate the location of

the centre marker (i.e. the large 'ball-bearing'), by moving a circular mouse-driven cursor until the centre marker appears within its boundaries. The user is then prompted to press the right mouse button, thus verifying that the centre-marker is located within the area bounded by the circular cursor.

In order to measure the actual position of the centre marker, it must initially be segmented from its background, thus allowing its boundary and hence the required centre of gravity to be determined. Since the use of simulated control point distributions, in conjunction with the 'robot-mounted' frame, leads to high contrast calibration images (i.e. 'black' fiducials on a 'white' background), this segmentation process was implemented by applying a simple *adaptive thresholding* scheme. Accordingly, the JLGenial library function *imhisto* is used to scan the search region defined by the circular cursor. The outcome of this scanning process is an array containing the *histogram* data for this region of interest. In other words, the array contains the frequency of occurrence of each grey level in the region of interest.

Figure 4.18 shows the typical bimodal distribution obtained when the histogram data for an image region containing a small 'black' object on a 'white' background is plotted. This figure is actually for a 6-bits per pixel image (i.e. grey scale range: 0 to 63), and as such, grey level zero is assigned to black and grey level sixty-three is assigned to white. The small peak in the vicinity of grey level twenty-one therefore accounts for the pixels which make up the 'black' fiducial marker's image, while the large peak centred at grey level forty-four accounts for the background pixels. As the name suggests, thresholding techniques work be rejecting pixels with a grey level that is either greater or less than a predetermined value (i.e. a threshold). For the black object/white background scenario depicted by Figure 4.18, the aim is therefore to determine a threshold value which corresponds to a grey level in the trough between the two peaks. Pixels with a grey level greater than this threshold value can then be rejected (i.e. allocated as being background pixels), thus allowing the image analysis routines to work on the remaining pixels of interest (i.e. the fiducial marker's image).

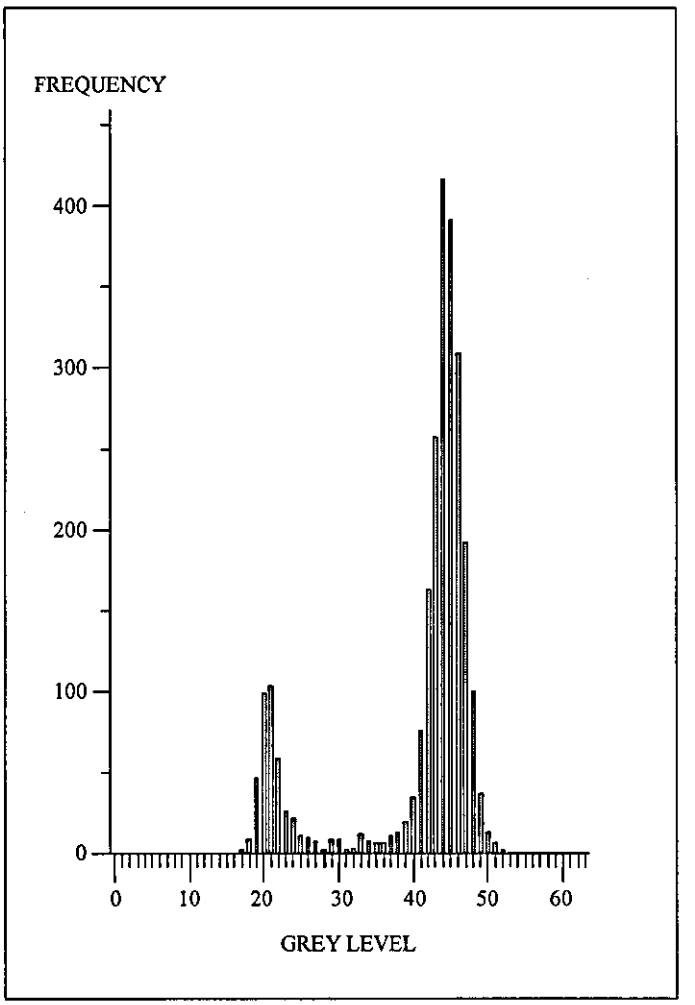


Fig. 4.18 : A Typical Histogram for a 'Black' Fiducial on a 'White' Background.

Consequently, having obtained the histogram data for the centre marker, an algorithm adapted from the software supplied by J-L Automation Ltd. is used to calculate an appropriate threshold value. In practice, this algorithm works by determining the grey level at which the difference between the actual pixel frequency, and the straight line joining the maximum frequency value to the origin of the histogram (i.e. a slope value), is a maximum. The JLGenial library function *slice* is then invoked to apply this threshold value to the region of interest. The output from this library function is a *binary image* in which object points are allocated a value of one, and background objects are allocated a value of zero. By graphically overlaying this binary image on top of the original grey level image, the user can therefore visualise the outcome of the thresholding process.

The 'object' pixels (i.e. pixels with value 1) in this binary image are then analysed using the JLGenial library functions *firstblob* and *nextblob*. These functions scan the region of interest until an 'object' pixel is encountered. The boundary of the object to which this pixel belongs is then traced by scanning its neighbouring pixels in an anticlockwise direction. At the end of the firstblob/nextblob routine, all of the 'object' pixels appearing in the binary image have therefore been allocated to unique 'pointsets' or 'blobs'. As explained in section 4.4.2, the 'blob' of interest can then be identified by applying three empirically determined evaluation criteria (i.e. area, perimeter, and circularity limits). In practice, these values are measured/calculated by invoking the *simple_measure* library function, which also provides the required object centre of gravity measurement.

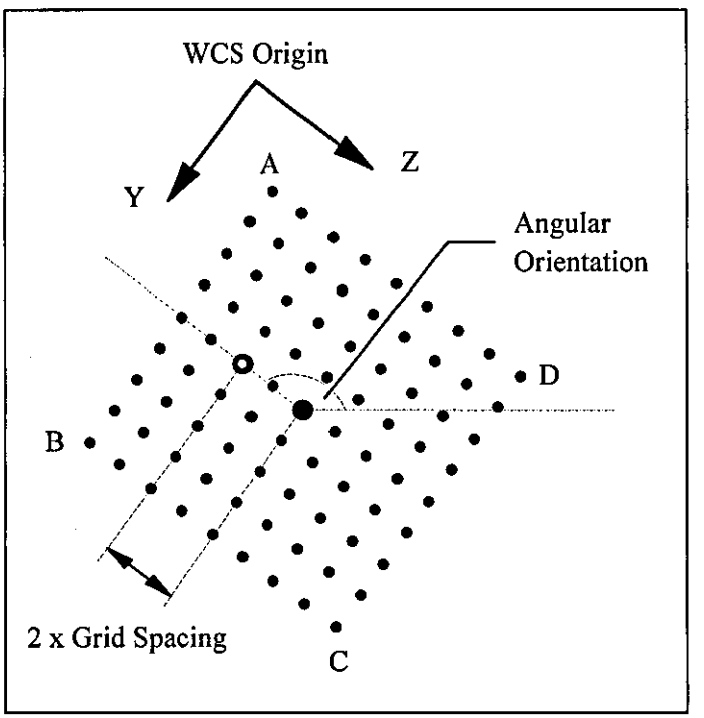


Fig. 4.19 : Calculation of Predicted Fiducial Marker Locations

Assuming that this image analysis routine successfully locates the centre marker, the entire process is then repeated to locate the orientation marker. In the unlikely event that one (or both) of these fiducials is not found, a back-up routine allows interactive indication of the approximate location of the object's centre of gravity. Having located this pair of reference markers, the predicted locations of the other fiducials in the plan view 'ball-bearing' calibration plane are calculated. In order to estimate these

positions, both the grid spacing (in pixels) and the in-plane angular orientation are initially determined from the pixel coordinates of the centre and orientation markers (refer to Figure 4.19). The locations of the fiducial markers at three corners of the grid (i.e. points A, B & D in Figure 4.19) are then calculated, thus allowing the incremental adjustments required to calculate the intermediate fiducial marker locations to be obtained.

The order in which these predicted locations are scanned is obviously important. As such, the numbering system shown in Figure 4.20 has been allocated to the plan view 'ball-bearing' plane fiducials. Adopting this approach allows each of the predicted locations to be individually scanned, with the resulting fiducial image coordinates being written to the corresponding array element. Accordingly, the scanning algorithm starts at fiducial (1, 1) (i.e. point A in Figure 4.19) and works along the top row towards fiducial (1, 9) (i.e. point D in Figure 4.19). The algorithm then scans the fiducials in the second row starting at fiducial (2, 1), and continues along the subsequent rows until fiducial (9, 9) (i.e. point C in Figure 4.19) is reached.

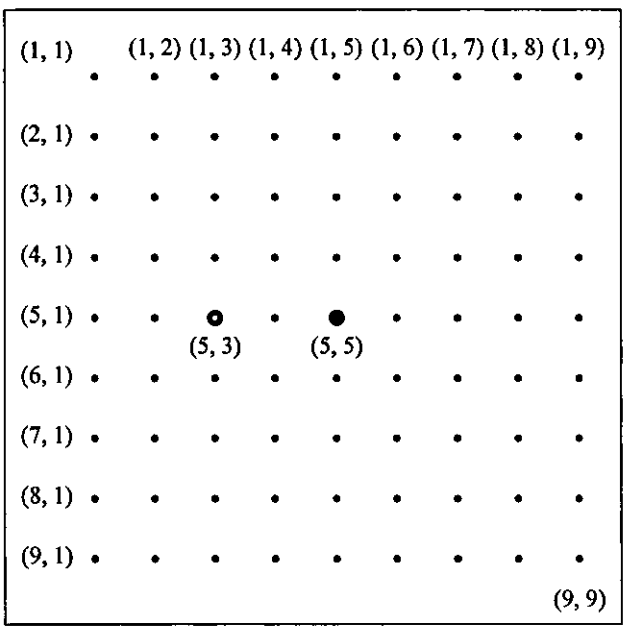


Fig. 4.20 : The Plan View 'Ball-Bearing' Plane Fiducial Numbering System

Given that the fiducials at the periphery of the grid are unlikely to appear within the field-of-view, the search routine initially tests the coordinates of the predicted fiducial locations to see if they are within the limits of the 512 pixel x 512 pixel image. If a

fiducial's predicted location is not within the field-of-view, its image coordinates are written to an array as (0, 0). However, when a fiducial's predicted location is within the required limits, a circular region of interest centred at the predicted point is defined. The adaptive thresholding routine is then applied to automatically determine an appropriate threshold value. The coordinates of the fiducial's centre of gravity can then be determined by supplying the fiducial identification/measurement routine with the appropriate search limits (i.e. area, perimeter & circularity values). Having scanned the entire 'ball-bearing' calibration plane in this manner, the array containing the resulting image coordinate data is written to file *P_BBS_I.*

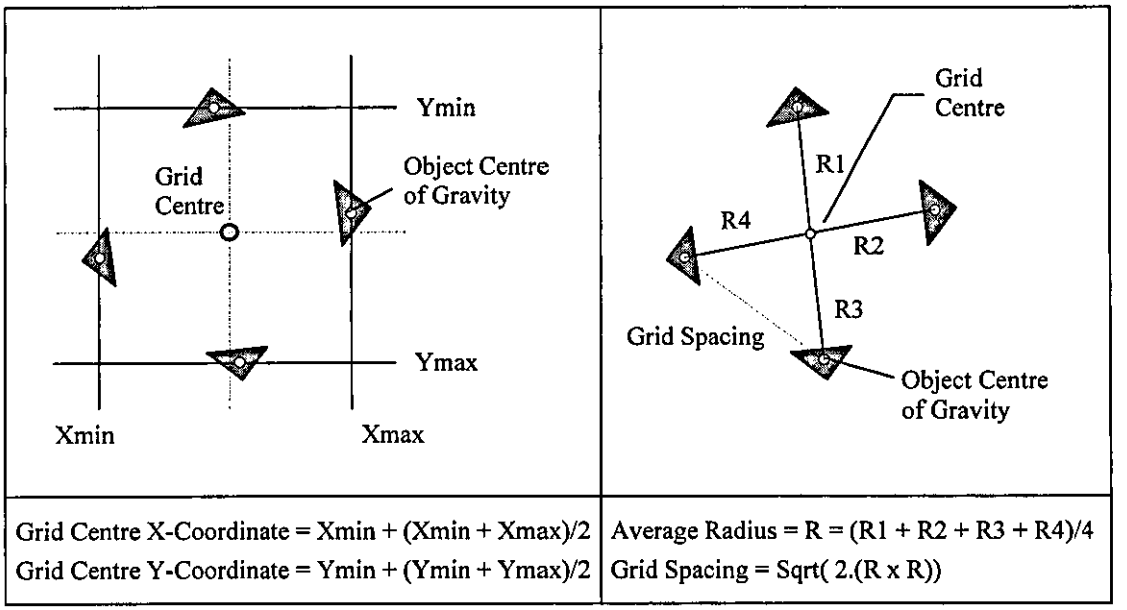


Fig. 4.21 : Calculation of 'Cross' Plane Grid Centre and Grid Spacing

The algorithm then switches attention to the plan view 'cross' plane (i.e. the calibration plane farthest from the plan view camera). Owing to the effects of perspective distortion, the grid spacing (in pixels) associated with this plane is not the same as that of the plan view 'ball-bearing' plane. As previously mentioned, the unconstrained imaging geometry also means the centres of the two plan view calibration planes are not coincident. In order to determine these values, four triangular-shaped fiducials are therefore provided at the centre of the 'cross' plane control point distribution. Accordingly, the program prompts the user to interactively indicate these reference markers, and then calls the adaptive thresholding routine, thus allowing the centre of

gravity of the four imaged triangles to be measured. The required information is then extracted in the manner shown in Figure 4.21.

Since the 'cross' plane grid centre, grid spacing, and in-plane angular orientation are now known, the program is able to calculate and automatically scan the predicted locations of the cross-shaped fiducials. The resulting image coordinate data is then written to file *P_CROS_I*. The program then indicates to the user that analysis of the plan view image is complete. The displayed plan view image is then cleared, and a live image is acquired from the side view camera. The entire image analysis procedure is then repeated, resulting in the appropriate image coordinate data being written to files *S_BBS_I* and *S_CROS_I*.

Having completed the image analysis stage of the camera calibration process, the program *twoplane.pas* can now be run to perform the actual camera calibration and photogrammetric reconstruction calculations. As described in Section 4.4.1, the camera calibration calculations can be performed using either a global or a local interpolation scheme. Accordingly, two versions of the program *twoplane.pas* have been produced. However, in order to avoid confusion, only the global interpolation based version will be described here. This version of the program requires four interpolation matrices (i.e. the [A] matrix of equation 4.10) to be pre-computed, and as such, provides the user with two main menu options:

- C)alibration.
- M)easurement.

Selecting the 'C)alibration' menu option instigates the sequence of calculations required to obtain the four interpolation matrices. In practice, this initially requires the relevant image coordinate data to be read from the files created by *adapt6.pas*. The program then calls the linked software unit *amatcalc.pas*, which calculates the corresponding world coordinates of the control points in the calibration planes. These calculations are obviously customised to the specific dimensions of the robot-mounted calibration frame, and are performed with respect to a frame of reference allocated to one of the corners of the frame (as indicated by Figure 4.14). The *amatcalc.pas* unit

then rearranges the corresponding world and pixel values into the format required by equation (4.11). The required interpolation matrices are then calculated by calling another linked software unit (*pseudo.pas*) which performs the least squares calculations required by equation (4.12). Control points which are outside the field-of-view, or whose fiducial markers have not been detected for one reason or another, are obviously not included within these calculations.

Having calculated the four interpolation matrices, the linked software unit *focal6.pas* is then called. The purpose of this unit is to estimate the location of the side view camera's *effective focal point*, which is required in order to implement the image matching scheme described later in this section. Given that the two-plane model does not impose a unique perspective centre, this effective focal point is obtained by using a linear least squares technique to determine the point in space at which the distance between a bundle of reconstructed lines-of-sight is a minimum.

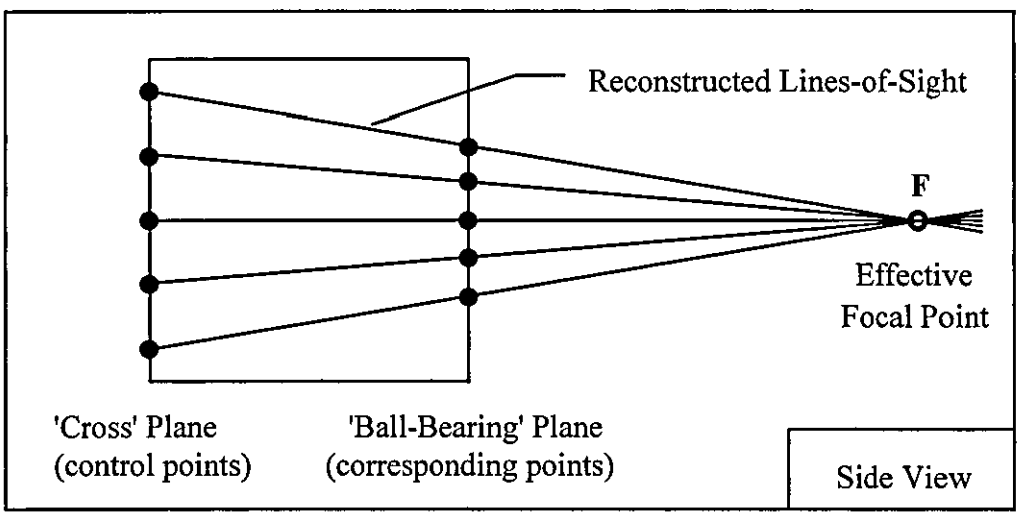


Fig. 4.22 : Estimation of the Side View Camera's Effective Focal Point.

As indicated by Figure 4.22, the input data for such a calculation can be obtained by initially selecting a symmetrical grouping of control points in the side view 'cross' plane. The pixel coordinates of these 'cross' plane control points are then substituted into equation (4.13), along with the interpolation matrix derived from the side view 'ball-bearing' plane calibration data. The world coordinates obtained from this calculation are therefore the points at which the lines-of-sight, which pass through the 'cross' plane control points, intersect the 'ball-bearing' plane.

These corresponding pairs of 'cross' plane control points and 'ball-bearing' plane intersection points, provide sufficient data to define lines-of-sight in the manner described by equation (4.16). The following theory can therefore be used to determine the effective focal point defined by this bundle of lines-of-sight:

With reference to Figure 4.23, the perpendicular distance (**d**) between a point **P**(X, Y, Z) and the line L, which passes through the point $\mathbf{P_1}(X_1, Y_1, Z_1)$ and has direction cosines $[a, b, c]^T$ (where $a^2 + b^2 + c^2 = 1$), can be obtained by applying Pythagoras' Theorem:

$$d^2 \equiv (PN)^2 = (P_1P)^2 - (P_1N)^2 \qquad .....(4.28)$$

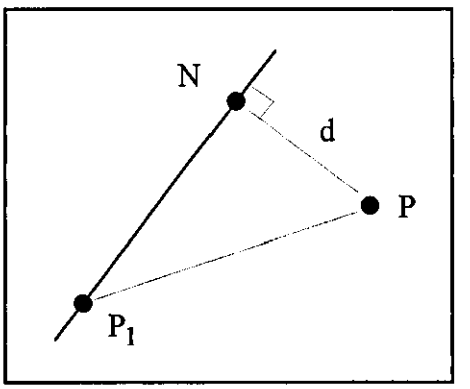


Fig 4.23 : Distance From a Point to a 3-D Line.

Since the P_1N term represents the projection of P_1P onto the line L, it can be shown that **[191]**:

$$P_1N = (X - X_1).a + (Y - Y_1).b + (Z - Z_1).c \qquad .....(4.29)$$

By applying Pythagoras' Theorem we can also obtain:

$$(P_1P)^2 = (X - X_1)^2 + (Y - Y_1)^2 + (Z - Z_1)^2 \qquad ....(4.30)$$

Noting that: $a^2 + b^2 + c^2 = 1$, equation (4.30) may be written as:

$$(P_1P)^2 = [(X - X_1)^2 + (Y - Y_1)^2 + (Z - Z_1)^2](a^2 + b^2 + c^2) \qquad .....(4.31)$$

Substituting for P_1P from (4.31) and P_1N from (4.29) into equation (4.28) yields:

$$d^2 = [(X - X_1)^2 + (Y - Y_1)^2 + (Z - Z_1)^2].(a^2 + b^2 + c^2) - [(X - X_1).a + (Y - Y_1).b + (Z - Z_1).c]^2 \quad .....(4.32)$$

Noting Lagrange's Identity **[191]**:

$$(l^2 + m^2 + n^2)(l_1^2 + m_1^2 + n_1^2) - (l.l_1 + m.m_1 + n.n_1)^2 \equiv (m.n_1 - m_1.n)^2 + (n.l_1 - n_1.l)^2 + (l.m_1 - l_1.m)^2 \quad .....(4.33)$$

Equation (4.32) may be re-written to provide the squared distance (d^2) from point P to line L, in the form:

$$\mathbf{d^2 = [c.(Y - Y_1) - b.(Z - Z_1)]^2 + [a.(Z - Z_1) - c.(X - X_1)]^2 + [b.(X - X_1) - a.(Y - Y_1)]^2} \quad .....(4.34)$$

As outlined by Gremban *et al.* **[26]**, this equation can be utilised to determine the required location of the effective focal point in the following manner. Expanding the terms of equation (4.34) yields:

$$d^2 = X^2.(b^2 + c^2) + Y^2.(a^2 + c^2) + Z^2.(a^2 + b^2) - 2.X.Y.a.b - 2.X.Z.a.c - 2.Y.Z.b.c + 2.X.(b.k_3 - c.k_2) + 2.Y.(c.k_1 - a.k_3) + 2.Z.(a.k_2 - b.k_1) + k_1^2 + k_2^2 + k_3^2 \quad .....(4.35)$$

Where:

- $k_1 = Z_1.b - Y_1.c$
- $k_2 = X_1.c - Z_1.a$
- $k_3 = Y_1.a - X_1.b$

Since $D = \sum d^2$ is a minima at the effective focal point, differentiating D with respect to X, Y & Z yields:

$$\begin{aligned}
\partial D/\partial X &= \sum 2.X.(b^2 + c^2) - \sum 2.Y.a.b - \sum 2.Z.a.c + \sum 2.(b.k_3 - c.k_2) \\
\partial D/\partial Y &= \sum 2.Y.(a^2 + c^2) - \sum 2.X.a.b - \sum 2.Z.b.c + \sum 2.(c.k_1 - a.k_3) \\
\partial D/\partial Z &= \sum 2.Z.(a^2 + b^2) - \sum 2.X.a.c - \sum 2.Y.b.c + \sum 2.(a.k_2 - b.k_1)
\end{aligned} \quad .....(4.36)$$

Where:

- The sums are taken over all of the line-of-sight vectors.
- The (a, b, c, k_1, k_2, & k_3) terms are functions of these line-of-sight vectors.

Setting these derivatives equal to zero in order to find the minima, and putting the equations into the matrix form:

$$h = [A].f \qquad .....(4.37)$$

yields:

$$\begin{bmatrix} \Sigma(c.k_2 - b.k_3) \\ \Sigma(a.k_3 - c.k_1) \\ \Sigma(b.k_1 - a.k_2) \end{bmatrix} = \begin{bmatrix} \Sigma(b^2 + c^2) & -\Sigma a.b & -\Sigma a.c \\ -\Sigma a.b & \Sigma(a^2 + c^2) & -\Sigma b.c \\ -\Sigma a.c & -\Sigma b.c & \Sigma(a^2 + b^2) \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad .....(4.38)$$

Hence the effective focal length (f) may be obtained from:

$$\mathbf{f = [A]^{-1}.h} \qquad .....(4.39)$$

Having generated the relevant line-of-sight data, the unit *focal6.pas* therefore initialises the [A] matrix and h vector terms, before calculating the inverse matrix $[A]^{-1}$ and then solving equation (4.39).

The calculation of the side view camera's effective focal point constitutes the end of the camera calibration process. The three-sided top section of the robot-mounted calibration frame can therefore be detached from the base section, and removed from the field-of-view. The photogrammetric reconstruction process is then instigated by selecting the 'M)easurement' menu option. Upon selecting this option, a group of algorithms in the linked software unit *desired2.pas* are called in order to perform the photogrammetric reconstruction process. Accordingly, a live image is acquired from the plan view camera, thus allowing a test object to be placed on the calibration frame's base unit (i.e. within the previously calibrated volume). When a satisfactory placement of the test object has been achieved, a still image of the frame base/test object combination is acquired.

In line with the MEDROSA project's aim of keeping the orthopaedic surgeon within the decision making loop, the program then prompts the user to interactively indicate a pair of target points in this plan view image. The pixel coordinates of each of these target points are then substituted into equation (4.13), along with the appropriate interpolation matrix, thus allowing the plan view plane intersection points (i.e. points P_1 and P_2 in Figure 4.10(A)) to be calculated. This pair of three-dimensional points then allows a plan view line-of-sight to be derived in the form of equation (4.16). In this way, the plan view line-of-sight corresponding to each of the target points is reconstructed.

In order to complete the photogrammetric reconstruction process, the locations of the two target points must also be indicated in the side view image of the test object (i.e. the *correspondence problem* must be solved). With a view to minimising the human error associated with this subjective indication process, an image matching technique, which is similar to the epipolar line-based method described in Section 2.2.2, has been implemented. The aim of this technique is to project the reconstructed plan view lines-of-sight into the side view image as graphical overlays. Since it is known that these lines-of-sight pass through the actual target points, the location of the corresponding points in the side view image must therefore lie on these projected lines. As a result, the projected lines can be used to guide the user's selection of the target point locations in the side view image.

With reference to Figure 4.24, this image matching scheme has been implemented as follows. Having previously estimated the location of the side view camera's effective focal point (F), lines-of-sight can be defined which pass through the relevant (P_1 & F) and (P_2 & F) combinations. The intersections of these lines-of-sight with the side view 'ball-bearing' plane (i.e. points SB_1 and SB_2 in Figure 4.24) can therefore be calculated, by utilising the fact that the Z-coordinate of this plane is known (i.e. Z = 225.4 mm everywhere within the plane) with respect to the calibration frame's coordinate system. The projection of these real world points into the side view image can then be performed by manipulating equation (4.10) to obtain:

$$[P] = [A]^{-1}.[W] \qquad .....(4.40)$$

Inverting the previously calculated side view 'ball-bearing' plane global interpolation matrix, therefore allows the pixel coordinates of points SB_1 and SB_2 to be calculated. A graphically overlaid straight line which passes through these calculated pixel locations can then be drawn into the side view image.

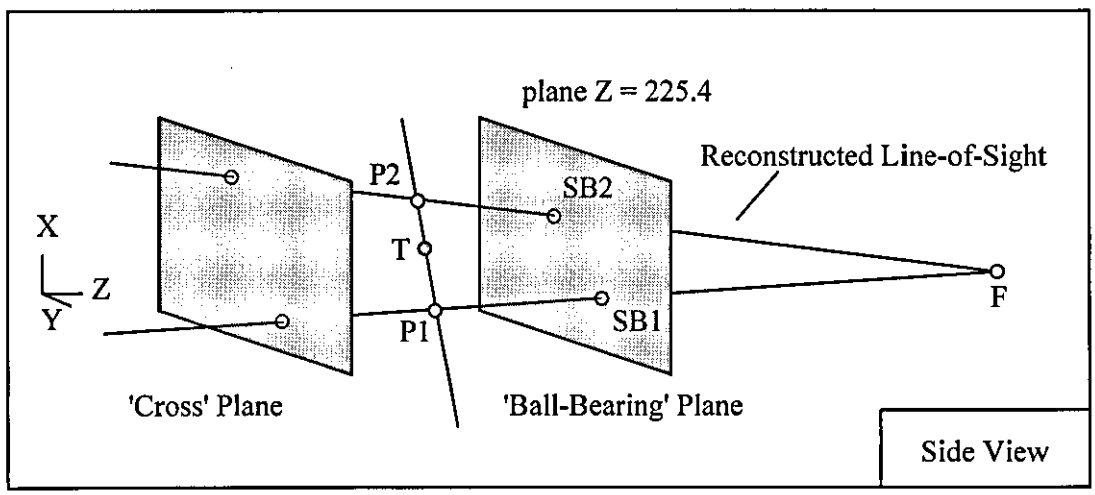


Fig. 4.24 : Overlaying the Reconstructed Plan View Line-of-Sight in the Side View Image.

The location of the two target points in the side view image are then interactively indicated with the aid of these graphically overlaid 'matching' lines. The pixel coordinates of each of these target points are then substituted into equation (4.13), along with the appropriate interpolation matrix, thus allowing the side view plane intersection points (i.e. points S_1 and S_2 in Figure 4.10(A)) to be calculated. The appropriate S_1 and S_2 pairs are then used to calculate the side view lines-of-sight which pass through the two target points. The three-dimensional reconstruction process described by equations (4.17 to 4.26) is then implemented, to yield the desired coordinates of the two target points with respect to the calibration frame's coordinate system.

In order to "debug" the software modules described above, test data was generated by imaging the steps calibration object with the robot-mounted calibration frame. Given that the base of the calibration frame is outside the field-of-view of the side view

camera, this configuration provided four test points per calibration image (i.e. only the upper four test points of the steps were visible in both views). As indicated by Table 4.8, repeated trials involving this calibration hardware, and the final versions of the global interpolation-based software modules, suggested reconstruction errors in the range of 1-2 mm. The localised interpolation-based software modules also produced reconstruction errors of this magnitude.

Table 4.8 : Typical Reconstructed Test Point Data (Global Interpolation).

	Reconstructed Coordinates (mm)			Line Seg. Length (mm)	Abs. Error (mm)
	X	Y	Z		
Test Point 1	121.85	119.62	83.91	0.61	1.51
Test Point 2	103.89	109.74	104.43	0.73	1.70
Test Point 3	82.28	99.02	122.37	0.53	1.37
Test Point 4	63.67	90.43	143.79	0.48	1.12
					Av. = 1.43

However, since the steps calibration object was not specifically designed to be used in conjunction with the robot-mounted calibration frame, positioning the object in relation to the frame is a highly subjective process. As such, the "nominal" locations of the test points, and hence the reconstruction errors, are only approximate values. Fortunately, a more reliable measure of the precision of the calibration software is provided by the "mismatch" (line segment length) values supplied in Table 4.8. Reassuringly, the sub-millimetre magnitude of this "accuracy indicator", suggests that the photogrammetric reconstruction process has indeed been successfully performed.

4.5 Summary of Chapter 4.

With a view to simulating the X-ray based intraoperative registration strategy of the MEDROSA system, the initial stages of the research work documented in this thesis have involved a CCD camera-based feasibility study. Although it is acknowledged that this simulation approach has a number of limitations, it has provided a low cost development route which avoids the potential health hazards associated with X-ray radiation. By exploiting the similarities between the optical and X-ray image formation process, it has therefore been possible to develop software modules that require only a minimal amount of modification in order to become X-ray 'compatible'.

Having implemented photogrammetric reconstruction schemes based upon both the Direct Linear Transformation and the two-plane camera calibration technique, a decision was made to pursue the two-plane option. Accordingly, two versions of this calibration technique, one based upon global interpolation and the other based upon localised interpolation, have been integrated into the following software package:

Program 1: *Adapt6.pas*

- *Description*: Image acquisition and semi-automatic image analysis routines.
- *Inputs*: Plan and side view images of the calibration frame.
- *Outputs*: The image (pixel) coordinates of the fiducial markers are written to file.

Program 2: *Twoplane.pas*

- *Description*: Camera calibration and photogrammetric reconstruction routines.
- *Inputs*: The image (pixel) coordinates of the fiducial markers are read from file.
- *Outputs*: The three-dimensional coordinates of reconstructed target points, with respect to the calibration frame's coordinate system.

Support Units:

- *Nptypes.pas* (global declarations common to all units/programs)
- *W_file.pas* (global declarations for file transfer routines)
- *Automate.pas* (simple image analysis/utility functions)
- *Amatcalc.pas* (calculation of global interpolation functions)
- *Pseudo.pas* (pseudo inverse linear least squares calculations)
- *Focal6.pas* (estimation of the effective focal point of the side view camera).
- *Desired2.pas* (image matching and photogrammetric reconstruction routines).

When used in conjunction with a custom-built robot-mounted calibration frame, this software has allowed laboratory trials of a camera-based version of the MEDROSA system to be performed. As described in Chapter 5, the success of these 'system integration' trials prompted a period of X-ray based research, which was facilitated by the acquisition of a mobile C-arm fluoroscopy unit.

5. X-RAY BASED FEASIBILITY STUDY.

5.1 Introduction.

In order for the MEDROSA Project to progress beyond the feasibility study stage, a point was eventually reached where unrestricted access to a mobile C-arm unit was essential. Accordingly, a 'second-hand' Siemens Siremobil model 2N was purchased from Stafford District General Hospital in 1996. At the time of acquisition, this particular C-arm unit was approximately fifteen years old, and was therefore deemed unsuitable for further diagnostic usage. However, in relation to the objectives of the MEDROSA Project, a C-arm model of this age was ideal for system development purposes, as it represented the worst case scenario which would be encountered in routine clinical use. In other words, developing calibration techniques which are able to compensate for the image distortion produced by older (Vidicon camera based) C-arm units, rather than state-of-the-art (CCD camera based) digital C-arm units, guarantees the versatility of the MEDROSA system.

Having established the compatibility of the Siremobil 2N with the IV120(c) frame grabber in a radiology suite at Stafford District General Hospital, and performing an initial set of checks following its delivery to Loughborough, the C-arm unit was then disabled to prevent it being used until an X-ray enclosure could be constructed. In compliance with the campus radiation protection standards, the lead-lined (2 mm thick sheeting) partition walls of this enclosure define a controlled area (2-3 metre radius) around the X-ray source. The single access door to the enclosure is also lead-lined, while a lead glass viewing window provides a safe vantage point. The approved system of work associated with this installation, states that the C-arm unit should be operated from outside the X-ray enclosure. Accordingly, the door to the enclosure, the mains electricity supply, and the C-arm foot pedal have all been incorporated into an interlocking/warning light system. Members of the MEDROSA group wishing to use the C-arm unit are also required to attend a training course for workers with ionising radiation, and to be registered as a radiation worker.

The remaining sections of this chapter outline the early stages of the X-ray based research conducted as part of the current thesis. As a consequence, Section 5.2 describes a preliminary feasibility study which was performed before the C-arm unit was purchased. This study gave an early indication of the practical problems that would eventually have to be overcome. Section 5.3 then covers the period of research immediately after the commissioning of the X-ray enclosure, which led to the development of X-ray compatible versions of the camera calibration techniques discussed in Chapter 4. Having developed this initial set of "working" software modules, a second period of research was then undertaken with the aim of improving both the reliability and accuracy of the system. As described in Section 5.4, this investigation was primarily concerned with the image analysis and image distortion problems. Finally, Section 5.5 summarises the findings of the chapter in relation to the follow-up studies discussed in Chapters 6 and 7.

5.2 Initial Demonstration of Feasibility.

In parallel to the camera-based research described in Chapter 4, an X-ray based feasibility study was also undertaken in a radiology suite at Glenfield General Hospital, Leicester. The primary aim of this investigation was to demonstrate the compatibility of the IV120(c) frame grabber card with existing mobile C-arm units. In other words, could 'digital' X-ray images be acquired by establishing a direct video link between the frame grabber and a C-arm unit. Accordingly, the relevant video plug adapter (10.94.572.B1913) was purchased from Siemens (Medical Engineering Group), thus allowing a standard (75Ω) RS-170 video lead to be wired-up.

In order to take full advantage of the radiographer's time and the facilities provided by Glenfield Hospital, two prototype X-ray calibration frames were manufactured. The first of these frames is depicted schematically in Figure 5.1, and will henceforth be referred to as the prototype *biplane* X-ray calibration frame. This biplane frame was designed to allow simulation of the distal interlocking stage of intramedullary nailing procedures. As a consequence, the dimensions of the frame were dictated by human anatomy, in accordance with the scenario depicted in Figure 5.2.

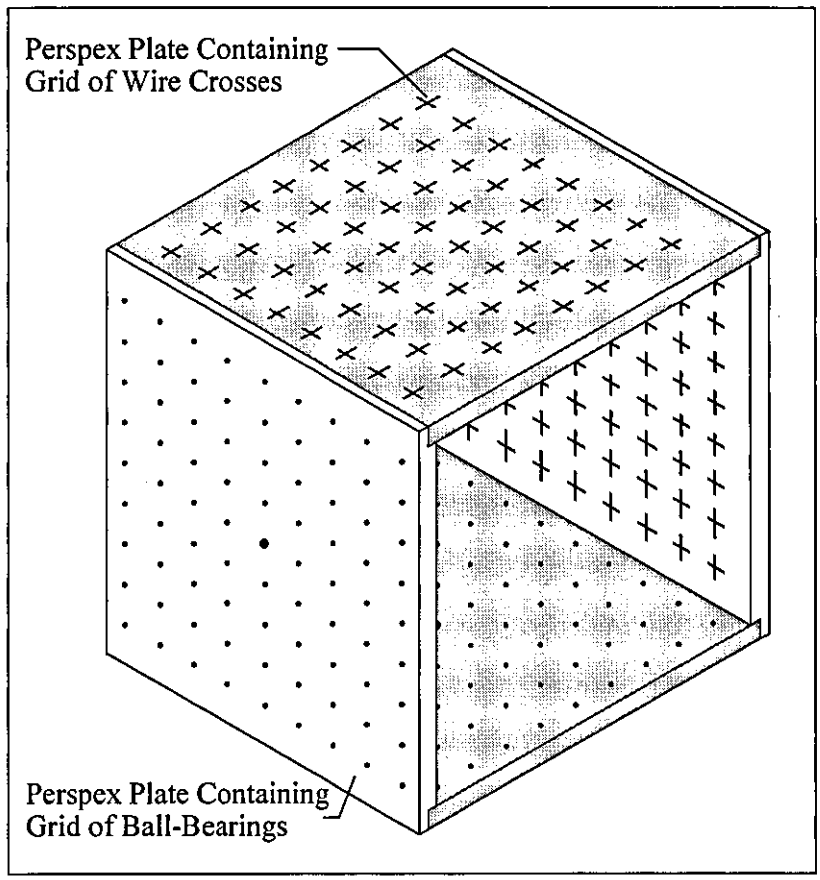


Fig. 5.1 : Prototype Biplane X-Ray Calibration Frame for Distal Interlocking.
(Height = 200 mm, Length = 200 mm, Width = 180 mm)

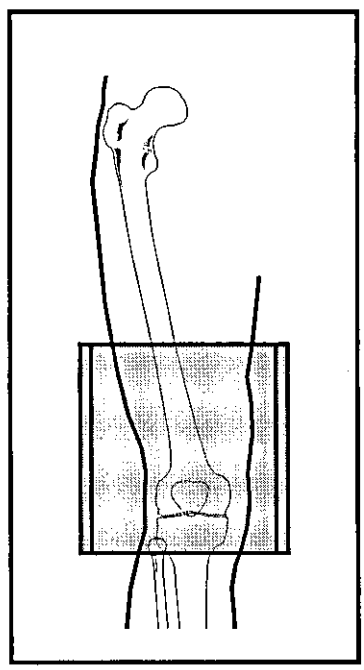


Fig. 5.2: Frame Positioning For Biplane Fluoroscopic Examination.

By contrast, the second calibration frame (depicted schematically in Figure 5.3) was designed with a view to simulating the internal fixation of hip fractures. In keeping with the findings of the RSA research discussed in Chapter 2, this frame supports a convergent imaging geometry which is more conducive to acquiring calibrated images of the hip/pelvis region. The dimensions of this prototype *convergent* X-ray calibration frame therefore comply with the scenario depicted in Figure 5.4, thus allowing images of either the left or right hip to be acquired.

Both of these prototype calibration frames were manufactured from 10 mm thick Perspex plates, into which radiopaque fiducial patterns were embedded. In the case of the calibration plate(s) closest to the X-ray source, this fiducial pattern consisted of a wire mesh, which was facilitated by milling a grid into the Perspex plate and then inlaying 0.75 mm diameter solder wire (chosen for its high lead content). The intersections of these wires could then be interpreted as control points. In order to be able to differentiate between the parallel pairs of calibration plates in the resulting X-ray images, a grid of 2 mm diameter ball-bearings was then embedded into the calibration plate farthest from the X-ray source (i.e. the image intensifier side of the frame). A 5 mm diameter ball-bearing was also introduced to provide a centre marker or reference point.

As indicated by Figure 5.5, the preliminary trials conducted at Glenfield successfully demonstrated both the desired frame grabber/C-arm compatibility, and the introduction of Perspex-mounted fiducial markers into 'digital' X-ray images. Figure 5.5 is actually a posterioanterior (PA) fluoroscopic image of a human (cadaver) femur inside the biplane X-ray calibration frame. Paying particular attention to the wire grid at the periphery of this image, the characteristic effects of pincushion distortion are readily noticeable. Given that both the wire and ball-bearing patterns have a grid spacing of 20 mm, the variable magnification caused by central projection can also be appreciated, with the wire grid (closest to the X-ray source) being subjected to more magnification than the ball-bearing grid. However, it must also be pointed out that the quality of an *in vivo* image of the biplane frame/lower thigh combination would be further degraded by the presence of the soft tissues surrounding the femur.

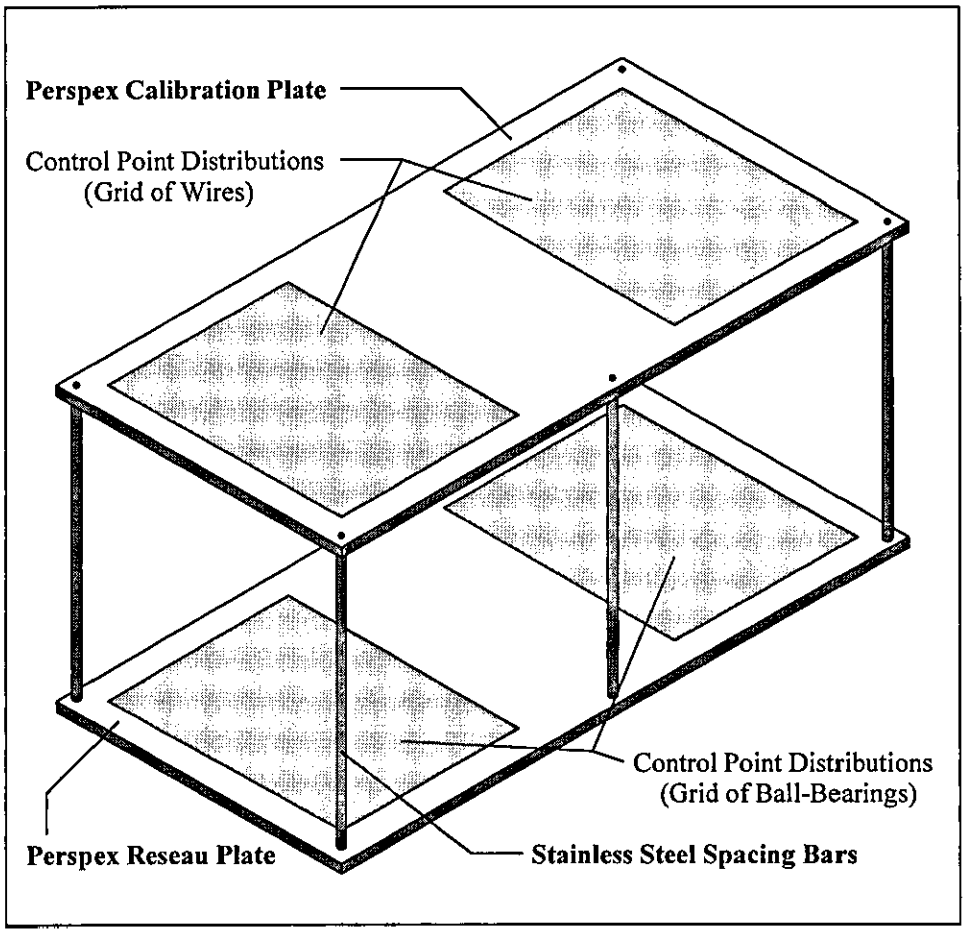


Fig. 5.3 : Prototype Convergent X-ray Calibration Frame for Hip Fractures.
(Height = 320 mm, Length = 650 mm, Width = 320 mm)

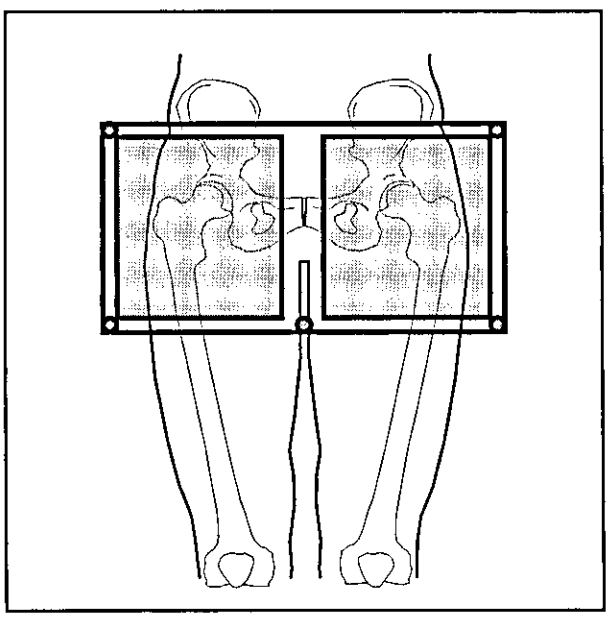


Fig. 5.4 : Frame Positioning For Convergent Fluoroscopic Examination:
The control point distributions are located over the left & right hip joints.

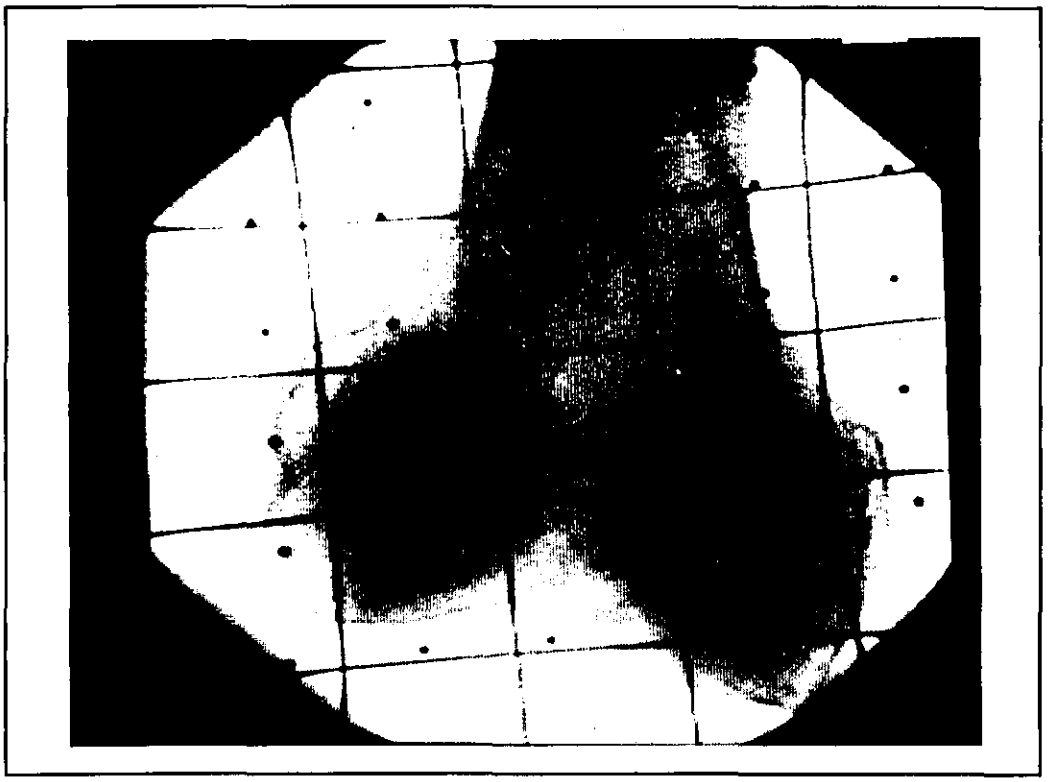


Fig. 5.5 : Typical X-ray Image of the Biplane Frame/Femur Combination.

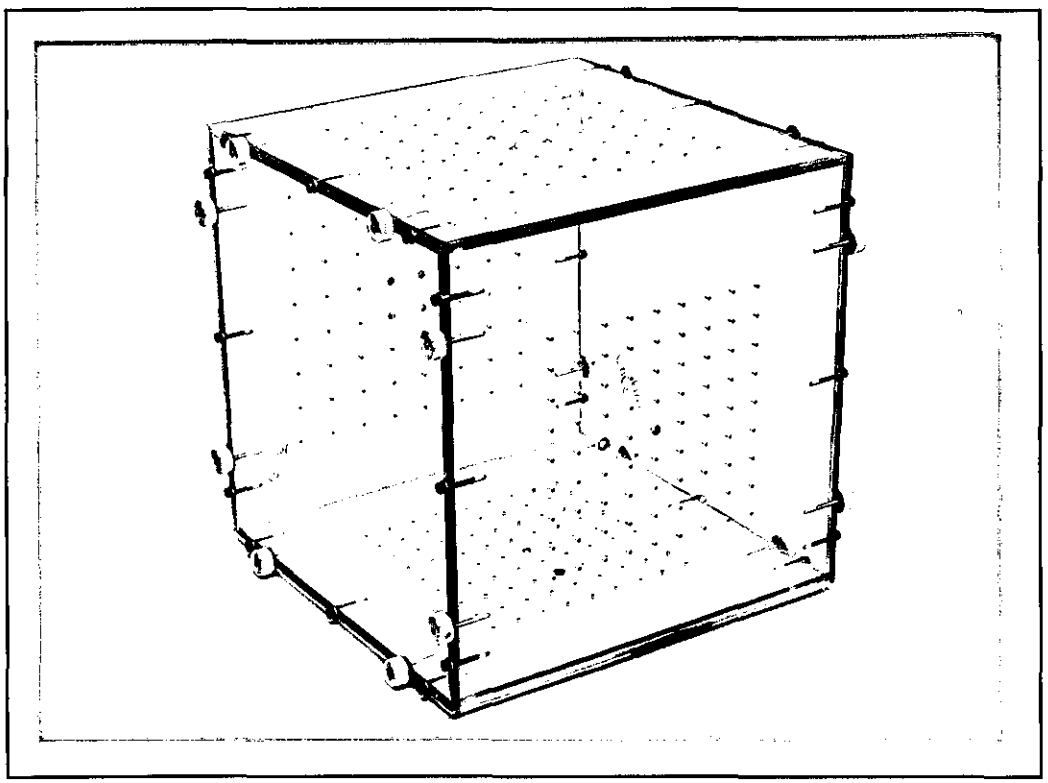


Fig. 5.8 : The "Hip Box" Biplane X-Ray Calibration Frame.
(Enclosed measurement volume = 225 mm x 225 mm x 225 mm).

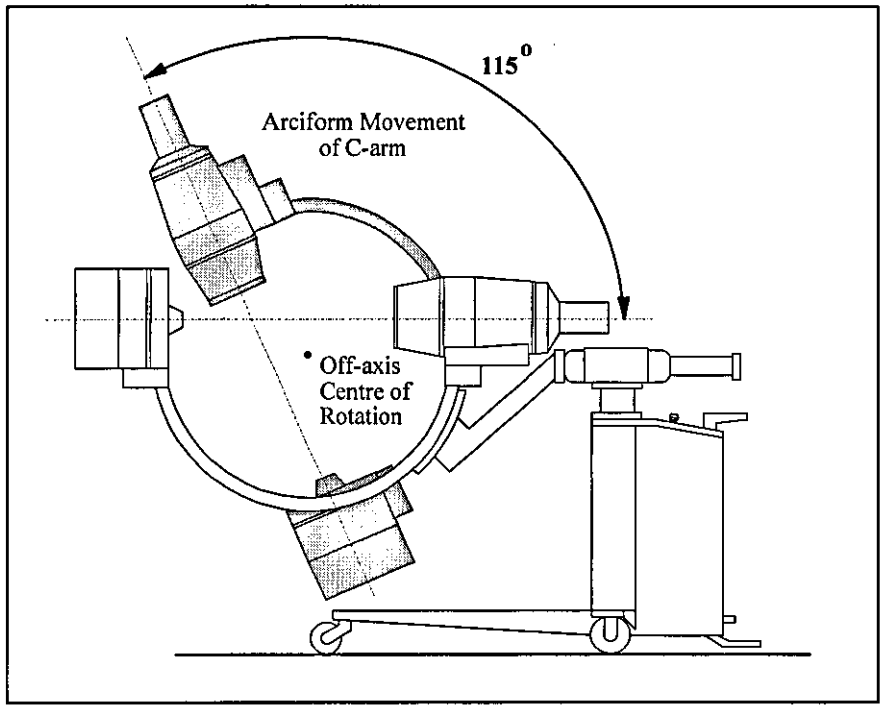


Fig. 5.6 : Mechanical Constraints Upon Arciform C-arm Movement.

Having successfully simulated the standard PA and lateral views of the biplane frame/cadaver femur combination, attention was then switched to the prototype convergent frame. However, when attempting to simulate the scenario depicted in Figure 5.4, it soon became clear that in relation to the requirements of convergent X-ray photogrammetry, the physical configuration of the mobile C-arm unit was far from ideal. As illustrated by Figure 5.6, the rotation of a typical C-arm's arc is mechanically constrained to 115°, thus precluding the acquisition of oblique views with an angle of incidence greater than 25°. Significantly, the X-ray source-to-image plane distance (SID) of a mobile C-arm unit is fixed (refer to Figure 5.7), and is typically of the order of 900 mm. Given that the height of the prototype convergent calibration frame is 320 mm, the base of the calibration frame is at most only 480 mm away from the source. The corresponding distance for the majority of the RSA studies covered in Chapter 2 is typically at least twice this value.

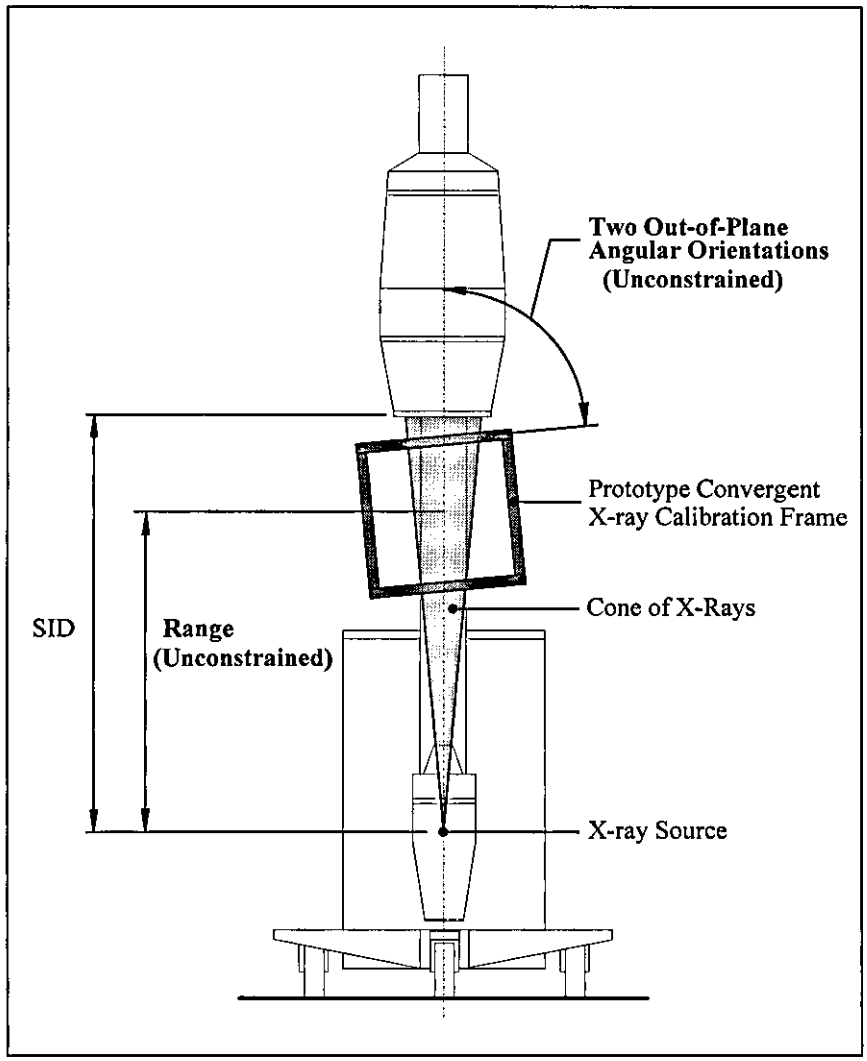


Fig. 5.7: The Unconstrained C-arm Imaging Geometry.

The implications of this relatively small SID value, combined with a large calibration plate separation, can be appreciated in relation to Figure 5.7. As depicted by this figure, unlike conventional RSA applications, the position and angular orientation of the calibration frame, with respect to the X-ray source, are both totally unconstrained when using a mobile C-arm unit. However, an ideal calibration image (such as the one depicted in Figure 4.15) will only be obtained when the calibration frame is perfectly aligned with the 'optical' axis of the C-arm unit. Since this perfect alignment is very hard to achieve in practice, angular misalignment can result in the wire and ball-bearing fiducial patterns being superimposed (i.e. overlapping) in the resulting X-ray image. Given that this eventuality severely complicates the semi-automatic image analysis task, it must be avoided at all costs. As a consequence, the

fiducial patterns must be designed in such a way that a certain amount of angular misalignment can be tolerated (i.e. offset grid patterns, customised grid spacings, etc.).

Unfortunately, as the separation of the calibration plates increases, the amount of angular misalignment required to produce an 'overlapping' calibration image decreases. When using the prototype convergent X-ray calibration frame in conjunction with a mobile C-arm unit, it was therefore found that excessive fluoroscopy times (and hence irradiation of the patient) were required in order to obtain a satisfactory calibration image.

In particular, alignment of the C-arm with the calibration frame was complicated by the fact that the centre of rotation of the arc is not on the 'optical' axis of the X-ray imaging system (refer to Figure 5.6). As such, the alignment process involves a combination of rotations and translations. Owing to the proximity of the wire calibration plate to the X-ray source, a 'penumbra' or blurring effect was also noticeable in the resulting image. On account of the large magnification factor, a much smaller grid spacing is also required in order to ensure that enough control points are introduced into the field-of-view. In view of these practical issues, and the fact that oblique views are not routinely used during orthopaedic surgery procedures, it was therefore decided that an alternative to convergent X-ray photogrammetry of the hip should be sought (refer to Chapter 7 for further details).

5.3 Preliminary Software Development.

Upon completion of the X-ray enclosure, it became necessary to transform the camera calibration software (discussed in Chapter 4) into an "X-ray compatible" form. However, in view of the lessons learnt during the preliminary trials at Glenfield Hospital, and a series of tests performed on calibration frame "mock ups" in the laboratory, it was decided that the image analysis process could be further simplified through the provision of a new X-ray calibration frame. Accordingly, the Perspex biplane X-ray calibration frame depicted in Figure 5.8 was manufactured to aid the software development process. Although this cube design is obviously incompatible

with the anatomy of the hip, the fiducial patterns embedded into the Perspex calibration plates were designed with a view to biplane imaging of the hip. As such, the frame depicted in Figure 5.8 will henceforth be referred to as the "Hip Box".

Having made the transition from a camera-based imaging system to an X-ray based one, it is convenient to adopt the terminology commonly used in radiology. Accordingly, the terms *plan view* and *side view*, as used in the camera calibration context in the previous chapter, will henceforth be replaced by *posterioanterior (PA) view* and *lateral view*. Similarly, the use of the terms *front* and *back plate* to indicate the calibration plates nearest and farthest from the camera are also superseded by *source plate*, which is used here to imply the calibration plate nearest to the X-ray source, and *XRII plate*, the calibration plate nearest to the X-ray image intensifier. Applying this terminology, the main improvements offered by the Hip Box over the prototype biplane X-ray calibration frame (depicted in Figure 5.1) are as follows:

- Fiducial Marker Patterns:

  Given the problems previously encountered with inlaid wire mesh/cross patterns, ball-bearings (2 mm diameter) were used to provide fiducial markers in all four calibration plates of the Hip Box. Owing to the increased magnification experienced by the source plate when the Hip Box is imaged, the ball-bearings embedded in this plate can be easily distinguished from those in the XRII plate. The centre of the two (PA and lateral) source grids is marked by a larger ball-bearing (5 mm diameter), while an orientation marker is provided by a large washer (6 mm outside diameter). Four smaller washers (4 mm outside diameter) at the centre of the grid pattern, as opposed to the previously used triangles or crosses, allow the grid measurements (grid spacing and grid centre) of the two XRII plates to be determined.

- Customised Grid Spacings:

  The ball-bearings embedded in the two source plates of the Hip Box define a nine-by-nine grid pattern, with a grid spacing of fifteen millimetres. By contrast, the ball-bearings in the two XRII plates define an offset eight-by-eight grid pattern

with a twenty millimetre grid spacing. When imaged at a range of 300 mm (i.e. with the source plate 300 mm from the X-ray image intensifier), owing to the variable magnification associated with central projection, the Hip Box should therefore ideally superimpose a calibration pattern of the form shown in Figure 5.9 into the standard diagnostic images.

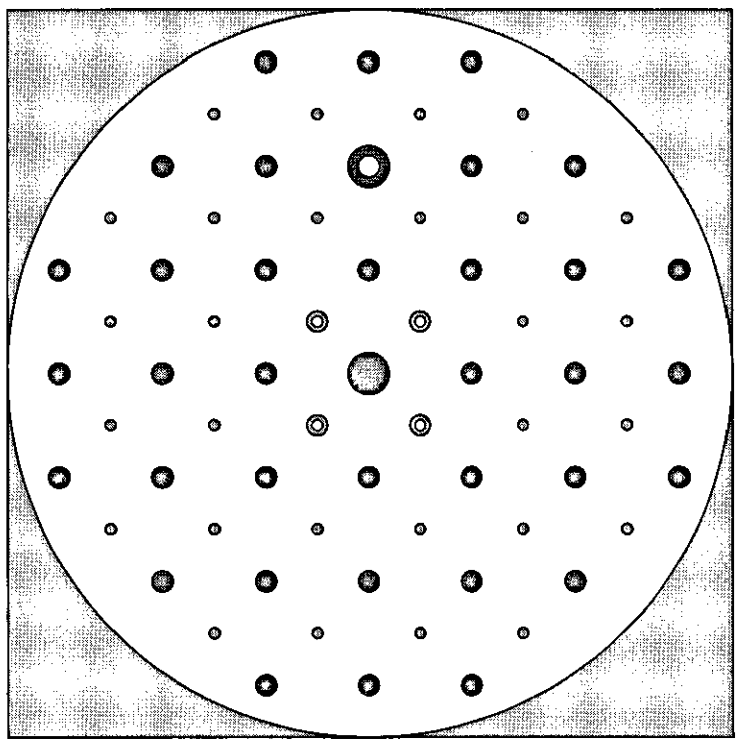


Fig. 5.9 : Ideal Hip Box X-Ray Calibration Image.

When imaged using the Siemens Siremobil 2N C-arm unit, the Hip Box produces a high contrast fluoroscopic image consisting of "black" fiducial markers on a "white" background. As a result, it was found that the adaptive thresholding approach to image segmentation, used during the camera calibration study, was equally applicable to X-ray imaging. An X-ray compatible image analysis program (*HB_Adapt.pas*) was therefore produced by modifying the camera-based program *Adapt6.pas* to take into account the new fiducial patterns of the Hip Box. However, since the X-ray imaging geometry is unconstrained (as depicted in Figure 5.7), it was found that in order to set appropriate limits for the magnification dependent fiducial identification criteria (i.e. area and perimeter), an automatic method of determining range must be provided.

In order to provide range independent image analysis software, an empirical technique involving an off-line calibration process was adopted. This calibration process required the Hip Box to be positioned at a known distance (range) from the input window of the X-ray image intensifier, prior to the acquisition of a fluoroscopic image. The process was then repeated several times, with the range being increased by a 20 mm increment on each occasion. Each of the resulting images was then analysed to extract both the area and the perimeter of the various fiducial marker types. This data was then plotted to yield graphs of the type shown in Figure 5.10, and continuous measurement/range functions were derived by performing polynomial curve fittings.

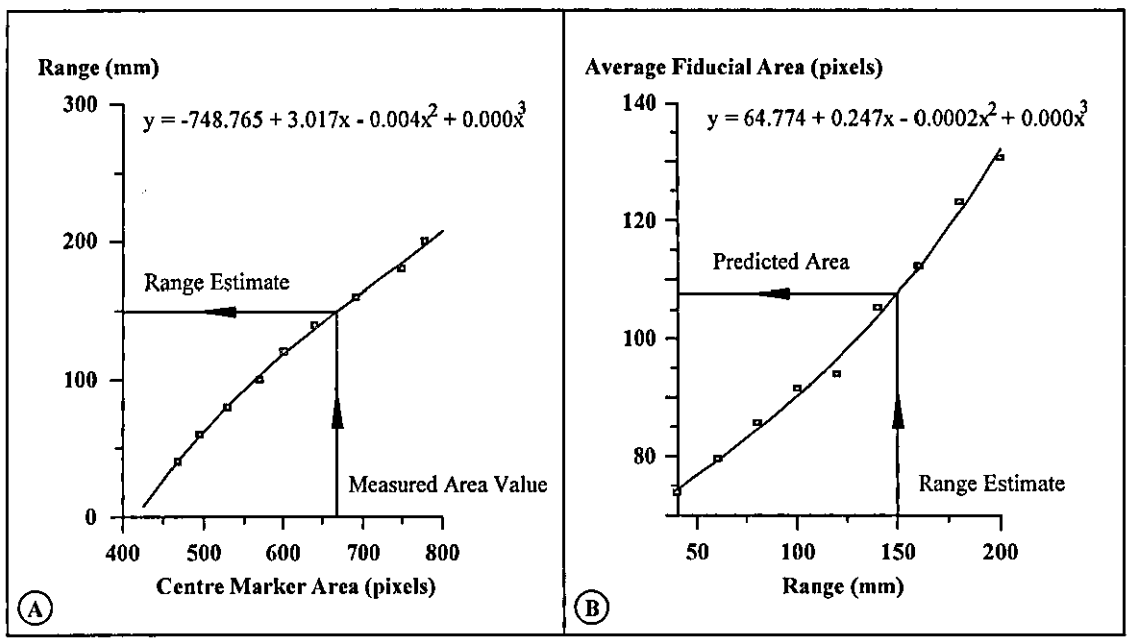


Fig. 5.10 : Automatic Determination of Range and Search Limit Values.

Once the system is calibrated, the image analysis scheme implemented by *HB_Adpat.pas* can proceed as follows. The user is prompted to indicate the location of the centre (largest ball-bearing) and orientation (largest washer) markers of the source plate. Both the perimeter and the area of these imaged fiducials are then determined using the adaptive thresholding and measurement algorithms. Assuming that these values are within the acceptable (range independent) limits, a range estimate is then determined by using the appropriate continuous measurement/range functions

in the manner shown in Figure 5.10(a). In practice, the average of the four values derived from the centre and orientation marker data is subsequently taken as the range estimate.

Having used the centre and orientation marker measurements to determine a range estimate, the search limits for the remaining fiducials are determined in the manner shown in Figure 5.10(b). By substituting the calculated range estimate into the appropriate continuous function, a value for the average predicted measurement (i.e. area or perimeter) is obtained. Upper and lower search limits are then calculated by adding/subtracting an "error band" or "safety margin" to this predicted average value. By adopting this range estimation approach, it was found that *HB_Adapt.pas* could successfully accommodate the unconstrained nature of the fluoroscopic imaging geometry. The measured pixel locations of the uniquely identified fiducial markers could therefore be written to user specified files, thus providing the input data for the photogrammetric reconstruction process.

By way of a first attempt at implementing this photogrammetric reconstruction process, the global interpolation version of the camera calibration program *Twoplane.pas* was adapted to the geometry of the Hip Box. Owing to the similarities between the X-ray and optical image formation processes, the resulting program, *HB_Prog.pas,* and its linked units (*HB_Amat.pas*, *HBPseudo.pas*, *HB_Focal.pas*, *HB_Surg.pas* and *HB_Types.pas*) required only minor modifications in order to achieve X-ray compatibility.

At this stage of system development, a means of introducing independent (non-control point) fiducial markers, into the measurement volume defined by the Hip Box, was required in order to establish the accuracy of the reconstruction process. In practice, this requirement was met by manufacturing the Perspex "test object" shown schematically in Figure 5.11. This narrow strip of Perspex fits diagonally into the Hip Box and is screwed to the blank side plates of its cubic structure. When imaged along with the Hip Box, the four ball bearings embedded into the test object therefore provide reference points whose locations are known with respect to the Hip Box's

arbitrarily defined coordinate system. As indicated by Figure 5.11 (lateral view), by modifying the Hip Box it was possible to attach the test object in three different locations ("upper", "middle" and "lower"), thus allowing a thorough investigation of the photogrammetric reconstruction process to be performed.

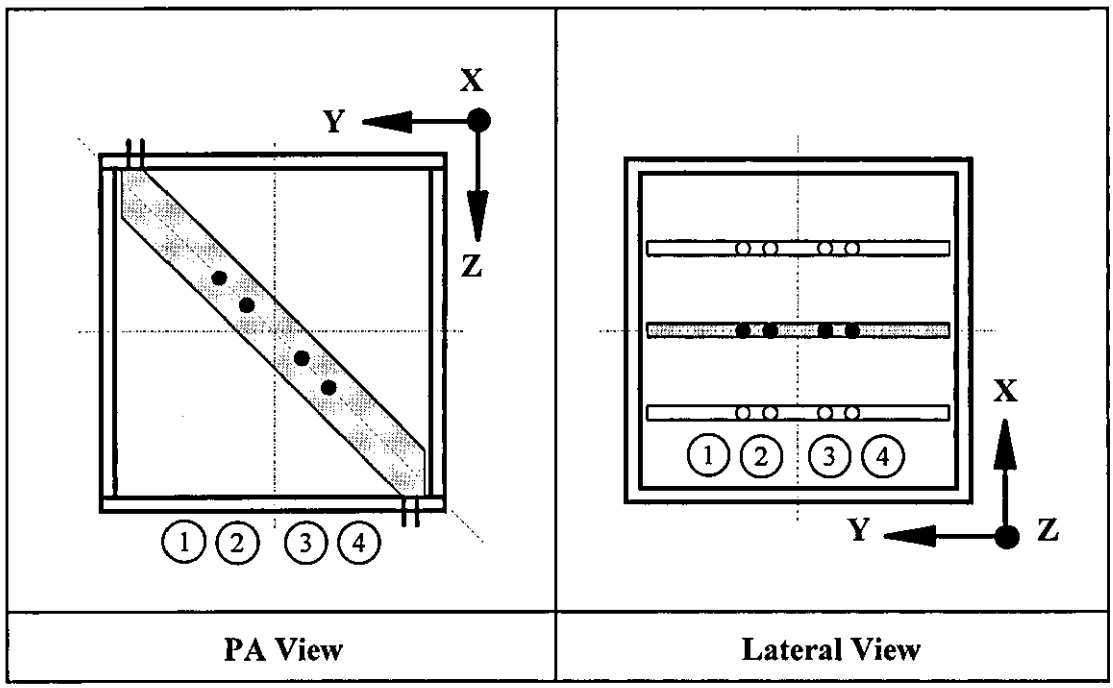


Fig. 5.11 : Perspex Test Object Inside Hip Box Calibration Frame

As indicated by Tables 5.1-5.3, the results obtained during trials involving the Perspex test object in conjunction with the Hip Box, suggested that the image analysis/photogrammetric reconstruction scheme implemented by *HB_Adpat.pas* and *HB_Prog.pas* did in fact lead to a satisfactory (sub-millimetre) level of precision. It is also interesting to note that in many of the calibration images acquired during these trials, the test object's ball-bearings were superimposed on top of Hip Box fiducial markers. As a result, when the image analysis program *HB_Adapt.pas* was used to scan the relevant images, incomplete sets of calibration data were obtained.

