Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author’s right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author’s permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage. [http://researchspace.auckland.ac.nz/feedback](http://researchspace.auckland.ac.nz/feedback)

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the [Library Thesis Consent Form](http://researchspace.auckland.ac.nz/forms/library-thesis-consent) and [Deposit Licence](http://researchspace.auckland.ac.nz/forms/deposit-licence).

Note : Masters Theses

The digital copy of a masters thesis is as submitted for examination and contains no corrections. The print copy, usually available in the University Library, may contain corrections made by hand, which have been requested by the supervisor.
3D Reconstruction of Patient Specific Bone Models for Image Guided Orthopaedic Surgery

Pavan Gamage

February 2012

Supervisor: Prof. Shane Xie
Co-Supervisors: Dr. Patrice Delmas and Prof. Peter Xu

Abstract

Visualisation of patient-specific fractured bone in 3D plays an important role in image guided orthopaedic surgery. Existing research often focuses on intra-operative registration of the patients’ anatomy with pre-operatively obtained 3D volumetric data (e.g. CT scans) utilising fiduciary markers. This expensive and invasive approach is not routinely available for diagnostics, and a majority of fracture reduction procedures currently solely relies on 2D x-ray/fluoroscopic images.

This research aims to assist orthopaedic surgeons in all steps of the femur fracture reduction procedure by introducing 3D anatomical visualisation. Studies conducted on femur fracture reduction has confirmed that computer-aided systems can significantly improve the accuracy of orthopaedic procedures by augmenting the current 2D image guidance with an interactive display of 3D bone models. This research indicates that the positioning errors, which generate bone misalignments and complications, will be reduced through the introduction of 3D bone fragment visualisation during surgical procedures. Consequently 3D visualisation of anatomy plays an important role in image-guided orthopaedic surgery and most importantly contributes to minimally invasive procedures.

The research goals of this thesis are achieved through the construction of a 3D model of a fractured bone, and the real-time tracking (pose estimation) of the bone segments intra-operatively. The first component of the research is the innovative 3D reconstruction technique proposed for pre-operative planning on procedures involving the femur, tibia and iliac. Pre-operative planning plays an essential role in the management of orthopaedic injuries because many of the technical problems that may arise during surgery can be anticipated during this preparatory phase. The novel reconstruction algorithm is based on two conventional orthogonal (in anterior and lateral views) 2D radiographic images and a 3D model of an intact (healthy) bone. This intact model is customised through a non-rigid registration process to the shape of the patient’s bone. The customisation involves a fracture incorporation process that separates the bone into the proximal and distal segments and
identifies the pose (position and orientation) of each fragment. This generally applicable framework is a significant contribution over current literature which is hindered by proprietary models that limit usage, are only available for small regions of the bone and have time consuming feature matching requirements. Furthermore, tests conducted involving cadaveric bone models conveyed a millimetre level accuracy in reconstruction, which is superior to comparable literature.

The second component of the research is the intra-operative pose estimation conducted for cases involving bone segment motion tracking (e.g. femur shaft fracture reduction procedure). Here the pre-operatively reconstructed 3D model will be utilised intra-operatively in the 2D-3D registration process for real-time pose estimation. This novel intra-operative registration is performed solely utilising bony anatomical features extracted from fluoroscopic images. This contribution is an enhancement over the current literature which is a proponent of utilising invasive external fiduciary markers. Experiments conducted through phantom studies and cadaveric fractured bones identified a millimetre level accuracy in translation and less than a two degree level accuracy in rotation.
This thesis is dedicated to my parents, Hettiarachchie and Kumudini Gamage, and my brother, Nikini Gamage.

Thank you for always being there for me, without you none of this would have even been possible.
Acknowledgements

This research work would not have been accomplished without tremendous help and support I gratefully received from many individuals.

My supervisor, Professor Shane Xie, deserves special thanks. Our discussions never failed to provide me with encouragement and new direction. You provided me with all the resources and advice necessary to continue with this work. Also, thank you for offering me the freedom to explore my own path and ideas.

I would also like to thank Dr. Patrice Delmas, my co-supervisor. You have provided invaluable input into this work over the years and every time I presented my work to you resulted in new ideas and paths. I have always enjoyed the humour and mentoring you have presented. I am very much courteous to your scrupulous reading of my manuscripts that are doubtlessly time-consuming.

Professor Peter Xu, my co-supervisor, who has provided me with important input and resources for my work. I thank you for your generous and wise guidance during its development.

Further thanks go to my good friends and lab companions, Yun Ho Tsoi and Prashant Jamwal. Yun Ho, I’ve learnt a great deal through you over the years, and you deserve special thanks for always being there to help me with my work.

Foremost, without my parents, and brother by my side through some of the busiest and hardest times of this journey, accomplishment wouldn’t have been possible.

Thank you, everyone
# Table of Contents

Chapter 1 Introduction ........................................................................................................... 1  
 1.1 Medical Imaging ........................................................................................................... 2  
 1.2 Image Guided Surgery: Current Position ................................................................. 4  
 1.3 Importance of 3D Image Guided Surgery: A Case Study ........................................ 8  
 1.4 Principals of Surgical Treatment of Fractures ......................................................... 11  
 1.5 Research Objectives ................................................................................................... 12  
 1.6 Thesis Outline ............................................................................................................ 13  
 1.7 Chapter Summary ....................................................................................................... 15  

Chapter 2 Literature Review ................................................................................................ 17  
 2.1 Neurosurgical Procedures ......................................................................................... 17  
 2.2 Spinal Procedures ..................................................................................................... 20  
 2.3 Orthopaedic Procedures ........................................................................................... 21  
   2.3.1 Pre-Operative Planning ......................................................................................... 23  
   2.3.2 Intra-Operative Pose Estimation ......................................................................... 28  
 2.4 Discussion .................................................................................................................. 33  
 2.5 Chapter Summary ....................................................................................................... 38  

Chapter 3 Pre-Operative Planning: Intact Bone Applications .............................................. 39  
 3.1 Introduction ................................................................................................................ 39  
 3.2 Anatomical Generic Model Preparation ..................................................................... 42  
 3.3 Need for a Generic Model ......................................................................................... 45  
 3.4 Model Initialisation ..................................................................................................... 47  
 3.5 Contour Extraction ..................................................................................................... 47  
 3.6 Shape Customisation ................................................................................................ 54  
   3.6.1 Similarity Measure ............................................................................................... 55  
   3.6.2 Similarity Measure Feature Weight Selection ..................................................... 61  
 3.7 Full 3D Deformation .................................................................................................. 62  
 3.8 Experiments ................................................................................................................ 65  
   3.8.1 Hip and Knee Arthroplasties ................................................................................ 65  
   3.8.2 Femur Osteotomy ................................................................................................. 69  
   3.8.3 Discussion ............................................................................................................ 72  
 3.9 Extension: Inclusion of Interior Edges ....................................................................... 73  
 3.10 Chapter Summary ..................................................................................................... 77
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Pre-Operative Planning: Fracture Reconstruction</td>
<td>79-83</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>79</td>
</tr>
<tr>
<td>4.2</td>
<td>Contour Extraction</td>
<td>83</td>
</tr>
<tr>
<td>4.3</td>
<td>Fracture Modelling</td>
<td>84</td>
</tr>
<tr>
<td>4.4</td>
<td>Pose Estimation</td>
<td>86</td>
</tr>
<tr>
<td>4.5</td>
<td>Shape Customisation</td>
<td>88</td>
</tr>
<tr>
<td>4.6</td>
<td>Experiments</td>
<td>89</td>
</tr>
<tr>
<td>4.7</td>
<td>Chapter Summary</td>
<td>93</td>
</tr>
<tr>
<td>5</td>
<td>Segmentation of Radiographic Images under Topological Constraints</td>
<td>95-108</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>95</td>
</tr>
<tr>
<td>5.2</td>
<td>Edge Point Extraction</td>
<td>99</td>
</tr>
<tr>
<td>5.3</td>
<td>Registration</td>
<td>103</td>
</tr>
<tr>
<td>5.4</td>
<td>Contour Evolution</td>
<td>104</td>
</tr>
<tr>
<td>5.5</td>
<td>Experiments</td>
<td>105</td>
</tr>
<tr>
<td>5.6</td>
<td>Chapter Summary</td>
<td>108</td>
</tr>
<tr>
<td>6</td>
<td>Intra-Operative Pose Estimation</td>
<td>109-135</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>109</td>
</tr>
<tr>
<td>6.2</td>
<td>Potential Medical Applications</td>
<td>112</td>
</tr>
<tr>
<td>6.3</td>
<td>Problem Statement</td>
<td>114</td>
</tr>
<tr>
<td>6.4</td>
<td>Technical Issues</td>
<td>117</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Transformation Representation</td>
<td>117</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Number of Views</td>
<td>119</td>
</tr>
<tr>
<td>6.4.3</td>
<td>Nature of User Interaction</td>
<td>123</td>
</tr>
<tr>
<td>6.4.4</td>
<td>Parameterisation of Patient Pose</td>
<td>125</td>
</tr>
<tr>
<td>6.5</td>
<td>Acquisition Geometry</td>
<td>128</td>
</tr>
<tr>
<td>6.6</td>
<td>Extraction of Anatomical Surface Model</td>
<td>131</td>
</tr>
<tr>
<td>6.7</td>
<td>Pose Estimation System Overview</td>
<td>134</td>
</tr>
<tr>
<td>6.8</td>
<td>Chapter Summary</td>
<td>135</td>
</tr>
<tr>
<td>7</td>
<td>2D-3D Registration: Frontal and Lateral Alignment</td>
<td>137-143</td>
</tr>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>137</td>
</tr>
<tr>
<td>7.2</td>
<td>Fluoroscopic Image Processing</td>
<td>139</td>
</tr>
<tr>
<td>7.2.1</td>
<td>ROI Identification</td>
<td>139</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Feature Extraction</td>
<td>140</td>
</tr>
<tr>
<td>7.2.3</td>
<td>Shaft Line Pairing</td>
<td>141</td>
</tr>
<tr>
<td>7.3</td>
<td>Initialisation of the Cylindrical Segments</td>
<td>143</td>
</tr>
<tr>
<td>7.4</td>
<td>Frontal and Lateral Pose Estimation</td>
<td>143</td>
</tr>
</tbody>
</table>
# References

## Chapter 9 Conclusions and Future Works

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1 Pre-Operative Planning</td>
<td>181</td>
</tr>
<tr>
<td>9.1.1 Reconstruction Accuracy</td>
<td>183</td>
</tr>
<tr>
<td>9.1.2 Adaptability to Other Cases</td>
<td>184</td>
</tr>
<tr>
<td>9.1.3 Manual Intervention Required</td>
<td>184</td>
</tr>
<tr>
<td>9.2 Intra-Operative Pose Estimation</td>
<td>185</td>
</tr>
<tr>
<td>9.2.1 Feature Space</td>
<td>186</td>
</tr>
<tr>
<td>9.2.2 Pose Estimation Accuracy</td>
<td>187</td>
</tr>
<tr>
<td>9.3 Future Research Ideas</td>
<td>188</td>
</tr>
<tr>
<td>9.3.1 Coupling Segmentation and Registration</td>
<td>188</td>
</tr>
<tr>
<td>9.3.2 View of Fluoroscopic Acquisitions</td>
<td>190</td>
</tr>
<tr>
<td>9.3.3 Defining Automatic Stopping Criterion for Gradient Optimisation Protocols</td>
<td>190</td>
</tr>
<tr>
<td>9.3.4 Truncation/Limited Field of View</td>
<td>190</td>
</tr>
<tr>
<td>9.3.5 Distortion Effects and Dewarping</td>
<td>191</td>
</tr>
<tr>
<td>9.3.6 Code Optimisation</td>
<td>191</td>
</tr>
<tr>
<td>9.3.7 Clinical Trials</td>
<td>191</td>
</tr>
</tbody>
</table>

## Chapter 8 2D-3D Registration: Axial Alignment

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1 Introduction</td>
<td>152</td>
</tr>
<tr>
<td>8.2 Fluoroscopic Edge Extraction</td>
<td>153</td>
</tr>
<tr>
<td>8.3 Optimisation Procedure</td>
<td>153</td>
</tr>
<tr>
<td>8.4 Optimisation Formulation</td>
<td>155</td>
</tr>
<tr>
<td>8.5 Objective Functions</td>
<td>157</td>
</tr>
<tr>
<td>8.5.1 Projection Strategy</td>
<td>158</td>
</tr>
<tr>
<td>8.5.2 Back-Projection Strategy</td>
<td>162</td>
</tr>
<tr>
<td>8.5.3 Similarity Measure Evaluation: Protocol and Experimentation</td>
<td>165</td>
</tr>
<tr>
<td>8.6 Gradient Search Based Optimisation</td>
<td>169</td>
</tr>
<tr>
<td>8.7 In-Plane and Out-of-Plane Registration</td>
<td>174</td>
</tr>
<tr>
<td>8.8 Multiresolution Approach</td>
<td>174</td>
</tr>
<tr>
<td>8.9 Proof of Concept Experimentation</td>
<td>175</td>
</tr>
<tr>
<td>8.10 Experimentation</td>
<td>177</td>
</tr>
<tr>
<td>8.11 Chapter Summary</td>
<td>180</td>
</tr>
</tbody>
</table>

## Chapter 7 Summary

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.4.1 Experiments</td>
<td>146</td>
</tr>
<tr>
<td>7.5 Chapter Summary</td>
<td>149</td>
</tr>
</tbody>
</table>

## Chapter 8 Summary

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.11 Chapter Summary</td>
<td>180</td>
</tr>
</tbody>
</table>

## Chapter 9 Summary

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1 Pre-Operative Planning</td>
<td>181</td>
</tr>
<tr>
<td>9.1.1 Reconstruction Accuracy</td>
<td>183</td>
</tr>
<tr>
<td>9.1.2 Adaptability to Other Cases</td>
<td>184</td>
</tr>
<tr>
<td>9.1.3 Manual Intervention Required</td>
<td>184</td>
</tr>
<tr>
<td>9.2 Intra-Operative Pose Estimation</td>
<td>185</td>
</tr>
<tr>
<td>9.2.1 Feature Space</td>
<td>186</td>
</tr>
<tr>
<td>9.2.2 Pose Estimation Accuracy</td>
<td>187</td>
</tr>
<tr>
<td>9.3 Future Research Ideas</td>
<td>188</td>
</tr>
<tr>
<td>9.3.1 Coupling Segmentation and Registration</td>
<td>188</td>
</tr>
<tr>
<td>9.3.2 View of Fluoroscopic Acquisitions</td>
<td>190</td>
</tr>
<tr>
<td>9.3.3 Defining Automatic Stopping Criterion for Gradient Optimisation Protocols</td>
<td>190</td>
</tr>
<tr>
<td>9.3.4 Truncation/Limited Field of View</td>
<td>190</td>
</tr>
<tr>
<td>9.3.5 Distortion Effects and Dewarping</td>
<td>191</td>
</tr>
<tr>
<td>9.3.6 Code Optimisation</td>
<td>191</td>
</tr>
<tr>
<td>9.3.7 Clinical Trials</td>
<td>191</td>
</tr>
<tr>
<td>References</td>
<td>193</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1.1 The process of closed intramedullary nailing, reproduced from [5] .......................... 9
Figure 1.2 Prototype fracture reduction setup, reproduced from [6] ................................. 10
Figure 1.3 Fracture treatment procedure workflow from pre-operative to intra-operative ......... 12
Figure 2.1 Illustration of the research context of the thesis in relation to other work conducted in the area of orthopaedic IGS .......................................................... 22
Figure 2.2 Classification of rigid 2D-3D registration methodologies ................................. 29
Figure 2.3 Proximal and distal bone fragments with bone screws and active tracking devices attached to them, reproduced from [4] ...................................................... 30
Figure 2.4 Overview of proposed reduction process ......................................................... 38
Figure 3.1 The main process steps of the proposed intact bone reconstruction framework .......... 42
Figure 3.2 Femoral generic model and projections. The blue contour is the lateral projection while the red contour is the frontal projection ..................................................... 44
Figure 3.3 Illustrations to highlight the need of an anatomical generic model ...................... 46
Figure 3.4 Cross-sectional bony structure of a femoral head and the associated x-ray image .... 48
Figure 3.5 The segmentation process performed on a typical x-ray image that inhabits the problems seen in many x-ray images ......................................................... 51
Figure 3.6 Edge extraction on an example x-ray of a femoral head. The user defines lines to perform the segmentation. Segmentation time: 20s ................................................. 54
Figure 3.7 The main process steps of the non rigid registration process ............................ 55
Figure 3.8 Factors that the developed point matching algorithm must be able to handle to perform successful medical image processing ............................................................ 56
Figure 3.9 Illustrations of the shape context based measure ............................................ 58
Figure 3.10 Illustrations of the shape customisation stage ............................................. 61
Figure 3.11 Results of the testing performed on phantom data to identify the effectiveness of each feature measurement (Shape, Curvature, Edge Orientation and Euclidean Distance) against typical problems faced in radiographic image processing (Noise, Outliers, Deformations and Occlusions) .................................................. 62
Figure 3.12 Illustration of the FFD lattice warping to deform the model femoral bone (b) to fit the target bone (c). In (a) the red femur is the model while the blue femur is the target .......... 65
Figure 3.13 Several examples of the reconstruction framework being applied to long bones. The blue model represents the target model (patient-specific model) while the red model was customised to fit that target utilising only the anterior and lateral projections .................................. 67
Figure 3.14 Reconstruction framework being applied to an Iliac. This is an example of an irregular bone with little cross-sectional shape similarity ........................................ 68
Figure 3.15 Three examples of reconstruction error minimisation through successive iterations ...... 68
Figure 3.16 Comparisons of before and after results of the reconstruction procedure. The model (patient’s) profiles obtained from CT is given in blue while the as result of reconstruction procedure is given in red. This example is for the Left 1/ Left 3 model customisation test .......... 69
Figure 3.17 Examples of the deformed bones used in the reconstruction testing .................. 69
Figure 3.18 Several examples of the reconstruction framework being applied to deformed femora. The blue model represents the target model (patient-specific model) while the red model was customised to fit that target utilising only the anterior and lateral projections ........................................ 71
Figure 3.19 Example of the clearly identifiable edges on femur radiographic images ........................................ 74
Figure 3.20 Non-rigid registration on the sub-sampled edge contour data points for the anterior/lateral views and the associated translational field displayed on the generic model ........................................ 74
Figure 3.21 Comparison of the accuracy of the reconstruction technique that utilise only the outer contour (a) and the technique that utilise both the inner and outer contours (b). The colour bar indicates the absolute error in mm ................................................................. 76
Figure 4.1 Proposed femur fracture reduction pre-operative planning framework ........................................ 82
Figure 4.2 Two examples of the interactive x-ray segmentation process. From left to right: The user defines lines to perform the segmentation. Time taken for segmentation: (a) 20s, (b) 25s ............. 83
Figure 4.3 An example of the Active Contours based x-ray segmentation process ........................................ 84
Figure 4.4 Illustrations of the fracture surface modelling stage ................................................................. 86
Figure 4.5 Illustrations of the pose estimation stage ....................................................................................... 88
Figure 4.6 Illustrations of the shape customisation stage .............................................................................. 89
Figure 4.7 Three qualitative illustrations of testing performed. The left image of each pair shows one of the two x-ray images (DRRs) used in the fractured femur reconstruction. The image on the right shows the reconstructed fracture region after fracture modelling, pose estimation and shape customisation ............................................................................... 90
Figure 4.8 The potential of 3D models to be used in surgical simulation and implant selection. The images show example implants inserted into the bone fragments after alignment. Please note that these images are not intended to represent exact femoral nailing practices ......................................................... 93
Figure 5.1 The main process steps of the proposed segmentation framework ........................................ 99
Figure 5.2 Segmentation process applied to a femur with a shaft fracture (this was artificially fractured for illustrative purposes and is not meant to represent realistic x-ray images) ........... 99
Figure 5.3 The main process steps of the adaptive filtering applied ......................................................... 101
Figure 5.4 Comparison between Sobel, Canny and the proposed edge point extraction algorithms. 103
Figure 5.5 Example of the outlier removal process and regularised planar interpolation ............................ 104
Figure 5.6 Box plots of the benchmarking performed. The indices used in the figure are: 1-Proposed Framework, 2-CPD, 3- RPM, 4-ICP. The outlier tests had a range of outlier to point ratios (ranging from 1 to 2 times) ................................................................................. 106
Figure 5.7 Examples of segmentation results using the proposed framework. The blue point set is the projected contour which is non-rigidly registered with the red target edge points extracted from the x-ray/fluoroscopic image. The last image shows the final segmentation result ........................................ 108
Figure 6.1 The transformation parameter $\text{model}^\text{target}$ which relates the coordinate frames of the imaging environment and the data volume .................................................................................. 116
Figure 6.2 Geometrical setup of the registration of a 3D image to 2D x-ray projection images .................. 119
Figure 6.3 2D-3D registration based on projection strategy ................................................................. 120
Figure 6.4 2D-3D registration based on back-projection strategy ........................................................... 121
Figure 6.5 2D-3D registration based on reconstruction strategy .............................................................. 122
Figure 6.6 An illustration of a fractured 3D femur placed in context of two orthogonal fluoroscopy images. The projections of the 3D model are displayed in the imaging plane ........................................ 123
Figure 6.7 The 6-element parameterisation of patient pose comprises three consecutive rotations around the three coordinate axes, followed by a 3D translation ........................................ 125
List of Tables

Table 1.1 Comparison of radiograph based imaging methodologies .................................................. 6
Table 2.1 Pros and cons of intact bone customisation/reconstruction methodologies ....................... 35
Table 2.2 Pros and cons of bone pose estimation methodologies ....................................................... 37
Table 3.1 Intact femur reconstruction results ....................................................................................... 66
Table 3.2 Intact tibia reconstruction results .......................................................................................... 66
Table 3.3 Intact iliac reconstruction results .......................................................................................... 66
Table 3.4 Osteotomy surgical planning 3D femur model reconstruction results ................................... 70
Table 3.5 Patient-specific intact bone model reconstruction results in terms of mean and standard deviation of error ........................................................................................................ 76
Table 4.1 Patient-specific fractured bone model reconstruction results in terms of mean and standard deviation of error ........................................................................................................ 91
Table 5.1 Diagnostic radiograph segmentation testing results ................................................................. 107
Table 7.1 Patient-specific fractured bone frontal and lateral pose estimation results in terms of mean and standard deviation of error .................................................................................. 149
Table 8.1 Comparison of evaluation indices for similarity measures SM1 and SM2 ......................... 168
Table 8.2 Optimisation based 2D-3D registration results .................................................................... 176
Table 8.3 Errors between the recovered and actual rotations and translations. Rotational errors provided in degrees while the translational errors are provided in mm terms ........................................ 178
Table 8.4 Errors between the recovered and actual rotations and translations with modification to the pose initialisation. Rotational errors provided in degrees while the translational errors are provided in mm terms ........................................ 179
Chapter 1 Introduction

Surgical interventional techniques have radically changed over the past decade with the adoption of minimally invasive techniques [1, 2]. The aim of such techniques is to reduce the overall operating time, lessen the damage to surrounding tissue, condense both the post-operative pain and recovery time of the patient, and to be highly accurate [2]. These are achieved through the reduction of the physical stress applied to the human body by minimising the size and the number of incisions. Unfortunately, in comparison to open procedures, these approaches restrict the surgeon's view of the anatomy. This leads to an increasing need for advanced imaging techniques that would help the surgeons not only with diagnosis, but also with planning and guiding the procedure.

Minimally invasive interventions are being driven by these advanced medical imaging techniques used prior to and during surgical procedures. Typically referred to as Image Guided Surgery (IGS), these procedures provide 3D anatomical visualisation and track surgical alterations (and surgical instruments) in conjunction with pre-operative and intra-operative images. IGS aids in guiding the procedure to improve surgical manipulation and reduces mental strain.

IGS systems were initially used primarily in neurological and spinal applications [3]. They were recently adopted by the orthopaedic speciality, which has become the fastest growing market segment due to a population which is aging and an increasing availability of specifically designed imaging tools (e.g. intra-operative c-arm technologies) [3]. Orthopaedics is particularly well suited for IGS because bones and some soft tissue can be evaluated easily and accurately using diagnostic technologies such as radiography, fluoroscopy, computed tomography (CT), and magnetic resonance imaging (MRI). Subsequently, bony and soft tissue structures can be reconstructed to create 3D images that
can be used in pre-operative surgical planning or intra-operative guidance, tracking the effect of various surgical actions. Also, bone is a rigid structure that does not deform significantly when drilled or cut, thus rigid body mathematics can be applied to bony anatomy in contrast to elastic tissues and soft anatomical organs. Furthermore, with the recent advent of robotic fracture reduction devices there is a trend towards real-time pre and intra-operative 3D IGS [2, 4-6]. The development of 3D IGS technologies based on intra-operative radiographic imaging (c-arm fluoroscopy) permits pose estimation for robotic guidance [7-9].

This chapter briefly introduces the background of medical imaging modalities, discusses the current position in image guided surgical systems and identifies the problems associated with several cases in the orthopaedic speciality. The chapter concludes by establishing the research objectives and the motivation for conducting this research by giving consideration to all stakeholders (radiographers, radiologists and surgeons) involved in the surgical application of interest.

1.1 Medical Imaging

The history of medical imaging extends back more than 700 years to the 1300s when anatomy dissection theatres were common practice [10]. In such theatres the patients were surgically cut for visualisation of internal organs and structures. This was direct visualisation with an extremely high level of invasiveness.

Modern medical imaging began with the discovery of x-ray imaging by a German physicist, Wilhelm Conrad Röntgen, in 1895. An important contribution was made in 1917 by a mathematician named Radon who developed the mathematics for CT (nevertheless CT technology was not developed till much later). The Radon Transform is used in virtually all modern tomographic scanning systems. The use of radiography accelerated throughout the early 20th century, and by the 1940s real-time x-ray imaging (termed fluoroscopy) was developed and used. In the 1950s, nuclear medicine imaging techniques were developed. The
greatest advance in modern medical imaging came in the early 1970s with the advent of the CT scanner which began the era of digital imaging. These images were based on computational methods applied to numeric recordings of x-ray absorption from many angles of view through the body, using formulas based on Radon’s inversion mathematics. In the past three decades, dynamic spatial reconstruction capabilities have been developed and demonstrated, and currently CT, MRI, PET (Positron Emission Tomography), SPECT (Single Photon Emission Computed Tomography) and other scanning modalities provide high resolution, fast volume imaging capabilities. For a detailed history and timeline of medical imaging technologies, please refer to [10].

It is generally established that medical imaging techniques can improve diagnosis and treatment planning, as well as help the surgeons to control, conduct and evaluate a treatment. Image supported medical interventions can reduce risks, improve quality, and decrease invasiveness of current interventions. Each medical imaging technique has the ability to represent a specific physical property or modality, and acquire real world information about the patient and create a corresponding image.

On the basis of the information provided, medical imaging techniques can be divided into anatomical and functional techniques. Anatomical imaging provides information on a patient’s anatomy, while functional imaging provides information on its functionality.

On the basis of the number of image dimensions, imaging techniques can be split into spatial (2D/3D), spatial-temporal (4D), spatial-temporal-functional (5D) and hyper space (ND). Two-dimensional images are projections of 3D objects onto a 2D plane recording media (x-ray sheets). Three-dimensional imaging systems obtain true spatial 3D anatomy in the x, y and z dimensions and even 4D images when 3D anatomy is imaged rapidly over time. The current highest practical dimensionality is 5D, where the three spatial dimensions plus time are supplemented by synchronous functional information. Hyper space imaging, such as a vector image or a tensor image has multiple functional parameters all relating to the same object being imaged through time. These functions may not always be entirely independent of
Chapter 1 - Introduction

one another, but if they exhibit sufficient orthogonality they can be characterised as n-dimensional. An example would be a 3D heart beating through time with several measures of function at each myocardial muscle point. Such functional measurements could be electrical activity, pressure, temperature, elasticity or diffusion.

1.2 Image Guided Surgery: Current Position

As previously introduced, IGS refers to any surgical procedure involving the pre and intra-operative use of a representation of patient anatomy obtained via imaging techniques or computerised methods [10]. IGS systems enable the real-time visualisation of targeted tissues and structures being manipulated through surgery and also facilitate surgeons to undertake detailed planning of the surgery. It results in less invasive procedures through smaller physical insertions which ultimately leads to safer surgery with higher accuracies, decreased operating times and faster patient recoveries.

Two commercially available technologies that provide direct 3D imaging are CT and MRI. The equipment that performs the above imaging is commercially available pre-operatively, however intra-operatively, they are either not available or are only available in a limited scale (as will be highlighted in the next few paragraphs). Furthermore, the pre-operative use of CT and MRI imaging is restricted to a minority of complex procedures. This is mainly due to restriction placed by costs, availability and risks posed by unwarranted detailed imaging (such as radiation risk in the case of CT scanning) (Table 1.1). Conversely, intra-operative availability of direct 3D imaging is restricted due to mainly technological and imaging constraints [2, 11, 12].

Intra-operative MRI (IMRI) developed primarily for neurological surgery, utilises a very low magnetic field strength (0.12 Tesla) to ensure minimal interference with ferromagnetic surgical equipment [13]. Image quality is hence limited. IMRI is presently used for certain neurosurgical procedures, however several reviews conducted have highlighted their
inadequacies for routine intra-operative use [11]. Moreover, major surgical equipment modifications are also required to ensure the magnetic field does not interfere with the equipment and vice-versa to distort the image [2].

Intra-operative CT and motorised c-arms are another direct 3D imaging modality that has been mentioned in literature. The main concern with these devices is the radiation exposure involved with multiple usages, not only to the patient but also to the surgical team. The fluoroscopic c-arm produces lesser quality images due to the lower radiation dosage levels utilised while intra-operative CT creates very accurate 3D imaging with a consequential high radiation exposure as well as a longer 3D reconstruction time (Table 1.1). Several studies have identified that the use of such high detail is not required for therapeutic intra-operative purposes and is only required pre-operatively for diagnostic and planning [14]. As a reference, the measured radiation levels during the tests conducted by [14] for the motorised c-arm was 41.2 mGy (milligray, SI units of absorbed dose) and 389 mGy for the intra-operative CT scan, which highlights the vastly increased level of radiation exposure of the CT scan. However, a motorised c-arm’s field of view is limited (typically 12cm×12cm×12cm [14]) in comparison to CT scan. Another operational consideration that is inherent with the use of a motorised c-arm is the necessity for the operative site to be in an iso-centric position to the c-arm such that the relevant anatomy is properly imaged. Tests conducted by [12] have indicated that this intra-operative positioning and set up of a c-arm system required 187 seconds on average, the scanning procedure takes a further 120 seconds in all cases and the final analysis and reconstruction takes on average 247 seconds.

Furthermore, intra-operative 3D (and 4D) ultrasound imaging is also a possibility, achieved through expanding conventional 2D devices by tracking the position of the transducer during imaging with the aid of an electromagnetic marker attached to the transducer. Ultrasound images can thus be combined to create volumes, which cover an entire organ rather than just an individual layer [15]. However they suffer from low image quality inherent in 2D ultrasound and in many cases will only provide a partial 3D volume of the
anatomy of interest [16]. The ability of the ultrasound imaging to produce 4D images (spatial-temporal: 3D spatial and 1D temporal) is a distinctive feature over CT and MRI imaging which are limited to static 3D images. This ability ensures that ultrasound imaging is a highly researched area for imaging dynamic objects such as the heart [16].

Table 1.1 Comparison of radiograph based imaging methodologies.

<table>
<thead>
<tr>
<th></th>
<th>CT</th>
<th>Motorised C-arm</th>
<th>X-Ray</th>
<th>C-arm Fluoroscopic Images</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Modality</strong></td>
<td>3D</td>
<td>3D</td>
<td>2D</td>
<td>2D</td>
</tr>
<tr>
<td><strong>Time Taken for Imaging (Minutes)</strong></td>
<td>1-2</td>
<td>1-2</td>
<td>&lt;0.5</td>
<td>Real-Time</td>
</tr>
<tr>
<td><strong>Costs (NZ$)</strong></td>
<td>$800</td>
<td>$300</td>
<td>$100</td>
<td>$30</td>
</tr>
<tr>
<td><strong>Radiation Absorption Dosage (mGy)</strong></td>
<td>350</td>
<td>40</td>
<td>4</td>
<td>1.5</td>
</tr>
</tbody>
</table>

1 Time taken was confirmed after consultation with radiologists at Auckland Radiology and Specialist Radiology Group. The time figure does not include setup time of equipment.

2 The costs involved are those imposed by the private radiology service provider, Auckland Radiology.

3 Radiation exposure measurements were consolidated through [12, 14].

The next level of imaging performed intra-operatively is indirect 3D imaging where highly detailed 3D pre-operative information about the examined anatomy is fused with the more up-to-date, less detailed, intra-operative images (either 3D or 2D) or fiducial markers. There have been several projects attempting to merge intra-operatively acquired 2D and 3D ultrasound with pre-operative MRI to create higher quality 3D objects during neurosurgical interventions [17, 18]. Similarly, several projects have also registered intra-operative 2D fluoroscopic images with pre-operatively acquired CT data [19-22], to name a few. Another avenue taken is fiducial marker based tracking, utilising either pre-operatively obtained MRI or CT images and updating the pose of the anatomy of interest with the tracked fiducial markers. The general drawback of such systems is the guidance inaccuracies caused due to movement, deformation and changes in anatomy since the time of initial imaging. There are two main types of fiducial tracking systems: optical and electromagnetic tracking [23]. Optical tracking is widely used at present due its high tracking accuracies. Commercially available optical tracking systems have indicated tracking accuracies in translation of less than 0.3 mm and in rotation of less than 1 degree [23]. There are several
shortfalls of optical tracking, firstly direct line of sight between the marker assembly on the patient and the external sensor assembly has to be maintained at all times. This is a cumbersome requirement in a crowded operating room. Moreover, since the orientation data are derived from measuring several coordinate pairs (at least three LED’s or reflective markers), the markers need to span a certain physical dimension, thus increasing the size of the instrument. Acquiring data from a very small marker assembly results in unreliable orientation data. Electromagnetic tracking systems have found some acceptance as position measurement devices for 3D ultrasound [23]. The main advantage of these systems is the fact that the electromagnetic fields used for position determination penetrate non-conductive materials and therefore do not depend on an unobstructed line of sight. These systems however face distortions and decrease in accuracy due to conductive or ferromagnetic materials [23]. Another benefit of fiduciary based tracking systems is the ability to track rigid surgical instruments such as rigid endoscopes [23].

Thus, as outlined above, there are currently prohibitive restrictions on the use of 3D medical imaging intra-operatively and in some cases pre-operatively. Although most of the imaging techniques can be used for both diagnosis and treatment, different criteria such as limitations or requirements of specific medical procedures, image quality, invasiveness, required speed and costs of imaging, define their specific use. Projection 2D x-ray imaging techniques, such as radiography and fluoroscopy, are usually cheaper and faster in comparison to 3D imaging modalities. These imaging modalities are hence commonly used for both, diagnosis and planning in the pre-operative stage of medical interventions and for control of surgical or radiological interventions in the intra-operative phase. However, 2D imaging lacks spatial information contained in 3D modalities such as CT or MRI. Thus, more and more research work has been focused on creating a fusion between pre-operatively obtained 3D images with the intra-operatively obtained 2D images, as alluded to in the previous paragraph. It is this type of reconstruction that will be the focus of this thesis as outlined in the proceeding sections/chapters.
1.3 Importance of 3D Image Guided Surgery: A Case Study

Section 1.2 highlights several limitations with the current position in 3D medical image guidance. Thus many surgical interventional procedures still rely on 2D images for pre and intra-operative guidance. One such procedure is the intramedullary nailing performed for cases of femur shaft fractures. Femur fractures are commonly caused by high-energy injury mechanisms, like traffic accidents, predominantly in young males or by low-energy mechanisms, like falling, in elderly females [24]. Statistics gathered by [24] indicate an approximate femur fracture incidence rate of 37 per 100,000 persons per year, and thus it is a frequently encountered injury. In 2004, 3200 patients with fractures of the thigh bone have been counted in New Zealand [25, 26]. These figures include fractures in the proximal (hip side) femur as well as the shaft (the middle, diaphyseal) region. Information from the United Kingdom estimates the cost of an individual hip fracture to be approximately $60,000 (excluding costs to the family, travel and loss of income) [27]. Worldwide the total number of fractures to the hip in 1990 is estimated at 1.26 million, reaching 2.6 million by 2025 and 4.5 million in 2050 [28]. This increase can largely be attributed to the greater numbers of elderly as the population ages. Although these numbers are not directly for long bone fractures, it is reasonable to expect similar growth in fracture occurrences.

