Joint Multimodal Registration of Medical Images to a Statistical Model of the Lumbar Spine for Spine Anesthesia

by

Delaram Behnami

B.A.Sc., The University of British Columbia, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Biomedical Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2016

© Delaram Behnami 2016
Abstract

Facet joint injections and epidural needle insertions are widely used for spine anesthesia. Needle guidance is usually performed by fluoroscopy or palpation, resulting in radiation exposure and multiple needle re-insertions. Several ultrasound (US)-based guidance approaches have been proposed to eliminate such issues. However, they have not widely accepted in clinics due to difficulties in interpretation of the complex spinal anatomy in US, which leads to clinicians’ lack of confidence in relying only on information derived from US for needle guidance. In this thesis, a model-based multi-modal joint registration framework is introduced, where a statistical model of the lumbar spine is concurrently registered to intraprocedure US and easy-to-interpret preprocedure images. The goal is to take advantage of the complementary features visible in US and preprocedure images, namely Computed Topography (CT) and Magnetic Resonance (MR) scans. Two versions of a lumbar spine statistical model are employed: a shape+pose model and a shape+pose+scale model. The underlying assumption is that the shape and size of the spine of a given subject are common amongst all imaging modalities. However, the pose of the spine changes from one modality to another, as the patient’s position is different at different image acquisitions. The proposed method has been successfully validated on two datasets: (i) 10 pairs of US and CT scans and (ii) nine US and MR images of the lumbar spine. Using the shape+pose+scale model on the US+CT dataset, mean surface distance error of 2.42 mm for CT and mean Target Registration Error (TRE) of 3.14 mm for US were achieved. As for the US+MR dataset, TRE of 2.62 mm and 4.20 mm for the MR and US images, respectively. Both models models were equally accurate on the US+CT dataset. For US+MR, the shape+pose+scale model outperformed the shape+pose model. The joint registration allows augmentation of important anatomical landmarks in both intraprocedure US and preprocedure domains. Furthermore, observing the patient-specific model in preprocedure domains allows the clinicians to assess the local registration accuracy qualitatively. This can increase their confidence in using the US model for deriving needle guidance decisions.
Preface

This thesis is based on a manuscript, resulting from the collaboration between multiple researchers. This publication has been modified to make the thesis coherent. The research conducted in this study was undertaken under the approval of the UBC Research Ethics Board, certificate numbers H13-01968 and H16-00341 and Queen’s University Research Ethics Board, certificate number SCOMP-003-07.

A primary version of the framework presented in Chapter 2 as well as the study on the CT+US image pairs in Chapter 3 has been published in:


The author’s contribution was in developing and implementing the presented joint and weighted joint frameworks, constructing the shape+pose+scale model, as well as evaluating the proposed solution on two datasets of US+CT and US+MR clinical images. The original shape+pose model of the lumbar spine was previously developed by Dr. Abtin Rasoulian. Professors Purang Abolmaesumi and Robert Rohling helped with their valuable suggestions for extending the methodology.
Table of Contents

Abstract ................................................................. ii
Preface ................................................................. iii
Table of Contents ....................................................... iv
List of Tables .......................................................... vi
List of Figures .......................................................... vii

1 Introduction and Background ....................................... 1
  1.1 Clinical Background ............................................. 1
    1.1.1 Lumbar Spine Procedures ................................ 1
    1.1.2 Ultrasound Guidance and Its Challenges .............. 3
    1.1.3 Spinal US Augmentation ................................. 3
  1.2 Thesis Objective ............................................... 8
  1.3 Thesis Outline ................................................ 8

2 Model-based Joint Registration Framework ....................... 10
  2.1 Statistical Multi-vertebrae Models of the Lumbar Spine .... 10
    2.1.1 Model Construction ..................................... 10
    2.1.2 Shape+pose Model ...................................... 11
    2.1.3 Shape+pose+scale Model ............................... 11
  2.2 Point-set Registration of Statistical Models to Targets .... 13
    2.2.1 Registration of the Statistical Models to a Single Imaging Modality ........................................ 14
    2.2.2 Joint Registration of the Statistical Models to Multiple Imaging Modalities ................................. 15
    2.2.3 Parameter Optimization for Model-to-Target Registration .................................................... 16
# Table of Contents

3 Joint Registration of Model to US+CT Images of the Lumbar Spine .......................................................... 18
   3.1 US and CT Datasets ....................................................... 18
   3.2 Joint Registration for US+CT Images ................................... 18
      3.2.1 Target Point Cloud Generation ................................. 19
      3.2.2 Geometric Initialization of the Model Instances on Targets ................................................................. 19
      3.2.3 Optimization ........................................................... 20
      3.2.4 Model Transformation ............................................... 20
   3.3 Results ........................................................................... 22
      3.3.1 Registration Parameter Selection .................................... 22
      3.3.2 Validation ................................................................. 22
      3.3.3 US+CT Joint Registration Results and Analysis ............... 23
   3.4 Discussion and Summary ................................................ 33

4 Joint Registration of Model to US+MR Images of the Lumbar Spine .......................................................... 36
   4.1 US and MR Datasets ....................................................... 36
   4.2 Joint Registration for US+MR Images ................................... 38
      4.2.1 Target Point Cloud Generation ..................................... 38
      4.2.2 Model Initialization on Targets ..................................... 39
      4.2.3 Weighted Joint Registration ......................................... 39
   4.3 Results ........................................................................... 41
      4.3.1 Registration Parameter Selection .................................... 41
      4.3.2 Validation ................................................................. 41
   4.4 Discussion and Summary ................................................ 43

5 Conclusion ........................................................................ 50
   5.1 Contributions ............................................................... 50
   5.2 Future Work ................................................................. 51

Bibliography ........................................................................ 54
List of Tables

3.1 Breakdown of the fields of view of images in the CT dataset based on the visible lumbar vertebrae. 

3.2 Distribution of US TRE registration errors (mm) achieved using the shape+pose model for the US+CT dataset, per region. No significant improvement was observed using the joint registration.

3.3 US TRE (mm) achieved using the shape+pose model for the US+CT dataset. In each experiment (each column), the shape parameter of the model is optimized from the modalities indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, TP) and the average of all four.

3.4 CT SDE registration errors (mm) achieved using the shape+pose+scale model for the US+CT dataset. In each experiment (each column), the shape and scale parameters of the model are optimized from the modalities indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, VB) and the average of all four.

3.5 US TRE (mm) achieved using the shape+pose+scale model for the US+CT dataset. In each experiment (each column), the shape and scale parameters of the model are optimized from the modalities indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, TP) and the average of all four.

4.1 Breakdown of the fields of view of US images in the US+MR dataset based on the visible lumbar vertebrae.

4.2 Weights assigned to the different regions of the vertebrae.

4.3 Numbers of MR gold standard points used to report TRE values on MR space for the US+MR dataset. The numbers are provided per region for each subject.
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4</td>
<td>Numbers of US gold standard points used to report TRE values for the US+MR dataset. The numbers are provided per region for each subject. The landmarks were selected by an expert sonographer. Zeros indicate the region was not visible to the sonographer.</td>
</tr>
<tr>
<td>4.5</td>
<td>MR TRE (mm) achieved using the shape+pose model for the US+MR dataset. In each experiment (each column), the shape parameter of the model is optimized from the modalities indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, VB) and the average of all four.</td>
</tr>
<tr>
<td>4.6</td>
<td>US TRE (mm) achieved using the shape+pose model for the US+MR dataset. In each experiment (each column), the shape parameter of the model is optimized from the modalities indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, TP) and the average of all four.</td>
</tr>
<tr>
<td>4.7</td>
<td>MR TRE achieved (mm) using the shape+pose+scale model for the US+MR dataset. In each experiment (each column), the shape and scale parameters of the model are optimized using the methods indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, VB) and the average of all four.</td>
</tr>
<tr>
<td>4.8</td>
<td>US TRE achieved (mm) using the shape+pose+scale model for the US+MR dataset. In each experiment (each column), the shape and scale parameters of the model are optimized using the methods indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, TP) and the average of all four.</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Anatomy of human lumbar spine. The full vertebral column consists of 33 vertebrae, including five lumbar vertebrae highlighted in yellow (a). Different parts of an individual lumbar vertebra are shown in a cross-sectional view (b). (From www.biodigital.com and www.compelvisuals.com).  

1.2 Examples of typical hard-to-interpret US images of the lumbar spine. Only partial bone responses are present and different parts of the anatomy are hard to detect.  

1.3 An overview of the state-of-the-art methods for augmenting anatomical information with spinal US. Methods (a), (b) and (c) are proposed by [12, 21, 35, 36], [10, 18, 25, 27], and [5] as well as this thesis, respectively.  

2.1 First two modes of variation of shape ((a) and (b)) and pose ((c) and (d)) for the shape+pose model. Images on the left of each pair indicate $-3\sqrt{\lambda}$, and those on the right correspond with $+3\sqrt{\lambda}$. Values of $\lambda$ are the eigenvalues determined for each eigenvector (PC) at training.  

2.2 First two modes of variation of pose ((a), (b)) and the first mode of variation of the scale (c) for the shape+pose+scale model. This model decouples the scale from the pose. Images on the left of each pair indicate $-3\sqrt{\lambda}$, and those on the right correspond with $+3\sqrt{\lambda}$. The shape variations are similar to the shape+pose model in 2.2. Values of $\lambda$ are the eigenvalues determined for each eigenvector (PC) at training.  