Table 5.1 : Typical Global Reconstruction Data (Upper Test Object Location)

Test Point	X	Y	Z	Abs. Error
1	169.23 (168.75)	147.71 (147.5)	78.08 (78.5)	0.67
2	168.97 (168.75)	130.21 (130.0)	95.92 (96.0)	0.31
3	169.45 (168.75)	95.72 (95.0)	130.81 (131.0)	1.02
4	169.99 (168.75)	77.28 (77.5)	148.38 (148.5)	1.26
			Mean Absolute Error =	0.82

Table 5.2 : Typical Global Reconstruction Data (Middle Test Object Location)

Test Point	X	Y	Z	Abs. Error
1	112.81 (112.5)	147.28 (147.5)	77.62 (78.5)	0.91
2	112.76 (112.5)	129.47 (130.0)	95.91 (96.0)	0.59
3	112.62 (112.5)	95.75 (95.0)	130.93 (131.0)	0.76
4	112.87 (112.5)	77.33 (77.5)	148.54 (148.5)	0.41
			Mean Absolute Error =	0.67

Table 5.3 : Typical Global Reconstruction Data (Lower Test Object Location)

Test Point	X	Y	Z	Abs. Error
1	56.16 (56.25)	147.68 (147.5)	78.03 (78.5)	0.51
2	56.36 (56.25)	130.08 (130.0)	95.94 (96.0)	0.15
3	56.54 (56.25)	95.64 (95.0)	131.01 (131.0)	0.70
4	55.97 (56.25)	77.07 (77.5)	148.71 (148.5)	0.55
			Mean Absolute Error =	0.48

Note: Figures in Brackets = Reference (Expected) Coordinates.

However, is must also be pointed out that the measurement routines implemented by *HB_Prog.pas* require the user to interactively indicate the centre of the four test object points in both views. As a consequence, the results listed in Tables 5.1-5.3 provide an indication of the magnitude of the errors obtained by a subjective manual indication process. A test program (*HB_Test.pas*) was therefore created with a view to automating, via the adaptive thresholding scheme, the measurement of the four test object points. Unfortunately, acquiring an image in which the test object points are not superimposed on top of Hip Box fiducials, in either the PA or lateral view, proved to be very difficult. Nevertheless, in the few instances in which a satisfactory image was obtained, automatic detection of the test object points did tend to lead to improved measurement accuracy.

5.4 Second Generation X-ray Compatible Software.

Although the first set of X-ray compatible software allowed an "acceptable" set of reconstruction results to be obtained, there remained considerable scope for improvement, in relation to the reliability and accuracy of the overall system. More specifically, the following limitations were noted:

- Image Distortion:

  The program *HB_Prog.pas* implements a global interpolation scheme, and as such does not fully address the problem of localised image distortion effects. The calculation of interpolation functions based upon large numbers of control points, spread across the entire field-of-view, also involves a large computational effort often leading to stack overflow problems.

- Image Segmentation:

  The simple adaptive thresholding scheme implemented by *HB_Adapt.pas* was designed to work with high contrast ("black" objects on a "white" background) images. Unfortunately, the orthopaedic applications under investigation by this thesis require the X-ray calibration frame and the patient's anatomy to be imaged simultaneously. Given that the resulting images no longer display a well defined

bimodal grey level distribution, the adaptive thresholding scheme performs poorly under such conditions.

- Fiducial Measurement:

  The image analysis techniques implemented by *HB_Adapt.pas* calculate the centre of gravity of an imaged fiducial marker to the nearest (integer) pixel value. Previous photogrammetry research strongly suggests that it is desirable to extract these measurements to sub-pixel (real values) precision.

In view of these shortcomings, the software upgrades documented in the following sections (Sections 5.4.1 and 5.4.2) were deemed to be necessary.

5.4.1 Image Analysis Software.

The image segmentation and analysis schemes outlined in this sub-section were originally developed by Josef Schnabler, an MSc student affiliated to the MEDROSA Project. These software algorithms were then subsequently modified by the author in order to achieve compatibility with the Hip Box fiducial patterns. As such, the following sub-section provides a brief overview of the approach, with a more detailed account of the relevant background theory being provided in the appropriate report **[192]**.

The appearance of an object in an X-ray image is determined by the amount of attenuation that occurs as the X-ray radiation passes through the object; which in turn is related to the atomic number, density and thickness of the object's material. Given that the ball-bearings used as fiducial markers in the Hip Box calibration plates have a spherical shape, the thickness of material exposed to X-ray radiation varies across the sphere's cross-section. As such, it is logical to assume that a digital X-ray image of a ball-bearing will consist of a dark region (i.e. lowest grey levels) at its centre, surrounded by progressively lighter (i.e. higher grey levels) regions corresponding to the edges of the sphere. If this digital X-ray image of the ball-bearing is subsequently inverted, the resulting grey level distribution will therefore be similar to the form shown in Figure 5.12.

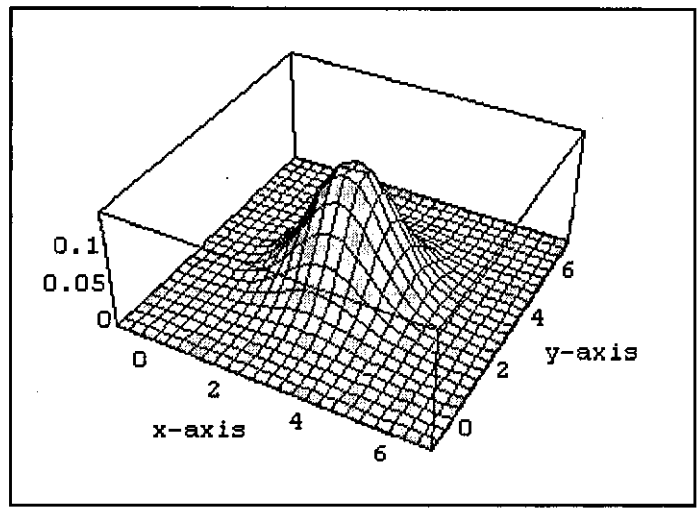


Fig. 5.12 : Predicted Grey level Distribution for an Inverted X-ray Image of a Sphere.
(Adapted from [192])

The new image analysis scheme uses a statistical method, based upon a 2D-normal distribution with two independent probability variables, to define an "ideal" model of this inverted X-ray image of a ball-bearing. The characteristic values of this model are then used as identification criteria when attempting to determined the presence, or absence, of a ball-bearing within a sub-region (area of interest) of an image. As such, in software terms, the image analysis task initially involves an image enhancement/segmentation process, which results in an inverted image similar to the form shown in Figure 5.12. The enhanced area of interest is then scanned to establish whether the remaining pixels do in fact represent the image of a ball-bearing (i.e. are the search parameters satisfied?).

Having defined a circular area of interest (AOI) surrounding a fiducial marker "candidate", the image enhancement routines are facilitated using the JLGenial library function *ImISample*. This software function scans the pixels within the AOI, and writes the corresponding grey level data to a dynamic array. The desired image enhancement processes are then applied to each pixel in the AOI by manipulating the grey level values in this array. Adopting this approach allows the required image inversion process to be achieved be applying the following function to each pixel in the AOI:

$$G_{inverted} = G_{max} - G_{original} \quad (5.1)$$

Where:

- G_{max} is the maximum grey level within the AOI.
- $G_{original}$ is the original grey level at a pixel location within the AOI.
- $G_{inverted}$ is the modified or "inverted" grey level at the corresponding AOI pixel location.

A segmentation process is then applied with the aim of eliminating the "background" pixels within the AOI. Given that the grey levels of the AOI have been inverted, the brightest pixels within the AOI will ideally be those corresponding to the ball-bearing. The segmentation function therefore "emphasises" or "magnifies" grey levels in the upper third of the range, and eliminates all other pixels by setting negative grey levels equal to zero. The function which achieves this segmentation process is of the form:

$$G_{segmented} = -2 \cdot b \cdot G_{inverted} + \frac{3 \cdot b}{G^{*}_{max}} \cdot G^{2}_{inverted} \quad .....(5.2)$$

Where:

- **b** is a magnification factor.
- G^{*}_{max} is the maximum grey level in the inverted AOI.
- $G_{inverted}$ is the grey level of a pixel in the inverted AOI.
- $G_{segmented}$ is the grey level at the corresponding pixel location in the segmented AOI.

Having enhanced/segmented the AOI, the fiducial identification/measurement process is instigated by calculating the following characteristic values of a 2D-normal distribution with two independent probability variances:

Sum of all grey levels:

$$G_{total} = \sum_{i=1}^{n} G_i \quad .....(5.3)$$

where:

- **n** is the number of pixels in the AOI.
- G_i is the grey level of the *i*th pixel location to be scanned in the AOI.

Mean coordinates ($\bar{x}$, $\bar{y}$):

$$\bar{x} = \frac{1}{G_{total}} \cdot \sum_{i=1}^{n} G_i . x_i \qquad .....(5.4)$$

$$\bar{y} = \frac{1}{G_{total}} \cdot \sum_{i=1}^{n} G_i . y_i \qquad .....(5.5)$$

Empirical variances (S_{xx}, S_{yy}, S_{xy}):

$$S_{xx} = \frac{1}{G_{total} - 1} \cdot \sum_{i=1}^{n} G_i (x_i - \bar{x})^2 \qquad .....(5.6)$$

$$S_{yy} = \frac{1}{G_{total} - 1} \cdot \sum_{i=1}^{n} G_i (y_i - \bar{y})^2 \qquad .....(5.7)$$

$$S_{xy} = \frac{1}{G_{total} - 1} \cdot \sum_{i=1}^{n} G_i (x_i - \bar{x}) \cdot (y_i - \bar{y}) \qquad .....(5.8)$$

Correlation coefficient ($-1 \leq R_{xy} \leq 1$):

$$R_{xy} = \frac{S_{xy}}{\sqrt{S_{xx}} \cdot \sqrt{S_{yy}}} \qquad .....(5.9)$$

The fiducial identification process then proceeds as follows. For an AOI containing a circular (spherical) object, the correlation coefficient R_{xy} should ideally be equal to zero. However, in practice this will obviously not be the case, and as such, the following condition is used to establish the presence/absence of a circular object within the AOI:

$$\left| R_{xy} \right| \leq \delta \qquad .....(5.10)$$

where:

- δ is an empirically determined value (typically in the range 0.1-0.3)

Assuming that a circular object has been "identified" (i.e. equation (5.10) is satisfied), the image analysis scheme then checks to see if the radius of the object is of the correct magnitude. This process takes advantage of the following relationships, which can be shown to be valid for a circular object **[192]**:

$$S_{xx} = \frac{R^2}{4} \quad \& \quad S_{yy} = \frac{R^2}{4} \qquad .....(5.11)$$

where:

- **R** is the radius (in pixels) of the circular object.

Prior knowledge of the actual radius (in millimetres) of the ball-bearing, and the image magnification factor at the calibration frame's typical working range, therefore allows selection criteria of the following form to be employed:

$$(R-a)^2 \leq 4S_{xx} \leq (R+a)^2 \quad \& \quad (R-a)^2 \leq 4S_{yy} \leq (R+a)^2 \qquad .....(5.12)$$

where:

- **a** is an empirically determined value, which dictates the allowable deviation from the predicted object (pixel) radius **R**.

If the selection criteria implemented by equations (5.10) and (5.12) are both satisfied, the fiducial identification process is deemed to be a success. The actual location of the newly identified fiducial is then obtained (to sub-pixel precision) from equations (5.4) and (5.5), with the centre of the fiducial (x_c, y_c) being equal to the mean coordinates $(\bar{x}, \bar{y})$ of the AOI.

In the case of a washer-shaped fiducial, the image analysis scheme described above can be implemented by making the following amendment to equation (5.11):

$$S_{xx} = \frac{R^2 + r^2}{4} \quad \& \quad S_{yy} = \frac{R^2 + r^2}{4} \qquad .....(5.13)$$

where:

- R is the outer diameter (in pixels) of the washer.
- r is the inner diameter (in pixels) of the washer.

All other aspects of the shape/size-based recognition and the fiducial measurement algorithms are then identical to the ball-bearing case.

In order to customise this image analysis scheme to the fiducial patterns of the Hip Box, a simple development program (*IB_Test2.pas*) was written to allow the influence of the important variables (AOI radius, expected fiducial radius, maximum R_{xy} limit δ, etc.) to be studied. Having empirically determined appropriate values for these variables, it was possible to identify both the advantages and the limitations of the scheme.

The most obvious advantage was the ability to locate fiducials to sub-pixel precision in low contrast images (e.g. when an object is imaged inside the Hip Box). However, it also became evident that in the rare case of the fiducial and background grey levels not differing significantly (i.e. the difference is less than say four grey levels) poor segmentation/measurement results can be obtained. In particular, during tests involving the Hip Box and femoral phantoms (plastic bones) it was found that problems could be encountered when attempting to locate the four small washers used to define the centre of the XRII fiducial patterns. The particular scenario in which a fiducial appears at the edge of an imaged object, and as such, the AOI background contains both dark and light areas, can also lead to problems.

Based upon these findings, it was decided that the four washers at the centre of the XRII grid should eventually be replaced by ball-bearings (i.e. fiducials of greater material thickness). However, given that the use of a larger washer as an orientation marker in the source plate allows the user to easily differentiate it from the centre marker, it was believed that this washer should be retained. In the short-term, a

software-based solution to the problems associated with the detection of washers was therefore implemented. Should the need arise, this "manual override" approach allows the user to interactively indicate the centre of the washer(s). The use of custom-made lead or even tantalum "washers" to provide higher contrast fiducials within the image has also been proposed as a long-term solution.

Improvements in the success rate of the fiducial detection scheme were also obtained by customising the AOI radius to the individual fiducial types. Reducing the AOI radius results in a significant reduction in the background-to-fiducial pixel ratio, and hence the potential for background pixels to corrupt the image segmentation process. However, as the AOI radius becomes smaller, it becomes increasingly more difficult to locate fiducials using the (global) predicted locations scanning technique adopted during the camera-based study (i.e. the likelihood of the fiducial lying within the predicted AOI decreases). A new grid scanning technique was therefore developed to maximise the potential of the normal distribution based image analysis scheme.

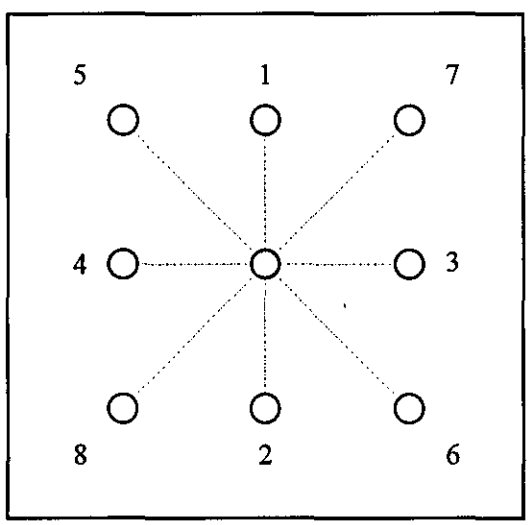


Fig. 5.13 : Fiducial Neighbourhood Scanning Order.

Assuming that the appropriate grid measurements (i.e. grid spacing, angular orientation, etc.) have been obtained, by semi-automatically locating either the centre/orientation markers or the four washers, this new scanning routine proceeds as follows. Starting at an appropriate point in the grid (e.g. the centre marker, or one of the four washers), AOI's centred at the predicted locations of the eight neighbouring

fiducials are scanned, in the order shown in Figure 5.13, until a successful fiducial identification is achieved. Having found this second fiducial position, the predicted locations of its eight neighbouring fiducials are then scanned in the same order, until another fiducial is located. The scanning process then continues in this manner until a point is reached where a search of all eight of a fiducial's predicted neighbours, does not lead to the location of a previously undetected fiducial.

In the event of the neighbourhood of the *n*th fiducial location not leading to the discovery of a new fiducial, the recursive algorithm returns to the (*n-1*)th fiducial location. If for instance the search of the (*n-1*)th fiducial's neighbourhood led to the discovery of the *n*th fiducial at location 3, the algorithm will then resume the scanning process at location 4 in the (*n-1*)th fiducial's neighbourhood. In this way, having checked the eight neighbouring locations of every fiducial (of the relevant calibration grid pattern) within the field-of-view, the scanning algorithm eventually returns to its starting point.

The main advantage offered by this recursive scanning routine is the fact that a fiducial does not have to be located at the first attempt. Since the predicted location of an AOI is now calculated using actual (previously located) fiducial coordinates, each subsequent attempt to find a given fiducial involves a slightly different search region. This localised approach to image scanning therefore significantly increases the chances of a given fiducial being located. The use of smaller AOIs also allows searches to be performed nearer to the periphery of the image, thereby leading to better coverage of the field-of-view.

5.4.2 Investigation of The Image Distortion Problem.

As previously discussed in Section 2.4, the image produced by an image intensifier tube is subject to several different forms of image distortion. In the case of pincushion distortion, which is arguably the most dominant form of image distortion, the distortion effects are self-evident. However, the effects attributable to the geomagnetic field, in particular the image shifting (translational) effect, are not readily noticeable in the output image from a C-arm fluoroscopy unit. Before attempting to

implement an image distortion correction (internal calibration) scheme, a full investigation of the imaging characteristics of the Siemens Siremobil 2N was therefore necessary. In-line with the "classical" approach to image distortion evaluation, a calibration object (grid), which could be attached to the input window of the image intensifier housing, was therefore designed.

The actual form of this calibration object is shown in Figure 5.14. In view of the large numbers of fiducial markers involved, it was decided that the Perspex sheet/ball-bearing approach to providing a calibration grid would not be used in this instance. Instead, the much simpler and hence less time-consuming option of accurately drilled holes in an aluminium plate was adopted. As such, the image distortion calibration plate introduces an almost circular (diameter 186 mm) fiducial grid pattern, consisting of 1 mm diameter holes (at a 3 mm pitch), into the field-of-view of the X-ray image intensifier. A larger (2 mm diameter) reference hole is also provided to mark the centre of the grid. When attached to the image intensifier housing, the calibration plate therefore produces an output image consisting of "white" circles on a "black" background.

By acquiring images of the calibration plate at a variety of C-arm arc settings, and then analysing the resulting images, it was possible to demonstrate the influence of image intensifier orientation, in relation to the Earth's magnetic field, upon the output image. In particular, the use of "difference" or "subtraction" images proved to be a simple and effective means of demonstrating the presence of S-shaped distortion and image shifting. Figure 5.15 was created by digitally subtracting a "PA" (C-arm arc rotation = 0°) image of the calibration plate from the corresponding "lateral" (C-arm arc rotation = 90°) image of the plate. Assuming that the plate has not moved, in relation to the image intensifier housing, these images should ideally be identical, thus producing a blank "difference" image. However, this is evidently not the case, with the "distorted" locations of both the "PA" (white) and the "lateral" (black) fiducials being clearly visible in Figure 5.15.

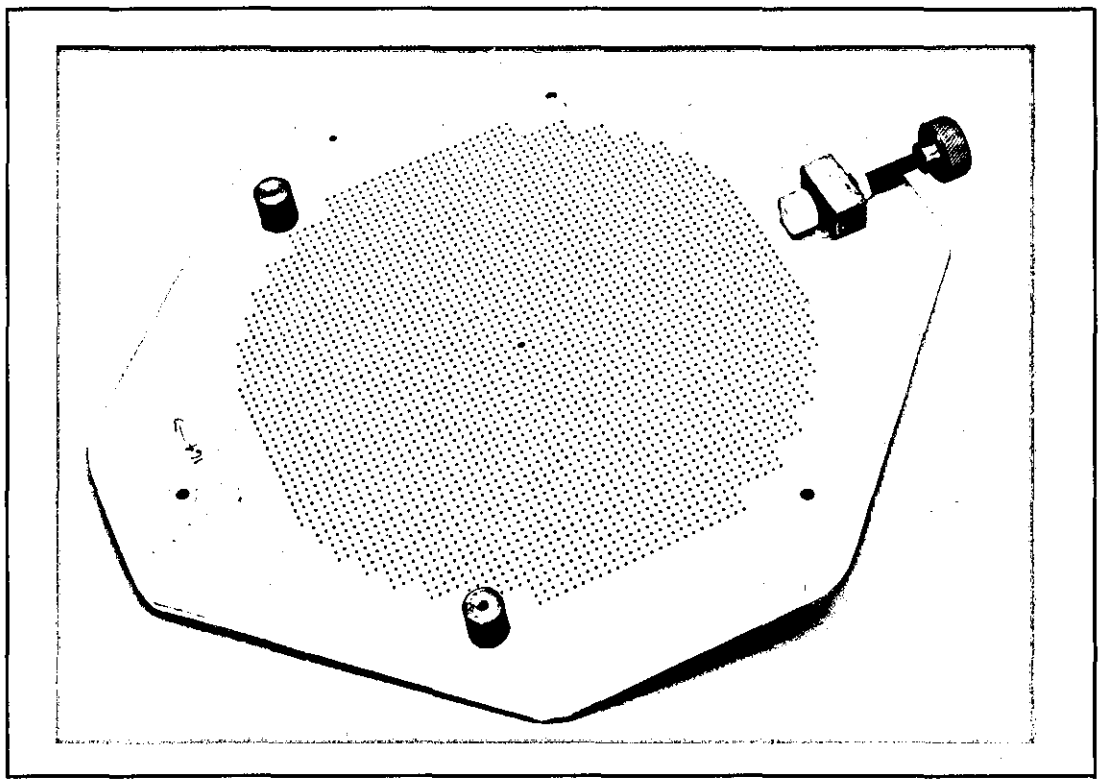


Fig. 5.14 : The Image Distortion Calibration Plate.

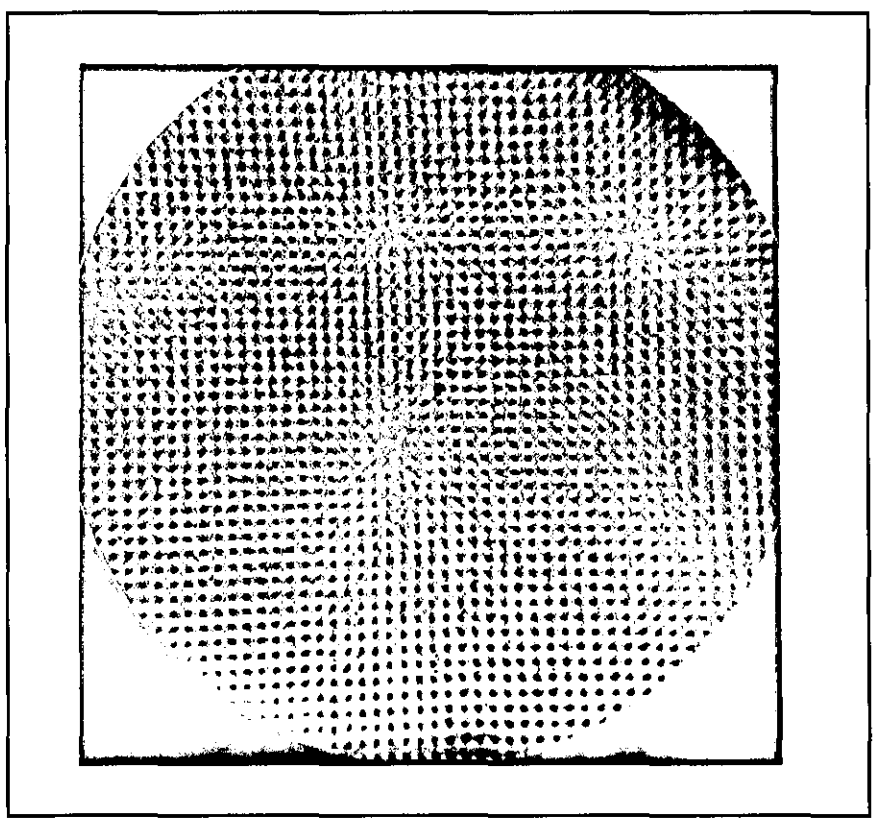


Fig. 5.15 : A Typical PA/Lateral Subtraction Image.

Since the relative locations of the corresponding "PA" and "lateral" fiducials vary significantly across the field-of-view, a simple translation/rotation of the calibration plate can obviously be eliminated as the cause of the discrepancies in Figure 5.15. It is therefore safe to assume that the geomagnetic distortion effect has successfully been demonstrated. The "blank" regions within the field-of-view (i.e. areas of "zero difference") also confirm both the presence and the localised nature of this distortion effect. By measuring the location of the grid's centre marker in corresponding "PA" and "lateral" images, it was also possible to estimate the magnitude of the image translation caused by the geomagnetic distortion effect. Given that this shift was typically of the order of eight or more pixels (i.e. 2-3 mm), a re-evaluation of the proposed image distortion correction scheme therefore became necessary.

Having empirically determined that geomagnetic image distortion effects were too large to be ignored, implementing one of the "classical" (off-line) image distortion correction schemes outlined in Section 2.4.2 was no longer a viable proposition. As a consequence, an on-line internal calibration technique capable of compensating for the localised effects of geomagnetic distortion was required. Given that a calibration frame (i.e. the Hip Box) was already being used to support the external calibration process, it seemed logical to attempt to use the same frame to provide reference data for the internal calibration scheme. An investigation of the on-line use of a localised interpolation scheme was therefore performed.

The starting point for the software implementation of this internal calibration scheme was the localised interpolation version of the camera calibration program *Twoplane.pas*. As described in Section 4.4.1, this software uses the calibration frame fiducials to define triangle-shaped image sub-regions. The known world coordinates and measured image coordinates of the three fiducials defining each triangle are then used to derive localised interpolation functions. By modifying the algorithm which determines the three closest fiducials to a particular point in the image, thus ensuring that the three points do not lie along a straight line, it was also possible to implement an extrapolation scheme for points at the edges of the image (i.e. points not surrounded by three fiducials).

Upon completion, the accuracy of the resulting program *HBL_Prog.pas* and its supporting units (*HBL_Foc.pas*, *HBL_Surg.pas* & *HBL_Type.pas*) was then tested, using the previously described Hip Box/Perspex text object arrangement. The image data used during these tests was derived using the program *HB_Corel.pas*, which implements the sub-pixel image analysis and scanning routines described in Section 5.4.1. Once again, following calibration of the system, the locations (in both the PA and lateral views) of the four ball-bearings embedded in the test object were interactively indicated.

The random errors associated with the manual indication process mean that direct comparisons with the results obtained by the original global interpolation (*HB_Prog.pas*) software cannot be made. However, as indicated by Tables 5.4-5.6, repeated tests did tend to suggest an improvement in reconstruction accuracy when using the new localised interpolation software (*HBL_Prog.pas*). The results summarised in Table 5.7 also demonstrate that the average "mismatch" value is significantly reduced when using the localised approach. The introduction of image distortion compensation and sub-pixel image analysis routines would therefore appear to have produced the desired improvement in the accuracy of the photogrammetric reconstruction process.

Table 5.7 : Average "Mismatch" Values.

Test Object Location	Global Interpolation	Localised Interpolation
Upper	0.55 mm	0.18 mm
Middle	0.61 mm	0.19 mm
Lower	0.37 mm	0.21 mm

It must also be noted that the problem of test object ball-bearings being superimposed on top of Hip Box fiducials, will undoubtedly have led to a reduction in the accuracy of the localised interpolation scheme. As such, the true accuracy of the system, under ideal conditions, may be even better than the results supplied in Tables 5.4-5.6 suggest.

Table 5.4: Typical Localised Reconstruction Data (Upper Test Object Location)

Test Point	X	Y	Z	Abs. Error
1	170.37 (169.75)	148.12 (148.5)	78.56 (78.5)	0.73
2	170.53 (169.75)	130.76 (131.0)	96.23 (96.0)	0.85
3	170.29 (169.75)	96.09 (96.0)	131.05 (131.0)	0.55
4	170.18 (169.75)	78.59 (78.5)	148.44 (148.5)	0.44
			Mean Absolute Error =	0.64

Table 5.5 : Typical Localised Reconstruction Data (Middle Test Object Location)

Test Point	X	Y	Z	Abs. Error
1	113.62 (113.5)	148.09 (148.5)	79.30 (78.5)	0.48
2	113.60 (113.5)	130.79 (131.0)	95.82 (96.0)	0.30
3	113.44 (113.5)	96.33 (96.0)	131.01 (130.0)	0.33
4	113.72 (113.5)	78.48 (78.5)	148.30 (148.5)	0.30
			Mean Absolute Error =	0.35

Table 5.6 : Typical Localised Reconstruction Data (Lower Test Object Location)

Test Point	X	Y	Z	Abs. Error
1	57.54 (57.25)	148.47 (148.5)	78.34 (78.5)	0.33
2	57.45 (57.25)	130.91 (131.0)	95.95 (96.0)	0.22
3	57.45 (57.25)	96.34 (96.0)	130.86 (131.0)	0.42
4	57.47 (57.25)	78.69 (78.5)	148.46 (148.5)	0.29
			Mean Absolute Error =	0.32

Note: Figures in Brackets = Reference (Expected) Coordinates.

The reasoning for this loss of accuracy is as follows. A superimposed test object ball-bearing obscures one of the fiducials making up the ideal interpolation triangle. The algorithm that determines the three closest points is therefore forced to allocate another point, which is further away from the test object ball-bearing than the true (superimposed) fiducial location. As a consequence, the subsequent interpolation process takes place over a larger triangular sub-region of the image. Since the localised interpolation scheme is essentially using a linear interpolation function to approximate a non-linear (image distortion) effect, this increased sub-region size will undoubtedly result in reduced reconstruction accuracy.

5.5 Summary of Chapter 5.

A full investigation of the digital X-ray photogrammetric reconstruction process has been undertaken using a Siemens Siremobil 2N mobile fluoroscopy unit, which was operated from the outside of a purpose-built lead-lined X-ray enclosure. Having initially adapted the techniques developed during the camera-based study to achieve X-ray "compatibility", a further period of research has subsequently led to the development of software-based solutions to the following problems:

- Semi-automatic segmentation of fiducial markers from "low" contrast X-ray images.
- Measurement of fiducial marker locations to sub-pixel levels of precision.
- Compensation for the effects of both pincushion and geomagnetic image distortion.
- Solution of the photogrammetric reconstruction problem.

The development of these solutions has been facilitated by manufacturing a Perspex biplane X-ray calibration frame, referred to as the "Hip Box". When imaged simultaneously with the "patient" (as represented by a phantom femur) this frame superimposes a fiducial marker distribution, which has been specifically designed to simplify the image analysis task, into the resulting fluoroscopic images. By extracting the appropriate fiducial location data from these images, both the external

(photogrammetric reconstruction) and the *internal* (*image distortion compensation*) calibration processes can then be performed *on-line*.

Trials involving the use of a custom-made test object, in conjunction with the Hip Box, indicate that sub-millimetre levels of precision can be obtained via this calibration frame-based approach. As a consequence, the technique could potentially be very beneficial to a wide variety of orthopaedic surgery procedures. However, in the context of the MEDROSA Project, the investigation is initially limited to femoral osteosynthesis procedures. Chapters 6 and 7 therefore outline the methodology adopted during follow-up studies aimed at customising this frame-based technique to satisfy the requirements of two specific orthopaedic procedures.

Software Summary:

Program 1: *HB_Corel.pas*

- *Description*: Statistical (semi-automatic) image analysis routines.
- *Inputs*: PA and lateral images of the Hip Box calibration frame.
- *Outputs*: Fiducial (sub-pixel) image coordinate data written to file.

Program 2: *HBL_Prog.pas*

- *Description*: On-line (localised) internal and external calibration routines.
- *Inputs*: Fiducial (sub-pixel) image coordinate data read from file.
- *Outputs*: The 3-D coordinates of reconstructed target points, with respect to the Hip Box calibration frame's coordinate system.

Support Units:

- *HBL_Type.pas* (global declarations common to all units/programs)
- *HBL_Foc.pas* (estimation of the lateral view effective focal point).
- *HBL_Surg.pas* (image matching and photogrammetric reconstruction routines).

6. INTRAMEDULLARY NAILING OF FEMORAL SHAFT FRACTURES.

6.1 Femoral Anatomy.

A clear understanding of the internal fixation of femoral fractures requires familiarity with the anatomy of the femur, and the associated orthopaedic terminology. Accordingly, this section provides a brief description of the important aspects of femoral anatomy, with a view to clarifying both the discussion of femoral shaft fractures presented in this chapter, and the discussion of hip fractures presented in the next chapter.

In general, when describing the location of anatomical structures, the following terminology is used in relation to a standing person facing the examiner **[193]**:

- *Anterior*: denotes the front surface of the structure.
- *Posterior*: denotes the back surface of the structure.

- *Superior*: denotes the upper side of the structure.
- *Inferior*: denotes the lower side of the structure.

- *Lateral*: denotes the side furthest away from the midline of the body.
- *Medial*: denotes the side closest to the midline of the body.

- *Proximal*: denotes the part of the structure closest to the centre of the body.
- *Distal*: denotes the part of the structure furthest away from the centre of the body.

The femur (thigh bone) is a prime example of a *long bone*. That is to say, it consists of an elongated tubular shaft (or *diaphysis*) connecting two extremities (or *epiphyses*), which are wider than the shaft itself (refer to Figure 6.1). Applying the above terminology therefore leads to the use of the term *proximal femur*, to signify the femoral extremity in the vicinity of the hip joint, and the term *distal femur*, to signify the femoral extremity in the vicinity of the knee joint.

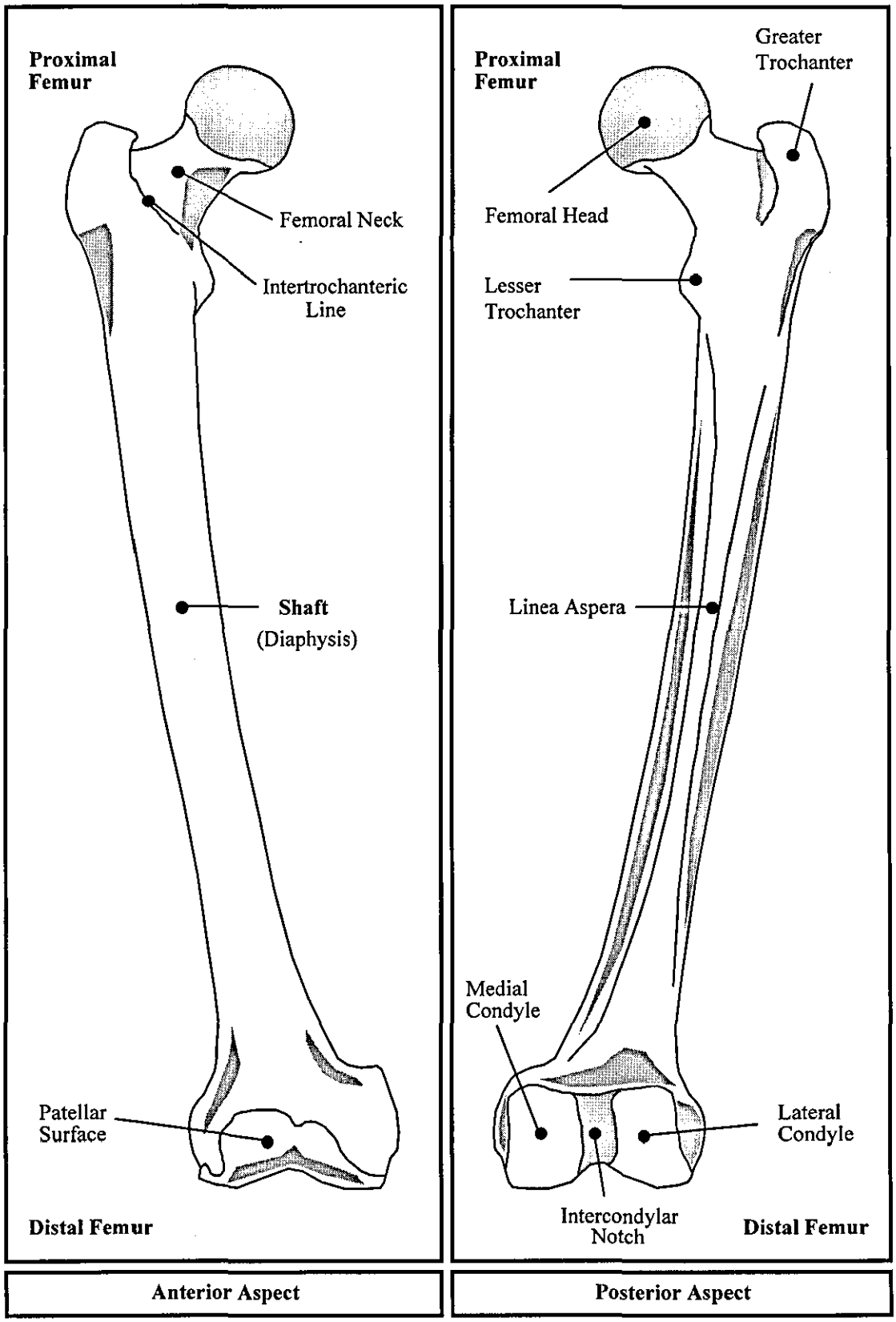


Fig. 6.1 : Anatomy of the Femur.

The proximal femur incorporates an almost spherical surface (the *femoral head*), which articulates with the *acetabulum* (a cup-shaped depression in the pelvis) to form the ball-and-socket of the hip joint. In order to allow free movement of the leg about the hip joint/pelvis, the femoral head is attached to the shaft of the femur, at an angle of approximately 125°, by a short piece of bone referred to as the *femoral neck* **[194]**. Immediately distal to the femoral neck, the bone widens into two large prominences, which act as attachment sites for the muscles of the upper leg. The lower of these prominences (on the medial side of the femur) is referred to as the *lesser trochanter*, while the larger upper prominence (on the lateral side of the femur) is referred to as the *greater trochanter*. On the anterior aspect of the proximal femur, the junction between the neck and the shaft is also marked by a ridge of bone (the *intertrochanteric line*) running from the greater to the lesser trochanter.

The distal femur widens significantly to form two bony masses, known as the *lateral* and *medial femoral condyles*. These knuckle-like prominences articulate with the *patella* (kneecap) and the *tibia* (shin bone), to provide the kinematics of the knee joint. Accordingly, the anterior aspect of the distal femur consists of a continuous articular surface, known as the *patellar surface*, extending across both condyles. However, in order to make room for the cruciate ligaments within the capsule of the knee joint, the two condyles are separated posteriorly by the *intercondylar notch*. The distribution of weight from the femur to the tibia is therefore divided between two articular surfaces.

In keeping with most long bones, the shaft of the femur is composed of a hollow "tube" of compact (*cortical*) bone, which surrounds a central region (*medullary cavity*) of soft, fatty bone marrow and spongy (*cancellous*) bone **[195]**. With the exception of a ridge of bone (the *linea aspera*) running down the posterior aspect, the exterior of the femoral shaft is essentially smooth. The purpose of this ridge (i.e. increased bone thickness) is to withstand the concentration of forces that arises, due to the fact that the femoral shaft is slightly bowed anteriorly. As can be seen from Figure 6.1, in the standing position, the femoral shaft normally inclines about 10° from the vertical axis of the tibia.

6.2 Femoral Shaft Fractures.

The femur is both the longest and the strongest bone in the human body **[194]**. While its length is obviously associated with mankind's striding gait, the strength of the femur reflects its role as a primary weight-bearing bone. In addition to this inherent strength, the massive muscle groups which surround the femur also afford a certain amount of protection against externally applied forces. As a consequence, fractures of the femoral shaft are usually associated with direct high-energy trauma, and in particular, road traffic accidents. Understandably, this form of injury is therefore most commonly suffered by children and young adults.

The most common site for a femoral shaft fracture is in the middle third of the shaft (i.e. the narrowest part of the shaft) **[196]**. However, as indicated by Figure 6.2, fractures can also occur at the proximal or distal ends of the femoral shaft. The actual geometry of the fracture is closely related to the type of force(s) applied to the bone. Angular forces, resulting from a violent blow to the bone, produce a *transverse fracture*, which is approximately perpendicular to the axis of the femoral shaft (refer to Figure 6.2). Given that the femoral shaft is essentially a tubular structure, torsional forces, arising from violent twisting about the axis of the bone, are also a common cause of fracture. In such cases, either an *oblique* or a *spiral fracture* is generally produced (refer to Figure 6.2).

The severity of the fracture is related to the magnitude of the applied force(s). As indicated by Figure 6.3, a large angular force can produce either a *segmental fracture* (i.e. two transverse fractures leaving a section of the femur isolated) or a *wedge fracture* (i.e. two transverse fractures producing a "butterfly" fragment). More severe forms of *comminuted fracture* (i.e. fractures resulting in more than two bone fragments) can also result from a large impaction force, caused for example by a direct blow to the flexed knee joint.

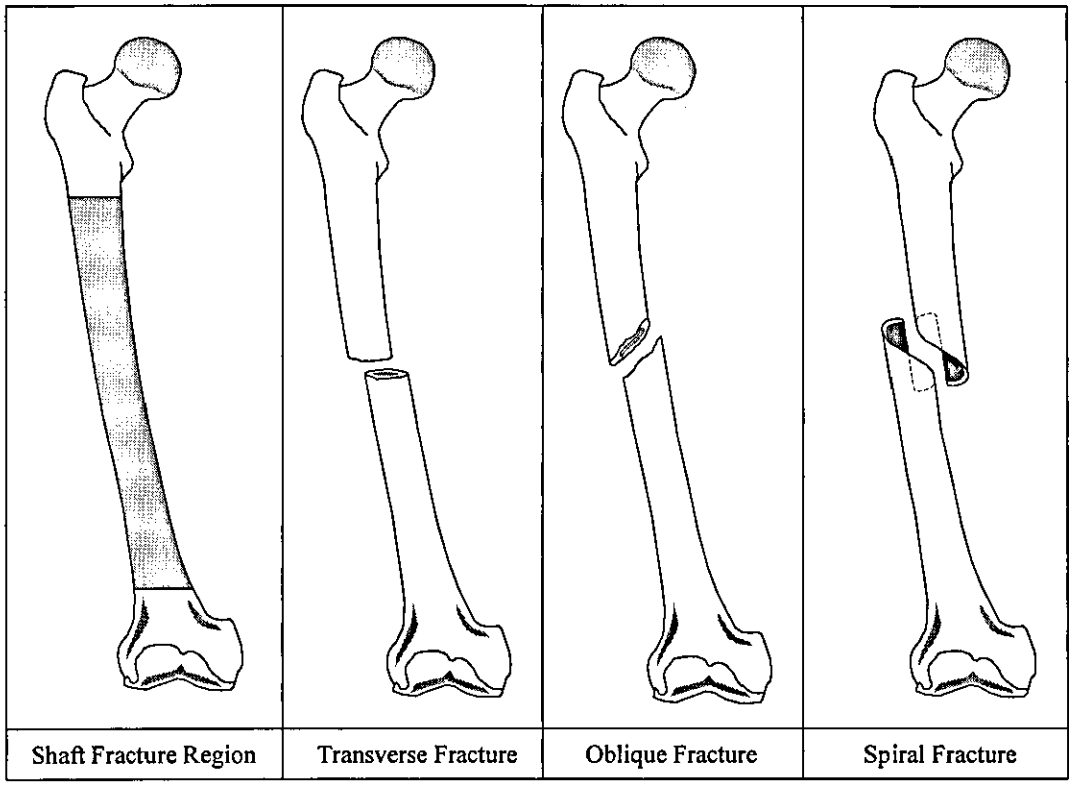


Fig. 6.2 : Simple Femoral Shaft Fracture.

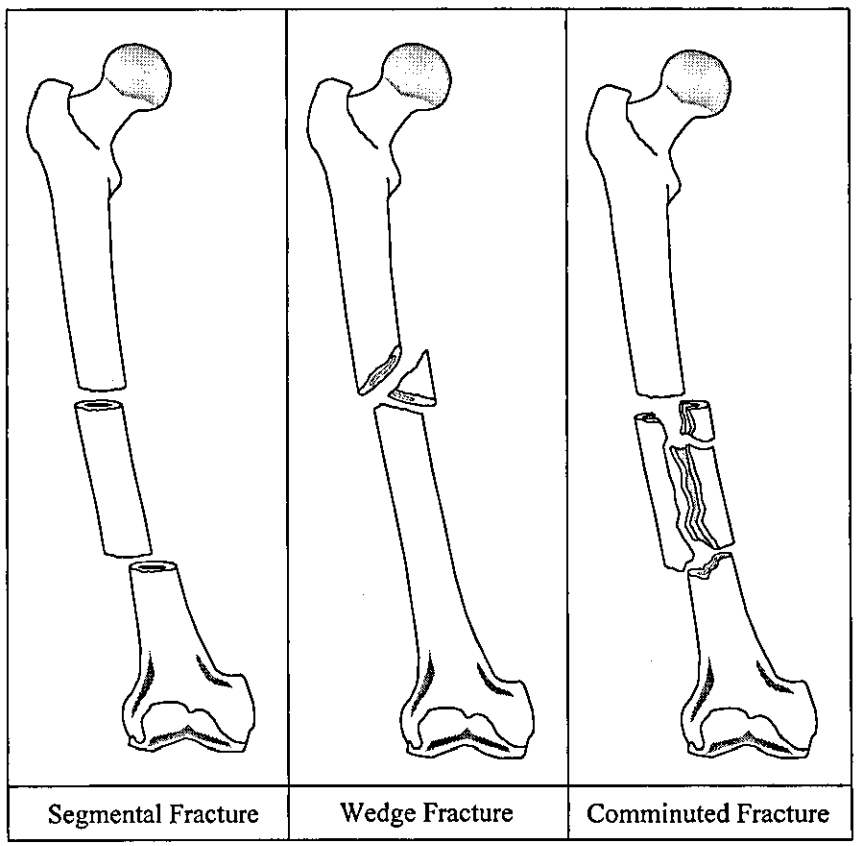


Fig. 6.3 : Complex Femoral Shaft Fractures.

Because of the levels of violence associated with femoral shaft fractures, it is not uncommon for the patient to suffer additional fractures and significant soft tissue damage. In severe cases, bone fragments can also penetrate the skin from within, thus producing a *compound* or *open fracture*. Immediately after sustaining the injury, the patient must therefore be treated for shock. A blood transfusion and prophylactic antibiotics will also be subsequently administered, as part of the emergency treatment aimed at stabilising the patient's condition **[197]**. An appropriate orthopaedic technique can then be applied to deal with the actual shaft fracture.

6.3 Locked Intramedullary Nailing.

The successful treatment of bone fractures, initially requires the restoration of the broken bone fragments to their normal anatomical positions and alignment. This process is referred to as *reduction* of the fracture, and involves either a manipulative or a surgical procedure. A *stabilisation* technique is then applied, in order to ensure that the bone fragments maintain this anatomically correct position throughout the natural bone healing process. The continuity of the bone is then restored by the formation, and progressive hardening, of a *callus* at the fracture site.

In the case of femoral shaft fractures, poor reduction/stabilisation of the fracture site can lead to serious complications. When the femoral shaft is fractured, the muscles attached to the bone fragments contract, thereby shortening the leg. In order to prevent prolonged disability it is therefore essential that normal limb length is restored. Malalignment of the bone fragments, which can alter the biomechanics of the hip and knee joints, must also be avoided, in order to prevent joint disorders such as osteoarthritis developing. Because the thigh is so heavily padded with muscle, an external plaster cast cannot provide the required level of stabilisation. As such, the selection of an appropriate treatment is limited to one of the following techniques, with the final decision being made on the grounds of several fracture (type and location, degree of comminution, open or closed, etc.) and patient (age, general health, severity of any additional injuries, etc.) related variables **[198]**:

- Skeletal traction.
- Cast bracing.
- External fixation.
- Internal fixation.

Skeletal traction of a femoral shaft fracture involves the insertion of a metal pin through the proximal tibia **[199]**. Weights are then attached to this pin, via a series of ropes and pulleys, to provide a traction force which is strong enough to overcome the contraction of the leg muscles. A prolonged period of hospitalisation is then required, in order to allow the bone healing process to take place. Skeletal traction is therefore an expensive form of treatment, and can also incur a number of the complications associated with prolonged bed rest (i.e. muscle atrophy, pressure sores, etc.). As such, its use is generally restricted to severely comminuted fractures, which cannot be treated surgically. The period of hospitalisation can however be shortened, by applying a cast brace when the fracture becomes stable, as indicated radiographically by the appearance of a callus. This protective device is similar in appearance to a prosthetic limb, and includes a hinged knee joint to allow free leg motion, thus aiding muscle recovery.

Fixation techniques achieve stabilisation of the fracture site by surgically inserting metallic devices (e.g. screws, pins, nails, rods, plates, etc.) into the bone. These devices then hold the bone fragments in a fixed relationship to one another while the fracture heals. A distinction is made between *external fixation* and *internal fixation* (*osteosynthesis*), to indicate whether the bulk of the fixation device is situated outside or inside the patient's body. In the context of femoral shaft fractures, the use of external fixation devices tends to be limited to contaminated or unstable open fractures **[197]**. In such cases, the external fixator provides temporary stabilisation of the fracture while the contaminated wound is being treated. Upon closure of the wound, the external fixator is then replaced by skeletal traction. However, in recent years, internal fixation has become the preferred stabilisation technique for most long bone fractures.

The popularity of internal fixation devices stems from the fact that they permit load-bearing across the fracture site, and as such, allow early postoperative mobilisation of the patient. A number of these internal fixators can also be inserted using closed surgery techniques, thereby minimising surgical trauma by avoiding exposure of the fracture site. In relation to femoral shaft fractures, the most effective form of internal fixation is closed intramedullary nailing. As the name suggests, the *intramedullary nail* is inserted into the medullary cavity of the femoral shaft (i.e. inside the femur). Prior to insertion of the nail, in order to establish an interference fit between the bone and the nail, the medullary cavity is surgically prepared using a series of flexible reamers. The resulting friction between the bone and the nail therefore ensures that both the length and the alignment of the broken bone fragments are maintained.

Although conventional intramedullary nailing provides adequate stabilisation of mid-shaft fractures, problems can be encountered when the technique is applied to more distal/proximal shaft fractures. In particular, when torsional forces are applied to the femur, friction between the nail and the bone can be insufficient to control rotation of the bone fragments. In order to rectify these problems, and hence extend the scope of intramedullary nailing, the concept of *interlocking* has therefore been introduced. This technique involves rigidly attaching the nail to the bone fragments, in order to provide more effective fixation.

Interlocking nail designs are generally categorised as using either an internal or an external locking mechanism **[200]**. In the case of internal systems, the locking mechanism is integral to the nail's design. As such, following insertion of the nail into the medullary cavity, the locking mechanism is deployed from the proximal end of the nail. The actual locking system generally comprises of either internally deployed fins/wings, as used by the Brooker-Wills **[201]** and Derby **[202]** nails, or a set of flexible rods/pins, as used by the Marchetti/Vicenzi and LIFO **[203]** fixation systems, which spread out to grip the distal femur. However, because problems can be encountered when attempting to deploy an internal locking mechanism, external locking is currently the more widely used technique.

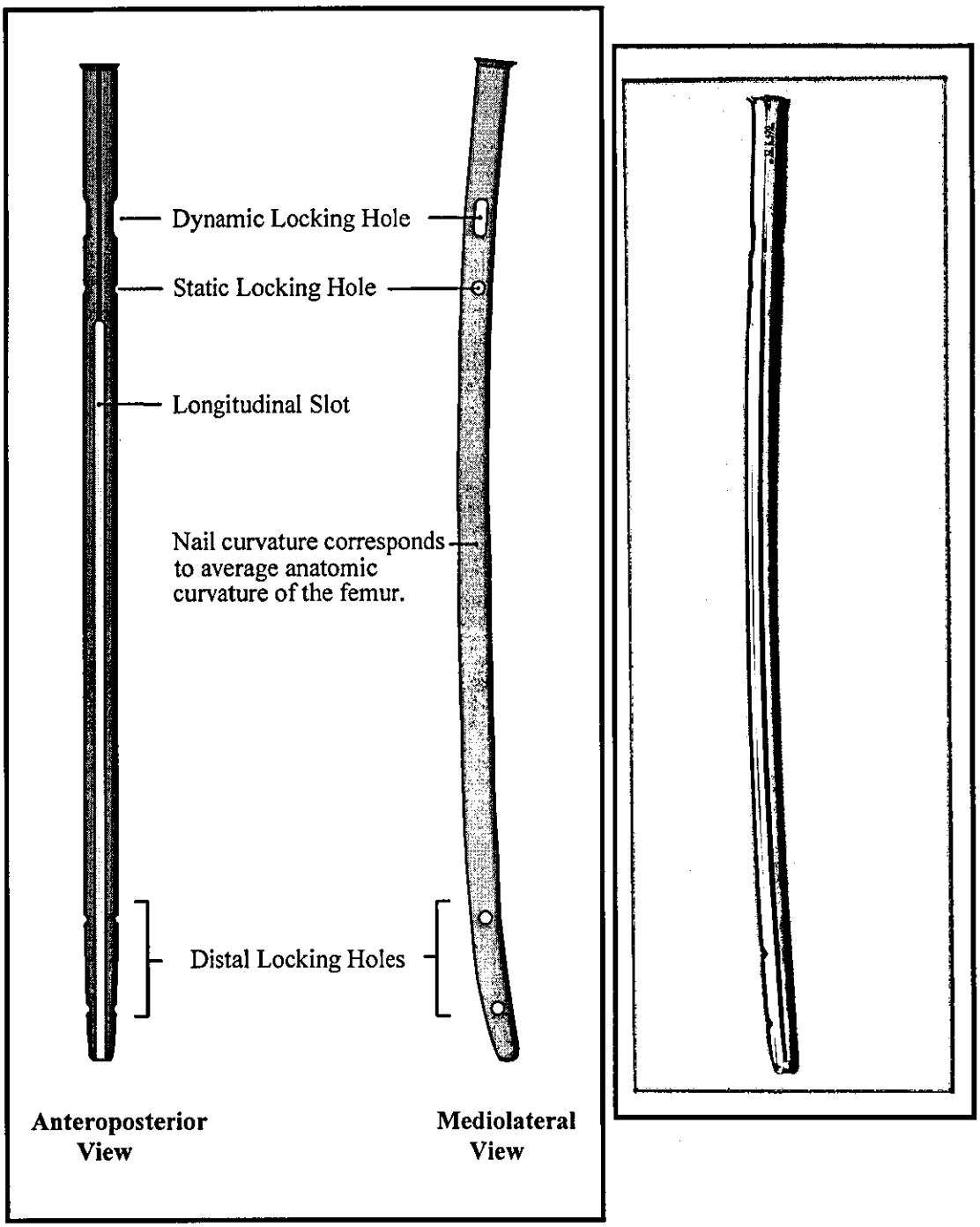


Fig. 6.4 : The AO/ASIF Universal Femoral Nail.

The AO/ASIF Universal Femoral Nail **[204]** shown in Figure 6.4 is a typical example of an externally locked femoral intramedullary nail. It consists of a hollow stainless steel tube, which is pre-bent to match the curvature of the femur, and incorporates a continuous longitudinal slot to prevent stress concentrations at the upper end of the nail. Interlocking of the AO/ASIF nail is achieved by inserting transfixion screws in the manner shown in Figure 6.5. Accordingly, the nail's design includes a pair of holes at both the distal and proximal ends. This arrangement allows two possible modes of fixation, which are referred to as *static* and *dynamic* locking **[205]**.

Static locking (as depicted in Figure 6.5) involves the insertion of screws through both ends of the nail, and is used to prevent axial compaction of severely comminuted fractures. Dynamic locking on the other hand, involves the insertion of a screw(s) through only one end of the nail, and is applied to more stable fractures in which there is sufficient contact between the bone fragments to prevent shortening. Given that dynamic locking allows impaction of the fracture site, and hence early weight-bearing, conversion from static to dynamic fixation is frequently performed, by the removal of a screw(s), when a satisfactory callus has been produced **[206]**.

6.4 Closed Intramedullary Nailing Technique.

The precise form of an intramedullary nailing procedure is influenced by several factors, and as such, can vary considerably from hospital to hospital. Given that an inventory of nails must be maintained in order to accommodate anatomical variations, intramedullary nailing is a lucrative market for manufacturers of surgical instrumentation. Accordingly, a wide variety of nail designs are currently available, each of which is supplied with its own customised set of surgical insertion tools. A second major source of discrepancies is the surgeon's personal preferences with regard to surgical techniques. Owing to the freehand nature of orthopaedic surgery, the training of orthopaedic surgeons involves a long learning curve. As an individual surgeon's level of experience with a particular surgical procedure increases, he therefore refines his surgical technique to suit his own personal capabilities.

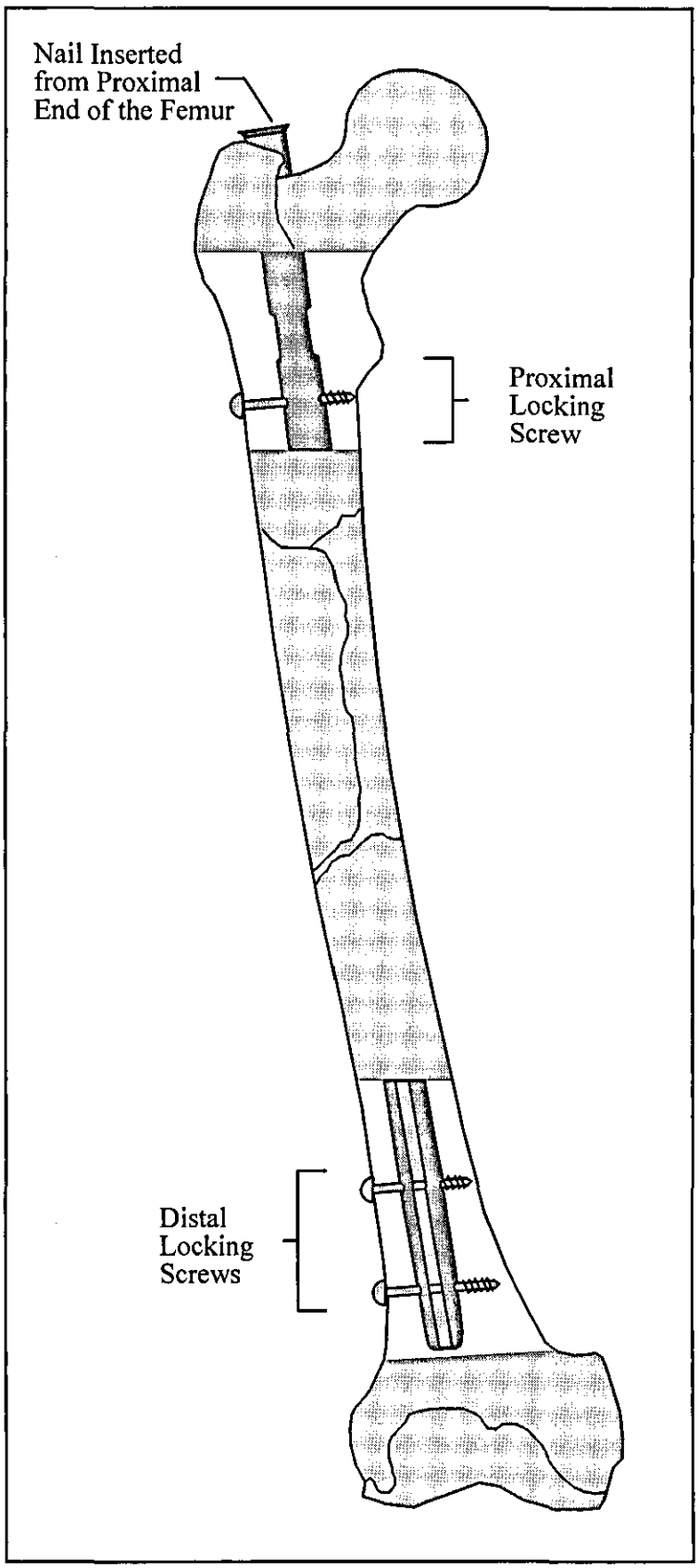


Fig. 6.5 : Locked Intramedullary Nailing of a Femoral Shaft Fracture.

In view of this lack of standardisation, the following overview of the locked intramedullary nailing procedure is based upon textbook recommendations **[198, 204, 207]** and, to a lesser extent, the clinical experiences of Kempf *et al.* **[206]** and Winquist *et al.* **[208]**. Ideally, the insertion of the intramedullary nail is performed as a closed surgery procedure, with the nail being introduced via a small incision at the proximal end of the femur. In practice, this closed surgery procedure involves the following key stages:

- Preoperative Planning.
- Patient positioning and preparation.
- Closed reduction.
- Opening of the medullary cavity.
- Reaming of the medullary cavity.
- Nail insertion.
- Interlocking.

Given the significant potential for additional injuries, the patient initially undergoes a thorough radiographic examination of both limbs. Assuming that only one of the femurs is fractured, the X-ray images of the uninjured femur are then used to obtain an estimate of the size (i.e. length and width) of the required intramedullary nail. As a consequence, it can be ensured that an appropriate range of nail sizes, some larger and some smaller than the estimated size, are available in the operating room at the time of surgery. In order to simplify the intraoperative fracture reduction process, and minimise pain at the fracture site, normal leg length is then restored by applying preoperative skeletal traction.

In line with most closed osteosynthesis procedures, intramedullary nailing is heavily reliant upon intraoperative fluoroscopy. As a consequence, the patient must be positioned on the operating table in a "C-arm compatible" pose. In other words, the operating table and the patient's uninjured limb must not obstruct movement of the mobile C-arm unit, and must not appear within the field-of-view when fluoroscopic images (AP and lateral) of the fractured femur are acquired. The need to apply skeletal traction intraoperatively, to allow reduction of the fracture, also creates

problems in relation to patient positioning when using a conventional operating table. Ideally, an orthopaedic fracture table should therefore be used during closed intramedullary nailing procedures. These modified operating tables incorporate adjustable extensions for securing the patient's legs, and a padded perineal post to prevent movement of the pelvis when traction is applied. The bottom one-third of the table is also significantly shortened to allow direct fluoroscopic visualisation of the fracture site.

The nailing procedure can actually be performed with the patient in either the *lateral decubitus* or the *supine* position. As depicted by Figure 6.6, the lateral decubitus position involves lying the patient on his/her uninjured side, thereby affording easy access to the greater trochanter of the fractured femur (i.e. the nail's insertion point). In practice, obtaining this position can be difficult, and special table attachments may be required to safely support the patient's weight. By contrast, the supine position (depicted in Figure 6.7) merely involves lying the patient on his/her back, and is therefore more suitable for patient's with multiple injuries. However, the supine position affords limited access to the trochanter of the fractured femur, making insertion of the intramedullary nail more difficult.

Once the patient is in position, skeletal traction is reapplied (via a femoral or tibial pin), and a closed reduction of the fracture is attempted. This fluoroscopically monitored process involves the manual application of localised pressure to both sides the fracture site, and can require assistance from a second surgeon. If a satisfactory reduction can be obtained, the usual skin preparation and patient draping processes are performed, and a nail insertion is then attempted. Access to the greater trochanter is gained by making a lateral skin incision (4 to 8 cm long), and retracting the relevant muscle groups. An opening into the medullary cavity is then created using a manual surgical instrument (e.g. a sharp awl, a centering pin and cannulated cutter, etc.). Prior to making the actual opening, the position of the cutting tool is checked fluoroscopically to ensure that it is centred over the medullary cavity in both the AP and lateral views.

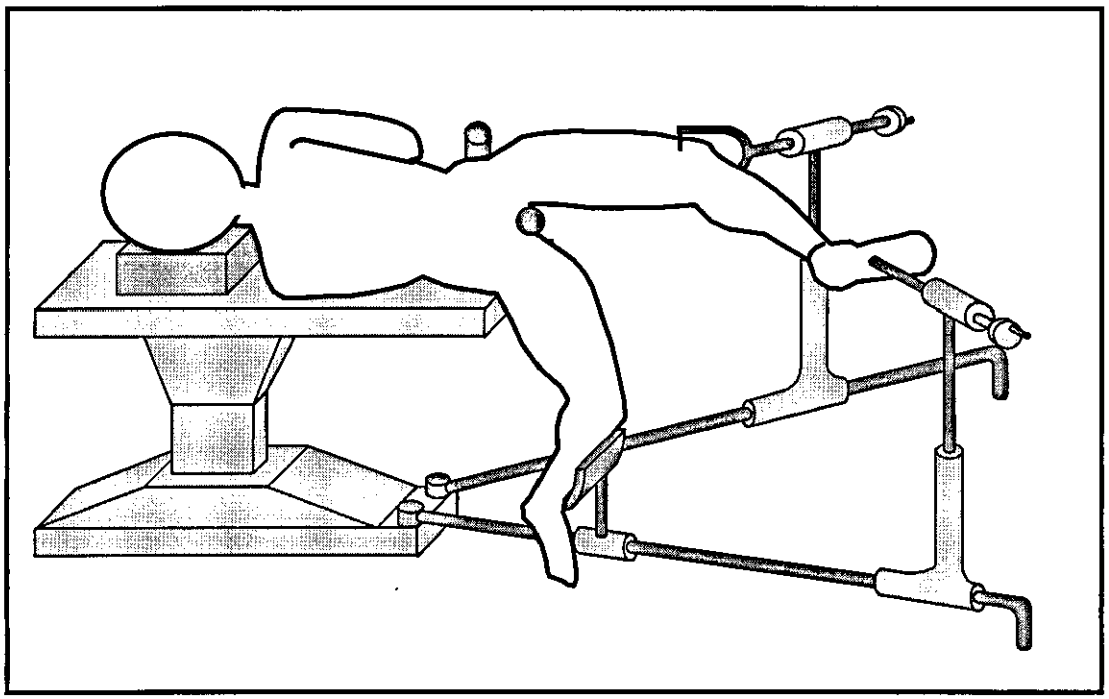


Fig. 6.6 : Patient in Lateral Decubitus Position on Fracture Table.
(Adapted from [204])

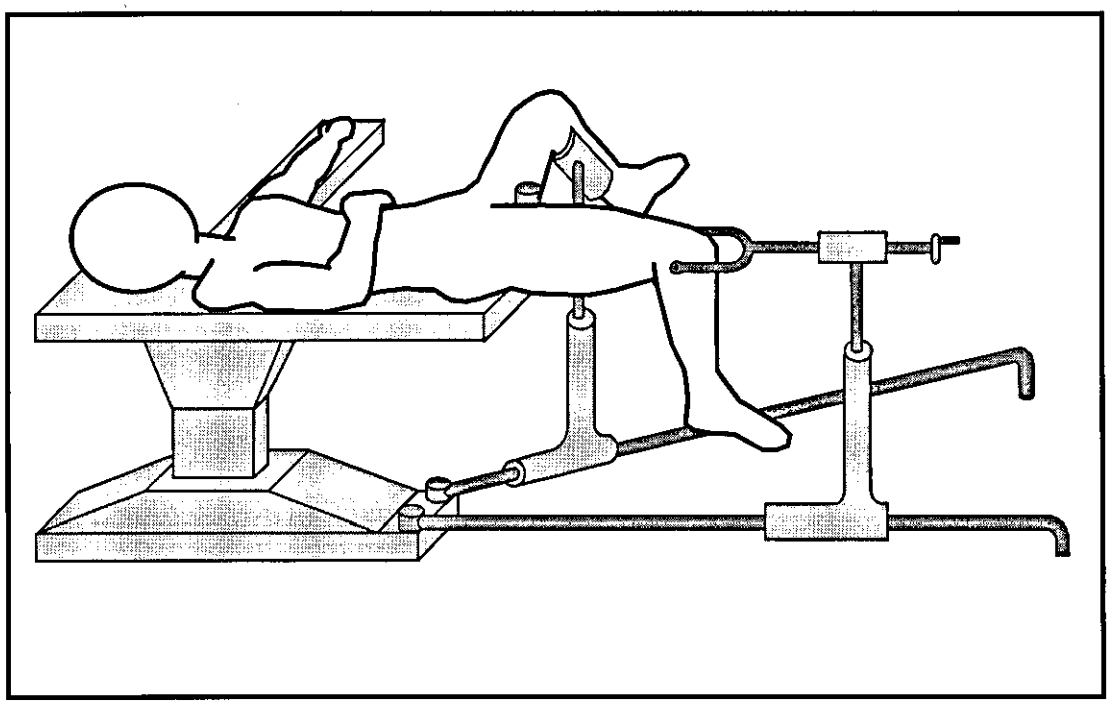


Fig. 6.7 : Patient in Supine Position on Fracture Table.
(Adapted from [204])

Reaming of the medullary cavity initially requires the insertion of a bulb-tipped guide wire/reaming rod. As this guide wire must pass across the fracture site and into the distal femoral bone fragment, it is essential that reduction of the fracture is maintained throughout its insertion process. Having used a mallet to lightly tap the guide wire into the distal fragment, the required intramedullary nail length is confirmed by measuring the length of guide wire remaining outside the femur. Reaming of the medullary cavity is then performed by gently driving cannulated flexible reamers over the guide wire. Typically this process will begin with an eight or nine millimetre reamer, and will continue in one millimetre increments until contact is made with cortical bone. Reaming will then continue in 0.5 mm increments until the desired cavity size has been prepared. Throughout the entire reaming process, fluoroscopy is used to check for eccentric reaming and further comminution of the fracture site.

Prior to insertion of the pre-bent intramedullary nail, the (reaming) guide wire is replaced by a larger diameter nail insertion guide. The nail insertion instrumentation is then assembled, and is attached to the proximal end of the intramedullary nail, which usually incorporates an internal screw thread. These instruments include an insertion handle to control nail alignment, which also doubles as a drill guide for the proximal interlocking holes. Following manual introduction of the nail into the proximal medullary cavity, a ram mounted on a linear ram guide is then used to gently drive the nail over the guide rod. Once again, it is essential that accurate reduction is maintained as the nail passes across the fracture site. Fluoroscopic monitoring of this stage of the insertion process is therefore mandatory.

Depending upon the stability and type of fracture present, either static or dynamic interlocking may then be required to control rotation of the bone fragments. In cases requiring static or proximal dynamic locking, a single screw will be inserted through the proximal end of the nail with the aid of a drill guide incorporated into the insertion handle. This "open" interlocking process takes place via the original (insertion) skin incision. Similarly, in cases requiring static or distal dynamic locking, two transfixion screws will be percutaneously inserted through the distal end of the nail. As further discussed in Section 6.4.1, a wide variety of surgical techniques are currently applied

to perform this fluoroscopically-guided distal interlocking process, all of which leave considerable room for improvement.

During the first or second postoperative week, the patient is encouraged to start walking with the aid of crutches (i.e. walking without weight-bearing). In the following weeks, a programme of physiotherapy exercises is undertaken to strengthen the quadriceps (thigh muscles) and regain the full range of knee joint motions. Following radiographic confirmation of the formation of a satisfactory callus, progressive weight-bearing can then begin. Assuming that full weight-bearing is achieved, whenever possible, a second surgical procedure to remove the interlocking screw(s) and the intramedullary nail, will then be performed during the second postoperative year.

6.4.1 The Distal Interlocking Problem.

The insertion of a distal transfixion screw requires the implementation of a drilling trajectory that passes through the lateral femoral cortex, one of the intramedullary nail's distal screw-holes, and finally the medial femoral cortex. Since the intramedullary nail is obviously concealed within the femur, this "blind" drilling process can only be performed under fluoroscopic guidance. Unfortunately, the distal interlocking instruments that are currently supplied with intramedullary nails are difficult to use, and can also be inaccurate. As a consequence, distal interlocking is often the most technically demanding and time-consuming stage of the entire intramedullary nailing procedure. Dosimetric studies, discussed at length in Appendix 1, have also shown that inexperienced orthopaedic surgeons can be exposed to excessive levels of X-ray radiation while performing distal interlocking. In recent years, a wide variety of techniques have therefore been proposed with the aim of simplifying the distal interlocking process, and making it safer in relation to irradiation of the surgeon.

All of the existing distal interlocking techniques involve the following stages, with the main variable being the choice of distal targeting device:

- Acquisition of the *perfect circles* fluoroscopic image.
- Fluoroscopic localisation of the incision site.
- Distal targeting to align the drill bit with the axis of the screw-hole.
- Drilling through both cortices and the screw-hole in the nail.
- Measurement of the required transfixion screw length.
- Insertion of the transfixion screw.
- Fluoroscopic verification of the insertion.

In order to perform the drilling process under fluoroscopic guidance, the X-ray axis of the mobile C-arm unit must first be brought into alignment with the drilling trajectory. A preliminary coarse alignment of the C-arm unit with the patient's thigh, produces a lateral fluoroscopic image in which the distal screw-hole has an elliptical profile. The shape of this profile is explained by the fact that the distal screw-hole is essentially a cylinder, whose ends are defined by a pair of corresponding holes in the lateral and medial sides of the hollow intramedullary nail. When viewed obliquely, the resulting projection distortion means that both of these holes produce oval-shaped profiles, which are then superimposed to form the elliptical profile of the screw-hole. However, as illustrated by Figure 6.8, when the C-arm unit is aligned with the axis of the screw-hole (i.e. the drilling trajectory), the profiles of the two holes appear as concentric circles at the centre of the image. Consequently, alignment with the drilling trajectory can be obtained by adjusting the C-arm's position until this so called "perfect circles" image is acquired. In practice, this four degrees-of-freedom alignment process can prove to be very difficult and time-consuming.

Having obtained the perfect circles image, the location of a suitable incision site above one of the screw-holes is determined by introducing a surgical instrument (e.g. scalpel, forceps, awl, etc.) into the field-of-view of the C-arm unit **[207]**. This tool is then centred over the screw-hole in the fluoroscopic image, thereby allowing the incision site to be marked. The process is then repeated for the other distal screw-hole. A longitudinal incision is then made down to the bone, over each of the distal screw-holes.