Intramedullary nailing has been established as a standard technique for a definite stabilising treatment in diaphyseal fractures of the lower extremities. A detailed description of this surgical procedure can be found in [29]. The complete process of intramedullary nailing is shown as a sketch in Figure 1.1. The process starts with the opening of the medullary cavity, and then a small soft tissue cut of about 5 cm has to be placed at the proximal end of the femur. In extension of the femoral shaft, the bone's cavity has to be opened, which is achieved with a surgical drill. Now the intramedullary nail is inserted into the bone's medullary cavity until it reaches the fracture region. Subsequently the two major bone fragments are aligned accordingly to their correct anatomical position.
For femur fracture intramedullary nailing, currently the diagnostics, pre-operative planning and surgical tracking of fracture reduction rely primarily on 2D radiographs typically acquired from the anterior and lateral viewpoints [29]. A recent study conducted on femur fracture reduction has confirmed that computer-aided systems can significantly improve the accuracy of orthopaedic procedures by augmenting the current 2D image guidance with interactive display of 3D bone models [4]. This research indicates that the positioning errors and misalignment that are seen in 18% of femur fracture cases along with the misalignments created, would be reduced with the introduction of an interactive display of 3D bone models into the fracture reduction process [4]. The study directly attributes the currently limited visualisation of structures on the bone surface and the restricted display window as the main problems for achieving a correct rotational alignment. The financial costs associated with such misalignments are high, including the lengthened and costly rehabilitation during which the patient might be inhibited from working and potentially having to perform secondary surgery to correct the initial complications. Studies done by a German group [30] provide further information on the high impact on functional biomechanics and socio-economic effects of femur fracture reduction complications.
Chapter 1 - Introduction

The case of femur shaft fracture surgery has been used throughout this thesis as the application of interest. Other researchers within the Medical and Rehabilitation Research group have extensively worked on surgical technologies behind femur shaft fracture [6].

The prototype device developed by the group (Figure 1.2) consists of a 6 degrees-of-freedom parallel platform mechanism and reduction table. The platform is mounted horizontally and attaches to the manual degrees-of-freedom arm of the reduction table. In this example a foot holster is used to attach the platform to the patient’s leg and perform reduction. Alternatively the robot may be attached directly to the femur with a pin through the femur head. Image guidance and pose estimation work presented in this thesis is vital for trajectory planning and positional feedback for closed loop control of the trajectory.

Figure 1.2 Prototype fracture reduction setup, reproduced from [6].
1.4 Principals of Surgical Treatment of Fractures

The principals behind the treatment of fractures along with the aforementioned current shortcomings are important to understand the need for further image guidance in the orthopaedic speciality. The AO (Arbeitsgemeinschaft für Osteosynthesefragen) foundation for orthopaedic research stress four critical stages of surgical treatment of fractures through its courses [31]. These are pre-operative planning (including fixation selection), anatomical reduction, stable internal fixation and finally, post-operative evaluation and mobilisation of the injured anatomy. The above AO principles [31] and the feedback received through the research’s medical clinical collaborators (at the Auckland City Hospital) have resulted in the formalisation of the following guidelines for fracture treatment,

1) Fracture visualisation: Visualisation and surgical exposure is required to develop a 3D perspective of the fracture configuration. The identification of fracture displacement and the reduction planning is also critical. Radiography or fluoroscopy based imaging permits fracture visualisation without soft-tissue dissection. Commonly, two planar radiographs are acquired in the frontal and lateral planes to provide fracture visualisation.

2) Surgical reduction of the fracture: Surgical reduction is typically achieved by initially aligning the fragments in the frontal and lateral planes followed by the axial alignment. This procedure is guided through c-arm fluoroscopy images.

3) Stability of the fracture: The fracture is temporarily held in place and the alignment of the fragments checked prior to completion. Typically the reduction alignment is checked utilising external anatomical land marks. Surgeons observe the contralateral anatomy to ensure anatomical conformance. Finally a mechanical constraint (rod or plate) is used to fasten the fracture.

4) Post-operative evaluation: Proceeding the stabilisation, the success of the surgery is reviewed by observing the motion of the injured anatomy in comparison to the contralateral, to ensure biomechanical conformance.
Chapter 1 - Introduction

Figure 1.3 Fracture treatment procedure workflow from pre-operative to intra-operative.

Figure 1.3 illustrates the stages involved in an emblematic orthopaedic surgical procedure. The proceeding chapters will often refer back to the aforementioned principals behind the treatment of fractures when the clinical motivation and application behind the work needs to be highlighted.

1.5 Research Objectives

The author’s main goal in conducting this research is to assist orthopaedic surgeons in all steps of femur fracture procedures. Thus the system proposed has to be able to perform pre-operative 3D fractured bone reconstruction, for planning purposes, and pose-estimation / tracking of the bone fragment intra-operatively for guidance. It must also be stated that a secondary goal of the research was to ensure that the work conducted will be applicable to as many orthopaedic applications as possible, on top of the femur fracture case.

The objectives of this research have been grouped according to the surgical use of the proposed IGS systems. Thus the goals are twofold are discussed herein.

1) Develop a pre-operative fractured bone reconstruction algorithm for the purpose of diagnostics and planning the orthopaedic procedure. The reconstruction must be performed with the aid of only 2D x-ray radiographic images. There are no hard real-time constraints placed on this customisation process, however it is important that the customisation is accurate. The pre-operative reconstruction algorithm will also be extended to cater for not only femur fracture cases but also other potential orthopaedic cases. This includes hip
arthroplasty and femur osteotomy. These procedures are also currently planned through 2D images and would benefit greatly from 3D image guidance.

2) Develop an automatic 2D-3D registration algorithm for the purpose of image-guided orthopaedic surgery. The algorithm should be able to obtain the pose of the pre-operatively customised bone of interest, in the physical space of the surgical theatre. The pose should be obtained by performing rigid 2D-3D registration between the pre-operatively customised 3D fractured segments and a set of intra-operative fluoroscopic images, acquired with an interventional c-arm. To be applicable to surgical procedures, the method should be fast, robust, and accurate. An extension to the aforementioned 2D-3D registration algorithm development would also need to be able to handle alternative imaging modalities that are in common use in alternate orthopaedic procedures. For instance in several spinal procedures a compulsory pre-operative CT scan is conducted, prior to the operation. This adaptability to a wide variety of imaging modalities will ensure that the developed algorithms will also be applicable to other orthopaedic procedures.

1.6 Thesis Outline

Following this first introductory chapter, Chapter 2 discusses and reviews the state-of-the-art in existing image guided surgery literature. The shortcomings of existing literature are presented, primarily the lack of research work in fractured bone reconstruction and intra-operative tracking. A background discussion on neurological and spinal IGS literature is followed by a thorough detailing of orthopaedic IGS systems (particularly involving the femoral bone).

Chapter 3 details the novel 3D bone reconstruction framework, which utilises bi-planar 2D images to conduct 3D shape customisation. Two orthopaedic applications are used as cases to test the proposed framework: hip and knee arthroplasty and femur osteotomy. Pre-operative planning is vital for both these cases involving implant fixation and bone incisions.
Chapter 1 - Introduction

In both these applications proper pre-operative planning is vital to reduce complication including dislocations, accelerated wear, and loosening of the implants.

This leads onto Chapter 4, which details the novel pre-operative 3D fractured bone reconstruction framework and algorithms proposed, which is the main application of interest of this thesis. The framework utilises frontal and lateral 2D radiographs to model the fracture surface, separates the bone into the two fractured fragments, identifies the pose of each fragment, and automatically customises the shape of the bone fragments. The use of 3D imaging permits full spatial inspection of the fracture by allowing for special views through manipulation of the interactively reconstructed 3D model, and ultimately better pre-operative planning. Extensive tests conducted are also detailed in both the aforementioned chapters. The proposed software tool allows orthopaedic surgeons to visualise, diagnose, plan and simulate femur shaft fracture reduction procedures in 3D.

Chapter 5 looks at a segmentation technique that was an off product of the reconstruction framework detailed in Chapter 3. It provides detailed analysis of the non rigid registration process that is capable of segmenting images. The chapter details a radiographic image segmentation algorithm which is conducted under topological control. The system is intended for use in common radiological tasks including fracture treatment analysis, osteoarthritis diagnostics, and osteotomy management planning.

Chapter 6 introduces the problem of pose estimation and 2D-3D registration. The chapter presents the motivation behind this work and the potential medical application of this research. The chapter also details the problem statement and the technical considerations behind the pose estimation problem. A high level of detail is provided on the transformation components that need to be recovered as a result of the rigid body registration algorithm.

Chapter 7 details the proposed frontal and lateral registration algorithm. The chapter documents a component of the full registration procedure. The frontal and lateral plane
registration is conducted through a novel analytical feature based technique. This ensures that
the registration in these two planes is conducted in a robust and timely manner.

Chapter 8 details the optimisation based pose estimation along with the objective
functions. The chapter documents the final component of the registration procedure. It
presents an algorithm for recovering the position and orientation of the target anatomy in 3D
space based on an iterative comparison of 2D planar radiographs with the pre-operative 3D
model, through an optimisation process. The system uses x-ray fluoroscopic images acquired
intra-operatively, and iteratively compares them with projection image generated from the 3D
model. Both the aforementioned chapters detail the registration experiments conducted using
both synthetic and real datasets, as well as detailed analysis of their results.

The thesis concludes with Chapter 9, which summarises the project and the contributions.
The conclusions are also presented with several related future research ideas that would
require investigation.

This work has contributed to the Orthopaedic IGS research and medical community and
provided the mechanisms for which further work can be carried out, as will be discussed
throughout the various chapters. Many of the chapters are directly based on published work
by the author of this thesis [7-9, 32-41].

1.7 Chapter Summary

This chapter provided an overview of Image guided surgery and medical imaging
modalities. It highlighted the need for 3D IGS by documenting the problems that arise due to
the lack of guidance during orthopaedic procedures, from diagnostics to surgery. A brief
background to femurs shaft fracture surgery was also presented as the main application of
interest of the thesis. Currently this procedure remains a challenge with many accounts of
malrotation and incorrect union. The chapter also formally documented the research
objectives and finally outlined the thesis structure.
Chapter 2 Literature Review

Since the discovery of radiography, IGS has played a major role in the guidance of surgical procedures. While medical imaging began with analogue x-ray sheets, the advent of the computer and digital imaging has been a major factor in the recent development of this field. IGS is currently undergoing rapid development with a strong emphasis being placed on the use of imaging technology to render surgical and therapeutic procedures less and less invasive and to improve the accuracy with which a given procedure can be performed, compared with conventional methods [1, 2].

This chapter documents the current state-of-the-art research in 3D IGS systems. A brief discussion of neurological and spinal IGS literature is followed by a thorough consideration of orthopaedic IGS systems. Here, particular emphasis is placed on work carried out for IGS systems involving the femoral bone. The chapter highlights shortcomings in the existing work and concludes with an overview of the system developed in this research. The background material in neurological and spinal IGS literature is provided as a precursor to orthopaedic IGS systems, purely to highlight the continuity between the different medical specialities.

2.1 Neurosurgical Procedures

The need for image guidance during neurosurgical operations has always been a concern for surgeons and has evolved through several stages over the last 50 years. Image guidance in neurosurgery is required to plan and guide surgical procedures and instruments down to lesions (brain tumours) in a safe manner through a narrow channel, to monitor the surgical resection, to finally control / evaluate the result.
The introduction of CT and MRI led to neuro-navigation systems for surgical planning [42-51]. Frame-based or frameless neuro-navigational systems use 3D pre-operative data merged with the patient’s anatomy by registration.

Frame-based navigational systems have the advantage of being extremely accurate because a rigid head frame is fixed to the skull. Disadvantages include patient discomfort during frame placement, time taken to calculate the trajectory, and inability to project or image where the biopsy probe is located [42].

Frame-based stereotactic systems were gradually replaced by frameless stereotactic systems [43, 44] as the sensing and computer technology matured. Frameless neuro-navigation systems have proven to be very useful over the last decade, especially in the pre-operative planning phase, and are among the most common systems in use today [43, 44].

Current literature available on non-invasive neurosurgical IGS differ in the way they integrate pre-operative image data with the physical space of the operating room (patient registration), the kind of tracking technology they use to follow the surgical tools that are used (optical, magnetic, ultrasonic or mechanical) and in the way the image information is presented to the surgeon [43, 44]. The main clinical output of intra-operative frameless neuro-navigational systems is to track the movement of surgical instruments in space so their relative position to the lesion can be projected from pre-operative imaging.

The main concern for most current research is the potential sources of registration error. Registration may be inaccurate because of scalp movement, especially if the patient changes from supine to prone position (due to the lack of fiducials). Other possible sources of error are geometrical distortion in the images, movement of the patient with respect to the system during surgery and brain shift (movement of the brain relative to the cranium between the time of scanning and the time of surgery) [45, 46].

Brain shift, the leading concern, is caused after opening of the dura (the outermost of the three layers surrounding the brain) and results in fluid shifts, tumoural volume resection, and
Chapter 2 - Literature Review

fluid leakage. It makes pre-operative data inaccurate. Intra-operative brain surface deformation greater than 10 mm has been documented within one hour of opening the dura prior to actual tumour resection, in over half of the patients studied [45, 46]. The error induced by this type of shift may be even larger in certain cases due to the presence of hydrocephalus or pre-existing loss of parenchymal volume. Due to these sources of error, the usefulness of surgical navigation may diminish during the surgical procedure.

The brain shift problem [46] can only be solved adequately by integrating intra-operative imaging with navigation technology. Literature in this area makes use of intra-operative CT [47] and MRI [48, 49] scanners in order to update the images.

IMRI allows imaging of changes during surgery, accurate navigation, immediate assessment of such complications as haemorrhage, and verification of the planned resection [48, 49]. Computed tomography has a greater tissue resolution in most instances, but it is limited to uniplanar imaging and has the disadvantage of radiation exposure to the patient and surgical team [47]. For these reasons, intra-operative CT has not been widely adopted. With both intra-operative MRI and CT it is possible to obtain close to real-time 2D images defined by the location of the surgical tool used, in addition to update the 3D map in minutes without moving the patient. However, these systems require high investments, high running costs, and a special operating room, as well as surgical equipment.

Another option presented in current literature is intra-operative ultrasound, which is a cheaper and less bulky solution [50, 51]. However ultrasound images have major limitations for intra-operative use in the brain, where images are often difficult to interpret because of echogenic structures, that cannot reliably discern normal from abnormal tissue, and air bubbles and blood products in the surgical field may cause misinterpretation of ultrasound images [50]. One additional drawback with ultrasound compared with intra-operative MRI is the limited volume it covers in the brain, which may cause an orientation problem [51].
2.2 Spinal Procedures

Spinal surgical procedures require intricate guidance due to the complex micro and macro anatomy of the spine, which is not fully visible to the surgeon during the operation. Furthermore, due to the sensibility of the neuro-vascular structures surrounding the spine accuracy and surgical approach has always been a key issue in all types of spinal surgery.

Before the advent of image guidance, spinal navigation was based on the surgeon’s knowledge, experience, and judgment combined with information gathered from 2D radiographs or fluoroscopy. Monitoring of nerve or spinal cord status provided surgeons information regarding the function of these structures (and whether or not a spinal instrument or implant may be placing this function at risk) but is only an indirect indicator of instrument position.

Early spinal surgery guidance procedures were adopted from neuro-navigational systems, where skin surface markers were used. Unfortunately, significant registration inaccuracy due to relative movement between the mobile skin surface and the underlying bony anatomy occurred [52, 53]. These problems with registration inaccuracy were solved with the introduction of intra-operative CT and more importantly 3D fluoroscopy based systems.

Work conducted by [54] first proposed the use of intra-operative CT based spinal image guidance. The scan can be obtained after the spine has been exposed and fiducials have been implanted. This greatly simplifies registration and increases registration accuracy as compared with the use of anatomic fiducials. The use of a CT scanner intra-operatively removes the problem of inter-segmental motion between scan acquisition and operative positioning. Disadvantages include the cost of purchasing a dedicated intra-operative CT system (in addition to an IGS system), the need to use a specially designed operating room table and the difficulty faced with implanting fiducials.

Fluoroscopy based navigation proposed by [55] became popular to mitigate problems associated with intra-operative CT. “Virtual fluoroscopy” was introduced as a novel method
of intra-operative navigation that combines computer-aided surgical technology with typical c-arm fluoroscopy. The major drawback of fluoroscopy is the amount of occupational radiation exposure. However, 3D c-arm fluoroscopy is seen as a significant advancement in the rapidly developing field of image guidance and potentially represents the future of intra-operative spinal navigation [55].

With the aforementioned intra-operative imaging technologies, research expanded the use of image guidance to more complex procedures throughout the entire spine, such as thoracic pedicle screw insertion [56, 57] and C1–C2 transarticular screw placement [58, 59]. Moreover, literature work also emerged proposing technology for reducing the surgical trauma of spinal surgery, resulting in the development of thoracoscopic techniques, retractor systems for lumbar surgery and ventral fusion techniques via dorsal approaches [60, 61]. One major concern for present literature is the accuracy of pedicle screw placement, which is vital as misplacement rates of up to 30% in the lumbar spine and up to 55% in the thoracic spine have been reported [60]. The technical problems of integrating image information, spinal anatomy, and the action of surgical instruments in a computer system operating in a real-time mode were solved by [61], resulting in the first report on the successful clinical application of an image guidance system for pedicle screw placement in the lumbar spine.

2.3 Orthopaedic Procedures

Even though image guided surgery has been commonplace for the past 50 years for other applications, literature on orthopaedic IGS (specifically 3D) systems have only emerged over the past 10-15 years. Moreover, articles with a focus on long bone and femur orthopaedic fracture treatment are currently limited. This section provides detail into the current state of the art research in orthopaedic IGS. It will also highlight the clinical constraints placed on the use of medical imaging modalities and will build on the facts mentioned in the previous chapter (Section 1.2).
Orthopaedic IGS can be broken down into two main segments depending on the clinical use (Figure 2.1). Diagnostics and planning will utilise a static pre-operative image while intra-operative surgical guidance will require real-time tracking of bone movements. Moreover these areas can be split into two further divisions depending on the surgical procedures involved: intact or fractured bone IGS. Bone deformation correction and joint corrections would require the model reconstruction of an intact bone. This type of intact bone reconstruction is done primarily for pre-operative planning and in a number of cases, intra-operatively for pose estimation. Orthopaedic procedures involved with fracture reduction and internal fixture insertion will require the reconstruction of the fractured or broken bones. Similar to the procedures requiring intact bone reconstruction, fractured bone reconstruction will be performed pre-operatively for surgical planning. However the most important feature of fractured bone reconstruction will be the ability to track/register the multiple fracture segments intra-operatively, ideally in real-time. Figure 2.1 summarises these segments of use, branching off with the pre and intra-operative uses. Literature works associated with these segments are presented herein.
Chapter 2 - Literature Review

2.3.1 Pre-Operative Planning

Pre-operative planning for fractured bone reduction procedures can currently be accomplished only through direct 3D imaging modalities (typically CT scan [31]). This approach imposes major limitations as CT scans incur a high health care cost on the patients, and are currently prescribed only for a few complex procedures. Thus in many situations the surgeons will be limited to using 2D x-ray and fluoroscopic images. This shortfall provides a gap where the ability to create 3D fracture models from 2D images will be advantageous.

On the other hand, pre-operatively planning for intact bone procedures can be performed in two distinct ways. Firstly, the use of a direct 3D imaging modality (such as CT), and secondly, a reconstruction of a 3D model based upon information gathered through 2D imaging modalities (x-ray, fluoroscopy or ultrasound).

Reconstruction of 3D models from CT or MR data volumes is trivial and consists of a two-step process. Firstly an image slice segmentation is required for the extraction of the contours of the anatomical structures of interest and secondly surface reconstruction is performed by utilising a surface triangulation technique such as the marching cubes algorithm (based on the extraction of a polygonal mesh of an iso-surface from a 3D scalar field) [62, 63]. CT is the most frequently used pre-operative 3D imaging modality because of its high resolution, high contrast between the bone and its surrounding soft tissues, long scanning range, and short scanning time. Segmentation of bones is easy on CT and images of long skeletal segments such as the entire spine and lower extremities can be obtained in a single scan. Distortion of 3D skeletal models due to motion artefacts of patients can be avoided by using high-speed CT machines. This type of navigation system has been used successfully in pelvic osteotomy [64], total hip and knee arthroplasty [65], and the reconstruction of knee cruciate ligaments [66].

The more complex intact bone reconstruction methodology generates the 3D model of anatomical structures using 2D imaging modalities. These methods can be further divided
Chapter 2 - Literature Review

into two subgroups given the different inputs required. The first group contains methods based purely on the 2D images, while the second group contains methods based on a-priori knowledge of the considered anatomical structure as well as 2D images. Research conducted by [67, 68] illustrates 3D modelling methodologies based purely on information extracted from 2D images where the back-projection of the target object contours previously extracted from the x-ray images is used to calculate the cross-sections for surface reconstruction. A typical assumption is that the bone shape is axially symmetric and the reconstruction of the 3D model of a femur consists in merging a collection of Hermite surface patches based on this symmetry [69].

Furthermore, several studies have utilised a-priori knowledge of the anatomy coupled with the 2D images to reconstruct a 3D model. The main difference here is the availability or the definition of a template model. Its role is to supply the information that is missing or is hard to extract from 2D images alone. This template model is modified according to the information extracted from the patient-specific images.

These studies have utilised four main types of a-priori knowledge (template models) which can be listed as:

**Statistical Models - [19, 70-76]**

Statistical model based reconstruction methodologies are the predominant type of research being carried out in intact bone reconstruction. A statistical model gives an effective parameterisation of the shape variations found in a collection of sample models of a given population. It establishes the legal variations of shape of the anatomy under study. These statistical models are typically referred to in literature as Point Distribution Models (PDM).

The building of a PDM model is a simple task, albeit time consuming. Firstly, it requires the annotation of landmarks of the anatomy to encapsulate the shape of the model. Thus enough corresponding landmarks in all the models of the dataset must be annotated to ensure sufficient representation of the variability of the anatomy. Secondly, Principle Component
Analysis (PCA) is applied to the dataset to compute the eigenvectors and eigenvalues of the training set covariance matrix. Each eigenvector describe a principal mode of variation along the set. The corresponding eigenvalue indicates the importance of this mode in the shape space scattering. A different mode of shape variation \( x \) is generated through the equation, 
\[
    x = \bar{x} + \varphi b,
\]
where \( \bar{x} \) is the mean of the anatomical landmark points, \( \varphi \) is the matrix of eigenvectors and \( b \) is a vector that defines the set of shape parameters of the deformable model.

Once the parametric statistical model has been built it can be customised to suit the patient-specific radiograph. This is performed by identifying the edge contour of the patient’s radiographs and optimising the model’s shape and pose parameters to minimise the distance between the contour and the projected edge of the optimised model. The shape and pose optimisation can be performed separately as done by [72, 73] or together as done by [19].

Due to the high manual input required during the landmark identification stage typically only small specific regions of the bone are statistically modelled. Work carried out by [19] only modelled the proximal segment of the femur while [70] only modelled the distal segment of the femur. Furthermore it is vital to have an accurate edge contour of the patient-specific model that the statistical template will attempt to deform into. A common disadvantage of all these PDM statistical model based reconstruction methods lies in the fact that they require either knowledge about anatomical landmarks [74], which are normally obtained by interactive reconstruction from the input images, or an interactive alignment of the model with the input images [75, 76]. Such a supervised initialisation is not appreciated in a surgical navigation application, largely due to the strict sterilisation requirement. Moreover, a large database of anatomical models is required to create the necessary variability in the statistical model.
**Free Form Deformation (FFD) - [20, 21]**

This reconstruction methodology utilises a generic surface model of the anatomy of interest and performs deformations based on the FFD control lattice movements. The basic idea of FFD is that, instead of deforming the object directly, the object is embedded in a rectangular space that is deformable. There are two approaches to deforming the template shape, firstly using a higher-order deformation function with one control lattice or using a linear deformation function with many control lattices and control points. The work presented by [21] conveys a method where FFD is used to reconfigure the template model of a femur using separate controls for the head, shaft and condylar regions. In [20], a template model of a tibia is reconfigured using a hierarchical FFD. To control the deformation process, the bounding box is subdivided iteratively until the deformed shape of the bone back-projects images similar to the input x-ray images.

The bounding box that is created around the model and the number of control points selected depends on the control the user wishes to have over the deformation and object anatomy. This will further dictate the degrees-of-freedom available. The FFD method is limited by the fact that it does not consider the topology of the object it is deforming. Unless the bounding box subdivides the mesh into individual parts, the transformations will apply to all nearby objects. This tends to unrealistically distort the objects, due to the limited 2D information available in the radiographic images used.

**Morphological Measurements - [77]**

The use of morphological measurements to drive the customisation is of considerable interest, as it has the same effect of providing constraints to the deformation as seen with the case of a statistical model. One such literature of interest is the methodology presented in [77] which utilises a simple linear scaling of the generic model, based on the width of a subject’s femoral condyles as measured on the planar radiographs. The validation given by the authors on the choice of morphological measurement is that the measurement endpoints must be
clearly identifiable on both the radiographs and on the generic 3D model. Thus the condylar width was chosen because the structures used during measurement were easier to identify on radiographs than those used to measure the depth of the condyles. On most radiographs the anterior contour of the condyles could not be identified due to an overlap by the patella.

The main limitation of morphological measurement based customisation is that the deformation is driven by the relative scaling between the patient’s 2D image and the same measurements on the 3D generic model. The work outlined above has utilised a simple linear scaling with a single measurement, but an extension to this where multiple measurements are taken (with non linear scaling) would not be complex. However, this fails to take into consideration the deformation variations that fall outside these key measurements used, which the FFD based methodology (above) and the anatomical database based methodology (below) will cover.

**Anatomical Database Based - [22, 78]**

The concept presented in [22] builds a customised volumetric model of the patient’s tibia based on two orthogonal images and a database of CT scans of the cadaveric tibiae. Using a shape similarity detection process, the 2D orthogonal radiographic images of the bone of interest (patient’s bone) are compared with all the 2D radiographic images stored in the database (cadaver bones). By this procedure, the most similar bone is selected and the corresponding CT data set is retrieved. Further 2D image warping is performed on the slices of the best matched CT data prior to 3D reconstruction. The benefit of such a system is that the generated 3D model can be expressed as a volumetric model or a geometric surface model. However this requires a large database to ensure that a sufficient close match is found to perform the customisation. Furthermore, [78] presents a study where a generic femur was employed and was altered on the basis of bone boundaries visible on patient-specific radiographs. This is similar to the system proposed by [22] but utilises a single generic pre-built 3D model of the femur to perform the customisation.
Chapter 2 - Literature Review

The basic assumption of this type of customisation method is that the femoral cross-sections remain similar among different patients if a fillet is cut perpendicularly to the geometrical axis formed by the shaft axis and the neck axis. Work presented by [78] has confirmed this assumption by quoting authors who analysed sections cut along the femoral diaphysis.

2.3.2 Intra-Operative Pose Estimation

In IGS, pre-operatively, 3D medical data are used to diagnose, plan, simulate, or otherwise assist a surgeon (or possibly a robot) in planning a surgical or therapeutic procedure. The plan is constructed in the coordinate system relative to pre-operative data, while the surgical procedure is performed in the coordinate system relative to the patient. The relationship or spatial transformation between pre-operative data/plan and physical space occupied by the patient during treatment is established by registration (pose estimation or 2D-3D registration) conducted intra-operatively.

Mathematically, registration is a process to determine a geometrical transformation that aligns points in two frames of reference, so that the attributes associated with those points can be viewed and analysed jointly. Clinically, registration is an important step in computer assisted surgical navigation to correlate morphological information collected in different surgical stages, before, during and after the operation.

2D-3D registration measures can be classified into two groups depending on the information utilised to perform the registration (Figure 2.2). On one side is feature based registration, which uses a few selected points or features, and on the other is intensity based registration, which uses the pixel and voxel intensities in both the data sets to perform the registration.
Chapter 2 - Literature Review

Figure 2.2 Classification of rigid 2D-3D registration methodologies.

Feature based registration algorithms match selected geometric features from each data set by finding the transformation that minimises the sum of distances between paired features. Features can be implanted markers, anatomical landmarks, or surface contours.

Feature based registration typically consists of four steps: 1) Feature extraction (choosing the features of interest in each data set); 2) Feature pairing (establishing correspondences between features of each data set); 3) Dissimilarity formulation and outlier removal (quantifying the dissimilarity between paired features); and 4) Dissimilarity reduction (finding the transformation that optimally minimises the dissimilarity).

Feature based algorithms presented in literature can be classified into three categories: fiducial marker based, point based and surface based.

Fiducial marker based techniques [4, 5] involve tracking the bone fragments in real-time by means of an optical tracking system via a set of infrared markers (Figure 2.3). Thus to track the position of the distal and proximal bone fragments, the surgeon must implant fiduciary markers into the bone segments. This is currently pursued by groups working in the area of robotically assisted fracture reduction [4-6]. They utilise pre-operative CT scans of the fractured bones and update the pose of the bone through the tracked external markers. As previously mentioned however, this technique for merging of pre-operatively acquired 3D images with fiduciary markers tracked intra-operatively has inherent weaknesses. The main drawback of such systems is the guidance inaccuracies caused due to movement, deformation and changes in anatomy since the time of initial imaging. Furthermore optical tracking requires a direct line of sight between the LED assembly on the patient and the external
sensor assembly to be maintained at all times. This is a cumbersome requirement in a crowded operating theatre. The implantation of such fiduciary markers would involve several invasive insertions. Moreover, the marker size and the implantation requirements would also burden the surgical process. This is currently the only methodology presented in literature to tackle the problem of the intra-operative pose estimation (tracking) of the fracture segments. The other techniques have only been applied to intact bone procedures.

![Image of bone fragments with markers]

Figure 2.3 Proximal and distal bone fragments with bone screws and active tracking devices attached to them, reproduced from [4].

In point based techniques, features used for registration are corresponding landmark points found on the 2D image and 3D volume (typically 5–10 points in each [79]). Landmark points can be defined manually or by some automatic method, and can be anatomically descriptive (lying on anatomical structures that can be identified on both images), or can represent centres of fiducial markers. Registration is concerned with minimising the distance between the two point sets after the 3D landmark points are projected onto the 2D plane. Extraction of anatomical descriptive points in 2D and 3D images can be a difficult task and is typically achieved through a skilled human operator or automated by an intensity weighting method of locating centres of markers [80]. However, a rough manual determination of corresponding anatomical points, in both 2D and 3D images, can provide a good registration estimate. This estimate can be further used as a starting position for other automatic registration algorithms.

In surface based techniques, features used in this registration approach are surfaces of an object of interest (extracted from a 3D model), and contours that outline the same object on one or more 2D images [80-83]. Registration is concerned with finding the object pose that
minimises the distance between surface and the contour. One way to achieve this minimum is by extending 3D projection lines from points on the contour in the 2D image to a point representing the x-ray source. Then the distance to be minimised can be represented as the Euclidean distance between a projection line and the nearest point on the 3D surface [82]. Another methodology is to project 3D surface points onto the 2D plane and then attempt to minimise the distance between the contour of the projected image and the initial 2D contour. The work done by [80, 82, 83] have presented another methodology, where projections of the volumetric data gradients are computed and compared with x-ray image gradients. The volumetric data pose is adjusted to minimise this gradient difference.

The key characteristic of feature based methods is that they use a small fraction of the image data, usually fiducial centres and anatomy surface points. Thus the registration is performed quicker than in intensity based methodologies. Feature based registration however requires segmentation as well as an efficient feature pairing scheme (with outlier removal). However, accurate and automatic segmentation of bone structures in fluoroscopic images is technically challenging. In practice, robustness is achieved by first performing coarse registration with landmarks followed by fine registration with surfaces. Feature based registration between fluoroscopic images and CT has not yet reached the market, most likely due to the challenges posed by robust segmentation of fluoroscopic images.

On the other side of the 2D-3D registration continuum lies intensity based registration between 2D radiographs and CT which was initially proposed by [84]. Here the 2D-3D registration is based solely on pixel or voxel intensity and spatial information extracted from images. Intensity-based algorithms match the intensities of one image data set with the intensity of the other by maximising a similarity measure between them [85]. The matching can be restricted to regions of interest (ROIs) in the image, such as regions around bone surfaces in CT and fluoroscopy or an entire acquired image.

Intensity-based registration consists of three steps: 1) Generation of digitally reconstructed radiographs (DRRs) for each pose; 2) Measurement of the pose difference by
comparing the DRRs with the real fluoroscopic x-ray images; 3) Computation of a pose that minimises that difference.

Intensity based algorithms presented in literature can be classified into two categories, ROIs/ROIs [86-89] where regions of interest (ROIs) in both data sets, usually in the vicinity of the anatomy surface for each data set are used in the registration similarity measure or Image/Image [84] where the entire 3D CT scan and/or the 2D radiographic images are used in the similarity measure.

The calculation of a DRR by numerical summation of CT image intensities along projection rays involves high computation cost and is thus time consuming. Usually more than a hundred DRRs have to be generated during the search to achieve reliable and accurate registration. A number of methods have been proposed that simplify and consequently speeds up the calculation of DRRs, without losing information that is needed for registration [90].

The similarity measure performed between the DRR and the acquired radiograph is important as it dictates the success of the optimisation process. Work has been done by [85], where several similarity measures for registration of 3D CT and 2D x-ray images were tested providing some insight to the possible choice of an adequate similarly measure. Their conclusion was that similarity measures that use solely image intensity information, such as normalised cross coefficient of intensity, entropy of the difference image, and mutual information, are least appropriate for registration of real interventional medical images. Similarity measures that comprised information on spatial changes of intensities, like gradient correlation, pattern intensity, and gradient difference, were shown to be more accurate and more robust.

The key characteristic of intensity-based registration is that it does not require segmentation. The rationale is that using as much information as available and “averaging it out” reduces the influence of outliers and is, thus, more robust. However, this approach is computationally expensive since it requires generating high-quality DRRs and searching a six
degree-of-freedom space with local minima, which depends on the similarity measure employed. The Cyberknife radiosurgery system [91] is the only commercial system in routine clinical use that uses this registration method.

2.4 Discussion

While existing methods discussed above are to a certain extent effective, they have a number of limitations with regards to accuracy of reconstruction, adaptability to other orthopaedic cases and the level of manual intervention required. Work presented in this thesis is motivated by these limitations.

Currently available pre-operative planning systems for fracture reduction are all restrictive (time and cost) as they require a CT scan of the injured bone. This is not acceptable by our clinical partners. They have claimed that the increased radiation exposure and extra costs imposed by CT scanning is not warranted for femur fracture surgery. As previously mentioned CT scanning is conserved for complex orthopaedic interventions (e.g. pedicle screw insertion and intra-articular fractures). Consequently in many situations the surgeons will be limited to using 2D x-ray and fluoroscopic images. In order to ensure clinically applicability the research presented in this thesis focuses on 2D image based reconstruction techniques currently presented in intact bone reconstruction literature. Thus a novel fractured bone reconstruction technique was developed to be able to reconstruct 3D fractured bone models pre-operatively. The main criteria required when developing and building a customisation methodology is the accuracy of reconstruction, adaptability to other orthopaedic cases and the level of manual intervention required. The time taken was not of a main concern during this initial customisation phase. All the above criteria meant that not one methodology currently available in the literature would satisfy all the requirements. Currently available intact bone reconstruction literature, detailed in Section 2.3.1, present certain shortfalls that prohibit them being directly used for fractured bone customisation.
The current prevalent method proposed in literature is based on techniques that utilise a-priori knowledge about the anatomy along with the 2D images to perform the customisation. Statistical model based reconstruction techniques provides deformation constraints that can be utilised to ensure global shape conformance. These techniques ensure a high level of localised shape warping in comparison to other methodologies available. However the PDM creation, which is required for the statistical model, involves analysing a large database of anatomical models. The dataset also requires a time consuming corresponding landmark identification process. Thus typically only a small specific portion of the anatomy is modelled. This user intensive analysis means that models built through statistical techniques are not easily expandable to other orthopaedic procedures. Free form deformation based techniques can be utilised with sparse deformation information, however they provide no shape deformation constraints and are thus prone to unrealistically deformed shapes. They also require an initial setup of the control points and a bounding box. Morphological measurement based techniques provides deformation constraints that can be utilised to ensure localised shape conformance. However they do not provide the same level of shape constraints as seen with a statistical model. They are however quickly adaptable to other orthopaedic procedures. Moreover, morphologic measurement based techniques provide no local shape warping capabilities. Nevertheless, the extendibility of this methodology to provide such local shape customisation would not be a difficult task through a combination with the anatomical database based method. Anatomical database based reconstruction techniques provide a high degree of localised shape warping in comparison to other methodologies. However they provide no shape deformation constraints and thus will tend create unrealistically deformed shapes. Many of the literature that has used this methodology, all required a sufficiently large database to ensure the model and target anatomies are closely similar prior to deformation. These concerns are summarised in Table 2.1.
Table 2.1 Pros and cons of intact bone customisation/reconstruction methodologies.

<table>
<thead>
<tr>
<th>Method</th>
<th>Positives</th>
<th>Negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical Models</td>
<td>Realistic deformation constraints</td>
<td>Requires a large database of anatomical models</td>
</tr>
<tr>
<td></td>
<td>High level of localised shape warping</td>
<td>Time consuming</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not easily expandable to other orthopaedic procedures</td>
</tr>
<tr>
<td>Free Form Deformation</td>
<td>Ideal to be used with sparse deformation info</td>
<td>No shape deformation constraints</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requires an initial setup</td>
</tr>
<tr>
<td>Morphological Measurements</td>
<td>Provides local deformation constraints</td>
<td>No local shape warping capabilities</td>
</tr>
<tr>
<td></td>
<td>Quick adaptability to other procedures</td>
<td></td>
</tr>
<tr>
<td>Anatomical Database Based</td>
<td>local shape warping capabilities</td>
<td>Provides no shape deformation constraints</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requires a sufficiently large database</td>
</tr>
</tbody>
</table>

Thus it is proposed that this research develops a hybrid customisation methodology merging the anatomy database based method with the incorporation of the constraints placed by the morphological measurement based method. The morphological measurements will provide constraints on the key deformation criteria that will need to be adhered to. This will depend on the specific surgical procedure of concern and is easily accessible through the segmented patient-specific images. The non-rigid registration performed with the anatomical database based methodology will provide the overall shape customisation capability.