3.1 An overview of the proposed framework for joint registration of US, preoperative images and a statistical model.
List of Figures

3.2 Example snapshots of the registration results using the shape+pose model on CT (a) and US (b) images. The joint method (blue) performs better than the model-to-US technique (red). The yellow annotations on US depict the gold standard segmentation obtained by the sonographer. ................................................................. 24

3.3 Colormaps showing the difference between shape SDE values per surface point computed from US-to-atlas and joint registrations, using segmented reference CT as the gold standard for patients 1 (a) to 10 (j). Blue regions show improvements achieved by the joint registration, where US-to-atlas error is larger than joint registration error. ................................................................. 27

3.4 Box plots showing the break-down of TRE values obtained using the shape+pose model for US images (US+CT dataset) at different regions of the vertebrae (a) and the five different vertebrae (b). No significant improvements are observed. .................. 28

3.5 Example snapshots of the registration results using the shape+pose+scale model on CT (a) and US (b) images. The joint method (blue) performs better than the model-to-US technique (red). The yellow annotations on US depict the gold standard segmentation obtained by the sonographer. ................................................................. 29

3.6 Colormaps showing the difference between shape SDE values per surface point computed from US-to-atlas and joint registrations, using segmented reference CT as the gold standard for patients 1 (a) to 10 (j). Blue regions show improvements achieved by the joint registration, where US-to-atlas error is larger than joint registration error. ................................................................. 30

3.7 Box plots showing the break-down of TRE values obtained using the shape+pose+scale model for US images (US+CT dataset) at different regions of the vertebrae (a) and the five different vertebrae (b). No significant improvements are observed. ................................................................. 31

3.8 Colormaps showing the difference between shape SDE values per surface point computed from US-to-atlas and joint registrations, using segmented reference CT as the gold standard for patients 1 (a) to 10 (j). Blue regions show improvements achieved by the joint registration, where US-to-atlas error is larger than joint registration error. ................................................................. 32
3.9 Snapshots of a CT scan belonging to a scoliotic patient (P8), which was proven challenging to register. The vertebrae seem to diverge from the vertical yellow line (left). The yellow line on the coronal view also marks the sagittal slice on the right. White arrows indicate the failed pose registration of the model. .................................................. 34

4.1 An example MR image from the present dataset. The yellow dashed line on the sagittal image (right) indicates the location of the axial slice (left). The TPs are not captured in the field-of-view of these MR images. The sagittal slices are 3.5−4.4 mm thick in this dataset. ................................. 37

4.2 An example US image from the present challenging dataset, shown in three different views. The yellow and green dashed lines on the sagittal image (middle) indicate the location of the axial (left) and coronal slice (right), respectively. The shadow of the rib attached to T12 is used to locate L1. The blue lines are drawn to help visualize the landmarks. ........ 38

4.3 The lumbar spine in the sitting position. Figure in (a) shows that in the sitting pose, lumbar spine is almost straight but the L5-sacrum curve is sharper (courtesy of ergonut.com [22]). Figures in (b) and (c) demonstrate the straight pose used for registering the shape+pose and shape+pose+scale models to the US images, respectively. ................................. 40

4.4 Example snapshots of the registration results using the shape+pose+scale model on MR (a) and US (b) images. The weighted joint method (blue) performs better than the model-to-US technique (red). The yellow annotations on US depict the gold standard segmentation obtained by the sonographer. .... 44

4.5 Box plots showing the break-down of TRE values obtained using the shape+pose model for US images (US+MR dataset) at different regions of the vertebrae (a) and the five different vertebrae (b). No significant improvements are observed. ... 47

4.6 Example snapshots of the registration results using the shape+pose+scale model on MR (a) and US (b) images. The weighted joint method (blue) performs better than the model-to-US technique (red). The yellow annotations on US depict the gold standard segmentation obtained by the sonographer. .... 48
4.7 Box plots showing the break-down of TRE values obtained using the shape+pose+scale model for US images (US+MR dataset) at different regions of the vertebrae (a) and the five different vertebrae (b). Significant improvements were observed at AP, SP, TP regions, as well as at L1 and L3. . . . . 49
Chapter 1

Introduction and Background

1.1 Clinical Background

1.1.1 Lumbar Spine Procedures

A human spine consists of 33 vertebrae including the cervical, thoracic and lumbar vertebrae as well as the sacrum and coccyx. The lumbar spine is located halfway at the bottom of the spinal column and is consisted of five vertebrae L1-L5 (Fig. 1.1(a)). The bone of each vertebra consists of the spinous process, the laminae, articular processes, and transverse processes (Fig. 1.1(b)).

Facet joint injections and epidural needle insertions are common procedures performed on the lumbar spine, which aim to deliver anesthesia and analgesia [8, 11]. Facet joints are the paired joint connecting the articular processes of two adjacent vertebrae, which allow the vertebrae to extend and flex [8]. The cartilage on facet joints can wear out and cause chronic back pains in some patients, a problem experienced at least once by 80% of the adult population [28]. To treat this condition, facet joint injections are performed on patients suffering from it. In these procedures, a local anesthetic is injected in the affected joints to numb them and hence, relieve the pain. These injections can also include steroids which help reduce the inflammation of the joints [4]. Epidural needle insertions, or simply epidurals, involve inserting specialized needles in the epidural space of the spine, the area between the vertebral walls and dura mater [19]. Every year, approximately 3.2 million people undergo these procedures in North America alone [11]. The purpose of these procedures is to help alleviate chronic pains by blocking the nerve impulses and hence the sensation of pain in the desired regions. Depending on the target of the procedure, epidurals can be performed on different parts of the spine. Lumbar spine epidural anesthesia is also very popular amongst women as 50% of women going in labor choose to receive this treatment to ease the pain of labor.
1.1. Clinical Background

Figure 1.1: Anatomy of human lumbar spine. The full vertebral column consists of 33 vertebrae, including five lumbar vertebrae highlighted in yellow (a). Different parts of an individual lumbar vertebra are shown in a cross-sectional view (b). (From www.biodigital.com and www.compelvisuals.com).

Lumbar spine needle insertion procedures are particularly challenging due to the proximity to nerve tissue, the deep location of the target, the small and narrow size of the channel between the articular processes of the joint, and the oblique entry angle [8]. Hence, careful needle placement is very important to ensure no damage is done to the nerves and nearby soft tissue, and the therapy is effective [37]. In traditional epidural anesthesia, needle insertion is performed with the help of manual palpation and the loss-of-resistance technique [19]. In this method, the clinicians rely on the loss of pressure they feel as they insert the needle, to know that needle has entered the epidural space. This blind and consequently inaccurate technique can lead to longer needle paths, multiple insertions or painful insertions through the muscle [5]. In both facet joint and epidural injections, inaccurate needle placement can cause complications such as accidental dural puncture (2.5% of all cases [1], and 3-5% in procedures performed by inexperienced operators [14]) and post-puncture headache (86%) [31]. Performing these procedures becomes even more challenging in the cases of
obese patients, patients with abnormally curved spines (scoliotic patients), and those with previous back surgery [37].

1.1.2 Ultrasound Guidance and Its Challenges

To prevent the above complications from happening, image-based guidance methods may be beneficial. The current standard of care for guiding facet joint injections is fluoroscopy, which provides real-time contrast images of the anatomy during a procedure [32]. However, this imaging modality imposes health risks as it exposes both the patient and the surgeon to ionizing radiation. Moreover, the specialized equipment for fluoroscopy are expensive and have limited portability [37].

Consequently, ultrasound (US) has been suggested as an alternative modality to guide these procedures. US imaging has many advantages including its non-ionizing technology, real-time 3D acquisition capability, accessibility and low cost [37]. Many studies demonstrate the feasibility of US guidance for spine interventions [13, 20, 23, 34, 37]. However, at present, these methods are not commonly used in clinics because of the challenges in interpreting spinal US. Due to their physical nature, US images in general are often corrupted by significant amounts of noise and artifacts. Interpreting images of bony structures is specially challenging as bone responses can appear blurry and disconnected. Also, as US waves do not penetrate bones, only bone surfaces closest to the surface of the skin can be visible in US images. Finally, the complex anatomy of the spine adds to the difficulties of making sense of US images and the imaged anatomies. Figure 1.2 shows two example slices from a typical sagittal spinal US scan. As it can be seen in these images, only little bone responses are visible and it is not easy to identify which part of the lumbar spine they represent.

1.1.3 Spinal US Augmentation

With the goal to improve interpretation of sonographic images, many works have proposed US augmentation by fusing these images with other forms of data. These fusion techniques can be divided into two categories: (i) fusion with anatomical information from preoperative Computed Tomography (CT) or Magnetic Resonance Imaging (MRI), and (ii) statistical anatomical models.
1.1. Clinical Background

Figure 1.2: Examples of typical hard-to-interpret US images of the lumbar spine. Only partial bone responses are present and different parts of the anatomy are hard to detect.

**US Augmentation by Multi-modal Fusion**

Several methods have focused on registering easy-to-interpret preprocedure CT images, to intraprocedure US [12, 21, 35, 36] (Fig. 1.3(a)). For patients receiving routine injections for back pain, preprocedure CT or MR images are often available as they were acquired for initial diagnosis [30]. Hence, such multi-modal methods can be feasible for US augmentation. Prior works on fusion of preoperative CT with US mostly focus on individual bones [29]. For registration of multiple bones, such as that of multiple vertebrae, Brendal et al. considered the ensemble of bone structures as a single rigid object [9]. Yan et al. attempted to independently register each individual bone [36]. Gill et al. proposed a multi-object registration, which used a biomechanical model for regularization purposes [15]. Bø et al. attempted US augmentation for spine surgery by fusion with MR images [6]. This was performed by rigidly registering the bone responses in US to the pre-segmented MR image. An automatic segmentation followed by manual correction was done for this purpose. CT-based and MR-based methods require an accurate segmentation of the vertebrae in the CT data, which cannot effectively fit in the clinical workflow. This is because manual segmentation of preprocedure images is very time-consuming and labour intensive, and hence not practical for anesthesiologists or interventional radiologists in busy clinics. On the other hand, automatic segmentation of preprocedure data for accurate intraoperative registration to US within the
operation time has been proven challenging. Moreover, preprocedure images often cover only a portion of the anatomy that is imaged with US during the procedure for targeting. The limited fields-of-view mainly of these images is mainly intended to minimize radiation dose in CT or reduce acquisition time in MR. Hence, even in the presence of accurate segmentation of the spine from the preprocedure CT or MR data, these images may not cover the entire region of interest on the lumbar spine for guiding anesthesia. Finally, although CT images can provide improved anatomical representation to US, CT image acquisition is harmful due to the ionizing nature of this modality.
1.1. Clinical Background

Figure 1.3: An overview of the state-of-the-art methods for augmenting anatomical information with spinal US. Methods (a), (b) and (c) are proposed by [12, 21, 35, 36], [10, 18, 25, 27], and [5] as well as this thesis, respectively.
Model-based US Augmentation

An alternative approach for augmenting spinal US images is to fuse them with statistical anatomical models of the imaged anatomy (Fig. 1.3(b)). Several works have developed statistical models of the spine for augmentation purposes. Boisvert et al. [7] constructed a statistical pose model that focused on detecting relative positions of individual vertebrae with respect to their neighbors. In this work, the shape of the full spine were further studied and the pose model was used to register to radiographic images. Our group has previously proposed several methods to register a statistical atlas of the lumbar spine to the spinal US [10, 18, 25, 27]. Khallaghi et al. [18] focused on creating statistical shape models for each vertebra. In order to impose constraints on the relative pose of neighboring vertebrae, a biomechanical model was further used. The constrained shape atlas was then registered to US images. The problem with this method is that having multiple shape models ignores the similarities between shape of different vertebrae and makes the computations unnecessarily expensive. To address these problems, Rasoulian et al. [25] proposed a more complex multi-body shape+pose statistical model for the lumbar spine. This statistical model was developed by viewing the spine as a single unified structure constructed of five bones, which are allowed to move within a certain range. This multi-vertebrae shape+pose model was then used to segment CT images of the lumbar spine. In a later study, the same model was used for guiding anesthesia in real-time in tracked 2D US [27]. Brudfors et al. further investigated registering this model to tracker-less 3D US [10]. While being promising, the above model-to-US registration methods have not been clinically accepted. This is because it is difficult to evaluate the patient-specific models that have been registered to US since the problem of interpreting spinal US still exists. As a results, it is challenging for clinicians to decide whether or not the registration results are reliable and can be used as a basis for their guidance decision. Furthermore, registration errors obtained using these methods vary in magnitude and location and tend to be sensitive to several factors such as the registration parameters, initialization and regularization, image quality, etc. This sensitivity is mainly due to the presence of noise and the sparsity of US images in terms of including anatomical information. Hence, for such methods to gain clinical acceptance, their accuracy and reliability should be improved. Furthermore, these methods should provide the clinicians with additional confidence regarding the accuracy of the US augmentation in order for the clinicians to determine whether or not the guidance decisions can be based on the registration results.
1.2 Thesis Objective

The overall objective of this thesis is to propose an effective augmentation framework for guiding facet joint and epidural injections using preprocedure data and intraprocedure US. To ensure reliability of spine registration in US, we present a joint framework for registration of statistical models of the lumbar spine to preoperative images (CT or MR) and intraoperative US (Fig. 1.3(c)). Essentially, this framework combines the two fusion techniques mentioned in section 1.1.2 and attempts to eliminate the downfalls of each one. Including a statistical model, eliminates the need for segmentation of the vertebrae, a challenge faced in US-CT or US-MR fusion. A joint framework makes it possible to take advantage of all available information, i.e., two or more imaging modalities of the same patient as well as prior information about the anatomy coming from a statistical atlas. This can lead to a more accurate localization of the lumbar spine in US images. On the other hand, involving multiple imaging modalities in the registration allows for a simultaneous visualization of the model on both US and preoperative image spaces. Finally, this can help the clinician assess the registration accuracy and decide whether to base guidance decisions on the registration results.