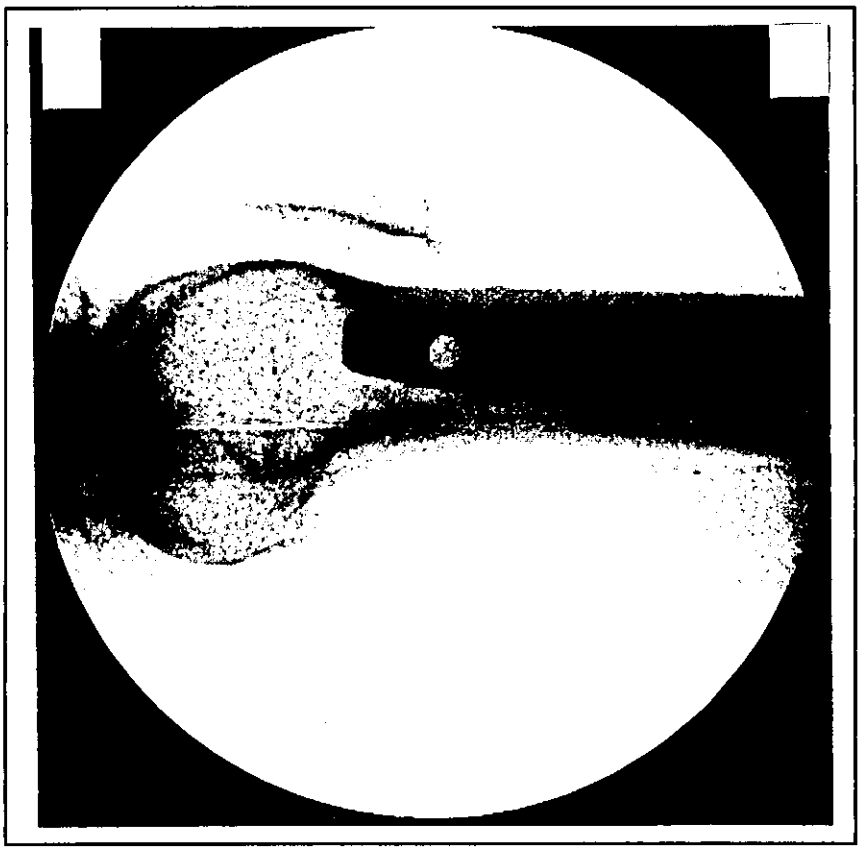


Fig. 6.8 : The "Perfect Circles" Fluoroscopic Image.

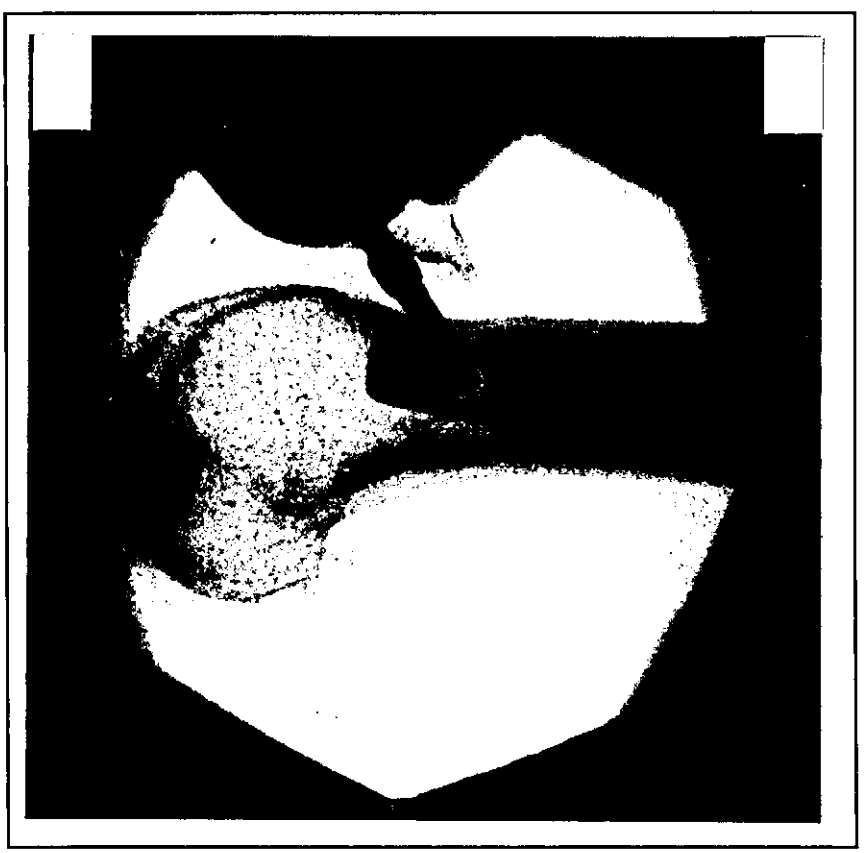


Fig. 6.15 : Direct Exposure of the Surgeon's Hand to X-ray Radiation.

Immediately after making the incision(s), distal targeting is performed using either the nail manufacturer's recommended technique, or an adaptation of the *freehand* technique. The distal interlocking instrumentation supplied with the intramedullary nail tends to fall into one of three categories:

- Nail-mounted targeting devices.
- Hand-held targeting devices.
- C-arm mounted targeting devices

Nail-mounted targeting devices originally took the form shown in Figure 6.9(A), which depicts the instrumentation supplied with the Grosse-Kempf Interlocking Nail (Howmedica, Rutherford, NJ). This mechanical jig system was initially assembled prior to insertion of the intramedullary nail. The assembly was then screwed into the proximal end of the nail, and adjustments were made to bring the holes in the drill guide block into alignment with the distal screw-holes in the nail. The assembly was then removed, and reapplied following insertion of the nail. Theoretically, the drill guide block could then be used to assist drilling of the holes for the distal transfixion screws. Unfortunately, in reality, a slotted intramedullary nail can deform significantly during the insertion process, resulting in the holes in the drill guide block no longer being aligned with the distal screw-holes in the nail. As a consequence, the use of "rigid" nail-mounted targeting devices has generally been abandoned **[198]**.

In order to overcome these problems, the Russell-Taylor Interlocking Nail (Richards Medical Company, Memphis, TN) is supplied with the adjustable nail-mounted targeting device depicted in Figure 6.9(B) **[209]**. This instrumentation is assembled intraoperatively, with the proximal drill guide initially being attached to aid insertion of the intramedullary nail. If required, the drill guide incorporated into the handle of this assembly can then be used to facilitate proximal interlocking. Following acquisition of the perfect circles image, the shaft of the distal targeting device is then inserted through the adaptor block, until a position which corresponds to the individual nail's length is obtained. The shaft is then locked into position (parallel to the nail) by tightening a grub-screw in the adaptor block. The distal targeting device

can then be aligned with the more proximal of the two distal screw-holes, by using the adjustment instrument to make several fluoroscopically-guided fine adjustments.

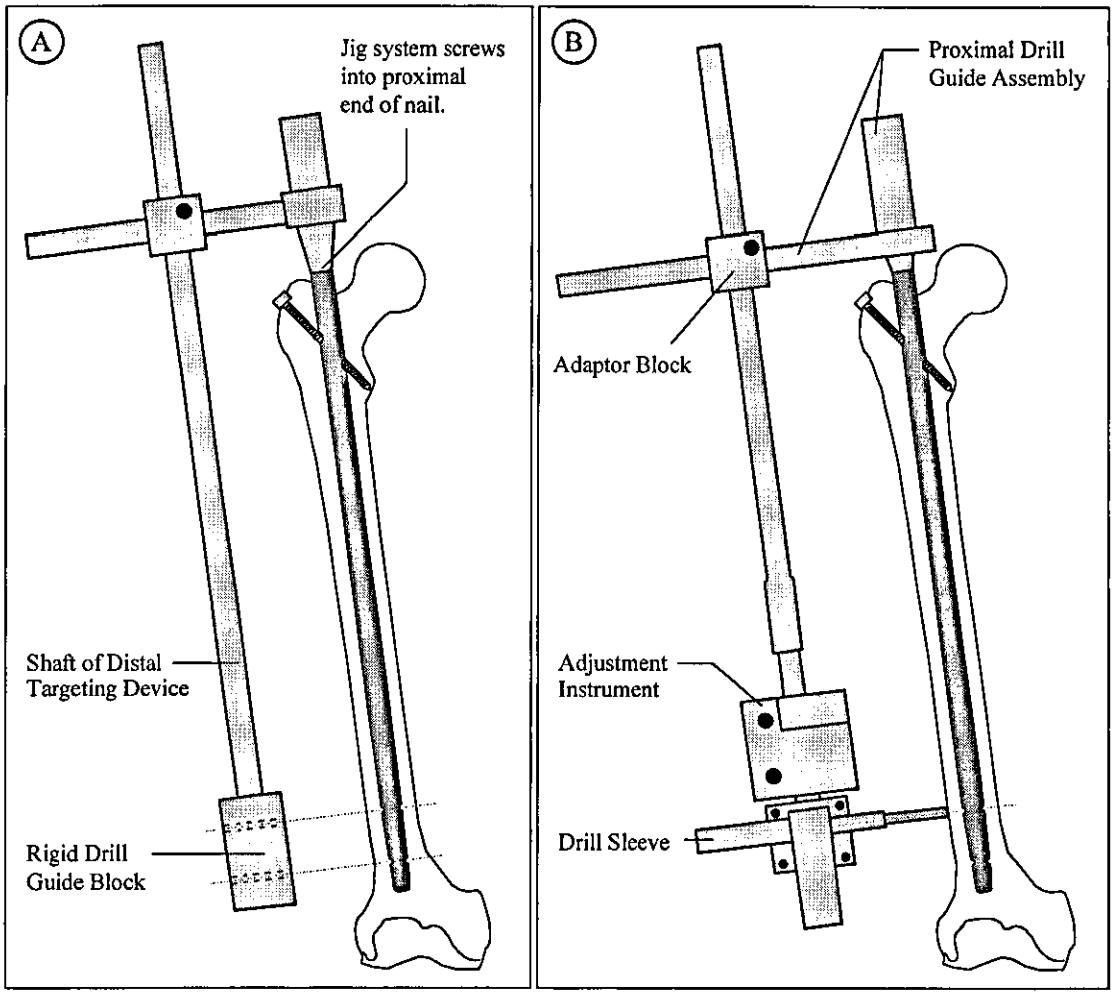


Fig. 6.9 : Nail-Mounted Distal Targeting Devices.
(Adapted from [207])

One method of avoiding the prolonged set-up times associated with nail-mounted targeting devices is to provide the surgeon with a hand-held aiming device. A prime example of this type of device is currently supplied with the AO/ASIF Universal Femoral Nail **[204]**. As depicted in Figure 6.10, the AO distal aiming device is a three-piece instrument, consisting of a handle, a detachable direction finder, and a variety of insert designs. Following acquisition of the perfect circles image, and creation of the surgical incision(s), distal interlocking is achieved using the techniques summarised in Figure 6.11.

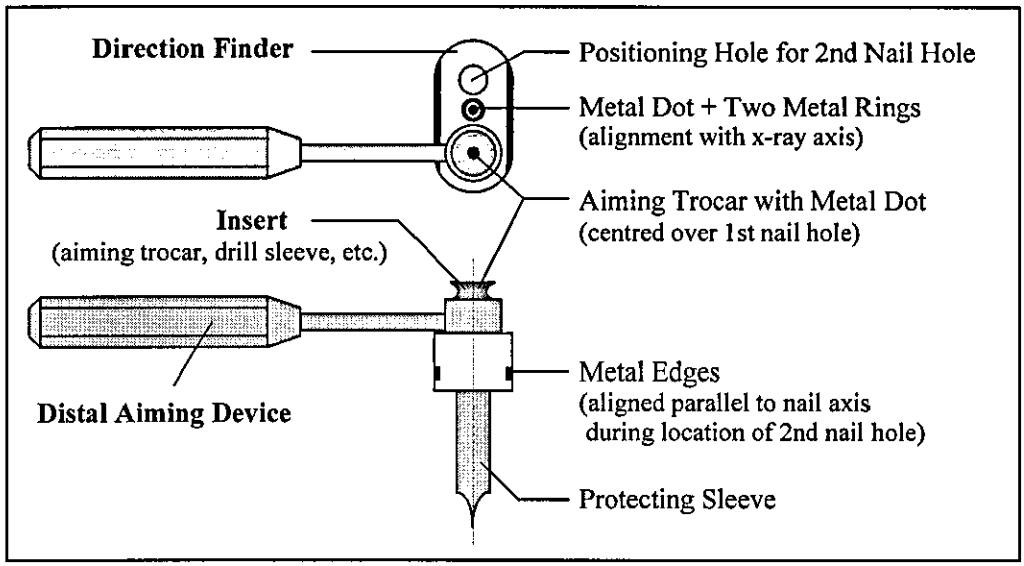


Fig. 6.10 : The AO Distal Aiming Device.
(Adapted from [204])

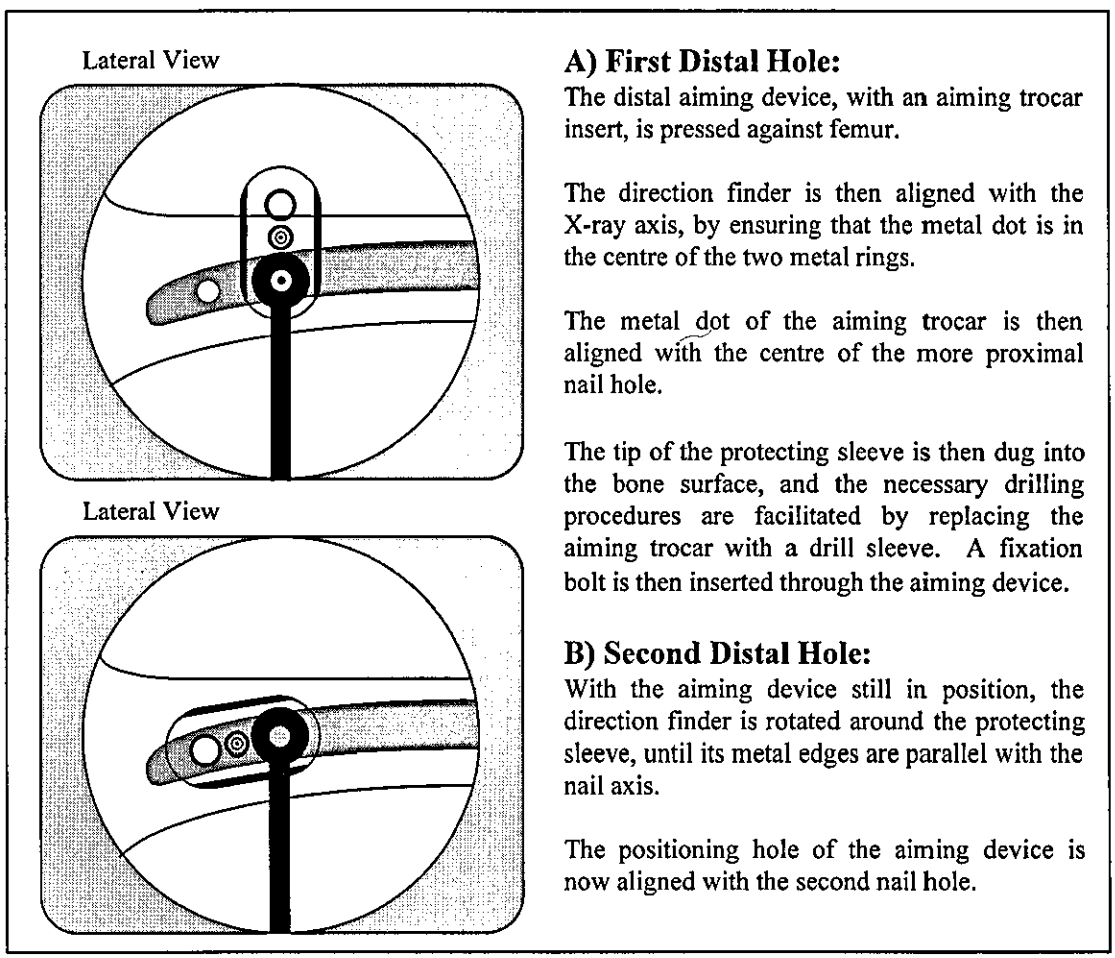


Fig. 6.11 : Distal Interlocking Technique Using The AO Distal Aiming Device.
(Adapted from [204])

The obvious disadvantage of a hand-held aiming device is that it leaves the surgeon with only one free hand. A second surgeon/assistant may therefore be required to hold the aiming device in position, while the primary surgeon performs the distal interlocking process. Given that distal targeting is fluoroscopically-guided, the use of a hand-held aiming device also creates a potential radiation safety problem. Although the handle of the aiming device should be long enough to prevent direct exposure of the surgeon's hand to the primary X-ray beam, significant levels of scattered radiation will undoubtedly be received due to the proximity of the surgeon's hand to the patient **[210]**.

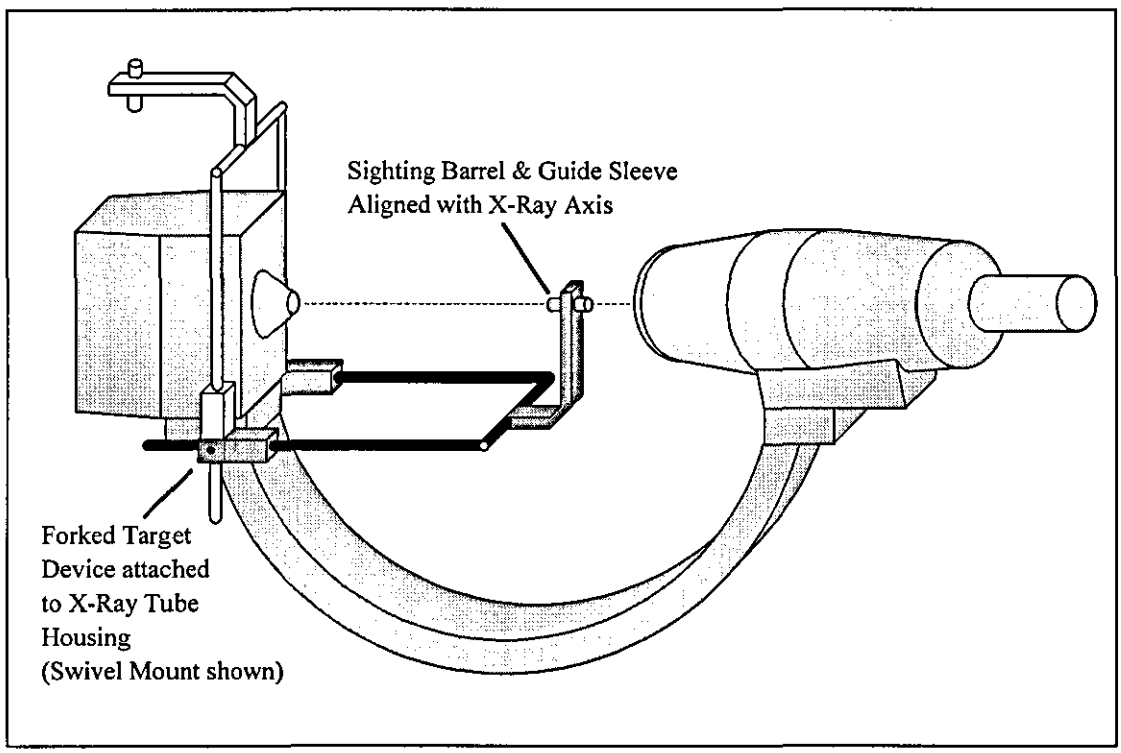


Fig. 6.12 : The Grosse-Kempf Fluoroscope-Mounted Targeting Device.
(Adapted from **[198 & 206]**)

C-arm mounted targeting devices were introduced as a means of reducing the radiation exposure levels received by orthopaedic surgeons. The most noteworthy example is supplied with the Grosse-Kempf Interlocking Nail (Howmedica, Rutherford, NJ), and takes the form shown in Figure 6.12. This sterile targeting device is attached to the X-ray tube housing of the C-arm unit immediately prior to the acquisition of the perfect circles image **[206]**. The surgical incision is then made, and the drill guide incorporated into the targeting device is brought into alignment with one of the distal

screw-holes. In practice, this is made possible by mounting the targeting device on a hinged bracket, in such a way that the drill guide is aligned with the X-ray axis when the targeting device is lowered into a horizontal position. A linear slide then allows the drill guide to be brought into contact with the lateral femoral cortex. Unfortunately, the long lever arm of the targeting device, coupled with the lack of rigidity of the C-arm gantry, results in the drill guide deviating from the desired trajectory when manual pressure is applied during the drilling process. As acknowledged by Hindley *et al.* **[211]**, perfection of the technique therefore requires "some practice".

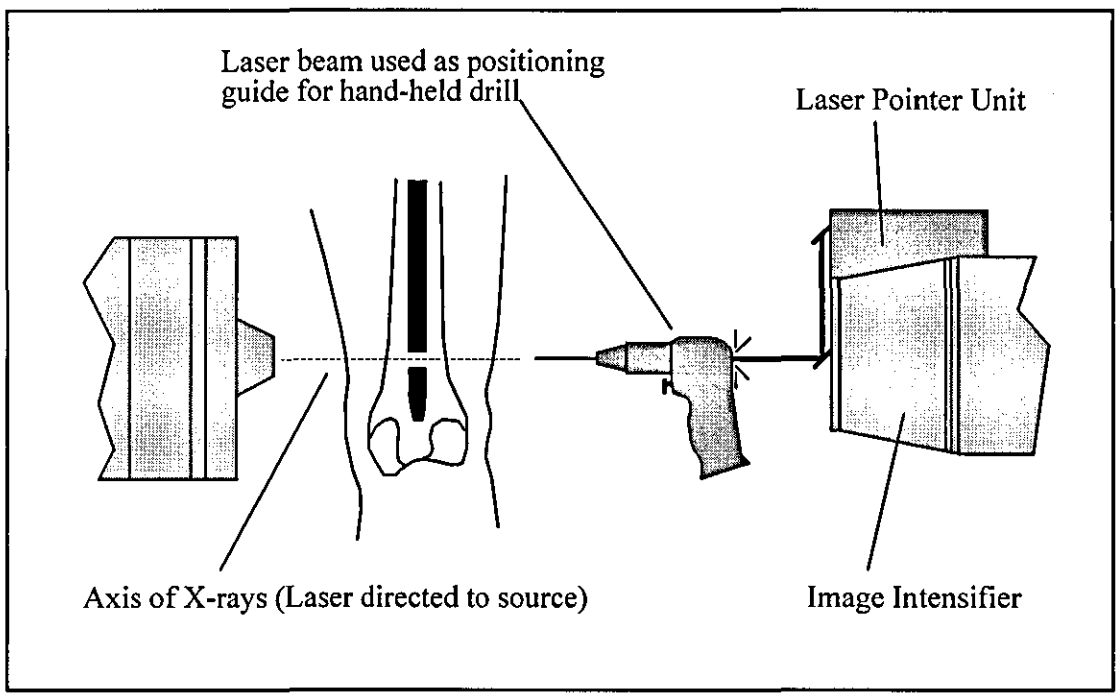


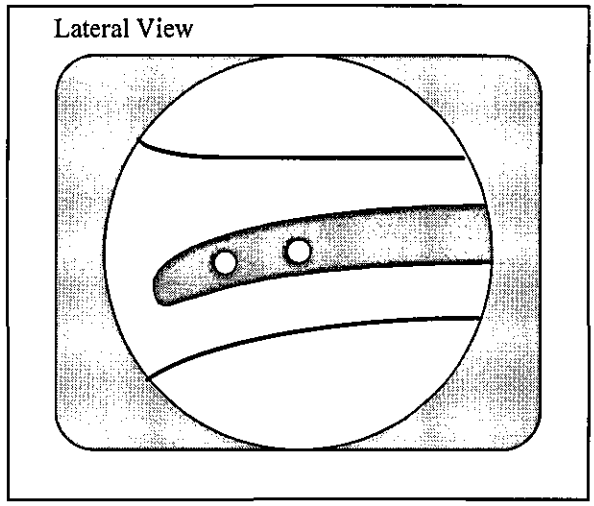
Fig. 6.13 : The Goodall Laser Guidance System for Distal Interlocking.
(Adapted from [212])

A more recent variation upon the C-arm mounted targeting device is the use of laser guidance systems. Goodall *et al.* **[212]** have developed a system which uses a visible laser beam aligned with the X-ray axis of the C-arm unit. As depicted by Figure 6.13, when the perfect circles image is acquired, this image intensifier-mounted laser system indicates not only the incision site, but also the drilling trajectory. Drilling of the distal hole can therefore be performed without the use of fluoroscopic guidance, leading to significant reductions in surgeon irradiation. Goulet *et al.* **[213]** have also developed a system which uses intersecting perpendicular planes of laser light to

indicate the X-ray axis. The laser diodes used by this system are mounted on the X-ray tube housing. In order to aid acquisition of the perfect circles image, Goulet *et al.* also use a perpendicular pair of wires attached to the front of the image intensifier housing. These wires are superimposed into the fluoroscopic image as cross-hairs, allowing an alignment technique similar to the one described by Saw **[214]** to be adopted. Finally, the latest models of the Siremobil® 2000 (Siemens Medical Engineering Group, Erlangen, Germany) mobile C-arm unit are also known to incorporate a virtually identical laser guidance system to the one described by Goulet *et al.* **[213]**.

However, when using a laser guidance system, care must be taken to avoid direct exposure of the eyes to the laser beam(s). Modification of the surgical drill may also be necessary to avoid any unwanted reflections of the laser beam(s). In addition, alignment of the surgical drill with the drilling trajectory, as indicated by the laser beam(s), is a subjective process. Given that no drill guide is being used, owing to the lack of physical constraint, there is also a tendency for the drill bit to slip on the surface of the femur when drilling commences.

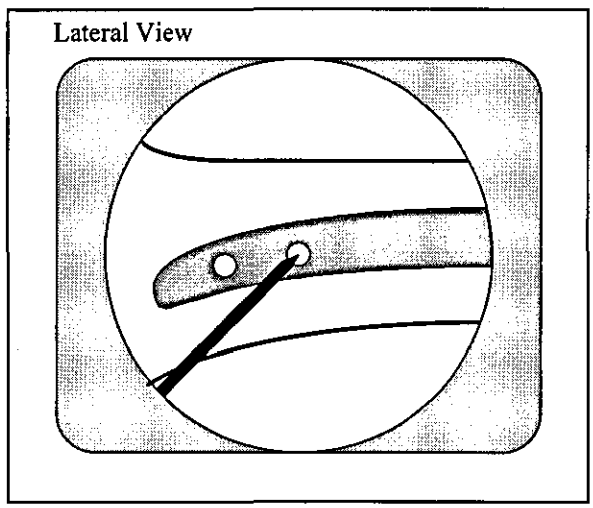
The freehand aiming technique was originally developed to allow completion of the distal interlocking process, when an initial attempt using the instrumentation supplied with the intramedullary nail had failed. However, because of its relative simplicity, the freehand method has since become the most commonly used distal targeting technique. This simplicity stems from the use of a surgical instrument as the aiming device. In the majority of cases, this "targeting tool" will be some form of drill bit or guide wire, which is inserted into the chuck of an air-powered surgical drill **[215 - 217]**. Alternatively, a hand-held surgical awl may also be used to perform distal targeting **[198]**. Following acquisition of the (lateral) perfect circles image and creation of the surgical incision, the technique therefore involves the stages summarised by Figure 6.14.



A) Perfect Circles Image:

The C-arm's position is adjusted until the two distal screw-holes appear as near-perfect circles.

A surgical tool (awl, K-wire, drill bit or guide wire) is then held within the image intensifier's field-of-view, with its tip in contact with the lateral femoral cortex.

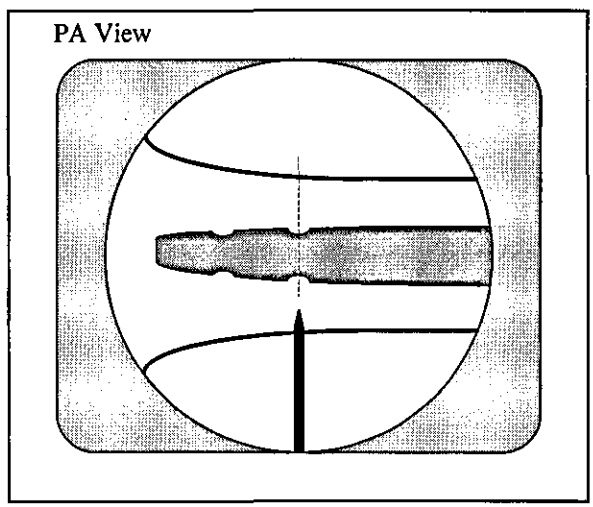


B) Entry Point Determined:

Under image intensifier control, the tip of the surgical tool is centred over one of the distal screw-holes.

During this procedure the tool must be held at approximately forty-five degrees to the femur, in order to allow visualisation.

A small starting hole is then made in the femur at this position.



C) Alignment Determined:

The surgical tool is then held perpendicular to the femur without moving its tip, and its orientation is checked on an PA view.

Having checked the drilling trajectory on both views, the screw-hole can then be drilled.

The entire process is then repeated for the second screw-hole.

Fig. 6.14 : The Freehand Aiming Method for Distal Interlocking.

The required entry point on the lateral femoral cortex is identified, under fluoroscopic guidance, by centering the tip of the targeting tool over one of the distal screw-holes. This entry point is then marked by drilling a few millimetres into the lateral femoral cortex. In order to fluoroscopically guide alignment of the targeting tool with the axis of the screw-hole, the C-arm unit must then be rotated through ninety degrees to acquire an PA view of the distal femur. The targeting tool is then adjusted until it is perpendicular to the distal screw-hole, and the required distal hole is drilled.

Figure 6.15 highlights the most obvious disadvantage of the freehand technique. The dark mass at the top of this fluoroscopic image is actually the surgeon's hand, and as such, the surgeon has been directly exposed to the primary X-ray beam. As noted by Hudson *et al.* **[218]**, this type of direct exposure is not uncommon when using the freehand technique. Ideally, lead-impregnated gloves should be worn by surgeons performing the freehand distal targeting process. However, these gloves are unpopular because they limit manual dexterity. A compromise solution involving the use of a long-handled tool holder has therefore been proposed by a number of surgeons **[219 - 221]**.

As illustrated by Figure 6.16, this type of device ensures that the surgeon's hands are removed from the primary X-ray beam. It has also been claimed that further reductions in surgeon irradiation can be obtained by making use of the known distance between the two distal screw-holes. Having drilled the first distal hole using the freehand technique, rather than repeat the process again, the second hole is therefore drilled with the aid of an appropriate targeting jig/guide **[222, 223]**. However, these modified or "safer" versions of the freehand technique still incur the problems associated with drilling without the aid of a drill guide.

Assuming that distal targeting has successfully been achieved using one of the four techniques discussed above, problems can still be encountered during the drilling and screw insertion procedures. If the first attempt to drill one of the distal holes fails, subsequent attempts are hindered by the previously drilled hole(s). Repeated attempts to drill a hole can also lead to weakening of the bone, or poor purchase by the

transfixion screw. In view of these potential problems, a pilot hole is generally created using a drill bit or guide wire, whose diameter is smaller than the required distal transfixion hole. A larger drill bit, or even a cannulated drill bit, is then used to increase the diameter of the correctly drilled pilot hole. Finally, as noted by Browner *et al.* **[205]**, successful drilling of the distal holes does not necessarily guarantee satisfactory insertion of the self tapping transfixion screws. The location of the transfixion screw must therefore be confirmed by acquiring both a lateral and a PA (as depicted in Figure 6.17) view.

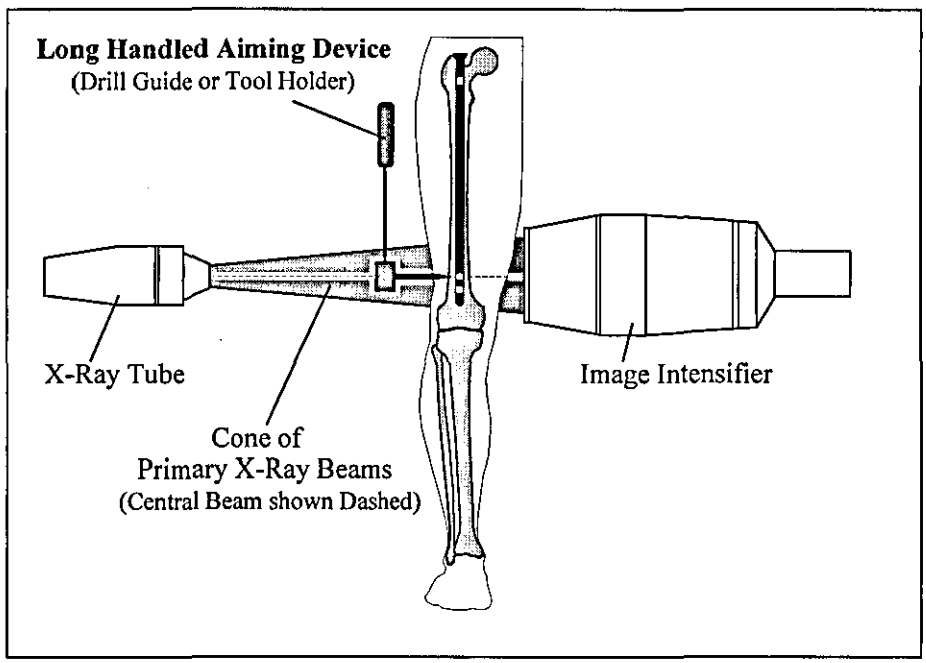


Fig. 6.16 : Removal of Surgeon's Hands from the Primary X-ray Beam
(Adapted from Skjeldal *et al.* **[219]**)

In summary, existing surgical techniques to perform the distal interlocking of femoral intramedullary nails are far from ideal. A time-consuming fluoroscopically-guided alignment process must initially be undertaken, in order to acquire the "perfect circles" image. The technically demanding distal targeting process is then performed using a technique which is either inaccurate, time-consuming, or exposes the surgeon to unacceptably high levels of radiation. Finally, drilling of the distal holes must then be performed, frequently without the aid of a drill guide, using a hand-held surgical drill. As a consequence, optimisation of the distal interlocking process remains an active and potentially lucrative topic of research, from both the orthopaedic and engineering perspectives.

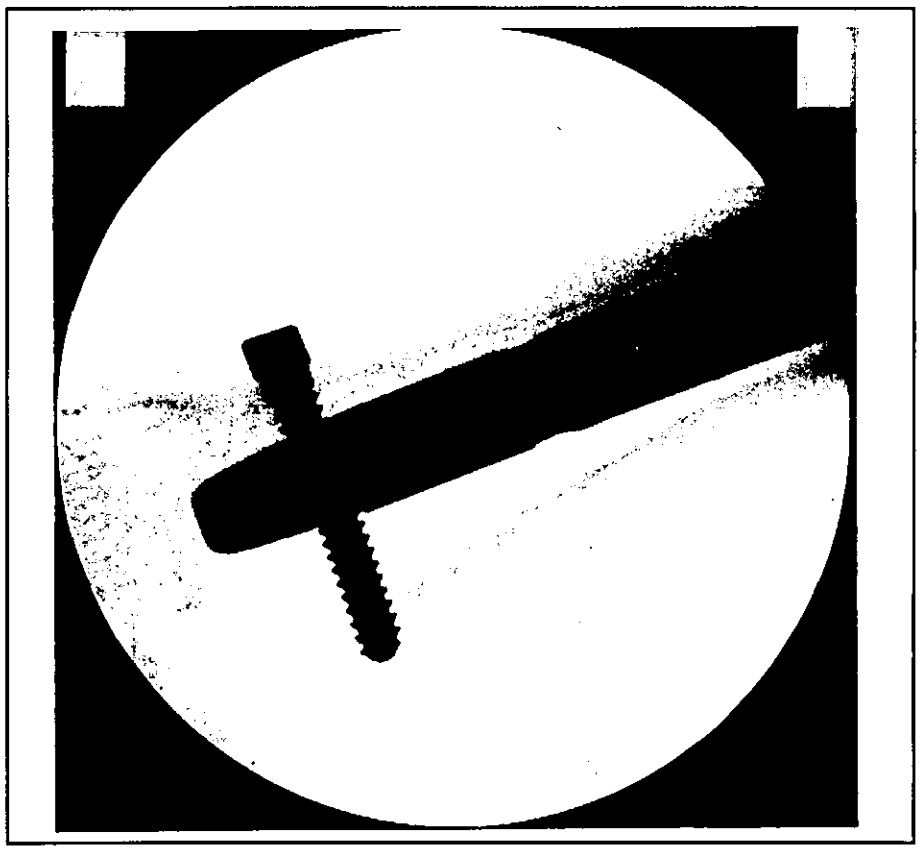


Fig. 6.17 : Fluoroscopic Verification of Transfixion Screw Insertion.

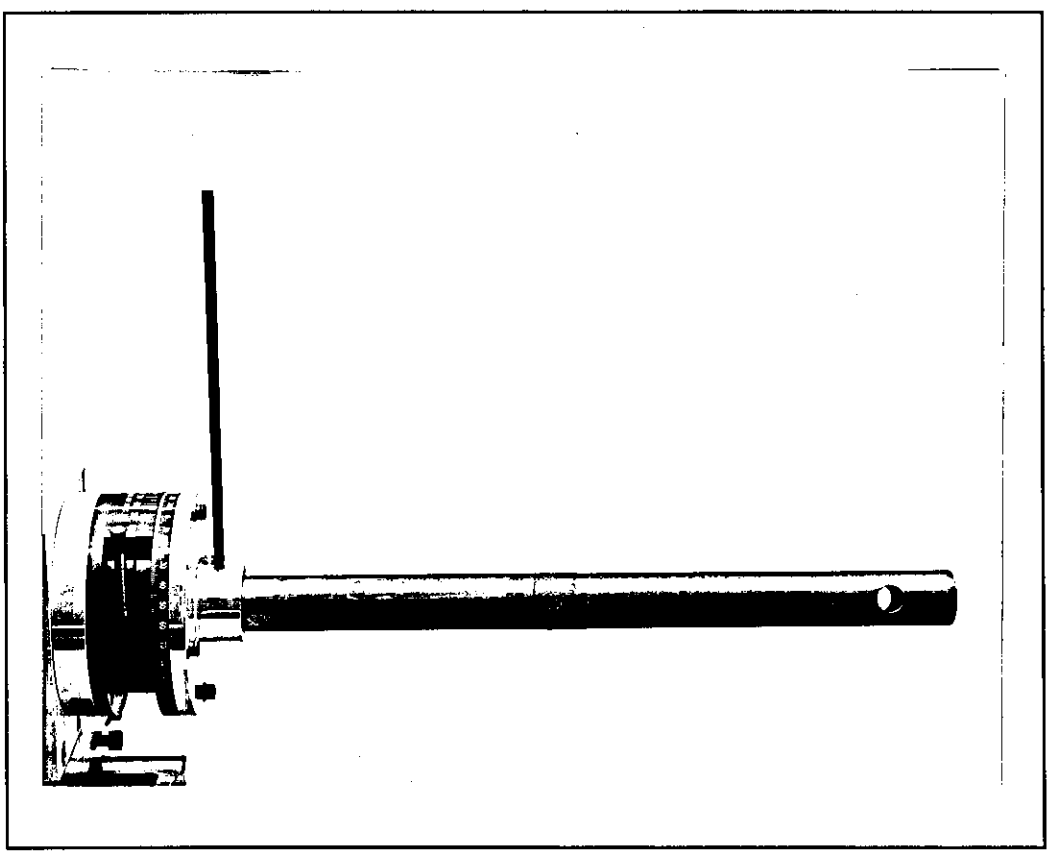


Fig. 6.22 : Test Rig Used To Simulate Nail Rotation.

6.5 Robotic-Assisted Distal Interlocking.

As explained in Appendix 2, the MEDROSA research group is currently investigating the intraoperative use of an X-ray image-guided robotic manipulator, with a view to improving both the accuracy/repeatability and the safety of the distal interlocking process. In relation to the current thesis, the important aspects of this robotic-assisted surgery application obviously relate to the intraoperative registration and trajectory planning processes.

Under the proposed surgical protocol, the "robotic" stages of the nailing procedure are instigated following a successful insertion of the nail into the medullary cavity of the femur. In-line with the current thesis, registration between the patient (femur) and the robotic manipulator will then be established using a robot-mounted X-ray calibration frame, which complies with the scenario depicted in Figure 5.2. Given that this frame is introduced at an intermediate stage of the surgical procedure, in order to prevent contamination of the surgical site, it must be either sterilisable or capable of being isolated from the patient.

Ergonomic considerations dictate that the calibration frame must have a two-piece configuration, which can be assembled around the patient's lower thigh. As such, the use of sterile custom-made plastic "packets", to isolate the individual frame components, is one potential solution to the sterility problem. Alternatively, an autoclave-compatible calibration frame could also be developed. The choice of materials from which such a frame can be manufactured is severely limited by the harsh conditions inside an autoclave; effective steam sterilisation requires items to be exposed to saturated steam in the range 121-134 °C for up to fifteen minutes. Nevertheless, by autoclaving test specimens, consisting of corrosion resistant (surgical stainless steel) ball bearings inserted into Tufnol® (phenolic fabric) sheets with an interference fit, it has been possible to successfully demonstrate the "autoclaveable" calibration plate concept at Loughborough University.

Having obtained registration using a "sterile" calibration frame, a pair of drilling trajectories, which pass through the distal holes in the intramedullary nail, will be

reconstructed from the two calibrated intraoperative images (i.e. the PA and lateral views). As further discussed in Sections 6.5.1 and 6.5.2, this surgical planning process will be facilitated by providing the surgeon with a graphical user interface. A combination of interactive indication processes and automatic image analysis routines will then be used to reconstruct the desired drilling trajectories, in relation to the coordinate system of the calibration frame. This drilling data will then be downloaded to the robot controller, where it will be transformed into the operating room (manipulator) coordinate system.

At this stage of the procedure, the calibration frame will be detached from the robotic manipulator, and removed from the vicinity of the patient. The manipulator will then be used to accurately align a drill guide, or alternatively an actuated drilling unit, with the reconstructed drilling trajectories. The invasive (bone drilling) stages of the distal interlocking process can then be performed either manually, with the aid of the drill guide, or "robotically", using the actuated drilling unit. Upon successful drilling of the second distal hole, the robotic-assisted stage of the intervention is complete, and the manipulator can be withdrawn, thereby increasing the workspace around the patient. The remaining stages of the distal interlocking process (i.e. increasing the diameter of the screw holes, screw insertion, etc.) will then be performed in the conventional manner.

6.5.1. "Perfect Circles" Based Approach.

In order to quantify the existing distal interlocking process, it was initially proposed that the trajectory planning scenario depicted in Figure 6.18 would be implemented. This scenario assumes that the (lateral view) perfect circles image has already been acquired, prior to positioning the calibration frame around the patient's lower thigh. As such, the axis of one of the distal interlocking holes is "perfectly" aligned with the X-ray axis of the C-arm unit. The centre of the imaged hole's projection, which can either be interactively indicated or automatically measured, can therefore be used as both an "entry" and a "target" point. In other words, the line-of-sight which passes through the centre of the projected hole, also passes through the centres of the actual holes in both the lateral and medial sides of the nail. By reconstructing this lateral

view line-of-sight, and projecting it into the PA view as a graphical overlay, the image matching problem can therefore be solved in the manner indicated in Figure 6.18.

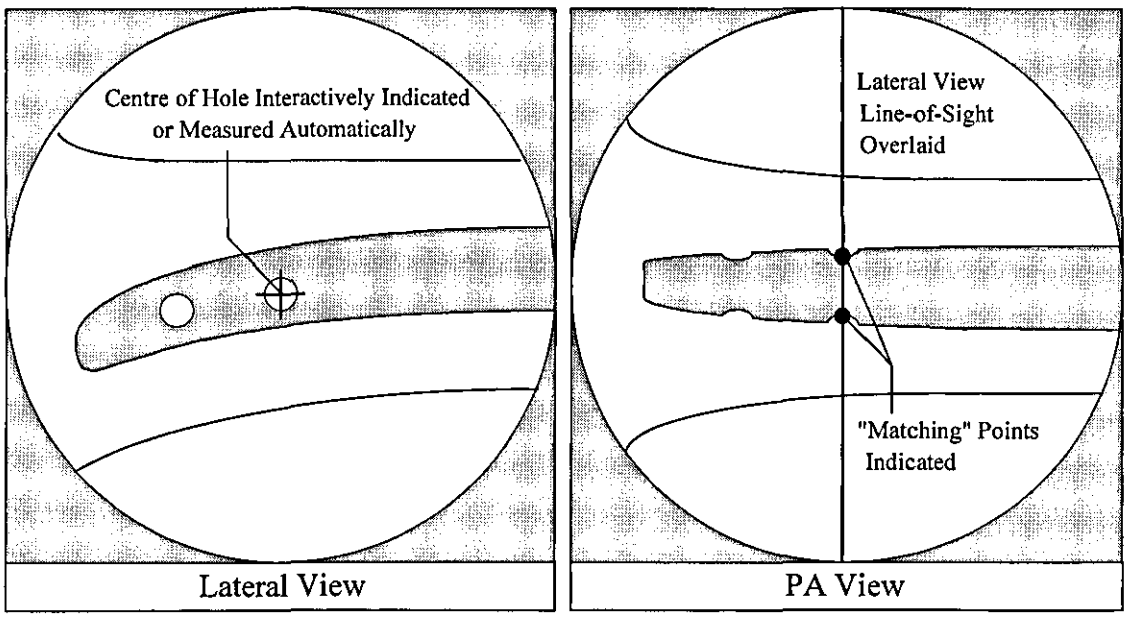


Fig. 6.18: The Trajectory Planning Scenario (Assuming Perfect Circles).

In relation to the Hip Box study discussed in Chapter 5, the only major problem raised by this proposed scenario is the presence of the intramedullary nail within the field-of-view. As discussed in Chapter 2, during X-ray photogrammetry studies involving surgical (internal fixation/prosthetic) implants, it is inevitable that a number of fiducials will be obscured by the metallic component(s). Assuming that the perfect circles approach is to be adopted, in the case of distal interlocking, the metallic component is located at the centre of the lateral field-of-view. Unfortunately, this is the region of the Hip Box fiducial patterns that contains the "critical" fiducial markers (i.e. the centre marker, orientation marker and four washers). Given that scanning of the relevant calibration images cannot proceed if these fiducials are not detected, an alternative set of fiducial patterns was therefore required.

As illustrated by Figure 6.19, these new fiducial patterns overcome the metallic implant problem by moving the critical fiducials away from the centre of the image. Ideally, the centre and orientation markers will therefore appear within the field-of-view on opposite sides of the intramedullary nail. However, in making these changes, the image scanning process is obviously significantly altered. As a consequence, the

relevant image analysis program (*K_Correl.pas*) differs from its Hip Box counterpart (*HB_Corel.pas*) in a number of respects. Most notably, since the lateral view patterns have been rotated by 90°, the PA and lateral scanning processes are no longer identical.

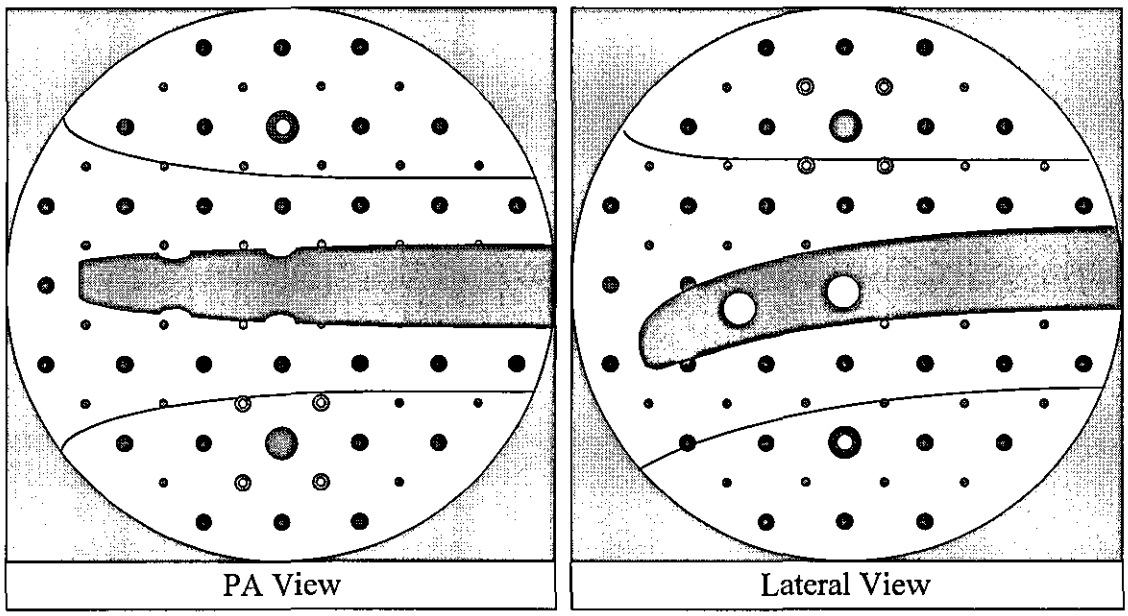


Fig. 6.19 : Ideal Calibration Images For Distal Interlocking.

For the purposes of a laboratory-based feasibility study, Perspex calibration plates incorporating these modified fiducial patterns, were manufactured in accordance with the dimensions of the existing robot-mounted calibration frame (previously shown in Figure 4.13). When assembled in this configuration, owing to its intended role in relation to the patient's anatomy, the resulting X-ray calibration frame is referred to as the "Knee Box".

By acquiring images of this frame and a femoral phantom/nail combination, it was then possible to develop the necessary photogrammetric reconstruction software. The resulting program (*KLocProg.pas*) and its linked units (*KLoc_Foc.pas*, *KLocSurg.pas* & *KLocType.pas*) are essentially modified versions of their Hip Box (HBL_*.pas) counterparts. A localised interpolation-based reconstruction scheme, incorporating on-line distortion compensation, is therefore implemented. However, since the central fiducials are obscured, the unit *KLoc_Foc.pas* uses a more peripheral set of fiducials to estimate the effective focal point locations. Given that all of the fiducials in the

XRII plane are not guaranteed to be found, the number of lines-of-sight in the reconstruction bundle was also increased from sixteen to a potential maximum of twenty-five.

Having already established the accuracy of the Hip Box and its supporting software, it was decided that a similar evaluation of the Knee Box would not be undertaken. This decision was primarily made in view of the following facts:

- Both frames support a biplane imaging geometry with an ideal range of 300 mm.
- Both frames define a 225 mm x 225 mm x 225 mm measurement volume.
- The software used in conjunction with both frames is virtually identical.

Modifying the previously used Perspex test object, in order to achieve compatibility with the Knee Box, was therefore deemed to be an unnecessary drain upon the resources of the MEDROSA project. Following the usual software checks, the emphasis of the distal interlocking study was therefore switched to the development of a more versatile trajectory planning scenario.

6.5.2. Enhanced Trajectory Planning.

As outlined in the previous section, the Knee Box was initially developed as a means of enhancing existing (i.e. perfect circles-based) distal interlocking techniques. However, it was soon realised that the calibration/registration data provided by the Knee Box could also be used to implement a trajectory planning scenario involving "non-perfect" images. In other words, by quantifying the biplane (PA and lateral) fluoroscopic examination process, the need to acquire the perfect circles image could be avoided. As such, a technically demanding and time-consuming stage of the intramedullary nailing procedure could potentially be eliminated.

This "non-perfect circles" scenario exploits the fact that a single point and a rotation about that point, provide sufficient information to define a line (vector) in three-dimensional space. Applying this mathematical concept to the geometry of the distal interlocking problem, drilling trajectories can therefore be reconstructed by extracting the following measurements from the relevant fluoroscopic images:

- The location of the centre of the nail hole (i.e. the intersection of the long axis of the nail with the hole axis).
- The rotation of the nail about its long axis.

As illustrated by Figure 6.20, determining the location of the centre of the nail hole (point C) is a reasonably straightforward process. By indicating (or automatically measuring) the centre line of the nail in both the PA and lateral views, it is possible to reconstruct the location of the long axis of the nail in relation to the Knee Box 3-D coordinate system. The line-of-sight which passes through the centre of the nail hole must then be reconstructed.

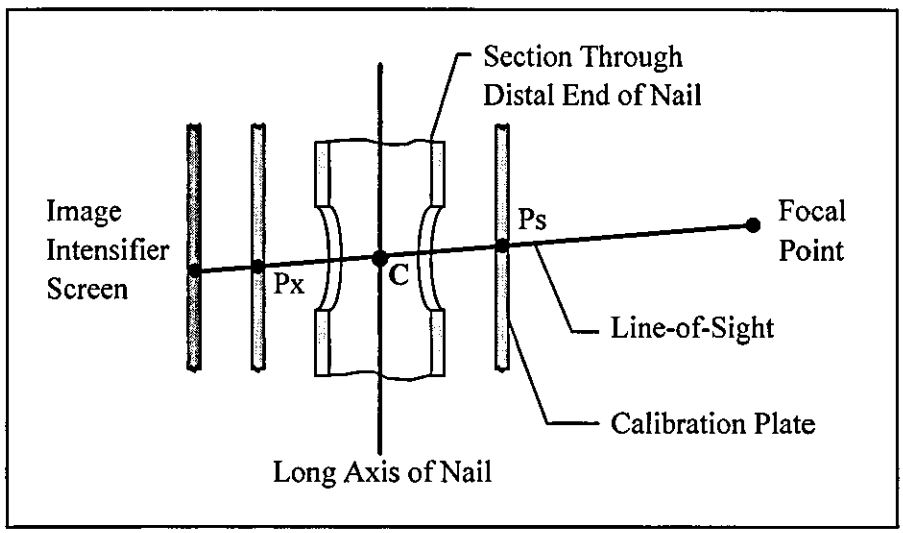


Fig. 6.20 : Reconstruction of The Nail Hole's Centre (C)

Accordingly, the image (pixel) coordinates of the centre of the hole's oval-shaped projection are measured in the "non-perfect" lateral view. The desired line-of-sight can then be reconstructed by applying the appropriate interpolation functions to this measured pixel location, thus allowing the source (Ps) and XRII (Px) calibration plate intersection points to be calculated. Alternatively, the line-of-sight can also be calculated using the known effective focal point location, and either Ps or Px. The intersection of the nail axis with this line-of-sight then yields the desired centre of the nail hole (i.e. point C).

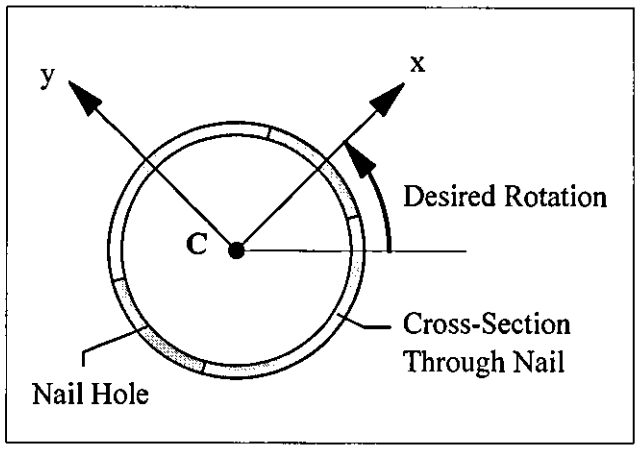


Fig. 6.21 : Rotation About the Long Axis of the Nail.

By contrast, determining the nail rotation depicted in Figure 6.21, involves a more complex mathematical modelling process. The oval-shaped projection of the nail hole must initially be extracted from the "non-perfect" lateral view. This hole profile is then analysed to extract one or more "characteristic" measurements, whose response to changes in the angular orientation of the nail have been empirically modelled off-line. Unfortunately, this image/model matching process is further complicated by the presence of image distortion and the unconstrained nature of the imaging geometry. As a result, the chosen characteristic measurements must be scale invariant. Compensation for the effects of image distortion must also be applied to either the imaged profile or the model.

In order to identify potential candidates for these "characteristic" measurements, a means of accurately setting the rotation of the nail was required. Accordingly, the test rig depicted in Figure 6.22 was manufactured from a vernier-type rotary adjustment system. However, since intramedullary nails are pre-bent to match the natural curvature of the femur, this system was not designed to be used in conjunction with an actual nail. Instead, the distal end of the nail was simulated using a one inch diameter metal tube, into which a pair of "nail holes" with appropriate diameters were drilled. Starting with the perfect circles alignment case (0° rotation), fluoroscopic images of this model of the nail were then acquired at one degree increments, until the hole profile was no longer visible.

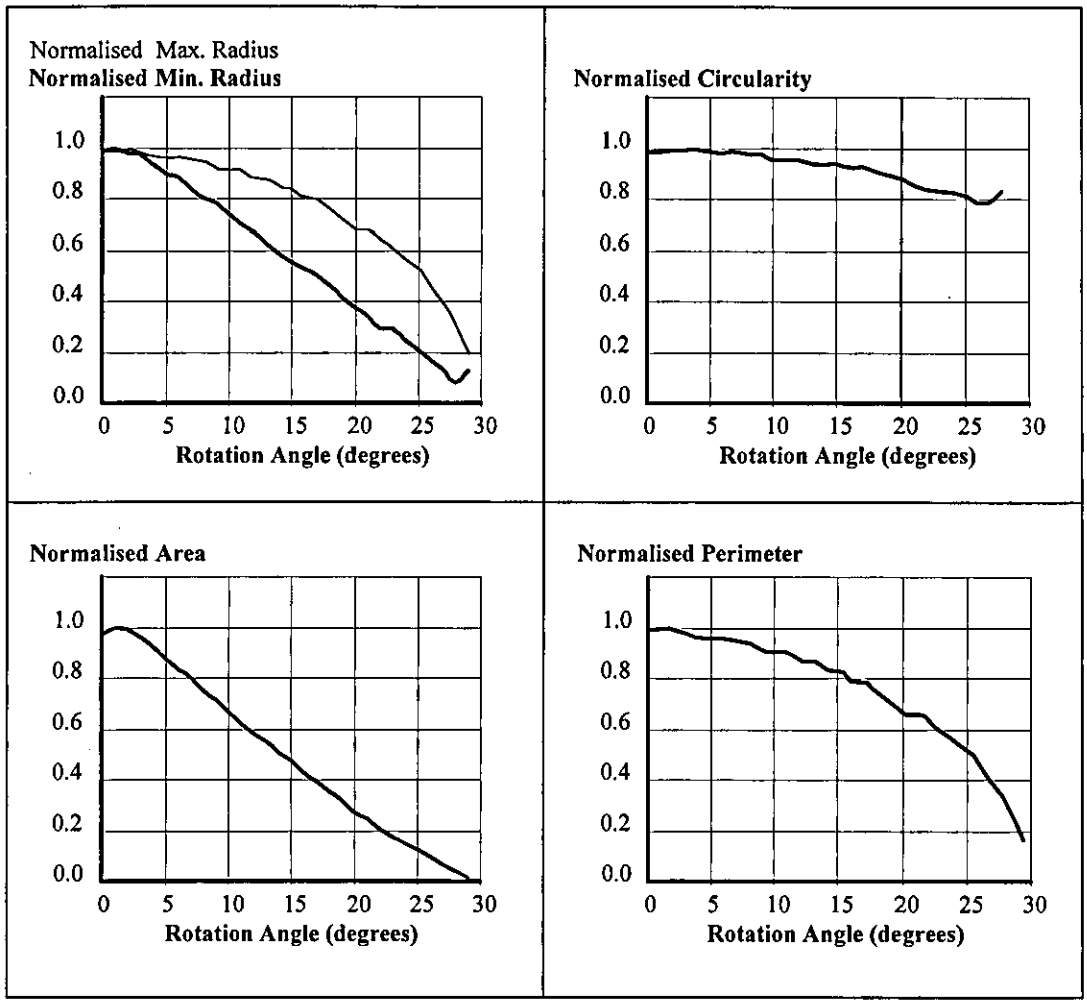


Fig. 6.23 : Graphs of Characteristic Hole Measurement Vs. Rotation Angle

For the particular nail model under investigation, the "cut-off" point occurred after thirty degrees of rotation. A sequence of thirty images was therefore generated, from which the relationship between projected nail hole characteristics and angular orientation could be studied. Having deliberately opted to use a nail model that was larger than actual size, in order to minimise the influence of errors in the image analysis process, the standard JLGenias segmentation and measurement routines were deemed to be adequate for the image analysis task. The data obtained using this software was then plotted to produce the measurement versus angular rotation graphs depicted in Figure 6.23. These graphs were then evaluated in relation to the requirements of a "characteristic" measurement. Due to the high degree of linearity in the appropriate graph, a decision was subsequently made to further investigate the use of a scale invariant area-based measurement as the "characteristic" measurement.

Having formulated the basis of a technique for "non-perfect circles" trajectory planning, the task of writing image analysis software to implement the proposed scenario was offered as an MSc project. A detailed account of this research can therefore be found in the appropriate report **[224]**. However, in view of the relevance to the current thesis, a brief summary of the main findings is presented in the following paragraphs.

By developing appropriate segmentation and sub-pixel measurement algorithms, it was found that the imaged profile of a nail hole could be reliably extracted, from a "non-perfect" lateral view, to a high level of precision. The area enclosed within the extracted boundary of the hole can then be calculated using a numerical approximation technique. The desired scale invariant "characteristic" measurement is then obtained by dividing this calculated area by the square of the nail's diameter, which in turn is also measured from the appropriate lateral image.

As indicated at the start of this section, information extracted from both (PA and lateral) views can then be used to establish the location of the long axis of the nail and the centre of the nail hole. A matching process between the actual image of the nail, and a computer generated model of the distal end of the nail, can then be performed to determine the rotation of the nail about its long axis. This process initially requires the graphical model of the nail to be translated to the known location, in relation to the Knee Box coordinate system, of the actual nail. An initial estimate of the actual angular orientation of the nail is then used to rotate the graphical model.

In order to provide compensation for the effects of image distortion, the graphical model of the nail must then be projected into one of the calibration planes of the Knee Box. The area-based "characteristic" measurement can then be extracted from this projection of the graphical model. If this value matches the corresponding value from the actual image, the (known) angular orientation of the model and the (unknown) angular orientation of the actual nail are identical. If the two "characteristic" measurements are not identical, an iterative optimisation process is performed. This process starts by adjusting the angular orientation of the graphical model, and then

repeats the entire projection/measurement process. In this way, each successive iteration reduces the error between the angular orientations of the model and actual nail, until a satisfactory solution is eventually obtained.

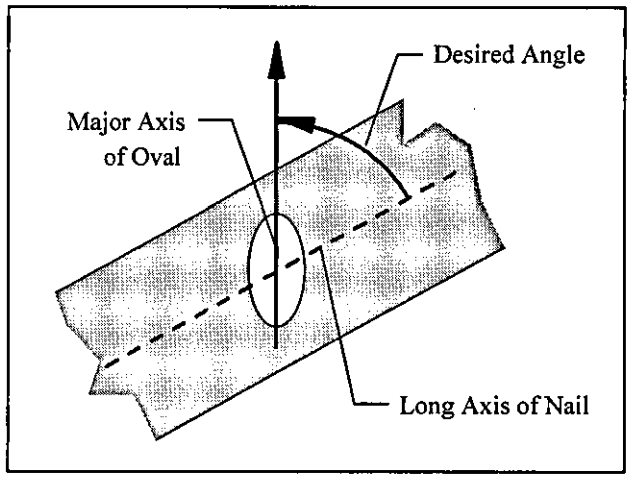


Fig. 6.24 : The Angle-Based "Characteristic" Measurement
(Adapted from [224])

Using the test rig depicted in Figure 6.22, it was possible to demonstrate that the area-based "characteristic" measurement approach allowed the rotation of the (test rig) nail model to be measured to a high level of accuracy (1° or better). However, it was also discovered that the area-based approach does not provide a unique solution with regard to the rotation angle (i.e. a ± 180° scenario exists). An "angle-based" approach was therefore investigated as a more appropriate alternative. As indicated by Figure 6.24, the "characteristic" measurement used by this scenario is the angle between the major axis of the oval-shaped hole profile and the long axis of the nail. Given that the location of the long axis of the nail does not change as the nail is rotated, this angle-based approach provides the desired unique solution to the nail rotation problem.

Trials involving the test rig suggest that this angle-based approach is more accurate than the area-based approach. However, there is one obvious scenario under which the angle-based technique will not work; the major axis of the oval-shaped hole projection being parallel to the long axis of the nail. As such, it is recommended that the area-based approach is maintained as a backup to be used in the unlikely event of this limiting scenario ever being encountered.

6.6 Summary of Chapter 6.

The orthopaedic management of femoral shaft fractures frequently involves the surgical insertion of an intramedullary nail across the fracture site. Depending upon the complexity of the fracture, additional stabilisation may then be obtained by rigidly attaching the nail (using a pair of screws) to the distal end of the femur. Unfortunately, since the nail is located inside the femur, this distal interlocking process can only be performed under fluoroscopic guidance. As outlined in this chapter, the "blind" insertion of distal locking screws therefore involves a technically demanding, time-consuming sequence of processes, which significantly prolong the duration of the operation, and can also lead to excessive levels of patient and surgeon irradiation.

In view of these shortcomings, the MEDROSA Project has targeted the distal interlocking process as a prime candidate for robotic-assistance. The X-ray photogrammetry based registration techniques described in Chapter 5, have therefore been adapted to meet the requirements of this fluoroscopically-guided drilling process. The outcome of this research is a custom-built calibration frame (i.e. the Knee Box), which allows quantitative data to be extracted from fluoroscopic images containing a metallic (internal fixation) implant. Software modules which implement a simple trajectory planning scheme, involving the perfect circles image, have also subsequently been developed.

However, in relation to the distal interlocking procedure, the most significant performance-related benefit offered by the MEDROSA system is undoubtedly the potential to eliminate the perfect circles C-arm alignment process. Given that the Knee Box supports on-line internal and external calibration, it has been possible to develop image analysis routines which allow the location and angular orientation of the intramedullary nail to be accurately determined from a "non-perfect" pair of fluoroscopic views. As a consequence, the use of this computer-assisted trajectory planning software, could potentially lead to significant reductions in both the procedural and fluoroscopy times currently associated with conventional intramedullary nailing.

7. INTERNAL FIXATION OF PROXIMAL FEMORAL FRACTURES.

7.1 Proximal Femoral Fractures.

Fractures of the proximal femur, more commonly referred to as hip fractures, are initially classified as being either *intracapsular* or *extracapsular*. Intracapsular proximal femoral fractures occur within the capsule of the hip joint, and as depicted in Figure 7.1, are usually designated as being either femoral head or neck fractures. The most common site for an intracapsular fracture is the junction between the head and neck of the femur, with a fracture in this region being referred to as a *subcapital* fracture. Fractures at the narrowest part of the femoral neck, referred to as *transcervical* fractures, are also common. Extracapsular proximal femoral fractures occur near, but outside, the capsule of the hip joint. The most noteworthy types of extracapsular fractures are *intertrochanteric* fractures, which occur along a line joining the greater and lesser trochanters, and *pertrochanteric* fractures, in which the fracture line passes through both trochanters. A transverse fracture just below the lesser trochanter may also be referred to as a *subtrochanteric* fracture, or alternatively, the fracture may be considered to be a proximal femoral shaft fracture.

The treatment of intracapsular fractures is generally associated with high complication rates. If the fracture is *displaced*, that is to say the bone fragments move away from each other, the blood supply to the femoral head can be completely cut off. In such cases, a condition known as *avascular necrosis* (i.e. "bone death") occurs as a result of the absence of an adequate blood supply. If a large enough proportion of the bone cells die, the femoral head becomes very brittle, thus preventing adequate purchase being obtained by internal fixation devices. In addition, following an intracapsular fracture, and in particular a subcapital fracture, the femoral head is free to rotate inside the acetabulum. Obtaining an accurate reduction of the fracture can therefore prove to be very difficult. As a consequence of these complications, *non-union* (i.e. failure to unite) can occur in up to one third of all surgically treated femoral neck fractures **[199]**.

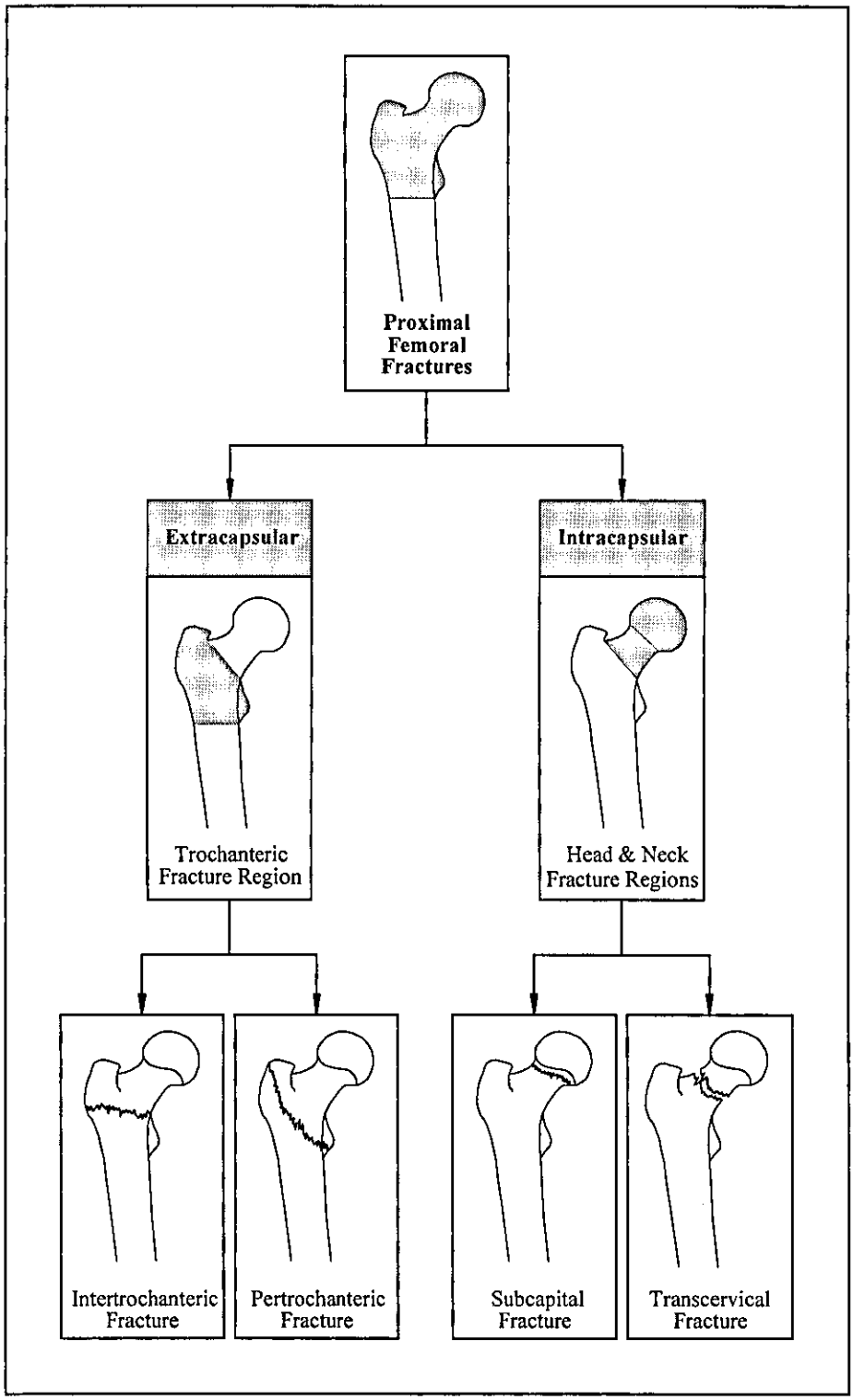


Fig. 7.1 : Proximal Femoral Fracture Types.

By contrast, the prognosis for extracapsular fractures is much better. Since the fracture is outside the joint capsule, the blood supply to the femoral head usually remains intact. A larger contact area also exists between the broken bone fragments,

thus making reduction easier to establish. As a result, extracapsular fractures tend to unite very easily, and seldom cause avascular necrosis.

Proximal femoral fractures are rare in adults under the age of fifty years old **[225]**. However, as indicated by Figure 7.2, above this age the incidence of hip fractures increases exponentially, and eventually reaches very high rates of fracture among adults in their seventies and eighties **[226]**. Consequently, these *age-related fractures* are most commonly associated with elderly patients with brittle bones **[227]**. Most femoral neck fractures are therefore caused by minor trauma, typically a fall within the home. Similarly, trochanteric fractures tend to occur as a result of minor trauma involving rotation/twisting of the femur. However, the relative importance of the factors contributing to these fractures remains controversial **[228]**.