The intra-operative pose estimation is the next part of the research. Prior to the development of the novel 2D-3D registration algorithm there are several factors that were considered, each with its unique constraints. These four factors, listed below, culminate from current literature work [92].

1) Feature space (information extracted): There are a number of different sets of information that can be chosen from images, that includes, pixel or voxel intensities, edges, contours, surfaces and line intersections. Hence the information extracted will dictate if the registration algorithm is feature based or intensity based.

2) Similarity measure: The similarity measure gives a numerical value, which indicates the alignment of the two sets of information extracted. The choice of similarity measure is
affected by the type of information that is used for registration as well as the optimisation process used to maximise the similarity.

3) Search space: Defines the types of image transformations that are allowed to register the images. A larger search space requires more extensive search of the similarity measure optimum through the space and can increase the number of local similarity measure optima. With respect to speed and reliability of registration, it is important to keep the search space to the minimal dimension required to achieve an accurate registration.

4) Search strategy: Determines the best way to move through the search space in order to reach the optimum value of similarity measure. Each calculation of the similarity measure has a computation cost and therefore a search strategy should reach the optimal value of similarity measure in a minimum number of iterations. The choice of a search strategy greatly depends on the smoothness of the similarity measure within the search space.

Current literature in 2D-3D registration utilises either feature based or intensity based methods. Several shortfalls exist in the application of these techniques and they have been highlighted below (Table 2.2).

The key characteristic of feature based methods is that they use a small fraction of the image data, usually fiducial centers and anatomy surface points, whose location is assumed to be known very accurately. Feature based registration works best with high quality segmentation, an efficient feature pairing scheme, and a good outlier removal. In current literature robustness is achieved by first performing coarse registration with landmarks followed by fine registration with surfaces. Contrarily, the key characteristic of intensity based registration is that it does not require segmentation. The rationale is that using as much information as available and averaging it out reduces the influence of outliers and is, thus, more robust. However, this approach is computationally expensive since it requires generating high-quality DRRs and searching a six-dimensional space with local minima.
which depend on the similarity measure employed. It requires an initial pose guess close to the final pose and the definition of ROIs.

Table 2.2 Pros and cons of bone pose estimation methodologies.

<table>
<thead>
<tr>
<th></th>
<th>Positives</th>
<th>Negatives</th>
</tr>
</thead>
</table>
| **Feature Based** | Fast convergence to optimal solution due to less information (specific number of features) being used in the optimisation.  
Accurate convergence with the availability of accurate feature information. | Feature extraction is crucial and would need to be highly accurate.  
Image pre-processing (segmentation) would be needed to extract the feature information required. |
| **Intensity Based** | Requires no pre-processing of the images. | Slow registration as the DRRs that are created to perform the registration require a high level of CPU resources and is time consuming.  
Less accurate registration. |

Although related registration methodologies have been developed in the past, there remains significant challenges as these methodologies have not yet been introduced into fractured bone pose estimation. The anatomy feature based intra-operative pose estimation framework proposed in this thesis is motivated by these limitations as listed below.

The 2D-3D registration algorithm proposed by this thesis can be separated into two distinct parts: 1) frontal and lateral alignment, and 2) axial alignment.

The initial partial registration to conduct the only the frontal and lateral alignment is conducted through anatomical feature based registration. This algorithm does not involve any iterative optimisation, thus the pose is analytically solved. It has been identified through discussion with surgeons that this is a currently acceptable solution and will provide adequate guidance. In this work the features used are extremely robust and are extracted through a novel image processing algorithms that have been tested for reliability and proof of concept.

Full six degree-of-freedom registration is next achieved by conducting an axial alignment. This final alignment is carried out through an optimisation based technique. Again the algorithm has been developed to be timely, with the dimensionality of the pose estimation separated, for fast and robust optimisation.
Chapter 2 - Literature Review

Based on review of literature and consultation with orthopaedic surgical staff this research proposes a new system for orthopaedic IGS, focusing not only on long bone fractures but also osteotomy and arthroplasty. The developed pre-operative planning framework was developed in strict guidance to the requirements presented by the AO guidelines [31]. Figure 2.4 shows a high level overview of the procedure and the surgeon interaction.

![Figure 2.4 Overview of proposed reduction process. The patient is imaged by the surgeon to form a patient-specific model, which is used for pre-operative reduction path planning. The surgeon can then review the path, visually inspecting the result before submitting the specific plan for surgery. During operation the surgeon is guided by the intra-operative 3D model which ensures ease of surgery and accuracy.](image)

### 2.5 Chapter Summary

This chapter presented a background to image guided surgery starting with neurosurgery, spinal applications and orthopaedic applications. The chapter highlighted the shortcomings of the current literature in orthopaedic IGS focusing on long bones and specifically femur fractures. The chapter concluded with an overview of the technology developed in this research and related work by others in the same group.
Chapter 3 - Pre-Operative Planning: Intact Bone Applications

Chapter 3 Pre-Operative Planning: Intact Bone Applications

Three-dimensional pre-operative planning is vital for several orthopaedic cases involving implant fixation and bone incisions. The research presented in this chapter focuses on two such orthopaedic applications: hip and knee arthroplasty and femur osteotomy. In both these applications proper pre-operative planning is vital to reduce complication including dislocations, accelerated wear, and loosening of the implants [31]. Correct implant orientation is the most important factor in preventing impingement, which is a major cause of dislocation and wear following total hip replacement surgery [31]. This implant orientation is dependent upon patient-specific factors such as pelvic anatomy bone coverage and level of femoral shape, and can affect leg length and offsets. In long bone osteotomy, the bone incision angle recognition would require extensive spatial planning preceding the procedure. Thus for both these cases, it is vital that the surgeon conducts pre-operative planning prior to surgery and utilise 3D models of bones for spatial and shape feedback. This chapter tests femur, tibia and iliac reconstruction, which are the three primary bones of interest in the aforementioned cases. The work presented this chapter has been published in [32-34].

3.1 Introduction

Pre-operative planning has been identified as an essential requirement for successful surgical outcomes. The AO foundation for orthopaedic research, instructs on the importance of pre-operative planning through its courses [31]. Three-dimensional knowledge of the bony anatomy is vital for the pre-operative planning of several orthopaedic IGS procedures. The surgeon should emerge from the planning phase with a clear idea of the patient’s bone dimensions, detailed plan of the internal fixation and the overall surgical approach. In current orthopaedic cases most pre-operative diagnostics and planning is conducted through 2D x-ray
imaging. Two-dimensional images lack significant spatial information that is present in 3D modalities. Imaging modalities such as CT have the ability to provide direct 3D volumetric images. However, the use of such imaging is restricted to a minority of complex procedures due to constraints placed by cost, availability and risks posed by unwarranted detailed imaging. Thus an alternative to direct 3D imaging must be developed to augment procedures that currently rely on pure 2D radiographs. The work presented in this chapter is motivated by this requirement, and proposes a 3D bone reconstruction framework from 2D radiographic images. There are several procedures that will be benefited from such pre-operative 3D visualisation. These include hip and knee arthroplasties [93, 94], osteotomy [20] and several fracture reduction procedures [4, 37]. The first two of these cases will be addressed in this chapter. While the subsequent chapter will focus on the primary case of femur fracture reduction planning and diagnostics (for fracture assessment, nail selection etc).
operative planning also allows the surgical team to prepare the instrumentation required for each operation, have the proper inventory of implants available, and predict complications and needs that may arise during surgery. During the planning session the surgeon integrates the general goals of arthroplasty surgery with the particular patient’s anatomy. Goals include conserving bone stock, optimising implant position and fit, equalising leg length, and avoiding complications.

Furthermore, another benefit of reconstructing pre-operative 3D bone models is the ability to custom design implant components based on patient-specific anatomical measurements (to overcome existing shortcomings of current designs) [92, 95]. The longevity of implant components are highly dependent on the initial fit between the bone surface and the implant. Currently, most custom knee implants are designed using a generic shape of the surfaces and the bone-implant interface is created using planar cuts.

The second case of interest involves the use of patient-specific anatomical models for planning of an osteotomy. This involves the straightening of a bone that has healed crookedly following a fracture or to correct congenital deformity (bow-legged) of the bone [20, 92, 95]. Long bone deformity may consist of a combination of coronal, sagittal and oblique plane distortion. The most common of which are coronal and sagittal plane angular deformities [96]. For these cases, the planar angular and torsional misalignment of the bone axis can be identified through the frontal and lateral radiographs. However several key angles including the acetabular anteversion, are not measurable through 2D imaging. Moreover, the plane on which the surgical cut has to be performed and the magnitude of the angular correction can only be accurately defined by measurements viable through 3D spatial visualisation [20, 96].

The aforementioned concerns for the lack of 3D pre-operative planning in arthroplasty and osteotomy motivated the work presented in this chapter. The proposed framework to achieve 3D reconstruction is outlined in Figure 3.1. The 2D x-ray images are initially processed to extract the edge points that potentially form the femur boundary. A non-rigid registration is then performed between the edges identified in the x-ray image and the
projected contour points of the generic model. The identified point correspondence will next be interpolated to create a 2D planar translational field in both the anterior and lateral viewpoints. This translational field will identify the deformations required by the 3D anatomical model in the equivalent viewpoint. Finally a full 3D translational field will be created through interpolation and the 3D generic anatomical data will be deformed accordingly.

![Diagram](image)

**Figure 3.1 The main process steps of the proposed intact bone pre-operative reconstruction framework.**

### 3.2 Anatomical Generic Model Preparation

The anatomical generic surface models utilised in this thesis were segmented and extracted from CT scan data. Each bone was positioned in the centre of the scan field to minimise beam hardening and the partial volume effect. This was achieved by aligning the longitudinal axis of the bone with the scanning axis. The images were acquired in a helical mode with a HiSpeed CT/i, General Electric Medical Systems CT scanner, operating at 120 kV, 150 mA, with a pitch of 1.5:1 and a gantry rotation time of 0.8 seconds. With all slices overlapping by 50% the effective slice spacing was 0.5 mm for the condylar region and 1.5 mm for the diaphysis. From each femur about 155 transverse image slices with a pixel size of 0.33 mm were obtained.

Bones scanned include six femora, six tibias and six iliacs. Next, each of the individual bone models were aligned with a common coordinate system by adjusting the position of the models until the posterior aspects of the condyles appeared superimposed in the lateral/frontal view. After alignment, the models were scaled uniformly to the average cross-sectional width.
Chapter 3 - Pre-Operative Planning: Intact Bone Applications

of the model groups. The aligned and scaled models were next sectioned in planes orthogonal to the z-axis. The models were then sectioned at 1 mm intervals. The external bone contour was automatically extracted from each image slice using a border tracing algorithm [10]. Subsequently, in order to determine the average outline of the bone, the extracted contours were combined for each plane of section. To obtain a better alignment for the contours prior to averaging, the contours were translated such that their centroids were aligned at a common point. This was due to the contours in the shaft not being closely superimposed. The average of all the centroid positions was chosen as the common point for the alignment of the contours. The generic 3D model was reconstructed from the average contour outlines by firstly creating a volume (3D) array from the images of the average contour outlines. Then a list of vertices and polygons describing the contour surface was generated. Finally, the minimum amount of spatial smoothing was performed on the polygon mesh in order to preserve the local morphology. After their reconstruction the 3D models were saved in Virtual Reality Modelling Language (VRML) file format so that they can be imported as polygon mesh objects for further processing.

The generic models required certain automated pre-processing to identify the outer surface edge points that would be clearly identifiable on radiographic images. The outer contours in the frontal and lateral directions were identified through projection ray-tracing (also referred to as ray-casting). Projection rays are constructed between points of the imaging plane and the imaging source. Then the binary projection silhouette is computed by identifying whether the volume element intersects a particular ray. A silhouette image is a binary image that represents whether an image point, projected as a visual ray from the camera centre, intersects the imaged object’s surface in the scene. Each pixel is either classified as a silhouette in the foreground, or as belonging to the background. Thus for a particular instance of the 3D surface data \( S(T) \) with transformation \( T \), two 2D projection images, \( I_p \), are created in the frontal and lateral directions (Figure 3.2). The projection image is generated by means of rays that are emitted by a point source and propagated onto an image acquisition plane. The calibration of the point source with respect to the image plane is
known and is consistent with the camera calibration that was used for the acquisition of the reference image. If the path of a ray to the image plane intersects the surface model then $I_p(x)$ is set to one, where $x$ is the projected intersection. Otherwise, $I_p(x)$ is set to zero. Once the projection silhouette is created the outer contour is extracted from the binary image.

Figure 3.2 Femoral generic model and projections. The blue contour is the lateral projection while the red contour is the frontal projection.

The projective rays detailed above of the generic bone were created through a basic pinhole model [97] that represents mapping from a 3D scene onto a 2D image. The relationship between the coordinates of a 2D image pixel and its corresponding 3D object point can be expressed through the equation below.

$$x = MX \text{ where } M = K[R \mid t]$$ (3.1)

Here the $3 \times 4$ projection matrix $M$, relates any 3D point $X$ to its corresponding projection $x$ in the acquired image. The intrinsic projection parameters of the x-ray tube (focal length and principal point coordinates), are represented through $K$, and the extrinsic parameters (rotation and translation of the acquisition system in a world coordinate system) through $[R \mid t]$. 

44
Chapter 3 - Pre-Operative Planning: Intact Bone Applications

The camera calibration mythology is not an integral part of this research. The intrinsic camera calibration method simply relies on the technique described in [97] where a set of coplanar points visible on the image is used to compute the camera parameters. We utilised a radio-opaque board with spherical metal markers applied to it on both sides.

The extrinsic parameters relating to the orientation $R$ and position $T$ of the acquisition system are computed for different x-ray tube orientations. As previously mentioned, there are two main viewpoints that typical pre-operative images are viewed in: frontal and lateral. Thus the extrinsic parameters are two preset orientations. However care must be taken during the image acquisition to ensure that the femoral bone lies iso-centrally to the two viewpoints. Clinically this is achieved by examination of the knee joint (patella) and the ankle joint. Prior to imaging, the patient’s leg must lie with the iliac spine lined up with the patella and the cleft between the first and second toes.

3.3 Need for a Generic Model

The rationale behind the use of the generic model is attributed to the ill defined problem posed by the 3D reconstruction of an object from two 2D projective views (it inherently lacks any axial cross-sectional information on the object of interest).

Several early work in volumetric model reconstruction (from images) employ techniques that reconstruct the 3D object based on geometric intersection and do not utilise any a-priori knowledge of the cross-section of the object [67, 68]. By determining voxel occupancy, the task is to decide, with the scene constructed as a set of 3D voxels, whether each individual voxel is empty or occupied. The most typical solution is by silhouette intersection, either from multiple views of a single object or using a single camera with the object rotating on a turntable to approximate the visual hull of the imaged object (Figure 3.3(a)). Thus it is the intersection of all back-projected silhouette cones. The idea behind geometric intersection and shape from silhouettes is that since each silhouette point defines a ray in scene space that
intersects the object at some unknown depth along the ray, the union of all such rays should produce a visual hull in which the actual object must lie within. Thus the visual hull is always guaranteed to enclose the true object. Shape from voxel occupancy and silhouette intersection techniques has two major disadvantages in that they fail to exploit the consistency of the scene between different views and reconstruction from silhouettes alone can in general only accurately reconstruct convex objects and cannot sufficiently describe surface concavities (Figure 3.3(a)).

For example, Figure 3.3(b) illustrates a cross-sectional slice of a cylindrical object with two orthogonal projective views. Through the intrinsic calibration of the projective matrix and the extrinsic identification of the transformation from one view to the other, one can at best identify the four intersectional corners (in 3D) highlighted in Figure 3.3(b). Thus it is not possible to predict the cross-sectional geometry through these means. The only technique seen in literature to “predict” the cross-sectional shape is through a spline based curvature approximation. This technique can only produce artificially smoothed circular/oblique cross-sectional objects, which makes it flawed for bony anatomy reconstruction. Hence this dictates having a-prior knowledge (generic model) on the object to provide the missing cross-sectional information.

![Figure 3.3 Illustrations to highlight the need of an anatomical generic model. From left to right: (a) a cross-sectional object is viewed from three cameras, and the visual hull is illustrated. Notice that the concavity cannot be reconstructed using volume silhouettes intersection; (b) cross-sectional illustration of the projection process.](image-url)
3.4 Model Initialisation

Model initialisation, with regard to position, orientation and scale might be necessary in certain cases. Typically, femur x-ray images are acquired in only the anterior and lateral viewpoints. Thus the pose (position/orientation) of the bone when imaged is standard on many femur x-ray images. The manual pose initialisation will only be required if a different angle of acquisition is employed during the x-ray imaging. This initialisation involves a six degree-of-freedom movement of the 3D model, with three translational and three rotational parameters.

The pose estimation employs the pair of 2D x-ray images and the corresponding 3D anatomical model. It then aims to identify the transformation of the model so that its projections on the frontal and lateral planes match the acquired x-ray images. This registration can be considered as determining the equivalent affine transformation which includes a set of translational \((T_x, T_y, T_z)\), and rotational \((R_x, R_y, R_z)\) parameters where the x, y and z axis are aligned with the frontal, lateral and axial axis of the bone segment. The user would interactively identify the pose of the 3D bone model. There are seven parameters involved in the pose estimation, which include control over the six degrees-of-freedom movement as well as the centre of rotation.

3.5 Contour Extraction

The main objective of the proposed edge extraction technique is to identify the bone object boundaries with sufficient continuity to be successfully employed in the proceeding shape based point correspondence estimation. There are a multitude of medical imaging segmentation algorithms presented in literature, as summarised in [10]. However x-ray images inhabit several complexities that prohibit the use of such classical “low level” segmentation algorithms. Issues faced during the segmentation process of x-ray images
include poor contrast, ill-defined boundaries, noise and acquisition artefacts (such as illumination gradients) [10, 98].

These issues are due to the non-homogeneous internal structure of long bones. For instance, in a femur x-ray image, the femoral head region contains a non-uniform texture pattern due to the trabeculae spongy bone, and the femoral shaft region has non-uniform intensity due to the hollow interior within solid bony walls (Figure 3.4 (a)). Moreover, the femoral head overlaps with the pelvis bone which causes the extraction of the femur boundary to become a more complex problem than a classical image segmentation problem (Figure 3.4 (b)).

Due the aforementioned issues, segmentation is a complex task and cannot be achieved using grey-level information alone. Thus to achieve successful segmentation, low-level segmentation algorithms (thresholding, clustering and region growing algorithms) have to be combined with higher level techniques such as deformable and active models. Thus an interactive segmentation technique was utilised to achieve robust and accurate segmentation. Interactive medical image segmentation literature suggests techniques such as active contours (or snakes) and semi-autonomous boundary tracing tools (LiveWire, Intelligent Scissors etc).
Two segmentation algorithms are presented, both an active contours and a LiveWire segmentation technique. The pros and cons of each technique are presented and the distinct scenarios, in which they can be utilised in, are discussed herein. The main point of difference is that the active contours technique can be fully autonomous while the LiveWire technique requires manual user interaction.

Active contour introduced in [3] is an energy-based method which evolves through the minimisation of its functional. The latter is a balanced combination of internal energy constraints (based on bending and stretching physical properties of thin plates) and external energy constraints (based on image information which best describe the features to extract). The work presented in [99] was primarily introduced as an interactive method. Using an input device, users were able to attract (via the adjunction of an additional energetic term known as string) the active contour to some interesting part of the image or instead repel (via an energetic term known as volcano) it away from other specific regions such as incorrect image edges. Although, active contours provided an influential approach to image segmentation it had several limitations. Firstly, local optimisation techniques were used to minimise the Snake energy, forcing the user to initialise the Snake close to the target object boundary to ensure convergence to the desired solution. Secondly, the internal deformation energies, made it incapable of deforming to fit complex shapes with rapid changes. Finally Snakes have prohibitive editing capabilities where user interaction based solely on virtual springs does not provide the degree of control and precision required to accurately segment complex objects.

The active contours segmentation algorithm proposed in this chapter is a combination of adaptive thresholding and a level-set based segmentation [100]. The segmented image obtained after the adaptive threshold process is utilised as the initial contour for the level set methodology which fine tunes the segmentation. Figure 3.5 illustrates the initial use of adaptive thresholding to identify approximately the object of interest followed by the level set based segmentation to obtain the final segmented image.
Chapter 3 - Pre-Operative Planning: Intact Bone Applications

The active contours segmentation methodology implemented is a variation of the one proposed by [101].

The proposed algorithm begins by defining a contour parameterised by arc length $s$ as

$$C(s) = \{(x(s), y(s)) : 0 \leq s \leq L\} : \mathbb{R} \rightarrow \Omega$$

(3.2)

where $L$ denotes the length of the contour $C$, and $\Omega$ denotes the entire domain of an image $I(x,y)$. The corresponding expression in a discrete domain approximates the continuous expression as

$$C(s) \approx C(n) = \{(x(s), y(s)) : 0 \leq n \leq N, s = 0 + n\Delta s\}$$

(3.3)

Where $L = N\Delta s$. An energy function $E(C)$ can be defined on the contour such as,

$$E(C) = E_{\text{int}} + E_{\text{ext}}$$

(3.4)

where $E_{\text{int}}$ and $E_{\text{ext}}$ respectively denote the internal energy and external energy functions. The internal energy function determines the regularity and smooth shape, of the contour. The implemented choice for the internal energy is a quadratic functional given by

$$E_{\text{int}} = \sum_{n=0}^{N} \left[|C'(n)|^2 + \beta |C''(n)|^2\right] \Delta s$$

(3.5)

Here $\alpha$ controls the tension of the contour and $\beta$ controls the rigidity of the contour. The external energy term determines the criteria of contour evolution depending on the image $I(x, y)$, and can be defined as

$$E_{\text{ext}} = \sum_{n=0}^{N} E_{\text{img}}(c(n)) \Delta s$$

(3.6)

where $E_{\text{img}}$ denotes a scalar function defined on the image plane, so the local minimum of $E_{\text{img}}$ attracts the active contour to the edges. An implemented edge attraction function is a function of the image gradient, given by
Chapter 3 - Pre-Operative Planning: Intact Bone Applications

\[ E_{\text{img}}(x,y) = \frac{1}{\lambda \sum_{G_o \times I(x,y)}} \]

(3.7)

where \( G_o \) denotes a Gaussian smoothing filter with the standard deviation \( \sigma \), and \( \lambda \) is a suitably chosen constant. Solving the problem of active contours is to find the contour \( C \) that minimises the total energy term \( E \) with the given set of weights \( a, \beta \) and \( \lambda \). The contour points residing on the image plane are defined in the initial stage, and then the next position of those snake points are determined by the local minimum \( E \). The connected form of those points, are considered as the contour to proceed with. Figure 3.5 shows an example of the above method in operation over a series of iterations.

Figure 3.5 The segmentation process performed on a typical x-ray image that inhabits the problems seen in many x-ray images. From left to right: (a) original image; (b) adaptive thresholding (mean intensity was utilised as the thresholding limit); (c) the initial contour used for the level set segmentation process; (d) the level set contour after 20 iterations; (e) the level set contour after 50 iterations resulting in the final segmented image; (f) image of a femoral head with manually initialised contour; (g) the level set contour after 20 iterations; (h) the level set contour after 35 iterations resulting in the final segmented image.
Following the introduction of Snakes, several related semi-autonomous boundary tracing tools ("LiveWire" or "Intelligent Scissors" [102, 103]) have emerged as effective segmentation techniques. These tools allow the user further interaction and control over the segmentation process. LiveWire methods are much faster, accurate and are more robust to initialisation than Snakes, which rely on manual tracing of object boundaries. Thus an interactive LiveWire technique was employed in this chapter to provide a higher degree of interactive segmentation, which would cater for x-ray images.

The LiveWire technique proposed in this chapter is based on [102] with modifications to suit greyscale x-ray imaging. The interactive segmentation requires the user to initially specify a seed point on the object boundary utilising an input device (mouse). The input device must subsequently be moved to advance the cursor to a point further along the object boundary. A globally optimum path from the initial seed point to the current point is computed and displayed in real-time. The optimal path is determined by assigning cost functions to boundary elements (edge strength/edge orientation), and then finding the minimum cost path. As the user moves the cursor slightly, different paths are computed and displayed. If the cursor moves close to the boundary, the LiveWire snaps to the edge. If the user is satisfied with the computed boundary segment, the cursor point can be set. This point becomes the new seed point and the recursive process continues.

The minimum cost path should correspond to an image object boundary that exhibit strong edge features. Thus, the local cost matrix is created from two edge features: gradient magnitude and gradient direction. The overall local cost matrix is computed as a weighted sum of these components. \( C(m,n) \) represents the local cost of the link from pixel \( m \) to a neighbouring pixel \( n \), where \( f_G \) and \( f_D \) are the gradient magnitude and gradient direction cost functions respectively and \( w_G \) and \( w_D \) are the corresponding scalar weights.

\[
C(m,n) = w_G f_G + w_D f_D \quad (3.8)
\]
Gradient magnitude provides a direct correlation between edge strength and local cost. If \( F \) is the greyscale x-ray image, then the gradient magnitude \( G \) is approximated with,

\[
G = \|\nabla F\| = \sqrt{\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2}
\]

(3.9)

The gradient is next scaled and inverted to ensure high gradients produce low costs to facilitate minimal cost path optimisation.

\[
f_G = 1 - \frac{G}{\max(G)}
\]

(3.10)

Gradient direction adds a smoothness constraint to the boundary by associating a high cost for sharp changes in boundary direction. The gradient direction is the unit vector defined by the partials of an image \( I \) in \( x \) and \( y \) (\( I_x \) and \( I_y \)). The gradient direction feature cost can be defined as \( f_d \), where \( D(m) \) is the unit vector perpendicular (rotated 90 degrees clockwise) to the gradient direction at point \( m \), and \( n-m \) is the bidirectional link or edge vector between neighbouring pixels \( m \) and \( n \) pointing towards \( n \). Links are either horizontal, vertical, or diagonal (8-connectivity).

\[
f_d(m,n) = \frac{1}{\pi} \left[ \arccos(c \cdot D(m) \cdot (n-m)) + \arccos(c \cdot D(n) \cdot (n-m)) \right]
\]

(3.11)

Where,

\[
D(m) = [I_y(m), I_x(m)] \quad D(n) = [I_y(n), I_x(n)] \quad c = \text{sign}(D(m) \cdot (n-m))
\]

(3.12)

The link direction forces high cost to an edge between two pixels that have similar gradient directions but are near perpendicular to the link between them. Thus the direction feature cost is low when the gradient directions of the two pixels are similar to each other and the link between them.

Finally, 2D graph searching is employed for an optimal minimal cost path selection. This work utilises the optimal graph search presented by [104]. Dijkstra's algorithm is a graph search algorithm that solves the shortest path (path with the lowest cost) problem for a graph
with nonnegative edge path costs. The sequence of images in Figure 3.6 shows an example deformed femoral head x-ray being segmented through the LiveWire technique.

![Figure 3.6 Edge extraction on an example x-ray of a femoral head. The user defines lines to perform the segmentation. Segmentation time: 20s.](image)

Proceeding the edge point extraction, a certain level of sub-sampling has to be performed on the edge points identified to provide a realistic number of points as input to the point correspondence estimation algorithm. The sub-sampling is weighted in accordance to the connected components of the extracted edges. Thus the final relative proportion of pixels after sub-sampling will remain the same. These sub-sampled points are then utilised in the point matching algorithm as detailed in the proceeding section.

### 3.6 Shape Customisation

A non-rigid registration between the 2D projective contours of the 3D generic model and the extracted edges of the patient-specific x-ray images is performed as the first step of the customisation process. The registration is performed through a point correspondence estimation between the two point sets (Figure 3.7). Correspondence algorithms generally consist of two parts, firstly a similarity measure which provides a measure of “correspondence” between two pairs of points or features in different images, and secondly, a cost function that analyses the values produced by the similarity measure to identify a series of one-to-one matches between the points on those images. Figure 3.7 illustrates the proposed
process steps of the non-rigid registration. Each of these steps will be detailed in the proceeding pages.

Figure 3.7 The main process steps of the non rigid registration process.

3.6.1 Similarity Measure

The similarity measure utilised for this point correspondence estimation needs to be robust to several factors that are inherent problems seen in edges extracted in x-ray images. These factors include,

1) Disturbance (Noise): extracted edge points will not form a smooth outline of the anatomy of interest.

2) Deformation: extracted edge points will have shape variations as well as affine alteration (translation, rotation and scale).

3) Outliers: Edges that do not form the outline of the anatomy of interest will erroneously be extracted.

4) Occlusions: Edge points that form the outer contour of the anatomy of interest might be occluded or might not have been detected.

A novel similarity measure was developed, in order to cater for the variety of factors discussed above and is separated into several components that utilise four key features: topology information, edge orientation, curvature and continuity (distance). The final similarity measure will be a weighted combination of all these individual similarities. Using multiple features increases the discrimination power of the individual features and leads to improved results, as seen empirically. The individual components of the similarity measure are discussed herein.
In the proposed approach we treat the model and target objects as two point sets. These point sets represent either the internal or the external contours of the model (generic model) and target (patient) bone. They do not need to correspond to key feature points of the shape, however the inclusion of key points such as maxima of curvature or inflection points would strengthen the registration. The point set can be sampled with uniform spacing, although this is also not critical. Figure 3.8 show that sampled points from two shapes. Assuming that the points are piecewise smooth, we can obtain a good approximation to the underlying continuous shapes by picking a sufficiently large number of points. The requirement of the similarity measure is to identify the best matching target point, \( p_t \), for each of the model contour points, \( p_m \). The reasoning behind the use of multiple descriptors is that empirical experience has suggested that rich local descriptors that take into account the local gray-scale window or a vector of filter output performs better than a single pixel brightness based measure.

Shape based point matching is critical as it drives the correspondence between salient features of the objects. The shape descriptor introduced in this thesis is one based on shape histograms which describe the distribution of a series of points with respect to a given point on a shape. Termed Shape Context, this descriptor was introduced in [105, 106].

**Figure 3.8** Factors that the developed point matching algorithm must be able to handle to perform successful medical image processing. From left to right: (a) point matching with noise; (b) point matching with deformation; (c) point matching with outliers; (d) point matching with occlusions.
For a point $p_i$ on the shape, a histogram $h_i$ of the relative coordinates of the remaining $n-1$ points is computed as,

$$h_i(n) = \#\{q \neq p_i : (q - p_i) \in \text{bin}(n)\}$$ \hspace{1cm} (3.13)

where $N$ is the total bin number. This histogram is defined as the shape context of $p_i$. The bins that are uniform in log-polar space make the descriptor more sensitive to the positions of nearby sample points than to those of points further away. An example is shown in Figure 3.9. The amount of the segmented bins affects the similarity result. If a segmented bin area is too small, the similarity information is too noisy. If a segmented bin area is too big, the similarity information contains the rough global information without the local difference information. As shape contexts are distributions represented as histograms, it is natural to use the Chi-squared ($\chi^2$) test statistic.

The shape context-based cost function ($C_{\text{shape}}$) to match a point $p_m$ on the model contour to a point $p_t$ on the target contour can be expressed as:

$$C_{\text{shape}}(p_m, p_t) = \frac{1}{2} \sum_{n=1}^{N} \frac{(h_m(n) - h_t(n))^2}{h_m(n) + h_t(n)}$$ \hspace{1cm} (3.14)

where $h_m(n)$ and $h_t(n)$ denote the N-bin histogram (normalised) at $p_m$ and $p_t$, and $N$ is the total number of points under consideration.
Figure 3.9 Illustrations of the shape context based measure. From left to right: (a) model shape used to illustrate the shape context measure, reference samples marked by $\bigcirc$ and $\Delta$; (b) target shape used to illustrate the shape context measure, reference sample marked by $\bigdiamond$; (c) log-polar histogram of the point marked by $\bigcirc$; (d) log-polar histogram of the point marked by $\Delta$; (e) log-polar histogram of the point marked by $\bigdiamond$. The log-polar histogram is for the coordinates of the rest of the point set measured using the reference point as the origin (dark segments indicate large values). Note the visual similarity of the shape contexts for $\bigcirc$ and $\bigdiamond$ which were computed for relatively similar points on the two shapes. In contrast, the shape context for $\Delta$ is quite different.

Due to the Chi-squared test used as the matching cost between the two shapes, the $C_{shape}$ similarity measure is intrinsically bounded in $[0,1]$. The primary benefit of utilising this measure is that it is invariant to translation, rotation, scale, and shape. Due to this the shape context descriptor is widely used in many object recognition applications [105]. The descriptor is however vulnerable to outliers, noise and occlusions and thus is supported by other features.

Edge orientation is introduced as a similarity feature to ensure robustness against outliers. Literature work presented by [107] illustrate the effective use of edge orientation information to group points that belong to a certain region of interest. This same idea was applied in this
situation to ensure point correspondence is only performed between points with similar gradient orientations. The edge orientation cost function ($C_{\text{edge}}$) is calculated as:

$$C_{\text{edge}}(p_m, p_t) = \sum_{n=1}^{N} |\theta_m - \theta_t|$$  \hspace{1cm} (3.15)

where $\theta_m$ and $\theta_t$ denote model and target edge orientations. The measure is subsequently normalised to be bounded in [0, 1].

The curvature feature information is utilised to ensure further robustness against outliers. The curvature of edge point $p_i$ is calculated as:

$$c(p_i) = \left| p_{i-1} - 2p_i + p_{i+1} \right|$$  \hspace{1cm} (3.16)

where $p_{i-1}$ and $p_{i+1}$ represent the neighbouring points to $p_i$.

Since the curvature is used to compare structurally different objects, normalisation is used to stretch the dynamic range onto a [-1, 1] interval. This normalisation facilitates matching of structurally different regions by assigning curvature values according to rank order, rather than using absolute curvature and thus is scale invariant. The curvature cost function ($C_{\text{curv}}$) is calculated as:

$$C_{\text{curv}}(p_m, p_t) = \sum_{n=2}^{N-1} |\hat{c}_m - \hat{c}_t|$$  \hspace{1cm} (3.17)

where $\hat{c}_m$ and $\hat{c}_t$ are the normalised curvature values of the model and target point sets. $C_{\text{curv}}$ is intrinsically bounded in [0, 1].

The Euclidean distance feature information is utilised to ensure continuity of the matched points. The assumption here is that neighbouring points on the model should map onto target points which are also close to each other. The importance of figural continuity has been shown by [106]. The Euclidean distance based measure employed can be denoted as below. The distance term has to be normalised through the external reference term $d_0$.  


Thus the final similarity measure is a weighted combination of the individual similarity measures as shown in the equation below. All measures are bounded between [0, 1]. Empirical weights \( w_{\text{shape}} \), \( w_{\text{edge}} \), \( w_{\text{curv}} \) and \( w_{\text{dist}} \) are assigned to shape context, edge orientation, curvature and Euclidean point distance respectively.

\[
C_{\text{total}} = w_{\text{shape}}C_{\text{shape}} + w_{\text{edge}}C_{\text{edge}} + w_{\text{curv}}C_{\text{curv}} + w_{\text{dist}}C_{\text{dist}}
\]  

(3.19)

The total cost of matching the point sets can be minimised through a Bi-partite graph searching methodology. The minimisation of the total cost matrix \( C_{\text{total}} \) between point sets, \( p_m \) and \( p_n \), is subject to the constraint that the matching is one-to-one. This square assignment problem is solved through the Hungarian algorithm with a time complexity of \( O(N^3) \) [108] (Figure 3.10).

The correspondences identified through the Hungarian algorithm are subsequently processed and filtered to remove any outliers (misidentified correspondences). The filtering is performed on a moving window of the identified translational values (of the correspondences) and by ensuring that they are within \( \pm 3 \) standard deviations from the local mean (Figure 3.10).

\[
\left| T(n) - \mu(T(n)) \right| \geq \beta \sigma(T(n)) \rightarrow \text{Outlier}
\]  

(3.20)

The filtering is followed by a regularised TPS smoothing to interpolate the outer contour translations identified (through the non-rigid registration) to the entire projective view. Subsequent to the planar interpolation, a 3D translational field used to deform the generic model is created by performing an interpolation on the sparse translational data identified. This interpolation is described in Section 3.7 (Figure 3.12).

The algorithm iteratively refines and deforms the 3D generic model by continuously performing correspondence estimations, creating the 2D planar translational field (anterior...
and lateral) and deforming the 3D generic model (Figure 3.10). Typically the algorithm runs for a fixed number of iterations or till it converges to a user specified tolerance error level.

