# Model-Based Segmentation and Registration of Multimodal Medical Images

**ZHANG JING**

*(B.Eng. Tsinghua University, M.Eng. NUS)*

**A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY**

DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2009

# Acknowledgements

First and foremost, I wish to express my sincere appreciation to my supervisors, A/Prof Ong Sim Heng and Dr. Yan Chye Hwang. Their motivation, guidance and instruction are deeply appreciated. I would like to thank them for giving me the opportunity to pursue my interest in research.

I also want to thank Dr. Chui Chee Kong for his patience, encouragement and tremendous help offered. I am grateful for his trust and belief in me.

I am thankful to Prof. Teoh Swee Hin for his advice and assistance.

I thank all the staff and research scholars of Biosignal Lab who have been a terrific bunch to work with. These individuals are: Wang Zhen Lan, Jeremy Teo and Lei Yang.

Last but not least, I would like to thank my family for all their support and encouragement.

# Table of Contents

# Summary

Registration helps the surgeon to help overcome the limitation of relying on a single modality for image-guided surgery. There is a need for an accurate registration system which will improve surgical outcomes. The work described has involved the investigation and development of a new registration system based on computational model. Preoperative CT images of patient are segmented using an adaptive thresholding method, which takes into consideration the inhomogeneity of bone structure. A patient-specific surface model is then constructed and used in the registration process.

We proposed and developed a new automatic surface-based rigid registration system using neural network techniques for CT/CT and CT/MRI registration. A multilayer perceptron (MLP) neural network is used to construct the bone surface model. A surface representation function has been derived from the resultant neural network model, and then adopted for intra-operative registration. An optimization process is used to search for optimal transformation parameters together with the neural network model. In CT/CT registration, since no point correspondence is required in our neural network (NN) based model, the intra-operative registration process is significantly faster than standard techniques.

We proposed a weighted registration method for CT/MRI registration, which can solve the CT/MR registration problem and MR image segmentation problem

simultaneously. This approach enables fast and accurate CT/MR feature based registration, accurate extraction of bone surface from MR images, and fast fusion of the two different modalities. Since the bone surface in CT images can be extracted quickly and accurately, the CT segmentation result is used as the reference for MR image segmentation. The process starts with a coarse extraction of bone surface from MR images, and the coarse surface is then registered to the accurate bone surface extracted from CT images. The CT bone surface is re-sampled according to the registration result. It is used as the initial estimation for MR image segmentation. The MR segmentation result is subsequently registered to CT bone surface. The segmentation result of MR images is improved at each iterative step using the CT segmentation result. In the iterative segmentation-registration process, since the goal boundary is close to the initial one, only fine adjustment is needed. Computational time is hence saved and unreasonable segmentation due to poor scans can be effectively avoided.

We also investigated the application of statistical methods to assist CT/CT and CT/MR registrations. CT/CT and CT/MRI registration methods were integrated into a generic software toolkit. The toolkit has been used in segmentation of various human and animal images. It has also been applied to register human bone structures for image-guided surgery. The successful completion of the weighted registration method greatly enhances the state-of-art for CT/MRI registration.

# List of Tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| 2D | two dimensions |
| 3D | three dimensions |
| BMD | Bone Mineral Density |
| CT | computed tomography |
| DFLS | double-front level set |
| DFT | discrete Fourier transform |
| DXA | Dual-Energy X-Ray Absorptiometry |
| FE | finite element |
| FPGA | field-programmable gate array |
| GD-DTPA | gadolinium- Diethylene triamine pentaacetic acid |
| HD | Hausdorff distance |
| ICP | iterative closest point |
| MI | mutual information |
| MLP | multilayer perceptron |
| MR | Magnetic resonance |
| MRI | Magnetic resonance imaging |
| NMI | normalized mutual information |
| NMR | nuclear magnetic resonance |
| NMRI | nuclear magnetic resonance imaging |
| NN | neural network |

| | |
|---|---|
| PCA | principal component analysis |
| QCT | quantitative computed tomography |
| QUS | quantitative ultrasound |
| RF | radio-frequency |
| SNR | signal to noise ratio |
| VHD | Visible Human Dataset |
| VOI | volume of interest |