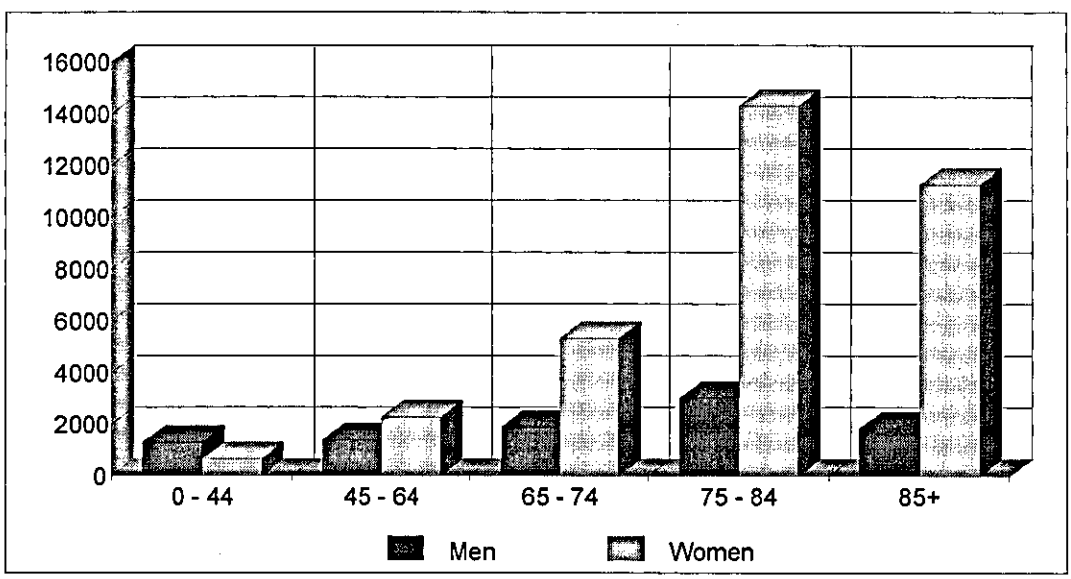


Fig. 7.2 : Distribution of Femoral Neck Fracture Cases by Age & Sex, in England
(Adapted from 1985 data supplied by Hoffenberg *et al.* **[229]**)

Examination of Figure 7.2, and the follow-up study data supplied in Table 7.1, reveals two important facts. The first is that the average age of a patient suffering a proximal femoral fracture is approximately eighty years old. The second is that on average, women are four times more likely to suffer a proximal femoral fracture than men. A partial explanation for these figures has been provided by studying *osteoporosis*, an

age-related disorder characterised by decreased bone mass and increased susceptibility to fractures.

Table 7.1 : Typical Follow-Up Data from Proximal Femoral Fracture Studies: Average patient age, male-to-female ratio & fracture fixation failure rates.

Study	Year	Region	Study Size	Av. Age	M:F	Fixation Failure Rate
Hoffenberg *et al* [229]	1985	England	43,220	NA	1:3.54	NA
Mainds *et al.* [230]	1989	Glasgow	385	74	1:3.50	9/385 (2.3 %)
Davis *et al.* [231]	1990	Durham	114	80.6	1:4.75	22/114 (19.3 %)
Thomas *et al.* [232]	1991	Wolverhampton	90	82	1:5.20	9/90 (10.0 %)
Gregg *et al.* [233]	1993	Leicester	656	79.7	1:3.50	NA
Keene *et al.* [234]	1993	Birmingham	1000	79.0	1:4.26	NA
Gundle *et al.* [235]	1995	Oxford	89	82	NA	12/89 (13.4 %)

The strength of bone is determined by mineral density, which in turn is dictated by the mass of calcium salts per unit volume **[236]**. However, as indicated by Figure 7.3, bone mass varies considerably throughout an individual's lifetime, with three distinct phases being observed: skeletal growth, consolidation, and bone loss **[227]**. As highlighted by Figure 7.3, at all ages, women have a lower bone mass than men. However, a more significant factor demonstrated by this figure is the accelerated bone loss experienced by women at an early age, as a result of *postmenopausal osteoporosis* **[226]**. Thus, although both sexes suffer from *senile osteoporosis*, women already have significantly lower bone mass prior to the onset of this age-related form of osteoporosis. Women are therefore more likely to reach the "fracture threshold", the bone mass below which a fracture can occur in the absence of severe trauma, than men. Given that women also tend to live longer than men, their risks of sustaining age-related fractures, such as hip, vertebral, and distal radius fractures, are also much higher than the corresponding values for men.

However, a direct link between osteoporosis and proximal femoral fractures has not been proven **[233]**. Recent studies have failed to show consistent differences in bone mass between patients suffering a proximal femoral fracture, and members of the

general public of the same age **[228, 237]**. As a result, the incidence of proximal femoral fractures may be more closely related to the increased risk of falling experienced by the elderly. A reduced ability to break the impact of a fall, due to failing eyesight, impaired reflexes, and less fat and muscle around the hip, may also be a significant factor, particularly among patients over the age of seventy-five years old **[228]**.

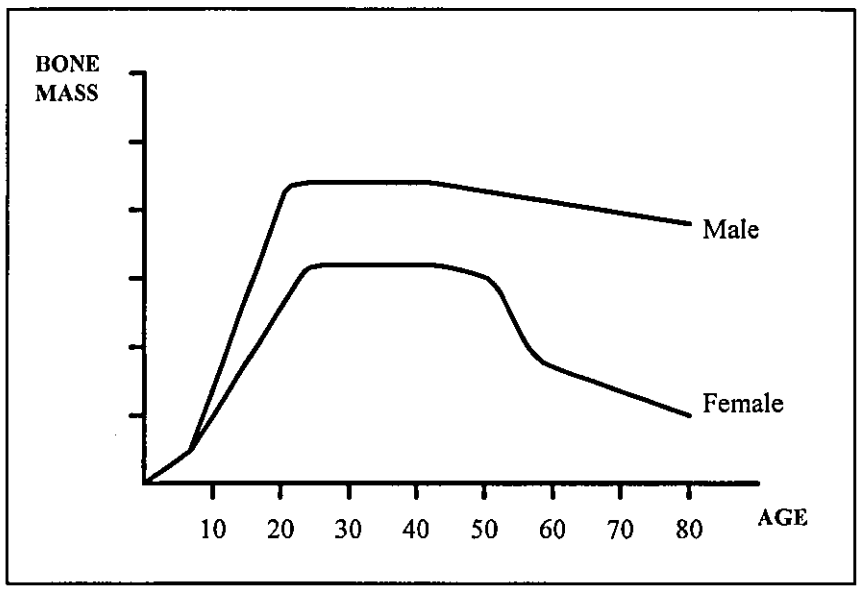


Fig. 7.3 : Diagrammatic Representation of Changes in Bone Mass with Ageing.
(Adapted from **[227]**)

The importance of determining the causes of proximal femoral fractures becomes self-evident, in light of the burden placed upon orthopaedic resources by patients suffering this type of fracture. Given the advanced age of a typical patient, proximal femoral fractures are associated with high rates of morbidity and mortality; up to a third of all patients die as a result of conditions such as bronchopneumonia or pulmonary embolism **[234]**. Unfortunately, the fact that only minor trauma is required to produce a hip fracture in the elderly means that this serious injury is also very common. In 1985, the reported number of proximal femoral fracture cases reached 43,200 in England alone **[229]**. The average length of hospitalisation for these patients was thirty days, and as such, on any given day 3,500 orthopaedic beds were occupied by patients with proximal femoral fractures. The total cost of treating hip fractures in the United Kingdom was therefore estimated as being £160 million per annum (at 1987 prices) **[229]**. Looking further afield, the cost of treating the quarter of a million hip

fractures occurring in the United States has been estimated as $8 billion per year (1988 prices) [238]. Similarly, the number of hip fractures occurring world-wide in 1990 was approximately 1.6 million [239].

However, the most significant feature of this type of data is the fact that successive studies reveal large increases in the numbers of hip fractures. By 1991/92 the figures for England had risen to an estimated 56,613 proximal femoral fractures, treated at a total cost of £288 million [240]. It therefore appears, that as the life expectancy of the population increases, the incidence of proximal femoral fractures is set to reach epidemic proportions [241]. If the current rates of increase are maintained, the world-wide incidence of hip fractures could therefore be as high as 6.26 million by the year 2050 [239]. In order to alleviate the enormous burden that will be placed upon orthopaedic and geriatric services, major initiatives are therefore required to both prevent hip fractures, and improve existing treatment techniques.

7.2 Treatment of Proximal Femoral Fractures.

In view of the advanced age of most hip fracture patients, prolonged periods of traction are not advisable [229]. Unless the patient has a life-threatening condition, which makes administering anaesthesia dangerous, osteosynthesis is therefore the preferred treatment option. The primary aims of this surgery are reduced mortality rates, stabilisation of the fracture site, and early mobilisation of the patient. Ideally, over a period of months, the patient can then regain the walking ability that he/she had prior to sustaining the hip fracture. As with most osteosynthesis procedures, selection of the internal fixation device depends upon the personal preference of the surgeon. However, the location of the fracture site is also a major determinant, with differing forms of fixation usually being applied to intracapsular and extracapsular fractures.

The most common site for a hip fracture is in the trochanteric region [242]. As these extracapsular fractures are usually caused by more severe trauma, their associated mortality rates are much higher than those of intracapsular fractures [243]. Over the years, stabilisation of extracapsular fractures has been achieved using a wide variety of

internal fixation devices, ranging from one piece nail-plates, such as the Jewett Nail **[244]**, to intramedullary devices, such as the Gamma Locking Nail **[245]** or Ender Nails **[243]**. However, the most successful implant design to date is the *sliding compression hip screw*, which is typified by the ASIF Dynamic Hip Screw (DHS) **[246]** shown in Figure 7.4, and the Richard's Compression Screw and Plate.

As can be seen from Figure 7.5, the sliding compression hip screw is a two piece device consisting of a *lag screw* and an *angled plate*. The cannulated lag screw is designed to be inserted into the femoral neck/head with the aid of a guide wire. The angled plate is then attached to the lateral cortex of the proximal femur, using a series of screws, with its barrel containing the unthreaded portion of the lag screw. As such, the lag screw is free to slide within the barrel of the angled plate, thus allowing controlled impaction of the fracture site, while at the same time maintaining a constant neck-shaft angle (typically 135°) **[247]**. It is believed that this impaction improves fracture stabilisation and speeds up the bone healing process.

Owing to the high incidence of avascular necrosis, internal fixation of an intracapsular fracture should ideally take place within twenty-four hours of the patient being admitted into hospital **[229]**. Given that it is essential to control rotation of the proximal bone fragment (i.e. the femoral head), in relation to the rest of the femur, stabilisation of intracapsular fractures is generally achieved using two or more fixation devices. A group of pins, for example Knowles pins **[243]** or Hansson pins **[248]**, may therefore be inserted into the femoral neck/head in a parallel configuration which allows impaction of the fracture site. Alternatively, lag screws, such as the Garden Screw, the Von Bahr Screw **[249]**, or the Uppsala Screw **[248]**, may be inserted in either a parallel or a crossed configuration. In order to allow insertion over a guide wire, these screws are frequently cannulated, as in the case of the AO/ASIF Cancellous Bone Screw **[246]**.

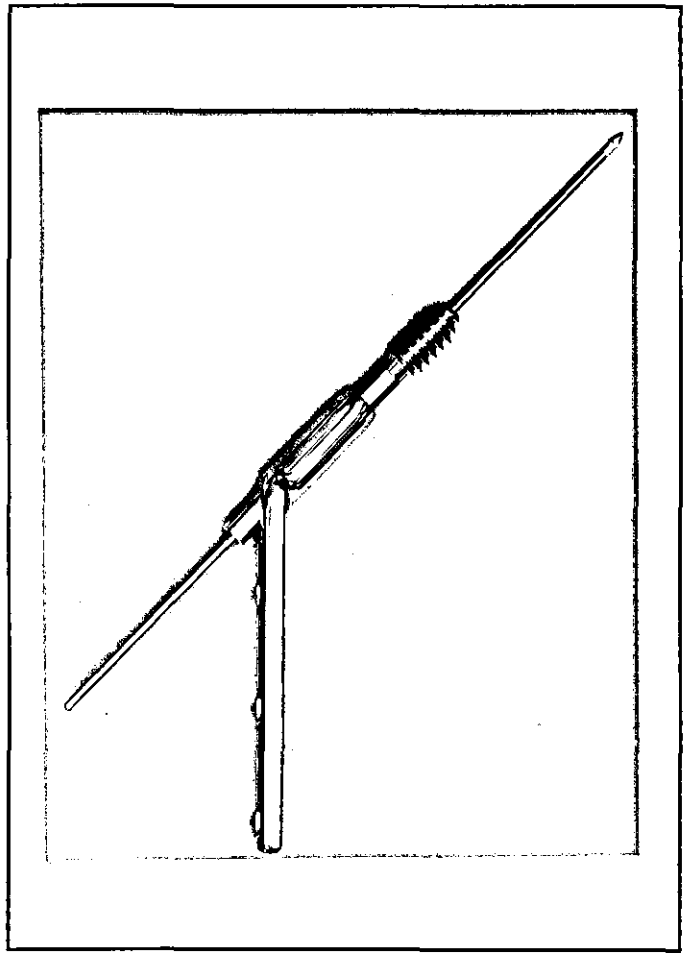


Fig. 7.4 : The AO Dynamic Hip Screw.
(Shown with insertion guide wire.)

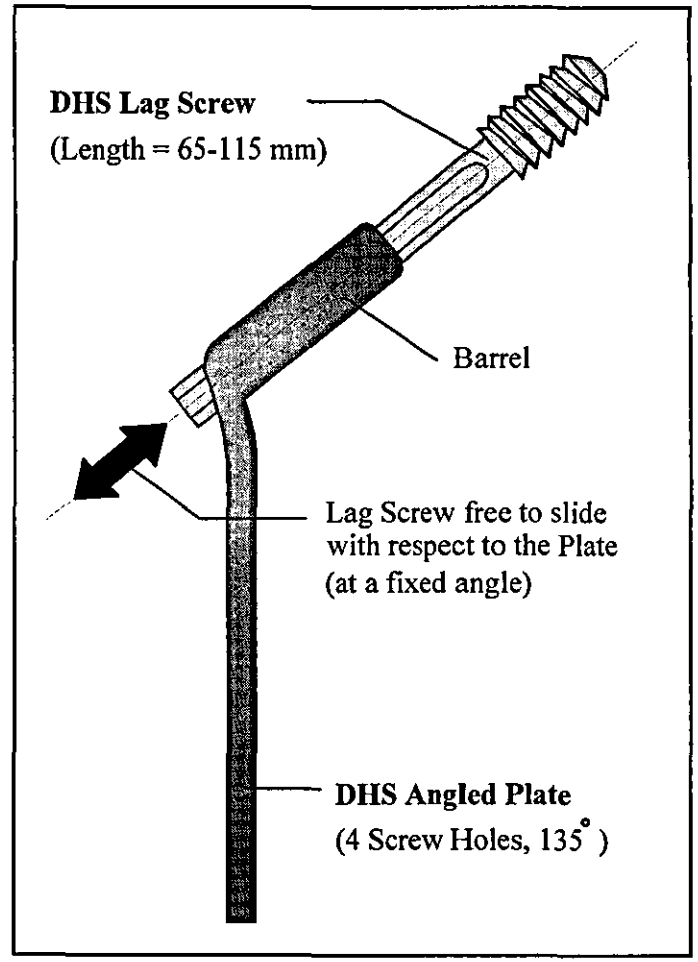


Fig. 7.5 : AO DHS Components.

Finally, although the sliding compression hip screw gives poor rotational stability in comparison to multiple pin/screw techniques, it is still occasionally used to provide fixation of intracapsular fractures. In particular, the use of a sliding compression hip screw is advantageous for intracapsular fractures involving comminution of the lateral femoral cortex, or severe osteoporosis **[243]**. However, an angled plate which is smaller than those used to achieve fixation of extracapsular fractures is generally required. In order to improve rotational stability, it may also be necessary to insert a second lag screw parallel to the sliding compression hip screw **[242]**.

7.3. Insertion of the Sliding Compression Hip Screw.

As previously discussed in Section 6.4, the many variables associated with an osteosynthesis procedure mean that a definitive description of surgical technique cannot be provided. The following account of sliding compression hip screw insertion is therefore based upon textbook recommendations **[243, 246, 250]**, and surgical procedures witnessed by the author at Leicester Royal Infirmary and Mayday University Hospital, Croydon, Surrey. The description applies to both the internal fixation of extracapsular (trochanteric) and intracapsular (neck) fractures, with any noticeable differences being highlighted. Owing to the success of the sliding compression hip screw, several design variants are currently in routine clinical use. It must therefore be appreciated that the description provided here applies primarily to the ASIF Dynamic Hip Screw (DHS) and the Richard's Compression Screw and Plate. Minor deviations from the outlined surgical technique may therefore be required when using other sliding compression hip screw systems.

The internal fixation of a proximal femoral fracture using a sliding compression hip screw involves the following stages:

- Positioning of the patient.
- Reduction of the fracture.
- Exposure.
- Guide wire insertion.
- Reaming/tapping of the femur.

- Insertion of the lag screw and plate.
- Attachment of the plate.

The anaesthetised patient is placed in the supine position on a fracture table, with his/her perineum firmly against a radiolucent counter-traction post. The patient's feet are then fastened to the traction foot pieces of the fracture table, and the uninjured leg is flexed and abducted (i.e. moved away from the midline of the body) to achieve compatibility with the mobile C-arm unit. Posteroanterior (PA) and lateral-to-medial fluoroscopic images can therefore be acquired throughout the surgical procedure in the manner shown in Figure 7.6. Following the application of traction to the injured leg, a closed reduction of the hip fracture is then attempted. If this manipulation process is unsuccessful, an open reduction will be performed. When fluoroscopic verification of a satisfactory reduction is obtained, the invasive stages of the surgical procedure are then initiated.

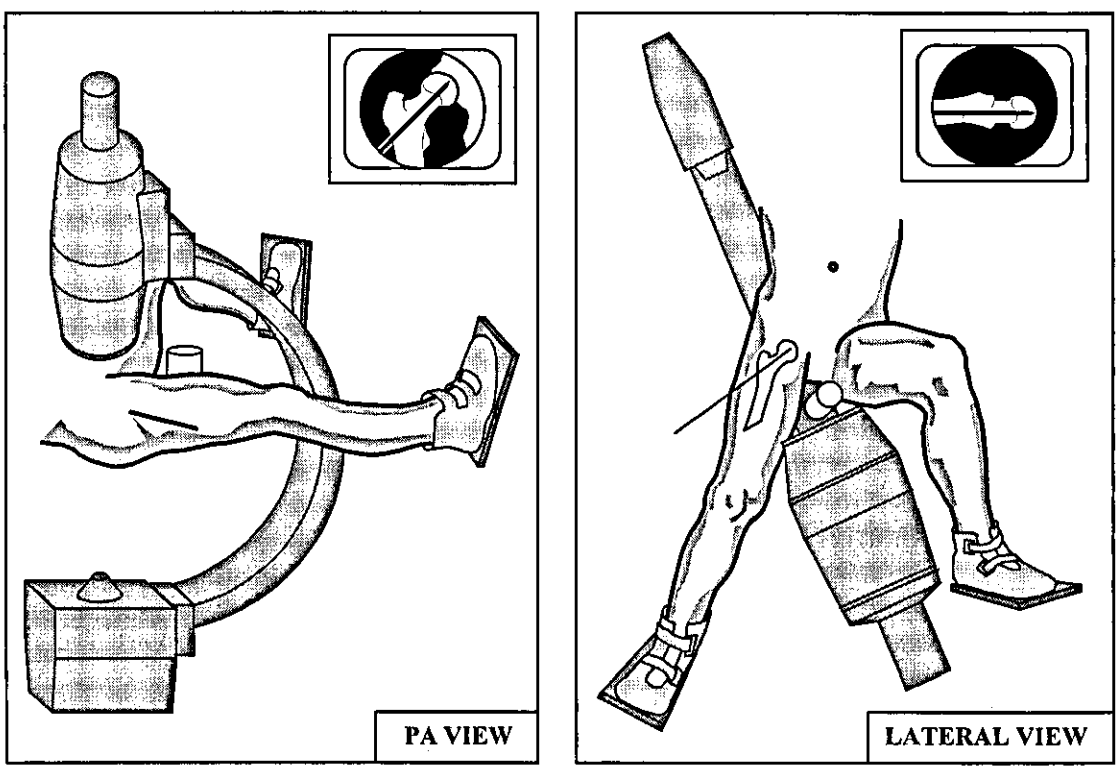


Fig. 7.6 : C-Arm Position for Fractures of the Proximal Femur.
(Adapted from **[251]**)

The usual skin preparation and draping procedures are followed by the creation of a lateral incision (approximately 15-20 cm long) over the proximal femur. Access to the

lateral femoral cortex is then gained by splitting and retracting the surrounding muscle groups. In order to allow attachment of the angled plate of the sliding compression hip screw, any muscle tissue that remains attached to the lateral femoral cortex is then stripped away.

As depicted in Figure 7.7(A), a pilot hole must then be drilled into the femoral neck/head, using a guide wire mounted in an air-powered surgical drill. As this guide wire must maintain a fixed angle to the femoral shaft (typically 135°), the pilot hole is created with the aid of a hand-held angle (drill) guide, which is placed against the lateral femoral cortex. In order to further simplify the guide wire insertion process, a second (external) guide wire is frequently introduced into the soft tissue above the femoral neck. The position of this external guide wire is then manually adjusted until it appears to be aligned with the ideal drilling trajectory in the PA fluoroscopic view. Using this external guide wire as a reference marker, the guide wire is then drilled into the lateral femoral cortex at a point just opposite from the tip of the lesser trochanter. Drilling then continues under fluoroscopic guidance (both views) until the threaded tip of the guide wire is approximately one centimetre from the articular surface of the femoral head. The external guide wire can then be removed.

Having successfully inserted the guide wire, the length of the required lag screw is determined by measuring the length of guide wire protruding from the femur. Manual reaming of the pilot hole must then be performed to allow insertion of the chosen lag screw. A triple reamer assembly, adjusted to the correct depth, is therefore introduced over the guide wire, thus allowing reaming for the lag screw, the barrel of the angled plate, and the barrel/plate junction to take place. Following fluoroscopic verification (both views) of the reaming process, a screw thread will then be cut into the femoral head using a tap. However, as an elderly patient's femur is usually significantly weakened by osteoporosis, this tapping process is not always required. Manual insertion of the cannulated lag screw, over the guide wire (refer to Figure 7.7(B)), is then performed using a T-handled wrench assembly. Once again, fluoroscopy (both views) is then employed to check the lag screw's final position.

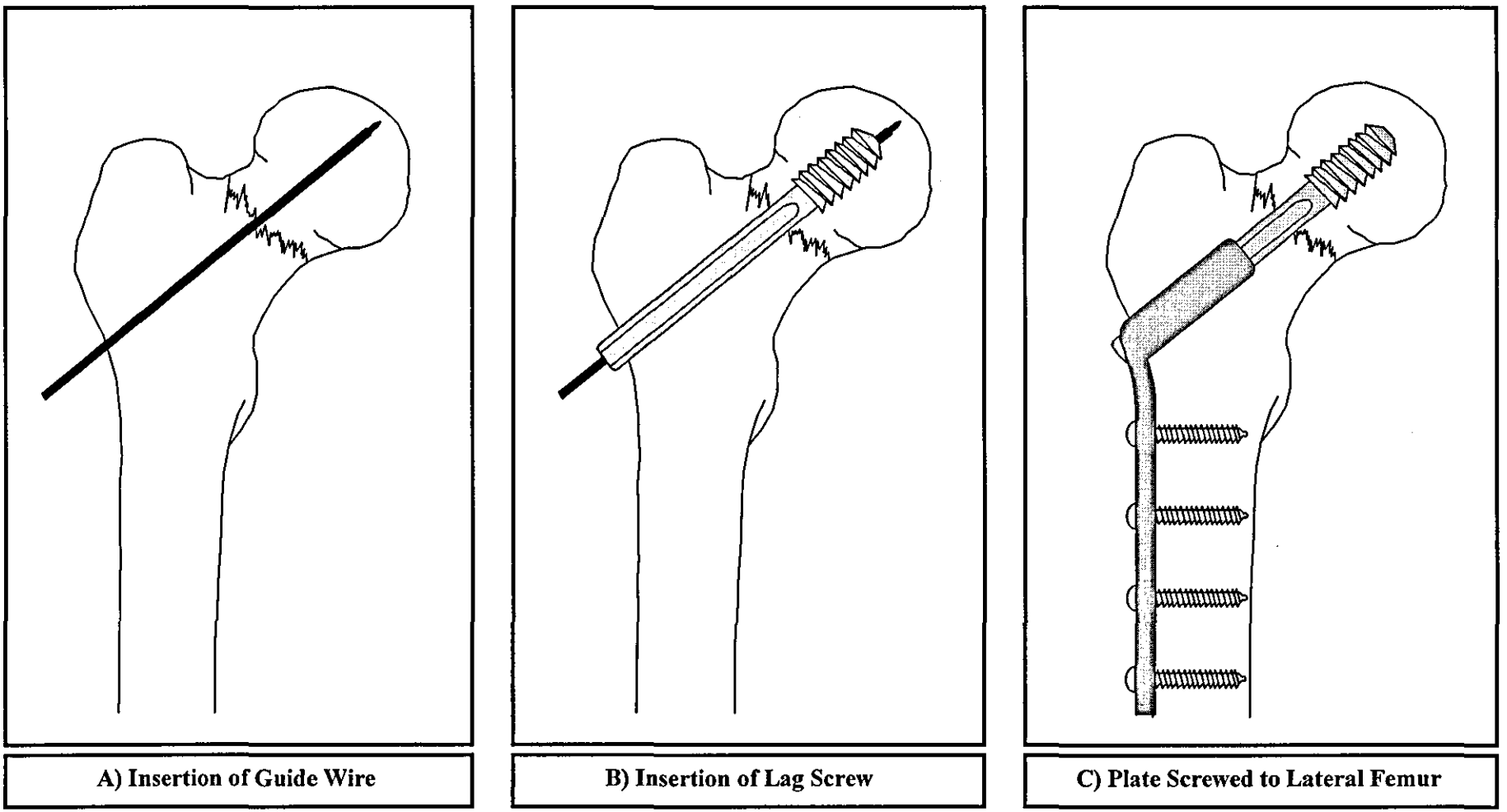


Fig. 7.7 : Surgical Insertion of the Sliding Compression Hip Screw.

Insertion of the sliding compression hip screw is then completed, by attaching the angled plate to the lateral femoral cortex in the manner shown in Figure 7.7(C). Accordingly, the angled plate is inserted over the lag screw, and is seated against the femoral cortex with the aid of an impaction instrument. Four holes are then drilled into the proximal femoral shaft through the screw-holes in the angled-plate, and the length of the required cortical screws is determined using a depth gauge. Following manual insertion of the cortical screws, the guide wire is then removed, and traction is released from the injured leg. If required, impaction of the fracture site may then be achieved by either manually tapping the femur, or inserting a compression screw into the end of the lag screw. The usual wound closure procedures (i.e. suturing of muscle tissue and external stapling) are then performed.

Postoperative mobilisation of the patient begins as soon as possible. Bedside sitting on the first postoperative day is therefore followed by walking (protected weight-bearing) with the aid of a frame, and then crutches. A programme of physiotherapy exercises is also undertaken to build up the leg muscles. Following radiographic evidence of satisfactory callus formation, typically after 8-12 weeks, progressive weight-bearing is then allowed. Ideally, follow-up radiographic examinations are then continued for up to one year after surgery, with a typical proximal femoral fracture taking 3-5 months to heal.

Apart from the previously mentioned use of a smaller angled plate, typically with two screw-holes instead of the usual four, the only significant difference between the fixation of extracapsular and intracapsular fractures occurs during the guide wire insertion process. Owing to the need to control rotation of the femoral head, placement of the guide wire is more critical during the internal fixation of intracapsular fractures. Prior to drilling the pilot hole, it is therefore common practice to insert a second (stabilising) guide wire into the femoral neck/head. This stabilising guide wire is then removed following insertion of the primary guide wire and the lag screw.

7.3.1 Guide Wire Insertion Problems.

The "blind" insertion of a guide wire into the femoral neck/head is a prime example of a three-dimensional orthopaedic task performed under two-dimensional fluoroscopic control. In order to compensate for the resulting loss of depth perception, a trial-and-error insertion technique must therefore be adopted. As such, the guide wire is tentatively advanced into the femur in small increments, with progress being monitored fluoroscopically (both views) at intermediate stages. Owing to the subjective nature of this insertion process, inexperienced surgeons can require several attempts to achieve a satisfactory guide wire placement. This failure to perform right-first-time guide wire insertions has many implications.

Repeated attempts to insert the guide wire obviously prolong the duration of the surgical procedure. The patient therefore receives more anaesthetic, and loses more blood intraoperatively. The risk of infection is also increased due to the relatively large incision, which is required to accommodate the angled plate of the sliding compression hip screw, being open for a longer period of time. Given that the surgeon is required to hold the air-powered drill in position during the acquisition of each fluoroscopic image, additional insertion attempts also increase his/her cumulative exposure level. However, it is the relationship between fixation failure rates and guide wire positioning accuracy that gives the most grounds for concern.

Although the sliding compression hip screw provides more reliable stabilisation of extracapsular fractures than any other osteosynthesis device, follow-up studies frequently reveal fixation failure rates of the order of 10% (refer back to Table 7.1). Failure to maintain fixation of the fracture until union, generally occurs as a result of one of the following complications **[232, 250, 252]**:

- Cutting-out of the lag screw through the femoral head or neck.
- Failure of the device's sliding mechanism.
- Screw failure (angled plate pulling away from the femoral shaft).
- Device failure (bending or breaking).

Of these fixation failure modes, the lag screw cutting-out of the femoral head/neck is by far the most common **[231, 251]**. By evaluating radiographs taken immediately after surgery (i.e. before weight-bearing), follow-up studies have been able to demonstrate that a direct relationship exists between the position of the lag screw in the femoral head and failure to maintain fixation **[230-232, 235, 253]**. In other words, lag screws inserted into some parts of the femoral head are more likely to cut-out than others, and as such, fixation failure rates are directly related to the accuracy of the surgical procedure (i.e. the guide wire insertion). Unfortunately, as demonstrated by Figure 7.8, the repeatability of the technically demanding guide wire insertion process is poor. As a consequence, a less than "ideal" guide wire placement must frequently be accepted after several insertion attempts, leading to an increased probability of fixation failure.

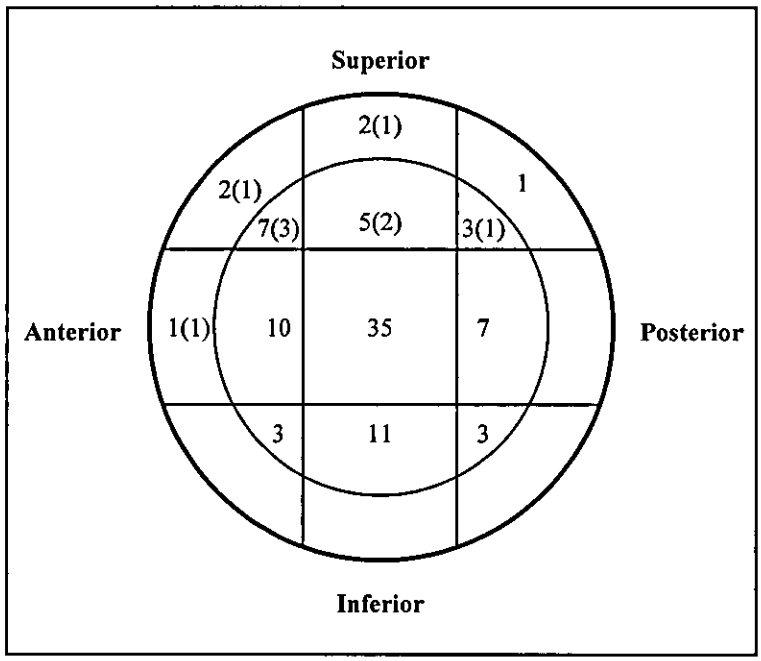


Fig. 7.8 : Fixation Failure as a Function of Lag Screw Position:
Number of lag screws placed in each zone of femoral head (Failures in brackets).
(Adapted from Thomas **[232]**)

Finally, it is also worth noting that the "ideal" position for the guide wire within the femoral neck/head remains the subject of much debate within orthopaedic surgery. Traditionally, orthopaedic textbooks have recommended a posteroinferior placement of the guide wire **[243]**. In this position, the guide wire appears to be slightly inferior of the centre of the femoral head in the PA view, and slightly posterior of the centre of

femoral head in the lateral view. However, as demonstrated by Noordeen *et al.* **[254]**, when the guide wire is placed outside the central two-thirds of the femoral head, in one or both of the views, unrecognised penetration of the femoral head can occur. Essentially, this means that when the spherical femoral head undergoes central projection by a cone of X-ray beams, it is possible for a protruding guide wire to appear to be within the femoral head. As such, a posteroinferior guide wire placement incurs a certain amount of uncertainty with regard to penetration of the femoral head at the time of surgery. A number of recent studies have also demonstrated that the posteroinferior placement leads to increased fixation failure rates due to cutting-out of the lag screw **[231, 235, 253]**. Thus, while there is general agreement that the guide wire should be placed centrally or inferiorly in the PA view, in order to take advantage of the strong cancellous bone in the inferomedial region of the femoral head **[250]**, the guide wire's position in the lateral view remains controversial, with many surgeons now preferring a more central placement.

Thus in summary, reductions in fixation failure rates, and hence the required number of corrective surgical procedures, can potentially be obtained by improving the accuracy and repeatability of the guide wire insertion process. Owing to the huge numbers of proximal femoral fractures that occur annually, even a small reduction in the overall fixation failure rate would prove to be very significant in terms of orthopaedic resources. A partial solution to this problem can be obtained by allocating a small team of specialist "hip" surgeons to treat all patients admitted with a hip fracture. Follow-up studies, such as the one reported by Johansson *et al.* **[255]**, have shown that this dedicated approach leads to improved surgical accuracy, and hence reduced complication rates. However, in the long term, automation of this X-ray guided drilling process may prove to be a more accurate, cost-effective and safer alternative.

7.4 Robotic-Assisted Guide Wire Insertions.

As explained in Appendix 2, automation of the guide wire insertion process, via robotic-assistance, is one of the prime objectives of the MEDROSA research project. However, in relation to the X-ray photogrammetry based methodology adopted by the

current thesis, such an application raises a significant problem: the provision of a calibration frame that is compatible with both the anatomy of the hip/pelvis region and the mobile C-arm fluoroscopy unit.

Traditionally, the imaging geometry employed by an X-ray photogrammetry study is dictated by patient anatomy, with the body part of interest being surrounded by calibration plates in a manner that ensures an interpolation-based reconstruction process. Examinations of smaller body parts (i.e. the extremities), which can be totally enclosed by a four-sided calibration frame, tend to adopt the biplane (PA and lateral) imaging geometry used during orthopaedic surgery procedures. However, this is obviously not a viable option when examining the bulky region of the lower torso. As outlined in Section 2.3.2, the established approach to X-ray photogrammetry of the hip joint therefore involves a convergent imaging geometry, which is supported by a single pair of calibration plates (i.e. one above and one below the patient).

Unfortunately, as discussed in Section 5.2, the mobile C-arm fluoroscopy unit is not well suited to convergent X-ray photogrammetry of the hip joint. In particular, a combination of the following factors make the acquisition of a suitable calibration image difficult:

- The mechanical limits placed upon the C-arm unit's arc motion.
- The relatively small source to image intensifier distance (SID) of the C-arm unit.
- The unconstrained nature of the C-arm unit's location in relation to the calibration frame.

The intraoperative use of a convergent calibration frame, which is large enough to encompass the lower torso, also incurs a number of practical problems. In addition to the necessary deviation from the standard (PA and lateral) fluoroscopic examination process, assembling/introducing such a frame around the anaesthetised and draped patient could prove to be very difficult and time-consuming. The weight and bulk of the frame would also lead to compatibility problems in relation to both the robotic manipulator and the operating/fracture table.

Having established these facts, it became clear that an "operating room" compatible method of quantifying the (biplane) hip fracture fluoroscopic examination process, could only be provided by deviating from established photogrammetric principles. In other words, an extrapolation-based reconstruction technique would have to be adopted, in order to satisfy intraoperative (anatomical and C-arm compatibility) criteria. However, in spite of the fact that extrapolation-based techniques inevitably lead to increased reconstruction errors, it was hoped that "acceptable" errors, in relation to the accuracy requirements of the guide wire insertion process, could still be obtained by implementing such an approach.

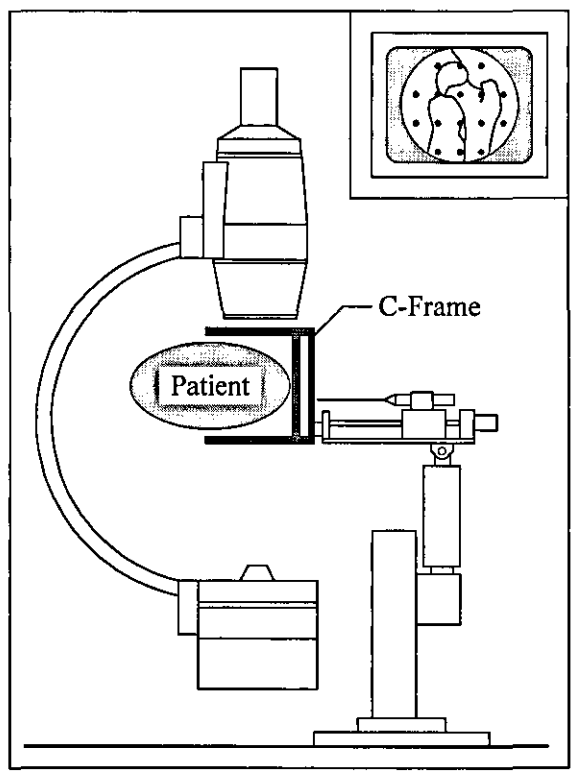


Fig 7.9 : The C-Frame Based Registration Scenario

Figure 7.9 depicts the calibration frame design that was originally proposed as a solution to the intraoperative "hip fracture" registration problem. In this "C-Frame" configuration, calibration of the PA view is achieved by positioning a single calibration plate above and below the patient, thus constituting an interpolation-based reconstruction scenario. However, in order to gain anatomical compatibility, calibration of the lateral view is supported by a pair of calibration plates situated next to the lateral aspect of the patient's hip, thus constituting an extrapolation-based

reconstruction scenario. Unfortunately, in practice, positioning a calibration plate below the patient's hip proves to be difficult due to the presence of the fracture table. Consequently, an L-shaped calibration frame design was considered to be a more promising alternative.

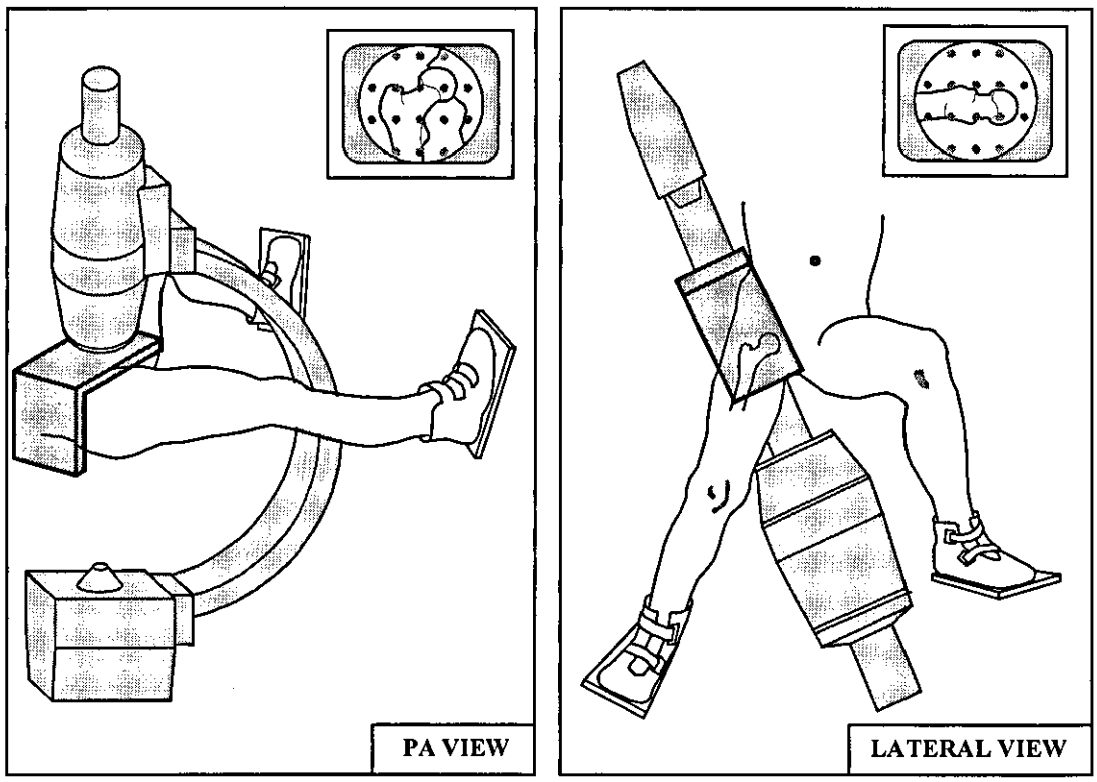


Fig. 7.10 : "L-Frame" Positioning in Relation To Patient Anatomy.

As depicted in Figure 7.10, this "L-Frame" approach involves the positioning of a pair of calibration plates above the hip, and a second pair of calibration plates to the lateral side of the hip. Calibration of both the PA and the lateral views therefore involves an extrapolation-based process. In other words, the patient's hip joint is not located between the calibration plates, and as such does not lie within the true calibrated volume. However, from the orthopaedic perspective, such an arrangement has a number of distinct advantages, most notably:

- Quicker Set Up Times:

  Introducing the L-Frame into both the PA and lateral fields-of-view of the C-arm fluoroscopy unit does not require a time-consuming assembly/dismantling process. Robotic positioning of a lightweight version of the frame could also be achieved

with relative ease. Since the separation of the calibration plate pairs is small, acceptable calibration images should also be obtainable without the need for excessive repositioning or irradiation.

- Compliance With Surgical Sterility:

  The L-Frame configuration could potentially be isolated in a custom-made sterile plastic drape. As a result, many of the material selection and manufacturing problems associated with the design of an autoclavable X-ray calibration frame could be avoided.

Sections 7.4.1 and 7.4.2 therefore outline the details of a laboratory-based evaluation of this L-Frame configuration, in relation to the requirements of robotic-assisted guide wire insertions.

7.4.1 Initial L-Frame Investigation.

The feasibility of the L-Frame registration scenario was established using the calibration frame shown in Figure 7.11. This particular frame design employs a pair of L-shaped aluminium side plates, with a connecting cross-piece in the base, to maintain four Perspex calibration plates in the desired locations. The fiducial patterns used are similar to those of the Hip Box:

- XRII (Inner) Plates:

  A nine-by-nine grid (15 mm grid spacing) of 2 mm diameter ball-bearings. Grid centre (5 mm diameter ball-bearing) and orientation (6 mm outside diameter washer) markers are also provided.

- Source (Outer) Plates:

  An offset eight-by-eight grid (13 mm grid spacing) of 3 mm diameter ball-bearings. Four smaller (4 mm outside diameter) washers mark the centre of the grid.

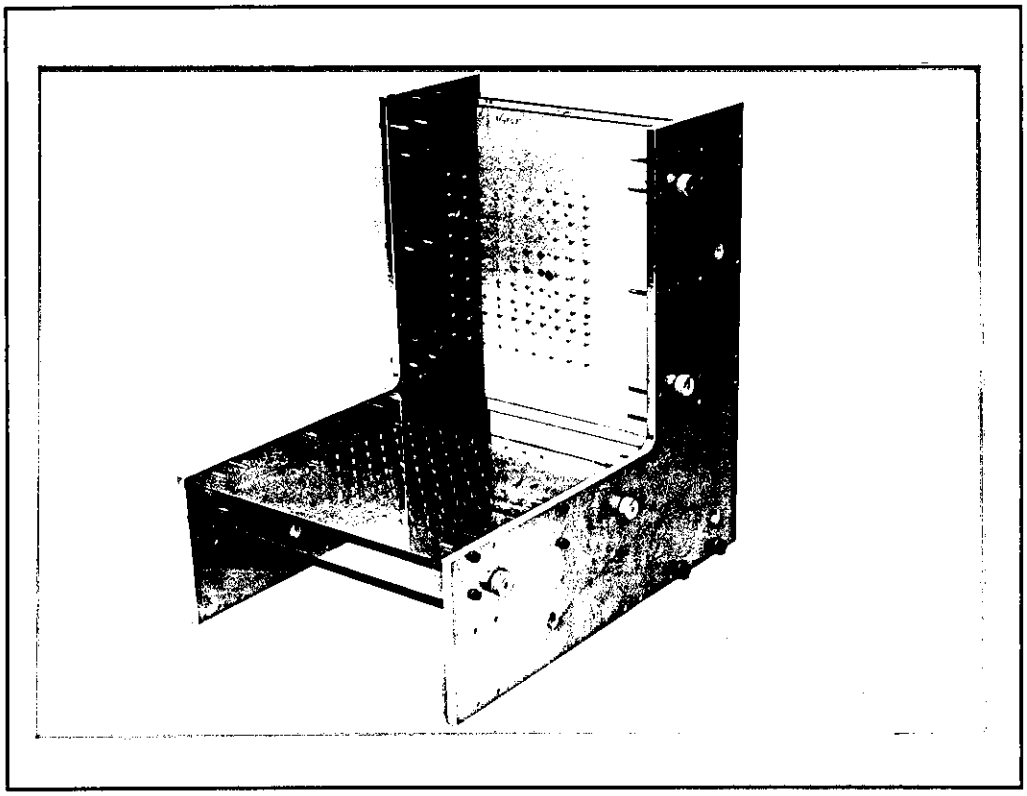


Fig. 7.11 : The "L-Frame" X-Ray Calibration Frame.
(External Measurement Volume = 225 mm x 225 mm x 225 mm).

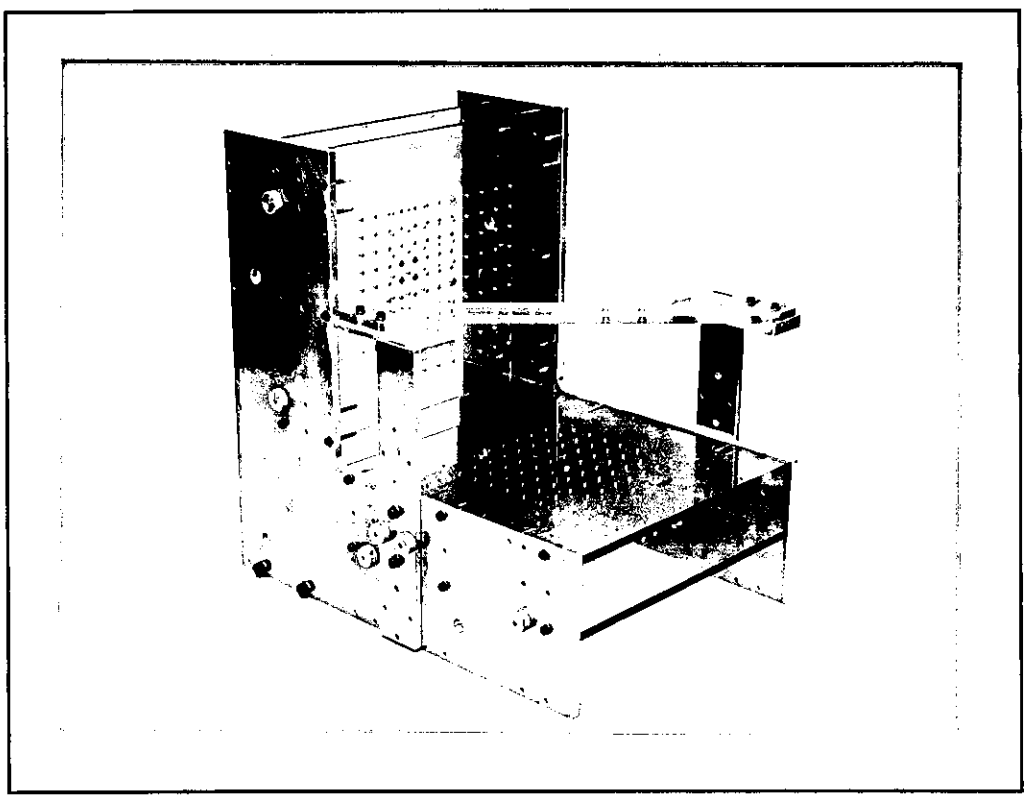


Fig. 7.13 : Perspex Test Object Used in Conjunction With The L-Frame.

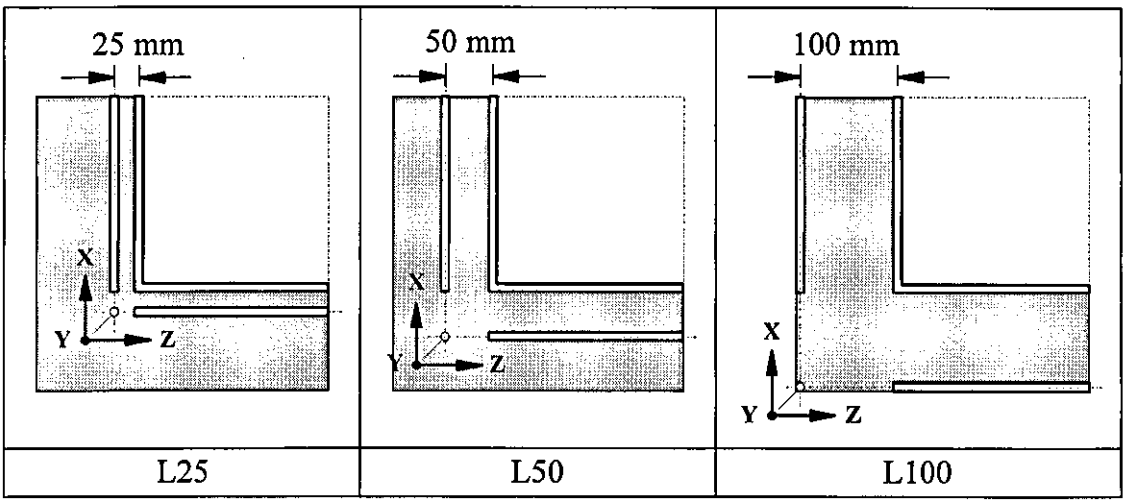


Fig. 7.12 : The Three L-Frame Calibration Plate Configurations.

However, it must be stressed that this design reflects the frame's role as a development "tool", and as such, the frame deviates from the intraoperative scenario depicted in Figure 7.10 in the following ways:

- Increased Size/Weight:

  The L-Frame evaluation trials were primarily conducted to allow comparisons to be made with the reconstruction errors obtained when using the "classical" interpolation-based (Hip Box/Knee Box) approach. The influence of calibration plate separation upon reconstruction accuracy was therefore an important aspect of these L-Frame trials, and was investigated by incorporating the three calibration plate configurations shown in Figure 7.12 into the frame's design (i.e. the location of the outer plates can be adjusted in 25 mm increments). As a consequence, size and weight restrictions were not major design criteria.

- Robot Compatibility:

  Having opted to implement the three calibration plate configurations, the resulting weight of the L-Frame precluded its use in conjunction with the prototype MEDROSA manipulator. Supporting the L-Frame in the upright position depicted in Figure 7.10, also proved to be difficult. The test results supplied in this section therefore apply to an imaging geometry in which the L-Frame is located with its PA view calibration plates acting as a base (i.e. as depicted in Figure 7.12).

As an interesting research exercise, the initial set of tests involving this L-Frame were performed using software modules which were equivalent to those used during the original Hip Box study. It was hoped that this approach would effectively demonstrate the pitfalls associated with implementing an extrapolation-based reconstruction scheme in conjunction with a non-linear imaging system. Accordingly, image analysis was performed by a program (*L_Adapt.pas*) which implemented a nearest (integer) pixel adaptive thresholding scheme. The photogrammetric reconstruction process was then performed using a global interpolation program (*L_Prog.pas*) and its support units (*HB_Types.pas*, *L_Amat.pas*, *HBPseudo.pas*, *L_Focal.pas* & *L_Surg.pas*). The addition of a "Select frame configuration" menu option into program *L_Prog.pas*, allowed all three calibration plate configurations to be accommodated by a single set of software modules.

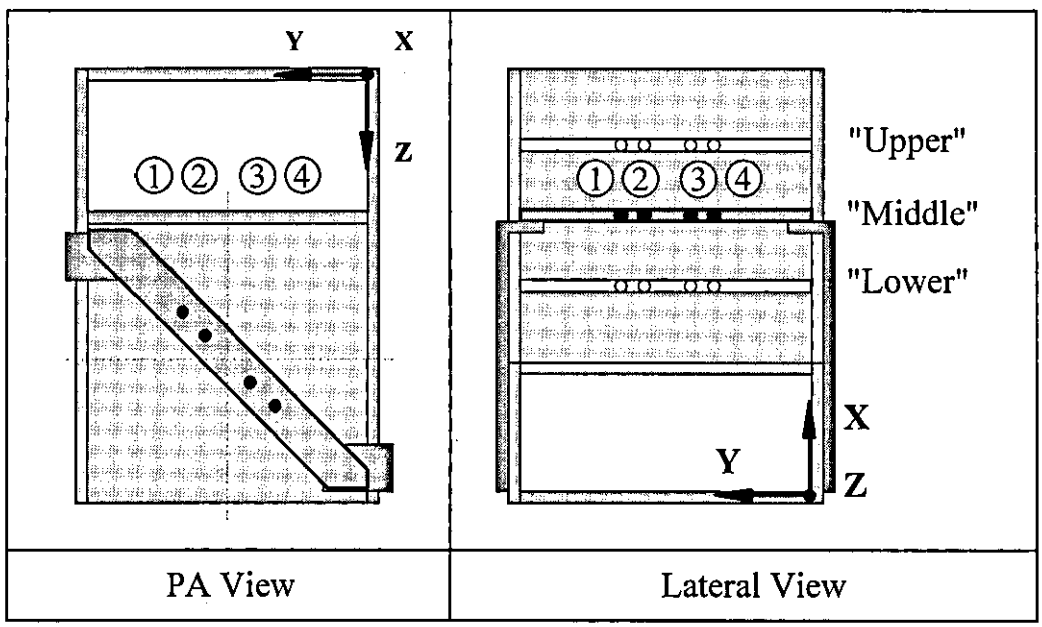


Fig. 7.14 : The Three Test Object Locations in Relation to the L-Frame

By manufacturing an appropriate pair of brackets (refer to Figure 7.13), it was possible to use the existing Perspex test object in conjunction with the L-Frame. As indicated in Figure 7.14, an adjustment on these brackets allows three different test object positions ("middle", "upper" and "lower") to be obtained. Thus in total, nine different test object/calibration plate setting combinations were evaluated.

Table 7.2: Typical Global Reconstruction Data (L25M Configuration)

Test Point	X	Y	Z	Abs. Error
1	135.51 (137.5)	147.85 (147.5)	99.90 (103.5)	4.13
2	134.92 (137.5)	130.38 (130.0)	118.50 (121.0)	3.61
3	133.84 (137.5)	96.02 (95.0)	153.26 (156.0)	4.68
4	133.34 (137.5)	77.90 (77.5)	170.81 (173.5)	4.97
			Mean Absolute Error =	4.35

Table 7.3 : Typical Global Reconstruction Data (L50M Configuration)

Test Point	X	Y	Z	Abs. Error
1	161.09 (162.5)	147.53 (147.5)	125.86 (128.5)	2.99
2	160.51 (162.5)	129.94 (130.0)	144.30 (146.0)	2.62
3	160.10 (162.5)	95.89 (95.0)	179.46 (181.0)	2.99
4	159.45 (162.5)	77.59 (77.5)	197.00 (198.5)	3.40
			Mean Absolute Error =	2.99

Table 7.4 : Typical Global Reconstruction Data (L100M Configuration)

Test Point	X	Y	Z	Abs. Error
1	211.78 (212.5)	146.81 (147.5)	176.86 (178.5)	1.92
2	211.17 (212.5)	129.63 (130.0)	195.23 (196.0)	1.39
3	211.18 (212.5)	95.69 (95.0)	229.90 (231.0)	1.85
4	211.32 (212.5)	77.94 (77.5)	247.14 (248.5)	1.85
			Mean Absolute Error =	1.76

Note: Figures in Brackets = Reference (Expected) Coordinates.

Tables 7.2-7.4 provide typical test results for the three calibration plate configurations when the test object is in the "middle" location. From this sample data it can clearly be seen that without effective compensation for image distortion, the extrapolation-based L-Frame reconstruction process leads to unacceptably high errors. The influence of calibration plate separation upon reconstruction accuracy is also self-evident, with improved results being obtained as the separation is increased.

In the case of the L25 configuration (25 mm plate separation), repeated test results also indicated that the accuracy of the reconstruction process varied considerably throughout the measurement volume. With the test object in the "lower" position, the mean absolute error for the four test object points was typically 3.2 mm. However, the corresponding value for the "upper" test object position was typically 5.6 mm. Fortunately, this measurement "instability" was not encountered when using the L50 and L100 frame configurations.

7.4.2 Follow-Up L-Frame Investigation.

Having demonstrated the problems associated with extrapolation-based reconstruction, a second batch of L-Frame compatible software was developed with a view to improving the accuracy of the calibration process. This enhanced software employs a customised version (*L_Correl.pas*) of the statistical image analysis routines, described in Section 5.4.1, to obtain the image coordinates of the L-Frame fiducials to sub-pixel precision. The photogrammetric reconstruction process is then performed using a modified version of the localised interpolation software described in Section 5.4.2. For reasons of convenience, a separate version of this reconstruction software was developed for each of the three calibration frame configurations. The relevant programs are saved with the format *L*Prog.pas*, where "*" is shorthand for 25_, 50_ or 100. The same format was also used for the appropriate linked units (*LLocType.pas*, *L*Foc.pas* & *L*Surg.pas*).

In addition to introducing compensation for the effects of image distortion, and improved image analysis routines, a number of modifications were also made to the actual photogrammetric reconstruction technique. When applied to the L-Frame, the existing reconstruction scheme (as described in Section 4.4.1) takes the form shown in Figure 7.15. To briefly recap, the 3-D location of a target point (T) is obtained, with respect to the calibration frame's coordinate system, by reconstructing two "intersecting" lines-of-sight. The first of these lines-of-sight is derived by using the measured (pixel) coordinates of the target point in the PA view, to calculate the calibration plate intersection points PA/X and PA/S. Applying the same method to the

lateral view then allows the intersection points LatS and LatX to be calculated, thus defining the lateral view line-of-sight. Unfortunately, as demonstrated by the results presented in the previous section, when the XRII and source plates are close together, large errors can be introduced into the line-of-sight calculations.

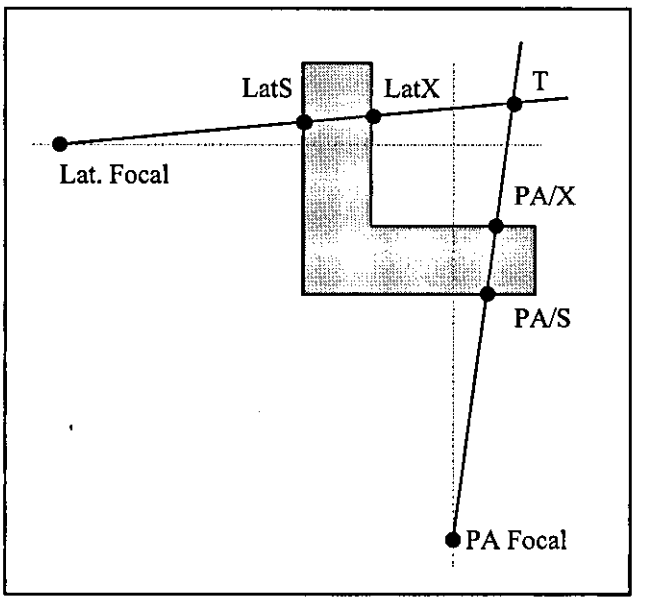


Fig. 7.15 : The Original L-Frame Reconstruction Scenario.

This problem is best illustrated by Figure 7.16, which depicts the PA line-of-sight reconstruction process. Acknowledging that the image analysis process will inevitably contain errors, the XRII and source plate intersection points, calculated from the measured (pixel) target point coordinates, will obviously not lie on the true line-of-sight passing through the target point (T). However, a given image measurement error (in pixels), produces a calibration plate intersection point error (in millimetres), whose magnitude depends upon the image magnification factor (i.e. the pixel-to-mm ratio).

The implications of this relationship with regard to the L-Frame are as follows. Since the three frame configurations all employ the same XRII calibration plate location, the error associated with the plate intersection point PA/X is common to all three scenarios. As such, it is the error associated with the source plate intersection point (PA/S) that accounts for the observed differences between the reconstruction accuracies of the three configurations. Since the pixel-to-mm ratio decreases as the

distance from the X-ray source increases, reducing the plate separation increases the error associated with point PA/S. As depicted in Figure 7.16, the L25 configuration therefore introduces a larger error into the line-of-sight calculations than the L100 configuration.

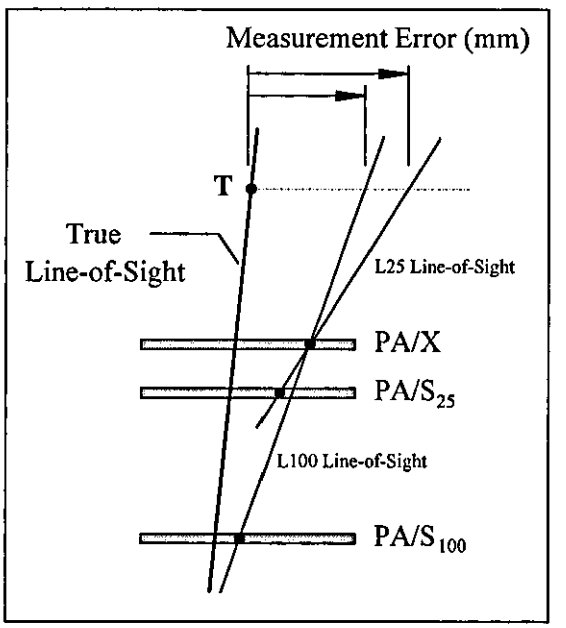


Fig. 7.16 : The Influence of Plate Separation Upon Measurement Errors.

In an attempt to reduce the influence of the point PA/S and its lateral view counterpart LatS upon reconstruction errors, a new method of calculating the PA and lateral view lines-of-sight was introduced. With reference to Figure 7.15, this new approach uses the relevant effective focal points (PA Focal and Lat. Focal) rather than the points PA/S and LatS to define the lines-of-sight. The PA line-of-sight is therefore defined by the points PA/X and PA Focal, and the lateral line-of-sight by the points LatX and Lat. Focal. In terms of software, implementing this change required the units *L*Foc.pas* to be modified so that both effective focal points were calculated (i.e. previously only Lat. Focal was calculated).

However, having made this change, it is essential that the estimated locations of the two effective focal point locations are accurate. As outlined in Section 4.4.3.2, these estimates are obtained by determining the point of closest approach for a reconstructed line-of-sight bundle. Since each of these lines-of-sight is derived using a point in the

XRII plate, and a corresponding point in the source plate, the influence of plate separation upon reconstruction errors cannot be totally eliminated. In particular, when using the L25 configuration, it was found that the focal point calculations were very sensitive, in relation to the number of lines-of-sight in the bundle and their distribution around the field-of-view. By way of compensation, the maximum number of lines-of-sight used to calculate the focal point locations was therefore increased from sixteen to twenty-five.

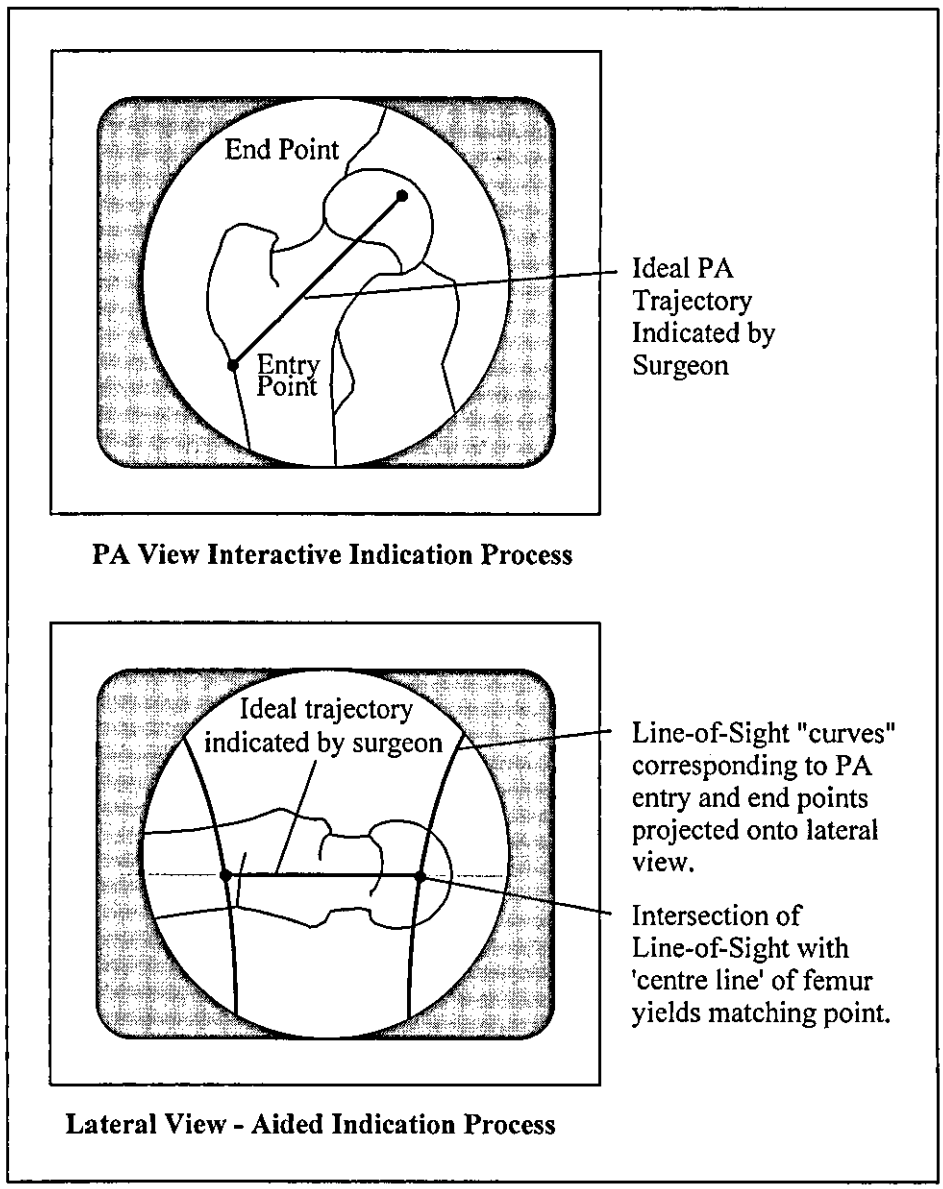


Fig. 7.17: Computer-Assisted Planning of the Guide-Wire Trajectory

The final software improvement that is worthy of comment, relates to the interactive trajectory planning scheme summarised by Figure 7.17. Having selected a suitable

entry point and an end point in the PA view, in order to define a three-dimensional drilling trajectory, the surgeon is required to select the corresponding points in the lateral view. As outlined in Section 4.4.3.2, this image matching process can be greatly simplified by projecting the relevant reconstructed PA lines-of-sight into the lateral view.

The location of each of these "image matching" lines was originally obtained by calculating two intersection points in the lateral view source plate. A straight line was then drawn through the corresponding image (pixel) locations. However, when projecting a 3-D line-of-sight into a fluoroscopic image, the effects of image distortion must also be taken into account. In other words, as a result of image distortion, the projected lines-of-sight become image matching curves. The units *L*Surg.pas* were therefore modified to allow multiple lateral view source plate intersection points, corresponding to several points on the relevant PA lines-of-sight, to be calculated. The desired image matching "curves" are then obtained by drawing line segments between the corresponding image (pixel) locations.

When the L-Frame/test object trials described in Section 7.4.1 were repeated, this time using the enhanced (second generation) software, reconstruction errors of the magnitude supplied in Tables 7.5-7.7 were obtained. Examination of these results allows the following conclusions to be made. Although an appreciable improvement in reconstruction accuracy was obtained, in relation to the guide wire insertion process, the L25 configuration still produced unacceptably high errors. However, the stability of the results did improve across the measurement volume when applying the new software (typical upper and lower test object results were 2.6 mm and 2.4 mm respectively).

By contrast, both the L50 and L100 configurations lead to reconstruction errors which are of the order of 1 mm or better. As indicated by Table 7.8, the average "mismatch" (accuracy indicator) values obtained when using the L50 and L100 configurations are also in the submillimetre range. Thus, although these results are not as good as those obtained when using the (interpolation-based) Hip Box, they are perfectly adequate in relation to the accuracy requirements of the guide wire insertion process.

Table 7.5: Typical Localised Reconstruction Data (L25M Configuration)

Test Point	X	Y	Z	Abs. Error
1	140.28 (138.5)	148.85 (148.5)	103.74 (103.5)	1.83
2	140.60 (138.5)	131.62 (131.0)	121.90 (121.0)	2.37
3	140.39 (138.5)	96.97 (96.0)	156.65 (156.0)	2.22
4	141.25 (138.5)	79.38 (78.5)	174.95 (173.5)	3.40
			Mean Absolute Error =	2.46

Table 7.6 : Typical Localised Reconstruction Data (L50M Configuration)

Test Point	X	Y	Z	Abs. Error
1	164.19 (163.5)	148.19 (148.5)	128.31 (128.5)	0.78
2	163.87 (163.5)	131.20 (131.0)	146.02 (146.0)	0.42
3	164.39 (163.5)	96.41 (96.0)	180.96 (181.0)	0.98
4	164.21 (163.5)	78.94 (78.5)	198.43 (198.5)	0.84
			Mean Absolute Error =	0.75

Table 7.7 : Typical Localised Reconstruction Data (L100M Configuration)

Test Point	X	Y	Z	Abs. Error
1	213.40 (213.5)	148.14 (148.5)	178.97 (178.5)	0.60
2	213.29 (213.5)	131.36 (131.0)	196.66 (196.0)	0.79
3	213.70 (213.5)	96.25 (96.0)	231.44 (231.0)	0.54
4	213.89 (213.5)	78.92 (78.5)	248.66 (248.5)	0.59
			Mean Absolute Error =	0.63

Table 7.8 : Average "Mismatch" Values.

L-Frame Configuration	Global Interpolation	Localised Interpolation
L25M	1.89 mm	1.35 mm
L50M	1.11 mm	0.72 mm
L100M	0.44 mm	0.38 mm

Note: Figures in Brackets = Reference (Expected) Coordinates.

However, with regard to the practicalities of intraoperative use, the L50 configuration offers a number of advantages (i.e. reduced size, weight and sensitivity to angular misalignment) over the L100 configuration. Future attempts to implement the robotic-assisted guide wire insertion process should therefore apply a lightweight, robot-compatible version of the L50 frame.

7.5 Summary of Chapter 7.

Having outlined that hip fractures currently place a huge burden upon orthopaedic resources, as well as detailing future predictions for the growth of the hip fracture "epidemic", this chapter has demonstrated that in relation to their internal fixation a need exists for improved surgical technique. Following an analysis of existing surgical practices, a critical stage (i.e. guide wire insertion) of the internal fixation process has accordingly been targeted for computer/robotic assistance. The second half of the chapter therefore describes how the X-ray photogrammetry based registration techniques developed by this thesis, have been adapted to meet the requirements of the guide wire insertion process.

The cornerstone of this proposed "hip fracture" registration scheme is the use of a custom-made X-ray calibration frame, referred to as the "L-Frame". In order to achieve compatibility with both the existing biplane (PA and lateral) fluoroscopic examination process, and patient anatomy, this L-Frame supports an extrapolation-based photogrammetric reconstruction process. However, laboratory tests have shown that the loss of accuracy associated with this deviation from the "classical" photogrammetry approach, is not significant in relation to the accuracy requirements of the guide wire insertion process. The L-Frame therefore represents a satisfactory "compromise" solution, on the one hand providing reconstruction errors of the order of 1 mm (or better), while at the same time offering significant advantages in terms of intraoperative ease of use.

8. CONCLUSIONS.

8.1 Introduction.

Owing to the "closed" nature of the osteosynthesis procedures under investigation, intraoperative registration of the MEDROSA system can only be established by quantifying the existing (biplane) fluoroscopic examination process. With a view to providing clinically acceptable solutions to the research problems identified in Section 1.4.1, the current thesis has therefore proposed the use of an *intraoperative* registration technique based upon *X-ray photogrammetry* principles. Given that photogrammetric reconstruction is traditionally a very time-consuming process, the key to the successful implementation of this proposal is the acquisition and rapid analysis of digital X-ray "calibration" images.