![Figure 3.10 Illustrations of the shape customisation stage. From left to right: (a) non-rigid registration on the sub-sampled edge contour data points for the anterior view and the associated translational field displayed on the generic model; (b) non-rigid registration on the sub-sampled edge contour data points for the lateral view and the associated translational field displayed on the generic model.](image)

### 3.6.2 Similarity Measure Feature Weight Selection

Several empirical tests were conducted to validate the combination of features used (Shape Context, Curvature, Edge Orientation, Euclidean Distance) as well as the best combination of weights, $W_{\text{shape}}$, $W_{\text{edge}}$, $W_{\text{curv}}$ and $W_{\text{dist}}$ for the scenarios faced (Noise, Deformation, Outliers and Occlusions). Figure 3.11 indicates the results of the testing that was done. The testing was performed on data that was synthetically modified into the four scenarios to test the algorithms effectiveness against noise, outliers, deformations and occlusions. Testing was performed on several levels of change per scenario, which is indicated on the x-axis of the graphs. The error quoted is the average Euclidean distance between the correspondences identified through the individual feature alone and the ground truth correspondence known prior to the test.

In cases where disturbance (noise) exists and the extracted edge points do not form a smooth outline of the anatomy of interest, the distance and shape context features show the best performance. The curvature and edge orientation do not perform highly in the disturbance testing since high disturbance will be classified as outliers. Thus conversely in cases of outliers where edges that do not form the outline of the anatomy of interest have
been erroneously extracted, then the curvature and edge orientation will perform extremely well. In these cases the distance feature will provide a means of filtering misidentified correspondences. Deformation cases, where extracted edge points have shape variations as well as affine alteration are best segmented through the use of a high shape context weight. This is not surprising since the measure was designed to be invariant to translation, rotation, scale and shape deformation. For occluded cases where edge points that form the outer contour of the anatomy of interest are hidden or might not have been detected all the features appear to be equally effective.

Figure 3.11 Results of the testing performed on phantom data to identify the effectiveness of each feature measurement (Shape, Curvature, Edge Orientation and Euclidean Distance) against typical problems faced in radiographic image processing (Noise, Outliers, Deformations and Occlusions).

### 3.7 Full 3D Deformation

Two deformation algorithms are presented, 1) TPS based interpolation; and 2) Free-form deformation based interpolation. TPS is a faster procedure and has shown to be statistically much more inferior to the free-form deformation based technique. The results section will highlight the differences of these two techniques further. This final 3D deformation is a 4D interpolation problem, with 3D coordinates and a deformation magnitude. For the TPS
technique this 4D problem is simplified to ensure that the reconstruction is performed in a realistic time frame. This is vital as the reconstruction process might go through several iterations. Thus the 4D problem is transformed into one which is 3D by making several assumptions. The interpolation is performed by initially transforming the cartesian data points into a spherical coordinate system. Following the transformation, the surface is separated into smaller local regions on which a TPS interpolation is performed on the inclination/azimuth angles as the independent data and the three translations as the dependant data. The interpolation is performed individually per single translational axis (x, y and z). The separation of the surface points into smaller local regions is conducted to ensure convexity that is needed to eliminate duplicate data on the radial axis in the spherical coordinate frame. Thus the radial dimension of the spherical coordinate system can be eliminated from the interpolation, which condenses the problem to one which is 3D. This local region interpolation also ensures that the deformation is performed adaptively.

The second interpolation technique presented is the cubic B-spline free-form deformation (FFD) model. FFD, introduced by [109], deforms an object by manipulating a regularly subdivided 3D parallelepiped lattice containing the object. By manipulating a mesh of control points a deformation function that specifies a new position for each point on the object is calculated. In this way the deformations of the FFD lattice are passed onto the object.

The FFD lattice is represented by an array of control points $P_{ijk}$. The control point values $P_{ijk}$ are the (x, y, z) coordinates of the imposed grid of control points. Thus the lattice grid is divided into $(3l+1)$ by $(3m+1)$ by $(3n+1)$ parts by letting $l$, $m$ and $n$ be the number of subdivisions along each of the three directions, $U$, $V$ and $W$.

Points which lie within the grid subdivision can be formulated as a sum of control points weighted by polynomial basis functions as below,

$$T(u, v, w) = \sum_{i,j,k=0}^{3} B_i(u)B_j(v)B_k(w)P_{ijk}$$

(3.21)
In the above equation $0 \leq u,v,w \leq 1$ and $B_i(u)$, $B_j(v)$ and $B_k(w)$ are defined as the uniform cubic B-spline basis functions evaluated at $u$, $v$ and $w$, respectively.

The main objective of the FFD based deformation technique is to utilise the scattered object point translations and to configure the control grid points such that the deformed location of the selected point matches the target point location.

This is achieved in a least squares sense, based on the manipulation algorithm proposed by [110], with additions to enable multiple point movements.

The deformed object point can be written as below,

$$ T = BP $$

where $B$ is a single row matrix of the blending functions, and $P$ is an array whose rows are the control point coordinates.

Thus a new deformed object point can be written as, where the task is to identify $\Delta p$.

$$ T_{\text{new}} = B(P + \Delta P) $$

This required change in position of the control grid points based on the movement of the target point can be expressed as below,

$$ \Delta P = B^+T_{\text{new}} $$

where $B^+$ is the pseudo-inverse of $B$ represented as below,

$$ B^+ = \frac{1}{\|B\|^2}B^T $$

Thus the identified $\Delta p$ can be added to the undisplaced position of $P$, and the FFD algorithm run.

The algorithm iteratively refines and deforms the 3D generic model by continuously performing correspondence estimations, creating the 2D planar translational field (anterior
and lateral) and deforming the 3D generic model (Figure 3.12). Typically the algorithm is run for a fixed number of iterations or till it converges to a user specified tolerance error level.

![Figure 3.12 Illustration of the FFD lattice warping to deform the model femoral bone (b) to fit the target bone (c). In (a) the red femur is the model while the blue femur is the target.](image)

3.8 Experiments

In conducting experiments to validate the proposed reconstruction technique the clinical requirements of the two cases were considered. Hip and knee arthroplasty requires the visualisation of the femur, tibia and possibly the iliac (hip). Femur osteotomy would require the visualisation of the deformed femur, or more specifically the region of deformation. The sections below highlight the experiments conducted to identify the reconstruction framework’s success in reconstructing these bony anatomies.

3.8.1 Hip and Knee Arthroplasties

The three bones of interest for this surgical case are the femur, tibia and iliac. Thus to validate the reconstruction framework, six CT scanned cadaveric femurs (three pairs) and the associated x-ray images in the anterior and lateral view points were utilised in a series of tests. The x-ray images were used for the reconstruction, and the corresponding CT scan data
for accuracy assessment. Table 3.1, Table 3.2 and Table 3.3 shows the quantitative results obtained from the six tests performed for each bony anatomy. The key accuracy measurement used during the testing is the Euclidean distance between the closest points of the reconstructed data and the ground truth CT scan data set. These results are within the accuracy requirements set by the authors and several other customisation studies [19, 70].

Table 3.1 Intact femur reconstruction results.

<table>
<thead>
<tr>
<th>Models Compared</th>
<th>Maximum Positive Error (mm)</th>
<th>Maximum Negative Error (mm)</th>
<th>Average Positive Error (mm)</th>
<th>Average Negative Error (mm)</th>
<th>Standard Deviation (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right 1/ Right 2</td>
<td>1.08</td>
<td>-0.91</td>
<td>0.22</td>
<td>-0.16</td>
<td>0.085</td>
</tr>
<tr>
<td>Right 1/ Right 3</td>
<td>1.12</td>
<td>-1.03</td>
<td>0.31</td>
<td>-0.27</td>
<td>0.075</td>
</tr>
<tr>
<td>Right 2/ Right 3</td>
<td>1.02</td>
<td>-1.16</td>
<td>0.29</td>
<td>-0.21</td>
<td>0.061</td>
</tr>
<tr>
<td>Left 1/ Left 2</td>
<td>0.95</td>
<td>-1.01</td>
<td>0.34</td>
<td>-0.23</td>
<td>0.094</td>
</tr>
<tr>
<td>Left 1/ Left 3</td>
<td>1.26</td>
<td>-0.98</td>
<td>0.32</td>
<td>-0.29</td>
<td>0.086</td>
</tr>
<tr>
<td>Left 2/ Left 3</td>
<td>1.09</td>
<td>-0.93</td>
<td>0.29</td>
<td>-0.25</td>
<td>0.081</td>
</tr>
</tbody>
</table>

Table 3.2 Intact tibia reconstruction results.

<table>
<thead>
<tr>
<th>Models Compared</th>
<th>Maximum Positive Error (mm)</th>
<th>Maximum Negative Error (mm)</th>
<th>Average Positive Error (mm)</th>
<th>Average Negative Error (mm)</th>
<th>Standard Deviation (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right 1/ Right 2</td>
<td>2.29</td>
<td>-0.40</td>
<td>0.60</td>
<td>-0.07</td>
<td>0.11</td>
</tr>
<tr>
<td>Right 1/ Right 3</td>
<td>1.20</td>
<td>-0.54</td>
<td>0.34</td>
<td>-0.11</td>
<td>0.08</td>
</tr>
<tr>
<td>Right 2/ Right 3</td>
<td>1.61</td>
<td>-1.06</td>
<td>0.34</td>
<td>0.36</td>
<td>0.11</td>
</tr>
<tr>
<td>Left 1/ Left 2</td>
<td>2.84</td>
<td>-0.50</td>
<td>0.59</td>
<td>-0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>Left 1/ Left 3</td>
<td>1.51</td>
<td>-0.52</td>
<td>0.61</td>
<td>-0.09</td>
<td>0.10</td>
</tr>
<tr>
<td>Left 2/ Left 3</td>
<td>1.38</td>
<td>-0.76</td>
<td>0.40</td>
<td>-0.15</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 3.3 Intact iliac reconstruction results.

<table>
<thead>
<tr>
<th>Models Compared</th>
<th>Maximum Positive Error (mm)</th>
<th>Maximum Negative Error (mm)</th>
<th>Average Positive Error (mm)</th>
<th>Average Negative Error (mm)</th>
<th>Standard Deviation (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right 1/ Right 2</td>
<td>2.33</td>
<td>-0.17</td>
<td>0.38</td>
<td>1.64</td>
<td>0.15</td>
</tr>
<tr>
<td>Right 1/ Right 3</td>
<td>3.04</td>
<td>0.86</td>
<td>1.64</td>
<td>1.21</td>
<td>0.17</td>
</tr>
<tr>
<td>Right 2/ Right 3</td>
<td>1.71</td>
<td>0.33</td>
<td>1.72</td>
<td>0.60</td>
<td>0.08</td>
</tr>
<tr>
<td>Left 1/ Left 2</td>
<td>2.57</td>
<td>0.44</td>
<td>0.68</td>
<td>0.01</td>
<td>0.16</td>
</tr>
<tr>
<td>Left 1/ Left 3</td>
<td>3.23</td>
<td>-0.35</td>
<td>2.15</td>
<td>0.80</td>
<td>0.16</td>
</tr>
<tr>
<td>Left 2/ Left 3</td>
<td>1.53</td>
<td>-0.09</td>
<td>1.66</td>
<td>0.28</td>
<td>0.24</td>
</tr>
</tbody>
</table>
Figure 3.13 Several examples of the reconstruction framework being applied to long bones. The blue model represents the target model (patient-specific model) while the red model was customised to fit that target utilising only the anterior and lateral projections. From top to bottom: (a) femur reconstruction; (b) tibia reconstruction.

Figure 3.13 and Figure 3.14 shows several qualitative results obtained from the one of the six tests performed for each anatomy. The colour bar indicates the absolute error in mm. Figure 3.15 and Figure 3.16 provide further measures on the customisation accuracies for several of the tested cases. Figure 3.16 is of the Left 1/Left 3 femur reconstruction, which inhabited the worst results in the testing conducted. Figure 3.15 demonstrate the convergence of the deformation and shows the reconstruction error minimisation through successive iterations for the Left 1/Left 3 anatomical reconstructions. For example Figure 3.15(a) indicates an initial maximum absolute error of 10.9mm and a final maximum absolute error
of 1.26mm. In all testing performed the generic bone model was successfully converged to the final patient-specific shape.

Figure 3.14 Reconstruction framework being applied to an Iliac. This is an example of an irregular bone with little cross-sectional shape similarity.

Figure 3.15 Three examples of reconstruction error minimisation through successive iterations. From left to right: (a) reconstruction iterations of the Left 1/ Left 3 femur model customisation; (b) reconstruction iterations of the Left 1/ Left 3 tibia model customisation; (c) reconstruction iterations of the Left 1/ Left 3 iliac model customisation.
Chapter 3 - Pre-Operative Planning: Intact Bone Applications

3.8.2 Femur Osteotomy

For femur osteotomy, the bone of interest is the deformed femur (or that region). Thus to validate the reconstruction framework, three cases of bone deformity were tested, femoral shaft deformity (Bowing), femoral head deformity (Shepherd’s Crook), femoral condyles deformity (Erlenmeyer Flask) (Figure 3.17).

![Figure 3.17](image)

Figure 3.17 Examples of the deformed bones used in the reconstruction testing. From left to right: (a) femoral head deformity; (b) femoral condyles deformity; (c) femoral shaft deformity.

Two CT scans of each type of deformity and the associated x-ray images in the anterior and lateral view points were utilised in the tests. The x-ray images were used for the reconstruction and the corresponding CT scan data for accuracy assessment. Table 3.4 shows...
the quantitative results obtained from the six tests performed. Figure 3.18 shows the qualitative results obtained from the one of the six tests performed. The key accuracy measurement used during the testing is the Euclidean distance between the closest points of the reconstructed data and the ground truth CT scan data set. A discussion of the results is presented below.

Table 3.4 Osteotomy surgical planning 3D femur model reconstruction results.

<table>
<thead>
<tr>
<th>Deformity Type</th>
<th>Maximum Absolute Error (mm)</th>
<th>Average Absolute Error (mm)</th>
<th>Standard Deviation (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Femoral Shaft</td>
<td>1.31</td>
<td>0.92</td>
<td>0.093</td>
</tr>
<tr>
<td>Femoral Shaft</td>
<td>1.69</td>
<td>1.12</td>
<td>0.089</td>
</tr>
<tr>
<td>Femoral Head</td>
<td>1.86</td>
<td>0.99</td>
<td>0.093</td>
</tr>
<tr>
<td>Femoral Head</td>
<td>2.36</td>
<td>1.14</td>
<td>0.094</td>
</tr>
<tr>
<td>Femoral Condyles</td>
<td>2.05</td>
<td>1.05</td>
<td>0.089</td>
</tr>
<tr>
<td>Femoral Condyles</td>
<td>2.56</td>
<td>1.21</td>
<td>0.091</td>
</tr>
</tbody>
</table>
Figure 3.18 Several examples of the reconstruction framework being applied to deformed femora. The blue model represents the target model (patient-specific model) while the red model was customised to fit that target utilising only the anterior and lateral projections. From top to bottom: (a) femoral condyles deformity; (b) femoral head deformity; (c) femoral shaft deformity.
3.8.3 Discussion

For the femur bone reconstruction the maximum positive individual Euclidean error identified was 1.26mm and the maximum negative individual error identified was -1.16mm. Results indicated by [19] convey an average mean surface reconstruction error of 2.4mm. This reconstruction targeted the proximal region of the femur. Furthermore work performed by [70] on distal femur reconstruction has indicated a mean RMS error of 3.27mm. Thus the results presented in this chapter are within the accuracy requirements set by the authors and several other customisation studies. Compared to other studies involving CT image 3D reconstructions, the current accuracy of the customised models is comparable to the mean error of 1 mm reported in the literature [62, 63].

The favourable results are due to the difference in the a-priori model used in the studies. Work conducted by [19, 70, 77] mainly utilised statistical shape models that place geometrical constraints on the deformation process. This means that complete local shape deformation is not possible, if that deformation falls outside the statistical constraints this would no-doubt add further error to the reconstruction. In this study the generic models were generated from averaging the contours of six femora, thus they contained ample a-priori information and were much faster to generate (in comparison to a statistical model). A benefit of using a generic model based technique is that the method can be easily extended to other bony anatomies, as has been highlighted with the tibia and the iliac.

During the testing, the maximum errors were clustered around the condyles for the distal fragment and the greater and lesser trochanters for the proximal fragment. This was due to the anatomical variability in these areas.

The main reasoning behind the errors is the ill-posed problem of attempting to reconstruct 3D models based on bi-planar 2D radiographic images. Making use of a-priori generic models provide cross-sectional shape constraints. However there are still certain patient-
specific cross-sectional variations that cannot be integrated with the use of generic models. These variations cause a majority of the reconstruction errors.

Long bones such as the femur and the tibia are cross-sectionally predictable and the generic model encompasses much of the axial information. However irregular bones such as the iliac (and vertebrae and mandible) have non uniform shapes that are not predictable axially. The results convey this, with the higher errors associated with the iliac reconstruction testing.

Another source of error is the bone segmentation inaccuracies and the error in defining of bone contours. Bone diameters are typically underestimated as femur x-ray edges are blunt due to bone roundness. Femoral diameter (cortical boundary) estimation errors are typically between 0.4 to 1.0 mm [78]. These erroneous contours that are extracted cause inaccuracies when used for shape customisation.

3.9 Extension: Inclusion of Interior Edges

The shape customisation algorithm documented in Section 3.6 utilises the outer contours of the projected model for the non rigid registration. An extension to this procedure would be to not only utilise the outer contours but also the inner contours. This would improve the reconstruction accuracies of the condyles and the trochanters. However, please note that these extensions were not included in as part of the main contribution because they lacked testing and robustness in identifying the interior edges.

To include the inner contours in the registration two minor changes are required. Firstly, the generic models require certain automated pre-processing, to identify the inner surface edge points that would be clearly identifiable on radiographic images. These interior contours on the generic model can be identified utilising the saddle points on the surface curvature (Figure 3.19(a)).
Chapter 3 - Pre-Operative Planning: Intact Bone Applications

Figure 3.19 Example of the clearly identifiable edges on femur radiographic images. From left to right: (a) automated 3D surface edge point extraction applied to a femur. The blue lines form the outer contour while the green lines form the interior edges in the bi-planar radiographic images; (b) edge extraction technique applied to actual femoral radiographic images.

As an example, the femoral generic model is shown in Figure 3.19(a) along with the outer and inner contours. Accurate identification of as many of these edges as possible on radiographic images is vital as the customisation process is driven by them, thus a new edge extraction algorithm is proposed (Figure 3.20). This is the second change that would be required to extract the inner contours from the patient x-ray images.

Figure 3.20 Non-rigid registration on the sub-sampled edge contour data points for the anterior/lateral views and the associated translational field displayed on the generic model.
Chapter 3 - Pre-Operative Planning: Intact Bone Applications

The proposed edge extraction technique has a similar objective as the one introduced in Section 3.5, however due to the multitude of edges that will need to be identified manual interactive segmentation is no longer an option. This new technique is briefly discussed below, but a further detailed description is available in Section 5.2.

Many of the classic first (Roberts, Prewitt, Sobel) and second order (LOG) derivative based edge segmentation methodologies extract isolated edge pixels and do not provide continuous edge contours. The proposed edge extraction technique attempts to link sets of potential edge pixels to create continuous boundaries. It employs two tactics to achieve this: thresholding with hysteresis and edge relaxation (Figure 3.19(b)).

Firstly, adaptive (local) hysteresis based thresholding is performed on the gradient image, where the two threshold values are set to be the 50th and 75th percentile of the gradient magnitude values in the local window. Adaptive thresholding adjusts the threshold level according to the intensity statistics of a local region. This technique is employed typically with x-ray/fluoroscopic images to counter the illumination gradients present on the radiographs.

Secondly, edge relaxation is performed through a region growing exercise. This region growing is conducted on the intermediate pixels (pixels that fall between the 50th and 75th percentile of gradient magnitude values) to ensure sufficient continuity of the edge contour. Edge relaxation involves the recursive re-labelling of intermediate pixels with one or more neighbouring edge pixels, utilising an eight neighbouring connectivity scheme. In order to be reclassified as a foreground edge pixel the difference of the magnitude and orientation of the intervening pixel with the surrounding edge pixels will be checked and has to be within a user specified tolerance level. Following the edge relaxation, a small component elimination morphological operation is performed to remove all connected pixel components with too few pixels. This is a noise elimination step that will remove any misidentified edge contour pixels.
Chapter 3 - Pre-Operative Planning: Intact Bone Applications

With the inclusion of the interior contours the testing was redone for the femur. Table 3.5 shows the quantitative results of the test. The reconstructions error was less than the previous testing conducted and the error spread was much more uniform throughout the femoral head and the condyles region (Figure 3.21).

Table 3.5 Patient-specific intact bone model reconstruction results in terms of mean and standard deviation of error.

<table>
<thead>
<tr>
<th>Models Compared</th>
<th>Maximum Positive Error (mm)</th>
<th>Maximum Negative Error (mm)</th>
<th>Average Positive Error (mm)</th>
<th>Average Negative Error (mm)</th>
<th>Standard Deviation (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right 1/ Right 2</td>
<td>1.02</td>
<td>-1.08</td>
<td>0.25</td>
<td>-0.24</td>
<td>0.075</td>
</tr>
<tr>
<td>Right 1/ Right 3</td>
<td>1.13</td>
<td>-1.04</td>
<td>0.29</td>
<td>-0.25</td>
<td>0.095</td>
</tr>
<tr>
<td>Right 2/ Right 3</td>
<td>1.04</td>
<td>-0.98</td>
<td>0.28</td>
<td>-0.31</td>
<td>0.083</td>
</tr>
<tr>
<td>Left 1/ Left 2</td>
<td>1.12</td>
<td>-1.05</td>
<td>0.31</td>
<td>-0.20</td>
<td>0.087</td>
</tr>
<tr>
<td>Left 1/ Left 3</td>
<td>1.13</td>
<td>-1.08</td>
<td>0.29</td>
<td>-0.26</td>
<td>0.075</td>
</tr>
<tr>
<td>Left 2/ Left 3</td>
<td>1.01</td>
<td>-0.98</td>
<td>0.29</td>
<td>-0.21</td>
<td>0.071</td>
</tr>
</tbody>
</table>

Figure 3.21 Comparison of the accuracy of the reconstruction technique that utilise only the outer contour (a) and the technique that utilise both the inner and outer contours (b). The colour bar indicates the absolute error in mm.
3.10 Chapter Summary

This chapter demonstrated that using a bi-planar shape customisation framework combined with several image processing methodologies can yield acceptable results for patient-specific femur reconstruction. The main criteria required when developing a visualisation (reconstruction) methodology for pre-operative planning is the accuracy of reconstruction, adaptability to multiple orthopaedic cases and the level of manual intervention required. The time taken is not of a main concern during this pre-operative customisation phase, since ample time is typically available prior to surgery. The aforementioned results have indicated that the proposed method has a high accuracy of reconstruction and is adaptable to a multitude of bony anatomies.

A novel registration framework that not only involves the outer contours of the bony anatomy in the reconstruction but also several key interior edges was proposed to enhance customisation accuracies and to ensure adaptability to other bony anatomies. The method is based on an iterative non-rigid 2D point matching process and thin-plate spline based deformation. The developed non-rigid registration is robust against disturbances (noise), deformation, outliers and occlusions in the radiographic images.

We have validated the proposed reconstruction framework, through a series of tests conducted with the aid of CT scan data. The accuracy of the reconstruction was adequate for orthopaedic surgical planning purposes, which is the target application of the proposed approach.

The proposed bi-planar method has the potential to be used in shape customisation of even more difficult bones and to be used in other applications (arthroplasties, osteotomy, etc). Previews of such applications were shown briefly in the results section. However, as previously mentioned the work presented here is used as a precursor to create 3D fractured bone models to be used intra-operatively by a robotic fracture reduction device. Thus there are several further stages subsequent to the shape customisation described here that will
Chapter 3 - Pre-Operative Planning: Intact Bone Applications

incorporate a fracture into the customised bone and will perform a 3D pose-estimation intra-operatively. The integration of this whole system will provide enhanced 3D image guidance to current orthopaedic procedures that rely on 2D fluoroscopic images for intra-operative assistance.
Chapter 4 Pre-Operative Planning: Fracture Reconstruction

The proposed software tool allows orthopaedic surgeons to visualise, diagnose, plan and simulate femur shaft fracture reduction procedures in 3D. The tool utilises frontal and lateral 2D radiographs to model the fracture surface, separate a generic bone into the two fractured fragments, identify the pose of each fragment, and automatically customise the shape of the bone. The use of 3D imaging allows full spatial inspection of the fracture providing different views through the manipulation of the interactively reconstructed 3D model, and ultimately better pre-operative planning.

The results obtained through the proposed methodology are benchmarked against 3D CT scan data to assess the accuracy of reconstruction. Tests conducted conveyed a sub-millimetre average reconstruction error.

This chapter presents work conducted on a new approach to pre-operative planning of femoral fracture surgery and has been published in [32-36].

4.1 Introduction

The femur (thigh bone) is the largest and longest bone in the human body and is essential for ambulation. Femur fractures must be treated accurately and swiftly to ensure rapid rehabilitation and minimal gait impediments. Pre-operative planning plays an essential role in the management of such injuries because many of the technical problems that may arise during surgery can be anticipated during this preparatory phase.

The pre-operative evaluation of the geometry and position of the fractured femur in current clinical practice is based on 2D x-rays acquired in the frontal and lateral viewpoints, where the surgeon has to mentally visualise the anatomy of interest. This is primarily owing
to direct 3D imaging modalities, such as CT, being restricted to only complex orthopaedic procedures due to their lower availability, increased radiation exposure and higher costs. Literature suggests, planning femur fracture surgical procedures on the basis of 2D images alone produces errors leading to costly secondary surgery or irreversible biomechanical defects [111]. Frontal and lateral malrotations cause not just cosmetic problems, but also biomechanical ambulation problems and changes to the weight-bearing axis, thus prematurely damaging the bone. A study conducted on femur fracture reduction confirmed that computer-aided systems can significantly improve the accuracy of orthopaedic procedures by augmenting the current 2D image guidance with an interactive display of 3D bone models [111]. This research indicated that the positioning errors generate bone misalignments and complications (about 18% of femur fracture reduction cases) which could be reduced with the introduction of 3D bone fragment visualisation during surgical procedures. Consequently 3D visualisation of anatomy plays an important role in image-guided orthopaedic surgery and most importantly contributes to minimally invasive procedures.

There are two major concerns with the current use of 2D imaging in the pre-operative planning process. Firstly, with 2D images it is difficult to visualise the bone fragment pose and the reduction path. Thus, a surgeon needs to mentally visualise and simulate the surgical reduction, which requires extensive experience with previous cases. Secondly, implant requirements are difficult to assess with 2D images. For instance, custom implants maybe necessary if a patient falls outside of the normal bone size range. Short or long femora, variations of neck shaft angle and excessive bone curvature are some of the problems that may require a change or delay in surgery to obtain the proper implant. These details are much clearly visible with 3D models than through planar 2D imaging. Ultimately, the surgeon should emerge from the planning phase with a firm idea of the patient's femoral dimensions and the specific implant to be used.
Orthopaedic teaching institutes have long identified the visualisation difficulty and 3D spatial skills required by surgeons to mentally reconstruct the pose of each of the fragments [31]. Mental reconstruction is a human component of spatial thinking that involves imagining the movement of objects, which requires broad experience with previous cases. For femur shaft fracture, there are three critical factors that make mental reconstruction a difficult process.

1) Mental reconstruction requires unique defining features that the observer can utilise for correspondence identification between the frontal and lateral images. This is however difficult with femur shaft fractures as typically only the cylindrical femoral shaft is visible (beside the fracture region), through the fluoroscopic c-arm. There are no key bony features that the surgeon will be able to utilise for mental pose estimation and reconstruction.

2) The surgeon needs to visualise the pose of both the proximal and distal segments. Thus the relative motion between both segments must be identified. Typically this problem is lessened by ensuring that the proximal bone does not move significantly during the procedure. Nevertheless, the need to identify the relative position of the two bone fragments complicates the mental reconstruction.

3) In order for a person to be able to continuously visualise an object in 3D by utilising only 2D images, ideally the same object must be at least seen once in a 3D spatial environment. This aids the mind to build a static image of the 3D object which can then be modified through continuous 2D imaging. Thus in a surgical environment, one solution would be to enable the surgeon to pre-operatively visualise a 3D model of the broken bone and then to update this mental visualisation intra-operatively through the 2D fluoroscopic images. However this 3D pre-operative visualisation is again not possible with femoral shaft fractured due to direct 3D imaging modalities, such as CT, being restricted to only complex orthopaedic procedures.
Chapter 4 - Pre-Operative Planning: Fracture Reconstruction

The aforementioned constrains with the current 2D pre-operative planning process and the resulting problems that arise during the surgery motivated the work presented in this chapter. The research proposes a procedure which allows the orthopaedic surgeon (or the radiologist) to interactively reconstruct patient-specific 3D models of a fractured femur. This is achieved by utilising bi-planar 2D x-ray images and a generic 3D model of the femur.

The proposed interactive framework to achieve 3D reconstruction from orthogonal 2D images is outlined in Figure 4.1.

![Figure 4.1 Proposed femur fracture reduction pre-operative planning framework.](image)

As conveyed in Figure 4.1, all patient-specific fracture details required for the framework are acquired from 2D x-ray images. Firstly, these pre-operative 2D images are processed to segment the bone contours. Two approaches to perform this segmentation are proposed: an automatic active contour based segmentation and an interactive LiveWire technique. Next, fracture modelling is performed to identify the fracture surface detail and to aid the separation of the generic intact bone. The fracture surface has to be accurately modelled for the pose estimation to correctly separate the fractured bone into the fragments. This chapter introduces an interactive feature-based modelling technique. The user must identify several corresponding points between the frontal and the lateral images to obtain 3D points which define the key landmarks on the fracture surface. Once identified, these 3D points define a model which approximates the fracture surface. Next, pose estimations and shape customisations are iteratively performed until the user is satisfied with the 3D visualisation presented. The pose estimation (2D-3D registration based) separates the bone into two fractured segments and identifies their respective pose (position and orientation). The shape customisation involves a non-rigid registration between the edges identified in the front and lateral x-ray image and the projected contour points of the generic model. A registration
process identifies planar point correspondences. They are interpolated to create a 3D translational field. The 3D generic anatomical data is deformed accordingly. The aforementioned stages are further detailed below.

### 4.2 Contour Extraction

The contour extraction used for the work presented in this chapter is detailed in Section 3.5. The proposed LiveWire technique is interactive and allows the user to specify the final segmentation quality. This is vital as for femur fracture pre-operative planning. The sequence of images in Figure 4.2 shows the interactive LiveWire segmentation technique applied to a frontal and lateral femur x-ray image.

![Figure 4.2](image.png)

**Figure 4.2** Two examples of the interactive x-ray segmentation process. From left to right: The user defines lines to perform the segmentation. Time taken for segmentation: (a) 20s, (b) 25s.

Apart from the LiveWire technique illustrated above the active contours technique detailed in Section 3.5 can also be used for the segmentation. This automated segmentation
technique (Figure 4.3) is not as robust as the interactive LiveWire procedure, as it is sensitive to initialisation.

Figure 4.3 An example of the Active Contours based x-ray segmentation process. From left to right: (a) original image; (b) the level set contour after 20 iterations; (c) the level set contour after 50 iterations resulting in the final segmented image.

4.3 Fracture Modelling

Fracture modelling is the first step of the fracture incorporation process. The aim is to model the surface of the fracture utilising only the frontal and lateral x-ray images. The complete fracture incorporation process is a combination of fracture modelling as well as fragment pose estimation (Section 4.4). The process aims to separate and identify the pose of the distal and proximal fragments.

The 3D fracture surface modelling is performed through a semi autonomous method. First, the user is required to manually identify several corresponding key feature points in both the frontal and lateral views that lie on the fracture surface of either the proximal or the distal segment. These manually identified landmarks are then utilised to create a series of 3D point sets that identify key landmarks on the fracture surface. The creation of the 3D
landmark points is a simple process, where the \( x-z \) coordinates are extracted from the frontal image and the \( y-z \) coordinates are extracted from the lateral image.

Following the landmark point identification, an interpolation is performed to fit a surface through the points (see Figure 4.4(b) and Figure 4.4(c)). This surface fitting is achieved through a thin plate spline (TPS) interpolation. The TPS function \( s(x, y) \) interpolates a surface that is fixed at the key feature points at a specific height. The TPS surface fitting minimises the bending energy (integral of the square of the second order derivatives of the mapping function) of the surface.

\[
E = \int \left[ \left( \frac{\partial^2 s}{\partial x^2} \right) + 2 \left( \frac{\partial^2 s}{\partial x \partial y} \right) + \left( \frac{\partial^2 s}{\partial y^2} \right) \right] dx \, dy \tag{4.1}
\]

The landmark identification step was not automated as this directly dictates the accuracy of the reconstruction and the pose estimation that is followed. Thus to ensure a robust and accurate fracture surface model it was deemed that a manual identification process was most suited.

The calibration of the x-ray images required to obtain the mm dimensions was performed utilising the Othomark calibration ball [112]. This calibration ball with a known diameter was placed in the frame of the x-ray image and the required intrinsic calibration parameters obtained through manual measurements.

Illustrations of corresponding point matching as well as an example of fracture surface reconstruction are displayed in Figure 4.4 below. Figure 4.4(a) shows a series of 3D points, identified interactively through the frontal and lateral images, which form the key landmarks on an example fracture surface.
4.4 Pose Estimation

The pose estimation framework employs the pair of 2D x-ray images, the corresponding 3D anatomical model and the 3D fracture surface. It aims to identify the transformation of the model so that its projections on the frontal and lateral planes match the acquired x-ray images. This registration can be considered as determining the equivalent affine transformation which includes a set of translational \((T_x, T_y, T_z)\), and rotational \((R_x, R_y, R_z)\) parameters where the \(x\), \(y\) and \(z\) axis are aligned with the frontal, lateral and axial axis of the bone segment.

The steps involved in the pose estimation between a model and a set of x-ray images are illustrated in Figure 4.5 and can be summarised as: 1) the creation of the frontal and lateral projections of the generic 3D model; 2) the evaluation of the visual similarity measure between projected contours above and the segmented contours from the patient’s x-ray images.

The user interactively identifies the pose of the 3D bone model. The similarity measure used for the alignment is the distance between the projected contours of the generic model.
and the segmented contours of the x-ray image. Once these contour lines are approximately overlaid then the pose of the bone has been successfully identified.

There are seven parameters involved in the pose estimation, which include control over the six degrees-of-freedom movement as well as the centre of rotation (Figure 4.5(a)). Typically, the centre of rotation for the proximal segment is the femoral head while it is the Condyles for the distal segment. The homogeneous transformation matrices that combine the rotations, $R$, translations, $t$, and the centre of rotation, $x_R$, are given in the equation below. Where, $[x_i, y_i, z_i]^{T}$ represents the original 3D model points (prior to transformation) and $[x_i', y_i', z_i']^{T}$ the transformed points.

$$
\begin{pmatrix}
    x_i' \\
    y_i' \\
    z_i'
\end{pmatrix} =
\begin{pmatrix}
    I_3 \\
    0 \\
    0
\end{pmatrix}
\begin{pmatrix}
    x_R \\
    1
\end{pmatrix}
\begin{pmatrix}
    R & t \\
    0 & 1
\end{pmatrix}
\begin{pmatrix}
    I_3 \\
    0 \\
    0
\end{pmatrix}
\begin{pmatrix}
    0 \\
    1 \\
    0
\end{pmatrix}
\begin{pmatrix}
    x_i \\
    y_i \\
    z_i
\end{pmatrix} - x_R
\begin{pmatrix}
    1
\end{pmatrix}
$$

(4.2)

The user has the ability to set numerical values or to use the mouse interactively to update the pose of the bone. When using numerical values, the user will initially set the centre of rotation and the six transformation parameters. For convenience, the centre of rotation can be set utilising the frontal and lateral x-ray images (similar to the fracture modelling key 3D landmark identification stage). When using the mouse, the user initially clicks on the centre of rotation and can move the mouse to rotate the model interactively. The translational values are however still set numerically as these additional degrees-of-freedom cannot be effectively expressed through a 2D input device such as the mouse. The pose estimation process is conducted alongside the shape customisation described next, which will attempt to ensure that the shape of the generic bone is adapted to be close to the patient-specific bone.
Figure 4.5 Illustrations of the pose estimation stage. From left to right: (a) controllable six degrees-of-freedom of the proximal segment; (b) previously modelled fracture surface included in the pose estimation.

4.5 Shape Customisation

A non-rigid registration between the 2D projective contours of the 3D generic model and the extracted edges of the patient-specific x-ray images is performed, to customise the shape of the generic bone to that of the patient. The registration is performed through a point correspondence estimation between the two point sets. Here, unlike Section 3.6, the point correspondence is driven only through the topology based measure (Shape Context). For comprehensiveness the similarity measure is again briefly detailed below (the shape customisation technique is fully detailed in Section 3.6).

\[
C_{\text{shape}}(p_m,p_t) = \frac{1}{2} \sum_{n=1}^{N} \frac{(h_m(n)-h_t(n))^2}{h_m(n)+h_t(n)}
\] (4.3)

Where \( C_{\text{shape}} \) is the shape context based cost function to match a point \( p_m \) on the model contour to a point \( p_t \) on the target contour. The total cost of matching the point sets can be minimised through a Bi-partite graph searching methodology. The minimisation of the total cost matrix between point sets is subject to the constraint that the matching is one-to-one. This square assignment problem is solved through the Hungarian algorithm [108].
Chapter 4 - Pre-Operative Planning: Fracture Reconstruction

The correspondences identified through the Hungarian algorithm are subsequently processed and filtered to remove any outliers (misidentified correspondences). Subsequent to the planar interpolation, a 3D translational field to deform the generic model is created. The full 3D deformation is based on a cubic B-spline free-form deformation model. By manipulating a mesh of control points a deformation function that specifies a new position for each point on the object is calculated. In this way the deformations of the FFD lattice are passed onto the object. Figure 4.6 shows an example full femur customisation through FFD based interpolation.