# List of Notations and Variables

| | |
|---|---|
| $\mu$ | mean |
| $\sigma$ | standard deviation |
| $B$ | bone |
| $\overline{B}$ | non-bone |
| $x$ | pixel |
| $I(x)$ | gray level of pixel $x$ |
| $E_B$ | the set of boundary pixels in $B$ |
| $W(x)$ | the window centered on pixel $x$ |
| $r$ | radius |
| $\mathbf{P}$ | the surface point clouds of a dataset |
| $\mathbf{Q}$ | the surface point clouds of a dataset |
| $\mathbf{p}$ | a point of dataset $\mathbf{P}$ |
| $\mathbf{q}$ | a point of dataset $\mathbf{Q}$ |
| $\overline{\mathbf{p}}$ | centroid of dataset $\mathbf{P}$ |
| $\overline{\mathbf{q}}$ | centroid of dataset $\mathbf{Q}$ |
| $\overline{\mathbf{P}}$ | a $3 \times N$ matrices $\overline{\mathbf{P}} = (\overbrace{\overline{\mathbf{p}} \quad \overline{\mathbf{p}} \quad \cdots \quad \overline{\mathbf{p}}}^{N})$ |
| $\overline{\mathbf{Q}}$ | a $3 \times N$ matrices $\overline{\mathbf{Q}} = (\overbrace{\overline{\mathbf{q}} \quad \overline{\mathbf{q}} \quad \cdots \quad \overline{\mathbf{q}}}^{N})$ |
| $\tilde{\mathbf{P}}$ | $\tilde{\mathbf{P}} = \mathbf{P} - \overline{\mathbf{P}}$ |
| $\tilde{\mathbf{Q}}$ | $\tilde{\mathbf{Q}} = \mathbf{Q} - \overline{\mathbf{Q}}$ |
| $D$ | the distance between any two points |

| | |
|---|---|
| $\mathbf{R}$ | a $3 \times 3$ rotation matrix |
| $\mathbf{t}$ | a 3D translation vector |
| $\mathbf{T}$ | a $3 \times N$ matrices $\mathbf{T} = (\overbrace{\mathbf{t} \quad \mathbf{t} \quad \cdots \quad \mathbf{t}}^{N})$ |
| $\mathbf{U}$ | Eigen matrix |
| $\mathbf{V}$ | Eigen matrix |
| $\mathbf{\Lambda}$ | Eigen matrix |
| $C(\mathbf{T})$ | the cost of a transformation function |
| $\mathbf{D}_r$ | the reference surfaces |
| $\mathbf{D}_c$ | the current surfaces |
| $f_r$ | reference surface function |
| $(x, y, z)$ | coordinates of a point |
| $d(x, y, z)$ | signed distance function |
| $\varphi(\cdot)$ | activation function in neural network hidden layers |
| $\mathbf{b}_1$ | the biases for the first hidden layer |
| $\mathbf{b}_2$ | the biases for the second hidden layer |
| $b_3$ | the biases for the output layer |
| $\varepsilon_R$ | the largest absolute eigenvalue of $(\hat{\mathbf{R}} - \mathbf{R})$ |
| $\varepsilon_T$ | the largest absolute eigenvalue of $\left\lVert \hat{\mathbf{T}} - \mathbf{T} \right\rVert_2$ |
| $\varphi_B$ | the back-propagating level set |
| $\varphi_F$ | the forward-propagating level set |
| $\mathbf{v}$ | velocity field |
| $\tau$ | threshold of gradient value |

| | |
|---|---|
| $\delta$ | pre-defined small constant |
| $S$ | DICE similarity coefficient |
| $\mathbf{m}$ | shape vector |
| $\boldsymbol{\omega}$ | covariance matrix |
| $\lambda_k$ | eigenvalue |
| $\boldsymbol{\alpha}$ | shape parameter |

# 1 Introduction

## 1.1 Motivation

Many surgical procedures require highly precise localization, often of deeply buried structures, in order for the surgeon to extract the targeted tissue with minimal damage to nearby structures. Image-guided surgery is a solution to address this clinical need. Segmentation and registration are important sub-tasks in image-guided surgery. The region of interest is extracted in segmentation. Registration is the process used to match the coordinate system of preoperative imagery with that of the actual patient on the operating table. After registration, possible image-based applications include interactive pre-operative viewing, determination of the incision line and navigation during surgery.