At the beginning of the research period it was demonstrated that digital X-ray images could be acquired using a PC-based frame grabber card linked to a mobile C-arm fluoroscopy unit. As discussed in Chapter 4, limited access to these fluoroscopy units then dictated that the original investigation of the registration process was performed using a CCD camera-based system. However, as outlined in Chapter 5, when unrestricted access to a C-arm unit did become available, the transition from a camera-based to an X-ray-based system was made with relative ease. With the benefit of hindsight, it is therefore possible to state that the use of CCD cameras, to simulate the (external) imaging geometry of the C-arm unit, was indeed a valid proposal.

Having developed a generic *digital* X-ray photogrammetry system, the relevant software modules and the calibration frame design have subsequently been customised, in order to meet the specific requirements of the osteosynthesis procedures targeted by the MEDROSA Project. As indicated in Chapters 6 and 7, although this research was restricted to *in vitro* laboratory-based studies, it has nevertheless been possible to fully address criteria such as anatomical compatibility and compliance with surgical sterility. The desired stage of development, that is to say the ability to perform advanced pre-clinical trials, has therefore been reached.

Acknowledging these general statements about research methodology, and the level of progress that has been made, the remaining sections of this chapter provide an in-depth evaluation of the X-ray photogrammetry based registration technique, in relation to the original objectives of the thesis (as stated in Section 1.4). Where appropriate, additional evaluation criteria, which were not fully appreciated at the start of the research period, are also introduced. Accordingly, Section 8.2 initially analyses the performance of the technique in relation to the four main stages of the intraoperative registration process: *image analysis*, *internal calibration* (compensation for image distortion), *external calibration* ("registration"), and *trajectory planning*. Sub-section 8.2.5 then covers the more subjective topic of "operating room compatibility". Finally, Section 8.3 makes a number of recommendations in relation to issues raised by the current thesis and the future development of the MEDROSA Project.

8.2 Evaluation of The X-Ray Photogrammetry Approach.

When attempting to analyse the X-ray photogrammetry based registration scheme as a whole, an initial set of evaluation criteria can obviously be generated by drawing comparisons with the intraoperative registration strategies reviewed in Chapter 3. In relation to some of the orthopaedic applications discussed in Section 3.3, the proposed use of a robot-mounted X-ray calibration frame can be seen to offer several performance-related benefits. The most significant of these benefits are highlighted by the registration graph for the MEDROSA system (Figure 4.3), and can be summarised as follows:

- The registration technique is *non-invasive.*
- A preoperative *CT* or *MRI* examination is *not required.*
- *Expensive position-sensing equipment* (e.g. infra-red optical digitisers, mechanical digitising arms, etc.) does *not* have to be *used.*
- Existing operating room equipment (in particular the *C-arm unit*) does *not* have to be *modified.*
- The registration process *will not prolong* the duration of the *surgical procedure.*
- *Existing* surgical and fluoroscopic *practices* are *not altered* significantly.

However, as acknowledged in Section 4.2, intraoperative registration schemes that do not employ real-time position-sensing have one notable limitation. In such cases, having successfully established intraoperative registration, the assumption must then be made that this initial registration "configuration" is subsequently *maintained* for the entire duration of the computer/robotic-assisted procedure. In other words, no relative motion occurs between the patient, intraoperative imaging system(s), position-sensing device(s), or robotic manipulator.

Given that the anaesthetised patient's injured limb is placed in traction, the general consensus of opinion, among the orthopaedic surgeons consulted on the matter, is that intraoperative patient motion is not a significant problem during the "MEDROSA" osteosynthesis procedures. The assumption has therefore been made that the drilling trajectory data, obtained from the calibrated intraoperative images, will still be valid during the invasive stages of the procedure. However, in the long-term, this assumption will obviously have to be fully investigated during clinical trials. Ideally, some form of patient motion monitoring should also be routinely implemented as a safety backup. In the unlikely event of motion being detected, the registration process could then be quickly repeated, thus avoiding significant errors in the robotic-assisted drilling process.

8.2.1 The Image Analysis Problem.

In order to perform the image distortion correction and photogrammetric reconstruction processes, two sets of accurate input data are required: the real-world (millimetre) coordinates of the fiducial markers in the X-ray calibration frame, and their corresponding (pixel) coordinates in the intraoperative images. Given the obvious time constraints associated with an intraoperative application, in particular the need to avoid prolonging the surgical procedure, the extraction of this image data must be performed with the minimum possible level of user intervention. A semi-automatic digital image analysis scheme has therefore been implemented to rapidly "scan" the four imaged fiducial patterns of each X-ray calibration frame.

The complexity of the image analysis task is obviously related to the informational content of the image being scanned. As a consequence, from the software perspective the image analysis process can be significantly simplified through the provision of a carefully considered fiducial pattern. With this goal in mind, and the need to uniquely identify every fiducial marker appearing within the field-of-view, reference markers (i.e. a centre marker(s) and an orientation marker) have been provided with differing sizes and shapes to the "standard" fiducial markers. Interactively indicating the locations of these reference markers allows the "predicted" locations of the remaining fiducials to be calculated. By automatically scanning areas of interest at these "predicted" locations, it is therefore possible to rapidly extract the locations of the vast majority of the imaged fiducial markers.

Having developed a scanning routine which ensures that the registration process does not significantly delay the surgical procedure, reliability and measurement accuracy become the important evaluation criteria. The reliability of an image analysis scheme is primarily related to the effectiveness of the relevant segmentation technique. The limitations of the segmentation scheme which is currently being used were previously discussed in Section 5.4.1. The most commonly encountered problem is the situation in which the grey level of the imaged fiducial does not differ significantly from that of the image's background. The obvious solution to this problem is to increase the contrast within the image, by using fiducial markers manufactured from a material (e.g. tantalum) which is known to produce a very dark image when X-rayed. A "manual override" technique, which allows the interactive indication of any reference markers not located by the automatic segmentation/measurement scheme, has also been implemented to provide a backup. As such, it should never be necessary to retake a fluoroscopic image in order to facilitate the image analysis process.

In practice, unless a second scheme is available for comparison purposes, it is very difficult to evaluate the accuracy of an image analysis (sub-pixel measurement) scheme. However, since the automatic measurement process works by analysing a circular area of interest (AOI), and successive measurement attempts tend to involve a slightly different AOI, it is possible to estimate the "repeatability" of the technique.

Accordingly, for the case of a "black" fiducial on a clear "white" background, the relevant value for the image analysis scheme used by this thesis is 0.33 pixels (approx. 0.08 mm). However, as indicated in Section 5.4.1, it must be noted that this value is influenced by both the size and the location of the AOI. The size of the AOI must therefore be kept as small as possible, in order to reduce the detrimental influence of background AOI pixels upon the segmentation/measurement process.

8.2.2 The Image Distortion Problem.

The non-linear imaging characteristics of the X-ray image intensifier tube are well documented in the relevant radiology literature. Having become aware of the pincushion distortion effect at an early stage of the investigation, it was therefore possible to empirically evaluate this easily observable phenomenon during the preliminary X-ray based trials discussed in Section 5.2. Given that the resulting image distortion was obviously too significant to be neglected, "classical" distortion *correction* schemes were then evaluated as a means of allowing accurate data to be extracted from intraoperative fluoroscopic images.

In view of the unconstrained imaging geometry of a mobile C-arm unit, the mathematical modelling approach to distortion correction (Section 2.4.2.3) is not well suited to the MEDROSA applications. Owing to the non-symmetrical (localised) nature of the observed image distortion, the use of a global polynomial correction function (Section 2.4.2.1) is also ill-advised. It was therefore initially intended that a calibration plate would be used in conjunction with local affine correction functions (Section 2.4.2.2), in order to implement a *preoperative* distortion correction scheme. Having used the calibration plate to determine the magnitude of the distortion effect, intraoperative images would then be "dewarped" by applying the corresponding correction functions.

Unfortunately, it has subsequently been determined that this off-line approach represents an over-simplification of the true image distortion problem. As discussed in Section 2.4.3, the trajectories of the electrons inside the image intensifier tube can be altered by the presence of a strong magnetic field. When the arc of a mobile C-arm

unit is rotated through ninety degrees, the resulting change in the orientation of the image intensifier tube, in relation to the geomagnetic field, has been shown (Section 5.4.2) to produce a significant (2-3 mm) *image shifting* effect. A characteristic *S-shaped warping* can also be observed, as a result of a geomagnetic image rotation effect being partially compensated for by the "magnetic" shield around the sides of the image intensifier tube.

In addition to the distortion effects caused by the Earth's magnetic field, the possibility of additional magnetic influences in the operating room cannot be overlooked. The image intensifier housing of the C-arm unit is by necessity frequently placed in close proximity to large metallic objects (e.g. the operating table, surgical instrumentation, etc.). Electromagnetic interference caused by operating room equipment is also widely acknowledged as being a significant problem. In relation to the mobile C-arm unit, the magnetic-related image distortion effects are therefore highly unpredictable, and as such, the use of an off-line correction scheme is *not* appropriate.

Having identified the need to implement *on-line/intraoperative* image distortion "correction", the use of a modified X-ray photogrammetry calibration frame to provide both external and *internal* calibration data has been investigated. Given that an image intensifier calibration plate is not employed to initially determine the magnitude of the localised distortion effects, strictly speaking, in the "classical" sense the resulting technique is not a true distortion *correction* scheme. The term distortion "*compensation*" therefore provides a more appropriate description of this novel approach.

The main advantage offered by this technique is that data from the actual intraoperative (PA and lateral) images is used to perform the compensation process. In other words, assumptions do not have to be made with regards to either the magnetic influences in the operating room, or the internal geometry of the image intensifier tube. Each individual fluoroscopic image is therefore calibrated using data that is valid for the specific imaging geometry used at the time of its acquisition. In

addition, applying this intraoperative compensation scheme also avoids the need to perform a preoperative internal calibration process.

In practice, the actual distortion compensation process is performed using a localised interpolation technique, which is similar to the local affine correction functions described in Section 2.4.2.2. When using this type of linear function to approximate a non-linear distortion effect, the errors associated with the approximating function are obviously closely related to the size of the sampled region. In-line distortion compensation therefore requires the provision of a calibration frame that provides more fiducial markers than would generally be required by a standard RSA/X-ray photogrammetry application.

However, while increasing the density of the fiducial pattern reduces the size of the sampled region, leading to improved distortion compensation, it also has implications in relation to the duration of the image analysis (scanning) process. When large numbers of fiducials are used, a subjective point can also be reached where "too much" of the anatomical content of the fluoroscopic image is obscured. A compromise solution must therefore be adopted, with the number of fiducial markers being reduced to the number that guarantees the level of accuracy required by a particular surgical application.

By performing tests with and "without" image distortion compensation, it has been possible to effectively demonstrate a marked improvement in reconstruction accuracy. The original X-ray compatible software implements a global interpolation scheme which has an "averaging" effect in relation to image distortion effects. While this ("without") approach surprisingly leads to acceptable reconstruction values when used in conjunction with the Hip Box (Tables 5.1 - 5.3), the same cannot be said in the case of the L-Frame (Tables 7.2 - 7.4). The ultimate confirmation of the distortion compensation scheme's effectiveness is therefore provided by the fact that, in both the Hip Box (Tables 5.4 - 5.7) and L-Frame (Tables 7.5 - 7.8) cases, sub-millimetre reconstruction accuracy has been obtained in spite of the presence of both magnetic and pincushion image distortion effects.

8.2.3 The Intraoperative Registration Problem.

From the mathematical perspective, having compensated for image distortion effects, the external calibration (i.e. intraoperative registration) process for a C-arm based system is identical to that of a conventional (film-based) X-ray photogrammetry application. Knowledge of the RSA photogrammetric reconstruction technique (Section 2.3.2.1), and firsthand practical experience with the Two-Plane camera calibration technique (Section 4.4), has therefore been applied to develop an intraoperative registration scheme based upon the reconstruction of an "intersecting" pair of lines-of-sight.

When this registration scheme is combined with the previously discussed image distortion compensation technique, the resulting *correction-calibration* approach offers two distinct advantages:

- By separating the distortion compensation and photogrammetric reconstruction processes into two independent stages, *linear mathematical techniques* can be applied to solve both problems.

- In comparison to global techniques employing least-squares optimisation (e.g. the DLT), the *computational effort* can be *significantly reduced* by implementing the distortion compensation scheme in such a way that the local interpolation functions do not have to be pre-computed.

However, in spite of the mathematical similarities, in practice, there is one obvious difference between conventional and fluoroscopy-based X-ray photogrammetry systems: the unconstrained nature of the C-arm imaging geometry (illustrated in Figure 5.7). In order to apply the correction-calibration approach described above to a C-arm based system, it has therefore been necessary to introduce practical measures which allow a variable imaging geometry to be accommodated.

The MEDROSA system is targeted at fluoroscopically-guided bone drilling processes, which are performed at intermediate stages of femoral osteosynthesis procedures. As a consequence, it must be assumed that the optimal position of both the patient and the

mobile C-arm unit will already be determined, prior to the introduction of the robot-mounted X-ray calibration frame. In view of these given constraints, when introducing the calibration frame into the field-of-view, a less than perfect alignment with the C-arm unit may have to be accepted.

Owing to the effects of central projection, in order to guarantee that sufficient fiducials are introduced into the field-of-view, the distance from the X-ray source to the calibration frame (i.e. the range) must be "known". Fortunately, by studying established fluoroscopic practices (illustrated in Figures 6.16 and 7.6), it has been possible to estimate typical range values for the two femoral osteosynthesis procedures under investigation. Using these estimates as design criteria, appropriate (different) fiducial pattern grid spacings can be determined for the XRII and source calibration plates. Offsetting this customised pair of fiducial patterns, then ensures that a "usable" calibration image is still acquired up to 100 mm from the ideal range value.

The use of an X-ray calibration frame in conjunction with a mobile C-arm unit also incurs problems in relation to angular orientation. Ideally, the imaged planes of the calibration frame should be parallel to the input window of the image intensifier unit. However, in practice, it is unlikely that this perfect alignment will be achieved, and as a consequence, under certain conditions the fiducial patterns in the two imaged calibration plates can become superimposed. Since this eventuality severely complicates the image analysis process, every attempt must be made to reduce the likelihood of its occurrence.

Of the solutions to this frame tilt problem developed by the current thesis, the use of a calibration frame consisting of closely spaced pairs of calibration plates (i.e. the L-Frame) is undoubtedly the most successful. Reducing the plate separation in this way has been shown to significantly reduce the sensitivity of the calibration frame to angular misalignment. The L-Frame configuration therefore represents a more user friendly design than the Hip Box/Knee Box configurations. However, accepting that these "Box" configurations yield better reconstruction accuracy, several measures have been introduced to alleviate the angular orientation problem when using these

frames. Offsetting and customising the fiducial patterns obviously has a positive impact in relation to angular misalignment. The use of a simple alignment device attached to the housing of the image intensifier unit has also been investigated as a potential solution. By adjusting the calibration frame until it comes into contact with this C-arm mounted device, the desired range and angular alignment can be achieved. However, whether or not the use of such a device is warranted, is something that will have to be investigated during future clinical trials.

8.2.4 Trajectory Planning.

Trajectory planning is a procedure-specific process, and as such, the relevant techniques for distal interlocking and guide wire insertion must be evaluated separately. However, before making these independent evaluations it is worth noting that in both cases, trajectory planning has been achieved without the need to perform a full image dewarping process. That is to say, by applying photogrammetric techniques, such as overlaying reconstructed lines-of-sight to aid image matching, it has not been necessary to reconstruct and display "true" (undistorted) fluoroscopic images. A global dewarping process of this nature involves grey scale interpolation, in order to restore grey scale continuity, and generally takes minutes rather than seconds. Given that a pair (PA and lateral) of images would have to be dewarped, a considerable intraoperative delay has therefore been avoided by only applying distortion compensation to individual points of interest.

As discussed in Section 6.5, from the trajectory planning perspective, the distal interlocking stage of the intramedullary nailing procedure is a prime example of an orthopaedic application that can significantly benefit from computer-assistance. Having initially quantified the fluoroscopic examination process using the Knee Box, accurate coordinate data can then be extracted from the relevant images of the nail/femur combination. Given that the object of interest (i.e. the nail) is a man-made metallic implant providing well-defined contours, this data extraction process can be performed with minimal user interaction. As such, the image analysis process associated with the "perfect circles" approach outlined in Section 6.5.1 is relatively straightforward.

However, as detailed in Section 6.5.2, the biggest incentive for any computer/robotic-assisted distal interlocking application is the possibility of eliminating the time-consuming "perfect circles" C-arm alignment process. A trajectory planning scheme, based upon the determination of a single 3-D point (i.e. the centre of the nail hole) and an angle of rotation (about the long axis of the nail), has therefore been proposed with a view to achieving this goal. In practice, it has been shown that this angle of rotation can be calculated by analysing the oval-shaped projection of the distal nail hole in the "non-perfect" lateral view. To be more precise, an iterative optimisation process is performed until a "characteristic" measurement of the model's hole, matches the corresponding measurement of the imaged hole in the actual nail. The (unknown) rotation of the actual nail is then equated to the known rotation of the model.

The critical aspect of this new approach is the selection of a satisfactory "characteristic" measurement. Ideally, as the nail is rotated about its long axis, the magnitude of the "characteristic" measurement should change in a predictable manner that allows a unique determination of the rotation angle. In order to overcome the effects of central projection (variable magnification), the "characteristic" measurement must also be scale invariant. Unfortunately, having proposed the use of an area-based "characteristic" measurement, it has been independently determined that this option does not provide a unique solution to the rotation problem. However, a more appropriate "characteristic" measurement, the angle between the major axis of the oval-shaped hole profile and the long axis of the nail, has subsequently been found to provide the desired unique solution **[224]**. It has therefore been possible to show that the X-ray photogrammetry-based registration techniques developed by this thesis, allow the distal interlocking process to be performed without initially having to acquire the "perfect circles" image.

Turning now to the guide wire insertion problem described in Section 7.4, it can hopefully be appreciated that the lack of man-made implants, within the field-of-view, makes automation of the trajectory planning process an extremely difficult task. Although an automatic bone contouring scheme could theoretically be implemented,

given the relative simplicity of the desired drilling trajectory, the prolonged development process associated with such a research effort is not deemed to be worthwhile. Instead, an interactive indication process has been adopted, which requires the selection of two (i.e. an entry point and a target point) points in both the PA and the lateral view.

Although this interactive indication approach obviously leaves scope for human error, because the accuracy requirements of the actual surgical procedure are not particularly severe, tests have shown (Tables 7.5 - 7.8) that the resulting errors can be tolerated. The use of reconstructed line-of-sight overlays, to guide the image matching process, has also been found to be an effective way of alleviating the depth perception problem currently experienced by orthopaedic surgeons during hip fracture fixation procedures. However, a more realistic demonstration of the scheme's effectiveness has been provided by independent laboratory trials involving the prototype MEDROSA manipulator **[192]**. Significantly, these trials show that the combined registration/trajectory planning/robotic system error lies within acceptable limits for the guide wire insertion process.

8.2.5 Compatibility with the Operating Room Environment.

The most important aspect of "operating room compatibility" is obviously the need to comply with surgical sterility. As explained in Section 6.5, the proposed use of an X-ray calibration frame in close proximity to the patient (surgical site), requires a frame design that can be either isolated or sterilised. For reasons of simplicity, isolation of the calibration frame is the preferred option. Covering the frame in sterile plastic draping, a technique which is currently employed to isolate the X-ray source and image intensifier housings of the C-arm unit, prevents contamination by isolating the frame from the "sterile" areas of the operating room. However, if the use of isolation draping proves to be problematic, a preliminary investigation has also shown that an "autoclave-compatible" calibration frame can be manufactured from materials with appropriate thermal, corrosion resistance and water absorption properties.

The remaining aspects of the "operating room compatibility" problem were previously discussed with reference to Figure 1.1 (Section 1.4.1), which depicts the existing operating room scenario for an osteosynthesis procedure, as well as the prototype MEDROSA manipulator. When evaluating the X-ray photogrammetry based registration scheme it would therefore seem to be appropriate to do so in relation to the same intraoperative scenario. Figure 8.1 therefore depicts the proposed robotic-assisted surgery scenario for a hip fracture (guide wire insertion) procedure, in an attempt to illustrate both the simplicity of the X-ray calibration frame-based technique, and the interrelationships of the various intraoperative components.

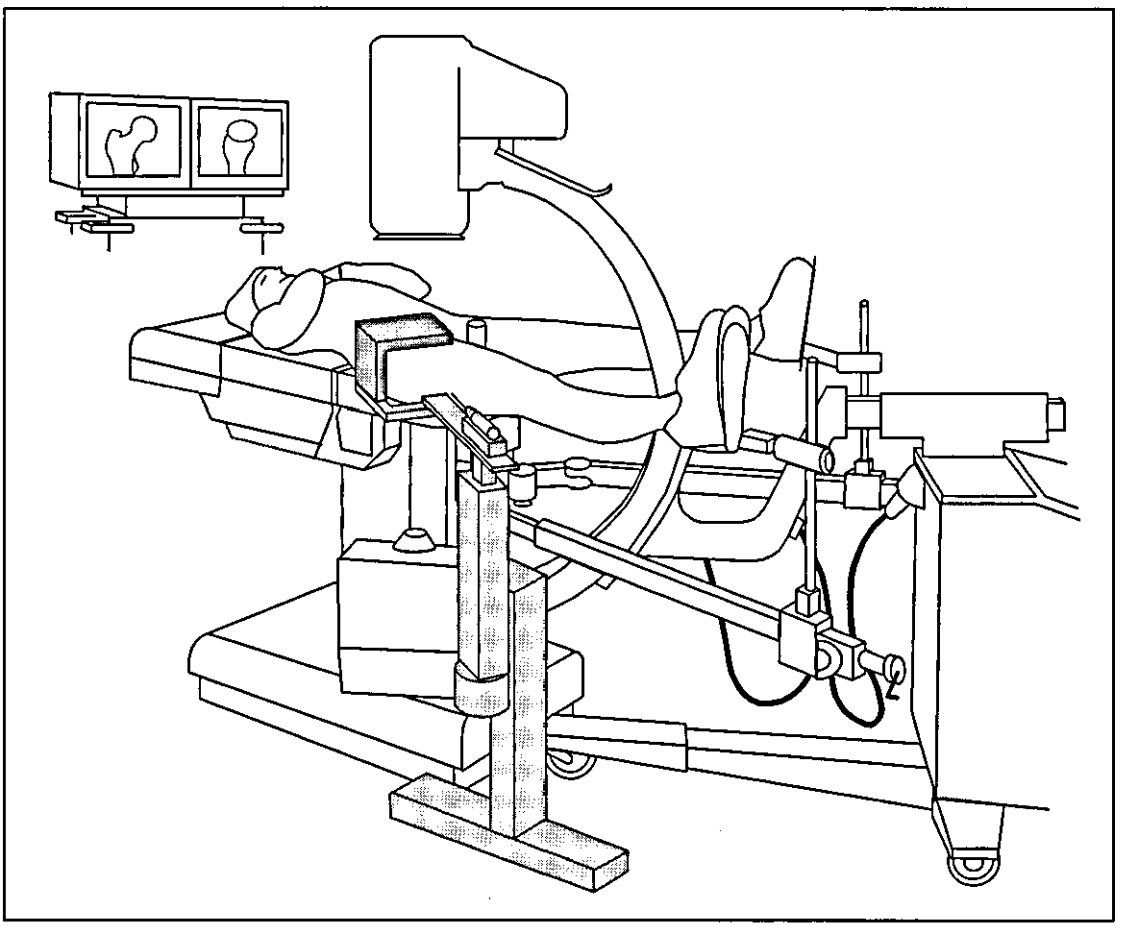


Fig. 8.1 : X-Ray Photogrammetry Based Intraoperative Registration.

It is hopefully self-evident from Figure 8.1 that the need to provide "robot-compatible" drilling trajectories has been satisfied by mounting the X-ray calibration frame on the end-effector of the manipulator. As described in Section 4.2, having developed an appropriate kinematic model for this frame/manipulator configuration,

trajectories defined with respect to the calibration frame's coordinate system can therefore be transformed into the manipulator's coordinate system. The mounting bracket for the calibration frame has also been designed in such a way as to ensure that the rotation of the C-arm unit's arc is not obstructed. However, as pointed out in Section 5.2, this "C-arm compatibility" only applies to the biplane (PA and lateral) imaging geometry; attempts to implement a convergent X-ray photogrammetry system having been abandoned for a number of practical reasons.

Figure 8.1 depicts the use of an orthopaedic fracture table, and as such, the patient's lower limbs are supported in "mid air". It can therefore hopefully be appreciated that when using the Knee Box to perform the distal interlocking procedure (in the manner illustrated in Figure 5.2), achieving anatomical and "operating" table compatibility is not a problem. Unfortunately, the same cannot be said for the guide wire insertion process performed during hip fracture procedures. An alternative to the "classical" technique of convergent X-ray photogrammetry of the hip has therefore had to be developed. As illustrated in Figure 8.1, this new approach involves the use of an L-shaped calibration frame which supports an extrapolation-based reconstruction process.

As this "L-Frame" obviously represents a deviation from "classical" (interpolation-based) photogrammetry theory, test results (Tables 7.5 -7.8) indicate that there is an associated loss of reconstruction accuracy in relation to results obtained with the "Box" configuration (Tables 5.4 - 5.7). It must also be noted that due to the narrower plate separation, a given machining error during the manufacturing process leads to a much greater reconstruction error in the case of the L-Frame. Nevertheless, it has still been possible to obtain sub-millimetre reconstruction accuracy when using the L-Frame. The L-shaped configuration also exhibits a number of distinct advantages over the "Box" configuration in relation to ease of use:

- The narrower calibration plate separation makes the L-Frame less sensitive to angular misalignment.
- The reduced weight and bulk of the L-Frame makes it more "robot-compatible".

- Compliance with sterility criteria (via isolation draping) is easier to achieve.
- The L-frame does not have to be assembled around the patient's anatomy, thus leading to shorter set-up times.

Significantly, the L-Frame is also more versatile in relation to anatomical compatibility, and as such, an interesting research question is raised: *can the L-Frame also be used to perform distal interlocking*? If it can be shown that the accuracy afforded by the L-Frame is adequate for the demanding distal interlocking process, its use instead of the more accurate Knee Box may (or may not) prove to be justifiable on ergonomic grounds.

8.3 Recommendations.

Having carried out the *in vitro* laboratory-based research documented in this thesis, it is obviously possible to make a number of informed recommendations about the future development of the vision-related aspects of the MEDROSA system. The following section therefore addresses issues relating to the user interface requirements of the system, as well as the integration of the vision and robotic system components. However, at a more general level, it is also important that the following points are appreciated:

- "Non-Femoral" Osteosynthesis Applications.

  In addition to femoral fracture procedures, the X-ray photogrammetry based registration scheme developed by this thesis is equally applicable to other osteosynthesis procedures involving the extremities. The scope of the MEDROSA Project could therefore quite easily be expanded to also include procedures such as locked humeral or tibial nailing.

- Scope For Industrial Applications.

  Owing to its relative simplicity, the on-line correction-calibration approach developed by this thesis could also be applied to industrial applications requiring quantitative fluoroscopic imaging (e.g. non-destructive testing).

In order to allow the various system components to be developed in parallel, the prototype MEDROSA system employs two personal computers. The first PC contains the IV120(c) frame grabber card, and is used to perform the vision-related tasks documented in this thesis. The other PC controls the prototype manipulator via an interface with an electronics ("robot controller") cabinet. However, in the long-term, this arrangement will obviously be replaced by a single PC system. Ideally, a vision system and a robotic system, programmed in the same software language, will therefore run in a multi-tasking environment.

Regrettably, although the existing digital image processing system is perfectly adequate for the "system development" role required by the current thesis, in relation to this future MEDROSA scenario, it exhibits major shortcomings:

- Obsolete DOS-based software (Microsoft Pascal 4.0).
- An image file format (.im) which cannot be directly viewed in a "Windows" environment.
- A two-monitor configuration which is not conducive to the development of an effective user interface.

Unfortunately, having opted to use the software library functions supplied with the IV120(c) frame grabber card, to facilitate "shortcuts" in the development process, a "point of no return" has been reached by the current thesis. In other words, switching to a "Windows-compatible" digital image processing system at this late stage would require not only the conversion of large sections of code into another software language, but also the development of new code to replace the JLGenial library functions.

The provision of frame grabber independent software is therefore beyond the scope of the current thesis. However, it is interesting to note that by programming in the Windows environment, the "join" between the image analysis and the photogrammetric reconstruction programs could be hidden using an icon-driven software format. Ideally, it would also be possible to limit user interaction solely to the PC-mouse, thus avoiding the need to isolate/sterilise the computer keyboard.

Switching attention to the more immediate future of the MEDROSA Project, the next logical step towards clinical acceptance is a series of advanced pre-clinical trials of the entire MEDROSA system. Accordingly, in order to perform a full evaluation of the guide wire insertion process, a lightweight "robot-compatible" version of the L-Frame will have to be manufactured. Ideally, the weight of the existing Knee Box design should also be reduced prior to performing distal interlocking trials. As indicated in Section 8.2.5, it may also prove to be highly beneficial to develop a second L-Frame customised to the geometry of the distal interlocking process.

Having manufactured these robot-compatible X-ray calibration frames, it will be possible to perform simulated fluoroscopically-guided drilling procedures on phantom (plastic) femurs containing pre-drilled holes. A more "realistic" set of drilling tests could then be performed on flesh-covered porcine femurs. The results from these tests will then allow the accumulation of errors from the individual system components to be evaluated, with respect to the accuracy requirements of the individual surgical procedures. The participation of orthopaedic surgeons in the testing process will also provide valuable "end user" feedback.

Finally, if access can be gained to a coordinate measurement machine (CMM), the X-ray calibration frames should be surveyed (probed) in order to obtain a more accurate knowledge of the fiducial marker locations. The main motivation for this CMM study would be to check the relative positions and orientations of the individual calibration plates (i.e. the frame's construction). The more straightforward task of determining the relative positions of the fiducials in an individual calibration plate could also be performed. In this way, the location of each individual fiducial could be determined, with respect to the arbitrarily defined calibration frame coordinate system, to a higher level of precision. The corresponding data could then be stored in a file which is subsequently read by the appropriate photogrammetric reconstruction software. Having improved the quality of the input data, a corresponding improvement should then be observed in the accuracy of the reconstruction process.

9. REFERENCES

[1] *Stedman's Medical Dictionary*, 26th edition, Williams & Wilkens, Baltimore, MD, USA, 1995. (ISBN 0-683-07922-0).

[2] **Dandy, D.J.** *Essential Orthopaedics & Trauma*, 2nd Edition, Churchill Livingstone, London, 1993. (ISBN 0-443-04722-7)

[3] **Hewit, J.R.** Mechatronics. In: *Mechatronics Design in Textile Engineering*, M. Acar (ed.), Kluwer Academic Publishers, 1995. (ISBN 0-7923-3204-0).

[4] **Bouazza-Marouf, K., Hewit, J.R.** and **Browbank, I.** Robotic-Assisted Orthopaedic Surgery. *Proceedings of the Joint Hungarian-British International Mechatronics Conference*, Budapest, Hungary, 21-23 September 1994, 589-595.

[5] **Bouazza-Marouf, K., Browbank, I.** and **Hewit, J.R.** Robotic-Assisted Internal Fixation of Femoral Fractures. *Proceedings of the Institution of Mechanical Engineers, Part H - Journal of Engineering in Medicine*, 1995, **209**(H1), 51-58.

[6] **Hewit, J.R., Slade, A.S., Bouazza-Marouf, K.** and **Browbank, I.** Mechatronic Assist Systems for Surgery. *Proceedings of the Second International Workshop: Mechatronics in Medicine and Surgery (MediMec'95)*, Bristol, UK, 6-9 September 1995, 126-144.

[7] **Bouazza-Marouf, K., Browbank, I.** and **Hewit, J.R.** Robot-Assisted Invasive Orthopaedic Surgery. *Mechatronics*, June 1996, **6**(4), 381-397.

[8] **Bouazza-Marouf, K.** and **Browbank, I.** Robotic/Computer Assisted Osteosynthesis. *Proceedings of the Conference on Intelligent Information Systems (IIS'97)*, Grand Bahama Island, Bahamas, 8-10 December 1997, 58-62.

[9] **Browbank, I.** and **Bouazza-Marouf, K.** Robotic-Assisted Repair of Femoral Fractures. *Proceedings of the 29th International Symposium on Robotics (ISR'98)*, N.E.C., Birmingham, 27th April - 1st May 1998.

[10] **Wolf, P.R.** *Elements of Photogrammetry*. McGraw-Hill, New York, 1983. (ISBN: 0-07-066637-7)

[11] **Beyer, H.A.** Calibration of CCD-Cameras for Machine Vision and Robotics. *Proceedings of SPIE - International Society for Optical Engineering*, 1989, **1197**, 88-98.

[12] **Fu, K.S., Gonzalez, R.C.** and **Lee, C.S.G.** *Robotics: Control, Sensing, Vision & Intelligence*. McGraw Hill, New York, 1987. (ISBN: 0-07-022625-3)

[13] **Robinson, M.** and **Ariyaeeinia, A.M.** An Active Co-ordinate Imaging System for Robot Vision. *Proceedings of SPIE - The International Society for Optical Engineering*, 1986, **657**, 144-151.

[14] **Gregory, A.** and **Lipczynski, R.T.** The Three-Dimensional Reconstruction and Monitoring of Facial Surfaces. *Medical Engineering and Physics*, 1994, **16**(3), 249-252.

[15] **Jarvis, R.A.** Perspective on Range Finding Techniques for Computer Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, March 1983, **5**(2), 122-139.

[16] **Martins, H.A., Birk, J.R.** and **Kelley, R.B.** Camera Models Based on Data From Two Calibration Planes. *Computer Graphics & Image Processing*, 1981, **17**(2), 173-180.

[17] **Duda, R.O.** and **Hart, P.E.** *Pattern Classification & Scene Analysis.* John Wiley & Sons, New York, 1973. (ISBN: 0-471-22361-1)

[18] **Abdel-Aziz, Y.I.** and **Karara, H.M.** Direct Linear Transformation from Comparator Co-ordinates into Object Space Co-ordinates in Close-Range Photogrammetry. *Proceedings of the ASP/UI Symposium on Close-Range Photogrammetry,* Urbana, Illinois, USA, 26-29 January 1971, 1-18.

[19] **Hatze, H.** High Precision 3-D Photogrammetric Calibration and Object Space Reconstruction Using a Modified DLT Approach. *Journal of Biomechanics*, 1988, **21**(7), 533-539.

[20] **Gazzani, F.** Comparative Assessment of Two Algorithms for Calibrating Stereophotogrammetric Systems. *Journal of Biomechanics*, 1993, **26**(12), 1449-1454.

[21] **Marzan, G.T.** and **Karara, H.M.** A Computer Program for Direct Linear Transformation Solution of the Collinearity Condition, and Some Applications of it. *Proceedings of the ASP Symposium on Close Range Photogrammetric Systems,* Champaign, Illinois, USA, 28th July - 1st August 1975, 420-476.

[22] **Tsai, R.Y.** An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition,* Miami Beach, FL, USA, 22-26 June 1986, 364-374.

[23] **Sid-Ahmed, M.A., Boraie, M.** and **Loy, W.** Photogrammetric Aerotriangulation Using Matrix CCD Cameras for Close-Range Position Sensing. *Computers in Industry*, 1989, **12**(4), 307-311.

[24] **Haigron, P., Lemehaute, A.** and **Collorec, R.** Multi-Plane and Thin-Plate Model for Camera Calibration. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* Paris, France, 29 Oct.-1 Nov. 1992, **14**, 2041-2042.

[25] **Champleboux, G., Lavallée, S., Sautot, P.** and **Cinquin, P.** Accurate Calibration of Cameras and Range Imaging Sensors: The NPBS Method. *Proceedings of the IEEE International Conference on Robotics & Automation,* Nice, France, 12-14 May 1992, **2**, 1552-1557.

[26] **Gremban, K.D., Thorpe, C.E.** and **Kanade, T.** Geometric Camera Calibration Using Systems of Linear Equations. *Proceedings of the IEEE International Conference on Robotics and Automation,* Philadelphia, PA, USA, 24-29 April 1988, 562-567.

[27] **Wei, G-Q.** and **Ma, S.D.** A Complete Two-Plane Camera Calibration Method and Experimental Comparisons. *Proceedings of the IEEE 4th International Conference on Computer Vision,* Berlin, Germany, 11-14th May 1993, 439-446.

[28] **Wei, G-Q.** and **Ma, S.D.** Implicit and Explicit Camera Calibration: Theory and Experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* May 1994, **16**(5), 469-480.

[29] **Fioretti, S., Germani, A.** and **Leo, T.** Stereometry in Very Close-Range Stereophotogrammetry with Non-Metric Cameras for Human Movement Analysis. *Journal of Biomechanics*, 1985, **18**(11), 831-842.

[30] **Wood, G.A.** and **Marshall, R.N.** The Accuracy of DLT Extrapolation in Three-Dimensional Film Analysis. *Journal of Biomechanics,* 1986, **19**(9), 781-785.

[31] **Challis, J.H.** and **Kerwin, D.G.** Accuracy Assessment and Control Point Configuration When Using the DLT for Photogrammetry. *Journal of Biomechanics,* 1992, **25**(9), 1053-1058.

[32] **Yu, B., Koh, T.J.** and **Hay, J.G.** A Panning DLT Procedure for Three-Dimensional Videography. *Journal of Biomechanics*, June 1993, **26**(6), 741-751.

[33] **Horn, B.K.P.** *Robot Vision,* MIT Press, Cambridge, MA, 1986. (ISBN: 0-262-08159-8)

[34] **Bose, C.B.** and **Amir, I.** Design of Fiducials for Accurate Registration Using Machine Vision. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1990, **12**(12), 1196-1200.

[35] **O'Gorman, L.,** *et al.* Subpixel Registration Using a Concentric Ring Fiducial. *Proceedings of the 10th IEEE International Conference on Pattern Recognition,* Atlantic City, NJ, USA, 16-21 June 1990, **2**, 249-253.

[36] **West, G.A.W.** and **Clark, T.A.** Survey and Examination of Subpixel Measurement Techniques. *Proceedings of SPIE - The International Society for Optical Engineering*, 1990, **1395**(1), 456-463.

[37] **Tian, Q.** and **Huhns, M.N.** Algorithms for Subpixel Registration. *Computer Vision, Graphics, and Image Processing*, 1986, **35**(2), 220-233.

[38] **Selvik, G.** Roentgen Stereophotogrammetry: A Method for the Study of the Kinematics of the Skeletal System. *Acta Orthopaedica Scandinavica, Supplement,* 1989, **60**(232), 1-51.

[39] **Stokes, I.A.F.** X-ray Photogrammetry. In *Three-Dimensional Analysis of Human Movement,* P. Allard, I.A.F. Stokes and J.-P. Blanchi (eds). Human Kinetics, Champaign, IL, 1995. (ISBN 0-87322-623-2).

[40] **Hay, G.A.** X-Ray Imaging. *Journal of Physics, Part E: Scientific Instrumentation,* 1978, **11**, 377-386.

[41] **Garrett, J.A.** and **Smithson, P.H.** Conventional X-Ray Imaging. *IEE Proceedings, Part A: Physical Science, Measurement and Instrumentation, Management and Education,* February 1987, **134**(2), 107-114.

[42] **Baumrind, S.** A System For Craniofacial Mapping Through The Integration of Data From Stereo X-Ray Films and Stereo Photographs. *ASP Symposium on Close Range Photogrammetric Systems.* Champaign, Illinois. 28th July 28 - 1st Aug. 1975, 142-166.

[43] **Turner-Smith, A.R., White, S.P.** and **Bulstrode, C.** X-Ray Photogrammetry of Artificial Hip Joints. *Proceedings of SPIE - The International Society for Optical Engineering,* 1990, **1395**(1), 587-594.

[44] **Lippert, F.G.,** *et al.* A Comparison of Convergent and Bi-Plane X-Ray Photogrammetry Systems used to Detect Total Joint Loosening. *Journal of Biomechanics*, 1982, **15**(9), 677-682.

[45] **Veress, S.A., Lippert, F.G.** and **Takamoto, T.** An Analytical Approach to X-Ray Photogrammetry. *Photogrammetric Engineering & Remote Sensing,* 1977, **43**(12), 1503-1510.

[46] **Selvik, G.** Roentgen Stereophotogrammetric Analysis. *Acta Radiologica,* 1990, **31**(2), 113-126.

[47] **Kärrholm, J.** Roentgen Stereophotogrammetry. Review of Orthopaedic Applications. *Acta Orthopaedica Scandinavica,* 1989, **60**(4), 491-503.

[48] **Ryd, L.** Roentgen Stereophotogrammetric Analysis of Prosthetic Fixation in the Hip and Knee Joint. *Clinical Orthopaedics and Related Research,* 1992, **276**, 56-65.

[49] **Fraser, C.S.** and **Abdullah, Q.A.** A Simplified Mathematical Model for Applications of Analytical X-Ray Photogrammetry in Orthopaedics. *Photogrammetric Engineering & Remote Sensing,* May 1981, **47**(5), 641-646.

[50] **Lippert, F.G., Veress, S.A., Tokamoto, T.** and **Spolek, G.A.** Experimental Studies on Patellar Motion Using X-Ray Photogrammetry. *ASP Symposium on Close Range Photogrammetric Systems.* Champaign, Illinois, 28th July -1st Aug. 1975, 186-208.

[51] **Lundberg, A., Svensson, O.K., Németh, G.** and **Selvik, G.** The Axis of Rotation of the Ankle Joint. *The Journal of Bone & Joint Surgery [Br],* 1989, **71**, 94-99.

[52] **De Lange, A., Kauer, J.M.G** and **Huiskes, R.** Kinematic Behaviour of the Human Wrist Joint: A Roentgen-Stereophotogrammetric Analysis. *Journal of Orthopaedic Research.* 1985, **3**(1), 56-64.

[53] **Jónsson, E.,** *et al.* Indications for Hemiarthroplasty of the Rheumatoid Shoulder. *Acta Orthopaedica Scandinavica,* 1988, Suppl. No. **227**, 77

[54] **Rune, B., Sarnäs, K.V., Selvik, G.** and **Jocobsson, S.** Roentgen Stereometry in the Study of Craniofacial Anomalies: The State of the Art in Sweden. *British Journal of Orthodontics,* 1986, **13**(3), 151-157.

[55] **Ragnarsson, J.I.** and **Kärrholm, J.** Stability of Femoral Neck Fracture: Roentgen Stereophotogrammetry of 29 Hook-Pinned Fractures. *Acta Orthopaedica Scandinavica,* 1991, **62**(3), 201-207.

[56] **Baldursson, H., Hansson, L.I., Olsson, T.H.** and **Selvik, G.** Migration of the Acetabular Socket After Total Hip Replacement Determined by Roentgen Stereophotogrammetry. *Acta Orthopaedica Scandinavica.* 1980, **51**(3), 535-540.

[57] **Mjöberg, B.** Fixation and Loosening of Hip Prosthesis: A review. *Acta Orthopaedica Scandinavica,* 1991, **62**(5), 500-508.

[58] **Hunter, J.C., Baumrind, S., Genant, H.K., Murray, W.R.** & **Ross S.E.** The Detection of Loosening in Total Hip Arthroplasty: Description of a Stereophotogrammetric Computer Assisted Method. *Investigative Radiology* 1979, **14**(4), 323-329.

[59] **Kiss, J., Murray, D.W., Turner-Smith, A.R.** and **Bulstrode, C.J.** Roentgen Stereophotogrammetric Analysis for Assessing Migration of Total Hip Replacement Femoral Components. *Proceedings of the Institution of Mechanical Engineers - Part H: Journal of Engineering in Medicine,* 1995, **209**, 169-175.

[60] **Ryd, L., Carlsson, L.** and **Herberts, P.** Micromotion of a Noncemented Tibial Component with Screw Fixation: An In-Vivo Roentgen Stereophotogrammetric Study of the Miller-Galante Prosthesis. *Clinical Orthopaedics and Related Research,* 1993, **295**, 218-225.

[61] **Jonason, C.O.** and **Hindmarsh, J.** Stereo X-Ray Photogrammetry as a Tool in Studying Scoliosis. *ASP Symposium on Close Range Photogrammetric Systems.* Champaign, Illinois, 28th July - 1st Aug. 1975, 245-259.

[62] **Hindmarsh, J., Larsson, J.** and **Mattsson, O.** Analysis of Changes in the Scoliotic Spine using a Three-Dimensional Radiographic Technique. *Journal of Biomechanics,* 1980, **13**, 279-290.

[63] **Kratky, V.** Analytical X-Ray Photogrammetry in Scoliosis. *ASP Symposium on Close Range Photogrammetric Systems.* Champaign, Illinois, 28th July - 1st Aug. 1975, 167-185.

[64] **Brown, R.H., Burstein, A.H., Nash, C.L.** and **Schock, C.C.** Spinal Analysis Using a 3-D Radiographic Technique. *Journal of Biomechanics*, 1976, **9**, 355-365.

[65] **Suh, C.H.** The Fundamentals of Computer Aided X-Ray Analysis of the Spine. *Journal of Biomechanics.* 1974, **7**, 161-9.

[66] **Lee, S., Harris, K.G., Nassif, J., Goel, V.K.** and **Clark, C.R.** In Vivo Kinematics of the Cervical Spine. Part I: Development of a Roentgen Stereophotogrammetric Technique Using Metallic Markers and Assessment of its Accuracy. *Journal of Spinal Disorders*, 1993, **6**(6), 522-534.

[67] **Olsson, T.H., Selvik, G.** and **Willner, S.** Postoperative Kinematics in Structural Scoliosis. *Acta Radiologica (Diagnosis).* 1977, **18**, 75-86.

[68] **Olsson, T.H., Selvik, G.** and **Willner, S.** Mobility in the Lumbrosacral Spine After Fusion Studied with the Aid of Roentgen Stereophotogrammetry. *Clinical Orthopaedics and Related Research*, 1977, **129**, 181-190.

[69] **Moores, B.M.** Digital X-Ray Imaging. *IEE Proceedings, Part A: Physical Science, Measurement & Instrumentation, Management & Education,* 1987, **134**(2), 115-125.

[70] **Cowen, A.R.** Digital X-Ray Imaging. *Measurement Science & Technology*, 1991, **2**(8), 691-707.

[71] **Cowen, A.R., Workman, A.** and **Price, J.S.** Physical Aspects of Photostimulable Phosphor Computed Radiography. *The British Journal of Radiology*, 1993, **66**(784), 332-345.

[72] **Murphey, M.D.,** *et al.* Computed Radiography in Musculoskeletal Imaging: State of the Art. *American Journal of Roentgenology*, 1993, **158**(1), 19-27.

[73] *The Physics of Medical Imaging.* Webb, S. (editor). Adam Hilger, Bristol, 1988 (ISBN 0-85274-349-1)

[74] **Vosburgh, K.G., Swank, R.K.** and **Houston, J.M.** X-Ray Image Intensifiers. *Advances in Electronics & Electron Physics,* 1977, **43**, 205 -244.

[75] **Lewall, D.B., Riley, P., Al Hassoon, A.** and **McParland, B.J.** A Fluoroscopy Credentialling Programme for Orthopaedic Surgeons. *The Journal of Bone and Joint Surgery*, 1995, 77(3), 442-444.

[76] **Van Der Eijk, B.** and **Kühl. W.** An X-Ray Image Intensifier with Large Input Formats. *Philips Technical Review*, 1984, **41**(5), 137-148.

[77] **Rudin, S., Bednarek, D.R.** and **Wong, R.** Accurate Characterisation of Image-Intensifier Distortion. *Medical Physics,* 1991, **18**(6), 1145-1151.

[78] **Fenner, E.,** *et al.* X-Ray Image Intensifiers: Image Quality and Possibilities for Enhancement. *Advances in Electronics & Electron Physics*, 1972, **33**, 1049-1059.

[79] **Casperson, L.W., Spiegler, P.** and **Grollman, J.H.** Characterisation of Aberrations in Image-Intensified Fluoroscopy. *Medical Physics,* 1976, 3(2), 103-106.

[80] **Fujita, H.,** *et al.* Basic Imaging Properties of a Large Image Intensifier-TV Digital Radiographic System. *Investigative Radiology,* 1987, **22**(4), 328-335.

[81] **Guggenheim, N.,** *et al.* Spatial Reconstruction of Coronary-Arteries from Angiographic Images. *Physics in Medicine and Biology,* 1991, **36**(1), 99-110.

[82] **Haaker, P., Klotz, E., Koppe, R.** and **Linde, R.** Real-Time Distortion Correction of Digital X-Ray II/TV Systems: An Application Example. *Proceedings of SPIE - The International Society for Optical Engineering,* 1989, **1027**, 261-266.

[83] **Baltzopoulos, V.** Videofluoroscopy Method for Optical Distortion Correction and Measurement of Knee-Joint Kinematics. *Clinical Biomechanics,* March 1995, **10**(2), 85-92.

[84] **Butler, D.A.** and **Pierson, P.K.** A Distortion-Correction Scheme For Industrial Machine-Vision Applications. *IEEE Transactions on Robotics and Automation,* August 1991, 7(4), 546 -551.

[85] **Boone, J.M., Seibert, J.A., Barrett, W.A.** and **Blood, E.A.** Analysis and Correction of Imperfections in the Image Intensifier-TV-Digitizer Imaging Chain. *Medical Physics,* 1991, **18**(2), 236-242.

[86] **Reimann, D.A.** and **Flynn, M.J.** Automatic Distortion Correction of X-Ray Image Intensifier Images. *Conference Record of the IEEE Nuclear Science Symposium & Medical Imaging Conference,* Orlando, FL, USA, 25-31 October 1992, 1339-1341.

[87] **Pietka, E.** and **Huang, H.K.** Correction of Aberration in Image-Intensifier Systems. *Computerised Medical Imaging & Graphics,* 1992, **16**(4), 253-258.

[88] **Ning, R., Barsotti, J.B.** and **Kido, D.K.** Clinical Image Intensifier-Based Volume CT Imager for Angiography. *Proceedings of SPIE - International Society for Optical Engineering,* 1991, **1443**, 236-249

[89] **Chakraborty, D.P.** Image Intensifier Distortion Correction. *Medical Physics,* 1987, 14(2), 249-52.

[90] **Duffin, W.J.** *Electricity and Magnetism.* McGraw-Hill Book Company (UK) Limited, Maidenhead, 1980. (ISBN 0-07-084111-X)

[91] **Solzbach, U., Wollschläger, H., Zeiher, A.** and **Just, H.** Optical Distortion Due to Geomagnetism in Quantitative Angiography. *Computers in Cardiology,* 1988, 355-357.

[92] **Gronenschild, E.** Effect of Spatial Orientation on the Image-Intensifier Distortion in Quantitative Coronary Angiography. *Computers in Cardiology,* 1993, 595-598.

[93] **Koppe, R., Klotz, E., Op de Beek, J.** and **Aerts, H.** 3D Vessel Reconstruction Based on Rotational Angiography. *Computer Assisted Radiology - Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy,* Berlin, Germany, 21-24 June 1995, 101-107.

[94] **Beier, J., Oswald, H.** and **Fleck, E.** Edge Detection For Coronary Angiograms : Error Correction and Impact of Derivatives. *Computers in Cardiology,* 1992, 513-516.

[95] **Van Der Zwet, P.M.J., Meyer, D.J.H.** and **Reiber, J.H.C.** Automated and Accurate Assessment of the Distortion, Magnitude, and Direction of Pincushion Distortion in Angiographic Images. *Investigative Radiology,* 1995, **30**(4), 204-213.

[96] **Ning, R., Riek, J.K.** and **Conover, D.L.** Image-Intensifier-Based Volume Tomographic Angiography Imaging System: Geometric Distortion Correction. *Proceedings of SPIE - The International Society for Optical Engineering,* 1994, **2163**, 199-210.

[97] **Onnasch, D.G.W.** and **Prause, G.P.M.** Geometric Image Correction and Iso-Center Calibration at Oblique Biplane Angiographic Views. *Proceedings of Computers in Cardiology,* Durham, NC, USA, 11-14 October 1992, 647-650.

[98] **Guggenheim, N.,** *et al.* Determination of Intra-Arterial Volumes from Biplane Coronary Angiography. *Computers in Cardiology,* 1989, 71-74.

[99] **Garreau, M., Coatrieux, J.L., Collorec, R.** and **Chardenon, C.** A Knowledge-Based Approach for 3-D Reconstruction and Labelling of Vascular Networks from Biplane Angiographic Projections. *IEEE Transactions on Medical Imaging,* June 1991,. **10**(2), 122-131.

[100] **Wahle, A., Oswald, H.** and **Fleck, E.** 3D Heart-Vessel Reconstruction from Biplane Angiograms. *IEEE Computer Graphics and Applications,* January 1996, **16**(1), 65-73

[101] **Guggenheim, N.,** *et al.* 3D Determination of the Intravascular Volume and Flow of Coronary Arteries. *International Journal of Biomedical Computing,* 1994, **35**(1), 13-22.

[102] **Metz, C.E.** and **Fencil, L.E.** Determination of Three-Dimensional Structure in Biplane Radiography Without Prior Knowledge of the Relationship Between the Two Views: Theory. *Medical Physics,* 1989, **16(1)**, 45-51.

[103] **Wahle, A.,** *et al.* Assessment of Diffuse Coronary Artery disease by Quantitative Analysis of Coronary Morphology Based upon 3-D Reconstruction from Biplane Angiograms. *IEEE Transactions on Medical Imaging,* June 1995, **14**(2), 230-241.

[104] **Prause, G.P.M.** and **Onnasch, D.G.W.** Steps of Image Correction and Processing for the Binary Reconstruction of the Ventricles from Biplane Angiocardiograms. *Proceedings of Computers in Cardiology,* Durham, NC, USA, 11-14 October 1992, 27-30.

[105] **Rougée, A.** *et al.* Three-Dimensional Coronary Arteriography. *International Journal of Cardiac Imaging,* 1994, **10(1)**, 67-70.

[106] **Rougée, A., Picard, C.L., Trousset, Y.L.** and **Ponchut, C.** Geometrical Calibration for 3D X-ray Imaging. *Proceedings of SPIE - The International Society for Optical Engineering,* 1993, **1897**, 161-169.

[107] **Drake, J.M.,** *et al.* Frameless Stereotaxy in Children. *Pediatric Neurosurgery,* 1994, **20**(2), 152-159.

[108] **Hassfeld, S.,** *et al.* Computer-Aided Skull Base Surgery. *Computer Assisted Radiology: Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR'95).* Berlin, Germany, 21-24 June 1995, 1168-1173.

[109] **Olivier, A., Germano, I.M., Cukiert, A.** and **Peters, T.** Frameless Stereotaxy for Surgery of the Epilepsies: Preliminary Experience. *Journal of Neurosurgery,* 1994, **81**(4), 629-633.

[110] **Glossop, N.D., Hu, R.W.** and **Randle, J.A.** Computer-Aided Pedicle Screw Placement Using Frameless Stereotaxis. *Spine*, 1996, **21**(17), 2026-2034.

[111] **Watanabe, E.** The Neuronavigator: A Computer-Controlled Navigation System in Neurosurgery. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds). MIT Press, Cambridge, MA, 1996, 319-327.

[112] **Reinhardt, H.F.** Neuronavigation: A Ten-Year Review. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds). MIT Press, Cambridge, MA, 1996, 329-341.

[113] **Adams, L.**, *et al.* Computer Assisted Surgery. *IEEE Computer Graphics and Applications*, May 1990, **10**(3), 43-51.

[114] **Galloway, R.L., Maciunas, R.J.** and **Edwards, C.A.** Interactive Image-Guided Neurosurgery. *IEEE Transactions on Biomedical Engineering*, December 1992, **39**(12), 1226-1231.

[115] **Giorgi, C.** Intraoperative Fusion of Field Images With CT MRI Data by Means of a Stereotaxic Mechanical Arm. *Minimally Invasive Neurosurgery*, 1994, **37**(2), 53-55.

[116] **Louhisalmi, Y.**, *et al.* Development of a Localisation Arm for Neurosurgery. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* Paris, France, 29th Oct.-1st Nov. 1992, **14**, 1085-1086.

[117] **Mazier, B., Lavallée, S.** and **Cinquin, P.** Computer Assisted Interventionist Imaging: Application to the Vertebral Column Surgery. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine & Biology Society*, Philadelphia, PA, USA, 1990, No.1, 430-431.

[118] **Simon, D.A., Herbert, M.** and **Kanade, T.** Techniques for Fast and Accurate Intra-Surgical Registration. *Proceedings of the First International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, Pittsburgh, PA, USA, 22-24 September 1994, 90-97.

[119] **Evans, R.** VISLAN Computer-Aided Surgery. *IEE Review*, March 1995, **41**(2), 51-54.

[120] **Grimson, W.E.L.**, *et al.* An Automatic Registration Method For Frameless Stereotaxy, Image-Guided Surgery, and Enhanced Reality Visualisation. *IEEE Transactions on Medical Imaging*, April 1996, **15**(2), 129-140.

[121] **Martelli, S.**, *et al.* A Laser Scanner System for Investigating Non-Invasive Matching Strategies in Computer-Assisted Orthopaedic Surgery. *Proceedings of the 13th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Orlando, FL, USA, 31 Oct.-3 Nov. 1991, **13**(4), 1757-1758.

[122] **Vassal, P.** *et al.* Image-Guided Accurate Set-Up of Patients in Radiotherapy. *Computer Assisted Radiology - Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR'95)*, Berlin, Germany, 21-24 June 1995, 1077-1082.

[123] **Galloway, R.L., Maciunas, R.J., Bass, W.A.** and **Crouch, D.** An Optical Device for Interactive, Image-Guided Neurosurgery. *Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, CA, USA, 28-31 October 1993, **15**, 954-955.

[124] **Nolte, L-P.,** *et al.* Clinical Evaluation of a System for Precision Enhancement in Spine Surgery. *Clinical Biomechanics*, 1995, 10(6), 293-303.

[125] **Lavallée, S.,** *et al.* Computer Assisted Spine Surgery: A Technique for Accurate Transpedicular Screw Fixation Using CT Data and a 3-D Optical Localiser. *Proceedings of the First International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, Pittsburgh, PA, USA, 22-24 September 1994, 315-322.

[126] **Glossop, N.** and **Hu, R.** Assessment of Vertebral Body Motion During Spine Surgery. *Spine*, 1997, **22**(8), pp.903-909.

[127] **Simon, D.A.,** *et al.* Validating Surface-Based Registration for Orthopaedic Surgery. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, MD, USA, 4-7 November 1995, 185-192.

[128] **Dessenne, V.,** *et al.* Computer Assisted Knee Cruciate Ligament Reconstructions: First Clinical Tests. *Proceedings of the First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed'95)*, Nice, France, 3-6 April 1995, 481-486.

[129] **Maurer, C.R.,** *et al.* The Accuracy of Image-Guided Neurosurgery Using Implantable Fiducial Markers. *Computer Assisted Radiology - Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR'95)*, Berlin, Germany, 21-24 June 1995, 1197-1202.

[130] **Edwards, P.J.,** *et al.* Augmented Reality in the Stereo Operating Microscope for Otolaryngology and Neurosurgical Guidance. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, MD, USA, 4-7 November 1995, 8-15.

[131] **Cutting, C.B., Bookstein, F.L.** and **Taylor, R.H.** Applications of Simulation, Morphometrics and Robotics in Craniofacial Surgery. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds). MIT Press, Cambridge, MA, 1995, 641-662.

[132] **Troccaz, J.,** *et al.* Patient Set-Up Optimisation for External Conformal Radiotherapy. *Proceedings of the First International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, Pittsburgh, PA, USA, 22-24 September 1994, 306-312.

[133] **Williams, R.** and **Leonard, J.** Surgery Benefits from Defence Technology. *IEEE Aerospace & Electronic Systems Magazine*, Oct. 1994, 9(10), 3-6.

[134] **Friets, E.M., Strohbehn, J.W., Hatch, J.F.** and **Roberts, D.W.** A Frameless Stereotaxic Operating Microscope for Neurosurgery. *IEEE Transactions on Biomedical Engineering*, 1989, **36**(6), 608-617.

[135] **Barnett, G.H.,** *et al.* Intraoperative Localisation Using an Armless, Frameless Stereotactic Wand. *Journal of Neurosurgery*, 1993, 78(3), 510-514.

[136] **Kato, A.,** *et al.* A Frameless, Armless Navigational System for Computer-Assisted Neurosurgery. *Journal of Neurosurgery*, 1991, **74**(5), 845-849.

[137] **Goerss, S.J., Kelly, P.J., Kall, B.** and **Stiving, S.** A Stereotactic Magnetic Field Digitiser. *Stereotactic and Functional Neurosurgery*, 1994, **63**(1-4), 89-92.

[138] **Gunkel, A.R., Freysinger, W., Thumfart, W.F.** and **Truppe, M.J.** Application of the ARTMA Image-Guided Navigation System to Endonasal Sinus Surgery. *Computer Assisted Radiology - Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR'95)*, Berlin, Germany, 21-24 June 1995, 1146-1151.

[139] **Amiot, L.P.,** *et al.* Computer-Assisted Pedicle Screw Fixation: A Feasibility Study. *Spine*, 1995, **20**(10), 1208-1212.

[140] **Dunn, S.M., Burdea, G.C.** and **Goratowski, R.** Robotic Control of Intraoral Radiography. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds). MIT Press, Cambridge, MA, 1995, 519-528.

[141] **Finlay, P.A.** and **Ornstein, M.H.** Controlling the Movement of a Surgical Laparoscope. *IEEE Engineering in Medicine and Biology Magazine*, 1995, **14**(3), 289-291.

[142] **Kazanzides, P.,** *et al.* An Integrated System for Cementless Hip Replacement. *IEEE Engineering in Medicine and Biology Magazine*, 1995, **14**(3), 307-313.

[143] **Bargar, W.L.,** *et al.* Clinical Abstract - Robodoc Multi-Center Trial: An Interim Report. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, MD, USA, 4-7 November 1995, 208-214.

[144] **Gottschalk, M.A.** Part Surgeon, Part Robot - It's Robodoc. *Design News*, June 1993, **49**(11), 68-75.

[145] **Lea, J.T., Santos-Munné, J.J.** and **Peshkin, M.A.** Diagramming Registration Connectivity and Structure. *IEEE Engineering in Medicine and Biology Magazine*, 1995, **14**(3), 271-278.

[146] **O'Toole, R.V.,** *et al.* Towards More Capable and Less Invasive Robotic Surgery in Orthopaedics. *Proceedings of the First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed'95)*, Nice, France, 3-6 April 1995, 123-130.

[147] **Simon, D.A.,** *et al.* Development and Validation of a Navigational Guidance System for Acetabular Implant Placement. *Lecture Notes in Computer Science*, 1997, **1205**, 583-592.

[148] **Matsen, F.A.,** *et al.* Robotic Assistance in Orthopaedics: A Proof of Principle Using Distal Femoral Arthroplasty. *Clinical Orthopaedics & Related Research*, 1993, **296**, 178-186.

[149] **Kienzle, T.C.,** *et al.* Total Knee Replacement. *IEEE Engineering in Medicine and Biology Magazine*, 1995, **14**(3), 301-306.

[150] **Fadda, M.,** *et al.* Computer-Assisted Knee Arthroplasty at Rizzoli Institutes. *Proceedings of the First International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, Pittsburgh, PA, USA, 22-24 September 1994, 26-30.

[151] **Martelli, S.,** *et al.* A Laser Scanner System for Investigating Non-Invasive Matching Strategies in Computer-Assisted Orthopaedic Surgery. *Proceedings of the 13th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Orlando, FL, USA, 31 Oct. - 3 Nov. 1991, **13**(4), 1757-1758.

[152] **Ho, S.C., Hibberd, R.D.** and **Davies, B.L.** Robot Assisted Knee Surgery. *IEEE Engineering in Medicine and Biology Magazine*, 1995, **14**(3), 292-300.

[153] **Aebi, M.** and **Webb, J.K.** The Spine. In *Manual of Internal Fixation: Techniques Recommended by the AO-ASIF Group*, 3rd ed., M.E. Muller, M. Allgower, R. Schneider & H. Willenegger (eds.), Springer-Verlag, Berlin, 1991, pp 627-682. (ISBN 3-540-52523-8)

[154] **Esses S.I., Sachs, B.L.** and **Dreyzin, V.** Complications Associated with the Technique of Pedicle Screw Fixation: A Selected Survey if ABS Members. *Spine*, 1993, **18**, 2231-2239.

[155] **Gertzbein, S.D.** and **Robbins, S.** Accuracy of Placement of Pedicular Screws In Vivo. *Spine*, 1990, **15**, 11-14.

[156] **Schwarzenbach, O.,** *et al.* Accuracy of Computer-Assisted Pedicle Screw Placement: An In Vivo Computed Tomography Analysis. *Spine*, 1997, **22**(4), 452-458.

[157] **Laine, T.,** *et al.* Improved Accuracy of Pedicle Screw Insertion With Computer-Assisted Surgery: A Prospective Trial of 30 Patients. *Spine*, 1997, **22**(11), 1254-1258.

[158] **Lavallée, S.,** *et al.* Computer-Assisted Spinal Surgery Using Anatomy-Based Registration. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds). MIT Press, Cambridge, MA, 1996, 425-449.

[159] **Nolte, L.P.,** *et al.* Image-Guided Insertion of Transpedicular Screws - A Laboratory Set-Up. *Spine*, 1995, **20**(4), 497-500.

[160] **Brodwater, B.K.,** *et al.* Extracranial Application of the Frameless Stereotactic Operating Microscope: Experience with Lumbar Spine. *Neurosurgery*, 1993, **32**(2), 209-213.

[161] Imperial College of Science, Technology & Medicine. Department of Mechanical Engineering. Review 1992-1994.

[162] **Lord, P.J.** and **Mann, R.W.** Computer Aided Intertrochanteric Osteotomy Planning and Surgery Simulation. *Proceedings of the American Society of Mechanical Engineers: Advances in Bioengineering*, 1993, **26**, 531-534.

[163] **Moctezuma, J.L., Gossé, F.** and **Schultz, H-J.** A Computer and Robotic Aided Surgery System for Accomplishing Osteotomies. *Proceedings of the First International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, Pittsburgh, PA, USA, 22-24 September 1994, 31-35.

[164] **Moctezuma, J.L., Schuster, D., Gossé, F.** and **Schulz, H.-J.** A New Surgery Tool for Robotic Aided Treatment Delivery. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, Maryland USA., 4-7 Nov. 1995, 304-311.

[165] **Radermacher, K., Staudte, H-W.** and **Rau, G.** Computer Assisted Orthopaedic Surgery by Means of Individual Templates - Aspects and Analysis of Potential Applications. *Proceedings of the First International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, Pittsburgh, PA, USA, 22-24 September 1994, 42-48.

[166] **Radermacher, K., Staudte, H.-W.** and **Rau, G.** Technique for Better Execution of CT Scan Planned Orthopaedic Surgery on Bone Structures. *Computer Assisted Radiology - Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR'95)*. Berlin, Germany, 21-24 June 1995, 933-938.

[167] **Potamianos, P., Davies, B.L.** and **Hibberd, R.D.** Intra-Operative Registration for Percutaneous Surgery. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, MD, USA, 4-7 November 1995, 156-164.

[168] **Santos-Munné, J.J.**, *et al.* A Stereotactic and Robotic System for Pedicle Screw Placement. (Presented at the Third Medicine Meets Virtual Reality Conference (MMVRIII), San Diego, CA, 19-22 January 1995). In: *Studies in Health Technology and Informatics*, Satava, R.M. *et al.*, (eds.), IOS Press, 1995, **18**, Ch. 66, 326-333.

[169] **Potamianos, P., Davies, B.L.** and **Hibberd, R.D.** Manipulator Assisted Renal Treatment. *Proceedings of the Annual International Conference of the Australian Robot Association & The International Federation of Robotics*, Brisbane, Australia, 14-16 July 1993, 214-227.

[170] **Potamianos, P., Davies, B.L.** and **Hibberd, R.D.** Intra-Operative Imaging Guidance For Keyhole Surgery Methodology and Calibration. *Proceedings of the First International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, Pittsburgh, PA, USA, 22-24 September 1994, 98-104.

[171] **Cinquin, P.**, *et al.* Computer Assisted Medical Interventions. *IEEE Engineering in Medicine and Biology Magazine*, 1995, **14**(3), 254-263.

[172] **Lavallée, S., Szeliski, R.** and **Brunie, L.** Anatomy-Based Registration of Three-Dimensional Medical Images, Range Images, X-Ray Projections, and Three-Dimensional Models Using Octree-Splines. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds), MIT Press, Cambridge, MA, 1996, 115-143.

[173] **Sautot, P., Cinquin, P., Lavallée, S.** and **Troccaz, J.** Computer Assisted Spine Surgery: A First Step Toward Clinical Applications in Orthopaedics. *Proceedings of the 14th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Paris, France, Oct. 29th - Nov. 1st. 1992, **14**, 1071-1072.

[174] **Hamadeh, A., Sautot, P., Lavallée, S.** and **Cinquin, P.** Towards Automatic Registration Between CT and X-ray Images: Co-operation Between 3D/2D Registration and 2D Edge Detection. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, MD, USA, 4-7 Nov. 1995, 39-46.

[175] **Lavallée, S.** and **Szeliski, R.** Recovering the Position and Orientation of Free-Form Objects from Image Contours Using 3D Distance Maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, **17**(4) , 378-390.

[176] **Lavallée, S.**, *et al.* Image-Guided Operating Robot: A Clinical Application in Stereotactic Neurosurgery. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds). MIT Press, Cambridge, MA, 1996, 343-351.

[177] **Hofstetter, R.** Fluoroscopy Based Surgical Navigation for Femoral Fracture Fixation. *Presented at the Second Computer-Assisted Orthopaedic Surgery Symposium (CAOS)*, Bern, Switzerland, 7-9 November 1996.

[178] **Finlay, P.A.** ORTHOSISTA: An Active Surgical Localiser for Assisting Orthopaedic Fracture Fixation. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, MD, USA, 4-7 November 1995, 203-207.

[179] **Phillips, R.**, *et al.* Image-Guided Orthopaedic-Surgery Design and Analysis. *Transactions of the Institute of Measurement and Control*, 1995, 17(5), 251-264.

[180] **Paul, H.A.,** *et al.* Development of a Surgical Robot for Cementless Total Hip Arthroplasty. *Clinical Orthopaedics and Related Research*, 1992, 285, 57-66

[181] **Shelley, T.** Robots Help Surgeons Extend the Possible. *Eureka on Campus*, Spring 1996, 36-37 + 42.

[182] **Holt, R.J.** and **Netravali, A.N.** Camera Calibration Problem: Some New Results. *CVGIP: Image Understanding*, 1991, **54**(3), 368-383.

[183] **Mohsen, A.M.M.A.,** *et al.* The Basic Orthopaedic Principle and the Non-Invasive Intelligent Orthopaedic Guide Concept. *Proceedings of the Second International Workshop: Mechatronics in Medicine & Surgery. (MediMec'95)*, Bristol, UK, 6-9 September 1995, 47-57.

[184] **Viant, W.J.,** *et al.* A Computer Assisted Orthopaedic Surgical System for Distal Locking of Intramedullary Nails. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 1997, **211**(4), pp 293-300.

[185] **Viant, W.J.,** *et al.* A Computer Assisted Orthopaedic System for Distal Locking of Intramedullary Nails. *Proceedings of the Second International Workshop: Mechatronics in Medicine & Surgery (MediMec'95)*, Bristol, UK, 6-9 September 1995, 86-91.

[186] **Mohsen, A.M.M.A.** *A Non-Intelligent Orthopaedic Guide: Concepts, Issues and Efficacy Assessment.* PhD Thesis, University of Hull, April 1995.

[187] **Phillips, R.,** *et al.* Steps Towards Computer Assisted Locking of Intramedullary Nails. *Proceedings of the 15th Annual International Conference of the IEEE Engineering in Biology & Medicine Society*, San Diego, CA, USA, 28-31 October 1993, 950-951

[188] **Benabid, A.L.,** *et al.* Potential Use of Robots in Endoscopic Neurosurgery. *Acta Neurochirurgica (Wien, Austria)*, 1992, Suppl. **54**, 93-97.

[189] **Craig, J.J.,** *Introduction to Robotics: Mechanics and Control.* Addison-Wesley Publishing Company, Reading, Massachusetts, 2nd Edition, 1989. (ISBN: 0-201-09528-9).

[190] **Nash, J.C.** *Compact Numerical Methods for Computers: Linear Algebra & Function Minimisation.* Adam Hilger Ltd, Bristol, 1979. (ISBN 0-85274-330-0)

[191] **Bell, R.J.T.** *An Elementary Treatise on Coordinate Geometry of Three Dimensions*, MacMillan & Co. Ltd., London, 3rd Edn., 1956.