![Figure 4.6 Illustrations of the shape customisation stage. From left to right: (a) non-rigid registration on the sub-sampled edge contour data points for the frontal/lateral views. The blue points are of the generic model projections while the red points are segmented contour points from the x-ray images; (b) TPS based interpolation of the transformation field to deform the 3D generic model. The blue model was customised to the shape of the red patient-specific model.](image)

**4.6 Experiments**

Several tests were performed to validate the developed pre-operative planning framework. A femoral CT scan data set was employed and synthetically broken with a variety of femur fracture types (transverse / oblique / spiral fractures). The CT data was utilised to represents a patient with a broken femur. Digitally reconstructed radiographs (DRRs) were created in the frontal and lateral viewpoints to provide the x-ray imaging. These digitally constructed x-ray images were highly realistic and were visually analogous to those obtained clinically on human subjects (see Figure 4.7). The reconstruction was based purely on the DRRs created.
Chapter 4 - Pre-Operative Planning: Fracture Reconstruction

The use of synthetically broken CT data enables the testing of the proposed framework against a large number of different fracture types. Subsequently, the CT data will provide a gold standard in 3D reconstruction which can be compared against the reconstructed model. This type of accuracy benchmarking is not possible with clinically available 2D x-ray images.

![Figure 4.7](image)

Figure 4.7 Three qualitative illustrations of testing performed. The left image of each pair shows one of the two x-ray images (DRRs) used in the fractured femur reconstruction. The image on the right shows the reconstructed fracture region after fracture modelling, pose estimation and shape customisation.

The aim of the testing was to attempt the reconstruction a 3D model of the fractured bone, through the proposed framework, by utilising only 2D DRR imaging. Following reconstruction, the accuracy was assessed in comparison to the CT data. Testing was conducted for three common variations of femur shaft fracture types: transverse, oblique and spiral. Several qualitative results of the testing performed are shown in Figure 4.7. The quantitative results are shows in Table 4.1. The key accuracy measurement used during the testing is the average Euclidean distance between the closest points of the reconstructed data and the ground truth CT scan data set.
Chapter 4 - Pre-Operative Planning: Fracture Reconstruction

Table 4.1 Patient-specific fractured bone model reconstruction results in terms of mean and standard deviation of error.

<table>
<thead>
<tr>
<th>3D Fractured Bone Reconstruction Error</th>
<th>Transverse Fracture Mean / SD (mm)</th>
<th>Oblique Fracture Mean / SD (mm)</th>
<th>Spiral Fracture Mean / SD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>1.75 / 0.45</td>
<td>2.05 / 0.55</td>
<td>1.91 / 0.43</td>
</tr>
<tr>
<td>Test 2</td>
<td>1.89 / 0.32</td>
<td>1.97 / 0.44</td>
<td>2.13 / 0.51</td>
</tr>
<tr>
<td>Test 3</td>
<td>1.66 / 0.39</td>
<td>2.16 / 0.46</td>
<td>2.08 / 0.54</td>
</tr>
<tr>
<td>Test 4</td>
<td>1.92 / 0.49</td>
<td>2.12 / 0.51</td>
<td>2.23 / 0.49</td>
</tr>
</tbody>
</table>

SD: Standard Deviation

The results in Table 4.1 are within the accuracy requirements set by the authors and several other customisation studies [19, 77]. During the testing the maximum errors were clustered around the condyles for the distal fragment and the greater and lesser trochanters for the proximal fragment. This was due to the anatomical variability in these areas. These areas are however not vital from a clinical point of view (for femur shaft fracture surgery).

Apart from the aforementioned areas, errors seen elsewhere are attributed to several sources.

1) The utilisation of orthogonal 2D radiographic images for 3D model reconstruction is ill-posed mathematically. Making use of a-priori generic models provide cross-sectional shape constraints. However there are still certain patient-specific cross-sectional variations that cannot be integrated with the use of generic models. These variations cause a majority of the reconstruction errors. One possible resolution is shown through the work conducted by [77], where the femoral generic models is generated by averaging multiple femora.

2) Bone segmentation and the defining of bone contours is also a difficult task leading to reconstruction errors. Bone diameters are typically underestimated as femur x-ray edges are blunt due to bone roundness. Femoral diameter (cortical boundary) estimation errors are typically between 0.4 to 1.0 mm [78]. These erroneous contours that are extracted cause inaccuracies when used for shape customisation.
3) Pose estimation errors are also common due to the difference in the centre of rotation of each fragment and the actual clinical fracture break mechanisms. The centre of rotation for each of the proximal and distal segments have to be set prior to the pose estimation process, which is a difficult task and require understanding of human joint motion dynamics.

The chapter has until now discussed stages involved in creating a 3D model of the fractured bone, which is the main focus. There are however several key benefits of a 3D model with regard to pre-operative planning that must be highlighted. As previously discussed, planning plays an essential role in the management of orthopaedic injuries. Many of the problems that arise during surgery can be anticipated during the planning phase. There are two parts to pre-operative planning: fracture segment visualisation (for diagnostics) and surgical simulation. The first part, which has been shown thus far, will be achieved through a static 3D model, where the surgeon will be able to visualise the fracture segments in 3D. The second part requires a tool to allow the surgeon to virtually manipulate the fracture segments and to plan the reduction path that will be taken during surgery. Additionally, implant selection will be required to identify which implant(s) should be used on the fracture after alignment. Both fragment manipulation and implant selection: aid surgical simulation, help to correctly restore leg length (femoral offset) and successfully align and secure the fragments.

With the use of the pre-operatively reconstructed 3D bone models, surgeons are able manipulate the bone fragments, select digital templates from a library, and to electronically overlay them in 3D space. The surgeon can then perform the necessary measurements on the templates and do the stepwise pre-operative planning process in a computer assisted environment. Examples of this type of surgical simulation possibilities are shown in Figure 4.8.
Figure 4.8 The potential of 3D models to be used in surgical simulation and implant selection. The images show example implants inserted into the bone fragments after alignment. Please note that these images are not intended to represent exact femoral nailing practices. They are purely artificial examples to convey what can be achieved with 3D model based pre-operative planning. From left to right: (a) femoral intramedullary nailing; (b) femoral plate locking.

4.7 Chapter Summary

This chapter demonstrated that using bi-planar radiographs combined with several image processing methodologies could yield acceptable results for patient-specific fractured femur reconstruction. The main criteria required when developing a visualisation (reconstruction) methodology for pre-operative planning were the accuracy of reconstruction, the adaptability to multiple orthopaedic cases and the level of manual intervention required. The framework completion time was not deemed a major factor during this pre-operative phase. The authors have validated the proposed reconstruction framework, through a series of tests conducted with the aid of CT scan data. The aforementioned results exhibit the proposed method’s performance in reconstruction accuracy and adaptability to a multitude of fracture types.
Chapter 5 - Segmentation of Radiographic Images under Topological Constraints

Chapter 5 Segmentation of Radiographic Images under Topological Constraints

This chapter details a radiographic image segmentation algorithm which is conducted under topological control. The system is intended for use in common radiological tasks including fracture treatment analysis, osteoarthritis diagnostics, and osteotomy management planning. The segmentation framework utilises a generic three dimensional (3D) model of the bone of interest to define the anatomical topology. Non-rigid registration is performed between the projected contours of this generic 3D model and extracted edges of the x-ray image to perform the segmentation. For fractured bones, the segmentation requires an additional step, where a region based active contours curve evolution is performed with a level set Mumford-Shah method to obtain the fracture surface edge. The application of the segmentation framework to analysis of human femur radiographs was evaluated. The proposed system has two major innovations. First, the definition of the topological constraints does not require a statistical learning process, so the method is generally applicable to a variety of bony anatomy segmentation problems. Second, the methodology is able to handle both intact bone segmentation and fractures. The work presented in this chapter has been published in [38].

5.1 Introduction

Image segmentation is defined as the separation of the image into constituent regions and is a pre-processing step leading onto image analysis and interpretation. Advances in medical imaging over the past thirty years have made segmentation one of the major components in medical image processing. Segmented medical images can be used to assist and to automate several radiological tasks. For instance, a segmented medical image allows medical
practitioners to better visualise shapes and relative poses of internal structures and accurately measure the quantitative volumes and distances between the anatomies of interest.

The work presented in this chapter will utilise two dimensional x-ray and fluoroscopic radiographic images as the modality of interest and will focus on the human femur as the anatomy of interest. Two dimensional radiographs are a popular diagnostic modality, and provide sufficient pathological information for fracture treatment analysis, osteoarthritis (arthroplasties) diagnostics and osteotomy management planning. Furthermore, two dimensional radiographs are preferred over three dimensional CT scans due to their lower cost, wider availability and lower radiation exposure to the patient.

Characteristically, in segmentation algorithms the constituent regions are homogeneous with respect to some feature attribute such as pixel intensities, gradient magnitudes or texture [98]. Radiographic medical images however, do not conform to this homogeneity criterion, made by classic segmentation algorithms. For example the intensity values of radiograph based x-ray image are dependent on the radiation absorption of the anatomy. In the case of bony anatomy this intensity will vary proportionally to the thickness of the bone and thus create non-homogeneous intensity patterns within the same anatomy. Further issues commonly faced during the segmentation of medical images include poor contrast, ill-defined boundaries, noise, and acquisition artefacts (such as illumination gradients) [98]. Thus successfully medical image segmentation cannot be achieved utilising these gray-level features alone. Hence literature in medical imaging segmentation typically utilise a-priori knowledge of the region of interest and augments that knowledge with gray-level segmentation algorithms [113-117].

Methods presented in literature, to segment anterior and lateral femur radiographs can be classified into three categories: 1) Statistical atlas based (both 2D and 3D) methodologies (e.g. statistical active shape model [113-116, 118, 119]; 2) Rule-based segmentation methods used to detect specific features on the femur contour [117]; 3) Deformable contour based methodologies (e.g. level set and active contours) [32, 117].
Chapter 5 - Segmentation of Radiographic Images under Topological Constraints

Work conducted by [114] presents a 3D active shape model based framework for the extraction of bone contours from x-ray images. In this algorithm the model initialisation is performed through the estimation of the Bayesian Network Algorithm to fit a pre-formed geometrical component model to the x-ray data. A 2D/3D registration between the statistical model and the x-ray images is then employed to extract the femur contours. Literature work of [113, 115, 116] present ideas based on 2D active shape models. In [113], a probabilistic segmentation methodology formulates explicit models of the overlapping object appearance and shape geometry from training images which is then used to segment the femur and tibia in knee x-ray images. Previous work conducted by the same authors have employed a simpler double contour 2D active shape model in order to simultaneously segment anterior and posterior contours of the knee region. The algorithm presented by [116] initialises the model contour through a template matching process based on the cross correlation and then employs the 2D active shape model to segment the proximal part of the femur. Moreover, the rule based approach presented by [117] initially detects prominent features of the femur, including the femoral shaft lines, femoral head and the turning point near the base of the greater trochanter. The algorithm then registers the model contour to the x-ray image in accordance to the features and is then refined through an active contour algorithm.

As previously mentioned there are a multitude of clinical uses for segmented femur contours. Literature presented by [113, 120] employs segmented bone contours for osteoarthritis management and tracking. Here the contours are used to provide measurement of unimpaired joint space width. Furthermore, research performed by [116, 117] utilises femur contours for osteoporosis assessment. It is used as a pre-processing step for non-invasive bone mineral density estimation. An example of contour extraction for pose estimation is seen in the work conducted by [114]. This application of pose estimation offers great benefits to medical practitioners as they can now visualise 3D patient-specific bony models without the need to perform CT scans.
While existing methods above have shown to be effective, they have a number of limitations. The work presented in this chapter is motivated by these limitations as listed below.

1) Not adaptable: The development of a statistical model of the anatomy (either 2D or 3D) requires building and processing a large training dataset. Hence this technique is not easily and quickly adaptable to the segmentation of other structures. However the benefit of a statistical model is that it provides topological constraints which are vital for anatomically realistic segmentation. The segmentation technique developed by the authors have aimed to work around this constraint by replacing the statistical constraints with geometrical and shape information.

2) Too restrictive: Active shape model based methods that aim at direct segmentation of structures using strict topological constraints require accurate initialisation of the model contour. This is essential due to imperfections seen in medical images (noise, non-homogeneities etc) which will lead onto geometrical errors in the segmented contour. Thus the work presented has eliminated this restrictive need for accurate initialisation.

3) Limited usage: Statistical model and rule based segmentation algorithms have been employed in the segmentation of intact bones. However segmentation of fractured bones have not been performed under these topological constraints. The techniques developed by the authors work not only for intact bone segmentation but also for fractured bones.

The flow chart in Figure 5.1 and Figure 5.2 illustrates the stages involved in the segmentation framework. The 2D radiographic image is initially processed to extract the edge points that potentially form the femur boundary. The novel non rigid registration is then performed between the edges identified in the x-ray image and the projected contour points of the generic model. The registration has been developed to be robust to occlusions, outliers, disturbance (noise) and deformations. This algorithm has been extensive tested and benchmarked against other point correspondence algorithms available in literature. This
process can be used to segment an intact femur or the proximal and distal portions of a fractured femur. In case of a fractured femur, a region based active contours evolutions is subsequently employed to expand this initial contour along the shaft to segment the fractured shaft region. Each of these stages will be detailed in the proceeding sections.

![Diagram](image)

**Figure 5.1** The main process steps of the proposed segmentation framework.

![Images](image)

**Figure 5.2** Segmentation process applied to a femur with a shaft fracture (this was artificially fractured for illustrative purposes and is not meant to represent realistic x-ray images). From left to right: (a) original x-ray image; (b) edge point extraction; (c) extracted edges overlaid with the projected contour points. The red points are the extracted edge points (sub-sampled) and the blue points are the projected contour points; (d) non rigid registration; (e) level set based contour evolution and final segmented x-ray image.

### 5.2 Edge Point Extraction

The main objective of the proposed edge extraction technique is to identify objects boundaries with sufficient continuity to be successfully employed in the proceeding shape based point correspondence estimation. Radiographic image edge extraction is hindered by poor contrast, ill defined boundaries, noise and acquisition artefacts. Atlas/model based approaches were not considered to conduct the edge contour extraction as they would require
statistical model building and also a certain level of manual intervention. Active contours based methodologies require accurate model contour initialisation and would be troubled with multiple overlapping bone contours. Furthermore, many of the classic first (Roberts, Prewitt, Sobel) and second (LOG) derivative based edge segmentation methodologies tested, extract isolated edge pixels and did not provide continuous edge contours. These edge segmentation techniques were sensitive to noise and non-homogeneities in the radiographic images. Furthermore they required threshold adjustments for individual images.

The proposed edge extraction technique attempts to counter the aforementioned problems and link sets of potential edge pixels to create continuous boundaries. It employs three tactics to achieve this: adaptive image enhancement, thresholding with hysteresis and edge relaxation. Let it be clear that this edge extracted is not intended to provide an exact contour of the bony anatomy, but to provide continuous potential bony anatomy edges with few outliers, noise and occlusions.

Adaptive image enhancement is conducted through the use of two filters: averaging filter and the unsharp filter. The unsharp filter is a simple contrast enhancement operator (high frequency components in an image) via a procedure which subtracts a smoothed (unsharp) version of an image from the original image. The averaging filter is a low-pass filter where each pixel is replaced by an average of its neighbourhood pixels used mainly for ensuring homogeneity and limiting the effects of noise.

The radiographic image is divided into overlapping blocks, each of which will have one of the aforementioned filters applied onto it to enhance the edges or eliminate noise. Each block will be classified as “Homogeneous” or “Edge Containing” dependent on two key image region statistics, standard deviation and the uniformity measure [121]. The values of these measures are used as an indicator for the presence of an edge structures and thus will dictate which of the filters will be applied to the blocked local region. A large standard deviation locally suggests existence of edges and thus should be enhanced through the unsharp filter, while a low standard deviation suggest homogeneous regions where the
averaging filter can be employed to reduce the contrast. The uniformity measure employed is aimed at eliminating the discriminating features from noise that may misidentify edges. For the uniformity measure each local block is set a threshold to maximise the homogeneity of the resulting regions. For an image thresholding at a level \( t \), the uniformity measure is defined as \( L(t) \), where \( \sigma_i^2 \) is the variance of segmented region \( I \) and \( C \) is a normalisation constant used to set the measure to zero for the worst case condition. The threshold \( t \) is chosen to maximise the uniformity measure. A bimodal local block histogram will reveal an edge containing block while a unimodal histogram should reveal a homogeneous region. This adaptive filtering procedure is illustrated in Figure 5.3.

\[
L(t) = 1 - \frac{\sigma_1^2 + \sigma_2^2}{C}
\]  

(5.1)

![Figure 5.3](image)

**Figure 5.3** The main process steps of the adaptive filtering applied.

The adaptive image enhancement stage of the algorithm uses two parameters. One defines the size of the blocks to be classified (width and height), and the other specifies the standard deviation range that will require added uniformity measure screening prior to classification, as “Homogeneous” or “Edge Containing.” In Figure 5.3 this range was set to be between the 20th and 80th percentile of standard deviation values. The application of the filters in a small neighbourhood (5×5 window) avoids the appearance of square pattern artefacts at the borders of differently classified blocks as well as minimal information loss.

The next two stages involve the actual edge extraction process. Firstly, adaptive (local) hysteresis based thresholding is performed on the gradient image defined below, where the
two threshold values are set to be the 50th and 75th percentile of gradient magnitude values in the local window. Adaptive thresholding adjusts the threshold level according to the intensity statistics of a local region. This technique is employed typically with x-ray/fluoroscopic images to counter the illumination gradients present on the radiographs. The local windowed region utilised in the thresholding phase is of a different size to that employed in the image enhancement phase. For the testing conducted, the local window sizes were set to be 5% of the total pixel width and height of the image.

\[ \| \nabla F \| = \sqrt{ \left( \frac{\partial F}{\partial x} \right)^2 + \left( \frac{\partial F}{\partial y} \right)^2 } \]  \hspace{1cm} (5.2)

Thresholding with hysteresis can be defined as (3), where the pixels greater than a certain threshold (T1) are classified as an edge (foreground), those below a particular threshold (T2) are classified as background and those in-between are classified depending on its connectivity to a foreground pixel. E(x,y) is a thresholded version of the gradient image.

\[
E(x, y) = \begin{cases} 
0 & \| F(x,y) \| < T_2 & Background \\
0/1 & T_1 > \| F(x,y) \| \geq T_2 & Intermediate \\
1 & \| F(x,y) \| \geq T_1 & Foreground 
\end{cases}
\]  \hspace{1cm} (5.3)

This process, which identifies the pixels that are definite bony matter (foreground) and the definite non-bony matter (background), is followed by an edge relaxation process applied onto the intermediate pixels (pixels that fall between the 50th and 75th percentile of gradient magnitude values). The edge relaxation is performed through a region growing exercise. This region growing is conducted on the intermediate pixels to ensure sufficient continuity of the edge contour. Edge relaxation involves the recursive re-labelling of intermediate pixels with one or more neighbouring edge pixels, utilising an eight neighbouring connectivity scheme. In order to be reclassified as a foreground edge pixel the difference of the magnitude and orientation of the intervening pixel with the surrounding edge pixels will be checked and has to be within a user specified tolerance level.
Chapter 5 - Segmentation of Radiographic Images under Topological Constraints

Following the edge relaxation a small component elimination morphological operation is performed to remove all identified components with too few pixels. This is a noise elimination step that will remove any misidentified edge contour pixels. Finally the labelled gradient image is filtered with the Laplacian of Gaussian filter to ensure that the edges are skeletonised. The Gaussian threshold is set to be zero to force closed contour edges.

Figure 5.4 Comparison between Sobel, Canny and the proposed edge point extraction algorithms. From left to right: (a) original image; (b) detected edges through the Sobel operator; (c) detected edges through the Canny operator; (d) detected edges through the authors’ edge point extraction algorithm.

Proceeding the edge point extraction, a certain level of sub-sampling has to be performed on the edge points identified to provide a realistic number of points as input to the point correspondence estimation algorithm. The sub-sampling is weighted in accordance to the connected components of the extracted edges. Thus the final relative proportion of pixels after sub-sampling will remain the same. These sub-sampled points are then utilised in the point matching algorithm as detailed in the proceeding section.

5.3 Registration

The registration is performed through a point correspondence estimation between the two point sets, which is detailed in Section 3.6. Similar to Section 3.6 the similarity measure utilised in this case needs to be robust to several factors that are inherent problems seen in edges extracted in x-ray images. These factors include: disturbances (noise), deformations
(shape and affine), outliers and occlusions. Thus the similarity measure employs four feature components: topology information, edge orientation, curvature and continuity (distance). The overall similarity measure is a weighted combination of the individual similarities. It has been noted empirically that the use of multiple features reduces the discriminatory power of the individual features and leads to improved results. Figure 5.5 shows an example of the non-rigid registration process.

![Figure 5.5 Example of the outlier removal process and regularised planar interpolation. From left to right: (a) correspondences identified through the Hungarian algorithm; (b) removal of outliers (misidentified correspondences); (c) regularised TPS based interpolation of the transformation field for the sub-sampled model contour; (d) regularised TPS based interpolation of the transformation field for the full model contour.](image)

### 5.4 Contour Evolution

Once the non-rigid registration of the proximal and distal segments has been performed a contour evolution is performed on fractured bones to segment the fracture region. The multi-phase active contour model proposed by [122] has been utilised to perform this multiple contour evolution. This technique is superior to other traditional region based active contours.
which partition an image into simply two subsets, either the inside or the outside of a single contour. The multiple active contours evolve independently based on the piecewise-constant model documented in [122]. This region based active contours model was preferred over edge based techniques (e.g. Geodesic Active Contours [123]) as it’s robust to initialisation and noise. Edge based techniques are sensitive to noise and artefacts, causing the evolution to hinder due to gaps in the boundary.

The multiple active contours evolve independently based on the piecewise-constant model shown below,

\[
\frac{\partial \phi(x, y)}{\partial t} = \delta_\epsilon (\phi(x, y)) [v k(\phi(x, y)) - \{I(x, y) - \mu_1\}^2 - \{I(x, y) - \mu_0\}^2] \tag{5.4}
\]

The function \( \phi(x, y) \) is termed the level set function and the zero level is defined as the contour of interest, through the equation below, where \( \Omega \) denotes the entire image plane.

\[
C \equiv \{(x, y) : \phi(x, y) = 0\}, \forall (x, y) \in \Omega \tag{5.5}
\]

\( \mu_1 \) and \( \mu_0 \) are the mean image intensities within the subset (inside and outside the contour). The final segmented image is represented as a set of piecewise-constants where each subset is represented as a contour. Furthermore, the curvature term \( k \) maintains the regularity of the contours, while the term \( v \) accelerates and keeps the contour evolution by either minimising or maximising the enclosed area.

### 5.5 Experiments

The point correspondence estimation algorithm proposed in this chapter was benchmarked against three state-of-the-art point matching algorithms. A sample point set presented by [124] was utilised for this experimentation. The algorithms involved in the test include Continuous Point Drift (CPD) [125], Robust Point Matching (RPM) [124], and the Iterative Closest Point (ICP) [81]. The data was synthetically modified into three variations to test the algorithms effectiveness against disturbances, outliers and deformations. Testing was
performed on several levels of change per criteria (disturbance, outliers and deformations) and 50 samples were collected per level. Figure 5.6 illustrates the performance of the algorithms under the various testing criteria. The error value quoted is the average Euclidean distance between the correspondences identified through the algorithm and the ground truth correspondence known prior to the test. The testing performed on the proposed algorithm was conducted with each feature similarity measure receiving equal weighting.

Figure 5.6 Box plots of the benchmarking performed. The indices used in the figure are: 1-Proposed Framework, 2-CPD, 3- RPM, 4-ICP. The outlier tests had a range of outlier to point ratios (ranging from 1 to 2 times). The disturbance and deformations tests had 5 levels of change each.

The developed algorithm performs best under the outlier and deformation criteria. This is primarily due to the specific features (curvature and edge orientation) utilised in the algorithm that ensures conformance to shape and outlier deteriorations. The proposed algorithm however does not perform highly in the disturbance testing. This is due to the disturbance being classified as outliers. Thus the algorithm logically rejects the matching between the model contour and an outlying point. Another reason behind the weak performance is due to the final correspondence outlier elimination step, which filters out much of the assignment correspondences. Firstly, this effect can be minimised by setting a lesser weight to the curvature and edge orientation criteria or vice versa a higher weighting to the topology or continuity similarity measures, which will ensure the removal of the outlier rejection affinity. Secondly the final correspondence outlier elimination filter criteria can be relaxed to pass through more assignments.

Furthermore, the segmentation procedure discussed in this chapter was tested on 15 sets of clinical x-ray and fluoroscopic images, both on distal and proximal femurs. In conducting the clinical tests, two generic models were employed. One generic model aimed at the
proximal femur segmentation and another for distal femur segmentation. These truncated models were required as the clinically acquired x-ray images are typically only for a certain portion of the full femur. Table 5.1 shows quantitative segmentation results. The key accuracy measurement used during the testing is the average root-mean-squared distance (RMSD), in mm terms, between the automatically segmented femur contour and the manual segmented ground truth. The average RMSD was 2.5 mm and the standard deviation was 0.80 mm. Average time taken for the 15 segmentation tests was 1.26 minutes. Figure 5.7 provides qualitative examples of several of the segmentation results. The calibration of the x-ray images required to obtain the mm error dimensions was performed utilising the Othomark calibration ball [112]. This calibration ball with a known diameter was placed in the frame of the x-ray image and the required intrinsic calibration parameters obtained through manual measurements.

### Table 5.1 Diagnostic radiograph segmentation testing results.

<table>
<thead>
<tr>
<th>Image ID</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSD Error (mm)</td>
<td>2.2</td>
<td>2.3</td>
<td>2.6</td>
<td>1.9</td>
<td>2.4</td>
<td>2.3</td>
<td>1.9</td>
<td>2.5</td>
<td>2.9</td>
<td>3.2</td>
<td>3.2</td>
<td>3.0</td>
<td>2.2</td>
<td>2.8</td>
<td>2.7</td>
</tr>
</tbody>
</table>
5.6 Chapter Summary

This chapter has introduced a topologically constrained automatic bone contour extraction framework. The testing performed on the framework has shown robustness to disturbance, outliers, deformation and occlusions. The only prior knowledge assumed by the framework is a 3D surface model of the anatomy of interest that is utilised to create projected 2D contours in the plane of the radiographic image. Thus the proposed technique eliminates much of the statistical and rule based learning requirements imposed by other segmentation techniques currently seen in literature.
Chapter 6 Intra-Operative Pose Estimation

This chapter provides a general introduction to the rigid-body pose estimation (2D-3D registration) problem applied specially to medical modalities. This registration is concerned with bringing the pre-operative 3D data (patient’s images or models of anatomical structures obtained from these images) and intra-operative data (patient’s images and positions of surgical tools) into the same coordinate frame. Registration of pre and intra-operative data is one of the key technologies for image guided surgery and minimally invasive interventions.

The chapter presents the motivation behind this work and the potential medical application of this research. The chapter also details the problem statement and the technical considerations behind the pose estimation problem. A high level of detail is provided on the transformation components that need to be recovered as a result of the rigid body registration algorithm.

6.1 Introduction

As previously stated, there has been a recent growth in the number of medical experts who advocate a minimally invasive approach to surgery [1, 2]. Their aim is to reduce the physical stress applied to the human body due to medical treatment/procedures and also to reduce treatment costs, by minimising the size and number of incisions. Unfortunately, in comparison to open procedures, these approaches restrict the surgeon's view of the anatomy. This leads to an increasing need for advanced imaging techniques that would aid in guiding interventions.

Pre-operative images provide an excellent source of detail about the anatomy in question. The research work presented in Chapter 3 and Chapter 4 are examples of such visualisation. Furthermore, direct 3D imaging with modalities such as MRI and CT are also possible in
certain cases. These 3D models greatly assist in establishing diagnosis and planning procedures pre-operatively or evaluating an intervention post-operatively. The same set of images can be conveniently utilised in surgery as well. However, they have the drawback that they may not completely reflect the surgical situation, since they are static.

In all minimal invasive surgical applications it is critical to use intra-operative images to follow the changes caused by the procedure or to visualise the location of a tool. In the operating room or interventional suite, it is mostly 2D images that are available to record details about the current anatomical state. x-ray, c-arm fluoroscopy and portal images are all examples of image modalities used for this purpose. Two-dimensional acquisitions are often taken instead of volumetric datasets because of timing, radiation related and technological arguments. First, acquiring several 3D volumetric images during a procedure takes too long to make it practical compared to 2D imaging. Second, the radiation dose to both the patient and the surgeon is reduced if only image slices are recorded rather than all the projections needed to reconstruct a 3D volume. Third, by using only 2D images, it is sufficient to have simpler imaging equipment in the operating rooms.

Unfortunately, 2D images lack significant information that is present in the 3D modalities. Hence, in order to relate the changes recorded by the 2D modalities to the detailed 3D model, medical experts need to fuse the information from the pre-operative and intra-operative images mentally, which can be a challenging task. Therefore, it is useful to find a way to both automate that procedure and to make it reliable. The combination of pre-operative and intra-operative images conveys the most information if the components are properly aligned in space. To achieve this it is necessary to determine their relative position and orientation. The procedure that identifies a geometrical transformation which aligns two datasets, or in other words locates one of them in the coordinate system of the other, is called registration, or more specifically 2D-3D registration.

2D-3D pose estimation is a crucial task for many problems in areas ranging from robot navigation to medical intervention. It involves estimating the relative position and orientation
of a known 3D object with respect to a reference camera system. In other words, we search for a transformation (the pose - position and orientation) of the 3D object such that the transformed object corresponds to 2D image data. For rigid objects, such a transformation can be the Euclidean transformation consisting of a rotation and a translation.

In practice, a registration algorithm tries to find among all allowed spatial transformations between two or more spaces in that transformation which fulfils or maximises the image alignment criterion. The criterion is usually a cost function or similarity measure describing the spatial alignment of features representing the registration basis. The aim of registration is then to find a spatial transformation that optimises the given criterion. However, in the particular case of 2D-3D registration the condition of one-to-one mapping is not fulfilled, since projection of a 3D space to a 2D image is a many-to-one mathematical relationship between the 3D and 2D space [126]. As such, 2D-3D registration is in general an ill-posed problem. The problem can be avoided by limiting the degree of geometrical transformation, most often to rigid transformations, and by increasing the number of non-correlated 2D projections used in the registration.

Registering pre-operative datasets to images acquired intra-operatively can provide up-to-date information at the treatment site, guiding surgery or other interventions. Majority of the medical applications for the proposed kind of registration task have emerged in the field of radiology. Alignment information is crucial in planning, guidance and treatment procedures. More specifically, the medical community has expressed interest in applying the 2D-3D alignment results in the following application areas: placement of pedicle screws in spine surgery [80, 95, 127], aortic endoprostheses in transfemoral endovascular aneurysm management [92, 95], verifying patient setup accuracy for radiotherapy [128] and acetabular implant position in case of total hip replacement [92-95]. These applications will be further detailed below.

Our collaborators (Auckland District Health Board and the Auckland City Hospital), in specific, are interested in applying the 2D-3D registration in the field of orthopaedics. One of
the major cases of interest is tracking of fractured femur bone segments during femur reduction surgery. Minimally invasive femur fracture surgery is presently guided through fluoroscopic c-arm images, due to the lack of direct inspection of the fracture movements and implant/tool locations. Compared to full x-ray images typically acquired pre-operatively, intra-operative fluoroscopic images have a lower resolution with a smaller field of view (typically 12cm × 12cm). These small 2D images create a challenging visualisation task for the surgeon. Typically during the procedure two orthogonal fluoroscopic images (frontal and lateral viewpoints) will be acquired and presented. Subsequently, the surgeon will require extensive 3D spatial skills to mentally reconstruct the pose of each of the fragments in 3D. External literature suggests conducting surgical procedures on the basis of 2D images alone produce errors leading to costly secondary surgery or irreversible biomechanical defects [111]. This research indicated that the positioning errors generate bone misalignments and complications (about 18% of femur fracture reduction cases) which could be reduced with the introduction of 3D bone fragment visualisation during surgical procedures. Consequently 3D visualisation of anatomy plays an important role in femur fracture surgery.

The work presented in the proceeding three chapters is motivated by the aforementioned concerns. The chapters present an algorithm for recovering the position and orientation of the target anatomy in 3D space, based on the evaluation of 2D planar radiographs with a pre-operative 3D model.

6.2 Potential Medical Applications

The orthopaedic speciality has had a tremendous increase in the number of minimal invasive interventions over the past two decades, due to an aging population and an increasing availability of sophisticated and specifically designed imaging tools (such as intra-operative c-arm technologies) [3]. Most of these interventions will be able to make use of a 2D-3D registration technique to provide intra-operative guidance. Below are several key
Chapter 6 - Intra-Operative Pose Estimation

surgical applications with existing literature highlighting the need of 3D intra-operative tracking.

Fracture Reduction: Tracking of the bone fracture segments intra-operative can be achieved through 2D-3D registration [19-22]. This is the main application of this research work and has been previously detailed in Section 1.3.

Total Hip Replacement: Hip joint replacement surgery has several uses for 2D-3D registration. One is implanting an acetabular cup into the pelvic bone during total hip replacement procedures. In order to verify the correct position and orientation of the metal cup before the operation terminates, 2D images are acquired. These need to be related to the 3D model of the anatomy. Another use concerns cases in revision surgery. Such a procedure is necessary if, following a total hip replacement procedure, the acetabular cup gets mislocated or gets detached from the pelvis. Thus 2D-3D registration can also be used for analysis of post-operative x-ray images for possible errors in cup positioning [92-95].

Spine Procedures: Spine procedures are another key application area for IGS, since back problems are very common and the potential complications of damage to the spinal cord are devastating. Planning may effectively use pre-operative CT, while the interventions may be most practically guided by the use of c-arm fluoroscopy equipment. One example procedure is vertebroplasty, which is the reinforcement of a failing vertebra by the placement of cement. In spine surgery, reinforcement of falling vertebra by placement of cement and pedicle screw placement are just some possible applications of 3D/2D image registration. Since the spine is a non-rigid anatomy, registration is based on single vertebra rigid registration [80, 83, 95, 127].

Metastatic Bone Cancer: Another application is related to an orthopaedics procedure, the treatment of metastatic cancer in the bones. The task here is to remove localised lesions from particular locations of the bones. Again, the treatment plan can be thoroughly designed prior to the operation using 3D CT volumes with high information content, but during the
intervention, guidance and verification is only practical by making use of intra-operative images. Utilising both of the two data sources requires the alignment of the intra-operative and pre-operative datasets.

Knee Arthroplasty: 3D/2D image registration of lower extremity bones and knee implants can be used to analyse knee kinematics before and after a total knee arthroplasty procedure [129]

Non Orthopaedic: Several other (non orthopaedic) areas for the use of intra-operative pose-estimation has been identified as Radiotherapy [128] and Endovascular Treatments [92, 95].

6.3 Problem Statement

As previously introduced, the goal of the chapter is to register pre-operative 3D data to intra-operative 2D images, through a procedure commonly known as 2D-3D registration. This thesis is particularly interested in examining the problem of aligning 3D fractured bone models to corresponding 2D x-ray fluoroscopy images.

The required task when attacking the 2D-3D registration problem is to return a geometric transformation that best specifies the position and orientation of the examined anatomy at the time of obtaining the 2D projection images. Hence, a way to align the images and the world coordinate systems or to determine the correspondence between the intra-operative imaging environment and the coordinates of the pre-operative volumetric data needs to be identified (Figure 6.1). The alignment transformation will aid in the fusing of pre-operatively acquired 3D surface data and bi-planar x-ray fluoroscopy images (acquired through an interventional c-arm).

According to the nature of spatial transformation and its degrees-of-freedom, registration can be classified as rigid and non-rigid. Non-rigid registration can further be divided into
affine, projection and high-order elastic registration. The majority of published 2D-3D registrations employ a rigid transformation model \([19-22]\), composed of translations and rotations. Rigid registration is generally applied when it is assumed that target anatomy fulfils the criterion of rigidity and no spatial distortions are induced in the image acquisition process. The limited number of degrees-of-freedom in rigid transformations ensures the 2D-3D registration problem is better defined, in spite of the fact that the condition for one-to-one mapping is not fulfilled.

The pose estimation of a rigid object from an image depends on eleven parameters. Extrinsically the object pose is identified through three rotational and three translational parameters needed to bring the model into the observed position. Throughout this thesis, we denote the transformation that aligns the two coordinate systems (pre-operative model coordinate system to the intra-operative target coordinate system) by transformation \(T_{\text{model} \rightarrow \text{target}}\) (or abbreviated \(T\)). Figure 6.1 illustrates the transformation involved in the 2D-3D registration problem. Moreover there are also five intrinsic camera parameters, the two principal point coordinates, the focal length and the two digitisation steps of the system of acquisition.