Traditional clinical practice utilizes only 2D magnetic resonance (MR) or computed tomography (CT) slices, and the surgeon must mentally construct the 3D object and compare the critical image information to the body of the patient. CT provides well-contrasted images of high-density biological objects such as bones and tumors but is usually not preferred for detailed soft tissue examination. MR imaging, with its moderate resolution and good signal-to-noise ratio is the modality of choice for soft tissues. Fusing CT and MR images will help overcome the limitation of relying on a single modality for image guided surgery. A typical fusion procedure comprises segmentation of the CT and MR images, followed by registration and spatial alignment/fusion. The region of interest in CT images (e.g., bone) or MR images

(e.g., kidney and liver) of a patient is first segmented. After spatial registration, the segmented CT and MR images are aligned to give a model comprising well-contrasted bone structure and the surrounding soft tissues. Such a composite model is important for surgical planning and education. For example, a vertebra, which is hard tissue, may have to be examined with the intervertebral disc, a soft tissue, for effective spinal surgery planning.

The objective of this work was the development of a system to produce a patient-specific hybrid model of the spine for image guided spinal surgery. The system should comprise CT/MR image segmentation, CT/CT and CT/MR image registration. It may also be employed for different anatomies, e.g., the ankle.

## 1.2 Background

### 1.2.1 CT and MRI

**Quantitative Computed Tomography**

In CT imaging, the two-dimensional internal structure of an object can be reconstructed from a series of one-dimensional projections of the object acquired at different angles as outlined in Figure 1.1.

The scanning for angles ranging from 0 ° to 360 ° is repeated so that sufficient data is collected to reconstruct the image with high spatial resolution. The reconstructed image is displayed as a two-dimensional matrix, with each pixel representing the CT number of the tissue at that spatial location. As the CT number and the attenuation

coefficient of a voxel related to the bone is a near-linear function of the bone density, CT imaging can be used to provide in-vivo quantitative analysis of bone density.

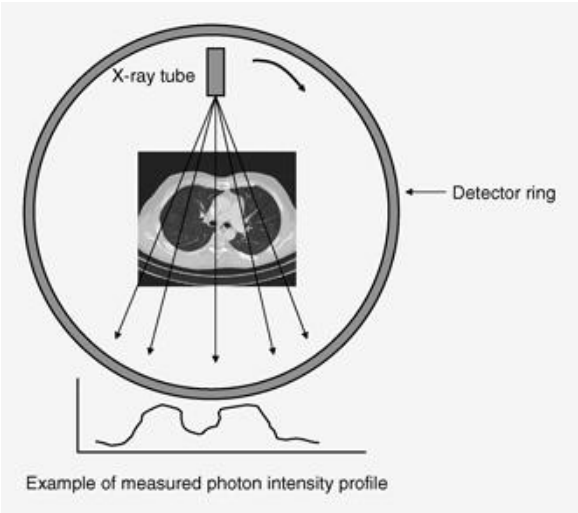



Figure 1.1. Principles of computed tomography image generation (adapted from [1]).

**Magnetic Resonance Imaging**

Magnetic resonance imaging (MRI) is an imaging technique used primarily in medical settings to produce high quality images of the inside of the human body. MRI is based on the principles of nuclear magnetic resonance, a spectroscopic technique used by scientists to obtain microscopic chemical and physical information about molecules. The technique was called magnetic resonance imaging rather than nuclear magnetic resonance imaging (NMRI) because of the negative connotations associated with the word nuclear in the late 1970's.