[192] **Schnabler, J.** *Robotic Assistance in Orthopaedic Surgery: Repair of Proximal Femoral Fractures.* MSc Thesis, Lougborough University, 1997.

[193] **Blauvelt, C.T.** and **Nelson, F.R.T.** *A Manual of Orthopaedic Terminology*, 5th edition, Mosby Year Book, St. Louis, Missouri, 1994. (ISBN 0-8016-6943-X)

[194] *Gray's Anatomy.* Warwick, R. and Williams, P.L. editors. 35th edition, Longman Group Ltd., Edinburgh, 1973. (ISBN 0-443-01011-0).

[195] **Moore, K.L.** *Clinically Oriented Anatomy*, Williams & Wilkins, Baltimore, 1983. (ISBN 0-683-06146-1)

[196] **Bender, C.E.,** and **Campbell, D.C.** Femoral Shaft Fractures. In *Imaging of Orthopaedic Trauma and Surgery*, T.H. Berquist (editor), W.B. Saunders Co., 1986. (ISBN 0-7126-1102-8).

[197] **Dandy, D.J.** *Essential Orthopaedics & Trauma*, 2nd Edition, Churchill Livingstone, London, 1993. (ISBN 0-443-04722-7)

[198] **Bucholz, R.W.** and **Browner, B.** Fractures of the Shaft of the Femur. In *Surgery of the Musculoskeletal System,* Volume 3, Second Edition, C.M. Evarts (ed.), Churchill Livingstone, New York, 1990, pp 2687-2729. (ISBN 0-44-08516-1).

[199] **Apley, A.G.** and **Solomon, L.** *Concise System of Orthopaedics and Fractures,* Butterworth Heinemann, London., 1991. (ISBN 0-7506-0343-7).

[200] **Weller, S.** Internal Fixation of Fractures by Intramedullary Nailing: Introduction, Historical Review & Present Status. *Injury*, 1993, **24** Supplement 3, 1-6.

[201] **Blumberg, K.D.,** *et al.* A Comparison of the Brooker-Wills and Russell-Taylor Nails for Treatment of Patients Who Have Fractures of the Femoral Shaft. *The Journal of Bone and Joint Surgery [Am]*, 1990, **72**(7), 1019-1024.

[202] **Papagiannopoulos, G.** and **Clement, D.A.** Treatment of Fractures of the Distal Third of the Femur. *The Journal of Bone and Joint Surgery [Br]*, 1987, **69**(1), 67-70.

[203] **De La Caffinière, J-Y., Pelisse, F.** and **De La Caffinière, M.** Locked Intramedullary Flexible Osteosynthesis: A Mechanical and Clinical Study of a New Pin Fixation Device. *The Journal of Bone and Joint Surgery [Br].* 1994, **76**(5), 778-788.

[204] **Weller, S.** and **Hontsch, D.** Medullary Nailing of Femur and Tibia. In *Manual of Internal Fixation: Techniques Recommended by the AO-ASIF Group*, 3rd ed., M.E. Muller, M. Allgower, R. Schneider & H. Willenegger (eds.), Springer-Verlag, Berlin, 1991, pp 291-365. (ISBN 3-540-52523-8)

[205] **Browner, B.D.** Pitfalls, Errors, and Complications in the Use of Locking Kuntscher Nails. *Clinical Orthopaedics & Related Research*, 1986, **212**, 192-208.

[206] **Kempf, I., Grosse, A.** and **Beck, G.** Closed Locked Intramedullary Nailing - Its Application to Comminuted Fractures of the Femur. *The Journal of Bone and Joint Surgery [Am]*, 1985, **67**(5), 709-720.

[207] **Taylor, J.C.,** Fractures of Lower Extremity. In *Campbell's Operative Orthopaedics*, Volume 2, 8th Edition, Crenshaw, A.H. (ed.), Mosby Year Book, St Louis, 1992, 786-893. (ISBN 0-8016-1096-6).

[208] **Winquist, R.A., Hanson, S.T.** and **Clawson, D.K.** Closed Intramedullary Nailing of Femoral Fractures: A Report of Five Hundred and Twenty Cases. *The Journal of Bone and Joint Surgery [Am]*, 1984, **66**(4), 529-539.

[209] *Russell-Taylor Interlocking Nail System,* Richards Medical Company, Memphis, TN, 1987.

[210] **Sugarman, I.D., Adam, I.** and **Bunker, T.D.** Radiation Dosage During AO Locking Femoral Nailing. *Injury*, 1988, **19**(5), 336-338.

[211] **Hindley, C.J., Evans, R.A., Holt, E.M.** and **Metcalfe, J.W.** Locked Intramedullary Nailing for Recent Lower Limb Fractures. *Injury*, 1990, **21**(4), 239-244.

[212] **Goodall, J.D.** An Image Intensifier Laser Guidance System for the Distal Locking of an Intramedullary Nail. *Injury*, 1991, **22**(4), 339.

[213] **Goulet, J.A., Londy, F., Saltzman, C.L.** and **Matthews, L.S.** Interlocking Intramedullary Nails: An Improved Method of Screw Placement Combining Image Intensification and Laser Light. *Clinical Orthopaedics and Related Research*, 1992, **281**, 199-203.

[214] **Saw, Y.** Closed Intramedullary Distal Locking Made Easier. *Injury*, 1993, **24**(3), 214-215.

[215] **Owen, T.D.** and **Coorsh, J.** Insertion of the Distal Locking Screws in Femoral Nailing : A Simplified Technique. *Injury*, 1993, **24**(2), 101-103.

[216] **Knudsen, C.J.M., Grobler, G.P.** and **Close, R.E.W.** Inserting the Distal Screws in a Locked Femoral Nail. *The Journal of Bone & Joint Surgery [Br]*, 1991, **73**(4), 660-661.

[217] **Noordeen, H.H., Sala, M.J.** and **Belham, G.J.** Insertion of Distal Screws in Interlocking Intramedullary Nails. *Injury*, 1993, **24**(5), 357-358.

[218] **Hudson, I.** Locking Nailing: An Aid to Distal Targeting. *Injury*, 1989, **20**(3), 129-130.

[219] **Skjeldal, S.** and **Backe, S.** Interlocking Medullary Nails - Radiation Doses in Distal Targeting. *Archives of Orthopaedic and Traumatic Surgery*, 1987, **106**(3), 179-181.

[220] **Medoff, R.J.** Insertion of the Distal Screws in Interlocking Nail Fixation of Femoral Shaft Fractures. *The Journal of Bone and Joint Surgery [Am]*, 1986, **68**(8), 1275-1277.

[221] **MacMillan, M.** and **Gross, R.H.** A Simplified Technique of Distal Femoral Screw Insertion for the Grosse-Kempf Interlocking Nail. *Clinical Orthopaedics and Related Research*, 1988, **226**, 252-259.

[222] **Hashemi-Nejad, A., Garlick, N.** and **Goddard, N.J.** A Simple Jig to Ease the Insertion of Distal Screws in Intramedullary Locking Nails. *Injury*, 1994, **25**(6), 407-408.

[223] **Rao, J.P., Allegra, M.P., Benevenia, J.** and **Dauhajre, T.A.** Distal Screw Targeting of Interlocking Nails. *Clinical Orthopaedics and Related Research*, 1989, **238**, 245-248.

[224] **Hochrainer, M.J.** *Robotic Assistance in Orthopaedic Surgery: Repair of Femoral Shaft Fractures*. MSc Thesis, Lougborough University, 1997.

[225] **Klenerman, L.** The Young Patient with a Fractured Neck of Femur. *British Medical Journal*, 1985, **290**(6486), 1928.

[226] **Riggs, B.L.** and **Melton, L.J.** Involutional Osteoporosis. *The New England Journal of Medicine*, 1986, **314**(26), 1676-1686.

[227] **Cooper, C.** Bone Mass Throughout Life: Bone Growth and Involution. In *Osteoporosis: Pathogenesis and Management*, R.M. Francis (ed.), Kluwer Academic Publishers, Dordrecht, Netherlands, 1990, pp 1-26. (ISBN 0-7462-0086-2)

[228] **Cooper, C., Barker, D.J.P., Morris, J.** and **Briggs, R.S.J.** Osteoporosis, Falls, and Age in Fracture of the Proximal Femur. *British Medical Journal*, 1987, **295**(6589), 13-15.

[229] **Hoffenberg, R.** *et al.* Fractured Neck of Femur: Prevention and Management. Summary and Recommendations of a Report of the Royal College of Physicians. *Journal of the Royal College of Physicians of London*, 1989, **23**(1), 8-12.

[230] **Mainds, C.C.** and **Newman, R.J.** Implant Failures in Patients with Proximal Fractures of the Femur Treated with a Sliding Screw Device. *Injury*, 1989, **20**(2), 98-100.

[231] **Davis, T.R.C.,** *et al.* Intertrochanteric Femoral Fractures: Mechanical Failure After Internal Fixation. *Journal of Bone and Joint Surgery [Br]*, 1990, **72**(1), 26-31.

[232] **Thomas, A.P.** Dynamic Hip Screws That Fail. *Injury*, 1991, **22**(1), 45-46.

[233] **Anderson, G.H., Raymakers, R.** and **Gregg, P.J.** The Incidence of Proximal Femoral Fractures in an English County. *Journal of Bone and Joint Surgery [Br]*, 1993, **75**(3), 441-444.

[234] **Keene, G.S., Parker, M.J.** and **Pryor, G.A.** "Mortality and Morbidity After Hip Fractures. *British Medical Journal*, 1993, **307**(6914), 1248-1250.

[235] **Gundle, R., Gargan, M.F.** and **Simpson, A.H.R.W.** How to minimise failures of fixation of unstable intertrochanteric fractures. *Injury*, 1995, **26**(9), 611-614.

[236] **Law, M.R., Wald, N.J.** and **Meade, T.W.** Strategies for Prevention of Osteoporosis and Hip Fracture. *British Medical Journal*, 1991, **303**(6800), 453-459.

[237] **Cummings, S.R.** Are Patients with Hip Fractures More Osteoporotic? Review of the Evidence. *The American Journal of Medicine*, 1985, **78**(3), 487-494.

[238] **Melton, L.J.** Hip Fractures: A Worldwide Problem Today and Tomorrow. *Bone*, 1993, **14** Supplement 1, 1-8.

[239] **Cooper, C., Campion, G.** and **Melton, L.J.** Hip Fractures in the Elderly: A World-Wide Projection. *Osteoporosis International*, 1992, **2**(6), 285-289.

[240] **Hollingworth W., Todd C.** and **Parker M.** The Cost of Treating Hip Fractures in the 21st Century. (Presented at the British Orthopaedic Association Meeting, Nottingham, 14-16 September 1994). *The Journal of Bone and Joint Surgery [Br], Orthopaedic Proceedings Supplement*, 1995, **77**(1), 9.

[241] **Rockwood, P.R., Horne, J.G.** and **Cryer, C.** Hip Fractures: A Future Epidemic? *Journal of Orthopaedic Trauma*, 1990, **4**(4), 388-393.

[242] **Ganz, R.** Proximal Femur. In *Manual of Internal Fixation: Techniques Recommended by the AO-ASIF Group*, 3rd edn., M.E. Muller, M. Allgower, R. Schneider & H. Willenegger (eds.), Springer-Verlag, Berlin, 1991, pp 519-534. (ISBN 3-540-52523-8)

[243] **Russell, T.A.** Fractures of Hip & Pelvis. In *Campbell's Operative Orthopaedics*, 8th Edition, A.H. Crenshaw (ed.), Mosby Year Book, St Louis, 1992, pp 895-987. (ISBN 0-8016-1096-6)

[244] **Esser, M.P., Kassab, J.Y.** and **Jones, D.H.A.** Trochanteric Fractures of the Femur: A Randomised Prospective Trial Comparing the Jewett Nail-Plate with the Dynamic Hip Screw. *The Journal of Bone and Joint Surgery [Br]*, 1986, **68**(4), 557-560.

[245] **Rosenblum, S.F., Zuckermann, J.D., Kummer, F.J.** and **Tam, B.S.** A Biomechanical Evaluation of the Gamma Nail. *The Journal of Bone and Joint Surgery [Br]*, 1992, **74**(3), 352-357.

[246] **Schatzker, J.** Screws and Plates and Their Application. In *Manual of Internal Fixation: Techniques Recommended by the AO-ASIF Group*, 3rd edn., M.E. Muller, M. Allgower, R. Schneider & H. Willenegger (eds.), Springer-Verlag, Berlin, 1991, pp 179-290. (ISBN 3-540-52523-8)

[247] **James, E.T.R.** and **Hunter, G.A.** The Treatment of Intertrochanteric Fractures - A Review Article. *Injury*, 1983, **14**(5), 421-431.

[248] **Olerud, C., Rehnberg, L.** and **Hellquist, E.** Internal Fixation of Femoral Neck Fractures: Two Methods Compared. *The Journal of Bone & Joint Surgery [Br]*, 1991, **73**(1), pp 16-19.

[249] **Rehnberg, L.** and **Olerud, C.** The Stability of Femoral Neck Fractures and its Influence on Healing. *The Journal of Bone & Joint Surgery [Br]*, 1989, **71**(2), 173-177.

[250] **Laros, G.S.** Intertrochanteric Fractures. In *Surgery of the Musculoskeletal System,* Volume 3, 2nd Edn., C.M. Evarts (ed.), Churchill Livingstone, New York, 1990, pp 2613-2639. (ISBN 0-44-08516-1)

[251] **Bannister, G.C., Ackroyd, C.E.** and **Langkamer, V.G.** Trochanteric Fractures of the Femur. In *Rob and Smith's Operative Surgery, Orthopaedics Part 1*, 4th edition. G. Bently & R.B. Geer (eds.), Butterworth-Heinemann Ltd., Guilford, Surrey, 1991, pp 209-215. (ISBN 0-7506-1480-3)

[252] **Doppelt, S.H.** The Sliding Compression Screw: Today's Best Answer for Stabilisation of Intertrochanteric Hip Fractures. *Orthopaedic Clinics of North America*, 1980, **11**(3), 507-523.

[253] **Parker, M.J.** Cutting-Out of the Dynamic Hip Screw Related to its Position. *Journal of Bone and Joint Surgery [Br]*, 1992, **74**(4), 625.

[254] **Noordeen, M.H.H, Lavy, C.B.D., Briggs, T.W.R.** and **Roos, M.F.** Unrecognised Joint Penetration in Treatment of Femoral Neck Fractures. *The Journal of Bone & Joint Surgery [Br]*, 1993, **75**(3), 448-449.

[255] **Johansson, A.,** *et al.* Improved Operations for Femoral Neck Fracture. *Acta Orthopaedia Scandanavica*, 1986, **57**(6), 505-509.

APPENDIX 1:

IONISING RADIATION EXPOSURE TO THE ORTHOPAEDIC SURGEON

CONTENTS

		Page
A1.1	INTRODUCTION	A1
A1.2	THE BIOLOGICAL EFFECTS OF IONISING RADIATION	A3
	A1.2.1 Classification of Radiation-Induced Effects	A4
A1.3	MEDICAL UTILISATION OF IONISING RADIATION	A6
	A1.3.1 Diagnostic Radiography	A7
	A1.3.2 Radiation Protection in Radiology	A9
A1.4	RADIATION PROTECTION STANDARDS	A11
	A1.4.1 Current Radiation Protection Standards	A13
	A1.4.2 The Controversy Surrounding Radiation Safety	A14
A1.5	CRITICAL REVIEW OF PUBLISHED EVIDENCE	A16
	A1.5.1 Dosimetric Study Methodology	A18
	A1.5.2 Discussion of Dosimetric Study Results	A21
A1.6	IMPLICATIONS FOR THE MEDROSA PROJECT	A27
A1.7	REFERENCES	A29

A1. IONISING RADIATION EXPOSURE TO THE ORTHOPAEDIC SURGEON

A1.1 Introduction.

As stated in Chapter 1, the aim of the MEDROSA Project is to enhance the performance of fluoroscopically-guided bone drilling procedures, via robotic-assistance. One of the main justifications for this robotic application is the provision of a "remote" operating capability, which allows the surgeon to distance himself from both the patient and the X-ray beam while fluoroscopic screening is taking place. As exposure to radiation obeys an inverse square law (i.e. decreases as the distance from the source increases), surgeon exposure during the robot-assisted stages of the surgical procedure could therefore be reduced to background levels. Consequently, if it can be established that internal fixation of either proximal or diaphysial femoral fractures involves a high risk of exposure to the orthopaedic surgeon, a stronger case can be made for the introduction of the MEDROSA System.

Diagnostic X-ray imaging is fundamental to the practice of modern orthopaedic surgery. Radiographic techniques allow non-invasive visualisation of internal patient anatomy, and can thus reveal the presence of fractures, foreign bodies, or even life threatening diseases. Unfortunately, X-ray radiation is also associated with a wide range of harmful effects. As a consequence, all forms of diagnostic radiography involve direct exposure of the patient to a potentially dangerous form of radiation. However, owing to the positive contribution to patient welfare provided by radiography, patient exposure is deemed to be justifiable. Radiographic examinations tend to be infrequent, and rigorously enforced safety protocols ensure that X-ray images are obtained using the lowest possible levels of irradiation to the patient. As such, the health risks associated with patient exposure are believed to be "acceptable".

In addition to the intentional irradiation of the patient, the medical staff performing radiographic examinations can also receive indirect exposures to X-ray radiation. Although the radiation doses involved are obviously only a fraction of those received by the patient, the detrimental effects of irradiation are known to be cumulative. Thus,

over the course of an entire career, occupationally exposed staff such as physicians and radiographers, could potentially receive significant levels of radiation exposure. For this reason, all staff working with X-ray equipment are required to wear protective lead-equivalent aprons and must also follow strict codes of practice. In the case of preoperative X-ray imaging in the radiology suite, the majority of examinations may be performed from behind lead-lined screens. However, adherence to safety guidelines is not as straightforward in the operating theatre. This has led to speculation that the parts of the surgeon's body not protected by the lead-equivalent apron, may be subjected to unacceptably high radiation doses.

In recent years, concerns over occupational surgeon exposure have increased as a result of the introduction of new "closed" surgical techniques. These procedures are in keeping with the current trend in the other branches of medicine towards minimally invasive surgery (MIS), which is often referred to as "keyhole" surgery, and are aimed at minimising the surgical trauma experienced by the patient. However, in order to obtain the patient-related benefits of closed surgery, the orthopaedic surgeon is frequently required to work without direct physical access to the surgical site. In such cases a direct line-of-sight to the surgical site does not exist, and as a consequence, intraoperative fluoroscopy is often the orthopaedic surgeon's only source of visual information.

The introduction of closed surgical techniques has therefore resulted in a corresponding increase in the use of intraoperative fluoroscopy during certain surgical procedures. One possible trade-off of this growing reliance upon fluoroscopic screening, may be an increased risk of radiation exposure to the orthopaedic surgeon. Current surgical techniques require orthopaedic surgeons to work in close proximity to intraoperative fluoroscopy units, while image acquisition is taking place. Indirect exposure of the surgeon to scattered X-ray radiation, which is caused by partial attenuation and deflection of the X-ray beam within the patient's body, is therefore unavoidable. Given the level of proximity involved, direct exposure to the surgeon, for instance a hand unintentionally straying into the collimated X-ray beam, is also a realistic possibility.

Fears that exposure to X-ray radiation may have become a serious occupational hazard for the orthopaedic surgeon, have prompted a number of independent studies to attempt quantification of the exposure levels received during fluoroscopically-assisted surgical procedures. The aim of this appendix is to evaluate the findings of these dosimetric studies with respect to the two femoral osteosynthesis procedures under investigation by the MEDROSA Project. Section A1.5 therefore presents an abridged version of a critical review of the published evidence **[1]**. The implications of the findings of this review for the MEDROSA Project are then interpreted in Section A1.6.

However, radiation exposure is a complex subject area, with its own units of measurement and associated jargon. For this reason, the first half of this appendix (Sections A1.2 to A1.4) outlines the background knowledge required to allow interpretation of dosimetric study results. Section A1.2 therefore provides a brief overview of radiobiology; the damage that can be caused to the human body by exposure to X-ray radiation. By way of balance, the beneficial diagnostic and therapeutic medical uses of radiation are then briefly outlined in Section A1.3, leading up to a detailed account of the practicalities of radiography. Finally, radiation protection standards, the yardstick against which dosimetric results are evaluated, and dosimetric quantities are then covered in Section A1.4.

A1.2 The Biological Effects of Ionising Radiation.

The term *ionising radiation* is used to describe a form of radiation which possesses sufficient energy to ionise the atoms of an irradiated material. Ionising radiation may be either a type of high frequency electromagnetic radiation (e.g. X-rays and gamma rays), or alternatively, fast moving subatomic particles (e.g. neutrons, alpha-particles or beta-particles) resulting from radioactive decay processes. The origins of ionising radiation can be either naturally occurring (e.g. cosmic radiation, radon gas etc.) or artificial (e.g. X-ray tubes, nuclear reactors etc.) sources.

When ionising radiation passes through biological tissue, damage can occur due to the transfer of energy to the tissue. Energy is absorbed by one of two possible forms of attenuation process; both of which result in ionisation of the atoms within the tissue:

- Photoelectric Absorption:

  Occurs when the energy of an incoming photon of ionising radiation is similar to the binding energy of the electrons in the atoms within the absorbing tissue. Under this set of circumstances, a photon may be fully absorbed by an electron in one of the absorbing tissue's atoms, resulting in ionisation of the atom via ejection of the electron. The liberated electron will then continue to cause further ionisation within the tissue until its kinetic energy is used up.

- Compton Scatter:

  Occurs when the energy of the incoming ionising radiation photon is much greater than the binding energy of the electrons in the atoms within the absorbing tissue. Under this set of circumstances, partial absorption of the photon's energy will occur, resulting in billiard-ball type collisions between incoming photons and the electrons. The outcome of such collisions, is an original photon travelling in one direction with reduced energy, and a recoil electron being liberated in another direction with the remainder of the photon's energy.

Ionisation causes tissue damage due to the associated breaking of the chemical bonds between atoms. If many atoms within a living cell are damaged by ionisation, the cell itself may die, lose its ability to repair and reproduce itself, or even become cancerous. If cell damage occurs on a large scale, a whole organ or body system may ultimately fail.

A1.2.1 Classification of Radiation-Induced Effects.

The amount of damage caused by ionising radiation depends upon both the dose received and the time scale over which the dose is delivered (i.e. the dose rate). In humans, two general categories of radiation-induced effects are therefore observed: *somatic* and *genetic* effects.

The term "somatic" when used in the context of radiation-induced effects, specifically refers to the consequences of exposure to the irradiated individual. By contrast, the term "genetic" is used to describe radiation-induced effects which may be introduced into future generations. These hereditary effects occur as a result of DNA damage, known as mutation, within the nuclei of reproductive cells. The removal of vital strands of DNA from the genes within a cell, impairs the cell's normal function and prevents it from accurately reproducing itself. Significantly, chromosomes containing the genetically impaired genes can also be passed on to children.

A further distinction can also be made, between *deterministic* and *stochastic* effects, with respect to the probability of a radiation-induced effect occurring. Somatic effects are *deterministic* (or non-stochastic), in that above a certain threshold value, the severity of the effect increases with the amount of radiation received by the individual. The threshold value and severity of the effect vary from one individual to the next. Deterministic effects are classified as being either *early* (acute) or *long-term* (chronic), depending upon the length of time before the radiation-induced effects begin to manifest themselves:

- Early (Acute) Effects:

  When substantial whole-body irradiation is received detrimental effects will appear within a few hours or days. Dose-dependent acute effects range from nausea and vomiting, to severe haemorrhaging, central nervous system damage, and ultimately death.

- Long-Term (Chronic) Effects:

  Due to complex delayed biological reactions, somatic effects may not appear until many years after the period of radiation exposure. Examples of long-term effects include cataract formation, sterility, hair loss, skin cancer and leukaemia.

In both cases, the deterministic effects are caused by high-dose exposures leading to either cell death or irreparable cell damage. Theoretically, exposure to these dose levels should never occur in either medical or industrial applications using ionising radiation.

Genetic effects are said to be *stochastic*, in that the probability of occurrence of the effect, but not the severity, increases with increasing exposure of individuals to radiation. Many forms of radiation-induced leukaemia and solid cancers (e.g. breast, lung, thyroid etc.) are also manifestations of stochastic effects. These conditions arise as a consequence of the radiation-induced activation of the mutant genes (oncogenes) which are present in some individuals. Once activated, these mutant genes cause unchecked cell replication, leading to the formation of a tumour or leukaemia. Stochastic effects occur as a result of cell damage rather than cell death. Accordingly, these chance effects are assumed to exhibit no clear threshold dose, below which they cannot occur. On this basis, there is no "safe" dose of ionising radiation exposure, with every exposure incurring an associated risk of stochastic effects.

A1.3 Medical Utilisation of Ionising Radiation.

Medical applications involving the use of ionising radiation, include a wide variety of both diagnostic and therapeutic procedures. The diagnostic applications are mainly confined to the field of medical imaging (*radiology*). Imaging techniques utilise the penetrative power of ionising radiation to produce images of internal patient anatomy, and as such the biological effects of ionising radiation are not directly involved. Consequently, ionising radiation for diagnostic purposes is kept at a minimum in order to avoid deterministic effects and reduce the risks of stochastic effects.

By contrast, large localised radiation doses are deliberately administered in radiation therapy (*radiotherapy*), with deterministic effects actually being the underlying principle of the treatment. Radiation therapy is particularly useful in the treatment of cancer, due to the increased radiosensitivity of many types of cancerous cell. This increased sensitivity arises from the fact that cancer cells proliferate more rapidly than healthy cells, and has allowed treatments to be developed which can kill cancer cells, yet at the same time cause minimal damage to healthy tissue.

A1.3.1 Diagnostic Radiography.

In Section A1.2, the adverse effects of ionising radiation on biological tissue were outlined in terms of attenuation processes and the subsequent ionisation of atoms. In diagnostic radiography, the same attenuation processes are utilised to produce X-ray "images" of the patient's anatomy. The amount of attenuation caused by a material is related to the thickness, atomic number and density of the material. Since the human body is heterogeneous, differential attenuation of an incident X-ray beam occurs during its passage through a patient. The required information about internal anatomical structures is carried in the parts of the beam that emerge unmodified from the patient. In order to visualise the emerging intensity distribution, fluorescent screens are required to convert incident X-ray photons into a visible image. This image may then be stored on photographic film, or alternatively in a digital format.

X-rays are produced when a heavy metal target is bombarded by high energy electrons. This process takes place in an X-ray tube, which consists of an evacuated glass envelope containing a thermionic cathode and an anode assembly. Electrons are emitted by the heated cathode and are then accelerated, in a high intensity electric field, towards a target in the anode assembly. As the electrons come to rest in the target, more than ninety-nine percent of the incident kinetic energy of the electron beam is converted into heat. The remainder is converted into a spectrum of X-ray quanta covering a wide range of energies. By regulating both the X-ray tube voltage and current, thus altering the energy of impact of the electrons striking the target, it is possible to match the properties of the spectrum of X-ray energy with the requirements of a specific diagnostic application.

At low X-ray photon energies (10 keV) **[2]**, photoelectric absorption is the predominant attenuation process, with the extent of absorption being proportional to the cube of the atomic number of the material. By contrast, at high X-ray photon energies (100 keV), attenuation is mainly attributable to Compton scattering and is proportional to the electron density of the absorbing tissue **[3]**. However, with the exception of a few highly specialised investigations, diagnostic radiography is performed at photon energies lying between these two extremes. This is because the

penetrating power of the X-ray quanta falls off as their energy decreases. For this reason, the low-energy components of the X-ray spectrum (below about 50 keV) are largely absorbed by the human body, playing no useful part in the image formation process. In order to minimise the radiation dose received by the patient, aluminium filters are therefore used to modify the X-ray spectrum by absorbing these unwanted "soft" X-rays. As a result, effects attributable to both photoelectric absorption and Compton scattering are present in the diagnostic range of energies.

The contrast observed between bone and soft tissue on a radiograph, is a consequence of the atomic number dependence of photoelectric absorption. The effective atomic number of bone is greater than that of soft tissue (refer to Table A1.1), and as such bone will absorb more X-ray photons. Fewer X-ray photons therefore reach the area of the film cassette behind a bone, reducing the number of photographic silver grains that are developed and creating a transparent or "white" image. For materials with a low effective atomic number the opposite case is true, with maximum film blackening occurring in areas where X-ray photons pass through air only. A radiograph is consequently analogous to a photographic negative.

Table A1.1 : Physical Properties of Selected Body Tissues
(Data adapted from [4])

Material	Effective Atomic Number (Z)	Density ($kg\ m^{-3}$)
Bone	11.6	1300-1850
Muscle	7.4	1000-1040
Water	7.4	1000
Fat	6.3	900
Air	7.6	1.293

Effects associated with Compton scatter can be both beneficial and detrimental to radiography. The density dependence of Compton scattering at higher kilovoltage settings, can be utilised to obtain contrast between tissues with similar effective atomic numbers, provided that they have significantly differing densities. High-kilovoltage chest radiography, is a prime example of a speciality utilising this

phenomenon. As the density of soft tissue is much higher than that of air (refer to Table A1.1), this technique allows air pockets within lung tissue to be visualised. However, in conventional diagnostic radiography, the presence of Compton scatter can result in a significant reduction in the contrast of a radiograph. The billiard-ball type collisions associated with Compton scatter, ensure that the number of deflected photons reaching the film is greater than the number that emerge undeflected from the patient. As these deflected photons have not travelled in a straight line from the X-ray source to the film, they contribute no anatomical information and merely introduce background noise.

In practice, contrast reduction due to scattered radiation is minimised by placing a grid structure between the patient and the film. By providing greater absorption for oblique scattered radiation than for undeflected radiation, the grid ensures that the contribution of scattered photons to the final radiographic image is significantly reduced. In order to prevent the grid from appearing in the X-ray image, it has to be deliberately oscillated during image acquisition. Appropriate measures are also taken to minimise scattered radiation with the aim of limiting patient exposure. In the low photon energy range where photoelectric absorption dominates, X-ray penetration is poor and yields a shallow dose distribution incurring high skin doses. However, high energy Compton scatter radiation is potentially far more dangerous, distributing the dose deeper into the body and effecting organs outside the original irradiated areas. "Beam definers" (lead shutters) are therefore used to collimate the incident X-ray beam, thus limiting the field of exposure to only the body parts which are to be imaged. Shielding of the more radiosensitive organs (e.g. thyroid gland, gonads etc.) may also be necessary, and can be achieved using either external radiopaque protective devices or via careful patient positioning to place less sensitive organs (e.g. muscle and bone) into the beam.

A1.3.2 Radiation Protection in Radiology.

While the dose levels associated with diagnostic procedures (refer to Table A1.2) make deterministic effects unlikely, the possibility of stochastic effects cannot be

ignored. An active system of radiation protection is therefore required in hospitals, in order to limit doses to both patients and staff.

Table A1.2 : Localised Effective Doses Given in Diagnostic Radiology
(Adapted from Plaut [5])

Modality	Procedure	Effective Dose (mSv)
X-ray:	Chest (PA view)	0.02
	Abdomen (AP view)	1.00
CT:	Chest	9.00
	Abdomen and Pelvis	9.50
	Lumbar Spine	6.00

The system of radiation protection recommended by the International Commission on Radiological Protection (ICRP) is based upon three general principles **[6]**:

- *Justification* (of each radiation procedure):
  No practice shall be adopted unless its introduction produces a positive net benefit.

- *Optimisation* (of radiation protection via ALARA):
  The ALARA principle of radiation protection states that "All exposures shall be kept *as low as reasonably achievable*, economic and social factors being taken into account" **[7]**.

- *Limitation* (via individual dose limits):
  The dose to individuals shall not exceed recommended limits.

Legislation based upon these recommendations is laid down in radiation protection standards, and is covered in Section A1.4. Strict adherence to protection standards means that most medical radiation workers receive doses far below recommended guidelines. However, a number of exceptional cases do exist in which high exposures to individual workers can occur as a result of the radiological techniques employed **[8]**. These potentially hazardous techniques include interventional angiography, cardiac catheterisation, and the use of *c-arm fluoroscopy* during certain procedures. In

the absence of safer alternative techniques, periodical monitoring of staff involved in these procedures is advisable.

A1.4 Radiation Protection Standards.

The general aim of radiation protection standards is to prevent detrimental deterministic effects and to limit the probability of stochastic effects to acceptable levels. Quantitative dose limits are therefore recommended, which represent levels of radiation exposure not expected to cause detectable bodily injury to an occupationally exposed person at any time during his lifetime. Such dose limits, are set in accordance with the information available at the time concerning the biological effects of radiation and the consequences for human health.

As knowledge of radiation-induced effects has improved over the years, dose limits have had to be frequently re-evaluated, with the prevailing trend being towards ever increasing risk estimates and a lowering of existing dose limits. Understandably, this has lead to speculation that previous dose limits were dangerously high. A further unfortunate consequence of continuous re-evaluation, has been the adoption of several different dosimetric units and conceptual frameworks over the years (refer to Figure A1.1), severely complicating matters.

The most recent re-evaluation of the way in which radiation is measured and expressed was initiated in 1977 [9]. A conceptual change was deemed necessary because the existing dosimetric quantity, *equivalent dose*, did not take into consideration the different radiation sensitivities of specific biological tissue types. As such, a one sievert equivalent dose does not incur the same risk wherever it is delivered within the body. A new method of calculating the biological effectiveness of radiation exposure, based upon weighted risk factors for individual tissues, was therefore introduced and became known as the *effective dose equivalent* (EDE).

Effective dose equivalent is defined as the sum, over all exposed tissues, of the equivalent dose to each region multiplied by a weighting factor for the tissue in the region:

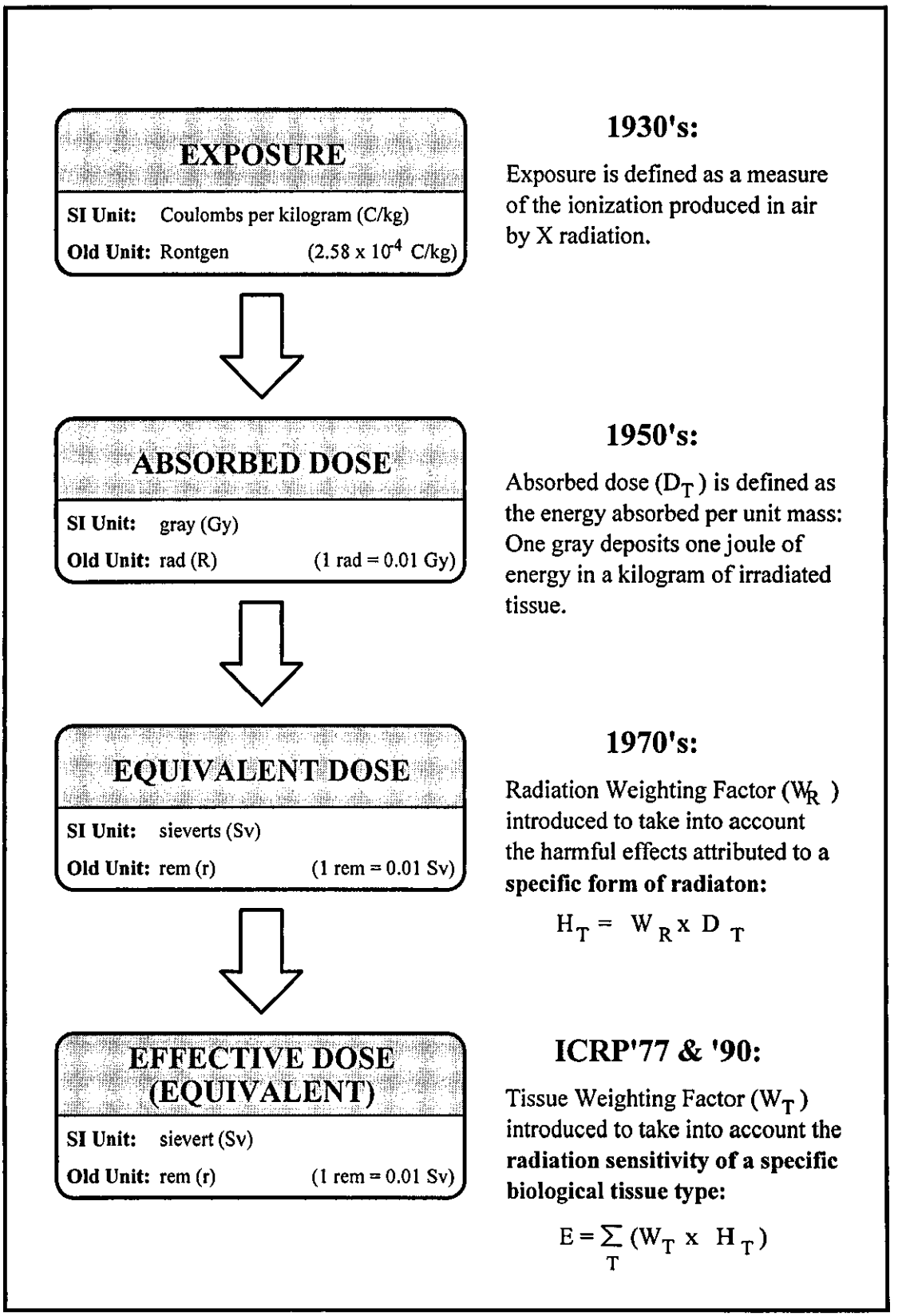


Fig A1.1 : The History of Radiation Protection Standards & Dosimetric Units

$$\text{Effective Dose Equivalent} = \sum_{T} W_T \times H_T$$

Where:

- Tissue weighting factors (W_T) represent the proportion of stochastic risk for the particular biological tissue (T) compared to the total risk for whole-body irradiation.
- H_T represents the equivalent dose in tissue (T).

This choice of definition, allows the risk associated with an exposure to be expressed as a single number, irrespective of whether the radiation is delivered uniformly over the whole body or is localised to one or more specific regions or organs. Consequently, a given effective dose equivalent has the same potential health risk regardless of how an individual received it, thus allowing comparisons between different types of exposures. Following a re-evaluation of the original (1977) tissue weighting factors in 1990 **[10]** the term *effective dose* (E) has superseded "effective dose equivalent" in order to stress and encourage the use of the new weightings.

A1.4.1 Current Radiation Protection Standards.

In the United Kingdom, effective dose limits for occupationally exposed personnel are set out in the Ionising Radiation Regulations 1985 (IRR'85) **[11]**. This legislation is based upon recommendations laid down by the International Commission on Radiological Protection (ICRP) in a 1977 report generally referred to as ICRP(26) **[9]**. As IRR'85 is currently under review with respect to the latest ICRP recommendations (ICRP(60) 1990 **[10]**), a complementary set of interim guidelines has also been issued by the National Radiological Protection Board (NRPB) **[12]**.

The annual effective dose limits for occupationally exposed individuals in the UK, applicable under IRR'85, may be summarised as follows:

Table A1.3 : IRR'85 Effective Dose Limits for Radiation Workers

Body Site	Dose Limit
Whole-body dose:	50 mSv
Dose limits for the lens of the eye:	150 mSv
Doses for individual organs or tissues:	500 mSv

However, the NRPB interim guidelines issued in response to ICRP60 recommend that the average whole-body dose over a five year period should not exceed 20 mSv per year.

Under IRR'85, workers who may be expected to receive more than three-tenths of the annual effective dose limit, are designated as "*classified*" radiation workers and are subject to statutory requirements such as continuous personal monitoring and an annual medical. In addition, the three-tenths level represents the upper limit for occupational exposure to radiation for persons below eighteen years of age. Annual effective dose limits for members of the general public are set at one-tenth of those for radiation workers. Further limitations are also applicable to foetal exposures, as mental retardation and development abnormalities can occur in children exposed during pregnancy.

A1.4.2 The Controversy Surrounding Radiation Safety.

Existing knowledge of radiation-related health risks is mainly derived from follow-up studies involving the survivors of the nuclear explosions at Hiroshima and Nagasaki, and to a lesser extent nuclear accidents such as Three Mile Island and Chernobyl. The extensive data collected during these studies, has lead to a moderate understanding of adverse effects of high-dose radiation delivered at high exposure rates. However, the effects of long-term low-dose occupational exposure are not well understood, due to the fact that only a limited amount of information is available for this type of exposure. As a consequence, existing risk estimates for occupationally exposed personnel have had to be obtained by extrapolating downward from data acquired at higher doses. This use of an extrapolation-based technique, coupled with arguments over the validity of the high-dose data, is the subject of much debate and as a result current radiation risk estimates are highly controversial.

This use of data from studies involving high-dose levels and high-dose rates, to predict radiation-induced effects at low-dose levels and low-dose rates, is subject to substantial scientific uncertainty. Extrapolation from higher doses requires the use of a predictive dose-response model which is believed to properly reflect the relationship

between the likelihood of a radiation-induced effect and the dose received. Quantitative risk estimates are therefore highly dependent upon the type of extrapolation model chosen. In recent years a more conservative linear model has replaced the previously advocated quadratic model, resulting in higher risk estimates for radiation-induced effects at lower doses **[13]**.

A further point of contention, is the fact that the data obtained from high-dose studies contains a number of acknowledged uncertainties. Recent evidence suggests that the average radiation dose received by Japanese A-bomb survivors has previously been overestimated **[10]**. If this is true, risk estimates may have to be re-evaluated by as much as a factor of two or three. Significant differences also exist between the populations exposed to high doses, and the work-force for which risk estimates are required. In Europe and the USA, radiation workers are primarily young or middle-aged males. However, the studied Japanese A-bomb survivors were mainly female, as most of the male population was away at war. The incidence of cancer is also known to vary significantly from one ethnic group to another.

Quantitative risk estimates for low-dose exposure cannot be obtained with absolute certainty, until statistically significant dose-response data is available. At lower dose levels however, the number of people affected by a given level of irradiation is small. For this reason, large-scale epidemiological studies are required to achieve statistically significant results. Regrettably, studies of this size are prohibitively expensive. It is therefore highly likely that the question of what constitutes an "acceptable" level of exposure to radiation, will remain unresolved for some time.

A1.5 Critical Review of Published Evidence.

In view of the uncertainties associated with occupational exposure to ionising radiation, the precise risks to the orthopaedic surgeon cannot be determined at the present time. However, by directly measuring the doses received by surgeons, during consecutive surgical procedures, it is possible to gain an understanding of the potential health hazards. In the belief that the results from this type of dosimetric study may (or may not) substantiate the case for "remote" robotic-assisted surgery, a critical review of the relevant orthopaedic literature has previously been undertaken **[1]**. The remainder of this section briefly outlines the methodology adopted by this review, and then summaries its main findings.

Having limited the search to dosimetric studies which are directly relevant to the MEDROSA applications, in other words, studies which quantify exposure to orthopaedic surgeons during fluoroscopically-assisted osteosynthesis procedures, twelve satisfactory articles were eventually identified. In order to systematically evaluate these articles, they were then divided into two main categories. The details of the eight studies making up the first grouping, which henceforth will be referred to as the *cumulative exposure studies*, are summarised in Table (A1.4).

Table A1.4 :The Cumulative Exposure Studies

1. Smith *et al.* **[14]**	1992	Middlesex & University College Hospitals, London (UK).
2. Goldstone *et al.* **[15]**	1993	Addenbrooke's Hospital, Cambridge (UK).
3. Noordeen *et al.* **[16]**	1993	Middlesex & University College Hospitals, London (UK).
4. Hynes *et al.* **[6]**	1992	Beaumont Hospital, Dublin, Ireland.
5. O'Rourke *et al.* **[17]**	1996	St. Vincent's Hospital, Dublin, Ireland.
6. Barry **[18]**	1984	Stanford University School of Medicine, California (USA).
7. Riley **[19]**	1989	St Luke's Hospital, Cleveland, Ohio (USA).
8. Sanders *et al.* **[20]**	1993	Tampa General Hospital, Tampa, Florida (USA).

These studies were performed with the aim of quantifying the long-term occupational exposure to orthopaedic surgeons. The cumulative dose received by a surgeon/group of surgeons, whilst performing his/their usual range of surgical procedures, has therefore been measured over an extended period of time.

The second category involves the five studies whose details are summarised in Table (A1.5). The aim of these studies was to investigate the radiation exposure to the primary surgeon during a specific fluoroscopically-assisted procedure: intramedullary nailing of femoral shaft fractures. Accordingly, the term *intramedullary nailing studies* will henceforth be used to refer to these studies.

Table A1.5 : The Intramedullary Nailing Studies

1. Miller *et al.* [21]	1983	University of Alabama, Birmingham, Alabama (USA).
2. Levin *et al.* [22]	1987	University of Texas Health Science Centre, Houston (USA).
3. Sugarman *et al.* [23]	1988	Queen's Medical Centre, Nottingham (UK).
4. Coetzee et al. [24]	1992	Univ. of Stellenbosch & Tygerberg Hospital, South Africa.
5. Sanders *et al.* [20]	1993	Tampa General Hospital, Tampa, Florida (USA).

Owing to the inherent nature of orthopaedic surgery, making direct comparisons between these published articles is difficult. Each dosimetric study has it own unique set of "surgical" variables:

- Patient variables (e.g. required surgical procedure, severity of injuries, etc.)
- Surgeon variables (e.g. level of experience, preferred surgical techniques, etc.)
- X-ray equipment variables (e.g. radiation output, performance features, etc.)

The scope of the various studies, in relation to the number of surgical procedures performed and the duration of the study, also varies considerably, as does the dosimetric methodology employed. Nevertheless, within both of the main review categories, similarities between small sub-groups of studies are sufficient to allow partial comparisons to be made. Thus, although the published data is somewhat lacking and occasionally contradictory, it has still been possible to spot the general trends discussed in Section A1.5.2. However, in order to analyse these results, it is important to have a basic understanding of the dosimetric methodology used to obtain them. Section A1.5.1 therefore provides a brief overview of dosimetric techniques and study design.

A1.5.1 Dosimetric Study Methodology.

The first logical step for a "surgeon-based" dosimetric study is to identify the parts of the body which are potentially most at risk. Using appropriate dosimetry techniques, measurements of the radiation doses received in the vicinity of these body parts can then be obtained. The recorded results can then be evaluated against radiation safety legislation and guidelines, in order to estimate the level of risk involved.

As it is standard practice for orthopaedic surgeons to wear a lead-equivalent apron, thus affording adequate protection to the radiosensitive bone marrow (i.e. liver and spleen) and reproductive organs, the whole-body dose to a surgeon should be minimal. As a result, the areas of the body which are not protected by the lead apron become the primary concern. Due to the "hands on" nature of orthopaedic surgery techniques, the surgeon's hands are the part of the body most likely to be directly exposed to the X-ray beam. On this basis, some form of extremity monitoring would seem to be obligatory. Of the remaining unprotected areas, the highly radiosensitive lens of the eye is an obvious candidate for dosimetric evaluation. Similarly, the thyroid gland, which is situated near the larynx, is also known to be particularly susceptible to radiation. Monitoring of the head/neck area would therefore seem to be advisable.

For intraoperative applications, dosimetry normally involves the use of either thermoluminescent dosimeters or photographic film badges. In both cases, these sensor can be incorporated into sterile tags or badges, which are then attached to the surgeon at the relevant body sites. Film badges or TLDs will also be permanently located in the operating theatre in order to measure the background radiation level. This value will then be subtracted from the surgeon-based measurements to yield true exposure levels.

Thermoluminescent dosimeters (TLDs) are based upon the properties of tissue-equivalent materials such as lithium fluoride. When irradiated, electrons in this material are raised from valence bands to optical bands, where they remain trapped. If the material is subsequently heated above 300 °C, the electrons absorb enough energy to return to the valence bands, and in doing so the energy acquired on irradiation is

released as photons of light. By measuring the intensity of the emitted light, and comparing it to that of calibrated TLDs, it is possible to determine the initial radiation dose.

Film badges consist of a piece of photographic film, covered by a variety of filters, housed in a plastic holder. Photographic film is blackened by ionising radiation in a dose- and energy-dependent fashion. The various filters employed, absorb different fractions of the radiation depending upon their constituent material and thickness. By determining the density of the exposed film under different filters, the energy of the radiation encountered can be calculated. The dose received may then be read from a calibration graph of film density versus exposure.

Table A1.6 : MPD Limits of Most Relevance to Orthopaedic Surgeons.
(All limits in millisieverts (mSv))

	Maximum Permissible Dose (MPD) Limits	Classification Limits (3/10ths of MPD)	Routine Monitoring Limits (1/10 th of MPD)
Whole-Body	50	15	5
Lens of the Eye	150	45	15
Thyroid	500	150	50
Extremities	500	150	50

Assuming that quantitative dose data has been obtained, using TLDs & film badges, the results must then be assessed against current radiation protection standards. The annual effective dose limits for occupationally exposed workers, as stated under IRR'85 **[11]**, were summarised in Table A1.3 (Section A1.4.1). Substituting the orthopaedic surgeon as the occupationally exposed worker, the dose limits with most relevance would appear to be those outlined in Table A1.6.

The figures in the first column of Table A1.6 represents the current maximum permissible dose (MPD) limits for occupationally exposed workers. Ideally, the orthopaedic surgeon should not have to be registered as a classified radiation worker. Hence, the 'classification' doses are a more appropriate set of upper limits, against

which the exposure to the orthopaedic surgeon can be evaluated. As stated in the second column of Table A1.6, the classification dose levels are set at three-tenths of the MPD for a given organ or body part. In order to avoid the need for classification, surgeon exposure must be kept below these levels.

The values in the final column of Table A1.6, which are one-tenth of the MPD values, have a double significance to orthopaedic surgery. They are the current legal limits for exposure to members of the general public (i.e. non-radiation workers), and as such represent the limits placed upon patient doses. Secondly, NRPB guidelines **[25]** aimed at limiting patient doses to acceptable levels, have recommended routine dosimetric monitoring for all staff who could reasonably be expected to receive doses in excess of one-tenth of the MPD.

By rigorously designing a dosimetric study, it is possible to collect data that will allow clarification of several pertinent issues. In the case of the prolonged cumulative exposure studies summarised in Table A1.4, answers were generally sought for the following questions:

- Is the orthopaedic surgeon exposed to significant levels of radiation with respect to the various annual MPDs?
- Should the orthopaedic surgeon be registered as a classified radiation worker?
- Should the orthopaedic surgeon be subject to routine personal monitoring?
- In general, what is the "limiting" dose in orthopaedics? (i.e. which body part receives the largest percentage of its own MPD?)

By contrast, dosimetric studies targeted at specific surgical procedures, such as the intramedullary nailing studies summarised in Table A1.5, generally address questions such as:

- For a given surgical procedure, which body part is most at risk?
- What is the average exposure to the high risk body part per procedure?
- How many procedures can be "safely" performed per year before the orthopaedic surgeon exceeds the MPD limits?

More importantly however, the data from quantitative dosimetric studies can result in an increased awareness among surgeons and theatre staff of the hazards of ionising radiation. Recommendations can therefore be made for "safer" surgical technique.

A1.5.2 Discussion of Dosimetric Study Results.

The results from the cumulative exposure studies are summarised in Table A1.7. Looking initially at the measured/extrapolated dosimetric data, it is immediately obvious that the surgeon's hands receive the largest dose of radiation. The available data therefore suggests that exposure to the hands is the limiting dose for orthopaedic surgery. However, while these hand doses may appear to be large in comparison with the torso and head/neck measurements, they represent only 5-10 % of the annual extremity MPD limit (500 mSv) currently applicable in the UK.

By measuring exposure to both the surgeon and the main surgical assistant, Riley **[19]** was also able to show that the surgeon receives an average hand dose per procedure (18.2 mrem) which is double that of the assistant (9.1 mrem); thereby demonstrating the increased risk of exposure to a surgeon resulting from the demands of surgical technique (i.e. proximity to the patient). A number of the studies have also noted a positive correlation between the relative level of experience a surgeon has with a particular surgical procedure, and the length of screening time or the actual radiation dose received. Fluoroscopically-assisted surgical procedures would therefore appear to present a greater risk to inexperienced junior surgeons.

On the basis of these results, fluoroscopically-assisted orthopaedic trauma surgery would appear to be "safe", providing that appropriate safety precautions are taken. Although the majority of the studies did record quantifiable levels of radiation exposure to the surgeon, the actual measured doses are not significant in comparison to current maximum permissible dose limits. In any given year, a surgeon involved with normal patient workloads, is therefore highly unlikely to exceed annual MPD limits; Sanders *et al.* **[20]** found that an orthopaedic surgeon could perform 7614 surgical procedures per year before exceeding the annual extremity MPD limit.

Table A1.7 : Results of Cumulative Exposure Studies

	STUDY NUMBER							
	1	2	3	4	5	6	7	8
STUDY PERIOD (Months):	1	1	1	3	6	12	-----	3
NUMBER OF PROCEDURES:	31	44	45	60	99	249	11	65
TOTAL SCREENING TIME (mins):	-----	-----	26.4	100	109	-----	-----	-----
AVERAGE SCREENING TIME PER PROCEDURE (mins):								
- All Study Procedures.	0.9	1.2	0.6	1.7	1.1	-----	5.1	2.85
- Intramedullary Nailing Procedures.	2.4	2.9	-----	-----	2.95	-----	-----	4.14
- Hip Fracture Procedures (DHS).	-----	1.4	-----	-----	-----	-----	-----	-----
MEASURED/EXTRAPOLATED ANNUAL DOSES: (mSv)								
- Whole-Body (under apron).	2.4	-----	0.6	1.6	1.2	0.05	-----	-----
- Hands.	47.4	30.0	4.1	-----	-----	-----	-----	25.6
- Eyes/Head.	2.4	-----	0.8	-----	1.8	2.27	-----	0.0
PERCENTAGE OF IRR'85 ANNUAL MPD LIMIT:	%	%	%	%	%	%	%	%
- Whole-Body (MPD = 50 mSv)	4.8	-----	1.2	3.2	2.4	0.1	-----	-----
- Hands (MPD = 500 mSv)	9.6	6.0	0.8	-----	-----	------	-----	5.1
- Eyes (MPD = 150 mSv)	1.6	------	0.5	-----	1.2	1.5	-----	-----
AVERAGE DOSE/PROCEDURE:								
- Whole-Body (under apron). (mRem)	-----	-----	-----	-----	-----	-----	0.0	-----
- Whole-Body (over apron). (mRem)	-----	-----	-----	-----	-----	-----	7.3	-----
- Hands. (mRem)	-----	-----	-----	-----	-----	-----	18.2	9.9
- Eyes. (mRem)	-----	-----	-----	-----	-----	-----	0.9	0.0
NUMBER OF PERMISSIBLE PROCEDURES PER YEAR	-----	-----	-----	-----	-----	-----	4000	7614

KEY: Study (1) : Smith *et al.* **[14]**
Study (2) : Goldstone *et al.* **[15]**
Study (3) : Noordeen *et al.* **[16]**
Study (4) : Hynes *et al.* **[6]**
Study (5) : O'Rourke *et al.* **[17]**
Study (6) : Barry **[18]**
Study (7) : Riley **[19]**
Study (8) : Sanders *et al.* **[20]**

Registration of the orthopaedic surgeon as a classified radiation worker would therefore also seem to be unnecessary. Nevertheless, most of the articles under review state that their results should not be interpreted as a licence to become complacent about the dangers of X-ray radiation. Owing to the uncertain nature of stochastic effects at low-dose levels, Hynes *et al.* **[6]** advocate "that it is wise to act on the basis that there is no safe dose of radiation". For similar reasons, Sanders *et al.* **[20]** conclude that the "cavalier use of fluoroscopic equipment ..... is to be condemned".

A more tangible warning of the potential dangers of fluoroscopic assistance, is provided by the data collected during closed locked intramedullary nailing procedures. The four cumulative exposure studies **[14, 15, 17, 20]** involving this nailing procedure, all state that it accounted for the longest fluoroscopic screening times. Sanders *et al.* **[20]** also found that statically interlocked femoral nailing incurred the highest radiation exposure levels (100 mrem/procedure) to the surgeon's hands; while the average exposure for the study as a whole was only 9.9 mrem/procedure. Intramedullary nailing, would therefore appear to represent one area of orthopaedic trauma surgery where significant exposure levels can be encountered.

This point is further validated by the results from the actual intramedullary nailing studies (summarised in Table A1.8). The average screening time data from these studies once again demonstrates the influence of surgeon expertise. The study by Coetzee *et al.* **[24]**, which was specifically conducted to investigate irradiation during nailing procedures performed by inexperienced surgeons, reports significantly longer average screening times than any other study. Conversely, the shortest average screening times were recorded by Sanders *et al.* **[20]**, who state that only experienced orthopaedic surgeons took part in their study.

However, the wide spread of fluoroscopy times recorded during these five dosimetric studies is not unusual. A number of non-dosimetric studies have also reported fluoroscopy times which are comparable to those of Coetzee *et al.* **[24]**. Hudson **[26]** states that in his experience, screening times for statically locked femoral nailing can be as high as thirty minutes. Likewise, Roberts *et al.* **[27]** report that the same

procedure can take up to twenty minutes. At the other extreme, Kempf *et al.* **[28]** reported an average screening time of three minutes forty-three seconds, for a mixed series of 452 locked nailings of the femur.

Turning now to the dosimetric data provided in Table A1.8, it can be seen that once again the extremity dose is the limiting dose. It is also apparent that even the excessive screening times recorded by Coetzee *et al.* **[24]** have led to average radiation doses per procedure which are small in comparison to annual MPD limits. Since locked intramedullary nailing is not a common procedure, at normal patient workloads, it is therefore unlikely that an individual orthopaedic surgeon will exceed the annual MPD or classification limits. However, in order to reach the level of exposure at which the NRPB **[25]** recommend routine dosimetric monitoring should be instigated (one-tenth of the MPD limit), only thirty-five procedures would have to be performed on the basis of the data provided by Coetzee *et al.* **[24]**.

The detailed dosimetric results provided by Sanders *et al.* **[20]** also indicate an important fact in relation to the distal interlocking process. During their study, five statically interlocked femoral nailing procedures incurred an average exposure to the surgeon's hands of 100 mrem (1 mSv) per procedure. However, the corresponding figure for six proximal dynamically locked femoral nailings was only 10 mrem (0.1 mSv). On the basis of this information, the use of distally interlocked intramedullary nails to stabilise complex femoral fractures, appears to involve unacceptably high levels of exposure. Locked intramedullary nailing must therefore be regarded as a potentially dangerous procedure, which is particularly hazardous for inexperienced surgeons.

In addition, the possibility of the surgeon's hands unintentionally straying into the primary X-ray beam cannot be overlooked. The dangers of direct exposure are highlighted by the patient-based dosimetry results. Sugarman *et al.* **[23]** recorded an average exposure of 1.81 mSv to the patient's groin. Since the annual whole-body MPD limit for members of the public is 5 mSv in the UK (i.e. one-tenth of the MPD for classified radiation workers), only two or three such procedures are allowable.

Table A1.8 : Results of Intramedullary Nailing Studies

	Study Number				
	1	2	3	4	5
Average Screening Time Per Femoral Nailing Procedure:	(mins)	(mins)	(mins)	(mins)	(mins)
- Unlocked.	0.4	-----	-----	6.90	-----
- Proximal Locking (Dynamic).	-----	3.64	-----	-----	2.37
- Distal Locking (Dynamic).	-----	5.12	-----	12.50	-----
- Proximal & Distal Locking. (Static).	-----	12.60	12.08	17.87	6.26
Average Radiation Dose Per Procedure: (mSv unless stated)	mR				
- Dominant Hand	6.3	0.25	0.71	2.10	1.00
- Non-Dominant Hand	4.7	-----	0.39	0.64	-----
- Collar (Thyroid)	7.7	0.15	0.44	0.14	0.00
- Eyes/Head	14.0	-----	-----	0.14	-----
- Body (under lead apron)	3.0	-----	0.11	0.05	-----
- Body (over lead apron)	22.3	-----	-----	-----	-----
Limiting Dose (mSv):					
- Body Part:	Eyes	Hand	Hand	Hand	Hand
- Quoted Annual MPD Limit.	50	750	500	750	750
- Average Dose Per Procedure.	-----	-----	0.71	2.10	1.00
Safe Permissible Number of Procedures Per Year :					
- At Quoted MPD Limit.	-----	-----	700	350	750
- At Classification Limit (3/10 ths).	----	-----	210	105	225
- At Routine Monitoring Limit (1/10 th).	-----	-----	70	35	75
Average Patient Dose Per Procedure: (mSv unless stated)	mR				
- Waist/Fracture Site	67.6	-----	1.81	17.90	-----
- Whole-Body MPD Limit	-----	-----	5.00	50.00	-----
- Safe Permissible Procedures/Year	-----	-----	2/3	3	-----

KEY: Study (1) : Miller *et al.* **[21]**
Study (2) : Levin *et al.* **[22]**
Study (3) : Sugarman *et al.* **[23]**
Study (4) : Coetzee *et al.* **[24]**
Study (5) : Sanders *et al.* **[20]**

Levin *et al.* [22] therefore state that the orthopaedic surgeon must be extremely careful how he positions his hands and body with respect to the X-ray beam. The inverse square law of radiation exposure has also been frequently cited, along with the need for the surgeon to distance himself from the X-ray beam whenever practicable. The wearing of additional protective clothing (e.g. thyroid shields and lead impregnated surgical gloves) has also been suggested by Miller *et al.* [21], and strongly recommended to inexperienced surgeons by Coetzee *et al.* [24].

Radiation protection concepts aimed at minimising the risks associated with external radiation sources, have traditionally involved three methods of protection: *shielding*, *time* and *distance*. Applying these concepts to intraoperative fluoroscopy, the knowledge derived from these dosimetric studies leads to the following recommendations:

- Shielding:

  A protective lead apron must be worn by orthopaedic trauma surgeons during all fluoroscopically-assisted surgical procedures. During high risk procedures, such as locked intramedullary nailing of the femur, the wearing of additional protective clothing is also advisable, particularly for inexperienced surgeons. Collimation of the X-ray beam should also be performed to minimise scatter radiation.

- Time:

  Several measures can be taken to reduce fluoroscopic screening times. Ideally orthopaedic trauma surgery should take advantage of the performance features of modern fluoroscopic equipment. The combination of pulsed mode fluoroscopy and image capture should be used whenever possible. If older continuous fluoroscopy systems are to be used, intermittent screening or even surgeon control of the footswitch are advisable.

- Distance:

  The intensity of X-ray radiation obeys an inverse square law with respect to distance. Any direct exposure to the surgeon is therefore significantly influenced by the tube-to-surgeon distance. Scatter radiation, which is the major contributor to

surgeon irradiation, has also been shown to rapidly decrease with increasing distance from the patient **[17, 21, 29, 30]**. As a consequence, whenever possible, surgeons should attempt to maximise their distance from the X-ray source and patient during screening.

The need for surgeons to distance themselves from the X-ray source and the patient during screening, is undoubtedly the most important of these recommendations.

A1.6 Implications for the MEDROSA Project.

The ideal solution to the radiation-related problems of fluoroscopically-assisted orthopaedic surgery, is the removal of the surgeon from the vicinity of the primary (i.e. the X-ray tube) and secondary (i.e. the patient) radiation sources at the time of screening. This "remote" surgery capability can be realised through the introduction of a robotic manipulator into the operating theatre. In addition to the many performance-related benefits associated with robotics, by taking the place of the surgeon during screening, a surgical manipulator also has the potential to reduce surgeon exposure to the background level of the operating theatre.

In order to gain clinical acceptance, a surgical robot must be safe, operating theatre compatible, easy to use, and cost-effective. An evaluation of cost-effectiveness, offsets the potential benefits of a system against any additional costs which are incurred. Before developing a robotic-assisted surgery system, it must therefore be established that a need exists for the benefits offered by the system. Performance capabilities which have immediate financial implications, such as shorter procedural times or improved surgical accuracy, which can lead to reduced patient morbidity and lower rates of corrective surgery, obviously take precedence when attempting to justify a system's initial capital investment and running costs. However, medical systems should not be judged purely in financial terms. The "perceived" value of the system is also an important factor.

In the short-term, irradiation of the orthopaedic trauma surgeon is neither life threatening nor measurable in financial terms. It is therefore difficult to justify the

expense of robotic-assisted surgery purely on the grounds of radiation exposure to the surgeon. Nevertheless, removal of the surgeon from the hazardous working volume surrounding the patient, is a commendable research goal. More importantly, a "safe" remote surgery capability, contributes significantly to the perceived value of the system. Hence, a robotic system which can match or exceed manually achievable surgical precision, whilst also offering reduced exposure levels, is an attractive proposition. As a consequence, the reviewed dosimetric studies provide valuable information, which may or may not partially corroborate the need for robotic-assisted orthopaedic surgery.

Since all exposures to ionising radiation are acknowledged to have an associated risk of stochastic effects, it could be argued that the radiation-based justification of robotic-assisted surgery is valid, irrespective of the actual doses received during surgery. However, annual classification MPD limits provide a more credible measure against which to assess the radiation-based argument. On the basis of this evaluation criterion, the internal fixation of proximal femoral fractures is not a prime candidate for robotic-assisted surgery. Radiation doses to the surgeon during DHS insertion have been shown to be minimal, and as such, even a surgeon who specialises in this very common surgical procedure is unlikely to perform sufficient procedures to exceed the classification limits. Semi-automation and improved surgical precession are therefore more realistic grounds for the justification of robotic hip fracture surgery.

By contrast, dosimetric data does substantiate claims for robotically-assisted locked intramedullary nailing of femoral shaft fractures. In relation to surgeon irradiation, locked femoral nailing appears to be the worst case orthopaedic procedure. Extremity doses of 1 mSv per procedure are not unusual for statically locked nailings, and higher dose levels are frequently received by inexperienced surgeons. Direct exposure to the surgeon's hands is also a realistic possibility during locked nailing procedures. However, the most significant dosimetric aspect with respect to robotic-assisted surgery, is the fact that distal interlocking accounts for at least half of the total surgeon irradiation during locked femoral nailing procedures. By coincidence, the most technically demanding and hazardous stage of the intramedullary nailing procedure for

a surgeon, is therefore also the simplest stage of the procedure to automate using a robot. As a consequence, robotic drilling of the distal interlocking holes, is warranted in terms of both the technical difficulty *and* the high levels of surgeon irradiation currently experienced during intramedullary nailing procedures.

A1.7 References.

[1] **Browbank, I.** Ionising Radiation Exposure to the Orthopaedic Surgeon: A Critical Review of the Published Evidence. *Internal Publication: Department of Mechanical, Loughborough University*, 1997.

[2] **Garrett, J.A.** and **Smithson, P.H.** Conventional X-Ray Imaging. *IEE Proceedings, Part A: Physical Science, Measurement & Instrumentation, Management & Education,* 1987, **134**(2), 107-114.

[3] **Hay, G.A.** X-Ray Imaging. *Journal of Physics, Part E: Scientific Instrumentation,* 1978, **11**, 377-386.

[4] **Moore K.L.** *Clinically Oriented Anatomy.* Williams & Wilkins, Baltimore, 1983 (ISBN 0-683-06146-1).

[5] **Plaut, S.** *Radiation Protection in the X-ray Department,* Butterworth Heinemann, Oxford, 1993. (ISBN 0-7506-0606-1)

[6] **Hynes, D.E., Conere, T., Mee, M.B.** and **Cashman, W.F.** Ionising Radiation and the Orthopaedic Surgeon. *The Journal of Bone and Joint Surgery [Br]*, 1992, **74**(3), 332-334.

[7] **Russell, J.G.B.** Radiation Protection in Radiography in the United Kingdom. The British Journal of Radiology, 1986, **59**(704), 747-748.

[8] **Hendee, W.R.** Estimation of Radiation Risks: BEIR V and its Significance for Medicine. *Journal of the American Medical Association*, 1992, **268**(5), 620-624.

[9] **ICRP(26):** Recommendations of the ICRP (Pergamon, Oxford, 1977). ICRP Publication No. 26.

[10] **ICRP(60):** 1990 Recommendations of the ICRP (Pergamon, NY). ICRP Publication No. 60 (Ann. ICRP v.21 1991)

[11] **IRR'85:** Ionising Radiation Regulations (1985) (HMSO).

[12] **NRPB'93:** Documents of the NRPB, 1993, Vol 4, nos. 1-3.

[13] **Hendee, W.R.** History, Current Status, and Trends of Radiation Protection Standards. *Medical Physics*, 1993, **20**(5), 1303-1314.

[14] **Smith, G.L., Lavy, C.B.D., Briggs, T.W.R.** and **Noordeen, H.** Ionising Radiation: Are Orthopaedic Surgeons at Risk? *Annals of the Royal College of Surgeons of England*, 1992, **74**(5), 326-328.

[15] **Goldstone, K.E., Wright, I.H.** and **Cohen, B.** Radiation Exposure to the Hands of Orthopaedic Surgeons During Procedures Under Fluoroscopic X-ray Control. *The British Journal of Radiology*, 1993, **66**(790), 899-901.

[16] **Noordeen, M.H.H.,** *et al.* Hazard of Ionising-Radiation to Trauma Surgeons: Reducing the Risk. *Injury*, 1993, **24**(8), 562-564.

[17] **O'Rourke, P.J.,** *et al.* Risks of Radiation Exposure to Orthopaedic Surgeons. *Journal of the Royal College of Surgeons of Edinburgh*, 1996, **41**(1), 40-43

[18] **Barry, T.P.** Radiation Exposure to an Orthopaedic Surgeon. *Clinical Orthopaedics & Related Research*, 1984, **182**, 160-164.

[19] **Riley, S.A.** Radiation Exposure From Fluoroscopy During Orthopaedic Surgical Procedures. *Clinical Orthopaedics and Related Research*, 1989, **248**, 257-260.

[20] **Sanders, R.,** *et al.* Exposure of the Orthopaedic Surgeon to Radiation. *Journal of Bone and Joint Surgery [Am]*, 1993, **75**(3), 326-330.

[21] **Miller, M.E.,** *et al.* Radiation Exposure and Associated Risks to Operating-Room Personnel During Use of Fluoroscopic Guidance for Selected Orthopaedic Surgical Procedures. *The Journal of Bone and Joint Surgery*, 1983, **65**(1), 1-4.

[22] **Levin, P.E., Schoen, R.W.** and **Browner, B.D.** Radiation Exposure to the Surgeon During Closed Interlocking Intramedullary Nailing. *The Journal of Bone and Joint Surgery [Am]*, 1987, **69**(5), 761-766.

[23] **Sugarman, I.D., Adam, I.** and **Bunker, T.D.** Radiation Dosage During AO Locking Femoral Nailing. *Injury*, 1988, **19**(5), 336-338.

[24] **Coetzee, J. C.** and **Van Der Merwe, E. J.** Exposure of Surgeons-in-Training to Radiation During Intramedullary Fixation of Femoral Shaft Fractures. *South African Medical Journal*, 1992, **81**(6), 312-314.

[25] **NRPB'88:** NRPB Guidance Notes for the Protection of Persons Against Ionising Radiations Arising from Medical & Dental Use. (HMSO, London, 1988).

[26] **Hudson, I.** Locking Nailing: An Aid to Distal Targeting. *Injury*, 1989, **20**(3), 129-130.

[27] **Roberts. P.J.** and **Jaffray, D.** Radiation Doses to Surgeons (Letter to the Editor). *Injury*, 1989, **20**(6), 383.

[28] **Kempf, I., Grosse, A.** and **Beck, G.** Closed Locked Intramedullary Nailing - Its Application to Comminuted Fractures of the Femur. *The Journal of Bone and Joint Surgery [Am]*, 1985, **67**(5), 709-720.

[29] **Giachino, A.A.** and **Cheng. M.** Irradiation of the Surgeon During Pinning of Femoral Fractures. *The Journal of Bone and Joint Surgery [Br]*, **62**(2), 227-229.

[30] **Skjeldal, S.** and **Backe, S.** Interlocking Medullary Nails - Radiation Doses in Distal Targeting. *Archives of Orthopaedic and Traumatic Surgery*, 1987, **106**(3), 179-181.

APPENDIX 2:

THE MEDROSA PROJECT

CONTENTS

		Page
A2.1	INTRODUCTION	A31
A2.2	PROJECT BACKGROUND	A32
A2.3	PROJECT AIMS AND LIMITATIONS	A33
A2.4	RESEARCH PROGRESS TO DATE (1992-1996)	A36
	A2.4.1 Purpose-Built Manipulator	A38
	A2.4.2 Drilling Unit	A40
	A2.4.3 Safety Features	A42
	A2.4.4 Sterility Issues	A43
A2.5	FUTURE DEVELOPMENT METHODOLOGY	A46
A2.6	CONCLUDING REMARKS	A48
A2.7	REFERENCES	A52

A2. THE MEDROSA PROJECT

A2.1 Introduction.

The research work documented by this thesis, also constitutes an integral part of a much larger robotic-assisted surgery project, which is known by the acronym-based title "MEDROSA". In order to put the thesis into a better context, it is obviously advantageous to have a working knowledge of the MEDROSA Project as a whole. However, it is difficult to describe the other aspects of the MEDROSA Project without detracting from the research topics included within the scope of the thesis, and as a consequence, coverage of the MEDROSA Project within the main thesis text was deliberately restricted to a single section of Chapter 1 (Section 1.3.1). For reasons of clarity, the "robotic" aspects of the MEDROSA Project are therefore described separately in this appendix, with the aim of expanding upon the details presented in Chapter 1.