Thus, the problem formulation can be generalised as, given: 1) the intrinsic parameters of the acquisition system, 2) the projective (pinhole) model for image formation, 3) the 3D object model, 4) a 2D target image, and 4) the extraction of the feature primitives used for the correspondence identification between the object and the 2D image, find the extrinsic pose parameters giving the position and orientation of the viewed object in the acquisition reference system.

Thus, the problem formulation can be generalised as find the extrinsic pose parameters, given: 1) the intrinsic parameters of the acquisition system, 2) the projective model (pinhole) for image formation, 3) the 3D object model, 4) a 2D target image, and 4) the extraction of the feature primitives used for the correspondence identification between the object and the 2D image.
Sections 6.4.1 details the pose estimation transformation definition composed of a $3 \times 3$ rotation matrix and a $3 \times 1$ translation vector that constitute the rigid body motion needed to transform a 3D object into agreement with 2D image data. Throughout this thesis the term pose estimation can be used synonymously with 2D-3D registration.

![Diagram of coordinate systems](image)

**Figure 6.1** The transformation parameter $T_{\text{target}}$ which relates the coordinate frames of the imaging environment and the data volume.

With a single 2D image the in-plane rotation and displacement transformations can be accurately recovered. It is however difficult to determine any out-of-plane transformations. Thus our experiments have used two 2D images to robustly recover all registration parameters required to properly position the examined anatomy in the 3D world. Two images provide sufficient information about the spatial location of the imaged object.

We assume that the two imaging views are related by a known transformation, hence it is necessary to recover the required transformation with respect to only one of them. This is a realistic assumption as bi-planar images are often acquired by rotating the imaging source by
Chapter 6 - Intra-Operative Pose Estimation

a pre-specified angle around one of the imaging axis. Also, bi-planar acquisitions are considered to be standard in orthopaedic applications.

In solving the proposed problem, the main challenges lie in identifying a similarity measure, or objective function, that can quantify the quality of the alignment between the images and defining a procedure to modify and refine current estimates of the problem parameters in a way that the similarity score is optimised.

An additional primary focus of this effort is finding 2D-3D alignment methods which have computational complexity that is compatible with the time constraints implied by the interventional applications.

Experimentally, we aim to demonstrate the performance characteristics of our registration algorithm on a wide variety of datasets. The experimental collection includes fluoroscopy and CT datasets of three fractured femora.

6.4 Technical Issues

6.4.1 Transformation Representation

In intra-operative IGS, registration is used to align the pre and intra-operative data during an intervention in such a way that corresponding anatomical structures in the two data sets are aligned. The data sets to be registered are defined in distinct spaces or coordinate systems. The 3D pre-operative data is defined in the model data (image) coordinate system $S_{pre}$. The intra-operative 3D data is defined either in some world (patient, treatment room) coordinate system $S_w$ or target data coordinate system $S_{intra}$, while the 2D intra-operative data is always defined in data coordinate system $S_{intra}$ (Figure 6.2). If the intra-operative data is defined in $S_{intra}$, a rigid transformation $T_{calib}$ has to be defined by calibrating the intra-operative data acquisition device to relate $S_{intra}$ to $S_w$ (Figure 6.2).
Chapter 6 - Intra-Operative Pose Estimation

Let \( x_A^{3D}, x_B^{3D}, \) and \( x_{Bj}^{2D}, j = 1,2,\ldots,N \), denote points of the pre-operative data \( A \), intra-operative data \( B \) and \( N \) intra-operative data \( B_j \), defined over the domains \( \Omega_A \subset R^3, \Omega_B \subset R^3 \), and \( \Omega_{Bj} \subset R^2, j = 1,2,\ldots,N \), respectively.

Besides let \( A(x_A^{3D}), B(x_B^{3D}), \) and \( B_j(x_{Bj}^{2D}), j = 1,2,\ldots,N \), denote values of data \( A \), \( B \) and \( B_j \), at positions \( x_A^{3D}, x_B^{3D}, \) and \( x_{Bj}^{2D}, j = 1,2,\ldots,N \), respectively.

The values can be intensities of the raw images or any feature data (this thesis has used feature data), even the points \( x_A^{3D}, x_B^{3D}, \) and \( x_{Bj}^{3D}, j = 1,2,\ldots,N \), themselves, extracted from these images or obtained by other means (optical tracking). Registration is concerned with finding the transformation \( mT_i \) that defines the pose of model \( S_{pre} \) in target \( S_w \) (Figure 6.2). The transformation \( T \) is found by transforming the data set \( A(x_A^{3D}) \) into \( A^T(x_A^{3D}) \) until \( A^T(x_A^{3D}) \) is best aligned with \( B(x_B^{3D}) \) or \( B_j(x_{Bj}^{2D}), j = 1,2,\ldots,N \).

A correct registration will allow any point defined in \( S_{pre} \) to be precisely located in \( S_w \) and/or \( S_{intra} \). Due to and dimensionality of data, the transformation \( mT_i \) is obtained by a 2D-3D registration.

When conducting an alignment of the 3D spatial data to projective data, the one-to-one correspondence between 3D and 2D data is not valid. In 2D-3D registrations of 3D spatial and 2D projective data, \( N \) 2D patient images \( B_j(x_{Bj}^{2D}), j = 1,2,\ldots,N \), are acquired intra-operative. Again, if the transformation \( T_{calib} \) is known, the relations of \( S_{intra,j} \) to \( S_w \) are also known. Because of the different dimensions of the data sets to be registered, the transformation \( mT_i \) is defined by transforming the data set \( A(x_A^{3D}) \) into \( A^T(x_A^{3D}) \) until the projections of \( A^T(x_A^{3D}) \) onto the domains \( \Omega_{Bj} \) are best aligned with images \( B_j(x_{Bj}^{2D}), j = 1,2,\ldots,N \),

\[
mT_i; P_j(A^T(x_A^{3D})) = A_j(x_{Aj}^{2D}) \iff B_j(x_{Bj}^{2D}), \quad \forall j \tag{6.1}
\]
or until \( A^T(x_A^{3D}) \) is best aligned with the back-projections or reconstructions of \( B_j(x_{Bj}^{3D}) \) into the domain \( \Omega_A \).

\[
m_T: A^T(x_A^{3D}) \Leftrightarrow B_j(x_{Bj}^{3D}) = BP_j(B_j(x_{Bj}^{2D})), \quad \forall j
\]

Or

\[
m_T: A^T(x_A^{3D}) \Leftrightarrow B(x_B^{3D}) = R[B_j(x_{Bj}^{2D})], j = 1,2,\ldots,N
\]

respectively. The \( P_j, j = 1,2,\ldots,N \), are projection matrices defining projections of points in 3D onto each of the \( N \) 2D planes and are obtained by calibrating the intra-interventional imaging device. The \( BP_j, j = 1,2,\ldots,N \), are corresponding back-projection matrices and \( R \) is a reconstruction function. The transformation \( T \) is again a complete mapping that incorporates image resampling and interpolation, and maps both positions and associated image values.

Figure 6.2 Geometrical setup of the registration of a 3D image to two 2D x-ray projection images. \( s_j \) and \( s_{j+1} \) are the positions of the x-ray sources related to the \( j^\text{th} \) and \( (j+1)^\text{th} \) 2D images defined in coordinate systems \( S_{\text{intra},j} \) and \( S_{\text{intra},j+1} \), respectively. \( S_w \) is the world coordinate system and \( S_{\text{pre}} \) is the coordinate system of the pre-interventional 3D image. \( T_{\text{calib},j} \) are the rigid transformations between \( S_{\text{intra},j} \) and \( S_{\text{intra},j+1} \). \( mT \) is defined by six parameters \( t_{x}, t_{y}, t_{z}, \theta_{x}, \theta_{y}, \theta_{z} \).

### 6.4.2 Number of Views

In order to detail the criteria utilised to consider the number of images it is necessary to understand the registration dimensionality and modes. To perform 2D-3D registration, the 3D
and 2D data have to be brought into dimensional correspondence. Dimensional correspondence can be achieved either by transforming the 3D data into 2D or by transforming the 2D data into 3D. While the former approach leads to 2D-2D registration(s), the latter approach leads to a 3D-3D registration. More specifically, dimensional correspondence can be achieved either by the projection, back-projection, or reconstruction strategy.

Through the projection based strategy (Figure 6.3), the 3D data is projected onto $N$ planes $\Omega_{B_j}, j = 1,2,\ldots,N$, of the intra-interventional 2D projection data $B_j(x^{2D}_{B_j})$ using the $N$ projection matrices $P_j$, associated with $B_j(x^{2D}_{B_j}), j = 1,2,\ldots,N$. The registration is then performed by optimising the sum of criterion functions, $CF_{j}^{2D}, j = 1,2,\ldots,N$:

$$mT_t = \arg\max_T \sum_{j=1}^{N} CF_{j}^{2D}$$

$$mT_t = \arg\max_T \sum_{j=1}^{N} CF_{j}^{2D} \left(P_j \left(A^T(x^{3D}_A), B_j(x^{2D}_{B_j})\right)\right)$$

(Figure 6.3 2D-3D registration based on projection strategy.)

Through the back-projection based strategy (Figure 6.4), each of the $N$ 2D intra-interventional data $B_j(x^{2D}_{B_j})$ is back-projected into the 3D space using corresponding back-projection matrices $BP_j, j = 1,2,\ldots,N$. Similarly to the projection strategy, the registration is
then performed by optimising the sum of criterion functions $CF_{j}^{3D}$, $j = 1,2,\ldots,N$, each calculated on the basis of the 3D pre-interventional data and one of the back-projected intra-interventional 2D data:

$$m\hat{T}_t = \operatorname{argmax}_T \sum_{j=1}^{N} CF_{j}^{3D}$$

$$m\hat{T}_t = \operatorname{argmax}_T \sum_{j=1}^{N} CF_{j}^{3D} \left( A^T(x_A^{3D}), B_j(x_{B_j}^{2D}) \right)$$

Figure 6.4 2D-3D registration based on back-projection strategy.

By the reconstruction strategy (Figure 6.5), a reconstruction function $R$ is used to reconstruct the 3D intra-interventional data $B(x_B^{3D})$ from $N$ 2D intra-interventional data, $B_j(x_{B_j}^{2D}) : B(x_B^{3D}) = R \left( B_1(x_{B_1}^{3D}), \ldots, B_N(x_{B_N}^{3D}) \right)$. Registration is then performed by optimising the criterion function $CF^{3D}$ calculated between the pre-interventional and the reconstructed 3D data:

$$m\hat{T}_t = \operatorname{argmax}_T \sum_{j=1}^{N} CF_{j}^{3D} \left( A^T(x_A^{3D}), B(x_B^{3D}) \right)$$

121
In clinical practice, it is desired to keep the number of x-ray images to a minimum due to acquisition and reconstruction times and also due to radiation exposure constraints. While in case of the projection and back-projection strategies one to two 2D images might be enough to achieve 2D-3D registration, the reconstruction approach requires much more than two to build a 3D image of sufficient quality that enables accurate and robust registration to the high-quality pre-operative image. Generally, the more intra-interventional images are used for reconstruction the better the registration accuracy.

The experiments presented in this thesis have employed a projection based strategy (Figure 6.6). It has been empirically identified that examining only a single 2D image is not sufficient to robustly recover all registration parameters required to properly position the examined anatomy in the 3D world. While in-plane rotation and displacement transformations can be accurately recovered, it is difficult to determine any out-of-plane transformations if no clear feature correspondences exist.

In order to establish all of the transformation components with a desired level of certainty, it has been empirically proven advantageous to use two or more 2D acquisitions for the
proposed alignment problem. Section 8.9 detail the tests conducted to analyse the effect of registration accuracy with varying number of images. Furthermore our medical collaborators have also specified that two orthogonal images are the norm within clinical practice.

A detailed description of the 2D-3D registration methods with respect to the described strategies and the nature of registration basis are provided in the proceeding chapters.

Figure 6.6 An illustration of a fractured 3D femur placed in context of two orthogonal fluoroscopy images. The projections of the 3D model are displayed in the imaging plane.

6.4.3 Nature of User Interaction

While research activities make every effort to develop fully automated 2D-3D registration algorithms, some user interaction is still necessary. The level of user interaction required can be interactive, semi-automatic or automatic. Fully automatic registration algorithms require that the user provides only the image data. Contrary to automatic registration, interactive registration is assisted by a user using software tools which provide visual feedback of the current transformation. Most often the level of user interaction is somewhere between these two extremes. For semi-automatic registration it is essential that the user provides some initialisation to the algorithm, such as segmentation, the initial guess of registration, and/or to validate the results of registration. It cannot be claimed that any of the published 2D-3D
registration algorithms is fully automatic. An interactive registration performed by tuning translation and rotation parameters was implemented by [36]. Unfortunately, accurate interactive registration is time consuming and depends on the skills of the human operator. However, even a course interactive 2D-3D registration can be fast and can provide an initial guess of registration for subsequent semi-automatic procedures. Most of 2D-3D registration methods are semi-automatic and require different levels of user interaction. In landmark-based methods, the user has to provide a set of corresponding anatomical landmarks [79]. For 2D-3D registration based on 3D surfaces and corresponding 2D outlines of anatomical structures [71, 130-134] a user may be needed to adjust the parameters of the automatic 3D and 2D segmentation methods and to validate and correct segmentation results. Minimal user interaction may be required by whole-image-content-based 3D/2D registration algorithms. In these registration algorithms it is essential that the initial guess of the registration is close to true registration. This condition may be fulfilled by accurate positioning of the patient with respect to imaging devices by means of laser markers or by a course interactive 2D-3D registration performed by the user [36, 85, 134]. Since the 2D-3D registration problem is by its nature an ill-defined problem, 2D-3D registration algorithms may provide registration results that are false. To avoid the risk of guiding a medical procedure with false registration results, the results of 2D-3D registration should be validated by the user. Initial interactive user registration and final validation of registration results may be supported by proper visualisation such as superimposing the outlines of a 3D surface [130, 132] or edges extracted from DRRs [83] onto 2D x-ray images.

The pose estimation work presented in the remainder of the thesis will highlight the key areas of user interactions required, they are tuning parameters involved with the 2D segmentation and feature extraction process (Section 7.2 and Section 8.2) as well as providing adequate initialisation to the optimisation algorithm documented in Chapter 8. The interactive nature of the algorithms will be detailed in the individual sections.
6.4.4 Parameterisation of Patient Pose

This section introduces the parameterisation guidelines for the pose parameter describing the position and orientation of the patient’s anatomy. In other words, the transformation $^mT_t$ (as detailed in Section 6.4.1) specify the coordinate transformation which maps coordinates from a stationary pre-operative data coordinate system into a coordinate system associated with the intra-operative patient. It is convenient to write this coordinate transformation $^mT_t(\gamma)$ as a 4×4 transformation matrix, where $\gamma$ is the pose parameter vector listing the three translational and three rotational parameters.

There exists several ways to encode the transformation parameters that need to be recovered. The displacement component of $^mT_t$ can be conveniently represented in a 3D vector format. However, the rotation parameter can be formulated in several different ways. Just to name a few of the options: roll-pitch-yaw; XYZ Euler angles; equivalent angle-axis, orthonormal matrices and quaternions. This thesis has employed an Euler angles based representation. This representation was appropriate for our needs as the Euler encoding is easy to formulate and the composition of rotation operators easily extracted.

![Figure 6.7](image)

*Figure 6.7 The 6-element parameterisation of patient pose comprises three consecutive rotations around the three coordinate axes, followed by a 3D translation.*

One convenient representation of the rigid body transformation between the pre-operative 3D coordinate system and the intra-operative coordinate system is the six parameter vector \([t_x, t_y, t_z, \theta_x, \theta_y, \theta_z]^T\), where $t_x$, $t_y$, and $t_z$ are orthogonal translations and $\theta_x$, $\theta_y$, and $\theta_z$ represent consecutive rotations around each of the three coordinate axes. Figure 6.7 illustrates the application of these rotations and translations.
Chapter 6 - Intra-Operative Pose Estimation

This parameterisation represents a rigid body transformation, which has six degrees-of-freedom, utilising the six parameters. The relationship between the pose estimation parameter representation and the matrix $T$ can be identified by writing the translation and each of the rotations as matrices, and then composing them.

The translation matrix representation, $T$, of the translation parameters $[t_x, t_y, t_z]$ is,

$$T(t_x, t_y, t_z) = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6.7}$$

Similarly, the matrices $R_x$, $R_y$, and $R_z$, which represent rotations around the X, Y, and Z axes, can be written as,

$$R_x(\theta_x) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_x) & -\sin(\theta_x) & 0 \\ 0 & \sin(\theta_x) & \cos(\theta_x) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_y(\theta_y) = \begin{bmatrix} \cos(\theta_y) & 0 & \sin(\theta_y) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\theta_y) & 0 & \cos(\theta_y) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6.8}$$

$$R_z(\theta_z) = \begin{bmatrix} \cos(\theta_z) & -\sin(\theta_z) & 0 & 0 \\ \sin(\theta_z) & \cos(\theta_z) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Composing these four matrices to find $T_{\text{target}}$ gives,

$$T_{\text{target}}^\text{model} = T(t_x, t_y, t_z) \ast R_x(\theta_x) \ast R_y(\theta_y) \ast R_z(\theta_z)$$

$$T_{\text{target}}^\text{model} = \begin{bmatrix} c_x c_z & s_x s_z c_x - c_x s_y & c_y s_x c_z + s_x s_y & t_x \\ c_y s_z & s_x s_y c_z + c_x s_z & -s_y c_x c_z - s_x s_z & t_y \\ -s_y & s_x c_y & c_x c_y & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6.9}$$

where,

$$s_x = \sin(\theta_x), c_x = \cos(\theta_x), s_y = \sin(\theta_y), c_y = \cos(\theta_y), s_z = \sin(\theta_z), c_z = \cos(\theta_z) \tag{6.10}$$
Chapter 6 - Intra-Operative Pose Estimation

Referring to the parameterisation detailed above, a 3D point \([w_x, w_y, w_z]^T\) in world coordinates corresponds to the model coordinate \([m_x, m_y, m_z]^T\) as follows:

\[
\begin{bmatrix}
  m_x \\
  m_y \\
  m_z
\end{bmatrix} = 
\begin{bmatrix}
  c_y c_z (w_x) + (s_x s_y s_z - c_x s_z)(w_y) + (c_x s_y c_z + s_x s_z)(w_z) + \tau_x \\
  c_y s_z (w_x) + (s_x s_y s_z + c_x c_z)(w_y) + (c_x s_y s_z - s_x c_z)(w_z) + \tau_y \\
  -s_y (w_x) + s_x c_y (w_y) + c_x c_y (w_z) + \tau_z
\end{bmatrix}
\]  
(6.11)

The disadvantage of this Euler angle parameterisation (in contrast to quaternions) is that it suffers from degenerations in certain areas of the parameter space. This is evident when \(\theta_x = 90\) degrees, and \(\theta_y = -90\) degrees. In this case, the rotations \(\theta_x\) and \(\theta_z\) are about the same physical axis. The degeneracy can be seen by computing the derivatives of the above equation with respect to the rotation parameters \(\theta_x, \theta_y,\) and \(\theta_z\):

\[
\frac{\partial}{\partial \theta_x} \begin{bmatrix} m_x \\ m_y \\ m_z \end{bmatrix} = 
\begin{bmatrix}
  (c_x s_y c_z + s_x s_z)(w_y) + (-s_x s_y c_z + c_x s_z)(w_z) \\
  (c_x s_y s_z - s_x c_z)(w_y) + (-s_x s_y s_z - c_x c_z)(w_z) \\
  (c_x c_y)(w_y) - (s_x c_y)(w_z)
\end{bmatrix}
\]

\[
\frac{\partial}{\partial \theta_y} \begin{bmatrix} m_x \\ m_y \\ m_z \end{bmatrix} = 
\begin{bmatrix}
  -(s_y c_z)(w_x) + (s_x c_y c_z)(w_y) + (c_x c_y c_z)(w_z) \\
  -(s_y s_z)(w_x) + (s_x c_y s_z)(w_y) + (c_x c_y s_z)(w_z) \\
  -(c_y)(w_x) - (s_x s_y)(w_y) - (c_x s_y)(w_z)
\end{bmatrix}
\]  
(6.12)

\[
\frac{\partial}{\partial \theta_z} \begin{bmatrix} m_x \\ m_y \\ m_z \end{bmatrix} = 
\begin{bmatrix}
  -(c_y s_z)(w_x) - (s_x s_y s_z + c_x c_z)(w_y) - (c_x s_y s_z - s_x c_z)(w_z) \\
  (c_y c_z)(w_x) + (s_x s_y c_z - c_x s_z)(w_y) + (c_x s_y c_z + s_x s_z)(w_z)
\end{bmatrix}
\]

When \(\theta_x = 90\) degrees, and \(\theta_y = -90\) degrees radians, the derivatives in the equations above become identical, and the three rotation parameters no longer represent independent rotations.

\[
\frac{\partial}{\partial \theta_x} \begin{bmatrix} c_{\theta_x} \\ c_{\theta_y} \\ c_{\theta_z} \end{bmatrix} \bigg|_{\theta_x=\frac{\pi}{2}, \theta_y=-\frac{\pi}{2}} = 
\begin{bmatrix}
  (s_z)(w_y) + (c_z)(w_z) \\
  (-c_z)(w_y) + (s_z)(w_z) \\
  0
\end{bmatrix}
\]

\(\frac{\partial}{\partial \theta_y} \begin{bmatrix} c_{\theta_x} \\ c_{\theta_y} \\ c_{\theta_z} \end{bmatrix} \bigg|_{\theta_x=\frac{\pi}{2}, \theta_y=-\frac{\pi}{2}} = 
\begin{bmatrix}
  (s_z)(w_y) + (c_z)(w_z) \\
  (-c_z)(w_y) + (s_z)(w_z) \\
  0
\end{bmatrix}
\]  
(6.13)

This type of degeneracy is known as gimbal lock because it mimics a physical limitation present in the mechanical rotational device known as a gimbal. Gimbal lock results in the loss
of one rotational degree-of-freedom at the degeneracy, and causes the parameterisation to be unstable in the neighbourhood of the degeneracy. The degeneracy in the \([t_x, t_y, t_z, \theta_x, \theta_y, \theta_z]^T\) parameter space are of no consequence, provided the actual pose of the patient is known not to lie in the neighbourhood of a degeneracy. This is guaranteed whenever all three rotation angles are small, as will be in this case of anatomical pose estimation.

### 6.5 Acquisition Geometry

As mentioned earlier, the calibration of the x-ray imaging device has to be known in order to perform a reasonable registration. Therefore the intrinsic camera parameters, reflecting the image source-detector arrangement have to be computed. If they cannot be retrieved solely based on technical specifications available for the device, some camera calibration is needed. The result of this procedure is the knowledge of the correct mapping of any 3D points in space onto the 2D image plane of the imager, or in other words the pixel coordinates in the resulting image. For x-ray devices it is always a perspective projection, which can be expressed in homogenous coordinates with a 3x4 matrix. A very comprehensive coverage of modelling and projective camera calibrating is provided by [97].

In this research the fluoroscopic image acquisition is described through a pinhole camera model [97], that represents mapping from a 3D scene onto a 2D image (Figure 6.8).

![Figure 6.8 Pinhole camera model showing camera centre (\(O_c\)), principal point (\(p\)), focal length (\(f\)), 3D point (\(X\)) and its image projection (\(x\)).](image)

The x-ray center of projection is defined as \(O_c\) and the imaging plane is located at \(z_c = \textit{focal length} (f)\). The ray from the center of projection to the image plane is called the
principal axis, and the image of the center of projection on the image plane is called the principal point \((p)\). The pinhole camera model dictates that 3D object points \(X_i\) map onto the imaging plane through the intersection of rays that originate from the centre of projection and pass through \(X_i\). Geometry and similar triangles shows that the image coordinates \((u,v)\) are related to the object coordinates \((x_o,y_o,z_o)\) through,

\[
\begin{align*}
    u_i &= \frac{f}{z_o} x_o \quad \text{and} \quad v_i = \frac{f}{z_o} y_o
\end{align*}
\] (6.14)

The relationship between the coordinates of a 2D image pixel, and its corresponding 3D object point can be expressed through the equation below,

\[
    x = MX \quad \text{where} \quad M = K[R|t]
\] (6.15)

Here the 3x4 projection matrix \(M\), relates any 3D point \(X = (x_o,y_o,z_o,1)^T\) to its corresponding projection \(x = (u,v,1)^T\) in the acquired image. The intrinsic projection parameters of the x-ray tube (focal length and principal point coordinates), are represented through \(K\), and the extrinsic parameters (rotation and translation of the acquisition system in a world coordinate system) through \([R|t]\). Here \((u_o,v_o)\) denote the coordinates of the principal point. The principal point is the point at which the light that passes through the image, and is perpendicular to the image. The pixel sizes along the \(u\) and \(v\) axes is denoted by \(pix_u\) and \(pix_v\), respectively.

\[
    K = \begin{bmatrix}
        \frac{f}{pix_u} & 0 & u_o & 0 \\
        0 & \frac{f}{pix_v} & v_o & 0 \\
        0 & 0 & 1 & 0
    \end{bmatrix}
\] (6.16)

The camera calibration methodology is not an integral part of this chapter. The intrinsic camera calibration method simply relies on the technique described in [97] where a set of coplanar points visible on the image, are used to compute the camera parameters. We utilised a radio-opaque board with spherical metal markers applied to it on both sides. The respective positions of the markers on the fluoroscopic images are identified through segmentation.
order to obtain the intrinsic parameters, a set of views of the calibration board at different
poses is used. First the parameters are approximated by a closed-form solution which
disregards lens distortion. Then the reprojection error is minimised using gradient descent
yielding the final values for the intrinsic parameters. These are used to undistort the
fluoroscopic images and to compute the extrinsic parameters for single radiographs.

The extrinsic parameters relating to the orientation $R$ and position $T$ of the acquisition
system are computed for different c-arm orientations. The c-arm orientation is described by
two anatomical angles, cranio-caudal ($\alpha$) and right/left anterior ($\beta$) orientation. And a
translational adjustment between the source and the imaging object (Source to Image
Distance, or SID). The SID and the $\alpha/\beta$ angles are measured in real-time by sensors (Figure
6.9). The extrinsic parameters of the c-arm are modelled as a function of $\alpha$ and $\beta$ angles.
Here, $R_o$ is the rotational matrix describing the initial ($\alpha=0$ and $\beta=0$) local camera frame with
respect to the global coordinate frame. $T_o$ is the initial ($\alpha=0$ and $\beta=0$) translation of the center
of projection in global reference frame coordinates. $R_{ai}$ is the rotational transformation due to
rotation of $\alpha_i$ about the resulting alpha axis. $R_{bi}$ is the rotational transformation due to rotation
of $\beta_i$ about the constant beta axis. It is assumed that both axes are orthogonal and intersect.

\[
R = R_o^T R_{\beta_i} R_{\alpha_i} T \quad R_{\alpha_i}^T \quad (6.17)
\]

\[
T = R_o^T (R_{\beta_i} R_{\alpha_i} - R_{\beta_i} R_{\alpha_i}) T_0 \quad (6.18)
\]
6.6 Extraction of Anatomical Surface Model

The patient-specific 3D anatomical bony structures required for the intra-operative registration may need to be pre-operatively extracted from CT data. Bony anatomies appear as high intensity regions in the CT slice images. Reconstruction of 3D models from CT volumes consists of a two-step process. Firstly an image slice segmentation is required for the extraction of the contours of the anatomical structures of interest and secondly surface reconstruction is performed by utilising a surface triangulation technique such as the marching cubes algorithm [62, 63]. Femoral contour extraction is a difficult process due to the variation in the bone intensity in the data and the lack of a distinct joint space. The interior structure of the femur contains trabecular bone with a shell of cortical bone. The image intensity varies between these two bone structure types. Furthermore, weakly defined joint space, around the femoral head and the pelvic bone also adds to the complexity of the contour extraction process.

Several techniques that appear in literature were considered to perform the segmentation and contour extraction. Atlas based segmentation techniques merge image segmentation with image registration [10]. However, fractured bone segmentation is not possible through this technique. Region growing approaches can segment whole skeletal structures, but cannot
separate individual bones [10]. Thus a Level-Set based segmentation technique was developed to provide adequate flexibility to segment fracture bones while ensuring the ability to extract only femoral bone fragments of interest. Level-Set methods are a type of finite element approach used for the modelling of evolving curves or surfaces. These methods have been widely used in the fields of fluid mechanics and material science for some time and have recently been applied within the field of machine vision for segmentation problems [100].

The principal idea behind level set methods is the definition a static, evenly spaced mesh in an image or volume space. The values at each point on the mesh relate to the proximity of the mesh point to an evolving curve or surface with the level set of zero defining the location of the curve or surface. Mesh points contained within the evolving surface are given negative values and mesh points outside the surface are given positive values. An evolving speed function for the movement of the curve or surface is defined and mesh values are updated (from an initial value) using a discrete time finite element update as described below. Let the moving interface be \( \Gamma(t) \in [0,1] \) and the level set function as, \( \phi(x) = \mp d \), where \( d \) is the distance between the point \( x \) and \( \Gamma(t) \), and the plus/minus sign is chosen depending on whether the point is outside/inside the interface.

\[
\phi_{t+1} + F|\nabla \phi| = 0 \tag{6.19}
\]

In the above equation, \( F \) is the speed of the interface point along its normal direction. The moving interface evolves under the effect of \( F \). It expands when \( F \) is positive, while it contracts when \( F \) is negative. When \( F \) is equal to zero, the interface stops and gives the segmentation result. The method implemented is a 3D level set which forms a single unconstrained 2D surface in 3D space. The formulation of the speed function is given in equation below. The speed function \( F \) controls the deformation of the interface.

\[
F(x,y) = (F_0 \nabla \phi_{x,y} + F_{\epsilon}(x',y') \nabla \phi_{x',y'}) e^{-F(x',y')} \tag{6.20}
\]
Chapter 6 - Intra-Operative Pose Estimation

In the above equation $F(x,y)$ is the force at mesh point $(x,y)$ and $(x',y')$ is the nearest point on the zero level set to $(x,y)$. $F_0$ is the Advection Force term, $F_c$ the Curvature Force term and $F_i$ the Image Force term based on the Gaussian derivative filters.

The image force term $F_i$ is based on a 1D Gaussian derivative filter orientated in the normal direction of the zero level set surface at the point of interest $(x_0, y_0)$. This filter is preferable over a simple edge detection such as a Sobel or Canny edge detector as its increased range means that less defined edges may be detected.

Level set methods have a high computational cost as the nearest point on the zero level set must be found for each mesh point. Narrow-band extensions to level set methods lower the computational cost of the algorithms by only updating the mesh in an area local to the zero level set. These methods require the mesh to be reinitialised every few time steps as the zero level set approaches the mesh points that have not been updated. This re-initialisation is in itself computationally expensive, however the overall computational cost over time was reduced substantially using this method.

The narrow band update method was used in the implementation with a band range of 5 in plane voxel widths. It is useful to note that mesh values may be used to determine how close to the level set the mesh point is and whether it is inside or outside the level set, based on the sign. A relatively large time step is used in our implementation as smaller time steps do not appear to affect the final result and increase the computational cost. The Advection Force is set to unity.

Following the segmentation, a 3D triangulated surface may be formed from the mesh by iso-surfacing with the value zero using the ‘Marching Cubes’ algorithm [62, 63]. This algorithm is based on examining each 8 adjacent points in the mesh and determining the triangulation required.
6.7 Pose Estimation System Overview

Figure 6.10 illustrates the proposed framework and conveys how the 2D-3D registration algorithm is used intra-operatively along with the pre-operatively reconstructed 3D fractured bone model. The 2D-3D registration framework can be split into two distinct phases: 1) frontal and lateral pose estimation, 2) axial pose estimation, which will be detailed in the following chapters. Since the registration is performed through images acquired from the frontal and lateral viewpoints, registration in these planes will be performed highly accurately through the first phase of the 2D-3D registration process. The axial alignment, which is critical as it has a high impact on functional biomechanics of the leg, will be conducted in the second phase. The intra-operative 2D-3D registration framework discussed is independent of how the patient-specific 3D data is acquired. Image processing is initially conducted on the fluoroscopic images to extract the required features of interest. This is followed by the analytical frontal and lateral alignment procedure. The optimisation based axial alignment process is finally conducted to provide the necessary full pose estimation. It must be stated that the optimisation based technique presented will be able to conduct all six degrees-of-freedom alignment if necessary. These stages and the experiments conducted are detailed in the proceeding chapters.

Figure 6.10 The main process steps of the 2D-3D registration framework.
6.8 Chapter Summary

This chapter presented a high-level description of the 2D-3D registration problem and provided some terminology and background information relevant to the proposed work. Additional details, including, specifics about medical image modalities, similarity functions, optimisation techniques and the transformation representation that was used to encode the searched pose parameters, are given. A short summary of the motivation and the basic framework of the alignment approaches investigated were also given.
Chapter 7 2D-3D Registration: Frontal and Lateral Alignment

The frontal and lateral plane registration detailed in this chapter is conducted through a novel analytical feature based technique. This ensures that the registration in these two planes is conducted in a robust and timely manner, as will be illustrated in the experimentation conducted.

The axial alignment algorithm is separately detailed in the subsequent chapter. Even though all three axes must be aligned for full registration, partial tracking through only frontal and lateral alignment is also seen as a clinically beneficial by our medical partners. Thus, bear in mind that even though Chapter 7 and Chapter 8 are sequentially followed they can be utilised individually as standalone components, as will be illustrated within the chapters.

The pose estimation work presented here has been published in [7-9, 37]. Moreover, as an extension, two publications also expand on the use of the pose estimation for robotic navigation [8, 39].

7.1 Introduction

The background to intra-operative pose estimation and the motivations behind it are provided in Sections 6.1 and 6.2. Our collaborators at the Auckland District Health Board and the Auckland Hospital have provided us with a set of user requirements and some key facts on how the procedure is currently performed. Hence when developing the proposed intra-operative 2D-3D registration algorithm these facts were carefully considered.

One key point that was revealed is that during the femur shaft fracture procedure the surgeon initially aligns the bone only in the frontal and lateral viewpoints. Then the
intramedullary nail is inserted and the axial alignment is conducted. The axial alignment is typically conducted utilising external anatomical landmarks (Figure 7.1). One clinical practice is to draw a string from the iliac crests to the toe and to rotate the distal fragment until the string lies over the patella. Another technique is to use the contralateral leg and the x-ray image of the lesser trochanter in relation to the proximal femoral shaft. Pre-operatively, the shape of the lesser trochanter of the uninjured limb (with the patella facing anterior) is analysed and stored. To ensure sufficient axial alignment, clinically the surgeon rotates the proximal fragment around the intramedullary nail using a Schanz screw (while the patella faces anterior), until the outline of the lesser trochanter matches the stored image of the uninjured side.

Figure 7.1 Current clinical axial alignment practices. From left to right: (a) use of a string to ensure biomechanical alignment; (b) use of the contralateral leg pose to ensure alignment, reproduced from [31].

Hence during the first part of the fracture alignment, surgeons only require visual tracking in the frontal and lateral planes, of the bone fragments. This ensures that the initial alignment can be conducted in real-time under a purely analytical feature based alignment. This chapter
proposes original features that are robust to extract and facilitate real-time registration. It also details the fluoroscopic image processing that will be conducted to extract the features of interest.

### 7.2 Fluoroscopic Image Processing

#### 7.2.1 ROI Identification

Region-of-Interest (ROI) identification is helpful as it can eliminate much of the acquisition noise and image artefacts. It further makes certain that unnecessary image processing is not conducted on background pixels ensuring timeliness. ROI identification can be split into two stages: 1) edge detection 2) fluoroscopic image circular ROI detection (Figure 7.2).

The edge detection is performed through a canny based multi stage process. The greyscale image is initially smoothed through a Gaussian convolution. Then the Roberts Cross operator is used to create a gradient image [10]. Through this first derivative operator, greyscale edges (first spatial derivatives) are highlighted and appear as ridges in the gradient image. Non-maximal suppression is subsequently utilised to repress any pixels not considered to be a credible edge. This non-maximal suppression is performed under an adaptive hysteresis based thresholding. Local or adaptive thresholding was used primarily due to intensity non-homogeneities of x-ray fluoroscopic imaging. This ensures that all edges are identified irrespective of background illumination levels. The local window size was set to be 5% of the image pixel dimensions. Hysteresis based thresholding, controlled through two threshold levels, ensures that noisy edges are kept intact and not broken up into multiple segments. The two threshold levels were empirically set to be the 50th and 75th percentile of the gradient magnitude values in the local window (Figure 7.2(b)).

The edge image was next processed to extract the circular viewing area of the fluoroscopic c-arm. Here a Hough transform was employed to identify the parametric circular
form. Circle detection using Hough transform identifies the coordinate of the circle center, \((a,b)\), and its radius, \(r\). The arbitrary edge points will be transformed into a right circular cone in the \(a\), \(b\) and \(r\) parameter space. Circle centers are detected through perpendicular lines emanating from the edge points identified by the edge detector. Peak points where a multitude of perpendicular lines intersect will provide the circle centers. Circle radii can be found by accumulating the edge strengths for discrete radii in the desirable range. Normalisation is performed by dividing through with the radius to ensure a non-dimensional quantity prior to selecting the strongest radius per circle centre. Only the strongest circle centre and radius is extracted from the image. The circular ROI is subsequently cropped to eliminate acquisition noise and artefacts present on the digital fluoroscopic images (Figure 7.2 (c)).