In this thesis, an improved and extended version of the existing model-based spine augmentation system is presented, which registers statistical atlases to multiple imaging modalities, i.e. preoperative CT or MR and intraoperative US images. The framework is set up based on the assumption that unlike the pose, the spine shape and size for a given subject are the same across multiple modalities. The statistical shape+pose model is improved by reformulating it to obtain a statistical shape+pose+scale model. The joint optimization of the shape and scale coefficients allows for flexibility to modify each parameter independently. The proposed method is validated on US+CT data to ensure feasibility. It is applied on US+MR data thereafter.

1.3 Thesis Outline

The rest of this thesis is divided into four chapters outlined below. Chapter 2 introduces the general framework for joint registration of medical images using multi-body statistical atlases. A statistical shape+pose model and a newly developed shape+pose+scale model of the lumbar spine are presented. A coherent point drift (CPD)-based registration framework is used for registering the statistical model to the target medical images of the
lumbar spine. Expressions for optimizing model parameters (shape, pose and scale) are derived thereafter. We then propose to modify the registration objective function to achieve a generalized version of the registration framework for two or more imaging modalities. This generalized framework allows for a joint registration of the statistical models to US with CT and/or MR images. In Chapter 3, the proposed framework is validated on a dataset of 10 preoperative CTs and corresponding intraoperative US images. The details of the preprocessing steps such as bone surface extraction and registration initialization are discussed. The registration results are reported on both US and CT spaces in terms of Target Registration Errors (TRE) and Surface Distance Errors (SDE). The results obtained by the joint framework are compared with existing methods. Chapter 4 includes a similar study of the method on a different dataset, which consists of nine intraoperative US and preoperative MR scans. In this chapter, the joint registration method is modified to a weighted joint registration framework, where weights are assigned to different regions of vertebrae in each modality, based on the visibility of the regions. Preprocessing steps specific to MR images are discussed. Several experiments are carried out and TRE results achieved on US and MR spaces are reported. Finally, Chapter 5 summarizes and concludes the thesis. The major contributions and future works are also highlighted in this chapter.
Chapter 2

Model-based Joint Registration Framework

This section focuses on the general mathematical framework for registering a statistical model of the lumbar spine to multiple imaging modalities. In Section 2.1, a shape+pose statistical model and a newly developed shape+pose+scale model of the lumbar spine are described. Section 2.2 introduces the joint registration framework by revisiting the optimization problem of registering a statistical atlas to a target image and extending it to multiple targets, i.e. multiple imaging modalities. Optimization of the parameters of the statistical models (shape, pose and scale) are further discussed.

2.1 Statistical Multi-vertebrae Models of the Lumbar Spine

2.1.1 Model Construction

In this thesis, two versions of statistical multi-vertebrae models are used for the lumbar spine: (i) a shape+pose model (Section 2.1.2) and (ii) a newly developed shape+pose+scale model (Section 2.1.3). These models are constructed based on the work presented by Rasoulian et al. [24]. Precisely, model construction is done by performing Principal Component Analysis (PCA) on a training set. The training set used in this thesis consists of 32 segmented CT images of the full lumbar spine. Each CT image includes lumbar vertebrae L1-L5. The data set was previously acquired and manually segmented in an previous study [24]. The PCA yields to the Principal Components (PCs) of the shape, pose and scale observed as observed in the training set. Theoretically, the PCs can describe the shape and pose of test images.
2.1.2 Shape+pose Model

Intuitively, the shape parameter of the model refers to the nonrigid transformation defining the shape of the bones. In the present method, the five lumbar vertebrae are assumed to be correlated in terms of shape. Hence, variations of the shape describe the shape of the overall multi-vertebrae structure, as opposed to those of individual bones. The model’s pose parameter, on the other hand, refers to the rigid transformation affecting the relative position of the vertebrae with respect to each other.

An instance of the shape+pose model is created by applying the desired coefficients (or weights) to each of the shape and pose PCs determined in model training and linearly adding them. Assuming that $\theta_s^k$ is the coefficient applied to the $k$th shape PC and $\theta_p^k$ is the coefficient applied to the $k$th pose PC, the $l$th object of the model can be instantiated as follows:

$$
\Phi(\theta_s, \theta_p) = \Phi_p^l(\Phi_s^l(\theta_s); \theta_p),
$$

where $\Phi^p_l(\cdot; \theta_p)$ represents a similarity transformation built as a weighted combination of the pose PCs and $\Phi^s_l(\cdot)$ describes the associated shape instantiation. The first and second modes of variation of the shape and pose for the shape+pose model are demonstrated in Fig. 2.1.

2.1.3 Shape+pose+scale Model

A shortcoming of the shape+pose model is that variations in the size of the anatomy are accounted for in the rigid transformation for pose [3]. Hence, the scale cannot be optimized independently of the pose at the time of registration. In order to better capture the pose and scale variations, we propose to decouple scale from pose in the model. Furthermore, decoupling the scale from the pose makes it possible to estimate the scale jointly and the poses independently. In other words, this guarantees that the jointly registered patient-specific models on multiple modalities have equal sizes.

Formally, for the shape+pose model, the pose can be formulated as the matrix $T_{n,l}$ which rigidly transforms the $n$-th point on the model to the $l$-th point on the target. More precisely,

$$
T_{n,l} = \begin{bmatrix}
k_{n,l} & R_{n,l} & t_{n,l} \\
0 & 1 & 1
\end{bmatrix}
$$

where $k_{n,l}$, $R_{n,l}$ and $t_{n,l}$ express the scaling, rotation and translation components of the rigid transformation, respectively. In order to separate the scaling from the rest of the pose transformations, we can rewrite $T_{n,l}$ as the
2.1. Statistical Multi-vertebrae Models of the Lumbar Spine

Figure 2.1: First two modes of variation of shape ((a) and (b)) and pose ((c) and (d)) for the shape+pose model. Images on the left of each pair indicate $-3\sqrt{\lambda}$, and those on the right correspond with $+3\sqrt{\lambda}$. Values of $\lambda$ are the eigenvalues determined for each eigenvector (PC) at training.

Product of two independent transforms $T_{n,l} = T_{n,l}^1 \times T_{n,l}^2$. Here, $T_{n,l}^1$ is the scaling transform and $T_{n,l}^2$ indicates the rotation and translation transforms. That is,

$$T_{n,l}^1 = \begin{bmatrix} k_{n,l} & 0 & 0 & 0 \\ 0 & k_{n,l} & 0 & 0 \\ 0 & 0 & k_{n,l} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.3}$$

$$T_{n,l}^2 = \begin{bmatrix} R_{n,l} & t_{n,l} \\ 0 & 1 \end{bmatrix} \tag{2.4}$$

An instance of the shape+scale+pose model can be created like below:

$$\Phi(\theta^s, \theta^k, \theta^p) = \Phi^p(\Phi_k^s(\Phi^s(\theta^s); \theta^k); \theta^p) \tag{2.5}$$

Here, $\Phi^s(\cdot; \theta^k)$ describes the scaling transformation with coefficients $\theta^k$.

Figure 2.2 demonstrates the scale variations (one mode) as well as the first two modes of variations for the pose in the shape+pose+scale model.
2.2. Point-set Registration of Statistical Models to Targets

The shape variations are not shown here as they are identical to those of the shape+pose model. Here, the size of the individual vertebrae does not change by varying the pose (Fig. 2.2 (a)). This is unlike what is observed with the shape+pose model shown previously (Fig. 2.1 (b)).

Figure 2.2: First two modes of variation of pose ((a), (b)) and the first mode of variation of the scale (c) for the shape+pose+scale model. This model decouples the scale from the pose. Images on the left of each pair indicate \(-3\sqrt{\lambda}\), and those on the right correspond with \(+3\sqrt{\lambda}\). The shape variations are similar to the shape+pose model in 2.2. Values of \(\lambda\) are the eigenvalues determined for each eigenvector (PC) at training.

2.2 Point-set Registration of Statistical Models to Targets

The registration of the model to targets is based on the Coherent Point Drift framework proposed by Myrorenko et al. [22]. Here, the registration problem essentially consists of finding the optimal parameters for the model to obtain minimized net distance between the model point cloud of the model and target images. The bone surfaces are represented by point clouds for
both the model and targets. Different bone surface extraction methods used for US, CT and MR will be discussed in Chapters 3 and 4.