In MR imaging, in order to selectively image different voxels (volume picture elements) of the subject, orthogonal magnetic gradients are applied. Although it is relatively common to apply gradients in the principal axes of a patient (so that the patient is imaged in $x$, $y$ and $z$ from head to toe), MRI allows completely flexible orientations for images. All spatial encoding is obtained by applying

magnetic field gradients which encode position within the phase of the signal. In one dimension, a linear phase with respect to position can be obtained by collecting data in the presence of a magnetic field gradient. In three dimensions (3D), a plane can be defined by "slice selection", in which an RF pulse of defined bandwidth is applied in the presence of a magnetic field gradient in order to reduce spatial encoding to two dimensions (2D). Spatial encoding can then be applied in 2D after slice selection, or in 3D without slice selection. Spatially-encoded phases are recorded in a 2D or 3D matrix; this data represents the spatial frequencies of the image object. Images can be created from the matrix using the discrete Fourier transform (DFT). Typical medical resolution is about 1 $mm^3$, while research models can exceed 1 $\mu m^3$. The three systems described above form the major components of an MRI scanner (Figure 1.2): a static magnetic field, an RF transmitter and receiver, and three orthogonal, controllable magnetic gradients.

The MR method has been one of the most powerful tools in medical field as well as in biological studies since the middle of last century. Magnetic resonance imaging is attractive in that not only high-density objects (e.g. bones), but also the soft tissues (e.g. brain, kidney) can be imaged with fair resolution and good signal to noise ratio (SNR) [2]. More encouraging is the fact that magnetic resonance can be applied to the live body safely in spite of the relatively high magnetic field.

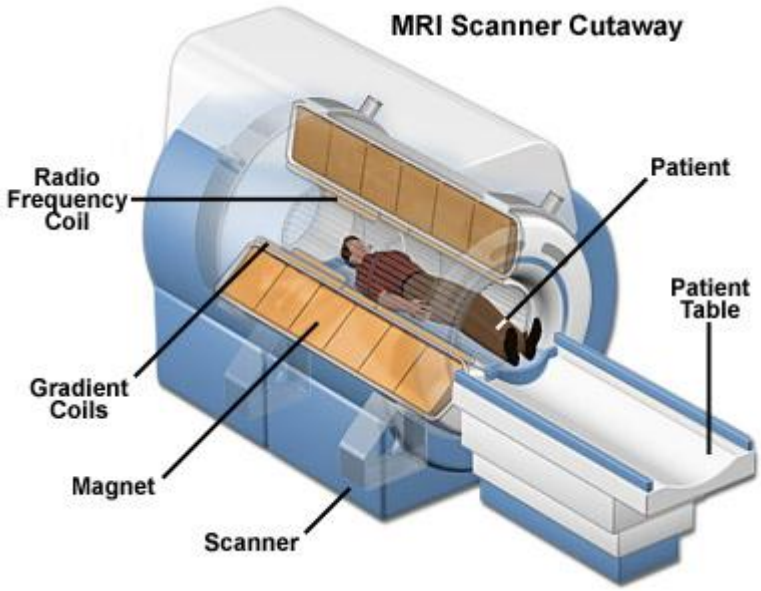



Figure 1.2. MRI scanner (adapted from [3]).

### 1.2.2 Image-guided Therapies for Vertebral Disease

In spinal surgery, it would be helpful for the surgeons to have a panoramic view of the vertebrae, the soft tissue, neural roots, and vessels around it. More care has to be taken in pre-surgery planning to reduce the possibility of damage during the actual operation. Thus there is a need to perform both CT and MRI scans on the patient. Due to the nature of CT and MRI, they provide advantages over each other under different circumstances. CT can give us well-contrasted images of high-density objects such as bones and tumors. However, it works poorly if we intend to examine soft tissue. MR images have the advantage under such circumstances in that both soft tissue and bones are visible, though the resolution and contrast is not as good as that of CT images. Thus these two modalities complement each other. After spatial registration, the results can be used to construct a model comprising clear bone structure and the surrounding soft tissues. This information can be used to plan the surgical procedure by the surgeon. It can also be used for education or training.

## 1.3 Proposed Medical Image Processing System

The proposed and developed system comprises CT/MR image segmentation, CT/CT and CT/MR image registration. As shown in Figure 1.3, segmentation is first performed on CT images to separate the region of interest (bone) from its surroundings. The bone surface is then used to construct the bone surface model using a MLP neural network. An initial MR image segmentation captures the general shape of the target object (the vertebrae). A coarse registration result is obtained by registering the MR and CT surfaces with a weighted surface-based registration algorithm. With the registered CT surface model as the reference, we use the intermediate results of MR image segmentation and registration to iteratively refine the suboptimal MR image segmentation. This iterative process is carried out until the segmented CT and MR surfaces match within a specified tolerance. The registered MR and CT dataset can be fused after this iterative process.