![Figure 7.2 Illustrations of the fluoroscopic ROI identification. From left to right: (a) original greyscale fluoroscopic image; (b) identified edges; (c) detected circle and the chosen ROI with the reduced circle radius.](image)

### 7.2.2 Feature Extraction

The features that are extracted from the fluoroscopic images are the femur shaft lines for the proximal and distal segments (Figure 7.3). The shaft lines are subsequently utilised to 1) initialise the cylindrical fragments (Section 7.3); 2) identify the frontal and lateral axis of rotation of the bone fragments (Section 7.4); and 3) to identify the axial location of the fragments (Section 7.4).

The shaft lines identification is performed through a Hough transform based line identification. The edge image created previously during the circular ROI identification is
utilised for this purpose. Line detection using the Hough transform can be defined in the polar form, where $\theta$ is the polar angle and $r$ is the radius. The arbitrary edge point $(x, y)$ will be transformed into an accumulator array in the $r$ and $\theta$ parameter space. The polar form ensures that the vertical lines are successfully represented (which is not possible through a Cartesian form due to the infinite gradient value). Every edge point $(x, y)$ defines a sinusoidal curve in the $(r, \theta)$ parameter space, which all intersect at the same point in the accumulator array covering the Hough space. The resolution of the accumulator determines the precision with which lines can be detected. For this work, a resolution of 1 pixel for $r$ and 1 degree for $\theta$ was used. The classical Hough transform detects lines given only by the parameters $r$ and $\theta$ and no information with regards to length. Thus, all detected lines are infinite in length. Only the edge points with a previously identified circular ROI was utilised for the Hough transform based line detection.

![Figure 7.3 Feature lines extracted through the Hough transform.](image)

### 7.2.3 Shaft Line Pairing

The feature lines extracted through the Hough transform will next be analysed to identify which pairs potentially form the outer edge of the proximal and distal segments of the femur. Three criteria were utilised to conduct this analysis and classify paired lines as a shaft edge or noise. For each pair of lines the following criteria are evaluated,
Chapter 7 - 2D-3D Registration: Frontal and Lateral Alignment

1) The inter-line distance is compared against the mean and variance of femur shaft cross-sectional anthropometry data gathered by [135]. Thus line pairs that fall within a certain range (mean ± standard deviation) of the anthropometric measurements will be potential shaft lines. The interline distance, \( d \), is calculated as below.

\[
\hat{n}_p = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \frac{b_2 - a_2}{\|b_2 - a_2\|} \\
\begin{bmatrix} n^p_1 \\ n^p_2 \end{bmatrix} = [\hat{n}_p \ -\hat{n}_1]^{-1} \begin{bmatrix} a_1 - \frac{b_2 + a_2}{2} \end{bmatrix}
\]

\[
d = |n^p_p| = \begin{bmatrix} 1 & 0 \end{bmatrix} [\hat{n}_p \ -\hat{n}_1]^{-1} \begin{bmatrix} a_1 - \frac{b_2 + a_2}{2} \end{bmatrix}
\]

Where \( a_1 \) and \( b_1 \) are end points of the longer line segment, \( a_2 \) and \( b_2 \) are the end points of the shorter line segment. \( \hat{n}_1 \) is the unit vector along the longer line segment. \( \hat{n}_p \) is the perpendicular unit vector for the shorter line.

2) The difference in gradient between the two lines in the pair, are examined. With this measure it is assumed that line pairs that have similar gradients are more likely potential shaft lines.

3) Intra-line distance is also used to eliminate noise as longer lines will more likely be part of the femur shaft. Since the lines detected through the Hough transform are infinite in length the technique proposed by [136] is used to identify finite lines. The edge image is searched along the line to identify the finite lines, and then all pixels on that line are removed from the edge image. In this way the algorithm returns finite lines, whose distance will be checked.

These three criteria are evaluated to identify the strongest pair of lines that form a one to one matching, with no duplicates. Depending on the clinical placement of the anatomy the line pair on one side will be classified as the proximal and the other side as the distal. In our clinical setup we have the right line pair classified as the proximal and the left as the distal.
Chapter 7 - 2D-3D Registration: Frontal and Lateral Alignment

Figure 7.4 Illustrations of the shaft line identification process. From left to right: (a) original lines identified through the Hough transform; (b) shaft line pairs short listed through the weighted criteria evaluation. (b) identified edges within the ROI; (c) detected lines through the Hough transform.

7.3 Initialisation of the Cylindrical Segments

Cylinders can be parameterised through a radius and a length. The length of the cylindrical fragments can be set depending on the field of view of the c-arm imaging system. The length was set to be 10 cm as empirical testing conducted illustrated that this be the best length for surgical visualisation purposes.

The radius required for the proximal and distal segments were calculated from the shaft lines previously identified. Thus the average inter-line pair distance (averaged from the frontal and lateral views) was used as the diameter for the cylindrical fragments. Anthropometrically this was typically between 1 cm and 2 cm.

7.4 Frontal and Lateral Pose Estimation

The task involved with the frontal and lateral pose estimation is to utilise the extracted features from the target fluoroscopic images and the projected contours of the model, to align the rotational axis of each bone fragment. Figure 7.6 illustrates the input data modalities involved. Figure 7.5(a) shows the projected contours and the mid shaft axis of the 3D model. The aim is to align the mid shaft axis in each plane with the extracted mid shaft axis of the target image, as shown in Figure 7.5(b).
Chapter 7 - 2D-3D Registration: Frontal and Lateral Alignment

Figure 7.5 Illustration of the data inputs required by the frontal and lateral pose estimation process. From left to right: (a) projected contours of the model bone; (b) features extracted from the fluoroscopic images. Features include the mid shaft line of the distal segment and the arbitrary point that lies on the mid line.

The frontal and lateral alignment is driven by the outer shaft lines identified through the fluoroscopic image processing. The inputs required for this alignment are:

1) 2D gradient directional vectors of the shaft in the frontal and lateral planes, and,

2) Two arbitrary points that lie anywhere on the mid shaft axis in those planes.

The mid shaft line is calculated by averaging the outer shaft lines in each plane. The required directional vectors and the arbitrary points are next extracted. This mid shaft line is then projected perpendicular to the plane. The intersection of the perpendicular projections from the frontal and lateral planes will form the axis of rotations for a particular segment (Figure 7.5). The gradient information of the mid shaft line is utilised to calculate the normal vectors of these projected planes by obtaining the cross product between the direction vectors of the mid shaft lines and the imaging plane normals. These normal vectors, $n_{\text{frontal}}$ and $n_{\text{lateral}}$ and the arbitrary points identified on the lines ($P_{\text{frontal}}$ and $P_{\text{lateral}}$) are used below to solve for the vector line equation of the shaft axis of rotation in 3D.
Chapter 7 - 2D-3D Registration: Frontal and Lateral Alignment

$M_{axis}$ defined below is the direction vector of the 3D axis of rotation.

$$M_{axis} = \text{Null} \begin{bmatrix} n_{frontal}^T \\ n_{lateral} \end{bmatrix} \quad (7.4)$$

$P_{axis}$ defined below is the anchoring point identified that lies on the 3D axis of rotation. This anchoring point is utilised to ensure that the translations in the frontal and lateral planes are identified.

$$P_{axis} = \left( \text{pinv} \begin{bmatrix} n_{frontal}^T \\ n_{lateral} \end{bmatrix} \right) \begin{bmatrix} n_{frontal}^T P_{frontal} \\ n_{lateral}^T P_{lateral} \end{bmatrix} \quad (7.5)$$

Once the axis of rotation is identified for the target images, the model axis of rotation can be aligned with it utilising the equation below. Here $modelT_{global}$ and $targetT_{global}$ are homogeneous transformation matrices which relate the local coordinate (reference) frames of the bone segments to a global coordinate frame. The z-axes of the local reference frame coincide with the axial rotation axes of the bone segments, while their origins are anchored on the corresponding feature points in the model and target bones. The final alignment transformation between the model and the target $modelT_{target}$ is then expressed as:

$$modelT_{target} = modelT_{global} target^{-1} \quad (7.6)$$

Figure 7.6 Notations involved in the frontal and lateral pose estimation process.
7.4.1 Experiments

We conducted two experiments to evaluate the robustness and accuracy of the proposed approach. The first experiment was to evaluate the fundamental proof of concept of the proposed registration technique. This experiment provided a controllable environment within which the technique’s accuracy and repeatability was tested. The second experiment was conducted with actual fluoroscopic images acquired through a cadaver. These images were identical to those acquired clinically and thus provided a practical testing environment for the proposed framework.

The first proof of concept study was performed with a distal fragment of an artificially fractured plastic bone (Figure 7.7). The fragment was attached to a Stewart platform and moved through a series of rotations in the frontal ($R_x$) and lateral ($R_y$) directions. Fluoroscopic imaging was emulated through a USB camera mounted on a rotary arm (similar to a fluoroscopic c-arm). Even though the effects of variation of x-ray absorption in tissues such as muscle, fat, and skin cannot be simulated with this test setup, effort was taken to closely match the actual clinical process steps undertaken. The proposed pose estimation technique was applied to this test setup and 20 iterations of the same test were conducted (Figure 7.8). The average absolute mean $R_x$, $R_y$ and $R_z$ errors were 0.567, 0.487 and 0.604 degrees, respectively. The standard deviation of errors was 0.098 degrees. The results further convey that the registration can be performed at sub-degree accuracy and with highly robust repeatability.
Chapter 7 - 2D-3D Registration: Frontal and Lateral Alignment

Figure 7.7 The test rig setup.

Figure 7.8 Stewart platform testing output. The top graphs indicate the results obtained in the frontal \( R_x \) and lateral \( R_y \) directions. The red lines (solid) indicate the rotation of the Stewart platform, while the blue lines (dashed) indicate the identified rotation in the \( R_x \) and \( R_y \) directions respectively. The graphs below indicate the corresponding absolute registration error in degrees.

The second test conducted is a clinical evaluation of the proposed system. The CT data and the fluoroscopic imaging required for the testing was provided by the Institute for Robotics and Process Control at the Technical University of Braunschweig, under the robotically assisted long bone fracture reduction project [5]. Initially a CT scan of a fractured femur was conducted to obtain an accurate model of the bone. This will ensure that any inaccuracies present are directly correlated to the pose estimation algorithm being tested. The fractured femur was next moved through a series of clinical reduction-like steps and frontal and lateral fluoroscopic images acquired at each pose. These sequential fluoroscopic images and the original model of the fractured bone was utilised in the testing. The key accuracy
measurement used during the testing is the average Euclidean distance between the closest points of the reconstructed data and the ground truth CT scan data set.

The quantitative results of the testing are displayed below in Table 7.1. Figure 7.9 shows two qualitative results for Bone A, Pose ID 1 and Pose ID 3.

Figure 7.9 Two qualitative results showing the identified 3D pose and the superimposition of the pose on the two fluoroscopic images.
Table 7.1 Patient-specific fractured bone frontal and lateral pose estimation results in terms of mean and standard deviation of error.

<table>
<thead>
<tr>
<th>3D Fractured Bone Pose Estimation Error</th>
<th>Bone A Mean (mm)</th>
<th>Bone A SD (mm)</th>
<th>Bone B Mean (mm)</th>
<th>Bone B SD^3 (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pose ID 1</td>
<td>2.09</td>
<td>0.56</td>
<td>2.05</td>
<td>1.03</td>
</tr>
<tr>
<td>Pose ID 2</td>
<td>2.38</td>
<td>0.56</td>
<td>2.75</td>
<td>0.47</td>
</tr>
<tr>
<td>Pose ID 3</td>
<td>1.73</td>
<td>0.87</td>
<td>2.55</td>
<td>0.63</td>
</tr>
<tr>
<td>Pose ID 4</td>
<td>2.15</td>
<td>0.61</td>
<td>2.87</td>
<td>0.70</td>
</tr>
<tr>
<td>Pose ID 5</td>
<td>2.90</td>
<td>0.52</td>
<td>2.54</td>
<td>0.62</td>
</tr>
<tr>
<td>Pose ID 6</td>
<td>2.42</td>
<td>0.65</td>
<td>2.17</td>
<td>0.76</td>
</tr>
<tr>
<td>Pose ID 7</td>
<td>1.74</td>
<td>0.74</td>
<td>3.05</td>
<td>0.77</td>
</tr>
<tr>
<td>Pose ID 8</td>
<td>3.15</td>
<td>0.95</td>
<td>3.26</td>
<td>0.91</td>
</tr>
</tbody>
</table>

7.5 Chapter Summary

This chapter presented one portion of the proposed 2D-3D registration framework. Frontal and lateral plane registration was detailed in this chapter. An analytical solution to the registration is presented which ensures that the alignment is completed robustly and in a timely manner. The feature extraction process from fluoroscopic imaging is thoroughly detailed and the testing conducted conveyed robustness. The second portion of the 2D-3D registration, the axial alignment, is subsequently discussed in Chapter 8.
Chapter 8 2D-3D Registration: Axial Alignment

This chapter presents a developed algorithm for recovering the position and orientation of the target anatomy in 3D space based on an iterative comparison of 2D planar radiographs with the pre-operative 3D model, through an optimisation process. More specifically, this system uses x-ray fluoroscopic images acquired intra-operatively, and iteratively compares them with projection image generated from the 3D model. The projections are generated based on an estimate of the position and orientation of the patient’s anatomy and this estimate is progressively updated throughout the course of the iterations.

The chapter provides details on the objective functions and the optimisation procedures used in the 2D-3D registration. The gradient search based optimisation algorithm, which is used to locate the optima of the objective function, is also discussed. Full details necessary to recalculate the update terms are given, which are vital to finding the desired alignment transformation. The chapter also details the image processing that is performed to extract the anatomical features of interest. The optimisation process is run under two conditions that ensure successful convergence and timeliness. Firstly, a multi-resolution framework is employed that involves first searching for the solution from a highly sub-sampled projected contour and then refining the solution through a series of reductions to the sub-sampling level. Secondly a variable step size based implementation is employed where the rotation and transformation steps start with large sizes and iteratively lessens in size as the algorithm nears its optimal solution.

The pose estimation work presented in this chapter has been published in [7-9, 37, 39], including areas relating to pose estimation for robotic navigations.
8.1 Introduction

Please refer to Sections 6.1 and 6.2 for a background into intra-operative pose estimation and the motivations behind the work. Figure 8.1 illustrates the proposed optimisation based 2D-3D registration algorithm. As previously mentioned this technique can be used solely for axial alignment or to perform all six degrees-of-freedom parameter identification.

For this registration, the target pose information is extracted through intra-operative 2D fluoroscopic imaging. The model information is obtained through the pre-operative 3D bone model, where images can be projected onto the intra-operative imaging plane (typically frontal and lateral). Registration is achieved by finding the right position and orientation with respect to the 3D model, so that a projected feature data computed at this pose matches features extracted from the target 2D image. The alignment of the projected contour and the target fluoroscopic image is assessed by a similarity measure.

For a feature based registration, this measure operates on extracted features returning a scalar value that informs us on how well the images are matching. This similarity measure is the objective cost function in the optimisation algorithm, which alters the rigid transformation in order to achieve a better alignment. It tries to find a transformation where the projected
features have the maximum similarity with the fluoroscopic image. This results in a search for a global optimum in up to a six-dimensional parameter space.

8.2 Fluoroscopic Edge Extraction

The edge extraction technique is detailed thoroughly in Section 5.2. The edge detection is performed through a canny based multi stage process. The greyscale image is initially smoothed through a Gaussian convolution. Then the Roberts Cross operator is used to create a gradient image. Through this first derivative operator greyscale edges (first spatial derivatives) are highlighted and appear as ridges in the gradient image. Non-maximal suppression is subsequently utilised to suppress any pixels not considered to be a credible edge. This non-maximal suppression is performed under an adaptive hysteresis based thresholding. Local or adaptive thresholding was used primarily due to intensity non-homogeneities of x-ray fluoroscopic imaging. This ensures that all edges are identified irrespective of background illumination levels. The local windows size was set to be 5% of the image pixel dimensions. Hysteresis based thresholding, controlled through two threshold level, ensures that noisy edges are kept intact and not broken up into multiple segments. The two threshold levels were empirically set to be the 50th and 75th percentile of gradient magnitude values in the local window.

8.3 Optimisation Procedure

By its definition, image registration is concerned with finding a geometrical transformation that brings one image into the best possible spatial correspondence with another image or physical space by optimising a registration criterion. The parameters that describe a geometrical transformation can be computed directly or searched for. Direct computation of transformation parameters is possible only when the one-to-one mapping between the limited set of points found in both images is known. Generally, point pairs are
not available. Therefore, registration parameters (similarity measures) have to be searched for iteratively by minimising the distance between corresponding feature sets.

Through a suitable similarity measure, the best alignment parameters can be located with the help of an optimisation procedure. Such a protocol is responsible for modifying the current parameter estimates in a way that the similarity function eventually takes on its (global) extremum. It must be noted that the objective functions can be a reward based similarity measures or a cost based dissimilarity measure. Throughout the chapter these two terms are used interchangeably, since one is the reciprocal of the other. Hence the best alignment is assigned the highest score and an optimisation procedure aims to maximise the objective function. In the case of whole-image-content-based registration the registration criterion is formulated as a similarity measure defined in multidimensional space of searched parameters. It is desired that the similarity measure is monotone and quasi-convex and in the vicinity of the true registered position.

There are two major types of strategies, seen in literature, that perform the optimisation task: non-gradient and gradient methods. Non-gradient strategies execute a local search in the parameter space by evaluating the objective function at different locations according to a pattern, while gradient procedures use the gradient information to indicate the direction to the desired optima. The former strategy might be easier to implement as it requires only the evaluation of the objective function and no additional computations to derive the consecutive search directions. However, the latter could potentially be much faster as its search is specifically guided towards the optima.

The commonly implemented optimisation techniques are Powell’s method [23, 80, 88, 90, 137, 138], the Downhill Simplex method [134, 139] and Gradient Decent [89, 137]). Moreover, the Monte-Carlo random sampling in conjunction with Powell’s optimisation method was investigated in [140], while [141] implemented the Kalman Filter and [129] employed simulated annealing optimisation. When the formulation of a similarity measure
allows an implicit estimation of its second order derivatives the Levenberg-Marquardt least square optimisation can be implemented [82, 133, 142].

To avoid registration being trapped in a false local optima and to speed up the registration process, different whole-image-content-based methods often implement a hierarchical coarse-to-fine, multi-scale and multi-resolution search strategy [19, 134, 138, 143].

8.4 Optimisation Formulation

The parameters of transformation $m^Tt$, are represented by a vector $a$ (This parameter vector was termed $x$ earlier in the thesis, however $a$ has been used in this chapter to avoid confusion with another notation). The parameter vector is specified as, $a = [\theta_x, \theta_y, \theta_z, t_x, t_y, t_z]$.

The inputs to the registration problem consist of the set $M$ of 3D points on the surface mesh of the 3D model and the set $C$ of 2D image points on the image boundary contour. The set $C \subset M$ of 3D model points, $m$, that project to the bone contour in the image plane depends on the rotation $R$, translation $T$, and projection $P$ of the 3D model. Thus, 3D-2D registration with unknown correspondence can be formulated as the problem of determining $R$ and $T$ that minimises error $E$. The problem can be written as follows,

$$E(a, \varphi) = \sum_{i=1}^{N} \|P(R(m_{\varphi(i)}) + T) - t_i\|$$

(8.1)

Where $m \in M$ are 3D model feature points and $t \in C$ are target 2D image feature points. The optimal $R/T$ gives the pose of the 3D model.

In order to measure alignment, correspondence between the projected model and target data points needs to be identified. This correspondence registration algorithm is denoted by the function, $\varphi(i)$, which selects, for each 2D target image point, $t_i$, the corresponding model point, $m_{\varphi(i)}$.

The algorithm to minimise this error can be summarised as follows,
Chapter 8 - 2D-3D Registration: Axial Alignment

1) Extraction of target image features (boundary contour) (Section 8.2).

2) Initialisation of the 3D model. This involves setting the initial pose. If the pose cannot be initialised then scale the 3D model so that the width of the bounding box of its 2D projection is equal to that of the image boundary contour and translate the 3D model so that the center of its 2D projection coincides with that of the image boundary contour.

3) Conduct the optimisation process until convergence:

- Apply rigid transformation on the 3D model.

- Project the 3D model onto the image plane.

- Extract the features from the projected model silhouette.

- Perform a similarity measure, 2D-2D (or 3D-3D), with unknown correspondence between model contour points and image contour points (Section 8.5).

- Identify the pose that ensures 2D-3D registration between 3D model and the target patient fluoroscopic image, and update rigid transformation parameters.

The algorithm is repeated until the error $E$ is minimised and does not change significantly. As will be shown through the experimentation, the algorithm can converge to a (global) optimal solution for most test images. For a small number of cases in which the algorithm does not converge, it can be terminated after running for a fixed number of iterations. In all cases, the best registration result is kept for each iteration, and it is reported as the final registration solution when the algorithm terminates.

Please note that the optimisation procedure discussed herein can be used for axial alignment only or to identify all six degrees-of-freedom. The experiments conducted have tested both these scenarios.


8.5 Objective Functions

The objective function must aid in the parameter estimation to solve the problem of registering model features with extracted target features. The model feature points are noted as a set $M$ features, $m_i = \{m_1, ..., m_M\}$ which have to be registered to a set of $C$ target feature points, $t_i = \{t_1, ..., t_C\}$, extracted from the fluoroscopic image.

In order to generate a projective image from a 3D model volume, each surface point of the volume $M_i^w$ in the world coordinate system is mapped onto a point $m_i^p$ in the pixel coordinate system.

$$m_i^p = P \cdot M_i^w = K \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot D \cdot M_i^w \quad (8.2)$$

The points $M_i^w$ and $m_i^p$ are given in homogeneous coordinates. The projection matrix $P$ contains the matrix $K$, which consists of the intrinsic camera parameters, a further matrix, which reduces the number of dimensions, and the matrix $D$, which consists of the extrinsic camera parameters.

Prior to the projection, a transformation according to the parameter vector $a$ is executed on the volume (The parameter vector is specified as, $a = [\theta_x, \theta_y, \theta_z, t_x, t_y, t_z]$). Therefore, the above equation is extended to,

$$m_i^p = P \cdot T \cdot M_i^w \quad (8.3)$$

Here the transformation matrix $T$, which models the translations and rotations of $a$, is performed onto the points of the model. Afterwards, the projected model features produced from the transformed volume (according to the current transformation $a$) is compared to the extracted fluoroscopy image by applying a similarity measure. The objective function discussed in this section document those similarity measures. By using a certain optimisation technique, the similarity between the projected model and extracted target features will be increased until the optimal parameter vector $a$ is found.
8.5.1 Projection Strategy

The background technical details for the projection and back projection based strategies are presented in Section 6.4.2. In a projection based strategy, the 3D data is projected onto \( N \) planes \( \Omega_{Bj} \), \( j = 1,2,...,N \), of the intra-interventional 2D projection data \( B_j(x^{2D}_{Bj}) \) using the \( N \) projection matrices \( P_j \), associated with \( B_j(x^{2D}_{Bj}) \), \( j = 1,2,...,N \). The registration is then performed by optimising the sum of similarity functions \( CF_j^{2D} \), \( j = 1,2,...,N \):

\[
m_T^* = \arg \max_T \sum_{j=1}^{N} CF^{2D}_j \left( P_j \left( A^T(x^{3D}_A), B_j(x^{2D}_{Bj}) \right) \right)
\]  

(8.4)

These similarity function(s) are applied on a 2D domain. In literature they are categorised as “shape matching” techniques, since they aim to judge the similarity of two 2D object shapes.

Literature on shape matching is extensive, with the biggest contribution coming from research involving content based image retrieval (for comparing images). There are a multitude of shape matching measures proposed in literature. Several of these were tested throughout this research and a final decision was made on a distance based similarity measure. A common shape similarity measure proposed in literature is one based on Fourier descriptors. Fourier descriptors are the normalised coefficients of the Fourier transformation, typically applied to a ‘signal’ derived from samples from the contour, such as the coordinates represented by complex numbers. Our experiments and external literature have shown that the centroid distance function, the distance from the contour to the centroid, is a signal that works better than many others [144]. Another similarity measure is the shape context based measure (introduced in Section 3.6.1). Shape context [105] is a method that first builds a shape representation for each contour point, using statistics of other contour points ‘seen’ by this point in quantised angular and distance intervals. The obtained view of a single point is represented as a 2D histogram matrix. To compute a distance between two contours, the correspondence of contour points is established that minimises the distances of corresponding
matrices. Furthermore, literature also suggests image edge orientation histogram based techniques [145]. The histogram is built by applying an edge detector to the image, then going over all pixels that lie on an edge, and making a histogram of the local tangent orientation. For a full review of the shape based similarity measure, please refer to [144].

The similarity measure proposed in this thesis is based on a distance transform: Hausdorff Distance. This distance based dissimilarity measure is ideal for situation where several feature points from the model, $m$, have no corresponding point in the target, $t$ (due to occlusion and noise). Typically, the two point sets are of different size, so that no one-to-one correspondence exists between all points. In that case, a dissimilarity measure that is often used is the Hausdorff distance. The Hausdorff distance is defined not only for finite point sets, but it is also defined on nonempty closed bounded subsets of any metric space.

The directed Hausdorff distance $h(m, t)$ is defined as the lowest upper bound (supremum) over all points in $m$ of the distances to $t$. The directed Hausdorff distance from $m$ to $t$, denoted by $h(m, t)$ is,

$$\max_{m \in M} \left( \min_{t \in T} \|m - t\| \right)$$  \hspace{1cm} (8.5)

The Hausdorff distance between $m$ and $t$, denoted by $H(m, t)$ is,

$$\max\{h(m, t), h(t, m)\}$$  \hspace{1cm} (8.6)

Intuitively, the function $h(m, t)$ finds the point $m \in M$ that is furthest from any point in $t$ and measures the distance from $m$ to its nearest neighbour in $t$.

For finite point sets, the Hausdorff distance can be computed using the Voronoi diagrams in time $O((m + n) \log(m + n))$.

However, as is apparent from its definition, the Hausdorff distance is sensitive to noise and outliers. A modification of the measure that is less sensitive to noise is the partial Hausdorff distance, defined as,
Chapter 8 - 2D-3D Registration: Axial Alignment

\[ H_k(m,t) = \max\{h_k(m,t), h_k(m,t)\} \quad (8.7) \]

where \( h_k(m,t) \) is the \( k^{th} \) value in increasing order of the distance from a point in \( m \) to \( t \). Thus, \( h_k(m,t) = k_{\text{in}} \text{min}_{t \in T} d(m,t) \). The experiments section will further document the proposed projective similarity measure.

Figure 8.2 illustrates the objective function for variations of the three rotational degrees-of-freedom and three translational degrees-of-freedom. The registration used anterior and lateral target images, acquired at the neutral position of \( \theta_x, \theta_y, \theta_z, t_x, t_y \) and \( t_z \) being set to zero. The objective function clearly shows a minimum at these three values.

![Figure 8.2](image-url)

(a) Rotation

(b) Translation

Figure 8.2 The objective function output of the projection strategy for rotational variation (\( \theta_x, \theta_y, \) and \( \theta_z \)) shown in (a) and translational variations (\( t_x, t_y \) and \( t_z \)) shown in (b).
Figure 8.3 The objective function output of the projection strategy for rotational variation (θₓ, θᵧ, and θz) and translational variations (tₓ, tᵧ, and tz) shown individually.

Analysing the similarity measure within the six dimensional search space is extremely difficult and usually just one or two dimensional plots are used to give a visual impression of
the measure’s behaviour. The plots of similarity measure were obtained from the data of an intact femur and two perpendicular x-ray images (frontal and lateral). For the plots five of the six transformation parameters were fixed to their “gold standard” values and similarity measures were calculated while varying the sixth parameter around its “gold standard” value.

The one-dimensional plots in Figure 8.3 illustrate the behaviour of similarity measure, when matching a 3D model with two perpendicular 2D images and varying one transformation parameter at a time. Peak values in these plots are very close to the “gold standard” (zero) position. This indicates that the similarity measure optimum is extremely close to true registration and that the registration method using this similarity measure can be very accurate. The behaviour of the similarity measure below 5 mm or 6° around “gold standard” position is very smooth. This leads to the conclusion that if local search strategy is used the capturing range of the registration method is at least 5 mm or 6°.

8.5.2 Back-Projection Strategy

Back-projection based techniques require each of the $N$ 2D intra-interventional data $B_j(x_{Bj}^{2D})$ to be back-projected into the 3D space using corresponding back-projection matrices $BP_j, j = 1,2,..,N$. Similarly to the projection strategy, the registration is then performed by optimising the sum of criterion functions $CF_j^{3D}, j = 1,2,..,N$, each calculated on the basis of the 3D pre-interventional data and one of the back-projected intra-interventional 2D data,

$$m\hat{T}_t = \arg \max_T \sum_{j=1}^{N} CF^{3D} \left(A^T(x_{A}^{3D}), B_j(x_{Bj}^{2D})\right)$$

(8.8)

The back projection strategy proposed, involves a set of 3D lines (bundles of x-ray paths leading from the point source to the imaging plane’s contour) whose equations are known in the intra-operative coordinate system and a set of corresponding 3D points (closest points on 3D model surface). The task is to determine the transformation that brings each point as close as possible to its corresponding line (Figure 8.5 and Figure 8.6).
Thus the optimisation minimises the sum of squared distance between a set of \( m \) 3D surface points, \( p_i \), and the corresponding \( m \) 3D line, \((c_i, v_i)\), where \( c_i \) is the center of perspective and a directional unit vector \( v_i, \|v_i\| = 1 \).

\[
\sum_{i=1}^{m} (p_i - (c_i + v_i \cdot d))^2
\]

**Figure 8.4** The distance measure used in the back-projection strategy.

The distance \( d_i \) between the 3D point, \( p_i \), and the \( i^{th} \) 3D line, \((c_i, v_i)\), is obtained by taking the norm of the cross product between the vector joining \( c_i \) and \( p_i \) and the unit vector, \( v_i \), that specifies the direction of the line (Figure 8.4).

\[
d_i = \| (p_i - c_i) \times v_i \| \quad (8.9)
\]

Figure 8.7 illustrates the objective function for variations of the three rotational degrees-of-freedom and three translational degrees-of-freedom. The registration used anterior and lateral target images acquired at the neutral position of \( \theta_x, \theta_y, \theta_z, t_x, t_y \) and \( t_z \) being set to zero.

**Figure 8.5** Illustration of the back-projected lines from the x-ray contour onto the 3D model surface.
Figure 8.6 Two illustrations showing visibly misaligned regions (a) where the projected rays are not tangential to the surface. The alignment will ensure that the projected rays encapsulate the model so that they are tangential to the surface, as in (b).

Figure 8.7 The objective function output of the back-projection strategy for rotational variation ($\theta_x$, $\theta_y$, and $\theta_z$) shown in (a) and translational variations ($t_x$, $t_y$ and $t_z$) shown in (b).
8.5.3 Similarity Measure Evaluation: Protocol and Experimentation

The evaluation criteria utilised for the similarity measures is necessary to judge the proposed similarity measures and identify the strengths and weaknesses in clinically applicable scenarios. The testing requires the “gold standard” registration of these images known prior to the experimentation. The spatial transformation $T$ that is supposed to bring the two featured datasets, the model and target, into correspondence is assumed to be rigid and is therefore composed of three translational ($t_x$, $t_y$, $t_z$) and three rotational ($\theta_x$, $\theta_y$, $\theta_z$) parameters. The six-dimensional parametrical space is first normalised so that equal changes of each of the six parameters in the normalised parametrical space will have approximately equal impact on the transformation magnitude. By normalising the parametrical space, Euclidean metrics may be used to determine distances from the position at which the images are “best” aligned and where a similarity function should have its optimum. Let the origin $X_0$ of the six-dimensional parametrical space be at the known “gold standard” position and let $SM(X)$ be the value of a similarity measure for the spatial transformation defined by location $X; X = [x_1, \ldots, x_K]$ in this space. Similarity measure values $SM(X_{n,m})$ are defined for image pairs, with the target image at the origin $X_0$ and the floating image transformed from the origin to location $X_{n,m}$. Values $SM(X_{n,m}), n = 1, 2, \ldots, N; m = -M/2, \ldots, M/2,$ of a similarity measure are defined on $N$ lines probing the six-dimensional parametrical space and at $M+1$ points evenly spaced along each line. Each of the $N$ lines is defined by a randomly selected starting position $X_{n,-M/2}$ at a distance $R, R = \text{abs}(X_{n,M/2})$, from the origin and its mirror point $X_{n,M/2}$. To make the similarity measure invariant to the absolute scale, each original similarity measure $SM_o(X_{n,m})$ is normalised to the interval [0, 1],

$$SM(X_{n,m}) = \frac{SM_o(X_{n,m}) - SM_{omin}}{SM_{omax} - SM_{omin}} \quad (8.10)$$

where $SM_{omin}$ and $SM_{omax}$ are the minimal and maximal values of $NM + 1$ similarity measure values before normalisation, respectively. If a similarity function is such that its minimum is sought for by optimisation, the similarity measure value $SM(X_{n,m})$ is changed to -
Chapter 8 - 2D-3D Registration: Axial Alignment

$SM(X_{n,m})$. Let $X_{n,opt}$ be the position and $SM(X_{n,opt})$ the value of the global optimum of the similarity measure along line $n$, and let $X_{n,loc}$ be the position of the local optimum closest to $X_{n,opt}$.

The behaviour of a similarity measure is assessed by five properties. All properties are statistical estimations, derived from the “gold standard” position, similarity measure values $SM(X_{n,m})$ and positive gradients $d_{n,m}$.

$$d_{n,m} = \begin{cases} 
SM(X_{n,m-1}) - SM(X_{n,m}) & \text{if } m < \text{opt} \& SM(X_{n,m-1}) > SM(X_{n,m}) \\
SM(X_{n,m+1}) - SM(X_{n,m}) & \text{if } m > \text{opt} \& SM(X_{n,m+1}) > SM(X_{n,m}) \\
0 & \text{otherwise}
\end{cases} \quad (8.11)$$

The five properties are:

1) Accuracy, ACC of a similarity measure is defined as the root mean square (RMS) of distances $\text{abs}(X_{n,opt} - X_0)$ between the origin $X_0$ and global optima $X_{n,opt}$, $n = 1,2, \ldots ,N$:

$$ACC = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \|X_{n,opt} - X_0\|^2 \text{[mm]}} \quad (8.12)$$

2) Distinctiveness of optimum, DO is defined as a function of distance $r$, $r = s \cdot \delta x$, from the optimum, where $\delta x$, $\delta x = 2R/M$ is the distance between two consecutive points along a line and $s$, $s = 1,2, \ldots$ is the number of steps:

$$DO(r) = \frac{1}{2rN} \sum_{n=1}^{N} 2 \times SM(X_{n,opt}) - SM(X_{n,opt-s}) - SM(X_{n,opt+s}) \times 10^{-3} \text{[mm]} \quad (8.13)$$

3) Capture range, CR is defined as the minimal distance between the position of optima $X_{n,opt}$ and the closest minimum $X_{n,loc}$:

$$CR = \min \|X_{n,opt} - X_{n,loc}\| \text{[mm]} \quad (8.14)$$

4) Number of minima, NOM($r$) is the sum of minima of the similarity measure within distance $r$ from each of the $N$ global optima, i.e., a cumulative number of minima as a
function of distance \( r \). The average number of minima per line \( NOM(R)/N \) is denoted by \( NOM \).

5) Risk of non-convergence, \( RON(r) \) is the property that describes the behaviour of a similarity measure around the \( N \) global optima. It is defined as the average of positive gradients \( d_{n,m} \) within distance \( r \) from each of the \( N \) global optima:

\[
RON(r) = \frac{1}{2rN} \sum_{n=1}^{N} \sum_{m=opt-s}^{opt+s} d_{n,m} \quad [10^{-6}/\text{mm}] \quad (8.15)
\]

A large value of \( RON(r) \) indicates that a similarity measure has distinctive (deep) or broader local optima in which optimisation may get trapped. \( RON(R) \) is shorter denoted as \( RON \). The better a similarity measure is, the smaller are the values of the accuracy, number of minima, and risk of nonconvergence and the larger the capture range and distinctiveness of optimum values. \( CR, NOM, \) and \( RON \) are the three properties that describe robustness. The better these values are, the more robust is the similarity measure.

In all the experiments, the number of lines \( N \) was set to 50, \( R \) to 35 mm, \( M \) to 140, \( \delta x \) to 0.5 mm, and \( s \) to 1. In all experiments, fluoroscopic images served as the target image and the projected features from the 3D volume of the bone as the model. The better the SM, the smaller its accuracy index is, number of maxima and risk of non-convergence and the higher its capture range and degree of distinctiveness. Comparisons of index values are given in Table 8.1. Here SM1 is the projection strategy Hausdorff distance measure, while SM2 is the back-projection based distance measure.

SM1 estimator is more robust, accurate, and convex than the SM2 estimator. The indices for evaluation given in Table 8.1 depend on the sampling step \( d \) of the hypersphere. Experiments carried out by decreasing the value of \( d \) to the assumed accuracy of SM1, have shown a simultaneous decrease of \( CR \) and increase of \( RON \). This could be explained by the fact that many local maxima are located within a distance \( d \) from the gold standard. Thus it appears that local search optimisation methods are not suited to this function.
Table 8.1 Comparison of evaluation indices for similarity measures SM1 and SM2.