2.2.1 Registration of the Statistical Models to a Single Imaging Modality

Rasoulian et al. previously showed the statistical shape+pose model could be registered to a target point cloud representing bone surfaces in CT \[24\] or US \[27\] images. In this method, an Expectation-Maximization (EM) algorithm is followed for the model registration. The expectation step consists of computing the probabilities of any $n$-th model point belonging to the $l$-th object ($t^l_n$) generating a target point $z^m$. In other words, the soft correspondences $P(t^l_n | z^m)$ between the $m$-th model instance point of $l$-th bone ($t^l_n$) and $n$-th target point of the image ($z^m$). The maximization step involves optimizing the problem’s objective function (Eq. 2.6) using the correspondences found in the expectation step. The following objective function is optimized

$$Q = \sum_{l=1}^{L} \sum_{m,n=1}^{M,N_1} P(t^l_n | z^m) \| z^m - \Phi(t^l_n, \theta^s, \theta^p) \|^2 + R^s + R^p \quad (2.6)$$

where $\Phi(t^l_n, \theta^s, \theta^p)$ denotes the transformation of point $t^l_n$ on the model with the shape and pose coefficients $\theta^s$ and $\theta^p$. $M$, $N_1$ and $L$ represent the total number of points in the model per bone, total number of points in the I-th target point cloud and the total number of bones in the atlas, respectively. The regularizers $R^s$ and $R^p$ constrain the shape and pose variations, respectively, and are computed according to \[24\]. For imaging modalities such as CT and MR, target points $z^m$’s are assumed to be edge points extracted from the bones and they are all equally weighted. However, when used for US, $z^m$’s are assigned probabilities indicating how likely the $m$-th target point is to represent a bone pixel. Calculation of $P(t^l_n | z^m)$ has been presented in \[24\]. Similarly, for the shape+pose+scale model, the cost function can be written as:

$$Q = \sum_{l=1}^{L} \sum_{m,n=1}^{M,N_1} P(t^l_n | z^m) \| z^m - \Phi(t^l_n, \theta^s, \theta^p, \theta^k) \|^2 + \text{Regularizers} \quad (2.7)$$

Here, $\Phi(t^l_n, \theta^s, \theta^p, \theta^k)$ represents the transformed model with shape, pose and scale coefficients $\theta^s$, $\theta^p$ and $\theta^k$, respectively.
2.2. Joint Registration of the Statistical Models to Multiple Imaging Modalities

In order to incorporate information from more than one modality, we modify the cost function to include the sum of all distances from the points belonging to the model and those of two (or more) target modalities of interest, such as US, CT and MR. For any imaging modality involved in the fusion, a model instance is created. These instances will all have the same shape coefficients, whereas the pose coefficients are specific to each corresponding modality. This is because the shape of the vertebrae remains the same before and during the procedure for each individual patient. The pose of the individual vertebrae, on the other hand, will change from one imaging modality to another. For example, preprocedure CT and MR images are acquired in the prone position, unlike intraprocedure US, which often requires the patient to be in the supine position.

Let \( m_d \in M = \{ \text{US}, \text{CT}, \text{MR} \} \) denote the target’s imaging modality. \( z_{m,md} \) will then represent the \( m \)-th point from the point cloud (of size \( M_{md} \)) of each given modality. Therefore, \( \Phi(t^l_n, \theta^s, \theta^p_{md}) \) represents the transformation of model points \( t^l_n \) with the common shape coefficients \( \theta^s \) and the individual modality-specific pose coefficients \( \theta^p_{md} \). To sum over the distances between the transformed model and the target points of both modalities, we extend the objective function to the following (Eq. 2.8).

\[
Q = \sum_{m_d \in MD} \sum_{l=1}^{L} \sum_{m,n=1}^{M_{md},N_l} P(t^l_n|z_{m,md})||z_{m,md} - \Phi(t^l_n, \theta^s, \theta^p_{md})||^2 + R^s + R^p_{md}.
\]

(2.8)

The above equation is the generalized objective function and holds for any subset \( MD \subseteq \{ \text{US}, \text{CT}, \text{MR} \}; MD \neq \emptyset \). In addition to the shape, the size of a patient’s spine remain the same in different imaging modalities. Therefore, for the shape+pose+scale model, we can jointly optimize the shape+scale coefficients from two or more imaging modalities, while independently optimizing the pose coefficients for each image (Eq. 2.9).

\[
Q = \sum_{m_d \in MD} \sum_{l=1}^{L} \sum_{m,n=1}^{M_{md},N_l} P(t^l_n|z_{m,md})||z_{m,md} - \Phi(t^l_n, \theta^s, \theta^p_{md})||^2 + \text{Regularizers}
\]

(2.9)
2.2.3 Parameter Optimization for Model-to-Target Registration

In order to optimize the model parameters, every iteration, the soft correspondences are computed in the estimation step. Practically, the maximization step is then carried out in two steps. First, the common shape (and scale) coefficients are jointly optimized. Then, \( n(MD) \) independent sets of modality-specific pose coefficients are estimated. The optimal values of the coefficients can be computed using the derivatives of the cost functions with respect to the model parameter. For the shape+pose model, differentiating Eq. (2.8) with respect to the common shape coefficients yields to Eq. 2.10 below.

\[
\frac{\partial Q}{\partial \theta_s} = \sum_{md \in MD} \sum_{l=1}^{L} \sum_{m,n=1}^{M_{md},N_{l}} P(t_n^l | z_{m,md}) \left[ \Phi(t_n^l)^\top - z_{m,md}^\top \frac{\partial \Phi(t_n^l)}{\partial \theta_s} \right] + R^s. 
\] (2.10)

Differentiation with respect to each set of pose coefficients \( \theta_p \) can be done in a similar fashion, and yields to the expression below (Eq. 2.11).

\[
\frac{\partial Q}{\partial \theta_p_{md}} = \sum_{md \in MD} \sum_{l=1}^{L} \sum_{m,n=1}^{M_{md},N_{l}} P(t_n^l | z_m) \left[ \Phi(t_n^l) - z_m \frac{\partial \Phi(t_n^l)}{\partial \theta_p} \right] + R_{p_{md}}^p. 
\] (2.11)

Calculation of partial derivatives \( \frac{\partial \Phi(t_n^l)}{\partial \theta_s} \) and \( \frac{\partial \Phi(t_n^l)}{\partial \theta_p} \) has been previously presented in [24].

We can similarly differentiate Eq. 2.13 the shape+pose+scale model parameters(Eqs. 2.12, 2.13).

\[
\frac{\partial Q}{\partial \theta_s} = \sum_{md \in MD} \sum_{l=1}^{L} \sum_{m,n=1}^{M_{md},N_{l}} P(t_n^l | z_{m,md}) \left[ \Phi(t_n^l) - z_{m,md}^\top \frac{\partial \Phi(t_n^l)}{\partial \theta_s} \right] + R^s. 
\] (2.12)

\[
\frac{\partial Q}{\partial \theta_p_{md}} = \sum_{md \in MD} \sum_{l=1}^{L} \sum_{m,n=1}^{M_{md},N_{l}} P(t_n^l | z_m) \left[ \Phi(t_n^l) - z_m^\top \frac{\partial \Phi(t_n^l)}{\partial \theta_p} \right] + R_{p_{md}}^s. 
\] (2.13)

\[
\frac{\partial Q}{\partial \theta_k} = \sum_{md \in MD} \sum_{l=1}^{L} \sum_{m,n=1}^{M_{md},N_{l}} P(t_n^l | z_{m,md}) \left[ \Phi(t_n^l) - z_{m,md}^\top \frac{\partial \Phi(t_n^l)}{\partial \theta_k} \right] + R^k. 
\] (2.14)
2.2. Point-set Registration of Statistical Models to Targets

The partial derivatives $\frac{\partial \Phi(t^i_n)}{\partial \theta_s}$ and $\frac{\partial \Phi(t^i_n)}{\partial \theta_p}$ are similar to before. The scale parameter partial derivative $\frac{\partial \Phi(t^i_n)}{\partial \theta_k}$ can be computed similar to that of the pose as presented in [24].
Chapter 3

Joint Registration of Model to US+CT Images of the Lumbar Spine

This chapter presents a study that applies the joint framework proposed in Chapter 2 to a dataset of US+CT images. In this chapter, the joint registration algorithm is revisited, with a focus on the preprocessing steps required for US and CT images (Section 3.2). Several experiments are performed on this data. Comparison is made between results obtained using the old model-to-US and the proposed joint methods, as well as between shape+pose and shape+pose+scale results. The results are reported and discussed in Sections 3.3 and 3.4.

3.1 US and CT Datasets

The dataset used in this study consists of volumetric US data and corresponding CT scans of 10 patients. US volume acquisition was performed with an electromagnetically 2D tracked US probe, in a zigzag sweep in a previous study [27]. Each US volume consists of 307 to 550 2D sagittal slices. 3D reconstruction of the US volume was performed using the PLUS software. Each US image includes all the lumbar vertebrae L1-L5, as well as the sacrum. The voxel size varies throughout the dataset between (0.5, 0.5, 0.5) to (1, 1, 1)mm.

3.2 Joint Registration for US+CT Images

Figure [3.1] shows the workflow for joint registration of US and CT image pairs. The steps are explained in details in Sections 3.2.1-3.2.4 below.
3.2. Joint Registration for US+CT Images

3.2.1 Target Point Cloud Generation

Prior to registration, US and CT images undergo a preprocessing step where visible bone surfaces are extracted from each imaging modality to form point clouds. Registration is later performed by minimizing the distances between the model and both of these target point clouds. The US volume is preprocessed by applying a phase-based bone enhancement technique [17] on individual native 2D slices. This bone enhancement results in a stack of bone probability maps from which a reconstructed 3D map can be obtained. The intensity value of each voxel on this map indicates the probability that this voxel represents bone tissue. A threshold is then applied to the map. Voxels above this threshold along with their corresponding probability values make up the probabilistic point cloud used in optimization. The probability assigned to each point helps determine the soft correspondences in the estimation phase of the EM algorithm. For the CT volume, the bone point cloud is obtained from the visible edges of the spine. Edges are extracted using a simple Canny edge detector [24]. We do not assign any predefined probabilities for these points.

3.2.2 Geometric Initialization of the Model Instances on Targets

Another essential step before registration is geometric initialization of the model instances on the two target point clouds. This is to ensure registration does not converge far off the optimal solution. In most cases, the assumption can be made that CT and US are acquired in supine and prone positions, respectively. This assumption allows the model to be aligned roughly at the correct anatomical position using a simple one-click initialization. For this purpose, approximate center of gravity of the L3 vertebra is manually selected on US and CT volumes. Two instances of the model (one for each modality) are then translated accordingly. The L3 vertebra is found by

<table>
<thead>
<tr>
<th>Field of View</th>
<th>Patient Numbers</th>
<th>Count (Number of Subjects)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2-L5</td>
<td>P2, P8</td>
<td>2</td>
</tr>
<tr>
<td>L3-L5</td>
<td>P1, P10</td>
<td>2</td>
</tr>
<tr>
<td>L4-L5</td>
<td>P3-P7, P9</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 3.1: Breakdown of the fields of view of images in the CT dataset based on the visible lumbar vertebrae.
identifying the sacrum and counting up. A rigid registration is performed thereafter to correct for rotations between the model and targets and further improve the initial alignments.

### 3.2.3 Optimization

Once the target point clouds are obtained and an instance of the model is placed on each, the iterative shape+pose or shape+pose+scale registration begins. Every iteration, the model parameters are estimated as described in Chapter 2. Pose coefficients are estimated and updated for each individual modality, whereas the shape and scale are computed jointly.

### 3.2.4 Model Transformation

At the end of each iteration, the model instances are updated according to the new model coefficients and the models are projected back to the CT and US spaces. This process continues until convergence. In the case of CT, a nonrigid registration is performed at the last step for a more accurate localization of the vertebrae [26].
3.2. Joint Registration for US+CT Images

Figure 3.1: An overview of the proposed framework for joint registration of US, preoperative images and a statistical model.
3.3 Results

3.3.1 Registration Parameter Selection

For choosing the number of PCs for the CT pose and the joint shape, a separate set of 10 segmented, publicly available CT data sets [16] were used and an exhaustive parameter search was performed with respect to the Surface Distance Error (SDE). The number of PCs for the US pose was chosen according to [27]. The US-related parameters are set according to prior work by Hacihaliloglu et al. [17].