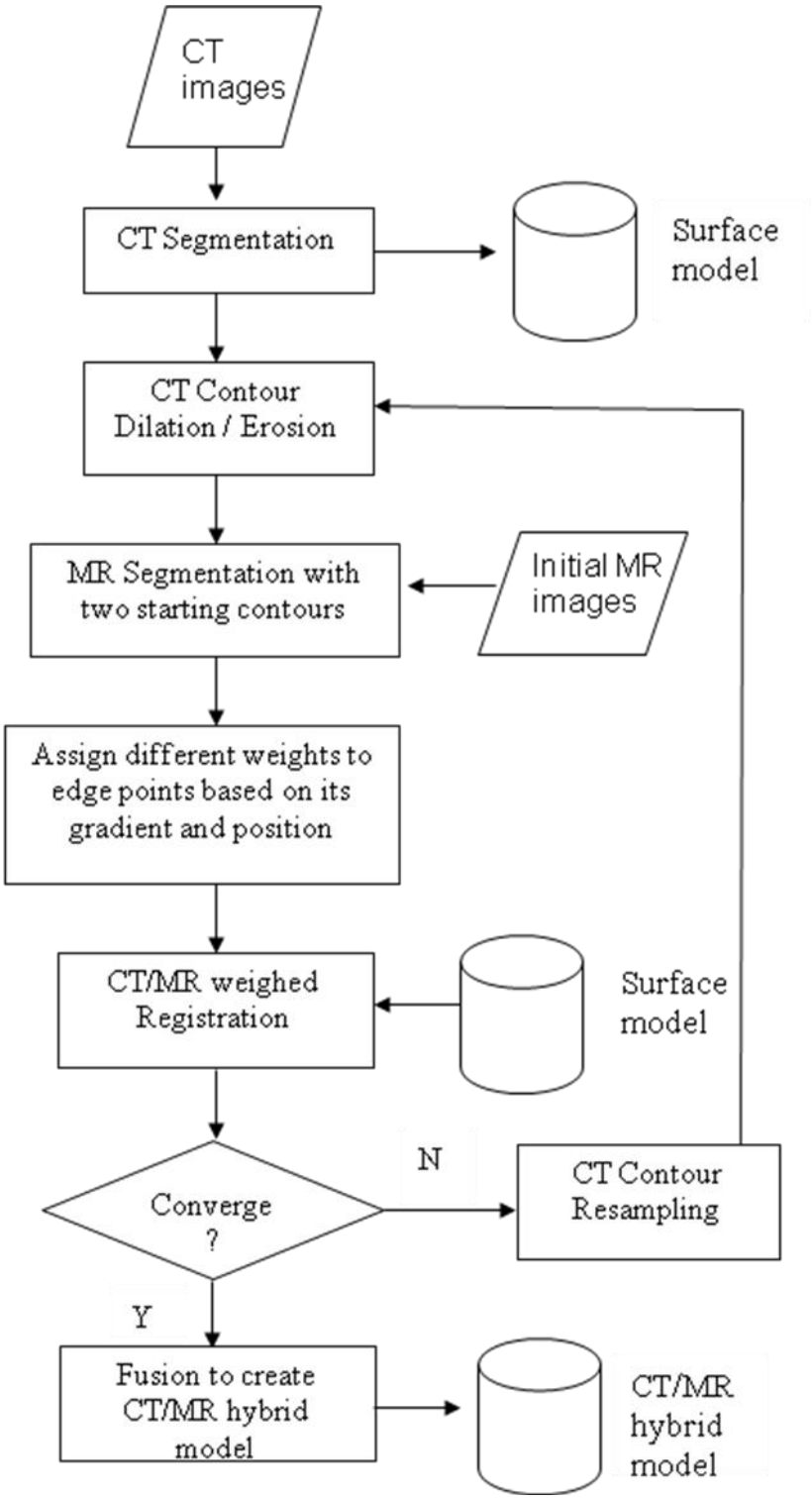



Figure 1.3. Flowchart of feedback segmentation-registration system.