<table>
<thead>
<tr>
<th>Datasets</th>
<th>SM</th>
<th>ACC (mm)</th>
<th>CR (mm)</th>
<th>NOM (δ)</th>
<th>DO(δ) (mm⁻¹)</th>
<th>RON(δ) (10⁻⁰ mm⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SM1</td>
<td>0.8</td>
<td>2.4</td>
<td>2</td>
<td>0.05</td>
<td>149.4</td>
</tr>
<tr>
<td></td>
<td>SM2</td>
<td>2.31</td>
<td>1.4</td>
<td>8</td>
<td>0.03</td>
<td>4522.4</td>
</tr>
<tr>
<td>2</td>
<td>SM1</td>
<td>0.12</td>
<td>5.11</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>SM2</td>
<td>0.63</td>
<td>0.63</td>
<td>1</td>
<td>0.02</td>
<td>182.5</td>
</tr>
<tr>
<td>3</td>
<td>SM1</td>
<td>0.25</td>
<td>7.19</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>SM2</td>
<td>0.25</td>
<td>1.79</td>
<td>4</td>
<td>0.04</td>
<td>42.56</td>
</tr>
</tbody>
</table>

Further experiments were conducted to assess the impact of the number of 2D images used for reconstruction. In this experiment, the properties of the similarity measures were evaluated with respect to the number of projections used for reconstruction. Target image numbers 1, 2, 4, 6 and 8 were used. The results shown in Figure 8.8 indicate that the accuracy and distinctiveness of optimum of all similarity measures improved when more images had been used for reconstruction.

However with more images, the measures became slightly less robust, as indicated by the increase of NOM and RON values. This can be explained with more details presented in 3D images reconstructed from more 2D projections. Because with more images the accuracy improved and the robustness slightly deteriorated, the optimal number of images used for reconstruction is between 2 and 4.
Chapter 8 - 2D-3D Registration: Axial Alignment

Figure 8.8 Five properties of the similarity measures (top to bottom) as a function of the number of two-dimensional images used for registration. ACC - accuracy, DO - distinctiveness of optimum, CR - capture range, NOM - number of minima, RON - risk of nonconvergence.

8.6 Gradient Search Based Optimisation

The task of finding the ideal transformation $T$ is formulated as an optimisation problem. An iterative search is conducted to locate the parameters that produce the highest score according to a similarity objective function. At each iteration of the algorithm, current estimate of transformation $T$ is used to simulate the creation of the projected contours. Next the quality of alignment between projected contours and the segmented fluoroscopic image is
judged through the similarity measure (or dissimilarity measure). Then, to improve the matching score, the transformation estimate is updated and the registration loop started over gain.

In order efficiently and reliably identify the set of updates, an optimisation procedure is needed. A gradient based optimisation strategy is used for this purpose.

In gradient descent techniques local search for the optimal parameter settings is guided by calculations of the objective function's gradient. It explores the parameter space by making steps in the directions defined by the gradients. As a result, the objective function does not need to be evaluated at each round, it is sufficient to only calculate the derivative terms. Nevertheless, in order to monitor the convergence behaviour of the algorithm, the similarity measure needs to be computed at each step of the minimisation phase.

The Levenberg–Marquardt (LM) algorithm which is an iterative non linear optimiser is used for this task [146]. This approach was shown to be fast, with an associated reduction of dependence on the initial estimate [146]. The LM method is detailed herein.

The parameters of $T$ are represented by a vector $a$. The parameter vector is specified as,

$$a = [\theta_x, \theta_y, \theta_z, t_x, t_y, t_z]$$  \hspace{1cm} (8.16)

As previously stated alignment is measured by an error function $E$, which can be quoted as,

$$E(a, \varphi) = \sum_{i=1}^{N} \| P(R(m_{\varphi(i)}) + T) - t_i \|$$  \hspace{1cm} (8.17)

This is used in order to measure alignment, correspondence between the projected model and target data points. This correspondence is denoted by the function $\varphi(i)$, which selects, for each 2D target image point, $t_i$, the corresponding model point, $m_i$. 
In general, the function $\varphi$ is considered part of the minimisation process. It is an essential part of the similarity measure. The estimate of the optimal registration is given by minimising over $a$,

$$\hat{a} = \min_a \sum_{i=1}^{N} \min_j \| P(R(m_j) + T) - t_i \|$$ (8.18)

The optimisation procedure proposed in this chapter is to directly minimise the model-data fitting error via non-linear minimisation. The LM algorithm is an optimisation procedure that is particularly suited to functions such as $E$, which are expressed as a sum of squared residuals.

The main component in the application of LM to the registration problem is how the derivatives of $E$ may be computed. Indeed the requirement for first derivatives might be seen as immediately disqualifying $E$ from LM optimisation, given the discrete minimisation over $j$ within the summation. However, it was shown by [146] that in fact the derivatives of $E$ may be easily and efficiently obtained, and that they are, or can be made, smooth. It can be computed via finite differencing [146], at a cost of $p$ extra function evaluations per inner loop. This means that each iteration's cost is increased by a factor of $1+p$, for the simplest form of LM. However, in typical cases, where LM requires fewer iterations to achieve a certain accuracy, this factor is an upper bound. In many cases, the reduction in number of iterations will exceed the increase in per-iteration cost. Since the derivatives are computed via finite differences, calculations of the form $(E(a+\delta) - E(a))/||\delta||$ are required. It is important that the closest point computations are repeated when making these calculations. It might be thought more efficient to compute closest points once, when calculating $E(a)$, and to use the same values of $\varphi$ to compute $E(a+\delta)$. However, this will compute the derivatives with fixed correspondences. It is better, particularly with a robust kernel, to allow the correspondences to change for each $\delta$.

Now follows a derivation of the LM algorithm. The error function $E(a)$ can be written as the sum of $N_d$ residuals as follows.
Chapter 8 - 2D-3D Registration: Axial Alignment

\[ E(a) = \sum_{i=1}^{N} E_i^2(a) \]  \hspace{1cm} (8.19)

where the residual for the \( i^{th} \) data point is given by

\[ E_i(a) = \min_{j} \| P(R(m_j) + T) - t_i \| \]  \hspace{1cm} (8.20)

An important concept in the derivation of LM will be the vector of residuals

\[ e(a) = \{ E_i(a) \}_{i=1}^{N} \]  \hspace{1cm} (8.21)

in terms of which the error function becomes,

\[ E(a) = \| e(a) \| \]  \hspace{1cm} (8.22)

The LM algorithm combines the gradient descent and Gauss–Newton approaches to function minimisation. Using the above notation, the goal at each iteration is to choose an update to the current estimate \( a_k \), say \( x \), so that setting \( a_{k+1} = a_k + x \) reduces the error \( E(a) \).

Expanding \( E(a+x) \) around \( a \),

\[ E(a + x) = E(a) + (\nabla E(a) \cdot x) + \frac{1}{2!} \left( (\nabla^2 E(a) \cdot x) \cdot x \right) + \text{h.o.t} \]  \hspace{1cm} (8.23)

Expressing this in terms of \( e \),

\[ E(a) = e^T e \]  \hspace{1cm} (8.24)

\[ \nabla E(a) = 2(\nabla e)^T e \]  \hspace{1cm} (8.25)

\[ \nabla^2 E(a) = 2(\nabla^2 e) e + 2(\nabla e)^T e \]  \hspace{1cm} (8.26)

We shall denote the \( N_a \times p \) Jacobian matrix \( \nabla e \) by \( J \), with \( ij^{th} \) entry \( J_{ij} = \partial E_i / \partial a_j \).

Introducing the Gauss–Newton approximation \[146], i.e. \( (\nabla^2 e) e \approx 0 \),

\[ E(a + x) \approx e^T e + 2x^T J^T e + x^T J^T Jx \]  \hspace{1cm} (8.27)

The task at each iteration is to determine a step \( x \), which will minimise \( E(a + x) \). Using the approximation to \( E \) that we have just derived, we differentiate with respect to \( x \) and equate with zero, yielding

\[ \nabla_x E(a + x) = 2J^T e + 2J^T Jx = 0 \]  \hspace{1cm} (8.28)
Solving this equation for $x$ yields the Gauss–Newton update, and gives the algorithm for one iteration of Gauss–Newton,

1) Compute the vector of residuals $e(a_k)$, and its $N_d \times p$ matrix of derivatives $J$ with respect to the components of $a$.

2) Compute the update $x = -\lambda^{-1}J^Te$

3) Set $a_{k+1} = a_k + x$.

Of course, the above strategy does not guarantee that the step taken will result in a reduced error at $E(a_k + 1)$. Whether or not it does so depends on the accuracy of the second-order Taylor series expansion at $a_k$, and on the validity of the Gauss–Newton approximation. However, it can be shown that when these approximations are good, as they tend to be when near the minimum, convergence is rapid and reliable.

By comparison, an accelerated gradient descent approach, used by some previous registration algorithms [146], is obtained by replacing Step 2 with computing the update $x = -\lambda^{-1}J^Te$ where the value of $\lambda$ controls the distance travelled along the gradient direction. For small $\lambda$, the iteration moves a long way along the downhill direction; large $\lambda$ implies a short step. In contrast to Gauss–Newton, gradient descent does guarantee to reduce $E$, providing $\lambda$ is sufficiently large. However, its convergence near the optimum is dismally slow.

The LM algorithm combines both updates in a relatively simple way in order to achieve good performance in all regions. Step 2 is replaced by, $x = -(J^TJ + \lambda I)^{-1}J^Te$. Now large $\lambda$ corresponds to small, safe, gradient descent steps, while small $\lambda$ allows fast convergence near the minimum. The art of a good LM implementation is in tuning $\lambda$ after each iteration to ensure rapid progress even where the Gauss–Newton approximations are poor. Details of such strategies may be found in [146].
8.7 In-Plane and Out-of-Plane Registration

This scheme only applies when conducting full six-degrees-of-freedom registration. Conducting axial registration alone will not need to adhere to this scheme.

In contrast to the world coordinate system, it is known for the camera coordinate system that translations and rotations along the x-axis and the y-axis are in-plane transformations and that the translation around the z-axis is an in-plane translation. The other transformation (rotation about the z-axis) is an out-of-plane transformation. As in-plane transformations cause the crucial movements in the projected image and out-of-plane transformations are less sensitive on the projected image, it is useful to distinguish between them. The optimisation strategy employed in this thesis divides the six-dimensional registration problem into a dimensional in-plane registration and an out-of-plane registration. In a first step, an in-plane registration is executed, in which only the in-plane parameters are estimated. All six parameters were initialised with zeros prior to the registration. These parameters are easier to estimate as they cause major changes in the projected image. In the meanwhile, the out-of-plane parameter remains unchanged. In the second step, the out-of-plane registration is executed, in which only the out-of-plane parameter is optimised. The starting point for the out-of-plane registration is the parameter vector that is returned by the in-plane registration. As the out-of-plane transformations cause only small effects in the projected image, it is difficult to obtain a stable estimation for them. Nevertheless, their estimation can significantly be improved if the in-plane transformations are already properly determined, as will be demonstrated in the experimentation section.

8.8 Multiresolution Approach

The optimisation is conducted in a hierarchical approach to achieve final 2D-3D registration. The idea behind this formulation stems from an approach originally offered in the field of image compression. Essentially, at the lower levels of the hierarchy the aim is to
eliminate superfluous information encoded in the dataset and attempt to represent it in a more compact manner. Since the introduction of this strategy, it has been widely used in the computer vision community. In medical imaging applications, for instance, excellent results have been presented in multi-modal 3D-3D head registration applications [147, 148].

The main motivation behind running the experiments on various levels of resolution was to increase the speed and the robustness of the alignment procedure. Even if timeliness is not paramount, this hierarchical approach is necessary to ensure robustness. The datasets used during the experimentation are generally at a higher than that of an acquisition for medical/treatment purposes would be. For example, one dataset indicated in the testing results had 390 slices and a slice thickness of 0.5 mm which is 2 or 3 times more than it would have been if requested for ordinary diagnostic purposes.

Handling such large datasets efficiently is a challenging task, especially when projection images are needed. Hence, given the 3D volumetric data was down sampled and smoothed (with a Gaussian kernel) to obtain versions of the original with lower resolution. Due to the high accuracy of our initial 3D model datasets, three to four levels of hierarchy was needed.

8.9 Proof of Concept Experimentation

Another fundamental proof of concept experiment was conducted (similar to the one done for the frontal and lateral alignment) to validate the proposed optimisation approach. The study involved a synthetically broken CT data set from which pairs of fluoroscopic images were acquired in the frontal and lateral planes. This acquisition was repeated for different poses of the bone data. Through these synthetic CT volume movements the authors were able to obtain highly realistic fluoroscopic images that were analogous to those obtained clinically on human subjects (see Figure 8.9). These acquired fluoroscopic images were next utilised in the testing phase where the 2D-3D registration was performed with a pre-built 3D surface model of the fracture region. This 3D model was extracted from the CT data set and the pose
initialisation was performed manually. Table 8.2 indicates the absolute average difference between the actual and identified rotation/translations through the 2D-3D registration process. DRRs which simulate the fluoroscopic images were created from the CT dataset utilising a ray-casting algorithm. This technique initially constructed rays between points on the imaging plane and the imaging source. The individual intensity values of the DRR images were computed by summing up the attenuation coefficients associated with each voxel along a particular ray in the CT data. The results (given in Table 8.2) convey that the registration was performed at sub-degree and sub-mm accuracy and with highly robust repeatability. This is in line with the accuracy requirements for image guidance required during present orthopaedic femur fracture reduction procedures.

Table 8.2 Optimisation based 2D-3D registration results.

<table>
<thead>
<tr>
<th>Image Set ID</th>
<th>Rotational Registration – Average Error (Degrees)</th>
<th>Translational Registration – Average Error (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.965</td>
<td>1.056</td>
</tr>
<tr>
<td>2</td>
<td>0.710</td>
<td>0.981</td>
</tr>
<tr>
<td>3</td>
<td>0.862</td>
<td>1.198</td>
</tr>
<tr>
<td>4</td>
<td>0.997</td>
<td>0.912</td>
</tr>
<tr>
<td>5</td>
<td>1.069</td>
<td>0.896</td>
</tr>
<tr>
<td>6</td>
<td>1.159</td>
<td>1.287</td>
</tr>
</tbody>
</table>

Figure 8.9 Qualitative results of the pose estimation performed on the series of phantom fracture reduction fluoroscopic images. Each set of paired images represents a pose that was estimated through the 2D-3D Registration algorithm. The left image of each set is the frontal fluoroscopic image generated while the right image is the registered 3D bone fragments overlaid on the fluoroscopic images.
8.10 Experimentation

This section documents the system validation experimentation performed. The CT data and the fluoroscopic imaging required for the 2D-3D registration testing was provided by the Institute for Robotics and Process Control at the Technical University of Braunschweig, under the robotically assisted long bone fracture reduction project [5].

CT datasets of three fractured femora were utilised in the testing. Testing was conducted to identify the impact of the number and projections of x-ray images, on the accuracy, and robustness of the 2D-3D registration method.

In the first test DRRs generated from the CT data was utilised and registration attempted for a wide range of starting positions and orientations around the “gold standard” registration position. For proximal femur, the ranges of rotational angles were $-15^\circ$ to $+15^\circ$ for $\theta_x$, $-15^\circ$ to $70^\circ$ for $\theta_y$, and $-10^\circ$ to $+10^\circ$ for $\theta_z$, which were observed in medical practice. For distal femur, the ranges of the rotational angles were $-20^\circ$ to $+20^\circ$ for $\theta_x$, $-30^\circ$ to $30^\circ$ for $\theta_y$, and $-10^\circ$ to $+10^\circ$ for $\theta_z$ as recommended by our medical collaborators. The angle resolution used for the testing was $2^\circ$. The algorithm successfully estimated the rotation translational parameters of all the test cases. Table 8.3 shows the mean and standard deviation of the absolute errors of the estimated values. The average angular error is about $2.03^\circ$. For proximal femur, the error in estimating $\theta_y$ is smaller than those of $\theta_x$ and $\theta_z$, because rotation of the proximal femur about the shaft axis produces a larger difference in the projected model contour than rotations about the x and z-axes. So, the algorithm is more sensitive in estimating $\theta_y$. For a distal femur, the error in $\theta_x$ and $\theta_y$ were similar due to the free movement of the bone. The out-of-pane rotation, $\theta_z$, was by far the biggest source of error.
When conducting the second set of experiments the optimisation parameter initialisation technique was modified by separating out the in-plane and out-of-plane registrations. The translations and rotations along the x-axis and the y-axis are in-plane transformations and the translation around the z-axis is an in-plane translation. Rotation about the z-axis is defined as an out-of-plane transformation.

In-plane transformations cause the crucial movements in the projected image and out-of-plane transformations are less sensitive on the projected image, hence it is useful to distinguish between them. The second optimisation strategy employed divided the six-dimensional registration problem into a three-dimensional in-plane registration and a three-dimensional out-of-plane registration. In a first step, an in-plane registration is executed in which only the three in-plane parameters are estimated. All six parameters were initialised with zeros before. These parameters are easier to estimate as they cause major changes in the projected image. Meanwhile, the three out-of-plane parameters remain unchanged. In the second step, the out-of-plane registration is executed in which only the three out-of-plane parameters are determined. The starting point for the out-of-plane registration is the parameter vector that is returned by the in-plane registration. As the out-of-plane transformations cause only small effects in the projected image, it is difficult to obtain a stable estimation for them. Nevertheless, their estimation can significantly be improved if the in-plane transformations are already properly determined, seen with the results below (Table 8.4). Based on our consultation with our medical collaborators and visual inspection of the registration results, we found that registration can be considered as successful with a
registration error of less than 2-3 degrees and translation error of 2-3 mm. Based on this criterion, the testing conducted had a perfect success rate.

Table 8.4 Errors between the recovered and actual rotations and translations with modification to the pose initialisation. Rotational errors provided in degrees while the translational errors are provided in mm terms.

<table>
<thead>
<tr>
<th>Error</th>
<th>$\theta_x$</th>
<th>$\theta_y$</th>
<th>$\theta_z$</th>
<th>$t_x$</th>
<th>$t_y$</th>
<th>$t_z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data 1: Mean</td>
<td>1.73</td>
<td>1.56</td>
<td>2.08</td>
<td>1.48</td>
<td>2.03</td>
<td>1.83</td>
</tr>
<tr>
<td>Data 1: Std Dev</td>
<td>1.44</td>
<td>1.75</td>
<td>1.90</td>
<td>1.72</td>
<td>1.42</td>
<td>1.68</td>
</tr>
<tr>
<td>Data 2: Mean</td>
<td>2.02</td>
<td>1.50</td>
<td>2.11</td>
<td>1.46</td>
<td>1.25</td>
<td>1.83</td>
</tr>
<tr>
<td>Data 2: Std Dev</td>
<td>1.04</td>
<td>1.60</td>
<td>1.87</td>
<td>1.97</td>
<td>1.58</td>
<td>1.69</td>
</tr>
<tr>
<td>Data 3: Mean</td>
<td>1.56</td>
<td>1.79</td>
<td>1.88</td>
<td>1.52</td>
<td>1.90</td>
<td>1.72</td>
</tr>
<tr>
<td>Data 3: Std Dev</td>
<td>1.45</td>
<td>1.98</td>
<td>1.81</td>
<td>1.85</td>
<td>1.65</td>
<td>2.35</td>
</tr>
</tbody>
</table>

There are three main sources of inaccuracies in the 2D-3D registration process. Firstly, image calibration can cause errors. Incorrect extrinsic calibration between image and device to be registered is one source. Intrinsic calibration errors can further lead to a lack of correspondences between the 2D and 3D data to be registered.

Secondly, image segmentation errors cause inaccuracies with the registration process since they negatively impact on the correspondence identification process between the contours to be registered. Incorrect segmentation eliminates bony contours and introduces noise and spurious points. One way to establish and quantify the errors is to perform a large number of experiments on various datasets, once a segmentation algorithm is established. This is however outside the scope of this chapter.

The registration algorithm presents another source of error. The registration algorithm establishes point correspondences between the projected and extracted contours. These correspondences could be incorrect. The use of a distance based measure implies that the closest points on the contours are the best match. However, this might not necessarily be the case. This correspondence identification can be improved in future works. Moreover, the registration would not be successful in the case where the pose initialisation causes the surfaces to be too distant from the correct position for the registration algorithm.
8.11 Chapter Summary

This chapter details an algorithm for recovering the position and orientation of the target anatomy in 3D space based on an iterative comparison of 2D planar radiographs with the pre-operative 3D model, through an optimisation process. The process uses x-ray fluoroscopic images acquired intra-operatively, and iteratively compares them with projection image generated from the 3D model. The chapter provided details on the objective functions and the optimisation procedures used in the 2D-3D registration.
Chapter 9 Conclusions and Future Works

The main motivation behind the research presented in this thesis is due to the fact that at present predominantly only 2D image guidance is available clinically (for the cases of interest), which lacks significant spatial information that is present in 3D modalities. Imaging modalities such as CT have the ability to provide direct 3D volumetric images. However the use of such imaging is restricted to a minority of complex procedures due to constraints placed by cost, availability and risks posed by unwarranted detailed imaging. Thus an alternative to direct 3D imaging needed to be developed to augment procedures that currently rely on pure 2D radiographs. The work presented throughout this thesis is motivated by this requirement.

This work proposes a 3D reconstruction technique utilising commonly available 2D radiographic imaging. As previously stated, orthopaedic IGS can be broken down into two main segments depending on the clinical use: pre-operative planning and intra-operative tracking. There are two main research contributions in this thesis. The first is involved with the pre-operative reconstruction framework presented in Chapter 3 and Chapter 4. The second is the intra-operative pose estimation algorithm presented in Chapter 6, Chapter 7 and Chapter 8.

9.1 Pre-Operative Planning

An intact bone reconstruction framework aimed at osteotomy and arthroplasty is proposed in Chapter 3. The framework provides a static pre-operative 3D model for diagnostics and planning. The pre-operative 2D x-ray images are initially processed to extract the edge points that potentially form the bony boundary. A non-rigid registration is then performed between the edges identified in the x-ray image and the projected contour points of the generic model. The identified point correspondences are next interpolated to create a 2D planar translational
field in both the anterior and lateral viewpoints. This translational field identifies the deformations required by the 3D anatomical model in the equivalent viewpoint. Finally a full 3D translational field is created through interpolation and the 3D generic anatomical data deformed accordingly.

Chapter 4 details the femur fracture pre-operative planning framework, which is the case of interest for this thesis. Currently available pre-operative planning systems for fracture reduction are all restrictive (time and cost) as they require a CT scan of the injured bone. This is not acceptable by our clinical partners. They have claimed that the increased radiation exposure and extra costs imposed by CT scanning is not warranted for femur fracture surgery. As previously mentioned CT scanning is conserved for complex orthopaedic interventions (e.g. pedicle screw insertion and intra-articular fractures). Consequently in many situations the surgeons will be limited to using 2D x-ray and fluoroscopic images. In order to ensure clinically applicability the research presented in this thesis focuses on 2D image based reconstruction techniques currently presented in intact bone reconstruction literature. Thus a novel fractured bone reconstruction technique was developed to be able to reconstruct 3D fractured bone models pre-operatively. All patient-specific fracture details required for the framework are acquired from 2D x-ray images. Firstly, these pre-operative 2D images are processed to segment the bone contours. Next, fracture modelling is performed to identify the fracture surface detail and to aid the separation of the generic intact bone. This chapter proposes an interactive feature-based modelling technique. Pose estimations and shape customisations are iteratively performed until the user is satisfied with the 3D visualisation presented. The pose estimation separates the bone into two fractured segments and identifies their respective poses. The shape customisation involves a non-rigid registration between the edges identified in the front and lateral x-ray image and the projected contour points of the generic model.
Chapter 9 - Conclusions and Future Works

Chapter 5 proposes a model based segmentation technique which is a by-product of the non-rigid registration used in the novel reconstruction technique proposed in the earlier chapters.

Both intact and fractured bone reconstruction algorithms presented several contributions over the current state-of-the-art literature. The two frameworks are individually original, however several end results are superior to current literature, as listed below.

9.1.1 Reconstruction Accuracy

The reconstruction accuracy of the proposed algorithm, shown through experimentation is far greater than comparable studies seen in literature. For the proposed intact bone reconstruction algorithm the experiments identified an average absolute Euclidean error of 1.15mm. The result is a significant improvement on other comparable literature work. Research conducted by [19] convey an average mean surface reconstruction error of 2.4mm (this work targeted the proximal region of the femur). Work performed by [70] on distal femur reconstruction indicate a mean error of 3.27mm. Furthermore, when the proposed reconstruction technique was supplemented with the inner contour driven algorithm, the average error dropped to a sub-millimetre level.

Fractured bone reconstruction experiments conducted conveyed an average absolute Euclidean error of 1.86mm. The errors presented are attributed to three main causes. The utilisation of orthogonal 2D radiographic images for 3D model reconstruction is ill-posed mathematically. Concave patient-specific cross-sectional variations cannot be integrated with the use of only generic models. Bone segmentation and the defining of bone contours is also a difficult task, and add to the reconstruction error. Pose estimation errors are also common due to the difference in the centre of rotation of each fragment and the actual clinical fracture break mechanisms.

The segmentation results indicated a root-mean-squared distance error (between the automatically segmented femur contour and the manual segmented ground truth) of 2.5 mm
with a standard deviation was 0.80 mm. This is well within the femoral bone segmentation errors typically noted in literature [78].

9.1.2 Adaptability to Other Cases

The proposed technique is highly adaptable to other orthopaedic cases as shown in Chapter 3 and Chapter 4. This generally applicable intact bone reconstruction algorithm presented is a significant improvement over the current techniques seen in literature including statistical model based techniques [19, 70-76], free form deformation based techniques [20, 21], and morphological measurement based techniques [77]. Statistical model based reconstruction techniques provides deformation constraints that can be utilised to ensure global shape conformance. However the PDM creation, which is required for the statistical model, involves analysing a large database of anatomical models. The dataset also requires a time consuming corresponding landmark identification process. Thus typically only a small specific portion of the anatomy is modelled. This user intensive analysis means that models built through statistical techniques are not easily expandable to other orthopaedic procedures. Free form deformation based techniques can be utilised with sparse deformation information, however they provide no shape deformation constraints and are thus prone to unrealistically deformed shapes. They also require an initial setup of the control points and a bounding box. Morphological measurement based techniques provide deformation constraints that can be utilised to ensure localised shape conformance. However they do not provide the shape constraints and local shape warping capabilities.

9.1.3 Manual Intervention Required

Similar to the aforementioned contribution, the level of manual intervention required with the proposed technique is minimal compared to other algorithms seen in literature. For instance the statistical model based techniques all require a highly time consuming corresponding landmark identification process in a large database of anatomical models [19, 70-76]. Furthermore the free form deformation based techniques require an initial setup of the
control points and a bounding box [20, 21]. Morphological measurement based techniques require the manual measurement between anatomically significant landmarks, which requires expert knowledge and is time consuming [77].

9.2 Intra-Operative Pose Estimation

The second contribution of this thesis is the intra-operative pose estimation algorithm. The second half of this thesis (Chapter 6, Chapter 7 and Chapter 8) introduces a novel framework to address the 2D-3D rigid-body registration problem between a pre-operative acquired 3D model and intra-operative x-ray fluoroscopy images. The main objective was to identify a geometrical transformation in 3D space that describes the relative position and orientation of coordinate systems of the imaging environment and the pre-operative data volume. This intra-operative guidance work is mainly aimed at fracture reduction surgery and internal fixture insertion.

The proposed 2D-3D registration algorithm can be split into two distinct phases: 1) frontal and lateral pose estimation (Chapter 7), and 2) axial pose estimation (Chapter 8). Since the registration is performed through images acquired from the frontal and lateral viewpoints, registration in these planes will be performed highly accurately through the first phase of the 2D-3D registration process. The axial alignment is critical, as it has a high impact on the functional biomechanics of the leg, and will be conducted in the second phase.

The frontal and lateral alignment is conducted analytically through extracted features from the fluoroscopic images. The task involved with this first phase of the registration is to utilise the extracted features from the target fluoroscopic images and the projected contours of the model, and to align the rotational axis of each bone fragment. Thus the model’s mid shaft axis must be aligned in each plane with the extracted mid shaft axis of the target image. Experimentation was done to test the robustness and repeatability of the developed algorithm.
Chapter 9 - Conclusions and Future Works

The axial alignment procedure utilises a distance based similarity measure to evaluate the alignment quality of the input datasets given a transformation estimate. The standard optimisation method and the gradient ascent strategies guide the estimate-update procedure by computing partial derivatives of the objective function with respect to the transformation components. In order to reduce the running time of these methods, they both operate on only a sparsely sampled version of the original inputs and approximate the true update terms. The experiments were run in a multi-resolution setting, which not only proved to decrease execution time, but also increased the robustness of the algorithms. The author examined the registration characteristics of these algorithms by both running them on simulated and real datasets.

Both algorithms presented in Chapter 7 and Chapter 8 have significant contributions over the current state-of-the-art literature. A few of the novelties are listed below.

9.2.1 Feature Space

The main contribution to the state of the art is demonstrating a feature based registration technique that is augmented by an optimisation based final registration. Our clinical collaborators have acknowledged the usefulness of the initial feature based alignment in the frontal and lateral planes as this mimics the process steps taken during surgery. The features of interest are extracted in real time and are extremely robust. These novel features are utilised in the analytical pose estimation as well in the optimisation based routine. The search space utilised in the optimisation routine has been streamlined to be able to be near real time and provide timely results.

This feature based technique is in contrast to the intensity based techniques typically seen in literature. Here the 2D-3D registration is based solely on pixel or voxel intensity and spatial information extracted from images [84-89]. This type of pose estimation requires three highly computational steps: 1) Generation of digitally reconstructed radiographs (DRRs) for each pose; 2) Measurement of the pose difference by comparing the DRRs with the real
fluoroscopic x-ray images; 3) Computation of a pose that minimises that difference. The calculation of a DRR by numerical summation of CT image intensities along projection rays involves high computation cost and is thus time consuming. Usually more than a hundred DRRs have to be generated during the search to achieve reliable and accurate registration. The key characteristic of intensity-based registration is that it does not require segmentation. The rationale is that using as much information as available and “averaging it out” reduces the influence of outliers and is, thus, more robust. However, this approach is computationally expensive since it requires generating high-quality DRRs and searching a six degree-of-freedom space with local minima, which depends on the similarity measure employed.

9.2.2 Pose Estimation Accuracy

The proposed pose-estimation algorithms are also highly accurate compared to similar literature works. For the frontal and lateral alignment, the proof of concepts testing conducted with a fragment of an artificially fractured plastic bone and a USB camera mounted on a rotary arm (similar to a fluoroscopic c-arm) identified average absolute mean $R_x$, $R_y$ and $R_z$ errors of 0.567, 0.487 and 0.604 degrees. The second set of clinically applicable testing conducted with CT data and the fluoroscopic imaging conveyed an absolute average error of 2.48mm with a standard deviation of 0.71 mm. These results are clinically acceptable and are an improvement over several other literature studies that have engaged in similar testing [140, 142].

For the optimisation based alignment, two types of experiments were conducted, one a controlled experiment, which used images of an artificially broken phantom femur, secondly, clinically significant experiments, where cadaveric CT and fluoroscopic images were utilised. For the experiments using real radiographic images, though somewhat preliminary, showed promising results. Consistent alignment of the input images according to the ground truth estimates was achieved. For the optimisation based testing conducted through CT data the average angular error was about 2.03 degrees with a translational error of 1.79mm. Based on our consultation with our medical collaborators and visual inspection of the registration
results, we found that registration can be considered as successful with a registration error of less than 2-3 degrees and translation error of 2-3 mm. Based on this criterion, the testing conducted had a perfect success rate. Two main sources of error exist with 2D-3D pose estimation. Image calibration can cause errors due to incorrect extrinsic calibration between image and device. Intrinsic calibration errors can further lead to a lack of correspondences between the 2D and 3D data to be registered. Image segmentation errors cause inaccuracies with the registration process since they negatively impact on the correspondence identification process between the contours to be registered.

9.3 Future Research Ideas

This final section of the thesis describes several ideas that emerged while working on the pre-operative reconstruction and intra-operative pose estimation projects. These are thoughts or techniques that could either improve or complement the performance of the existing system.

9.3.1 Coupling Segmentation and Registration

Segmentation, along with registration, is another major area within medical image processing. Its goal is to identify and group together image components representing the same region or having similar properties based upon some specific criteria. Traditionally, registration and segmentation methods have been investigated in relative isolation, but many times the assumption is made that the solution to one is known in order to tackle the other. That statement clearly applies to model based segmentation algorithms or feature based registration algorithms as discussed in Chapter 5 and Chapter 7 respectively.

In the case of feature based registration, the alignment procedure considers landmark points and identifies those depending on either some user interaction or some form of a segmentation technique.
Chapter 9 - Conclusions and Future Works

Segmentation could benefit from registration when multiple acquisitions are processed at the same time. The relative correspondence between the targeted images contains useful information for completing the task. Executing the two procedures simultaneously is motivated by the hope that overall computation time could be saved by not having to complete one before the other, and that the parallel registration and segmentation procedures can provide sufficient information to each other in order to improve their individual performances.

This thesis introduced the idea of coupling registration and segmentation in Chapter 5. This was from a segmentation point of view. However from a registration view point it is desirable to incorporate such a framework to improve or extend the performance of methods in more complex rigid-body or non-rigid registration problems. For example, the problem of registering spine datasets has long been a research challenge. Although the individual vertebrae behave as rigid bodies individually, the global movement of the spine cannot be described as that of a single rigid entity. A general approach to the problem is to piece-wise align single vertebrae and then to describe the spine movement as some combination of these. That scenario could be closely related to the study by [19, 134, 138, 143], which reduced the computational burden of its 2D-3D registration algorithm by only considering a small portion of the original CT dataset. The imaged anatomy (in their case, a single vertebra) was segmented out in the CT volume prior to any other processing. It was only then, that the alignment step took place with only the sub-volume considered.

With the new approach the segmentation and registration stages would not have to be isolated, but they could operate dynamically in parallel. Even the hierarchical and multi-resolution approaches could be incorporated in that framework. This could allow the algorithm to focus more and more on the targeted (parts of the) anatomies as the resolution level is increasing. That application could be useful in addressing the 2D-3D registration problem as reducing the size of the image data volume, could improve computation time significantly.
Chapter 9 - Conclusions and Future Works

9.3.2 View of Fluoroscopic Acquisitions

Another interesting question addresses the problem of how important the view is from which the fluoroscopic images are taken. It is an open question whether the registration results could be improved at all if these images were taken from a specific view and a given angle apart. Initial (synthesised) probing studies conducted by [95] demonstrated that, in case of the skull, the location of the projection source has a very insignificant impact on the performance of the objective function. In case of the femur (or pelvis region), however, these results might be different as it has a much more irregular geometry.

We would also like to investigate the related problem of the ideal imaging orientations of the images, that to accurately carry out the registration procedure. It would be very useful to know how much the robustness and the capture range of the algorithm would grow with various view modifications. This is not to identify the number of images, as this has been already documented in the thesis.

9.3.3 Defining Automatic Stopping Criterion for Gradient Optimisation Protocols

An algorithm that would automatically identify convergence to the optimal settings would be of great value if incorporated into the current registration framework. Presently the optimisation algorithm is executed, a predetermined number of times. This is however quite inefficient in some cases if alignment can be achieved earlier.

The complexity of this task originates from the fact that in case of our sampled stochastic approaches, the objective function evaluations are noisy estimates of the true values. Hence, any kind of stopping criterion would have to anticipate a large enough variance in the similarity values even in case of convergence.

9.3.4 Truncation/Limited Field of View

Further work and experimentation might be needed to study the robustness of our algorithms with respect to errors due to truncation and limited field of view. In case of the
presented distance based similarity measure, it has been known to be a weakness. However further work needs to be done to fully understand how much that problem surfaces in 2D-3D applications. This is an important question to be analysed as in case of real procedures, intra-procedural images are likely to contain images of surgical/treatment tools that could partially obstruct the anatomy of interest.

9.3.5 Distortion Effects and Dewarping

One of the main disadvantages of using x-ray fluoroscopy imaging is the geometrical distortion effects that are not corrected for at the time of image acquisition. This is particularly true for older c-arm devices that do not use “flat bed” imaging. The calibration technique presented in this thesis does not account for this problem. As the distortion effects can potentially be significant and qualitative evaluation of the algorithm on distorted images is not truly meaningful, therefore the pre-processing step could definitely benefit from this feature.

The c-arm devices clinically used by our medical collaborators do suffer from radial and pin-cushioning distortion. Thus distortion correction is useful from their perspective.

9.3.6 Code Optimisation

As of now, the Matlab codes utilised for the reconstruction and 2D-3D registration tasks are not optimised for clinical use. Ideally the algorithms can be implemented in C++ after eliminating all unnecessary computations and using optimised implementations of the algorithms.

9.3.7 Clinical Trials

Clinical trials are necessary to prove the claim that this application can be used for supporting or even replacing existing patient positioning methods. Ground Truth based registration has to be conducted on various target objects in order to make sure that the
overall precision is sufficient for regular clinical use. Among other things, it has to be studied if the non-rigidity of the specific target area is always negligible.
References

References

References

References

References

References

199
References