3.3.2 Validation

In addition to the joint method, the registration results are reported for a US-only method, where shape and scale optimization are done using US only, without involving any information from CT. This helps study the improvement achieved by the joint technique. Similarly, CT-only experiments are carried out to further investigate the reliability of the joint method. All experiments are done for each of the two (i.e. shape+pose and shape+pose+scale). The joint registration accuracy is evaluated quantitatively in both CT and US domain.

Experiments in CT Domain

In this step, each patient-specific model is projected to the CT space, i.e., the pose of the patient-specific model is computed from the CT image. The gold standard is obtained by manually segmenting the fully-visible vertebrae of each CT scan, which serves as the reference. For each visible vertebra in the CT volume, the shape SDE is computed between the registered model surface and the surface obtained from the gold standard segmentation. The correspondences between the model and segmentation point clouds are determined based on nearest neighbors. Hence, SDE is calculated as the Euclidean distance between the points on one surface to their closest neighboring points on the other. In order to investigate only the shape errors, each vertebra of the model is rigidly registered to the corresponding vertebra on the gold standard. This step align the model on the gold standard and eliminates potential pose errors. The shape errors are reported at several regions for more thorough and detailed analysis of the results. These regions include the spinous processes (SPs), transverse processes (TPs), anterior and posterior articular processes (APs) and vertebral bodies (VBs).
3.3. Results

The break-down of the different regions is done by manually selecting these regions on the gold standard on each vertebra, for each patient.

Experiments in US Domain

Similarly, the registered models are projected to the US space for this step. That is, the model’s pose is computed from the US scans only, while the shape and scale are optimized jointly. Landmarks were selected by an expert sonographer on the SPs and laminae (LAs) on sagittal slices, as well as TPs and APs on sagittal and axial slices. The US registration accuracy is measured using the Target Registration Error (TRE) metric calculated between pre-selected landmarks on the gold-standard and corresponding target points on the registered model’s surface. In the lack of a better alternative, the correspondences are assigned using the nearest neighbors approach.

3.3.3 US+CT Joint Registration Results and Analysis

The joint method was successfully applied on the US+CT dataset. Examples can be seen in Fig. 3.2. Registration results obtained in the different experiments are provided in Tables 3.2 and 3.3. The columns of these tables indicate the registration method, in terms of shape optimization. That is, CT shape refers to registration of the model by optimizing the shape using CT only. Joint shape refers to the proposed joint registration. US shape represents the method where the shape parameter are computed from US only. Colormaps in Fig. 3.3 illustrate the improvements achieved with the joint method and the shape+pose model. The joint registration technique in the CT domain resulted in the smallest errors at all regions, even out-performing the CT-only method. The largest errors can be seen at the SPs of the vertebrae. In the US domain, the joint shape improves the TRE at all regions except for the APs. Figure 3.4 shows the distribution of TRE for the US images using the three different shape optimization techniques at the different regions of the vertebrae (3.4(a)), as well as at the five vertebrae L1-L5 (3.4(b)). A statistical -test analysis was carried out. No statistically significant improvements were observed using the joint method over the other two. The joint registration process had an average run-time of 7.5 minutes, approximately 6.2 minutes of which was used for target point extraction, and 1.3 minutes for the actual shape+pose registration.

Tables 3.2 and 3.3 summarize the SDE and TRE differences of the two methods at different regions of the vertebrae, namely SP, TP, AP and VB. The joint approach leads to an overall improvement in the maximum and
3.3. Results

Figure 3.2: Example snapshots of the registration results using the shape+pose model on CT (a) and US (b) images. The joint method (blue) performs better than the model-to-US technique (red). The yellow annotations on US depict the gold standard segmentation obtained by the sonographer.

RMS errors.

Figure 3.3 highlights the differences between the joint and the US-only registration as colormaps projected on the individual vertebrae for each of the 10 patients.

The joint method combined with the shape+pose+scale model improves the overall SDEs and TREs compared to the US-only technique, i.e. the method that only uses US for shape optimization (Figure 3.5). The CT-only performs slightly better than the joint method. In US, the region with the smallest errors is the AP and largest errors are seen at LAs. In the CT domain, the errors are within the same range. SDE and TRE results obtained using the shape+pose+scale model are provided in Tables 3.4 and 3.5. Colormaps in Fig. 3.6 illustrate the improvements achieved with the joint method and the shape+pose+scale model. Box plots also show the break-down of TRE errors for US at different regions and vertebrae (3.7). The average run-time for the iterative registration was 1.9 minutes.

A comparison is made between the results obtained using the joint framework along with the two different models. Colormaps in Fig. 3.8 represent
### 3.3. Results

<table>
<thead>
<tr>
<th>Region</th>
<th>Reg. Type</th>
<th>CT-only</th>
<th>Joint</th>
<th>US-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td></td>
<td>2.88 ± 0.48</td>
<td>2.81 ± 0.56</td>
<td>2.97 ± 0.61</td>
</tr>
<tr>
<td>LA</td>
<td></td>
<td>2.71 ± 0.53</td>
<td>2.67 ± 0.65</td>
<td>2.86 ± 0.56</td>
</tr>
<tr>
<td>SP</td>
<td></td>
<td>3.26 ± 0.88</td>
<td>3.17 ± 0.80</td>
<td>3.24 ± 0.89</td>
</tr>
<tr>
<td>VB</td>
<td></td>
<td>2.92 ± 0.51</td>
<td>2.88 ± 0.54</td>
<td>3.07 ± 0.62</td>
</tr>
<tr>
<td>All</td>
<td></td>
<td>2.81 ± 0.50</td>
<td>2.40 ± 0.56</td>
<td>2.93 ± 0.49</td>
</tr>
</tbody>
</table>

Table 3.2: Distribution of US TRE registration errors (mm) achieved using the shape+pose model for the US+CT dataset, per region. No significant improvement was observed using the joint registration.

<table>
<thead>
<tr>
<th>Region</th>
<th>Reg. Type</th>
<th>CT-only</th>
<th>Joint</th>
<th>US-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td></td>
<td>1.65 ± 0.57</td>
<td>1.60 ± 0.56</td>
<td>1.42 ± 1.42</td>
</tr>
<tr>
<td>LA</td>
<td></td>
<td>2.99 ± 1.90</td>
<td>2.97 ± 1.88</td>
<td>3.18 ± 3.18</td>
</tr>
<tr>
<td>SP</td>
<td></td>
<td>1.81 ± 0.88</td>
<td>1.70 ± 0.78</td>
<td>2.03 ± 2.03</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>2.11 ± 1.23</td>
<td>1.99 ± 1.24</td>
<td>2.14 ± 2.14</td>
</tr>
<tr>
<td>All</td>
<td></td>
<td>2.94 ± 2.40</td>
<td>2.93 ± 2.40</td>
<td>2.99 ± 2.78</td>
</tr>
</tbody>
</table>

Table 3.3: US TRE (mm) achieved using the shape+pose model for the US+CT dataset. In each experiment (each column), the shape parameter of the model is optimized from the modalities indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, TP) and the average of all four.

...the difference in errors obtained using the two models...
3.3. Results

<table>
<thead>
<tr>
<th>Region Type</th>
<th>CT-only</th>
<th>Joint</th>
<th>US-only Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>2.48 ± 0.43</td>
<td>2.55 ± 0.50</td>
<td>3.17 ± 0.70</td>
</tr>
<tr>
<td>LA</td>
<td>2.37 ± 0.51</td>
<td>2.41 ± 0.53</td>
<td>3.04 ± 0.79</td>
</tr>
<tr>
<td>SP</td>
<td>2.63 ± 0.68</td>
<td>2.72 ± 0.70</td>
<td>3.31 ± 0.74</td>
</tr>
<tr>
<td>VB</td>
<td>2.44 ± 0.43</td>
<td>2.53 ± 0.50</td>
<td>3.11 ± 0.67</td>
</tr>
<tr>
<td>All</td>
<td>2.35 ± 0.37</td>
<td>2.42 ± 0.39</td>
<td>3.15 ± 0.53</td>
</tr>
</tbody>
</table>

Table 3.4: CT SDE registration errors (mm) achieved using the shape+pose+scale model for the US+CT dataset. In each experiment (each column), the shape and scale parameters of the model are optimized from the modalities indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, VB) and the average of all four.

<table>
<thead>
<tr>
<th>Region Type</th>
<th>CT-only</th>
<th>Joint</th>
<th>US-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>1.53 ± 0.61</td>
<td>1.49 ± 0.73</td>
<td>1.30 ± 1.42</td>
</tr>
<tr>
<td>LA</td>
<td>3.10 ± 2.08</td>
<td>3.08 ± 2.09</td>
<td>3.26 ± 2.21</td>
</tr>
<tr>
<td>SP</td>
<td>2.22 ± 0.95</td>
<td>2.12 ± 0.90</td>
<td>2.37 ± 0.93</td>
</tr>
<tr>
<td>TP</td>
<td>2.29 ± 1.18</td>
<td>1.96 ± 1.19</td>
<td>2.36 ± 1.17</td>
</tr>
<tr>
<td>All</td>
<td>3.16 ± 2.47</td>
<td>3.14 ± 2.47</td>
<td>3.22 ± 2.50</td>
</tr>
</tbody>
</table>

Table 3.5: US TRE (mm) achieved using the shape+pose+scale model for the US+CT dataset. In each experiment (each column), the shape and scale parameters of the model are optimized from the modalities indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, TP) and the average of all four.
Figure 3.3: Colormaps showing the difference between shape SDE values per surface point computed from US-to-atlas and joint registrations, using segmented reference CT as the gold standard for patients 1 (a) to 10 (j). Blue regions show improvements achieved by the joint registration, where US-to-atlas error is larger than joint registration error.
3.3. Results

Figure 3.4: Box plots showing the break-down of TRE values obtained using the shape+pose model for US images (US+CT dataset) at different regions of the vertebrae (a) and the five different vertebrae (b). No significant improvements are observed.
3.3. Results

Figure 3.5: Example snapshots of the registration results using the shape+pose+scale model on CT (a) and US (b) images. The joint method (blue) performs better than the model-to-US technique (red). The yellow annotations on US depict the gold standard segmentation obtained by the sonographer.
Figure 3.6: Colormaps showing the difference between shape SDE values per surface point computed from US-to-atlas and joint registrations, using segmented reference CT as the gold standard for patients 1 (a) to 10 (j). Blue regions show improvements achieved by the joint registration, where US-to-atlas error is larger than joint registration error.
3.3. Results