## 1.4 Thesis Contributions

### 1.4.1 3D Adaptive Thresholding Segmentation

A novel 3D adaptive thresholding segmentation method is proposed for 3D CT image segmentation. This fast and accurate method successfully segments the two kinds of bone structures (vertebrae and ankle) in our experiments. In 3D adaptive thresholding method, the thresholding of each voxel is updated up-to-date. For each voxel, a local window, which is a cylindrical region, is defined. The respective means and variances for bone and non-bone inside the corresponding region and similarly are calculated and used to classify all the voxels. The entire volumetric image is processed in an iterative process till it converges.

### 1.4.2 3D CT/CT Surface-based Registration

A novel automatic surface-based method using a neural network is used to perform the registration. The neural network is used to construct an invariant descriptor for human bone to speed up the registration process. Execution time and registration accuracy are the two important specifications for a registration system. The NN-based approach significantly improved computational.

### 1.4.3 MR Image Segmentation and CT/MR Image Registration

A new iterative methodology is proposed to perform fast and accurate multimodal CT/MR registration and segmentation of MR dataset in a concurrent manner. In MR image segmentation, we extend the ordinary single-front level set to the double-front level set. This effectively reduces computational time by limiting the search area around the target and enhances segmentation accuracy by avoiding leakage and

distraction by other objects. The iterative segmentation/registration method helps to refine the segmentation of MR images and the registration of MR to CT. The technique is fully automatic but still able to give results that are comparable to manual segmentation.

### 1.4.4 Statistical Modeling of Vertebrae

A statistical model-based framework is proposed to rapidly create FE meshes with patient-specific geometry using the CT images. These models can be used to create a human spine FE meshes especially lumbar FE meshes. A center firing searching method is implemented to find the correspondence control points for training the statistical shape model. This method has two advantages over conventional template-based mesh-generation methods. Firstly, a high mapping quality is ensured. A proper vertebral template is selected using statistical analysis of a pre-trained database instead of using a single template, which reduces the possibility of mapping error for a complex structure such as vertebra. Secondly, minimum preprocessing, e.g., pre-adjustment, is required.

## 1.5 Thesis Organization

This thesis brings together a 3D adaptive thresholding segmentation method in Chapter 3, CT/CT surface-based registration in Chapter 4, weighted CT/MR registration in Chapter 5 and statistical modeling of vertebrae in Chapter 6. These methodologies enable us to produce hybrid CT/MR model and the possible extension to spine structure.

In Chapter 2, the current image segmentation, registration and image-guided surgery are reviewed.

In Chapter 3, the 3D adaptive thresholding segmentation method is described in detail. The implementation of this method is presented. The experimental results are presented.

In Chapter 4, the surface-based registration method using neural network is presented. The coarse registration based upon principal-axes alignment method is described. Bone surface is modeled using MLP for registration. It is used to create a computationally efficient function for the cost calculation. This registration method achieves sub-voxel accuracy comparable to that of conventional techniques, and is significantly faster. These advantages are demonstrated using image datasets of the calcaneus and vertebrae.

In Chapter 5, a system that performs CT/MR rigid registration and MR image segmentation is presented. The segmentation/registration process progressively refines the result of MR image segmentation and CT/MR registration. For MR image segmentation, we propose a method based on the double-front level set that avoids boundary leakages. In order to reduce the registration error from the misclassification of the soft tissue surrounding the bone in MR images, we propose a weighted surface-based CT/MR registration scheme. The registration method achieves accuracy compatible to conventional techniques while being significantly faster. Experimental results demonstrate the advantages of the proposed approach and its application to different anatomies.

In Chapter 6, a study is proposed on statistical model-based framework to rapidly create FE meshes with patient-specific geometry. A center firing searching method was implemented to find the corresponding control points for training statistical shape model. The proposed framework can be used to generate FE models of complex geometrical structure such as human vertebrae from medical images.

Finally, the conclusion and recommendations for future work in this area of research are presented in Chapter 7.