The majority of robotic-assisted surgery applications involve two main stages:

- Preoperative Surgical Planning:
  Surgical trajectories are planned using graphical human-computer interfaces.
- Intraoperative Robotic Implementation of the Surgical Plan:
  The surgical plan is downloaded to a robot controller prior to accurate execution by a high precision manipulator.

For the osteosynthesis applications under investigation by the MEDROSA Project, surgical planning involves intraoperative quantification of the standard biplanar fluoroscopic examination process, using software-based calibration techniques. As this is essentially the main thrust of the current thesis, surgical planning will not be covered in great detail by this appendix. The appendix therefore concentrates upon research topics relating to the robotic execution of the surgical plan under operating theatre conditions. Accordingly, a description is provided of the design and manufacture of a purpose-built surgical manipulator. The critical areas of safety and

sterility are also addressed, along with integration of the vision and robotic components of the system.

A brief outline of the orthopaedic background of the MEDROSA Project is provided in Section A2.2. Included within this section is a description of a number of problems associated with current surgical practice, which have been targeted for solution by robotic-assistance. Section A2.3 then details the aims and limitations of the MEDROSA Project in relation to these problems, along with the potential benefits of the robotic system. The remaining sections of the appendix then cover the progress which has been made to date by the MEDROSA Project (Section A2.4), and the future research methodology which will be adopted to ensure that the system achieves full clinical acceptability and utilisation (Section A2.5).

A2.2 Project Background.

Orthopaedics is primarily concerned with the correction of deformity, diseases of bone and joints, and injuries to the musculoskeletal system **[1]**. As such, orthopaedic surgeons are now required to master a wide variety of surgical techniques, including microsurgery, arthroscopy (endoscopic examination of a joint) and ligament reconstruction surgery. However, despite the development of new surgical techniques and equipment, traditional bone machining processes using hand-held air-powered tools, still remain fundamental to the practice of orthopaedic surgery.

Drilling is the most widely used bone machining process in orthopaedic surgery, and often constitutes a preliminary stage in the insertion of pins/screws during the internal fixation (*Osteosynthesis*) of fractures by mechanical devices. Unfortunately, difficulties in maintaining freehand control of a drill, even when using a hand-held drill guide, can often result in less than optimum drilling results being obtained. A tendency for the drill bit to "walk" or slip on the surface of a bone is also often observed in practice. As the failure rates of many internal fixation procedures are directly related to the accuracy of the initial drill bit location, there is consequently considerable scope for improvement of the manual drilling process.

The problems associated with manual surgical drilling have been further compounded in recent years, as a result of changes in clinical procedure towards the use of "closed" surgical techniques. The aim of these procedures is to reduce the surgical trauma experienced by the patient, through gaining indirect access to the fracture site via a small incision. In theory, this results in lower intraoperative blood loss and reduced wound infection rates, speeding up postoperative recovery. However, as the fracture site is not directly visible, these clinical benefits have to be obtained at the expense of a heavy reliance upon intraoperative fluoroscopic X-ray imaging. Drilling processes are therefore frequently performed with near orthogonal pairs of X-ray images providing the only means of guidance. This has led to growing concerns among orthopaedic surgeons over their ever increasing occupational exposure to X-ray radiation (refer to Appendix 1 for an in-depth analysis of the surgeon irradiation problem).

The use of two-dimensional fluoroscopic images to guide a three-dimensional drilling process, also incurs a depth perception problem. Fluoroscopically-assisted drilling therefore relies heavily upon the surgeon's ability to mentally integrate 2-D views into a 3-D anatomical "model", and as a result, the outcome of these subjective procedures is heavily dependent upon the skill and experience of the individual surgeon. Owing to the trial-and-error nature of the task, it can take trainee surgeons several procedures to perfect this type of surgical skill. Throughout this long learning curve period, inexperienced surgeons are inclined to overuse fluoroscopy, and may also require several attempts at the procedure before satisfactory surgical results are achieved.

A2.3 Project Aims & Limitations.

The MEDROSA Project was instigated following an approach by Mr. R.G. Shedden, a consultant orthopaedic surgeon working at Mayday University Hospital (Croydon A.H.A.), to the Department of Mechanical Engineering at Loughborough University. Having experienced firsthand the problems outlined in Section A2.2, Mr. Shedden was interested to find out if manually obtained drilling results could be improved upon via the introduction of new technologies into the operating room. An initial investigation

of the feasibility of computer-assisted and robotic-assisted orthopaedic surgery, was therefore undertaken by Professor J.R. Hewit and Dr. K. Bouazza-Marouf [2].

This feasibility study found that a vision-guided robotic-assisted surgery system had much to offer the orthopaedic trauma surgeon. The depth perception problem encountered by inexperienced surgeons, was believed to be surmountable by the application of three-dimensional machine vision calibration techniques to the standard fluoroscopic examination. Quantification of intraoperative X-ray imaging using these techniques, reduces much of the subjectivity associated with surgical planning. De-skilling of the fluoroscopically-assisted drilling process was also proposed, via the use of a tool positioning robotic manipulator incorporating a drilling unit end-effector. Theoretically, such a system can provide a high precision drilling capability, resulting in right-first-time drilling scenarios, which have the potential to reduce both overall procedural times and fracture fixation failure rates. It was also noted that a robotic system would allow the surgeon to be distant from the patient at the time of fluoroscopic imaging, thereby significantly reducing his level of radiation exposure. Finally, given that the drilling of bone under bi-planar X-ray guidance is common to many orthopaedic surgery procedures, it was believed that a generic cost effective system could ultimately be implemented.

In the light of this positive feasibility study, an initial pilot study was proposed to evaluate the application of robotic-assisted surgery to two specific femoral fracture procedures:

- Locked Intramedullary Nailing of Femoral Shaft Fractures.

  This procedure initially involves the insertion of a nail (stainless steel tube) into the marrow (*medullary*) cavity of the femur, followed by the interlocking of the femur/nail using screws (refer to Chapter 6 for additional information). The 'blind' insertion of distal locking screws under image intensifier guidance, is widely recognised as being one of the most technically demanding orthopaedic surgery procedures. In addition to the complexity of the task, studies have also shown that the surgeon receives significant levels of radiation exposure during distal locking (refer to Appendix 1).

- Sliding Hip Screw Fixation of Proximal Femoral Fractures.

  Proximal femoral fractures (*hip fractures*) are normally associated with the elderly, and in particular women suffering from osteoporosis. As the average age of the world's populations increases, it is predicted that the already substantial burden placed upon medical resources by this very common injury (1.66 million hip fractures were reported world-wide in 1990 and 6.26 million are predicted for 2050 **[3]**) will increase considerably unless treatment patterns change. In addition to fracture prevention strategies, such as screening for osteoporosis, improved surgical techniques are therefore also required **[4]**. Details of the current surgical techniques used to insert sliding hip screw devices can be found in Chapter 7.

Having obtained partial funding for this pilot study from The Wishbone Trust of the British Orthopaedic Association, Professor Hewit and Dr. Bouazza-Marouf proceeded to develop and manufacture a prototype vision-guided robotic system at Loughborough University, with the aim of augmenting the performance of orthopaedic surgeons **[5]**.

The initial specification for this system, required satisfactory solutions to be obtained for problems relating to the following performance criteria:

- Patient and theatre staff safety.
- Prevention of contamination in the operating room by maintaining surgical sterility.
- Minimal deviation from existing surgical practice.
- No modification of existing operating room equipment.
- Intraoperative registration between the patient and manipulator.
- Compatibility with a wide range of drill bit and fixation device designs.
- Any additional cost or lengthening of the operation time, must be fully justifiable in terms of clinical benefits gained.

The overriding need to maintain patient safety, whilst at the same time achieving operating room compatibility, precluded the use of a commercially available robot. The mechatronics solution of a kinematic configuration specifically designed for orthopaedic surgery was therefore implemented (hence the acronym MEDROSA:

MEchatronic Design of a Robot for Orthopaedic Surgery Assistance), with the intention of performing the fluoroscopically-guided drilling stages of the two orthopaedic procedures under investigation.

A2.4 Research Progress to Date (1992-96).

Having performed a detailed evaluation of the two osteosynthesis procedures under investigation, specifications were produced for the necessary software and electro-mechanical components of the system. The mechatronics design philosophy was then put into practice to manufacture a custom-built prototype system, which was completed in time to be exhibited at the 1993 British Orthopaedic Association Annual Meeting (13-15th September 1993, Torquay). This early prototype system has since undergone refinement during a set of preliminary laboratory trials.

The prototype MEDROSA system consists of the following major components:

- A purpose-built manipulator incorporating a force feedback drilling unit.
- A PC-based robot controller unit and electronics cabinet.
- A PC-based machine vision sub-system which interfaces to existing X-ray imaging equipment (a video link is established between the C-arm unit and the PC-based frame grabber card).

The interaction of these components with the surgeon and the patient are summarised in block diagram form in Figure A2.1. With reference to this figure, a typical robotic-assisted drilling procedure will involve the following processes. The manipulator is initially wheeled over to the operating table, and is coarsely aligned with the patient's injured hip or femur. The manipulator's base unit is then pneumatically locked into position. As discussed in the main thesis text (refer to Chapter 4), intraoperative registration between the manipulator and the patient, is then established using the C-arm fluoroscopy unit as an intermediate sensing device. In practice, this requires the use of a perspex calibration frame containing radiopaque calibration markers, which is held in position around the patient by the manipulator.

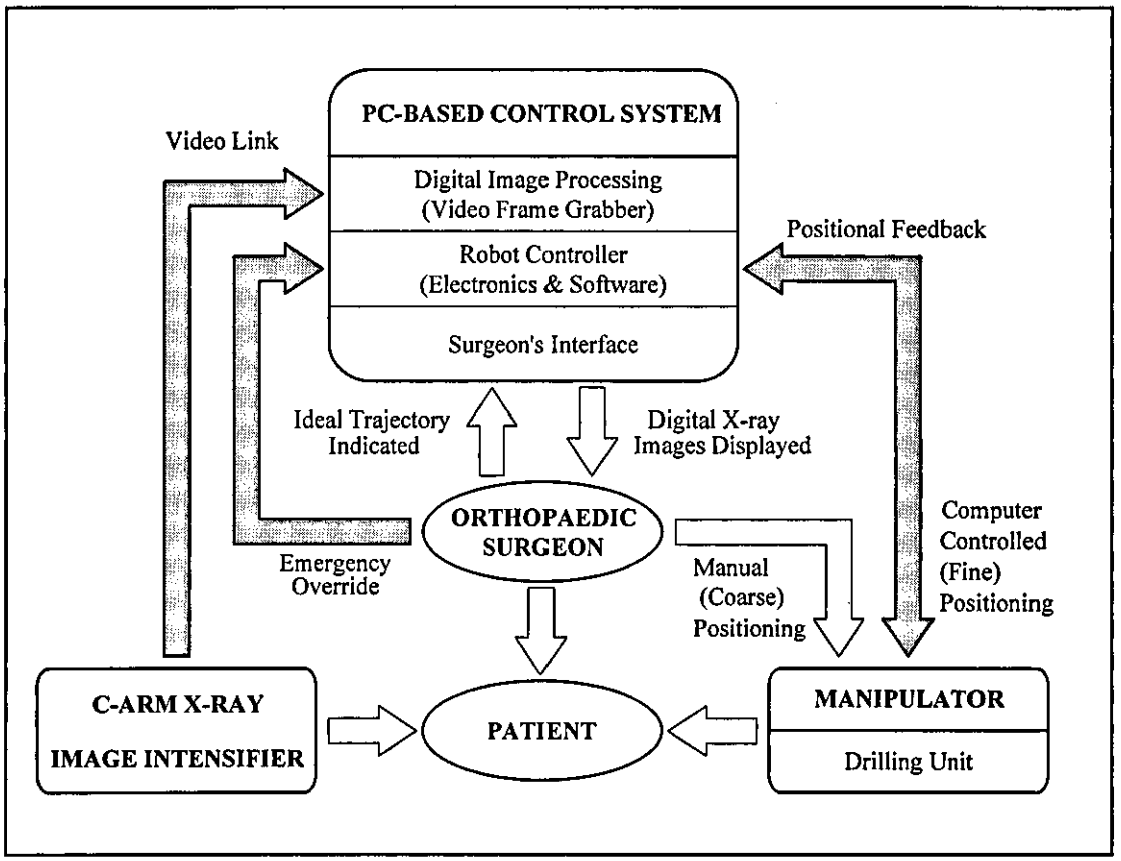


Fig. A2.1 : Block Diagram of the Prototype MEDROSA System

The surgeon is then required to indicate "ideal" drilling trajectories on both of the standard fluoroscopic views using a graphical human computer interface. A triangulation-based reconstruction process is then performed by the vision sub-system, in order to transform this 2-D drilling trajectory data into a 3-D trajectory in terms of the perspex calibration frame's coordinate system. Surgical planning is now complete, and the machine vision sub-system outputs this 3-D drilling trajectory data to the robot controller.

Before the surgical plan can be implemented, the drilling trajectory must first be transformed into the manipulator's coordinate system. This is achieved by the robot controller using a combination of preoperative robot calibration and intraoperative forward kinematics. Computer controlled fine positioning of the manipulator is then performed to align the drilling unit with the ideal drilling trajectory. The manipulator's degrees-of-freedom are then locked into position. At this stage of the procedure, the surgeon has the option of either using the drilling unit as a guide for

manual drilling, or alternatively, performing an automated drilling procedure using a surgical drill mounted on a motorised drill-feed incorporated into the drilling unit. If the latter option is followed, the surgeon monitors the drilling unit's progress and is provided with an emergency override device.

More detailed descriptions of the manipulator-related aspects of the prototype system are presented in the following sub-sections.

A2.4.1 Purpose-Built Prototype Manipulator.

The alignment of a manipulator-mounted surgical drill (or drill guide) with a three-dimensional drilling trajectory, requires a combination of two linear and two rotary positioning degrees-of-freedom. In order to facilitate an automated invasive drilling option, an additional linear degree-of-freedom is also required to feed the drill bit into and out of the bone. Having considered several different kinematic configurations incorporating five active degrees-of-freedom, the application specific arrangement depicted schematically in Figure A2.2 was eventually chosen for the prototype MEDROSA manipulator.

With reference to Figure A2.2, the degrees-of-freedom labelled one to four, are the manipulator's tool positioning motions. Each of the manipulator's degrees-of-freedom are driven by a single stepper motor. In the case of the two linear degrees-of-freedom, a horizontal drive (1) and a vertical drive (2), the stepper motors drive a backlash-free ballscrew, with the motions being constrained by a parallel pair of linear motion guides. The first rotary motion, designated as the pan drive (3), is driven by a stepper motor through a reduction gearbox fitted with an anti-backlash device (i.e. a harmonic drive). The second rotary motion, designated as the tilt drive (4), uses a similar ballscrew/motion guide configuration to the two linear degrees-of-freedom, and obtains the required rotary motion using a pivot and connecting rod arrangement.

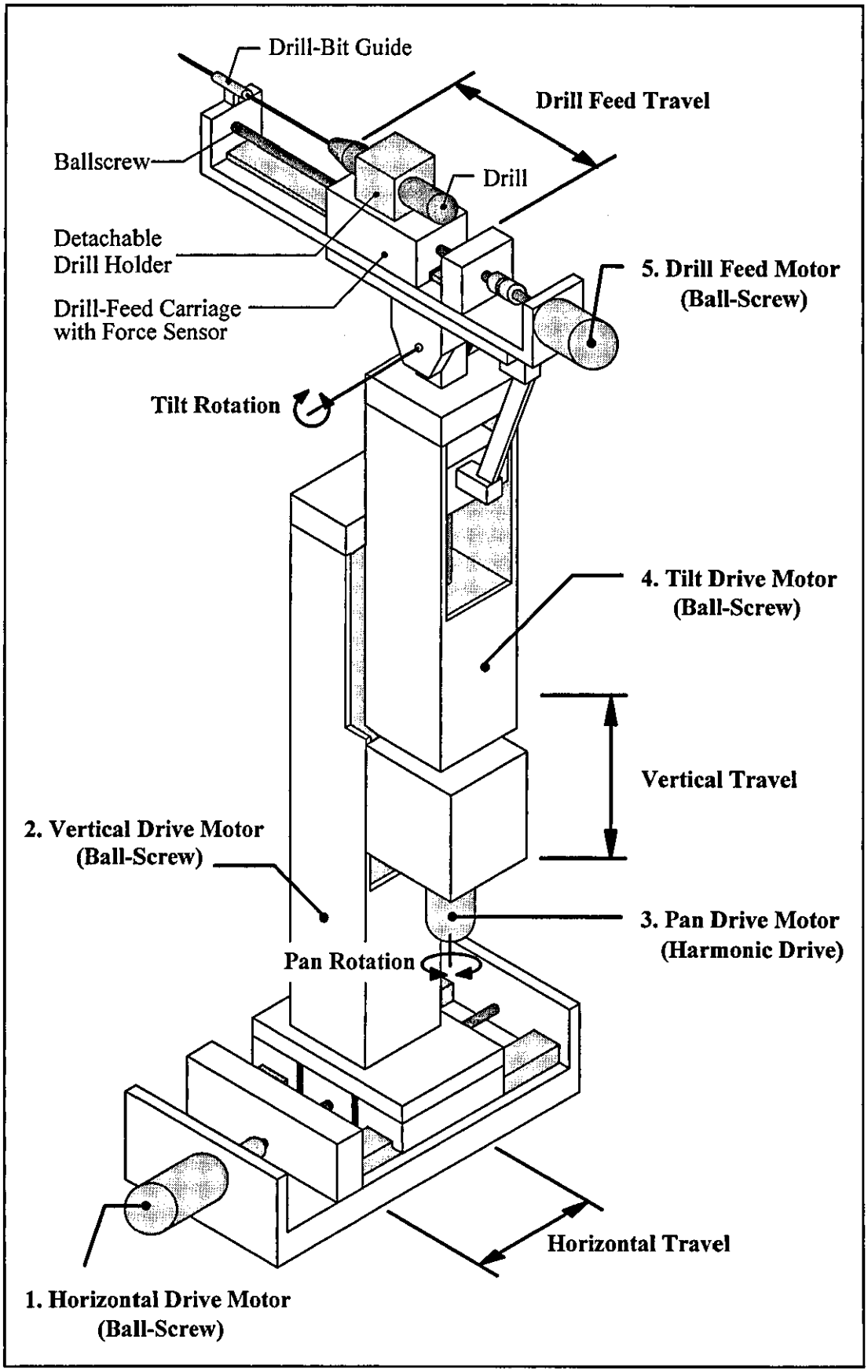


Fig. A2.2: Schematic Depicting the Prototype Manipulator's Degrees-of-Freedom

One aspect of the manipulator's "electro-mechanical" design which is not depicted in Figure A2.2, is the base unit. Ideally, the introduction of new technologies into the operating theatre should not prolong the duration of surgical procedures, and as such, the manipulator must have a quick and easy set-up mechanism. In order to satisfy this requirement of the original specification, a portable floor standing configuration was chosen, and the entire manipulator was mounted on a four-wheeled base unit. Positioning the manipulator next to the operating table and the patient, therefore involves the simple task of manually wheeling the combined unit into location. Once in position, the manipulator obviously has to be immobilised in order to maintain registration. A pneumatic/mechanical immobilisation mechanism, which is capable of raising and lowering the wheels, has therefore also been incorporated into the base unit.

A2.4.2 Drilling Unit.

The manipulator's fifth degree-of-freedom is incorporated into the drilling unit (end-effector), and is designated as the drill-feed travel (5) in Figure A2.2. The drill-feed actuator consists of a drill-feed carriage driven through a ballscrew by a stepper motor. In order to comply with surgical sterility requirements, the air-powered surgical drill is held in a detachable drill-holder unit, which in turn clamps onto the drill-feed carriage (this aspect of the drilling unit's design is further discussed in Section A2.4.4). In-line with current surgical practice, a drill bit guide is used to prevent the drill bit from slipping on the bone surface. A linear potentiometer has also been incorporated on the drill-feed actuator, in order to measure the drilling depth.

With a view to gaining initial clinical acceptance of the system, a manual drill bit guide (not shown in Figure A2.2) has also been integrated into the drilling unit. This allows the manipulator to be used as a non-invasive tool positioning device, aligning the manual drill bit guide with a calculated drilling trajectory. Manual completion of the drilling task can then be undertaken by the surgeon, using the guide to achieve optimum drilling results. However, the long-term aim of the drilling unit is to improve upon the results which may be obtained by manual drilling. This improvement is to be provided through quantification of the drilling process using

force feedback. A strain gauge force sensor, which monitors the axial drilling force, has therefore been integrated into the design of the drill-feed carriage.

The application of force feedback to automated surgical drilling is a promising area of computer/robotic-assisted surgery. Several research sites are currently investigating the use of cutting force information to anticipate and control drill bit breakthrough. Brett *et al.* **[6 & 7]** at the University of Bristol (AMARC research group), initially applied this technique to the micro-drilling stage of stapedotomy procedures (surgery of the middle ear). The same research group is also working in collaboration with Dario *et al.* **[8 & 9]** (Scuola Superiore Sant'Anna, Pisa, Italy) to develop a mechatronic drilling tool for orthopaedic procedures. The use of force-based penetration detection to monitor tool progress through adjacent tissue layers, has also been investigated as a means of training physicians to administer injections and perform biopsies. A telemanipulator demonstrator for the penetration of soft tissues has been developed by Davies *et al.* **[10]** at Imperial College, London (Mechatronics in Medicine Laboratory). Brett *et al.* **[11 & 12]** also report an experimental tactile tool force simulator, which demonstrates epidural puncture procedures.

In the case of the MEDROSA system, force feedback is utilised as part of the safety protocol of the automated drilling scenario. This "safety supervisor" algorithm works by continuously comparing the drilling forces, as measured by the strain gauge force sensor, with a theoretical drilling force profile. Experiments conducted on fresh porcine femurs **[13]** have demonstrated that variations in the composition of bone through a typical cross-section of the femur, produce drilling force profiles which include several easily identifiable features. The most noticeable of these characteristics, is the fact that the drilling force undergoes high rates of change at the boundaries between regions of cortical (average density: 1.3 g/cm^3) and cancellous (average density: 1.8 g/cm^3) bone.

Using empirically obtained results, it is therefore possible to accurately estimate the relationship between drilling force and drill bit depth of penetration into the femur. If the force measured during automated drilling deviates from this "expected"

relationship by more than a pre-determined threshold value, the safety supervisor algorithm will immediately stop the drilling process. In this way, it can be ensured that the tip of the drill bit remains within the femur, thereby avoiding unintentional drill bit breakthroughs and the associated damage to surrounding tissues. This problem is particularly relevant to the internal fixation of hip fractures, as it is not unknown for the drill bit to exit the femoral head and enter the body cavity. Drilling will also be terminated in the event of the drill bit coming into contact with a metal osteosynthesis implant. This scenario can occur if the drill bit misses the distal locking holes in an intramedullary nailing, and must be immediately corrected in order to minimise damage to the walls of the nail.

A2.4.3 Safety Features.

The overriding design constraint when introducing a robotic system into the operating room, is the safety of the patient, surgeon and theatre staff. The high level of human/manipulator interaction during robotic-assisted surgery is in complete contrast to industrial robot applications, where safety measures ensure that robots work in isolation from their human operators. Extensive modification of commercially available robots would therefore be required in order to meet the safety requirements of robotic-assisted surgery. However, the adoption of the mechatronic solution of a purpose-built manipulator, has allowed the MEDROSA Project to address these safety requirements from the outset of the design process.

The MEDROSA system's non-invasive application option, uses an active manipulator to align a manual drill bit guide with a pre-computed drilling trajectory. During the alignment process, software imposed motion limits confine the manipulator to a "safe" working envelope, thus avoiding unnecessary contact with the patient. Limit switches have also been fitted to all of the manipulator's joints as a backup to these software limits. The speed of all manipulator motions has also been deliberately restricted, giving the surgical staff adequate reaction time to press one of the system's emergency stop buttons. Once the desired alignment has been achieved, it must be maintained while the surgeon uses the drill bit guide to manually perform the drilling task. Fail-safe locking mechanisms (brakes) have therefore been fitted to all of the manipulator's

tool positioning drives. In the highly unlikely event of this breaking system being overridden, the use of non-backdrivable ballscrews ensures that the manipulator will still maintain its pose.

Using the manipulator's end-effector (drilling unit) to actively feed the drill bit into the patient's femur, obviously increases the risks involved with robotic-assisted procedures. However, as discussed in Section A2.4.2, drilling force feedback has the potential to remove human error from the drilling process, and as such, the invasive MEDROSA application option could theoretically be made to be safer than the non-invasive option. Enhanced performance can therefore be obtained by the provision of additional safety features, which ensure that the surgeon remains in complete control of the robotic system during all stages of the procedure. This has been achieved by providing the surgeon with a hand-held pendant, incorporating a dead man's handle type device, which is used in conjunction with a graphical human-computer interface.

Once the manipulator is aligned with the desired drilling trajectory, with all of the positioning motions locked, the surgeon is prompted by the system to initiate automated drilling. If the surgeon decides to proceed, the computer-controlled air-supply to the surgical drill is switched on and the drilling unit begins to feed the drill bit into the femur. A supervisory control strategy is then adopted, with the surgeon monitoring the progress of the procedure on the human-computer interface and using the dead man's handle to intervene should the need arise. The drilling forces are also monitored by the safety supervisor algorithm, which can also safely stop the automated drilling procedure. In the event of a system failure, the drill chuck can be loosened, allowing the manipulator to be wheeled away prior to manual completion of the surgical procedure.

A2.4.4 Sterility Issues.

Osteosynthesis procedures involve the insertion of foreign bodies (metallic internal fixation devices) into the human musculoskeletal system. Foreign bodies interfere with the normal anti-bacterial defence mechanisms, reducing the number of organisms which are necessary to cause infection. As a consequence, the risk of infection

associated with osteosynthesis procedures can be as high as 40% if extensive soft tissue damage is present **[14]**. Deep infection can lead to very serious complications, including failure of the fracture fixation, and frequently necessitates further surgery to remove the implant and necrotic tissue. Orthopaedic practice therefore aims to avoid infection at all costs. Prophylactic antibiotics are routinely administered in order to augment the normal host defence mechanisms **[15]**. Several measures are also taken to reduce the number of micro-organisms which can access the wound at the time of surgery.

The multiple sources of contamination in the operating room are summarised in Figure A2.3. The majority of surgical wound infections are *endogenous* in origin; that is to say they are caused by organisms originating from the patient **[16]**. Preoperative preparation of the patient (i.e. disinfecting the surgical site, skin drapes etc.) is therefore an essential measure to prevent infection. Cross-contamination (*exogenous* infection) caused by organisms originating from members of the surgical staff, may also occur as a result of direct contact or airborne dispersal. The airborne route to surgical infection was substantiated by the multicentre study reported by Lidwell *et al.* **[17]**, and has resulted in a renewed interest in ultra-clean air systems aimed at providing "sterilised" air in the operating room. Thus in addition to the conventional operating room ventilation system (i.e. filtered air at positive pressure), the use of ultraviolet lighting **[18]**, personnel isolation systems (whole-body exhaust-ventilated suits) **[19]**, and directional (laminar) flow control systems **[20]**, have all been investigated as a means of reducing the number of airborne organisms in the operating room. However, it is the direct contact route to surgical infection which is of most relevance to the MEDROSA Project.

Any item which comes into close contact with the surgical wound must be sterilised. During conventional surgery, this requirement applies to the surgeon's hands (sterile surgical gloves), the osteosynthesis implant(s), and the surgical tools used to perform the procedure. In the case of robotic-assisted surgery, the same requirement also applies to the manipulator, and in particular to the end-effector. The prevention of contamination in the operating room is therefore an overriding design criterion for the

MEDROSA system.

	ENDOGENOUS (Self Infection)	EXOGENOUS (Cross-Infection)
SOURCE	Patient.	Surgical Team & OR Staff. Equipment/Materials.
SITE OF SOURCE	Skin. Nose (Respiratory Tract). Gastro-Intestinal Tract.	Skin. Nose (Respiratory Tract). Equipment/Materials.
MODE OF SPREAD	Contact from Colonised or Infected Site	Air. Contact.

Fig A2.3 : Modes of Spread of Infection in the Operating Room
(Adapted from Ayliffe [16])

Sterilisation of the entire manipulator is not practicable due to size limitations imposed by the chamber dimensions of clinical sterilisation units. A modular manipulator design has therefore been adopted in order to achieve compatibility with existing sterilisation techniques. Sterilised draping is used to isolate the manipulator from the "clean" areas of the operating theatre. The drill holder unit, which has been autoclaved, then clamps onto the manipulator outside the isolation drape, thus maintaining surgical sterility. This draping technique is similar to the methods currently employed to isolate the image intensifier and X-ray tube housings of mobile intraoperative C-arm units.

The need to autoclave the drill holder unit imposes a number of design constraints. Effective steam sterilisation requires items to be exposed to saturated steam in the range 121-134 °C for up to fifteen minutes. These conditions preclude the use of many common engineering materials, and in particular most polymers. Special high temperature seals made out of virgin polytetrafluoroethylene (PTFE) have therefore been used. The drill bit guide has also been manufactured from stainless steel to ensure adequate corrosion resistance and biocompatibility.

System components which are to be handled by the surgeon and theatre staff, must also be either sterilised or isolated. These items may include a PC mouse, a hand-held pendant or even a keyboard. Electrical equipment obviously cannot be subjected to steam sterilisation conditions. Low temperature sterilisation by gaseous ethylene oxide (EO) is one possible solution to this problem. However, this process is more expensive to operate than steam sterilisation, and requires a long aeration time to reduce toxic EO residues to safe levels. Isolation of these items using sterilised plastic covers may therefore prove to be a more cost effective alternative to EO sterilisation.

A2.5 Future Development Methodology.

The ultimate aim of the MEDROSA research project is the development of a commercially available system which meets the requirements for authorisation by the relevant medical body (i.e. the Medical Devices Agency (MDA)). Laboratory trials have demonstrated the feasibility of the prototype MEDROSA system in a controlled environment, and continue to suggest that the application of robotic technology will benefit all of the parties involved in orthopaedic trauma surgery. The next logical step towards clinical acceptance, is therefore a two/three year programme of pre-clinical trials. These *in vitro* laboratory trials will be conducted on phantom and cadaver bone specimens under simulated operating room conditions, with the aim of further investigating the following areas:-

- Integration of the Complete System:

  The practicalities of introducing a vision-guided robotic system into the operating theatre are to be re-evaluated. Minor modifications of the prototype MEDROSA system will undoubtedly be necessary in order to ensure compatibility with existing surgical practices and equipment. The problem of electromagnetic interference in the operating room, which could lead to contamination of the MEDROSA system's control signals, will also have to be addressed. Manipulator "sterility" must also be closely looked at in conjunction with medical personnel.

- Robotic-Assisted Surgical Procedure:

  The step-by-step surgical protocols for both the manual and automated drilling scenarios, must be verified and improved upon in collaboration with orthopaedic surgeons.

- Human Safety:

  The fail-safe operation of the complete system must be re-evaluated under "operating room" conditions. Modifications to the existing safety protocol will be necessary in order to address problems which only become evident during realistic testing of the system.

- Accuracy:

  An error assessment is required in order to establish that the cumulative vision and robot inaccuracies lie within acceptable limits for individual orthopaedic procedures. In particular, the insertion of distal locking screws during intramedullary nailing of femoral shaft fractures is to be targeted for in-depth analysis. Further work on intraoperative registration between the robot and patient will also be undertaken as part of this investigation.

- Ergonomics:

  The ease of use of the complete system will be evaluated through close co-operation with the "end user", namely the orthopaedic surgeon. In particular, aspects of the human/computer interfaces will need to be addressed in order to ensure that the surgeon is provided with sufficient information in the required form.

- Alternative Applications:

  The system has initially been developed to augment a surgeons performance during the internal fixation of femoral fractures. However, similarities with internal fixation procedures for other parts of the body (spine, tibia, humerus etc.) are sufficient to allow a more widespread investigation of the system. The development of generic manipulator modules and procedure specific software are to be addressed.

Upon completion of this intermediate stage of the project, it is hoped that Ethics Committee approval can be gained for an initial set of clinical trials involving a limited number of patients. Surgeon feedback from these procedures will then influence a final round of system modifications. Permission will then be sought to carry out a larger programme of multicentre clinical trials, with the aim of demonstrating the benefits of the system in comparison to existing surgical technique, and establishing an adequate safety record for the system. It is hoped that favourable assessment of these clinical trials, will then lead to MDA approval, followed by routine clinical utilisation of the MEDROSA system.

A2.6 Concluding Remarks.

The potential benefits obtainable by introducing new technologies into the operating theatre, are just beginning to be realised through the routine clinical use of computer-assisted surgery systems. These non-invasive surgical navigation systems do not incur the inherent safety problems of invasive robotic-assisted surgery, and as a consequence, several have been granted approval to be commercially marketed. The Viewing Wand™ localiser (ISG Technologies, Mississauga, Ontario, Canada), approved by the US Food and Drug Administration (FDA), is typical of this trend, and has been used clinically in several hospitals **[21-24]**.

As a result of the less severe product liability implications, the major manufacturers of medical equipment are also beginning to show an interest in computer-assisted surgery. Philips Medical Systems (Best, The Netherlands), in collaboration with Aachen University of Technology (Germany), are reported to be developing an optical navigator for neurosurgery applications **[25]**. A similar joint venture between Johnson & Johnson Professional, Inc. (Randolph, MA, USA) and Vanderbilt University (Nashville, TN, USA), has also led to the development of the ACUSTAR™ I Advanced Neurosurgical Navigation System, which is currently under FDA review **[26]**.

Significant progress has also been made in the area of non-invasive robotic-assisted surgery. A limited number of these systems, such as the FDA-approved AESOP

endoscope manipulation system (Computer Motion, Inc., Goleta, CA, USA) [27], are now commercially available. Robotic tool guide positioning systems, which aid manual completion of the invasive stages of surgical procedures, are also reported to be in routine clinical use. The most notable example of this type of application, is the COMPASS™ robotically controlled arc-quadrant positioner (Compass International, Inc., Rochester, MN, USA), which has been applied to over 3000 stereotactic neurosurgery procedures [28]. Grenoble University's Image-Guided Operating Robot (IGOR), a semi-active robotic system which positions a probe holder during neurosurgery procedures, is also reported to have assisted over 600 interventions [29].

However, despite this growing clinical acceptance of technologically enhanced surgery, even the most advanced of the research projects investigating invasive robotic-assisted surgery are only at the clinical trials stage of development. The main reason for these invasive applications lagging behind their non-invasive counterparts, is obviously the safety problems created by the increased levels of autonomy. The development process for any invasive system is prolonged by the additional need to develop hardware and software solutions which ensure human safety at all times. Difficulties in securing research funding have also contributed to the delay. Manufacturers of medical equipment and government research councils are both wary of being associated with projects which could potentially generate adverse publicity, and as such funding has not been readily available. Finally, in the absence of safety legislation specifically relating to the surgical use of manipulators, licensing bodies are also reluctant to endorse invasive robotic-assisted surgery systems. Developers of invasive systems are therefore required to demonstrate both the clinical effectiveness and safety record of their system over a long period of time.

Examination of the progress made by some of the more ambitious research projects, reveals that the many justifiable obstacles which hinder the routine clinical use of invasive robotic-assisted surgery systems, necessitate a development process which can span many years. The ROBODOC® Surgical Assistant System (Integrated Surgical Systems, Inc., Sacramento, CA, USA), which is targeted at the milling of the femoral cavity during cementless total hip replacement procedures, is a prime example

of this trend. Research on this project began in 1986 (refer to Chart A2.1), and the system was first used on a human patient on 7th November 1992. The subsequent series of FDA-authorised multicentre clinical trials were still reported to be in progress in 1995 **[35]**.

A similar research time scale is also reported for the PROBOT system developed at Imperial College, London. This custom-built, active robotic system performs the invasive stages of transurethral resection of the prostate (TURP) procedures. Research into this application started in 1988 **[36]**, and led to what is claimed to be world's first use of an active robot to remove (soft) tissue from a human patient (25th March 1991) **[37]**. A second generation system was subsequently developed in order to gain Ethics Committee and MDA approval for the clinical investigations which are now reported as having begun **[38]**. The previously mentioned research at the University of Bristol, into the automation of stapedotomy procedures using a mechatronic micro-drilling tool, is also reported to be well advanced **[39]**. However, as this investigation did not start until 1990/91, this project is believed to be in the later stages of pre-clinical trials.

Having looked at the progress of these related research groups in this way, it is possible to make a number of conclusions about the MEDROSA Project's own progress and future development. Since the instigation of the project in 1991, a custom-built manipulator has been designed and manufactured in-house. All of the necessary control electronics and software modules are also now in place, along with the PC-based machine vision sub-system which falls within the scope the current thesis. A complete prototype system has therefore been developed and tested under laboratory conditions. Pre-clinical trials, which will make full use of the project's existing C-arm units and newly built lead-lined X-ray room, are set to follow upon obtaining additional research funding. As such, the progress of the MEDROSA Project is comparable to a number of the other research groups working in the area of robotic-assisted surgery.

RESEARCH PHASE	RESEARCH DETAILS
Limited Feasibility Study: Phase 1 (1986 - 1987)	Industrial robot used to mill simple shapes in synthetic bone [30].
***In Vitro* Laboratory Study:** Phase 2 (1987 - 1989)	ORTHODOC™ CT-based presurgical planning system developed [31]. Modified industrial robot used to mill human and canine cadaver bones [32].
***In Vivo* Canine Study:** Phase 3 (1989 - 1991)	IBM prototype surgical robot developed. Veterinary clinical trials: twenty-six robot-assisted operations performed on dogs (May 1990 - Sept. 1991) [33].
Clinical Applications (USA): Phase 4 (1991 -)	Second generation system developed to meet the requirements of human clinical trials [34]: **a) Human Feasibility Study:** FDA-authorised surgical trial involving ten human patients (7th Nov. 1992 - 20th Feb 1993). **b) Multi-Centre Study:** (Start Date: Sept. 1993) FDA-authorised surgical trials involving 300 human patients (150 patient study group & 150 patient control group). Interim results reported: Jan. 1995 [35].

Chart A2.1 : Development Process for the ROBODOC® Surgical Assistant
Phases 1 - 3: University of California at Davis (UCD) & IBM Thomas J. Watson Research Centre, NY
Phase 4 : Integrated Surgical Systems, Inc., Sacramento, California [Founded 1990].

The decision to initially opt for a non-invasive application of the MEDROSA system is believed to be critical to the long-term success of the project. The Compass™ and IGOR neurosurgery systems have already established a precedence for this tool guide positioning approach to robotic-assisted surgery, and have done so in a branch of surgery that requires much higher levels of precision than orthopaedics. Careful design of the system's manual drill bit guide, has ensured that the non-invasive MEDROSA scenario can be performed in such a way that the patient always remains outside the robot's working envelope during alignment with the desired drilling trajectory. The fail-safe braking system can then be used to ensure that the manipulator is immobilised during the manually performed drilling procedure. It can

therefore be argued that the manipulator is effectively not involved in the "dangerous" invasive stage of the procedure, which remains fully within the surgeon's control. As such, clinical acceptance of the non-invasive MEDROSA system should be dependant upon a satisfactory demonstration of surgical precision, reliability and cost effectiveness.

Assuming that the non-invasive MEDROSA system eventually gains MDA and Ethics Committee approval for clinical trials, and ultimately routine clinical use, the emphasis of the project will switch to the invasive application. It is still believed that the enhanced drilling performance achievable through the use of force feedback, warrants the additional research effort and expenditure required to make automated drilling "safe". However, in the absence of a precedence, it is unclear how long it will be before one of the pioneering invasive systems is finally granted approval by either the FDA or MDA. Uncertainties over the safety and legal implications **[40]** of such a system must be resolved before a commercial license can be granted. Standardisation of the development and evaluation processes, as noted by Mohsen *et al.* **[41]**, also seems to be a long way off. The level of commitment which has resulted in the development of prototype invasive robotic-assisted surgery systems over the past decade, must therefore be maintained over the next decade in order to ensure that the ultimate goal of routine clinical use is eventually realised.

A2.7 References.

1. **Blauvelt, C.T.** and **Nelson, F.R.T.** *A Manual of Orthopaedic Terminology,* 5th Edition. Mosby Year Book, St. Louis, Missouri, 1994. (ISBN 0-8016-6943-X)

2. **Hewit, J.R.** and **Bouazza-Marouf, K.** Robotic Assistance in Orthopaedic Surgery. *Internal Publication: Department of Mechanical Engineering, Loughborough University of Technology*, 1991.

3. **Melton, L.J.** Hip Fractures: A Worldwide Problem Today and Tomorrow. *Bone*, 1993, 14(Suppl.1), S1-S8.

4. **Hollingworth W., Todd C.** and **Parker M.** The Cost of Treating Hip Fractures in the 21st Century. *The Journal of Bone and Joint Surgery [Br]: Orthopaedic Proceedings Supplement,* 1995, 77(1), 9.

5. **Bouazza-Marouf, K.** and **Hewit, J.R.** Robotic Assisted Repair of Fractured Femur Neck. *Institution of Mechanical Engineers Seminar S133 - Robotics in Keyhole Surgery*, London, 5th May 1993.

6. **Baker, D., Brett, P.N., Griffiths, M.V.** and **Reyes, L.** A Mechatronic Drilling Tool for Ear Surgery: A Case Study of Some Design Characteristics. *Mechatronics*, June 1996, 6(4), 461-477.

7. **Brett, P.N., Baker, D.A., Reyes, L.** and **Blanshard, J.** An Automatic Technique for Micro-Drilling a Stapedotomy in the Flexible Stapes Footplate. *Proceedings of the Institution of Mechanical Engineers, Part H - Journal of Engineering in Medicine*, 1995, **209**(H4), 255-262.

8. **Allotta, B., Belmonte, F., Bosio, L.** and **Dario, P.** A Mechatronic Tool for Drilling in Orthopaedics. *Proceedings of the Second International Workshop: Mechatronics in Medicine & Surgery (MediMec'95)*, Bristol, UK, 6-9 September 1995, 92-107.

9. **Allotta, B., Belmonte, F., Bosio, L.** and **Dario, P.** Study on a Mechatronic Tool for Drilling in the Osteosynthesis of Long Bones: Tool/Bone Interaction, Modelling and Experiments. *Mechatronics*, June 1996, 6(4), 447-459.

10. **Yen, P.-L., Hibberd, R.D. and Davies, B.L.** A Telemanipulator System as an Assistant and Training Tool for Penetrating Soft Tissue. *Mechatronics*, 1996, 6(4), 423-436.

11. **Brett, P.N., Harrison, A.J., Thomas, T.A.** and **Stone, R.S.W.** A Computer Aided Tactile Simulator for Invasive Tools. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, MD, USA, 4-7 Nov. 1995, 324-328.

12. **Brett, P.N.,** *et al.* Automatic Surgical Tools for Penetrating Flexible Tissues. *IEEE Engineering in Medicine and Biology Magazine*, 1995, **14**(3), 264-270.

13. **Bouazza-Marouf, K., Browbank, I.** and **Hewit, J.R.** Robot-Assisted Invasive Orthopaedic Surgery. *Mechatronics*, June 1996, 6(4), 381-397.

14. **Matthews, I.P., Gibson, C.** and **Samuel, A.H.** Sterilisation of Implantable Devices. *Clinical Materials*, 1994, **15**(3), 191-215.

15. **Gillespie, W.** Prophylaxis Against Infection in Orthopaedic Practice. *Current Orthopaedics*, 1994, 8(4), 220-225.

16. **Ayliffe, G.A.J.** Role of the Environment of the Operating Suite in Surgical Wound Infection. *Reviews of Infectious Diseases*, 1991, **13** Supplement 10, 800-804.

17. **Lidwell, O.M.,** *et al.* Effect of Ultraclean Air in Operating Rooms on Deep Sepsis in the Joints After Total Hip of Knee Replacement: A Randomised Study. *British Medical Journal*, 1982, **285**(6334), 10-14.

18. **Berg, M., Bergman, B.R.** and **Horborn, J.** Ultraviolet Radiation Compared to an Ultra-Clean Air Enclosure. *Journal of Bone and Joint Surgery [Br]*, 1991, **73**(5), 811-815.

19. **Ross, I.** and **Clark, R.P.** The Ergonomics of Ventilated Operating Theatre Clothing. *Ergonomics*, 1988, **31**(7), 1103-1113.

20. **Lidwell, O.M.** Clean Air at Operation and Subsequent Sepsis in the Joint. *Clinical Orthopaedics and Related Research*, 1986, **211**, 91-102.

21. **Glossop, N.D., Hu, R.W.** and **Randle, J.A.** Computer-Aided Pedicle Screw Placement Using Frameless Stereotaxis. *Spine*, 1996, **21**(17), 2026-2034.

22. **Drake, J.M.,** *et al.* Frameless Stereotaxy in Children. *Pediatric Neurosurgery*, 1994, **20**(2), 152-159.

23. **Olivier, A., Germano, I.M., Cukiert, A.** and **Peters, T.** Frameless Stereotaxy for Surgery of the Epilepsies: Preliminary Experience. *Journal of Neurosurgery*, 1994, **81**(4), 629-633.

24. **Hassfeld, S.,** *et al.* Computer-Aided Skull Base Surgery. *Computer Assisted Radiology: Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR'95).* Berlin, Germany, 21-24 June 1995, 1168-1173.

25. **Adams, L.,** *et al.* Optical Navigator for Brain Surgery. *Computer*, January 1996, **29**(1), 48-54.

26. **Maurer, C.R.,** *et al.* The Accuracy of Image-Guided Neurosurgery Using Implantable Fiducial Markers. *Computer Assisted Radiology - Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR'95)*, Berlin, Germany, 21-24 June 1995, 1197-1202.

27. **Sackier, J.M.** and **Wang, Y.** Robotically Assisted Laparoscopic Surgery: From Concept to Development. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds.), MIT Press, Cambridge, MA, 1996, 577-580.

28. **Kall, B.A.** Computer-Assisted Surgical Planning and Robotics in Stereotactic Neurosurgery. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds). MIT Press, Cambridge, MA, 1996, 353-361.

29. **Lavallée, S.,** *et al.* Image-Guided Operating Robot: A Clinical Application in Stereotactic Neurosurgery. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds). MIT Press, Cambridge, MA, 1996, 343-351.

30. **Taylor, R.H.,** *et al.* Robotic Total Hip Replacement Surgery in Dogs. *Proceedings of the 11th Annual International Conference of the IEEE Engineering in Medicine & Biology Society*, Seattle, WA, USA, 9-12 November 1989, **11**, 887-889.

31. **Hanson, W.A.,** *et al.* ORTHODOC - An Image Driven Orthopaedic Surgical Planning System. *Proceedings of the 12th Annual International Conference of the IEEE Engineering in Medicine & Biology Society*, Philadelphia, PA, USA, 1-4 November 1990, **12**, 1931-1932.

32. **Paul, H.A.,** *et al.* Development of a Surgical Robot for Cementless Total Hip Arthroplasty. *Clinical Orthopaedics and Related Research*, 1992, **285**, 57-66

33. **Taylor, R.H.,** *et al.* Image-Directed Robotic System for Precise Orthopaedic Surgery. *IEEE Transactions on Robotics and Automation.* 1994, **10**(3), 261-275.

34. **Mittelstadt, B.D.,** *et al.* The Evolution of a Surgical Robot from Prototype to Human Clinical Use. *Proceedings of the First International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, Pittsburgh, PA, USA, 22-24 September 1994, 36-41.

35. **Bargar, W.L.,** *et al.* Clinical Abstract - Robodoc Multi-Center Trial: An Interim Report. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, MD, USA, 4-7 November 1995, 208-214.

36. **Davies, B.L., Hibberd, R.D., Timoney, A.G.** and **Wickham, J.E.A.** A Clinically Applied Robot for Prostatectomies. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds.), MIT Press, Cambridge, MA, 1996, 593-601.

37. **Ng, W.S., Davies, B.L., Hibberd, R.D.** and **Timoney, A.G.** Robotic Surgery: A First-Hand Experience in Transurethral Resection of the Prostate. *IEEE Engineering in Medicine & Biology Magazine.* 1993, **12**(1), 120-125.

38. **Harris, S.J.,** *et al.* The Probot - An Active Robot for Prostate Resection. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine,* 1997, **211**(4), pp 317-325.

39. **Baker, D.A., Brett, P.N., Baldwin, D.L., Blanshard, J.** and **Reyes, L.** A Mechatronic Micro-Drilling Technique for Controlling Drill Penetration Through Flexible Blind Bone Tissue Interfaces Using The Stapedotomy Procedure as an Example. *Proceedings of the Second International Workshop: Mechatronics in Medicine & Surgery (MediMec'95)*, Bristol, UK, 6-9 September 1995, 27-34.

40. **Davies, B.L.** A Discussion of Safety Issues for Medical Robots. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds.), MIT Press, Cambridge, MA, 1996, 287-296.

41. **Mohsen, A.M.M.A.,** *et al.* A Phased Approach to Efficacy Assessment of Computer Robotic Surgical Systems. *Proceedings of the Second International Workshop: Mechatronics in Medicine and Surgery (MediMec'95)*, Bristol, UK, 6-9 September 1995, 58-69.

APPENDIX 3:

INTRAOPERATIVE REGISTRATION FOR INTRACRANIAL NEUROSURGERY APPLICATIONS

CONTENTS

Page

A3.1 INTRODUCTION A56

A3.2 CONVENTIONAL STEREOTACTIC NEUROSURGERY A58

A3.3 ROBOTIC-ASSISTED NEUROSURGERY A62

A3.4 FRAMELESS STEREOTAXY A70

A3.5 THE BRAIN MOTION PROBLEM A75

A3.6 REFERENCES A77

A3. INTRAOPERATIVE REGISTRATION FOR INTRACRANIAL NEUROSURGERY APPLICATIONS

A3.1 Introduction.

Neurosurgery is the surgical speciality that deals with all aspects of the central nervous system. However, as most robotic-assisted surgery and CAS applications are targeted at intracranial procedures (i.e. brain surgery), the following discussion of registration strategies is restricted to this specific form of neurosurgery. A typical intracranial procedure requires a surgical tool to be accurately inserted into the brain via a small opening in the skull. This surgical tool may be a needle/cannula, which is used to remove a tissue sample (biopsy) for diagnostic analysis, or to drain excess fluid from a cyst, abscess, or haemorrhage. Alternatively, the surgical treatment of malignant tumours can involve resection of the tumour using forceps, or even the insertion of radioisotope seeds to facilitate internal radiotherapy (brachytherapy). Conditions such as chronic pain, epilepsy, and Parkinson's disease, can also require part of the brain to be either destroyed, or excited using implanted neurostimulating electrodes. Nevertheless, irrespective of the type of surgical tools which are to be introduced into the brain, intracranial procedures generally require a preoperative imaging stage.

The need for preoperative imaging, stems from the fact that procedures performed through a small opening in the skull, do not have direct visualisation of the surgical site. In order to minimise damage to healthy brain tissue, which can obviously have very serious implications, it is therefore essential that a neurosurgeon can preoperatively define his surgical strategy. Given that neurosurgery by its very nature deals with abnormal brains, and that substantial anatomical variations can exist from person to person, surgical planning cannot simply rely upon brain atlases which depict "typical" brain structure. Knowledge of the exact form of an individual patient's anatomy is therefore required.

This information is generally obtained by using a tomographic imaging technique (CT or MRI) to develop a three-dimensional model of the patient's brain. Since certain types of brain tissue die very rapidly when deprived of oxygen, a DSA examination may also be undertaken in order to ensure that the preoperatively planned surgical trajectories avoid important blood vessels. Depending upon the nature of the procedure, functional imaging techniques (PET, SPECT, etc.) may also be employed to highlight areas of the brain exhibiting abnormal behaviour or function.

Failure to accurately reproduce the preoperative plan during the actual neurosurgery procedure can have very serious consequences. In addition to damaging healthy brain tissue, tool-positioning errors can also lead to incomplete removal of a cancerous tumour. Early recurrence of the tumour, possibly in a malignant form, can therefore take place as a result of failing to cure the patient of a potentially benign condition. As a consequence, many neurosurgical procedures require levels of precision that cannot be easily achieved using free-hand surgical techniques. Current surgical practice therefore applies the well established *stereotaxy* method, which overcomes this problem through the provision of an intraoperative registration method and a surgical guidance system.

However, as explained in Section A3.2, conventional stereotactic neurosurgery techniques leave considerable scope for improvement. A number of robotic-assisted neurosurgery studies (discussed in Section A3.3) have therefore been conducted with the aim of improving surgical performance. The computer-assisted surgery (CAS) approach described in Section A3.4, which is generally referred to as *frameless stereotaxy*, is also beginning to receive considerable attention. Unfortunately, in spite of the significant progress that has been made in these areas, image-guided neurosurgery still exhibits a fundamental weakness.

Preoperatively planned neurosurgical procedures make the assumption that the actual patient anatomy observed in the operating room is identical to that appearing in the preoperative images. However, in reality this may not always be the case. Since the brain is essentially a collection of soft tissues, removing a tumour or draining excess

fluid can result in significant brain motion under the influence of gravity. Subsequent targets will therefore no longer be located at their preoperatively imaged positions, thus invalidating the registration process. Accordingly, Section A3.5 concludes this appendix by providing a brief discussion of potential solutions to this significant "brain motion" problem.

A3.2 Conventional Stereotactic Neurosurgery.

Stereotaxy is a precise surgical technique which works by initially establishing the location of a deep-seated brain structure, with respect to a three-dimensional coordinate system. A mechanical jig arrangement attached to the patient's head, which is referred to as a *stereotactic frame*, is then used to align a tool guide with a trajectory that passes through the calculated coordinates. The invasive stages of the surgical procedure are then performed by manually inserting surgical tools into the patient's brain with the aid of the rigidly locked tool guide.

Modern stereotactic neurosurgery is typified by procedures involving the use of the Brown-Roberts-Wells (BRW) guidance system. The BRW system was specifically designed to be fully compatible with a CT-scanner, and achieves intraoperative registration by rigidly attaching a *head or base ring* to the patient's skull using a series of bone screws. This base ring remains in position during preoperative imaging and the actual surgical procedure, and as such, establishes a common reference plane. By preoperatively determining the 3-D location of an imaged target point (e.g. a tumour) with respect to this common reference plane, it is therefore possible to establish the corresponding point in the operating room by attaching a stereotactic frame to the base ring.

Preoperative location of a target point with respect to the base ring coordinate system, is achieved by attaching a calibration object to the base ring. As depicted in Figure A3.1(A), this calibration object surrounds the patient's head with three sets of N-shaped fiducials, each of which consists of two vertical rods connected by a diagonal rod. Scanning this calibration object along with the patient produces CT

images of the patient's head, which also contain cross-sectional views of the fiducial rods (refer to Figure A3.1(B)).

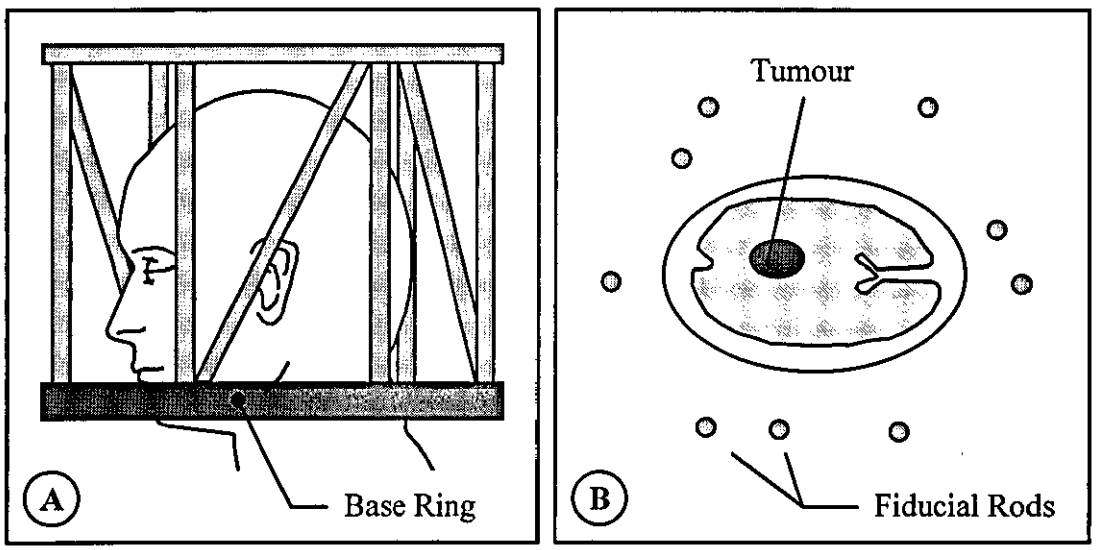


Fig. A3.1 : The Stereotactic Localisation Technique.

Depending upon the actual location of the CT scan plane, the cross-sectional image of each N-shaped fiducial varies in the manner shown in Figure A3.2. By measuring the relative distances between corresponding diagonal/vertical fiducial rod cross-sections, a given CT slice's location (i.e. a 3-D plane) can therefore be determined with respect to the base ring. The two in-plane coordinates of a target point appearing in the CT image are then established, by measuring its location with respect to those of the nine fiducial rods.

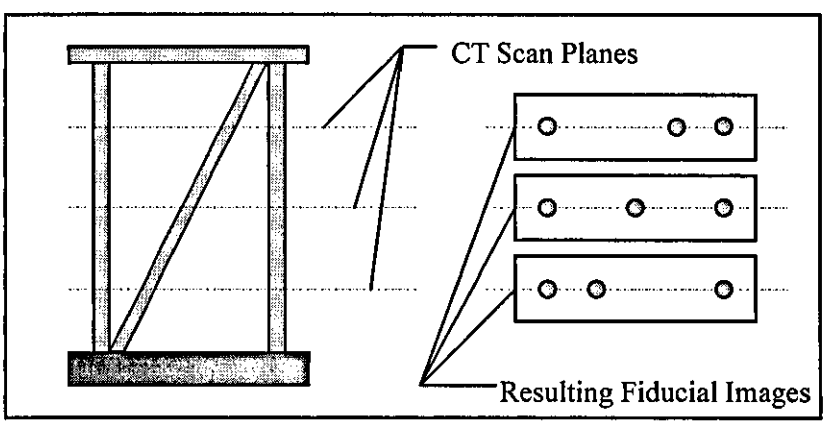


Fig. A3.2 : Location of CT Scan Slice Using N-Shaped Fiducials.

Having established the location of a target(s) with respect to the base ring coordinate system, preoperative planning is completed by defining a craniotomy (entry point)

site(s) which allows safe access to the target(s). The calibration object is then removed from the base ring, and the patient is transferred to the operating room. In order to immobilise the patient, the base ring is rigidly attached to the operating table. The sterile stereotactic arc guidance system depicted in Figure A3.3 is then attached to the base ring. The four positioning degrees-of-freedom of this device allow the probe holder to be aligned with an infinite number of trajectories. Calculation of the four frame settings which correspond to the trajectories defined by corresponding target/entry point-pairs, therefore allows the accurate reproduction of the surgical plan in the manner shown in Figure A3.4.

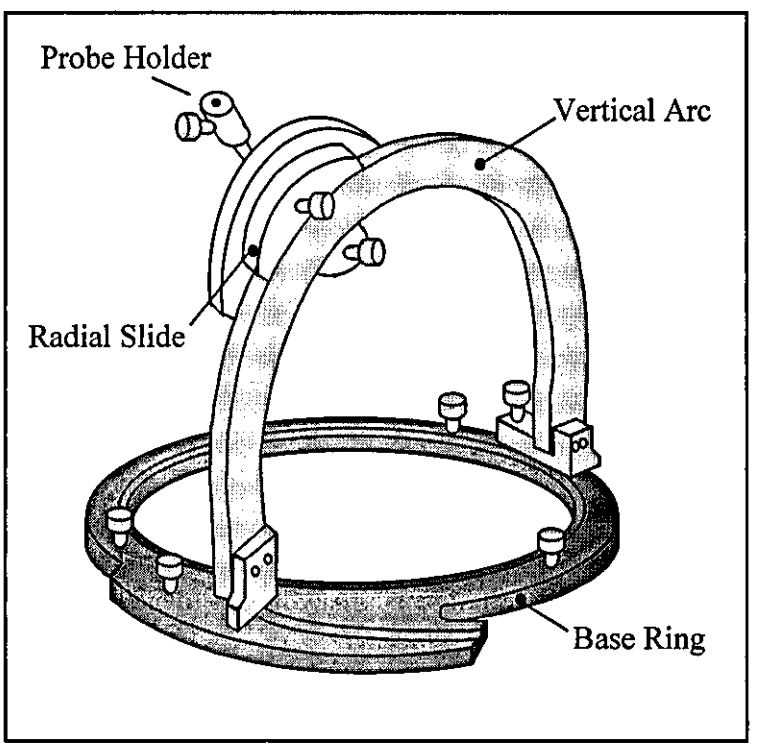


Fig. A3.3 : The BRW CT Stereotactic Arc Guidance System
(Adapted from Heilbrun [1])

Although stereotaxy is currently the most accurate means of performing neurosurgery, the technique also has a number of shortcomings. The patient must undergo a CT or MRI examination with the base ring and N-shaped fiducial arrangement in position. However, a CT/MRI examination will already have been undertaken in order to diagnose the patient's condition. Stereotactic neurosurgery therefore requires that the patient undergoes two MRI/CT examinations. Attaching the base ring to the patient's head also requires an additional surgical procedure, and can obviously involve a

certain degree of pain for the patient. Accordingly, this procedure is normally performed under a general anaesthetic. Since most base ring designs cannot be accurately replaced, the preoperative imaging process is normally performed immediately prior to surgery. As a result, only a limited amount of time is available for preoperative planning, and the patient is subjected to a prolonged period under anaesthesia.

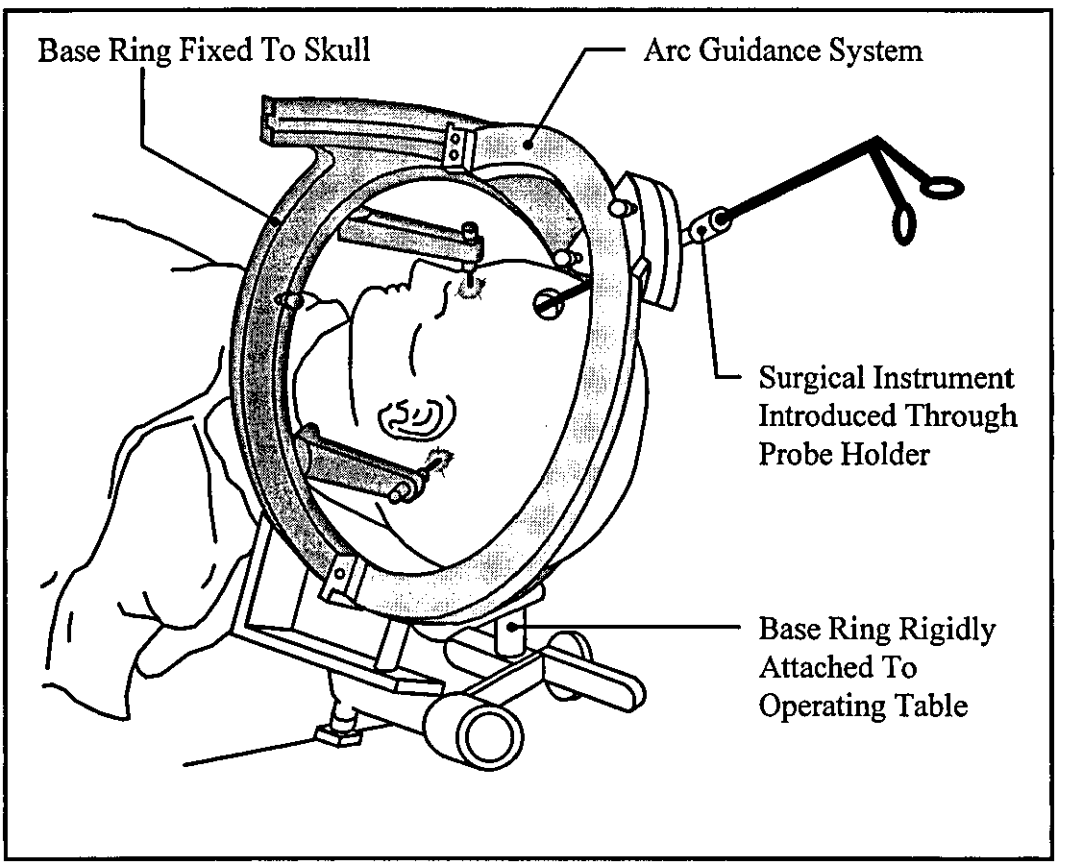


Fig. A3.4 : Conventional Stereotactic Neurosurgery.
(Adapted from Heilbrun [1])

The arc guidance system employed during the actual surgical procedure can also cause problems. Errors can occur during the manual adjustment of the four angular settings. As it is not uncommon for several frame adjustments to be required during a neurosurgical procedure, stereotaxy can also prove to be very time-consuming. During open surgery procedures, the arc may also interfere with the surgeon's access to the surgical field. The use of a robotic arm to replace the arc guidance system has therefore been investigated as a potential solution to these problems. Although the need to immobilise the patient with respect to the robot generally requires that some

form of base ring must still be used during surgery, robotic manipulators are ideally suited to perform the high accuracy, repetitive motions required to align a tool guide with a linear trajectory.

A3.3 Robotic-Assisted Neurosurgery.

In the context of the current discussion, the term "robotic-assisted" neurosurgery is used to imply a system that uses a computer-controlled *robotic arm* to perform tool positioning tasks during neurosurgical procedures. However, before proceeding to evaluate such systems, it is worth pointing out that several research sites have also investigated the use of motorised or "robotic" stereotactic frames as an alternative means of automating neurosurgical procedures. The most successful of these devices is the Compass™ Stereotactic System, which was developed by Kall *et al.* at the Mayo Clinic, Rochester, Minnesota **[2]**. During the period 1984-1993, it is reported that 3142 successful neurosurgical procedures were performed at the Mayo Clinic using this system **[3]**.

The first "robotic-assisted" neurosurgery procedure was performed in 1985 by Kwoh *et al.* **[4]** (Memorial Medical Centre, Long Beach, CA). During this procedure, a modified PUMA 200 robot (Unimation, Inc.) was used to accurately align a probe guide with a preoperatively planned trajectory. As depicted in Figure A3.5, an external clamp was then attached to the robot's end-effector in order to prevent unintended movements of the robotic arm. A biopsy needle was then manually inserted through the probe guide, and a tissue sample was removed from the patient's brain.

Given that the surgeon would be required to intervene in the unlikely event of a problem occurring with the robot, conventional stereotaxy methods were used during the preoperative planning stages of the procedure. Accordingly, immobilisation of the patient's head was accomplished using the base ring of a stereotactic frame. Should the need arise, an arc guidance system could therefore be attached to this base ring, thus allowing manual completion of the surgical procedure.

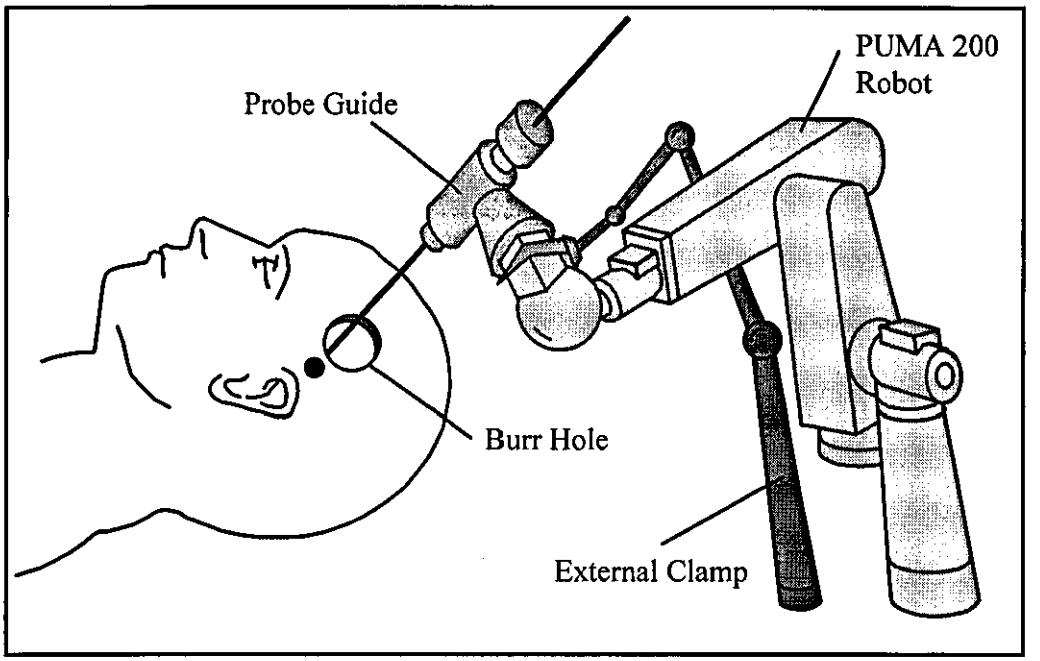


Fig. A3.5 : The Long Beach Robotic-Assisted Neurosurgery System.
Alignment of a Probe Guide with a Pre-Planned Trajectory (base ring not shown).
(Adapted from Kwoh *et al.* **[5]**)

The most interesting aspect of Kwoh et al's pioneering research, is the fact that the entire robot-assisted procedure was performed in a dedicated CT stereotactic operating suite **[6]**. The system therefore effectively allowed intraoperative CT scanning and immediate postoperative confirmation of surgical results. By performing neurosurgery in the vicinity of the CT-scanner, thus avoiding the need to transfer the patient back and forth between the CT scanning suite and the operating room, Kwoh *et al.* were therefore able to significantly reduce the overall duration of the procedure. Computer-controlled probe alignment using the robotic arm, also proved to be much faster than manually adjusting the stereotactic frame settings.

Unfortunately, during the course of the project the manufacturers of the PUMA 200 robot (Unimation, Inc.) were taken over by another company (Westinghouse) who, for legal reasons, no longer wished to be associated with potentially high-risk applications such as robotic-assisted surgery **[7]**. As the major manufacturers of CT-scanners were also reluctant to become involved in the commercial exploitation of the Long Beach system, Kwoh *et al.* were therefore eventually forced to abandon their research into robotic-assisted neurosurgery.

Drake *et al.* [8] (Hospital for Sick Children, Toronto, Ontario, Canada) have also used the PUMA 200 robot (Westinghouse) to successfully demonstrate the feasibility of robotic-assisted neurosurgery. However, as the actual surgical procedure was performed in a conventional operating room, up to one week after the preoperative imaging process, their approach was slightly different to that adopted by the Long Beach researchers. Following attachment of a BRW base ring to the patient's skull, the standard CT scanning procedure was performed using the N-shaped fiducial arrangement. An additional angiographic examination was also performed to allow the three-dimensional reconstruction of important blood vessels. The BRW base ring was then removed from the patient's head, and was reapplied in the operating room immediately prior to surgery. Intraoperative registration was then established by bolting the base ring to a rigid table, on which the robot had already been mounted. The robot was then used to align a surgical retractor (a metal cylinder of 20 mm diameter) with preoperatively planned trajectories. This surgical retractor was then manually introduced into the patient's brain.

The Toronto robotic-assisted neurosurgery system was used (circa 1990) to remove deep-seated tumours during six procedures involving child patients. Throughout these procedures the position of the surgical retractor was displayed in real-time, with respect to the preoperative images, on a three-dimensional display. However, although the robotic arm was shown to be more versatile than a fixed radius stereotactic arc guidance system, further system development was not undertaken. Owing to the modifications and expenditure that would be required to make the PUMA 200 robot fail-safe, Drake *et al.* decided that a frameless stereotaxy technique would be a more viable option. As such, they have subsequently adapted the Viewing Wand (ISG Technologies, Missasauga, Ontario) to meet the requirements of their patient workload. This device is a passive, articulated digitising arm, and was therefore considered to be "simpler to use with less inherent risk" [9].

The failure of the Long Beach and Toronto systems to achieve routine clinical use, was mainly attributable to safety problems arising from the use of a modified

industrial robot. The MINERVA Project which is currently being conducted by Glauser *et al.* **[10]** (Swiss Federal Institute of Technology of Lausanne) has therefore taken the logical step of developing a custom-built neurosurgical robot. The specification for this robot required that intraoperative CT scanning could be performed at any stage of the surgical procedure without having to move the robot **[11]**. As a consequence, the MINERVA robot was designed to actually work within the confines of a standard CT-scanner. Localisation of the target point therefore involves the scenario illustrated in Figure A3.6.