Figure 3.7: Box plots showing the break-down of TRE values obtained using the shape+pose+scale model for US images (US+CT dataset) at different regions of the vertebrae (a) and the five different vertebrae (b). No significant improvements are observed.
Figure 3.8: Colormaps showing the difference between shape SDE values per surface point computed from US-to-atlas and joint registrations, using segmented reference CT as the gold standard for patients 1 (a) to 10 (j). Blue regions show improvements achieved by the joint registration, where US-to-atlas error is larger than joint registration error.
3.4 Discussion and Summary

In this chapter, a joint registration framework was presented, where statistical models of the lumbar spine were registered to US and CT with the aim to improve interpretation of the anatomy in spinal US for facet joint injections and epidural needle insertions. The key contribution involved the concurrent optimization of the shape and scale coefficients, which accommodates multimodal fusion, despite the substantial pose differences between the modalities. It was demonstrated that by taking advantage of available easy-to-interpret preprocedure CT data for computing the shape of the statistical model, improved performance of the US registration can be achieved. Looking at the results at individual regions of the vertebrae, as opposed to the overall average, gives us better insight into the registration accuracy. The joint registration lead to a more accurate overlay of the anatomical information of the lumbar spine in both CT and US domain compared to the approach where only US information was used for registration. Moreover, as a by-product of the joint registration, the pose of the spine in CT is also estimated (Fig. 3.1). This allows the model to be transformed and overlaid onto the CT domain, helping the clinician to qualitatively evaluate the local registration accuracy of the patient-specific model constructed for the US image based on the detailed anatomy visible in the CT scan. Hence, including preprocedure CT makes it possible to simultaneously visualize the spine anatomy in both US and CT domains without the need for preoperative segmentation of CT. It can be seen from Tables 3.2–3.5 that the joint registration almost always performs better than the US-only method both in terms of RMS errors and standard deviation. The differences between the two methods are greater with the shape+pose+scale model, as opposed to the shape+pose model. This is most likely because in the absence of CT data, the false bone points lead to incorrect scale coefficient estimate, which is independently optimized. Although results are often improved by joint registration, the errors in isolated anatomical regions (shown in red) seem to have increased. We noticed that the registration accuracy of the joint model in more complicated areas of the processes was not high enough in the CT domain which propagated to an apparently higher error in the US domain. It is speculated that this issue stems from having a small size of training set (32 patients) used to generate the statistical spine model. Enlarging the training set can potentially lead to better results. Furthermore, correct edge detection for point cloud extraction in CT scans strongly depends on the quality and resolution of these images. As the examined CT volumes originate from routine clinical acquisitions with inhomogeneous voxel resolution,
3.4. Discussion and Summary

Figure 3.9: Snapshots of a CT scan belonging to a scoliotic patient (P8), which was proven challenging to register. The vertebrae seem to diverge from the vertical yellow line (left). The yellow line on the coronal view also marks the sagittal slice on the right. White arrows indicate the failed pose registration of the model.

The bone surface is not sharply visible in all imaging axes, making it difficult to obtain a dense bone point cloud ideal for model registration. Improving CT image quality during acquisition or enhancing bone surface detection might further increase the registration accuracy. Figures 4.5 demonstrated the registration error values vary across different vertebrae. For both US and CT, the model registers better to L1-L3, as opposed to L4-L5. This variation is likely to originate from the fact that in general, L1-L3 of the models undergo smaller pose variations when going from supine to prone. That is, the significant change in curvature happens at L4-L5. Moreover, the manual alignment of the centre of gravity of L3 can contribute to the registration results at L3. As mentioned above, an important limitation of the current models is their relatively small training set. Hence, it is expected for the model to fail in registering to the spine of certain subjects with shapes or poses unlike those of the subjects in the training set. For example, an observation easily made from the colormaps in Figs. 3.3, 3.5 is that errors obtained for patient P8 (Fig. 3.3(h), 3.6(h), and 3.8(h)) are high and inconsistent, compared to the other nine cases. This is most likely due to the fact that this particular subject seems to be suffering from scoliosis, as the CT image shows a lateral curvature. This indicates that the current models and methods are not suited for scoliotic patients (Fig. 3.9).
3.4. Discussion and Summary

This problem can be eased by increasing the number of subjects in the training set. Another possibility is to employ a two-step registration technique. In the first step the multi-vertebrae model is registered to the data as shown. In the second step, the registered model is used as an initialization to register smaller models locally to one or more vertebrae of interest. Nevertheless, the advantage of the proposed framework is that in such scenarios, the clinicians will still be able to look at the patient-specific model on the preoperative space and decide about the adequacy of the registration.

A limitation of the performed evaluation is that potentially inaccurate correspondence assignment may lead to inaccurate error evaluation. Currently, we rely on finding the closest neighbors for assigning correspondences between the gold standard segmentations and the model point because manual correspondence assignment is not feasible. Correspondences are improved in the case of SDE calculation for the CT, as a rigid registration is carried out on each vertebra prior to computing the errors. However, a rigid registration for the US is not particularly useful for US segmentations are partial and very sparse in this modality. The required clinical accuracy for facet joint injections and epidurals are estimated at around 5 mm and 3 mm, respectively. Hence, it was shown that the accuracy of the presented framework is sufficient for facet joint injections. However, further improvement of maximum errors is needed for epidural procedures. The statistical spine models used in this study have been designed for the five lumbar vertebrae, with different pose and shape variations changing the entire structure as a whole. However, the CT scans used for testing the method included variable fields-of-view with two to four visible vertebrae. The increase in the errors for patients P3 and P6 (red spots on APs and SPs in Figures 3.6(c) and 3.6(f)) in the shape+pose+scale results could be due to the fact that only two vertebrae were visible in their CT scans. For patients P1, P2, and P9-10, whose CT scans included three or four vertebrae, improvements were more pronounced.

Finally, though tolerable for preoperative CT registration, the computational complexity of the current method is still too high to allow real-time implementation for intraoperative procedures. However, the most time-consuming step of the workflow is extracting the bone surfaces, specially for the US. This is because in the current method, individual 2D slices are processed one at a time.
Chapter 4

Joint Registration of Model to US+MR Images of the Lumbar Spine

This chapter involves the study of applying the joint registration framework to intraprocedure US and preprocedure MR scans of the lumbar spine. A modification of the method presented in the previous chapters has been employed where predefined weights are assigned to different regions of the lumbar vertebrae based on their presumed degree of visibility in each modality. Results of different joint registration experiments performed on this dataset are presented and discussed.

4.1 US and MR Datasets

The dataset used in this study consists of volumetric US data and corresponding multi-slice MR scans of nine patients. Ethics approval and written consent were obtained from the subjects. This dataset includes both male and female subjects ranging between 17 to 75 years in age.

MR images used in this study are actual clinical images acquired at hospitals for diagnosis and treatment purposes. Data was downloaded from the hospitals’ servers and anonymized. The MR images used for the registration are acquired sagittally since bone detection can be performed more easily and accurately on sagittal slices. T1-weighted images are chosen for this study because the spinal cord appears dark in these images, making the bone outlines more easily detectable. Each 3D MR image includes the coccyx and sacrum, five lumbar vertebrae, as well as one or more vertebrae from the bottom of the thoracic spine. The sagittal field-of-view is limited in these images as the main focus has been on capturing the vertebral bodies, which are close to the median plane. Hence, TPs, which lie more laterally from the median plane, are not captured. The slice thickness is between 3.5 to 4.4 mm. Figure 4.1 depicts an example of sagittal and axial slices of
4.1. US and MR Datasets

Figure 4.1: An example MR image from the present dataset. The yellow dashed line on the sagittal image (right) indicates the location of the axial slice (left). The TPs are not captured in the field-of-view of these MR images. The sagittal slices are $3.5 - 4.4$ mm thick in this dataset.

Similar to the previous chapter, the US data has been acquired using an electromagnetically 2D tracked US probe and the 3D volumes are reconstructed using PLUS. However, the present dataset includes more challenging images compared to those used in Chapter 3. Firstly, the 2D image resolution are relatively lower, with voxel dimensions of $(1.3, 1.3, 1.3)$ to $(1.5, 1.5, 1.5)$ mm, and a total of 103 to 203 slices. Also, unlike the US scans in the US+ct study, the images used in this study were obtained in the sitting position, which puts the vertebrae in a completely different pose. Also, in this dataset, the sitting position limited the range of the scan, in terms of how far down the spine the US transducer could go without making the patient uncomfortable. As a result, present US scans have a limited axial
4.2 Joint Registration for US+MR Images

4.2.1 Target Point Cloud Generation

The MR images are processed in the steps described below:

1. **Intensity Bias Correction**: Initially, the biased intensity of the image is corrected using a modified fuzzy C-means algorithm [2]. The parameters for this step are set according to the work by Suzani et
2. **Anisotropic Diffusion:** An anisotropic diffusion is applied to remove the speckles and noise present in the image. This helps reduce the number of points that will be falsely detected as edge points in the next step.

3. **Edge Detection:** A Canny edge detection is performed on the processed MR volume. To further refine the extracted edges and remove false positive points, a morphological erosion is performed thereafter. The final edge points make up the logical target point cloud for the MR image.

Similar to Chapter 3, the US images are processed using a phase-based bone enhancement extract a probabilistic point cloud [17].

### 4.2.2 Model Initialization on Targets

The centers of gravity of L3 are manually selected on the US and MR images. Two instances of the model are roughly aligned on the targets accordingly. This step is followed by a rigid registration for further correction. Since the sacrum is not visible in the US images of this dataset, the shadows of the ribs attached to T12 are used to identify T12 and hence the lumbar vertebrae below it. As mentioned before, the US images in this dataset were acquired in the sitting position. To account for this, the model instance used on the US space is initialized with a straight spine pose, as opposed to the mean pose of the model. This is because in the sitting position, the five lumbar vertebrae align on a less curved line as shown in Fig. 4.3.

### 4.2.3 Weighted Joint Registration

In US and MRI data, the visibility of each part of a lumbar vertebra differs due to various factors. In MR images, the outlines of VBs can be seen very clearly in almost all slices, making the VB edges reliable points for registration. Similarly, LAs have high visibility due to the different intensity levels inside and outside the boundaries of these regions in the sagittal images. However, the SPs are more challenging to detect. This is because the current MR slices are very thick and hence, SP of each vertebra is typically only visible on one or no sagittal slice. Also, high-frequency details of tissues surrounding the SPs lead to falsely detected edge points in the target point clouds. AP edges are also relatively more difficult to accurately
4.2. Joint Registration for US+MR Images

Figure 4.3: The lumbar spine in the sitting position. Figure in (a) shows that in the sitting pose, lumbar spine is almost straight but the L5-sacrum curve is sharper (courtesy of ergonut.com [22]). Figures in (b) and (c) demonstrate the straight pose used for registering the shape+pose and shape+pose+scale models to the US images, respectively.

 extract as the neighboring tissues cause false positives. Finally, the TPs are not visible at all due to the limited sagittal field of view. US images, on the other hand, provide strong bone responses at LAs, TPs and SPs, relatively weaker responses at APs and almost no useful information from the VBs. To account for this variable visibility across different regions in each modality, we propose a weighted joint registration technique. In this method, predefined weights are assigned to each instance of the model depending of the visibility of the different regions in the given modality. The objective function (Eqs. 2.8 and 2.9) are modified by adding the weight matrix to the expectation step. the joint shape derivative becomes:

\[
\frac{\partial Q}{\partial \theta^s} = \sum_{md \in M} \sum_{l=1}^L \sum_{m,n=1}^{M_{md},N_l} W^l_{n,md} \times P(t^n_l|z_{m,md}) \left[ \Phi(t^n_l)^\top - z_{m,md}^\top \right] \frac{\partial \Phi(t^n_l)}{\partial \theta^s} + R^s.
\]

(4.1)

Here, \(W^l_{n,md}\) is a diagonal matrix, where the a \(n\)-th diagonal entry of the \(l\)-th object represents the weights assigned to point \(t^n_l\) of the model for the modality \(md\). These weights are assigned to Each point belonging to a given region (SP, LA, AP, TP and VB) is assigned the weight specific to that region. Intuitively, based on the presumption about the presence and correctness of information at these regions, \(W^l_{n,md}\) determines which regions the objective function will be in favor of for each modality.