# 2 Literature Review

## 2.1 Image-guided Surgery

Image processing is an important component of image guided surgery. Medical image analysis brings a revolution to the medicine of the 21st century. It introduces a set of powerful new tools designed to better assist the clinical diagnosis and to model, simulate, and guide more efficiently the patient's therapy. Image-guided surgery also requires input from other traditional disciplines like computer vision, computer graphics, artificial intelligence and robotics.

### 2.1.1 Simulation and Planning

A surgical plan in reconstructive surgery needs information of the shape, symmetry, dimension, and function of hard and soft tissue. At present, surgical plans and surgical outcomes are analyzed on 2D and 3D radiographs and photographs. Much of the challenge in image-guided surgery lies in understanding the relative spatial positions of critical vascular, neural and other structures in relation to the underlying bone and the facial surface. The recent developments in imaging techniques have allowed more effective pre-surgical diagnosis and surgical planning using patient-specific data.

Recently, much research emphasis has also been placed on computer-assisted surgical planning and augmentation systems. Scharver *et al*. [4] have developed an augmented reality system for craniofacial implant. A training system for simulating

temporal bone surgery was proposed by Agus *et al*. [5]. The system is based on patient-specific volumetric object models derived from 3D CT and MR imaging data. Real-time feedback is provided to the trainees via real-time volume rendering and haptic feedback. The performance constraints dictated by the human perceptual system are met by exploiting parallelism via a decoupled simulation approach on a multi-processor PC platform. Meehan [6] presented a system for 3D planning and pre-operative rehearsal of mandibular distraction osteogenesis procedures. Two primary architectural components are described: a planning system that allows geometric bone manipulation to rapidly explore various modifications and configurations, and a visuohaptic simulator that allows both general-purpose training and preoperative, patient-specific procedure rehearsal.

Jolez [7] proposed a method which clearly enhances the ability of the neurosurgeon to navigate the surgical field with greater accuracy, to avoid critical anatomic structures with greater efficacy, and to reduce the overall invasiveness of the surgery itself. Fischer [8] developed a 2D augmented reality image overlay device to guide needle insertion procedures. This approach makes diagnostic high-field magnets available for interventions without a complex and expensive engineering entourage. In preclinical trials, needle insertions have been performed in the joints of porcine and human cadavers using MR image overlay guidance; in all cases, insertions successfully reached the joint space on the first attempt. There are also some studies using robotic devices to aid surgery like needle placement or insertion [9, 10].

### 2.1.2 Validation

The validation process in the context of image-guided surgery is diverse and complex. Image-guided surgery systems involve many processing components, e.g., segmentation, registration, visualization, and calibration. Each component is a potential source of errors. Therefore, validation should involve the study of the performance and validity of the overall system, the performance and validity of the individual components, and error-propagation along the overall workflow. Clinical validation of image guided surgery systems (in terms of large-scale multi-site randomized clinical trials) is difficult, since image guided surgery is a recent technology and the required randomization is an ethical problem.

Validation is usually performed by comparing the results of a method or system with a reference that is assumed to be very close or equal to the exact solution. The main stages of reference-based validation are as follows. The first step is to clearly identify the clinical context and specify the validation objective. Then, the validation criteria to be studied and corresponding objective should be chosen, along with the associated validation metrics that quantify validation criteria. Validation data sets are chosen to provide an access to the reference. The method of computing the reference should be specified, as well as the format of the input and output of comparison between the reference and the results of the method applied to the validation data sets. The validation metric used for comparison is chosen according to its suitability for assessing the clinical validation objective. Quality indices are computed on the comparison output to characterize the properties of the error distribution. Finally, statistical tests are used to assess the validation objective.

A meta-analysis was conducted by Altedorneburg [11] out of clinical trials published between 1987 and 2001 in respect of the clinical pharmacology and safety as well as the diagnostic efficacy of gadolinium - Diethylene triamine pentaacetic acid (Gd-DTPA) for direct intra-articular injection before MRI examination. Binkert [12] compared the examination time with radiologist time and to measure radiation dose of CT fluoroscopy, conventional CT, and conventional fluoroscopy as guiding modalities for shoulder CT arthrography. Thakar [13] established their method validating the algorithm in an independent cohort of patients and black patients and compared two different definitions of renal outcome.