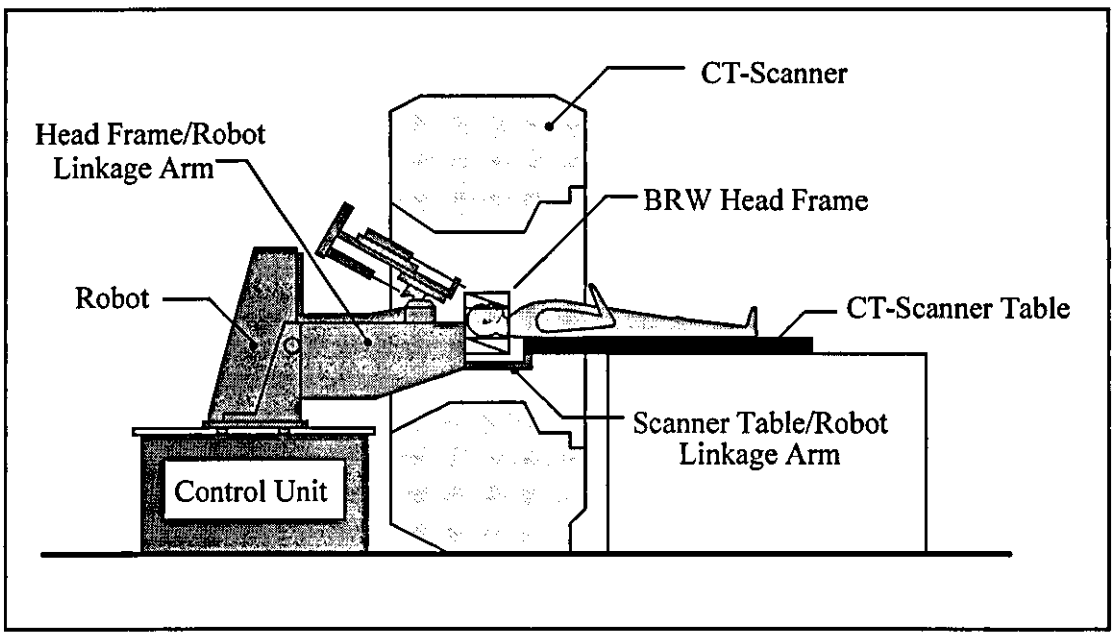


Fig. A3.6 : The MINERVA System
(Adapted from Burckhardt *et al.* **[12]**)

In line with conventional stereotactic practice, the patient lies on the motorised CT-scanner table with a BRW base ring securely fixed to his/her skull. The BRW base ring is also attached to the MINERVA robot using a carbon fibre linkage arm. As a result, the standard N-shaped fiducial arrangement can be used to locate the target with respect to not only the base ring, but also the MINERVA robot. However, in order to maintain a fixed transformation between the robot and the patient, the robot must replicate the movements of the CT-scanner table during image acquisition. The robot is therefore mounted on a passive translation guide and is directly coupled to the CT-scanner table.

The MINERVA system was used to perform an initial series of eight biopsy procedures in 1993. Postoperative CT scanning of the eight patients revealed that the accuracy of the robotic procedure (± 0.5 mm), was far superior to that of conventional stereotactic techniques (± 2 mm) **[10]**. By performing the entire procedure inside the CT-scanner, the duration of the procedure was also reduced from two hours to thirty minutes. Additional clinical trials are therefore currently being conducted in a specially designed surgical complex.

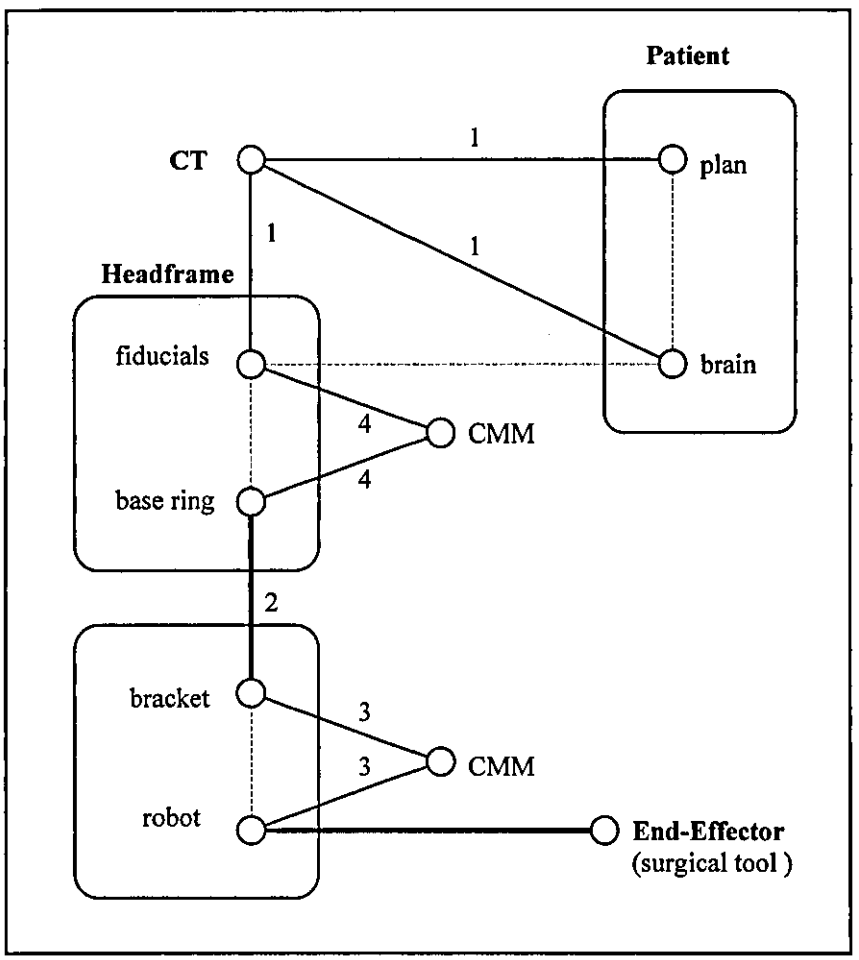


Fig. A3.7 : Registration Graph For Robotic-Assisted Neurosurgery Systems That Use Stereotactic Frames.
(Adapted from Lea *et al.* **[13]**)

In spite of the obvious differences between the Long Beach, Toronto, and MINERVA systems, all three employ essentially the same registration strategy. In order to emphasise this point, the diagrammatic technique proposed by Lea *et al.* **[13]** is used

here to provide an overview of the overall registration strategy of these systems. Readers who are unfamiliar with Lea et al's *registration graphs* should refer to the summary of the technique provided in Appendix 4. The registration graph for a robotic-assisted neurosurgery system that makes use of the stereotactic base ring, takes the form shown in Figure A3.7.

A preoperative CT examination of the patient's head, with the base ring and N-shaped fiducial arrangement firmly in place, establishes the location of the brain with respect to the fiducial rods and hence the base-ring. As this relationship is established indirectly using the CT-scanner as an intermediary sensor, the process of measuring the brain's location with respect to the fiducials is represented by the an induced link on the registration graph (i.e. the dashed line between the fiducials and brain nodes (small circles)). This induced link highlights the fact that subsequent brain motion relative to the fiducials will invalidate registration (i.e. the image data is only valid at the time of image acquisition). Immobilisation of the patient's head using the base-ring is therefore an essential part of the registration strategy.

Each of the three robotic-assisted neurosurgery systems achieves intraoperative registration by physically attaching the robot to the base ring. This "measurement" process is represented on the registration graph by a sustained link (i.e. the bold line running between the base ring and bracket nodes (small circles)). However, in order to perform the actual surgical procedure, the position of the base ring must be known with respect to the robot's coordinate system at all times. It is therefore necessary to determine the exact location of the robot's mounting bracket, to which the base ring is attached, with respect to the robot coordinate system.

As depicted in Figure A3.7, the easiest method of determining this relationship is to use a coordinate measurement machine (CMM). However, as the MINERVA researchers are the only group to actually stipulate the preoperative (off-line) use of a CMM to perform this task, it has had to be assumed that the other two groups also adopted this technique. Indirect measurement of the bracket's location, with respect to the robot coordinate system, is depicted by an induced link on the registration graph

(i.e. the dashed line between the bracket and robot nodes). Once again, this induced link highlights the fact that registration can be invalidated by relative motion between the robot and the bracket/base ring.

The final intraoperative measurement process required to complete registration, is determination of the surgical tool's position with respect to the robot coordinate system. However, since the surgical tool is held by the robot's end-effector, this measurement can be achieved in real-time by applying forward kinematics to the measured joint angles of the robot. As such, the end-effector's location is a continuously updated measurement, and is therefore represented by a sustained link on the registration graph (i.e. the bold line between the robot and end-effector nodes). Assuming that the fiducial locations have been measured with respect to the base ring (using a CMM for example), a continuous line exists on the registration graph through the following nodes [plan : brain : fiducials : base ring : bracket : robot : tool]. Accordingly, this subgraph depicts a chain of measurement processes that produce the required link between surgical planning and execution (i.e. [plan : tool]), thus validating the fact that registration has been established.

An alternative to this stereotaxy-based form of robotic-assisted neurosurgery, has successfully been developed by Lavallée *et al.* **[14]** (TIMC-IMAG, Faculté de Médecine de Grenoble, France). Their neurosurgery system is an example of the image-guided operating robot (IGOR) concept advocated by the Grenoble research group, which aims to improve surgical planning and execution via the use of multimodal image information and a non-invasive robotic manipulator. The system therefore employs both preoperative and intraoperative imaging techniques to define a surgical strategy, which is then implemented by using a modified industrial robot to position a tool guide. The surgeon can then introduce surgical tools into the patient's brain with the aid of this tool guide. It is reported that the IGOR system was first used in 1989 and has subsequently been employed during approximately 600 interventions **[14]**.

From published accounts of the IGOR neurosurgery system, it would appear that the preoperative imaging process (CT, MRI, or both) does not involve the use of the standard base ring/N-shaped fiducial technique **[15]**. It must therefore be assumed that the preoperative plan is defined with respect to the CT/MRI scanner's coordinate system. Consequently, an intraoperative registration process must be performed to transform this information into the operating room coordinate system. Given that intraoperative angiography is to be used to confirm that the preoperatively planned trajectories do not pass through any of the major blood vessels, intraoperative registration is achieved using the arrangement depicted in Figure A3.8.

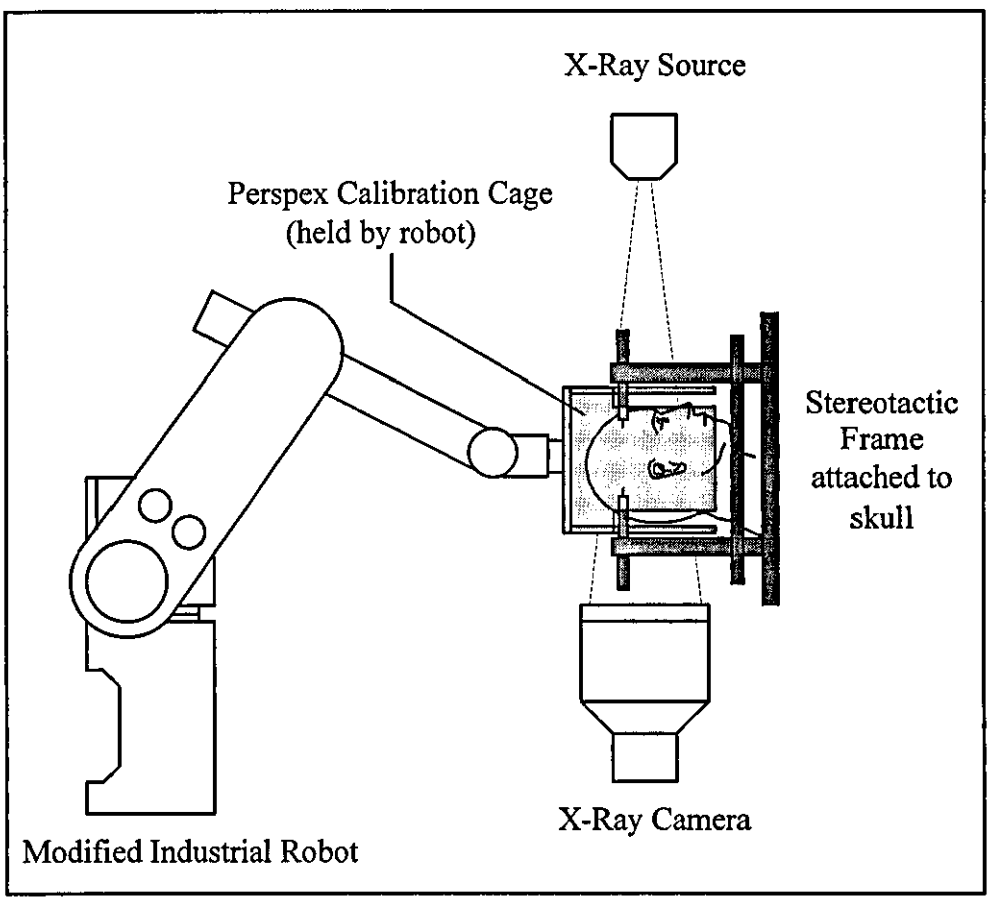


Fig. A3.8 : The Grenoble Neurosurgical Image-Guided Operating Robot (IGOR)
(Schematic of Intraoperative Registration)

The patient is immobilised using a stereotactic base ring, which in turn is attached to the operating table. The operating room coordinate system is then defined by surrounding the patient's head with a Perspex calibration cage; which is similar to those used by the X-ray photogrammetry applications discussed in Chapter 2. During

the registration process this cage is held in position by the robot, thus establishing the relationship between robot and operating room coordinates.

Intraoperative registration is then established on the basis of anatomical structures that can be imaged by both the preoperative CT/MRI systems, and the intraoperative X-ray system that is used to acquire the confirming angiograms. The structures selected for this *anatomy based registration* form part of the cerebral ventricular system. Intraoperative ventriculograms are therefore acquired by perpendicular X-ray systems, while the calibration cage is still in position. As a result, the metallic balls imbedded in the calibration cage are superimposed onto the intraoperative ventriculograms, thus allowing the position of the cerebral ventricles to be reconstructed with respect to the operating room coordinate system. By matching ventricular structures segmented from the 3-D preoperative images, with the corresponding structures in the 2-D intraoperative ventriculograms, it is therefore possible to transform the preoperative data into the operating room coordinates system.

Upon completion of the registration process, perpendicular angiograms are taken of the patient's head with the calibration cage in position. The vessels of interest are then reconstructed with respect to the operating room (calibration cage) coordinate system. Assuming that the surgical trajectories do not intersect these vessels, the calibration cage is then removed and the tool guide is attached to the robot's end-effector, prior to alignment with the surgical trajectory.

A3.4 Frameless Stereotaxy.

Given the safety problems associated with robotic-assisted neurosurgery, it is not surprising to find that most research sites have opted for the alternative computer-assisted surgery (CAS) approach. A number of the resulting "neuronavigation" systems employ an intraoperative position-sensing device in conjunction with a stereotactic base ring **[16, 17]**. Intraoperative registration therefore involves the relatively simple task of determining the location of the base ring with respect to the operating room coordinate system. Having performed this calibration task, the measured position of a hand-held surgical tool can then be displayed on the

preoperative images. However, as the stereotactic registration technique has already been covered in the previous section, further discussion of these "stereotactic" CAS systems is not warranted. As such, the aim of this section is to outline the registration strategies adopted by CAS systems that do not employ a stereotactic base ring during the preoperative imaging process. The term "*frameless stereotaxy*" is generally used to describe this type of system.

In frameless stereotaxy applications, the preoperative image data is defined with respect to the CT/MRI scanner coordinate system. Some form of intraoperative registration process must therefore be undertaken in order to transform the preoperative plan into the operating room coordinate system. The most straightforward solution to this problem is the *point-based registration* technique. The locations of three or more anatomical landmarks are initially measured using an intraoperative position-sensor. The corresponding points on a 3-D graphical model of the patient's head, which are derived from the preoperative image data, are then interactively indicated using a PC-mouse or similar pointing device. These corresponding point-pairs then allow a rigid transformation to be obtained that relates operating room and CT/MRI scanner coordinates. However, measurement of prominent anatomical features, such as the bridge of the nose, is a highly subjective task. Large errors can therefore arise when implementing a point-based registration scheme that relies purely upon anatomical landmarks.

Many of the early frameless stereotaxy systems overcame the inherent problems of anatomical landmarks, through a combination of skin markings and temporarily applied external fiducials. This technique requires that three or more points on the patient's face/scalp are marked with indelible ink. Imaging modality compatible fiducial markers are then attached to the patient's skin at these locations. When imaged along with the patient, these fiducial markers provide discrete points in the preoperative images. Once preoperative imaging is complete the fiducial markers are then removed. Intraoperative determination of the "corresponding" points is subsequently achieved using a position-sensing device to measure the locations of the ink markings on the patient's face/scalp. As each ink marking is measured, the

corresponding fiducial point in the preoperative images is simultaneously indicated, thus allowing a rigid transformation to be calculated.

The registration accuracy of a frameless stereotaxy system is obviously dependent upon both the type of position-sensing device used in the operating room, and the resolution of the preoperative imaging modality (i.e. the slice thickness). It is therefore difficult to make direct comparisons between individual systems that employ the point-based registration strategy. The published results from a number of the earlier research projects tended to suggest that the technique was not as accurate as conventional stereotactic surgery. The pioneering system developed by Roberts *et al.* **[18]** (Dartmouth-Hitchcock Medical Centre, Hanover, NH, USA) employed a sonic localiser, and was first used clinically in 1985. The reported registration error for this system was typically 5 mm. A similar value was also quoted for the passive digitiser arm system of Watanabe *et al.* **[19]** (Tokyo Metropolitan Police Hospital, Japan). Likewise, Kato *et al.* **[20]** (Osaka University Medical School, Japan) found registration errors of approximately 4 mm when applying an electromagnetic localisation system.

It is therefore generally acknowledged that motion of the skin markings with respect to the brain is a fundamental limitation of the technique. The most probable cause of this skin motion is the attachment of a head holding device immediately prior to surgery. However, more recent systems, particularly those employing infra-red optical digitisers, claim registration errors as small as 1 mm. The most notable example is the NeuroStation™ Computer-Aided Surgery Workstation developed by Bucholz *et al.* **[21]** (St Louis University School of Medicine). The ACUSTAR™ I Advanced Neurological Navigation System developed by Galloway *et al.* **[22]** (Vanderbilt University, Nashville, TN) also achieves this level of precision through the use of four bone implanted fiducials. Using a 1 mm thick CT slice and a sonic localiser, Barnett *et al.* **[23]** (Cleveland Clinic Foundation, Cleveland, Ohio) have also managed to register fiducial locations to within 1.5 mm ± 0.7 mm.

The precision that can be achieved using a point-based registration strategy is perfectly adequate for most open neurosurgery procedures. However, the reliance upon external fiducials means that the patient must undergo a second CT/MRI examination. An alternative technique known as the *surface-based registration* strategy is therefore becoming increasingly popular. This method works by using three or more anatomical landmarks to establish an initial point-based estimate of the registration transformation. An evenly distributed series of random surface points on the patient's face/scalp, are then digitised using an intraoperative position-sensing device. Refinement on the initial registration transformation is then performed by applying a least-squares fitting of these intraoperative surface points, to a preoperatively generated 3-D skin surface model of the patient's face/head. Given that the technique relies purely upon anatomical features, all of the necessary preoperative data can be obtained from the original diagnostic CT/MRI examination.

The application of surface-based registration to frameless stereotaxy is typified by the methodology adopted by Olivier *et al.* **[24]** (McGill University, Montreal, Canada). Using the Viewing Wand System (ISG Technologies, Missasauga, Ontario) to perform intraoperative registration, five or more facial features are digitised, thus allowing an initial estimate of the registration transformation to be obtained. Refinement of this initial transformation is then achieved by digitising forty random skin surface points, and then applying a least-squares fitting to intraoperative and preoperative point-pairs. It is reported that postoperative MRI scans indicate that the overall system error is typically of the order of 3 mm. As previously discussed, Drake *et al.* **[9]** (University of Toronto, Canada) also use the Viewing Wand System, and report a registration error of 2.3 mm when applying a surface-based registration scheme involving twenty random surface points. However, they also state that the error associated with the same surgical procedure when performed with a BRW stereotactic frame is only 1.8 mm.

An interesting variation upon the frameless stereotaxy principle has also been implemented by Pelizzari *et al.* **[25]** (University of Chicago, IL, USA). With the aim of eliminating patient immobilisation during neurosurgical procedures performed

under local anaesthesia, Pelizzari *et al.* have developed a real-time patient tracking system based upon the Flashpoint (Image Guided Technologies, Boulder, CO, USA) infra-red optical digitiser. Patient tracking is actually performed by attaching a Plexiglas plate to the patient's skull using a bone screw. The locations of a group of infra-red LEDs mounted on this plate are then measured in real-time using the Flashpoint system. However, the plate is also used to define a dynamic, patient-based reference system, which essentially acts as the operating room coordinate system. Registration of preoperative MRI data with this reference system requires the use of a hand-held stylus incorporating a group of IR-LEDs, which in turn can be tracked by the Flashpoint system. By using this stylus to digitise 150 random scalp points, and then applying a surface matching algorithm to the preoperative MRI data, it is reported that a registration accuracy of 4.8 mm ± 3.5 mm has been achieved without immobilisation of the patient's head. The corresponding value for free-hand surgery was found to be 14.0 mm ± 6.8 mm.

The three surface-based registration systems discussed up to this point, all achieve intraoperative position-sensing by measuring the location of a hand-held probe. In order to record the location of a random surface point on the patient's face/scalp, the probe must therefore be manually brought into contact with that point. A switch must then be pressed to indicate that the point is to be recorded. Given that the accuracy of the registration process improves as the number of surface points is increased, surface-based registration strategies generally require a large number of points to be digitised. Intraoperative data collection can therefore be a time-consuming process when using a probe based system; Pelizzari *et al.* quote a figure of ten minutes **[25]**, while Olivier *et al.* state that calibration of their system and data collection took between forty-five and sixty minutes **[24]**. Accordingly, a number of research sites are currently investigating the use of laser range-finders and structured light systems as a means of performing rapid data collection, prior to surface-based registration **[26-28]**. Upon completion of the registration process, these video based systems are also used to provide surgical guidance via the *image fusion* technique.

A conventional neuronavigation system works by superimposing the position of a surgical tool/probe onto preoperative images. As such, the tool's position with respect to the operating room coordinate system is effectively transformed into the preoperative coordinate system. Image fusion on the other hand performs the opposite transformation, thus allowing preoperative data to be superimposed onto intraoperative video images. Image segmentation techniques are initially used to identify the anatomical tissues of interest in the preoperative CT/MRI images. Three-dimensional graphical models are then generated from this segmented image data. Using the transformation established by the intraoperative registration process, these graphical models are redefined in terms of the operating room coordinate system. The required transformation from operating room to video display (pixel) coordinates, is then obtained by calibrating the video camera that acquires the live images of the surgical site. By applying this second transformation, the segmented preoperative models can then be "fused" with live images of the patient's anatomy.

The ability to accurately overlay preoperative data onto images of the actual surgical site offers many potential benefits. As the surgeon can effectively "see" inside the patient, the size and location of the craniotomy can be optimised. Graphical navigation aids, for example the boundary of a tumour, can also be superimposed onto the video images. Any anatomical changes that have occurred between preoperative imaging and surgery will also be self evident. However, in order to view the images displayed on the video monitor, the surgeon is required to look away from the surgical site. Alternative display formats are therefore being investigated as a means of overcoming this limitation of the image fusion technique. Since many neurosurgical procedures are performed with the aid of an operating microscope, several systems have been developed that allow the injection of synthetic images into the microscope eyepieces **[18, 29, 30]**. The "heads-up" display concept used by virtual reality technology has also been proposed as a possible solution **[31]**.

A3.5 The Brain Motion Problem.

Intracranial neurosurgery deals with the unique anatomical situation of a soft tissue mass (i.e. the brain) that is completely surrounded by bone (i.e. the skull). As a

consequence, it is reasonably safe to assume that accurate, repeatable determination of the brain's location can be achieved by immobilising the skull. Accordingly, the registration strategies discussed in Sections A3.3 and A3.4, all rely upon *preoperative imaging* of the patient, followed by intraoperative registration processes that derive *rigid coordinate transformations*. If the stereotactic base ring, fiducial markers, or skin surfaces employed during the registration process, have not moved with respect to the brain, the preoperative images should therefore accurately reflect the true nature of the patient's anatomy at the start of surgery.

Unfortunately, during the actual surgical procedure itself, significant brain tissue distortion can occur. In order to gain access to the surgical site, many open neurosurgery techniques require a large section of bone to be removed from the skull. Depending upon the nature of the procedure, some form of brain tissue (e.g. a tumour) may also have to be removed. As a result of these actions, soft tissue distortion will occur under the influence of gravity. Alternatively, brain shrinkage may occur when cerebrospinal fluid is drained from the brain. Due to brain structures no longer being at their preoperatively imaged locations, registration accuracy is therefore progressively degraded during surgery.

The obvious solution to this "brain motion" problem is *intraoperative* CT or MRI scanning. The potential benefits offered by intraoperative CT have already been discussed in relation to the robotic-assisted neurosurgery system of Glauser *et al.* **[10]**. However, since CT-scanning is a form of radiography, these benefits are obtained at the expense of increased patient irradiation. In order to gain access to the surgical site, is also necessary to slide the patient out of the CT-scanner. Interventional MRI would therefore appear to be a more practicable option. MRI provides better soft tissue discrimination than CT, and does not involve the use of ionising radiation. Significantly, "open" MRI scanners, such as the one developed by General Electric Medical Systems (Milwaukee, WI, USA) **[32]**, also allow physical access to the patient during image acquisition. Intraoperative imaging and real-time instrument tracking can therefore be performed without moving the patient **[33]**. MRI does however impose unique requirements upon the operating room environment. In

addition to the need for magnetic shielding, only MRI-compatible (i.e. non-ferrous) surgical instruments and fixtures can be used if artefacts are to be avoided.

The use of intraoperative CT or MRI scanning avoids the need for registration. However, due to the fact that a dedicated scanner is required, these techniques are currently a very expensive means of overcoming the brain motion problem. The short term solution to this problem may therefore lie in developing registration transformations that are able to take into account the tissue deformations caused by surgery. As these *elastic registration* strategies will involve *deformable models* of the patient, the obvious methods for their implementation are localised spline functions and finite element analysis. Given that some form of intraoperative imaging will be required to track deformations during surgery, the previously discussed image fusion and localised ultrasound techniques may become more important in the near future.

A3.6 References.

[1] **Heilbrun, M.P.** Computed Tomography-Guided Stereotactic Systems, Clinical Neurosurgery, 1983, **31**, 564-581.

[2] **Kall, B.A., Kelly, P.J., Stiving, S.O.,** and **Goerss, S.J.** Computer-Assisted Stereotactic Neurosurgery: Functional Design and Clinical Applications. *Proceedings of the 14th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Paris, France, Oct. 29th - Nov. 1st. 1992, 14, 1079-1080.

[3] **Kall, B.A.** Computer-Assisted Surgical Planning and Robotics in Stereotactic Neurosurgery. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds). MIT Press, Cambridge, MA, 1996, 353-361.

[4] **Kwoh, Y.S., Hou, J., Jonckheere, E.A.** and **Hayati, S.** A Robot With Improved Absolute Positioning Accuracy for CT Guided Stereotactic Brain Surgery. *IEEE Trans. on Biomedical Engineering*, 1988, **35**(2), 153-160.

[5] **Kwoh, Y.S.,** *et al.* A New Computerised Tomographic-Aided Robotic Stereotaxis System. *Robotics Age,* June 1985, 7(6), 17-22.

[6] **Young, R.F.** Application of Robotics to Stereotactic Neurosurgery. *Neurological Research*, 1987, **9**(2), 123-129.

[7] **Engelberger, J.** Viewpoint: Robots Come of Age. *International Journal of Service Robot*, 1995, **1**(1), 3.

[8] **Drake, J.M., Joy, M., Goldenberg, A.** and **Kreindler, D.** Computer-Assisted and Robot-Assisted Resection of Thalamic Astrocytomas in Children. *Neurosurgery*, 1991, **29**(1), 27-31.

[9] **Drake, J.M.,** *et al.* Frameless Stereotaxy in Children. *Pediatric Neurosurgery*, 1994, 20(2), 152-159.

[10] **Glauser, D., Fankhauser, H., Epitaux, M., Hefti, J.L.** and **Jaccotte, A.** Neurosurgical Robot Minerva: First Results and Current Developments. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, MD, USA, 4-7 November 1995, 24-30.

[11] **Glauser, D.,** *et al.* Configuration of a Robot Dedicated to Stereotactic Surgery. *Stereotactic and Functional Neurosurgery,* 1990, **54-55**, 468-470.

[12] **Burckhardt, C.W., Flury, P.** and **Glauser, D.** Stereotactic Brain Surgery. *IEEE Engineering in Medicine and Biology Magazine*, 1995, **14**(3), 314-317.

[13] **Lea, J.T., Santos-Munné, J.J.** and **Peshkin, M.A.** Diagramming Registration Connectivity and Structure. *IEEE Engineering in Medicine and Biology Magazine*, 1995, **14**(3), 271-278.

[14] **Lavallée, S.,** *et al.* Image-Guided Operating Robot: A Clinical Application in Stereotactic Neurosurgery. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds). MIT Press, Cambridge, MA, 1996, 343-351.

[15] **Lavallée, S.,** *et al.* Computer Assisted Interventionist Imaging: The Instance of Stereotactic Brain Surgery. *Proceedings of the 6th Conference on Medical Informatics (MEDINFO'89)*, Beijing, China, 16-20 October 1989, 613-617.

[16] **Reinhardt, H.F.** Neuronavigation: A Ten-Year Review. In *Computer-Integrated Surgery*, R.H. Taylor, S. Lavallée, G.C. Burdea and R. Mösges (eds). MIT Press, Cambridge, MA, 1996, 329-341.

[17] **Galloway, R.L., Maciunas, R.J.** and **Edwards, C.A.** Interactive Image-Guided Neurosurgery. *IEEE Transactions on Biomedical Engineering*, December 1992, **39**(12), 1226-1231.

[18] **Friets, E.M., Strohbehn, J.W., Hatch, J.F.** and **Roberts, D.W.** A Frameless Stereotaxic Operating Microscope for Neurosurgery. *IEEE Transactions on Biomedical Engineering*, 1989, **36**(6), 608-617.

[19] **Kosugi, Y.,** *et al.* An Articulated Neurosurgical Navigation System Using MRI & CT Images. *IEEE Transactions on Biomedical Engineering*, 1988, **35**(2), 147-152.

[20] **Kato, A.,** *et al.* A Frameless, Armless Navigational System for Computer-Assisted Neurosurgery. *Journal of Neurosurgery*, 1991, **74**(5), 845-849.

[21] **Smith, K.R., Bucholz, R.D.** and **Foley, K.T.** Free-Hand Stereotactic Computer-Assisted Surgery. *Computer Assisted Radiology - Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR'95)*, Berlin, Germany, 21-24 June 1995, 859-863.

[22] **Maurer, C.R.,** *et al.* The Accuracy of Image-Guided Neurosurgery Using Implantable Fiducial Markers. *Computer Assisted Radiology - Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR'95)*, Berlin, Germany, 21-24 June 1995, 1197-1202.

[23] **Barnett, G.H.,** *et al.* Intraoperative Localisation Using an Armless, Frameless Stereotactic Wand. *Journal of Neurosurgery*, 1993, 78(3), 510-514.

[24] **Olivier, A., Germano, I.M., Cukiert, A.** and **Peters, T.** Frameless Stereotaxy for Surgery of the Epilepsies: Preliminary Experience. *Journal of Neurosurgery*, 1994, **81**(4), 629-633.

[25] **Ryan, M.J.,** *et al.* Frameless Stereotaxy with Real-Time Tracking of Patient Head Movement and Retrospective Patient-Image Registration. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, MD, USA, 4-7 November 1995, 1-7.

[26] **Grimson, W.E.L.,** *et al.* An Automatic Registration Method For Frameless Stereotaxy, Image-Guided Surgery, and Enhanced Reality Visualisation. *IEEE Transactions on Medical Imaging*, April 1996, **15**(2), 129-140.

[27] **Evans, R.J., Colchester, A.C.F.** and **Roberts, P.T.E.** VISLAN: A Computer-Vision Based System For Surgical Guidance. *Proceedings of the British Computer Society Conference on Interactive Technology in Surgery and Medicine*, Leeds, UK, 10th March 1995.

[28] **Davey, B.,** *et al.* Applying Stereoscopic Visualisation to Image-Guided Neurosurgery. *Proceedings of the First International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, Pittsburgh, PA, USA. 22-24 September 1994, 264-271.

[29] **Edwards, P.J.,** *et al.* Augmented Reality in the Stereo Operating Microscope for Otolaryngology and Neurosurgical Guidance. *Proceedings of the Second Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, MD, USA, 4-7 November 1995, 8-15.

[30] **Luber, J.** and **Mackevics, A.** Multiple Co-ordinate Manipulator (MKM) - A Computer-Assisted Microscope. *Computer Assisted Radiology - Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR'95)*, Berlin, Germany, 21-24 June 1995, 1121-1126.

[31] **Robb, R.A., Hanson, D.P.** and **Camp, J.J.** Computer-Aided Surgery Planning and Rehearsal at Mayo Clinic. *Computer*, January 1996, **29**(1), 39-47

[32] **Blumenfeld, S.M.** Technology Introduction in a Cost Conscious Age: Image-Guided Therapy. *Computer Assisted Radiology - Proceedings of the International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR'95)*, Berlin, Germany, 21-24 June 1995, 611-619.

[33] **Moriarty, T.M., Kikinis, R., Jolesz, F.A., Black, P.M. & Alexander, E.** Magnetic Resonance Imaging Therapy: Intraoperative MR-Imaging, *Neurosurgery Clinics Of North America*, 1996, 7(2), 323.

APPENDIX 4:

REGISTRATION GRAPHS

CONTENTS

		Page
A4.1	INTRODUCTION	A80
A4.2	REGISTRATION GRAPH TERMINOLOGY AND SYMBOLS	A80
A4.3	SAMPLE REGISTRATION GRAPH	A83
A4.4	REFERENCES	A85

A4. REGISTRATION GRAPHS.

A4.1 Introduction.

Depending upon the nature of the surgical application, computer-assisted surgery (CAS) will involve some or all of the following components:

- 3-D Preoperative Imaging (e.g. CT, MRI, DSA, etc.).
- 2-D Intraoperative Imaging (e.g. fluoroscopy, ultrasound, endoscopy, etc.).
- 3-D Intraoperative Position-Sensing (e.g. optical tracking, sonic localisation, etc.).
- Surgical Guidance Systems (e.g. passive manipulators, image fusion, etc.).

In order to perform the actual surgical procedure it may therefore be necessary to establish spatial transformations between several components. The individual system components may also require calibration. As the number of registration/calibration processes increase, it can become very difficult to conceptualise the overall registration process from a purely text-based description. Further confusion can also arise, due to the fact that some of the registration/calibration processes will be preoperative (i.e. off-line), and the remainder intraoperative (i.e. on-line). The same interpretation problems also apply to a number of the more complex robotic-assisted surgery applications. The diagrammatic technique proposed by Lea *et al.* **[1]**, has therefore been adopted by the current thesis as an additional means of conveying information about registration strategies. Having used a number of these *registration graphs* within the main thesis text, it is necessary to provide some form of explanation for readers who are unfamiliar with Lea et al's paper. Section A4.2 therefore summarises the basic concepts associated with the technique, and is followed by an illustrative example of a registration graph in Section A4.3.

A4.2 Registration Graph Terminology and Symbols.

The registration graph is a very effective means of explaining the entire registration strategy of a CAS or robotic-assisted surgery system. A typical registration graph depicts all of the major components involved in the registration process, including

aspects of the patient's anatomy. The nature of the measurement processes used to co-register these components is also illustrated. As a result, complex registration strategies, involving a chain of registration/calibration processes, can be represented by an easily understood diagrammatic format. However, the registration graph does not merely demonstrate how registration is established, it also highlights the conditions which must be satisfied in order to maintain registration.

Registration graphs based upon the proposals of Lea *et al.* **[1]**, essentially consist of groups of *nodes* and *motion groups* interconnected by *links*:

- **A NODE** (shown as a *small circle* on the registration graph) represents a physical object which is to be co-registered, and is labelled accordingly. Alternatively a node may be thought of as representing the coordinate frame of reference attached to the object. *Examples include*: robot, fluoroscope, CT-scanner etc. For reasons of clarity, the surgical plan is also shown as a node, and should therefore be interpreted as being registered with the system components.

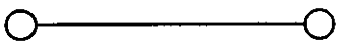
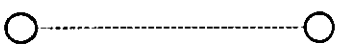
- **A MOTION GROUP** (shown as a *rounded rectangle* on the registration graph) is a visual aid used to highlight the fact that a group of nodes belong to a single rigid body. *Examples include*: patient (nodes: femur, fiducial pins, surgical plan) or surgical manipulator (nodes: robot and end-effector).

- **A LINK** (shown as a *line* on the registration graph) represents the act of measurement which establishes a known relationship between two objects (nodes). However, the link does not represent the actual spatial relationship itself. The three line types illustrated in Table A4.1 are employed to indicate different forms of measurement process.

If objects (nodes) move relative to one another, measurement operations (links) can become invalid resulting, in registration no longer being maintained. Patient motion during robotic-assisted surgery is a prime example of this scenario, as patient-to-robot registration is frequently obtained by indirect (induced link) methods (intermediate: imaging, 3D digitisers etc.). Unless real-time patient-to-robot (i.e. a sustained link)

registration can be established, patient immobilisation therefore becomes necessary to maintain registration.

Table A4.1 : Link Notation Employed by Registration Graphs.

SYMBOL	LINK TYPE
Thin Line: 	**A TRANSIENT LINK:** Represents a one-off measurement operation. *Example*: Acquisition of a single fluoroscopic image.
Bold Line: 	**A SUSTAINED LINK:** Represents an ongoing measurement process. *Examples*: Real-time infra-red optical tracking of the location of a marker on a moving bone. A sustained link is also formed when two objects are brought into contact and locked together.
Dashed Line: 	**AN INDUCED LINK:** Represents indirect knowledge of the spatial relationship between two objects, derived by intermediate measurement processes. *Example*: Patient position with respect to a surgical robot.

In order to indicate when a link is valid, and the conditions under which it will become invalid, the concept of *events* is therefore introduced:

- **AN EVENT TAG** (shown as a numbered link) indicates links that are simultaneously valid. *By convention:* links which are permanently valid, such as a link between two nodes belonging to the same rigid body (motion group), are left untagged as no relative motion can occur.

A4.3 Sample Registration Graph.

The easiest way to illustrate how these concepts are brought together, to form a registration graph, is to consider the measurement processes associated with a relatively simple CAS system. The registration graph for the Grenoble spine surgery system described by Lavallée *et al.* **[2]**, is therefore presented here as an illustrative example. Since an in-depth discussion of this system has already been provided in Section 3.3.2 of the main thesis, the system details are merely summarised here. The aim of this system is to preoperatively plan, and then accurately implement, the insertion of pedicle screws into one or more vertebra(e). In practice, this is achieved through preoperative CT-scanning and the intraoperative use of the Optotrak 3020 (Northern Digital Inc., Waterloo, Ontario, Canada) real-time infra-red optical digitiser. Intraoperative registration is therefore established by using the Optotrak system to track a group of infra-red light emitting diodes (IR-LEDs) on a hand-held probe, which in turn is used to digitise the locations of surface points on the vertebra. The surgical plan is then reproduced by using the Optotrak system to track IR-LEDs on a surgical drill, thus allowing the drill's position to be displayed with respect to preoperative images. Throughout the entire surgical procedure, a dynamic operating room coordinate system is defined by tracking a group of IR-LEDs mounted on a vertebral flag, which in turn is rigidly attached to one of the patient's vertebra.

Interpretation of the registration graph for the Grenoble spine surgery system (refer to Figure A4.1) should initially focus upon the major components involved in the registration process. The major system components are obviously the CT-scanner, the Optotrak position-sensor, the vertebral flag, the hand-held probe, and the surgical drill. These components are therefore depicted as nodes (small circles) on the registration graph. In addition, the patient's vertebra and the surgical plan must also be depicted as nodes. Since the vertebral flag is rigidly attached to the patient's vertebra, both the flag and the vertebra are depicted as belonging to the "patient" motion group (rounded rectangle). By convention the surgical plan is also depicted as belonging to this motion group.

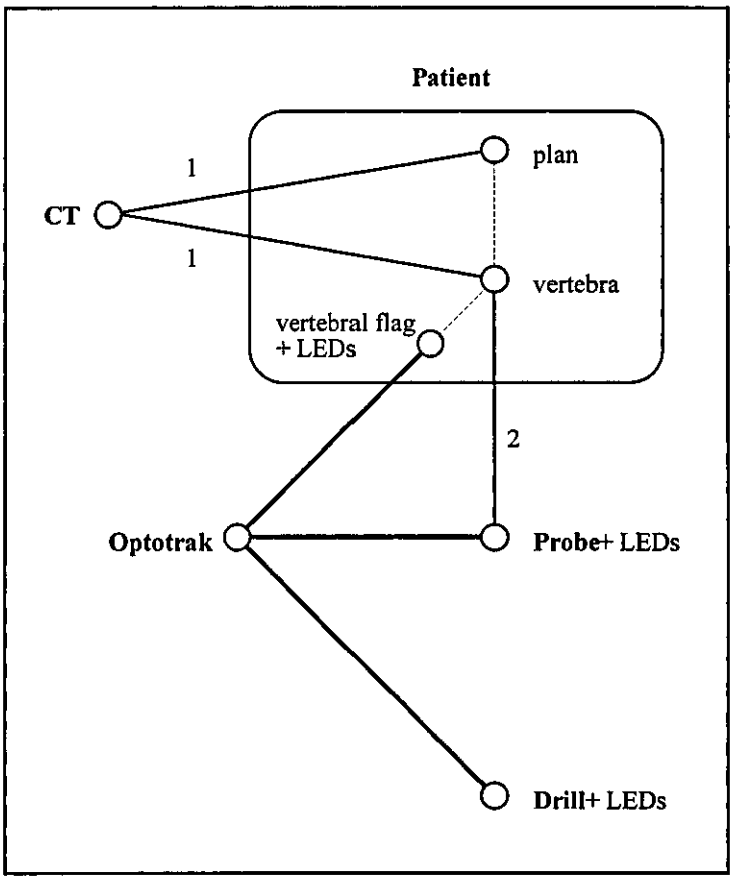


Fig. A4.1 : Registration Graph for the Grenoble Spine Surgery System.
(Adapted from Lea *et al.* [1])

Turning now to the preoperative measurement processes (i.e. the links), it can obviously be appreciated that CT imaging of the patient's vertebra is a one-off examination. A transient link is therefore established between the CT and vertebra nodes. By convention a similar transient link is also depicted as existing between the plan and CT nodes. However, as changes in the patient's anatomy can occur between preoperative imaging and surgery, the CT data can only be guaranteed to be valid at the time of image acquisition. Accordingly, the [CT : vertebra] and [CT : plan] transient links are marked with an event tag (in this case 1). Upon completion of the surgical planning stage of the procedure, an induced link is said to exist between the plan and vertebra.

Throughout the actual surgical procedure, the Optotrak system measures the position of the vertebral flag in real-time. A sustained link therefore exists between the Optotrak and vertebral flag nodes. Since the hand-held probe and surgical drill are

also tracked in real-time, [Optotrak : Probe] and [Optotrak : Drill] sustained links also exist. However, as the patient's spine will move during the breathing cycle, the measurements of the vertebral surface, performed for intraoperative registration purposes, are only valid while the probe tip is actually in contact with the vertebra. The [Probe : vertebra] link is therefore a transient one and also requires an event tag (in this case 2).

Measurement of the vertebral surface [Optotrak : Probe : vertebra], coupled with simultaneous measurement of the vertebral flag location [Optotrak : vertebral flag] allows a rigid transformation between the vertebra and the attached vertebral flag to be indirectly obtained. The [vertebral flag : vertebra] induced link is therefore created as a result of the intraoperative registration process. All of the measurement processes required to establish registration are then complete. The overall registration strategy of the system is therefore defined by the subgraph [plan : vertebra : vertebral flag : Optotrak : drill].

A4.4 References.

1. **Lea, J.T., Santos-Munné, J.J.** and **Peshkin, M.A.** Diagramming Registration Connectivity and Structure. *IEEE Engineering in Medicine and Biology Magazine*, 1995, **14**(3), 271-278.

2. **Lavallée, S., *et al.*** Computer Assisted Spine Surgery: A Technique for Accurate Transpedicular Screw Fixation Using CT Data and a 3-D Optical Localizer. *Proceedings of the First International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, Pittsburgh, PA, USA, 22-24 September 1994, 315-322.

APPENDIX 5:

SOLUTION OF THE LINEAR LEAST SQUARES PROBLEM

CONTENTS

		Page
A5.1	INTRODUCTION	A86
A5.2	THE NORMAL EQUATIONS	A87
A5.3	THE PSEUDO INVERSE METHOD	A88
A5.4	THE ORTHOGONAL MATRIX FACTORISATION APPROACH	A90
	A5.4.1 Q-R Decomposition	A90
	A5.4.2 Singular Value Decomposition	A95
A5.5	Q-R DECOMPOSITION SAMPLE CALCULATION	A98
A5.6	REFERENCES	A101

A5. SOLUTION OF THE LINEAR LEAST SQUARES PROBLEM.

A5.1 Introduction.

In Chapter 4, it was shown that the camera calibration problem can be described by a series of linear simultaneous equations, which may be expressed in matrix form as:

$$[A].x = b \qquad \text{.....(A5.1)}$$

Where:

- The vector **x** contains the unknown coefficients of the transformation matrix.
- The vector **b** contains the measured image coordinate data.
- The square matrix **[A]** contains the elements required to satisfy the model of the image formation process (including the world coordinates of the control points).

In the case of the Direct Linear Transformation (DLT) the vector **x** contains eleven unknown coefficients. Since each control point yields two simultaneous equations, in theory, six imaged control points therefore provide sufficient information to derive an exact solution to the camera calibration problem (as defined by A5.1). However, in practice, because of errors in the calibration data (e.g. measurement errors, digitising errors, noise, etc.) this exact solution often leads to poor three-dimensional reconstruction accuracy. As discussed in Chapter 2, over-sampling (i.e. the use of more control points than are theoretically required to give an exact solution) is therefore commonly employed as a means of overcoming such errors. This has the effect of 'averaging out' errors, thus providing a more stable set of calibration data. However, in doing so equation (A5.1) becomes:

$$[A].x \cong b \qquad \text{.....(A5.2)}$$

Where: [A] = a [m x n] matrix with (m > n).

x = a [n x 1] vector.

b = a [m x 1] vector.

Such a system of equations is described as being *overdetermined* (m conditions and n unknowns, m > n) and cannot logically be expected to provide an exact solution for all

the equations simultaneously. Consequently, a mathematical technique is required which gives a 'best fit' type of solution to problem (A5.2), not necessarily satisfying any of the equations exactly, but minimising any errors which may be present. The most commonly used 'best fit' methods are the least squares techniques outlined in this appendix.

A5.2 The Normal Equations.

For the overdetermined case, the matrix systems (A5.1) and (A5.2) may be stated as:

$$[\mathbf{A}].\mathbf{x} = \mathbf{b} + \mathbf{e} \qquad \text{.....(A5.3)}$$

Where: **e** is an error or residual vector.

The aim of any 'best fit' method is to obtain the **x** vector solution which minimises this residual vector. This is generally achieved by calculating **x** from a method which minimises a norm of **e**, most notably the Euclidean Norm:

$$\| e \|_E = (e^T.e)^{1/2} \qquad \text{.....(A5.4)}$$

and its square, referred to as the *sum of squares*:

$$e^T.e = \Sigma e_i^2 \qquad \text{.....(A5.5)}$$

The *least squares* solution **x** of (A5.2) is the set of parameters which minimises this sum of squares.

Applying this minimisation of $(e^T.e)$ with respect to **x**, to equation (A5.3) yields:

$$\begin{aligned} e^T.e &= (b - [A].x)^T(b - [A].x) \\ &= (b^T.b - b^T[A].x - x^T.[A]^T.b + x^T.[A]^T.[A].x) \\ &= b^T.b - 2.x^T.[A]^T.b + x^T.[A]^T.[A].x \end{aligned} \qquad \text{.....(A5.6)}$$

Differentiating (A5.6) with respect to the elements of **x**, and setting the result equal to zero to obtain a minimum value yields:

$$0 = [A]^T.[A].x - [A]^T.b \qquad .....(A5.7)$$

which may be re-written as:

$$\mathbf{[A]^T.[A].x = [A]^T.b} \qquad .....(A5.8)$$

The equations expressed in matrix form by (A5.8) are commonly referred to as the *normal equations*. The remainder of this appendix outlines three linear least squares techniques, which use the normal equations as the starting point for the solution of an overdetermined set of equations expressed in the form of (A5.2).

A5.3 The Pseudo Inverse Method.

Examination of the normal equations shows them to be n simultaneous linear equations in n unknowns, thus allowing a direct solution of the form:

$$x \cong [A^T.A]^{-1} . [A]^T.b \qquad .....(A5.9)$$

Since this equation is comparable with the solution of the exact problem (A5.1):

$$x = [A]^{-1}.b \qquad .....(A5.10)$$

the term:

$$[A]^* = [A^T.A]^{-1} . [A]^T \qquad .....(A5.11)$$

is generally referred to as being the *generalised* or *pseudo inverse* of [A]. Equation (A5.9) may therefore be alternatively stated as:

$$x \cong [A]^*.b \qquad .....(A5.12)$$

and the direct solution of the normal equations is commonly referred to as the *pseudo inverse method.*

The major computational task associated with the pseudo inverse method, is the calculation of the inverse matrix term $[A^T.A]^{-1}$. This calculation is normally

simplified/speeded up by the application of a triangular decomposition (or LU factorisation) technique of the form:

$$[A^T.A] = [L][U] \qquad \text{.....(A5.13)}$$

Where: [L] = a lower triangular matrix,
[U] = an upper triangular matrix.

For the purposes of this thesis, the Crout reduction technique has been adopted to perform this triangular decomposition. This particular decomposition technique requires the diagonal elements of the matrix [U] to be set to one. Hence, in general a matrix [A] is factorised into matrices [L] and [U] with the form:

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \begin{bmatrix} 1 & U_{12} & U_{13} \\ 0 & 1 & U_{23} \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{.....(A5.14)}$$

The [L] & [U] coefficients may then be calculated by solving simple linear equations:

$$A_{11} = L_{11} \quad A_{21} = L_{11}.U_{12} \quad A_{31} = L_{11}.U_{13}$$
$$A_{21} = L_{21} \quad A_{22} = L_{21}.U_{12} + L_{22} \quad A_{32} = L_{21}.U_{13} + L_{22}.U_{23}$$
$$A_{31} = L_{31} \quad A_{23} = L_{31}.U_{12} + L_{32} \quad A_{33} = L_{31}.U_{13} + L_{32}.U_{23} + L_{33}$$

Such a decomposition allows the required inverse matrix $[A^T.A]^{-1}$ to be obtained, by utilising the property that a matrix multiplied by its inverse results in the identity matrix [I]:

$$[A^T.A][A^T.A]^{-1} = [I] \qquad \text{.....(A5.15)}$$

Letting $[A^T.A] = [L][U]$

$$[L][U][A^T.A]^{-1} = [I] \qquad \text{.....(A5.16)}$$

Letting $[Y] = [U][A^T.A]^{-1}$

$$[L][Y] = [I] \qquad \text{.....(A5.17)}$$

Obtaining the solution for $[A^T.A]^{-1}$ therefore involves :

1) Obtaining the matrices [L] & [U] for $[A^T.A]$ using the Crout reduction.

2) Obtaining [Y] by forward substitution in (A5.17).

3) Obtaining $[A^T.A]^{-1}$ by backward substitution in (A5.16).

Having calculated the $[A^T.A]^{-1}$ matrix term, the pseudo inverse solution for **x** via equations (A5.9) involves a simple matrix multiplication process.

A5.4 The Orthogonal Matrix Factorisation Approach.

Given its simplicity, the pseudo inverse method has become the most commonly used technique to solve the linear least squares problem. Unfortunately the special structure of the $[A^T.A]$ matrix can, in certain circumstances, lead to unreliable solutions being obtained due to numerical instabilities and loss of information during its formation. For this reason, a number of references **[1 - 4]** recommend the use of a least squares method which avoids the formation of the $[A^T.A]$ matrix. Fortunately such a solution may be obtained from one of a number of orthogonal matrix factorisation approaches. This section outlines two of these orthogonal matrix factorisation methods: Q-R Decomposition and Singular Value Decomposition (SVD).

A5.4.1 Q-R Decomposition.

In general, orthogonal matrix methods are based upon the Q-R decomposition of the [A] matrix:

$$[A] = [Q][R] \qquad \text{.....(A5.18)}$$

Where: [A] = a [m x n] matrix.

[Q] = a [m x m] orthogonal matrix.

[R] = a [m x n] upper triangular matrix.

Applying this decomposition to the normal equations (A5.8), and substituting for $[A]^T$ yields:

$$0 = [A]^T.[A].x - [A]^T.b$$
$$0 = [A]^T (b - [A].x)$$
$$0 = [R]^T[Q]^T.b - [R]^T[Q]^T.[A].x \quad \text{.....(A5.19)}$$

Since $[R] = [Q]^{-1}.[A]$

& $[Q]^{-1} = [Q]^T$ (orthogonal matrix property)

Then: $[R] = [Q]^T[A]$

Hence the normal equations (A5.19) may be stated as:

$$0 = [R]^T([Q]^T.b - [R].x) \quad \text{.....(A5.20)}$$

Assuming that the Q-R Decomposition of [A] has been performed, a least squares solution for **x** can therefore be obtained (without forming the $[A^T.A]$ term) from the matrix system:

$$\mathbf{[R].x = [Q]^T.b} \quad \text{.....(A5.21)}$$

As with LU triangular decomposition, a number of methods are available to perform the Q-R decomposition of matrix [A] (e.g. Gram-Schmidt Orthonormalisation, Householder Reflection, etc.). However, for reasons of simplicity and storage requirements, the *Givens transformation method* has been adopted for this application. At the simplest level, the Givens method can be considered to be a number of eliminations performed on the rows of matrix [A] by a series of premultiplying orthogonal transformations. These premultiplying terms perform plane rotations, which transform two vectors so that an element of one becomes zero:

For a vector $\mathbf{v} = (v_1, v_2)^T$ there is a [2 x 2] orthogonal matrix [G]:

$$[G] = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \qquad \text{Where } c = v_1/(v_1^2 + v_2^2)^{½}$$
$$s = v_2/(v_1^2 + v_2^2)^{½}$$
$$c^2 + s^2 = 1$$

.....(A5.22)

Such that:

$$[G]v = \begin{bmatrix} (v_1^2 + v_2^2)^{1/2} \\ 0 \end{bmatrix} = \begin{bmatrix} R \\ 0 \end{bmatrix} \qquad \text{.....(A5.23)}$$

In this way, application of the relevant premultiplying terms (stored in the [Q] matrix), will result in the required [R] matrix being obtained from matrix [A]. Having calculated matrices [R] and [Q] using Givens method, solution of the **x** vector which satisfies the least squares equations (A5.21), simply requires a matrix multiplication operation followed by a backward substitution process. Given that a trapezoidal matrix is produced by the Q-R decomposition of an overdetermined set of linear equations, the equations (A5.21) actually reduce to the form:

$$[R].x = q_1 \qquad \text{.....(A5.24)}$$

Or:

$$\begin{vmatrix} R_{11} & R_{12} & R_{13} & R_{14} \\ 0 & R_{22} & R_{23} & R_{24} \\ 0 & 0 & R_{33} & R_{34} \\ 0 & 0 & 0 & R_{44} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix} \cdot \begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{vmatrix} = \begin{vmatrix} Q^T b_1 \\ Q^T b_2 \\ Q^T b_3 \\ Q^T b_4 \\ Q^T b_5 \\ Q^T b_6 \\ Q^T b_7 \end{vmatrix} = \left| \begin{array}{c} q_1 \\ \hline q_2 \end{array} \right|$$

Where for an [A] matrix of order [m x n]:

- vector $\mathbf{q_1}$ = [n x 1]
- vector $\mathbf{q_2}$ = [(m-n) x 1]

  Note: The solution for **x** is therefore independent of the vector $\mathbf{q_2}$ terms.

Since an element set to zero by one rotation must not be made non-zero by another, the ordering of these rotations is crucial. The Givens method therefore operates on the elements of matrix [A] in a column by column fashion, eliminating the elements below the diagonal of matrix [A] one by one (refer to Section A5.5 for a sample calculation). The generalised order of operation is therefore as follows:

v_1 = The diagonal element of column j = A[j,j] for j = 1 to n

v_2 = The elements in column j below diagonal = A[k,j]

.....(A5.25)

Where:

k = 2 to m for column 1 (j = 1),

= 3 to m for column 2 (j = 2),

= 4 to m for column 3 (j = 3), etc.

Once the two relevant elements of [A] have been assigned to $\mathbf{v_1}$ and $\mathbf{v_2}$, it is often wise to normalise the data prior to calculation of the **c** and **s** terms. This generally involves dividing the two values by the larger of their absolute values:

$$v_{1norm} = v_1/t \quad \& \quad v_{2norm} = v_2/t \qquad \text{.....(A5.26)}$$

Where: $t = abs(v_1)$ if $abs(v_1) > abs(v_2)$

or $t = abs(v_2)$ if $abs(v_2) > abs(v_1)$

Such a normalisation has the effect of keeping the data at a manageable size, whilst also minimising rounding errors. Its inclusion does however require a check for division by zero to be built into the algorithm. Having calculated the **c** and **s** values for a specific element's elimination, the relevant orthogonal premultiplication must be carried out. Starting from the general case for the Givens transformation matrix:

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} R \\ 0 \end{bmatrix} \qquad \text{.....(A5.27)}$$

and the matrix multiplication:

$$R[k][i] = \sum Q^T[k][j] * A[j][i] \quad (\text{for } j = 1 \text{ to } m) \qquad \text{.....(A5.28)}$$

a general case premultiplication orthogonal transformation matrix can be built up as follows. Letting:

$$v_1 = A[a][a] \quad \& \quad v_2 = A[b][a] \qquad \text{.....(A5.29)}$$

We require the following multiplications, in order to eliminate the $\mathbf{v_2}$ term:

$$R[a][a] = (c * A[a][a]) + (s * A[b][a]) \qquad \text{.....(A5.30)}$$

$$R[b][a] = (-s * A[a][a]) + (c * A[b][a]) \qquad \text{.....(A5.31)}$$

Thus the **c** & **s** values must occupy the following $[Q]^T$ locations if (A5.27) is to be satisfied:

$$R[a][a] = (Q^T[a][a] * A[a][a]) + (Q^T[a][b] * A[b][a]) \quad .....(A5.32)$$

$$R[b][a] = (Q^T[b][a] * A[a][a]) + (Q^T[b][b] * A[b][a]) \quad .....(A5.33)$$

Given that we do not wish to make previously eliminated terms non-zero again, the remaining terms of the premultiplication matrix $[Q]^T$ must be of the following form:

$$[Q]^T[A] = \left|\begin{array}{ccccc||ccc|} \mathbf{1} & 0 & 0 & 0 & 0 & x & x & x \\ 0 & \mathbf{c} & 0 & \mathbf{s} & 0 & 0 & \mathbf{v_1} & x \\ 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & x \\ 0 & \mathbf{-s} & 0 & \mathbf{c} & 0 & 0 & \mathbf{v_2} & x \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 & x & x \end{array}\right| \quad \begin{array}{l} \\ \\ \mathbf{Row} \\ \\ \mathbf{Row} \\ \\ \end{array}$$

.....(A5.34)

Rather than perform these matrix multiplications in full it is possible, due to the sparsity of the $[Q]^T$ matrix, to simply multiply the [A] matrix by the two rows of $[Q]^T$ which contain the **c** and **s** terms (all other terms remain unchanged). In this way the number of calculations may be greatly reduced for large systems of equations. The general case algorithm is thus:

For i = 1 to n:

$$R[a][i] = (c * A[a][i]) + (s * A[b][i]) \quad .....(A5.35)$$

$$R[b][i] = (-s * A[a][i]) + (c * A[b][i]) \quad .....(A5.36)$$

Since we require $[Q]^T$ here (and not [Q]) in order to be able to form the least squares solution $[R].x = [Q]^T.b$, exactly the same rotation algorithm may be applied to update the $[Q]^T$ matrix.

A5.4.2 Singular Value Decomposition.

The singular value decomposition (SVD) is derived from the Q-R decomposition by further factorising the [R] matrix:

$$[R] = [S][V]^T \quad \text{....(A5.37)}$$

Hence, replacing [Q] by [U] and substituting for $[R] = [S][V]^T$ in (A5.18) yields:

$$\mathbf{[A] = [U][S][V]^T} \quad \text{....(A5.38)}$$

Where:

[U] = a [m x n] matrix obtained by dropping arbitrary [Q] matrix elements.

[S] = a [n x n] diagonal matrix (diagonal elements chosen to be positive)

[V] = a [n x n] orthogonal matrix.

It can be shown by direct substitution **[1]**, that the generalised inverse of matrix [A] in (A5.38) is defined by:

$$[A]^+ = [V][S]^+[U]^T \quad \text{....(A5.39)}$$

Where:

$$S_{ii}^+ = \begin{cases} 1/Sii & \text{for } Sii > q \\ 0 & \text{for } Sii \leq q \end{cases} \quad \text{....(A5.40)}$$

And: **q** is a user defined tolerance.

Substituting $[A]^+$ from (A5.39) into (A5.12) therefore yields:

$$\mathbf{x = [V][S]^+[U]^T.b} \quad \text{.....(A5.41)}$$

Solution of the least squares problem using SVD consequently involves orthogonalisation of matrix [A] in accordance with (A5.38), followed by a sequence of matrix multiplication processes to satisfy (A5.41).

In practice, when attempting to perform the orthogonalisation of matrix [A], it is more convenient to calculate the [n x n] orthogonal matrix [V], which transforms the

[m x n] matrix [A] into another [m x n] matrix [B] whose columns are orthogonal. This alternative approach to the SVD problem therefore requires [V] such that:

$$[B] = [A][V] = (b_1, b_2, ..., b_n) \quad ....(A5.42)$$

Where:

$$b_i^T b_j = S_i^2 \delta_{ij} \quad ....(A5.43)$$

with Kronecker delta (δ_{ij}) taking values:

$$\delta_{ij} = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases} \quad ....(A5.44)$$

and (to satisfy orthogonality):

$$[V][V]^T = [V]^T[V] = [I] \quad ....(A5.45)$$

Since only the square of the quantities S_i are defined by (A5.43), they may be assigned as either negative or positive values. For the purposes of this thesis, these *singular values* of the matrix [A] will be arbitrarily assigned as positive values.

Defining unit orthogonal vectors $\mathbf{u_j}$:

$$u_j = b_j/S_j \quad ....(A5.46)$$

these vectors may be collected into a [m x n] matrix, and the singular values into a diagonal [n x n] matrix, allowing the statement:

$$[B] = [U][S] \quad ....(A5.47)$$

Where:

$$[U]^T[U] = [I] \quad ....(A5.48)$$

Combining (A5.42) & (A5.47) yields:

$$[A][V] = [U][S] \quad ....(A5.49)$$

Hence, applying (A5.45) to (A5.49), the orthogonality of [V] yields (A5.38):

$$[A] = [U][S][V]^T \qquad \text{....(A5.38)}$$

The matrix [V] calculated from equations (A5.42 - A5.49) can therefore be seen to satisfy the requirements of SVD.

The matrix [V] required to accomplish the orthogonalisation represented by (A5.42), is built up as a product of simpler matrices. The specific matrices used in this product are plane rotations. It can be shown that if $V^{(k)}$ is a rotation of angle ϕ in the ij plane, then all the elements of $V^{(k)}$ will be the same as those in a unit vector of order n except for:

$$V_{ii}^{(k)} = \cos\phi = V_{jj}^{(k)} \qquad \text{.....(A5.50)}$$

$$-V_{ij}^{(k)} = \sin\phi = V_{ji}^{(k)} \qquad \text{.....(A5.51)}$$

$V^{(k)}$ as a result affects only two columns of any matrix it multiplies. Labelling these columns as x & y, the effect of a single rotation is as follows:

$$(x, y)\begin{pmatrix}\cos\Phi & -\sin\Phi \\ \sin\Phi & \cos\Phi\end{pmatrix} = (X, Y) \qquad \text{....(A5.52)}$$

Hence:

$$X = x.\cos\phi + y.\sin\phi \qquad \text{.....(A5.53)}$$

$$Y = -x.\sin\phi + y.\cos\phi \qquad \text{.....(A5.54)}$$

If X & Y are to be orthogonal, then:

$$X^TY = 0 = -(x^Tx - y^Ty)\sin\phi.\cos\phi + x^Ty(\cos^2\phi - \sin^2\phi) \qquad \text{....(A5.55)}$$

A further condition on the choice of the rotation angle ϕ, results from the fact that it is convenient if the rotation can order the columns of the orthogonalised matrix [B] by length, so that the singular values are in decreasing order of size. Hence we also require:

$$X^TX - x^Tx \geq 0 \qquad \text{.....(A5.56)}$$

The specific formulae for the sine and cosine of the angle of rotation, which satisfy (A5.55) & (A5.56), are generally expressed in terms of the quantities :

$$p = x^Ty \qquad \text{....(A5.57)}$$

$$q = x^Tx - y^Ty \qquad \text{....(A5.58)}$$

$$v = (4p^2 + q^2)^{½} \qquad \text{....(A5.59)}$$

They are for $q \geq 0$:

$$\cos\phi = [(v + q)/(2v)]^{½} \qquad \text{....(A5.60)}$$

$$\sin\phi = p / (v .\cos\phi) \qquad \text{....(A5.61)}$$

And for $q < 0$:

$$\sin\phi = 1 \quad \& \quad \cos\phi = 0 \qquad \text{....(A5.62)}$$

In this way, the required ordering of the singular values is ensured. It can easily be appreciated that such rotations must be arranged in a specific sequence, in order to carry out the required orthogonalisation. Care is required to ensure that subsequent rotations do not undo the orthogonalisation effect of their predecessors.

A5.5 Q-R Decomposition Sample Calculation.

Input Data:

$$[A] = \begin{bmatrix} 0.4087 & 0.1594 \\ 0.4302 & 0.3516 \\ 0.6246 & 0.3384 \end{bmatrix} \qquad [Q]^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

i) Elimination of $[A_{21}]$ Term:

Givens General Form:

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} R \\ 0 \end{bmatrix}$$

$$v_1 = [A_{11}] = 0.4087 \qquad r = (v_1^2 + v_2^2)^{½} = 0.5934$$

$$v_2 = [A_{21}] = 0.4302 \qquad c = v_1/r = 0.6887$$
$$s = v_2/r = 0.7249$$

[R] Matrix Calculation: $[Q_1]^T[A] = [R_1]$

$$\begin{bmatrix} 0.6887 & 0.7249 & 0.0000 \\ -0.7249 & 0.6887 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \begin{bmatrix} 0.4087 & 0.1594 \\ 0.4302 & 0.3516 \\ 0.6246 & 0.3384 \end{bmatrix} = \begin{bmatrix} 0.5934 & 0.3647 \\ 0.0000 & 0.1266 \\ 0.6246 & 0.3384 \end{bmatrix}$$

$[Q]^T$ Matrix Calculation: $[Q]^T = [Q_1]^T[I]$

$$\begin{bmatrix} 0.6887 & 0.7249 & 0.0000 \\ -0.7249 & 0.6887 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.6887 & 0.7249 & 0.0000 \\ -0.7249 & 0.6887 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

ii) Elimination of $[A_{31}]$ term:

$$v_1 = [R_{11}] = 0.5934 \qquad r = (v_1^2 + v_2^2)^{1/2} = 0.8615$$
$$v_2 = [R_{31}] = 0.6246 \qquad c = v_1/r = 0.6887$$
$$s = v_2/r = 0.7249$$

[R] Matrix Calculation: $[Q_2]^T[Q_1{}^TA] = [R_2]$

$$\begin{bmatrix} 0.6887 & 0.0000 & 0.7249 \\ 0.0000 & 1.0000 & 0.0000 \\ -0.7249 & 0.0000 & 0.6887 \end{bmatrix} \begin{bmatrix} 0.5934 & 0.3647 \\ 0.0000 & 0.1266 \\ 0.6246 & 0.3384 \end{bmatrix} = \begin{bmatrix} 0.8615 & 0.4965 \\ 0.0000 & 0.1266 \\ 0.0000 & -0.0313 \end{bmatrix}$$

$[Q]^T$ Matrix Calculation: $[Q]^T = [Q_2]^T[Q_1]^T$

$$\begin{bmatrix} 0.6887 & 0.0000 & 0.7249 \\ 0.0000 & 1.0000 & 0.0000 \\ -0.7249 & 0.0000 & 0.6887 \end{bmatrix} \begin{bmatrix} 0.6887 & 0.7249 & 0.0000 \\ -0.7249 & 0.6887 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} = \begin{bmatrix} 0.4743 & 0.4993 & 0.7249 \\ -0.7249 & 0.6887 & 0.0000 \\ -0.4993 & -0.5256 & 0.6881 \end{bmatrix}$$

iii) Elimination of $[A_{32}]$ term:

$$v_1 = [R_{22}] = 0.1266 \qquad r = (v_1^2 + v_2^2)^{½} = 0.1304$$
$$v_2 = [R_{32}] = -0.0313 \qquad c = v_1/r = 0.9707$$
$$s = v_2/r = -0.2401$$

[R] Matrix Calculation: $[Q_3]^T[Q_2^TQ_1^TA] = [R_3]$

$$\begin{bmatrix} 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.9707 & -0.2401 \\ 0.0000 & 0.2401 & 0.9707 \end{bmatrix}\begin{bmatrix} 0.8615 & 0.4965 \\ 0.0000 & 0.1266 \\ 0.0000 & -0.0313 \end{bmatrix} = \begin{bmatrix} 0.8615 & 0.4965 \\ 0.0000 & 0.1304 \\ 0.0000 & 0.0000 \end{bmatrix}$$

$[Q]^T$ Matrix Calculation: $[Q]^T = [Q_3]^T[Q_2^TQ_1^T]$

$$\begin{bmatrix} 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.9707 & -0.2401 \\ 0.0000 & 0.2401 & 0.9707 \end{bmatrix}\begin{bmatrix} 0.4743 & 0.4993 & 0.7249 \\ -0.7249 & 0.6887 & 0.0000 \\ -0.4993 & -0.5256 & 0.6881 \end{bmatrix} = \begin{bmatrix} 0.4743 & 0.4993 & 0.7249 \\ -0.5838 & 0.7948 & -0.1652 \\ -0.6588 & -0.3448 & 0.6881 \end{bmatrix}$$

iv)Orthogonality Check on $[Q]^T[Q] = [I]$

$$\begin{bmatrix} 0.4743 & 0.4993 & 0.7249 \\ -0.5838 & 0.7948 & -0.1652 \\ -0.6588 & -0.3448 & 0.6811 \end{bmatrix}\begin{bmatrix} 0.4743 & -0.5838 & -0.6558 \\ 0.4993 & 0.7948 & -0.3448 \\ 0.7249 & -0.1652 & 0.6881 \end{bmatrix} = \begin{bmatrix} 0.9998 & 0.0000 & 0.0003 \\ 0.0000 & 0.9998 & 0.0003 \\ 0.0003 & 0.0003 & 0.9991 \end{bmatrix}$$

v) Decomposition check on [Q][R] = [A]

$$\begin{bmatrix} 0.4743 & -0.5838 & -0.6588 \\ 0.4993 & 0.7948 & -0.3448 \\ 0.7249 & -0.1652 & 0.6881 \end{bmatrix}\begin{bmatrix} 0.8615 & 0.4965 \\ 0.0000 & 0.1304 \\ 0.0000 & 0.0000 \end{bmatrix} = \begin{bmatrix} 0.4086 & 0.1594 \\ 0.4301 & 0.3515 \\ 0.6245 & 0.3381 \end{bmatrix}$$

A5.6 References.

[1] **Nash, J.C.** *Compact Numerical Methods for Computers: Linear Algebra & Function Minimisation.* Adam Hilger Ltd, Bristol, 1979. (ISBN 0-85274-330-0)

[2] **Rice, J.R.** *Matrix Computations & Mathematical Software.* McGraw-Hill. Tokyo, Japan, 1983. (ISBN 0-07-066509-5)

[3] **Jennings, A.** and **McKeown, J.J.** *Matrix Computation.* 2nd Edition. John Wiley, Chichester, England, 1992. (ISBN 0-471-93527-1)

[4] **Lawson, C.L.** and **Hanson, R.J.** *Solving Least Squares Problems.* Prentice Hall, New Jersey, 1974. (ISBN 0-13-822585-0)