Based on the rationale discussed above, weights given in Table 4.2 are
4.3 Results

selected for the US and MR data. For each modality, regions with high visibility are assigned weights of 1. More challenging regions, where false positives are expected, receive weights of 0.5. Zero weights are given to regions that are not at all visible in each image.

<table>
<thead>
<tr>
<th>Region</th>
<th>SP</th>
<th>LA</th>
<th>AP</th>
<th>TP</th>
<th>VB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight in MR</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Weight in US</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4.2: Weights assigned to the different regions of the vertebrae.

The iterative optimization and model transformation is as described in Chapters 2 and 3.

4.3 Results

4.3.1 Registration Parameter Selection

The registration parameters obtained in Chapter 3 were used for the MR+US data.

4.3.2 Validation

Registration results are assessed separately on both the MR and US spaces. Similar to Chapter 3, the results from the joint and weighted joint methods are compared with those of the US-only and MR-only techniques for both versions of the statistical model.

Experiments in MR Domain

In order to study the registration results on the MR domain, the pose parameter of the model is computed from the MR image. The shape is optimized jointly, or only from one of the modalities in different experiments. For each subject, TRE is calculated between the model and the landmarks depicted in the ground truth. Correspondences are assigned by locating the nearest models to each landmark point. For the MR images, the ground truth consists of landmarks selected on SPs, APs, LAs and VBs. More precisely, the following landmarks were segmented: the leading edge of visible SPs on the medial slice, APs and LAs on the most lateral slices which include visible APs and LAs, respectively, and the outline of VBs on the medial slice, as well as on those where LAs are marked. The landmark segmentation was
### 4.3. Results

<table>
<thead>
<tr>
<th>Patient Number</th>
<th>AP</th>
<th>LA</th>
<th>SP</th>
<th>VB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>353</td>
<td>326</td>
<td>265</td>
<td>2,469</td>
</tr>
<tr>
<td>2</td>
<td>227</td>
<td>275</td>
<td>131</td>
<td>1,944</td>
</tr>
<tr>
<td>3</td>
<td>361</td>
<td>270</td>
<td>233</td>
<td>1,587</td>
</tr>
<tr>
<td>4</td>
<td>546</td>
<td>565</td>
<td>413</td>
<td>3,425</td>
</tr>
<tr>
<td>5</td>
<td>390</td>
<td>325</td>
<td>135</td>
<td>2,190</td>
</tr>
<tr>
<td>6</td>
<td>254</td>
<td>277</td>
<td>182</td>
<td>1,192</td>
</tr>
<tr>
<td>7</td>
<td>525</td>
<td>568</td>
<td>280</td>
<td>3,637</td>
</tr>
<tr>
<td>8</td>
<td>331</td>
<td>343</td>
<td>155</td>
<td>2,250</td>
</tr>
<tr>
<td>9</td>
<td>607</td>
<td>395</td>
<td>233</td>
<td>1,345</td>
</tr>
<tr>
<td>All</td>
<td>3,594</td>
<td>3,344</td>
<td>2,027</td>
<td>20,039</td>
</tr>
</tbody>
</table>

Table 4.3: Numbers of MR gold standard points used to report TRE values on MR space for the US+MR dataset. The numbers are provided per region for each subject.

done manually on sagittal slices using MITK’s segmentation tools. Table 4.3 shows a breakdown of number of landmark points used for computing the TRE at each region in MR images.

### Experiments in US Domain

The analysis in the US domain is similar to that of Chapter 3. TRE values are computed for patient-specific models with US poses in the different experiments. The landmarks were selected by a sonographer using the same protocols described in Section 3.3.2. Table 4.4 shows the number of landmark points used on different regions of each US image.
4.4. Discussion and Summary

<table>
<thead>
<tr>
<th>Patient Number</th>
<th>Region</th>
<th>AP</th>
<th>LA</th>
<th>SP</th>
<th>TP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>220</td>
<td>0</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>125</td>
<td>0</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>173</td>
<td>33</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>242</td>
<td>30</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>164</td>
<td>0</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>220</td>
<td>61</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>197</td>
<td>46</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>174</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>270</td>
<td>16</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>47</td>
<td>1,785</td>
<td>186</td>
<td>426</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4: Numbers of US gold standard points used to report TRE values for the US+MR dataset. The numbers are provided per region for each subject. The landmarks were selected by an expert sonographer. Zeros indicate the region was not visible to the sonographer.

The shape+pose model was used for joint registration of MR and US. Figure 4.4 shows examples of the results of the weighted joint and the US-only methods on the two modalities. TRE values obtained in experiments performed are provided in Tables 4.5 and 4.6. U-tests were performed on the error distributions (4.7). No statistically significant improvements were observed. Execution time was approximately 1.1 minutes.

Similarly, results obtained using the shape+pose+scale model are provided below (Tab. 4.7-4.8 and Fig. 4.6). At AP, SP and TP regions p<0.05 were achieved. That is, the shape+pose+scale model combined with the weighted joint method lead to a significant improvement of TRE values. Also, L1 and L3 vertebrae improvements were obtained (p<0.05). An average run-time of 1.9 was achieved.

4.4 Discussion and Summary

In this chapter, the shape+pose and shape+pose+scale models were used in the joint and weighted joint registration methods to augment intraprocedure US and preprocedure MR. The US+MR dataset was shown to be challenging due the low quality and slice thickness, as well as limited fields of view of both modalities. Nevertheless, a consistent and statistically significant
improvement was seen using the joint methods using both models, specially in the US domain. The joint and weighted joint methods resulted in significant improvements in the TRE compared to the US-only method in both two modalities. U-tests revealed that significant improvements were made using the shape+pose+scale model combined with the weighted joint technique. Specifically, it was shown that APs, SPs and TPs TREs dropped using the joint methods. This suggest the joint methods are preferable over US-only ones for facet joint injections, were APs are the most critical regions. Errors at the most critical regions for epidurals, i.e. LAs, show much smaller changes across the different shape optimization methods and model versions. Nevertheless, using the shape+pose+scale, the weighted joint methods still achieve lower TREs compared to US-only. The joint and weighted joint methods were very similar in terms of accuracy for the shape+pose model. However, in the case of shape+pose+scale, the improvements made by the weighted joint method over the joint technique are more pronounced. This can be due to the fact that this model is less constrained in terms of pose deformation and is hence more prone to converge further from the optimal minima in presence of false positives. Overall, the shape+pose+scale model
### 4.4. Discussion and Summary

<table>
<thead>
<tr>
<th>Region Type</th>
<th>MR-only</th>
<th>Weighted Joint</th>
<th>Joint</th>
<th>US-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>2.80 ± 1.42</td>
<td>2.69 ± 1.44</td>
<td>2.70 ± 1.40</td>
<td>2.78 ± 1.54</td>
</tr>
<tr>
<td>LA</td>
<td>2.48 ± 1.37</td>
<td>2.49 ± 1.30</td>
<td>2.48 ± 1.31</td>
<td>2.52 ± 1.32</td>
</tr>
<tr>
<td>SP</td>
<td>4.21 ± 2.54</td>
<td>4.11 ± 2.55</td>
<td>2.24 ± 2.58</td>
<td>4.07 ± 2.44</td>
</tr>
<tr>
<td>VB</td>
<td>2.52 ± 1.38</td>
<td>2.50 ± 1.35</td>
<td>2.54 ± 1.40</td>
<td>2.54 ± 1.37</td>
</tr>
<tr>
<td>All</td>
<td>2.68 ± 1.68</td>
<td>2.75 ± 1.70</td>
<td>2.67 ± 1.64</td>
<td>2.60 ± 1.62</td>
</tr>
</tbody>
</table>

Table 4.5: MR TRE (mm) achieved using the shape+pose model for the US+MR dataset. In each experiment (each column), the shape parameter of the model is optimized from the modalities indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, VB) and the average of all four.

proved to be more successful for this dataset at throughout the different regions of the vertebrae. In general, the weighted joint framework combined with the shape+pose+scale model is more accurate than the shape+pose model at all the regions. With the shape+pose+scale model, the weighted joint method even outperformed the MR-only registration in some cases since the it benefits from the complementary features of both modalities. TRE values obtained for the MR+US chapter were proven to be higher than SDEs and TREs of the US+CT data in Chapter 3. This can be due to from several factors. Firstly, accurate and complete bone surface extraction is more difficult as for the US+MR was generally more challenging. Compared to the US+CT data, this dataset’s preprocedure image (MR) have thicker slices (3.5−4.4 mm). Hence, the best and most accurate obtainable error in the sagittal direction is 1.75−2.2 mm, which is already within the range of errors of the US+CT dataset. Furthermore, due to visibility of different tissues in MR images, facet joints and spinous process are not as clearly visible as in CT scans. This yields to relatively poorer and less accurate bone edge detection in MR, compared to CT. Also, the fields of view of both MR and US in this dataset are limited, with the clinical MR images not including the TPs and US images not capturing the full lumbar spine. Another important source of error, specially in the US space, is the subjects’ pose at image acquisition time. As mentioned before, the US images in this dataset were obtained in the sitting position, which is a more extreme pose compared to the prone position. To address this issue, a less curved pose was used to initialized the model. However, the observed errors at L4 and
4.4. Discussion and Summary

Table 4.6: US TRE (mm) achieved using the shape+pose model for the US+MR dataset. In each experiment (each column), the shape parameter of the model is optimized from the modalities indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, TP) and the average of all four.

<table>
<thead>
<tr>
<th>Region</th>
<th>MR-only</th>
<th>Weighted Joint</th>
<th>Joint</th>
<th>US-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>7.60 ± 4.05</td>
<td>7.92 ± 4.02</td>
<td>7.52 ± 3.63</td>
<td>7.98 ± 3.81</td>
</tr>
<tr>
<td>LA</td>
<td>3.37 ± 2.29</td>
<td>3.46 ± 2.48</td>
<td>3.44 ± 2.45</td>
<td>3.43 ± 2.43</td>
</tr>
<tr>
<td>SP</td>
<td>4.65 ± 1.74</td>
<td>4.39 ± 1.74</td>
<td>4.33 ± 1.60</td>
<td>4.11 ± 1.71</td>
</tr>
<tr>
<td>TP</td>
<td>3.92 ± 2.61</td>
<td>3.91 ± 2.59</td>
<td>3.93 ± 2.65</td>
<td>3.91 ± 2.63</td>
</tr>
<tr>
<td>All</td>
<td>4.25 ± 3.56</td>
<td>4.20 ± 3.64</td>
<td>4.24 ± 3.54</td>
<td>4.20 ± 3.58</td>
</tr>
</tbody>
</table>

L5 (the vertebrae affected most in the sitting position) were higher. This can be due the fact that the current models are based on CT images only, which may not actually contain the kind of pose deformation involved in the sitting position. Also, the initial poses were merely chosen by manually selecting a pose for each shape+pose and shape+pose+scale models that appeared straight, from the possible variations of pose in each model. In order to obtain a more accurate initial pose, an US training dataset is required, where the optimal pose parameters are manually determined. This prior knowledge about the pose of the patient can be integrated in the system to initialize the pose registration for a more accurate and faster convergence.

An important observation made from the data is that the variations of the pose of the sacrum are often quite noticeable in the different poses seen in the current datasets. Hence, we believe that including the sacrum in the statistical model can potentially help achieve a better pose convergence, specially for the sitting cases. Finally, there can potentially be errors in the process of calculating the TREs. In Chapter 3, each vertebra of the patient-specific model was rigidly registered to the corresponding vertebra on the gold standard to eliminate pose errors. This step cannot be done in the case of the MR+US dataset because full segmentation of the vertebra is not feasible. As a result, the reported TRE values include both shape and pose errors. Moreover, not performing a rigid registration on individual vertebrae increases the odds of assigning incorrect correspondences between the model points and points on the gold standard segmentation.
4.4. Discussion and Summary

Figure 4.5: Box plots showing the break-down of TRE values obtained using the shape+pose model for US images (US+MR dataset) at different regions of the vertebrae (a) and the five different vertebrae (b). No significant improvements are observed.

<table>
<thead>
<tr>
<th>Region Type</th>
<th>Reg.</th>
<th>MR-only</th>
<th>Weighted Joint</th>
<th>Joint</th>
<th>US-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>AP</td>
<td>2.73 ± 1.54</td>
<td>2.73 ± 1.50</td>
<td>2.80 ± 1.60</td>
<td>2.80 ± 1.63</td>
</tr>
<tr>
<td>LA</td>
<td>LA</td>
<td>2.52 ± 1.40</td>
<td>2.58 ± 1.61</td>
<td>2.60 ± 1.74</td>
<td>2.65 ± 1.67</td>
</tr>
<tr>
<td>SP</td>
<td>SP</td>
<td>4.10 ± 2.74</td>
<td>4.19 ± 2.85</td>
<td>4.44 ± 2.98</td>
<td>4.22 ± 3.04</td>
</tr>
<tr>
<td>VB</td>
<td>VB</td>
<td>2.45 ± 1.36</td>
<td>2.46 ± 1.39</td>
<td>2.54 ± 1.44</td>
<td>2.51 ± 1.42</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>2.61 ± 1.89</td>
<td>2.62 ± 1.63</td>
<td>2.71 ± 1.72</td>
<td>2.69 ± 1.70</td>
</tr>
</tbody>
</table>

Table 4.7: MR TRE achieved (mm) using the shape+pose+scale model for the US+MR dataset. In each experiment (each column), the shape and scale parameters of the model are optimized using the methods indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, VB) and the average of all four.
4.4. Discussion and Summary

Figure 4.6: Example snapshots of the registration results using the shape+pose+scale model on MR (a) and US (b) images. The weighted joint method (blue) performs better than the model-to-US technique (red). The yellow annotations on US depict the gold standard segmentation obtained by the sonographer.

<table>
<thead>
<tr>
<th>Region Type</th>
<th>MR-only</th>
<th>Weighted Joint</th>
<th>Joint</th>
<th>US-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>6.10 ± 2.18</td>
<td>6.28 ± 2.31</td>
<td>7.13 ± 2.16</td>
<td>8.26 ± 2.53</td>
</tr>
<tr>
<td>LA</td>
<td>3.40 ± 2.08</td>
<td>3.49 ± 2.11</td>
<td>3.61 ± 2.15</td>
<td>3.66 ± 2.32</td>
</tr>
<tr>
<td>SP</td>
<td>7.95 ± 4.89</td>
<td>7.57 ± 5.13</td>
<td>8.18 ± 4.96</td>
<td>7.76 ± 4.98</td>
</tr>
<tr>
<td>TP</td>
<td>5.38 ± 5.20</td>
<td>5.41 ± 5.19</td>
<td>5.54 ± 4.76</td>
<td>5.45 ± 5.05</td>
</tr>
<tr>
<td>All</td>
<td>4.15 ± 3.56</td>
<td>4.20 ± 3.64</td>
<td>4.24 ± 3.54</td>
<td>4.20 ± 4.58</td>
</tr>
</tbody>
</table>

Table 4.8: US TRE achieved (mm) using the shape+pose+scale model for the US+MR dataset. In each experiment (each column), the shape and scale parameters of the model are optimized using the methods indicated on the first row of the table. The errors are reported for different regions of the vertebrae (AP, LA, SP, TP) and the average of all four.
4.4. Discussion and Summary

Figure 4.7: Box plots showing the break-down of TRE values obtained using the shape+pose+scale model for US images (US+MR dataset) at different regions of the vertebrae (a) and the five different vertebrae (b). Significant improvements were observed at AP, SP, TP regions, as well as at L1 and L3.
Chapter 5

Conclusion

In this thesis, a novel solution was presented for augmentation of spinal US used in guiding anesthesia. A point-based shape+pose registration framework was modified and extended to jointly register multiple instances of the statistical model to multiple imaging modalities, with the underlying assumption that they vary in pose and not in shape and size. The proposed framework makes it possible to take advantage of complementary features existing in different modalities and hence, helps achieve a more accurate spine localization. The statistical model was improved by decoupling scale from pose in the atlas, yielding to a shape+pose+scale model. The joint registration technique using the two versions of the model was successfully applied to a US+CT and a US+MR dataset. Registration accuracy at articular processes, the most critical regions in facet joint injections, was significantly improved. As for the laminae, the most important visible bone regions in epidurals, no significant improvement was achieved. Nonetheless, it was demonstrated that the joint method allowed for a simultaneous visualization of the patient-specific models on intraoperative US and the preoperative images. Preoperative images are easier to interpret, hence the simultaneous visualization helps the clinicians assess the local registration accuracy of the model. Seeing the model on the preoperative images can help the clinician decide whether or not to base guidance decision on the registration results. This framework has the potential to improve the performance of delivering anesthesia by adding atlas overlays, eliminating the need to preprocedurely download or process data.

5.1 Contributions

The contributions made in this thesis are summarized below:

1. A novel joint framework is proposed for simultaneous registration of multiple available imaging modalities including intraoperative US, preoperative CT and MR.
5.2. Future Work

2. A modified version of the registration cost function is provided, which accounts for one common shape (and scale) and multiple poses for multiple imaging modalities.

3. The derivatives for fast optimization of shape and pose coefficients in the joint framework are derived and presented.

4. A weighted version of the joint framework was proposed to allow weights to be assigned to each imaging modality in the objective function. The weights of the different regions in each modality are chosen according to their presumed visibility. (They can be adjusted by the user according to the application, resolution of each image, number of visible vertebrae, and etc.)

5. A new shape+pose+scale statistical model of the lumbar spine has been developed. This was done by decoupling the scaling from the pose to allow the pose and scale to be optimized separately and independently. The separate parameter optimization lead to a more accurate fitting of pose to data.

6. The proposed method combined with both shape+pose and shape+pose+scale models has been validated on a US+CT dataset. The effectiveness of the method was demonstrated by carrying out several experiments for each dataset and comparing their registration errors at individual regions of the vertebrae.

7. The method has also been successfully applied to a challenging data set of low quality US and clinical MR images with limited field of view. To the best of our knowledge, previously no work of art had attempted US+MR fusion for the lumbar spine using a deformable registration approach with statistical models.

8. The presented joint registration method has been compared to a previous work that only register a statistical model to US.

9. The pose convergence has been improved in terms of speed and accuracy by initializing the registration according to prior knowledge of the pose in intraoperative US.

5.2 Future Work

Provided below are several possible paths to be taken for further research after this thesis.
5.2. Future Work

1. Optimal initial poses can be obtained for US images. In this thesis, for registering the model to the US+MR pair, we used an initial pose, in which according to the statistical model, the spine appeared straight. A statistical analysis on an US data set can be performed to find a better initial pose. Using such initial poses for the US helps reduce the risk of converging to local minima, and helps speed up the optimization.

2. The computation time can be improved by developing target point cloud extraction techniques that can be applied on 3D volumes, as opposed to individual 2D slices. Currently, these preprocessing steps very computationally expensive, specially for US images. This is because the bone enhancement algorithm is applied on individual 2D slices, one at a time. The computation time can also be improved through parallel processing on graphics processing units.

3. Increasing the size of test datasets will give more insight on the performance and robustness of the proposed method.

4. Increasing the size of the training dataset can also help increase the range of variability that the model is able to capture and hence, improve the statistical model.

5. Constructing a statistical model that includes the sacrum as well as the lumbar vertebrae can potentially help improve the pose registration for the lumbar spine.

6. A two-step registration approach can be used for improving the results. Here, the registered five-vertebra model can be used as an initialization for local registration of smaller models to individual vertebrae.

7. The proposed joint framework could be used for joint registration to both CT and MR scans together with the statistical model and US. We believe this can improve the final registration accuracy which, in turn, could further increase the confidence of the physician in using these patient-specific models for making guidance decisions in spine anesthesia.

8. A user study can be conducted, where sonographers will be asked to rate and comment on the effectiveness and viability of the presented framework for spinal ultrasound augmentation. This will be a helpful and essential step for the proposed method to gain clinical acceptance.
5.2. Future Work

9. US augmentation for scoliotic patients is additionally challenging due to abnormal curvatures in their spines. This method, which benefits from complementary features of multiple imaging modalities, can be applied on scoliotic patients.

10. Finally, the proposed framework can be thought of as a generic solution for the problem of registration of a statistical atlas to rigid multi-body anatomies, which may vary only in pose in different imaging modalities. Hence, we envision that such a multi-data fusion could be further used for applications such as wrist, femur, tibia, cervix, etc.
Bibliography


Bibliography


[24] Abtin Rasoulian, Robert Rohling, and Purang Abolmaesumi. Lumbar spine segmentation using a statistical multi-vertebrae anatom-


SPIE Medical Imaging, pages 90360P–90360P. International Society for Optics and Photonics, 2014.