## 2.2 Medical Image Segmentation

There are several established methods for CT image segmentation [14] but a robust, fast and general solution is lacking for MR images. The main difficulties are:

(1) Intrinsic limitations of image acquisition theory and system [15].

The spatial inhomogeneities in the radio-frequency (RF) gain lead to the overlapping of the intensities of two tissues, and thus blurred boundaries. On the other hand, the image acquisition system's failure to provide sufficient spatial resolution will add to the boundary fuzziness.

(2) Variability of object structure/shape/size/texture.

Various shapes and sizes of tissues, complicated topology and different tissue texture make it almost impossible to find universal criteria.

(3) Subject variability due to the operator.

This is due to the parameter settings in scanning and personal criteria of defining boundaries.

(4) Artifacts and noise [16].

Noise and artifacts are introduced in the process of image acquisition. These may be due to the system, hardware, physics or even the patient himself/herself.

All the variability and uncertainty contribute to the tremendous complications in medical image segmentation. Thus application-driven solutions are developed for a range of cases or even for some special cases. Most techniques are either region-based or surface-based, and can be further divided according to the information that is used and the classification method, e.g., intensity [15], morphology [17], probability [18, 19], clustering [20] and neural networks [21]. Surface-based techniques can be classified as parameter-based or geometry-based. There are also approaches that combine different techniques, within or across the classes.

### 2.2.1 Region-based techniques

Thresholding-based techniques are the most straightforward methods [19]. With a threshold value which is set manually or automatically, a point can be classified as object or background depending on its gray value. For example, in most MR images of the vertebrae, the intensity of the vertebral body is similar to the soft tissue and different from that of the spinal processes. Thresholding would thus classify the vertebral body and soft tissue into the same class and classifies the processes as another class. Nevertheless, it is highly subjective to set thresholding manually and

it is weak in error prevention. Much research has been conducted using adaptive thresholding.

Morphology-based techniques [17] always include the following operations: convolution, binarization/thresholding, classification/labeling, morphological operation (dilation/erosion/opening/closing), connected components analysis/region filling, logical operation (AND, OR, NOT, XOR, etc.). The system often has the following problems: (1) convolution with various structuring elements sometimes leads to the loss of details, (2) much manual interaction is often needed, and (3) it is sensitive to noise.

Probability-based techniques classify pixels according to the probability values or maximization of the expectation [18, 19]. Different constraints can be integrated to make the system more robust. However it still has difficulty in overlapped areas and thus misclassification may happen.

Clustering-based techniques are iterative processes of re-assigning pixels to different classes according to some fuzzy membership functions [20]. Clusters need to be carefully selected as they have crucial effect to the performance. The results also heavily depend on manual setting of parameters, which is highly subjective. The vulnerability to noise and high computational requirements are also considered to be shortcomings of clustering-based techniques.

Neural network-based techniques use training datasets to train a neural network for segmentation purposes [21]. However they are not adaptive - small changes in objects lead to re-training of the neural network, which is usually very time consuming. Therefore it is difficult to meet real-time requirements.

### 2.2.2 Surface-based techniques

Parameter-based techniques are derived from the original 2-D deformable model - snakes [9]. The idea of parameter-based deformable model is to locate the active contour to a position that minimizes its energy, external and internal. External energy is represented by image properties, while the snake itself decides on the internal energy. The details of the algorithm will be discussed in later chapters. However the active contour has intrinsic defects in that it has difficulty in tracing convoluted shapes, shapes that are not convex, sharp corners and bends. Snakes are also easy to be caught in local minima and are highly sensitive to noise.

Geometry-based techniques refer to Sethian's level set function [22, 23] and its variations. The level set is a time evolving function, and the so called "zero level curve" corresponding to a propagating front. The details of this algorithm will be discussed in later chapters. The level set method can deal with convoluted shapes, sharp corners or bends. Yet it also has some weaknesses. It is not good at growing bi-directionally, i.e., when the expanding front exits the goal boundary, it may not be able to "shrink" back. Furthermore, it is prone to leak into the background at a fuzzy boundary.

## 2.3 Medical Image Registration

Various medical image registration methods have been proposed for current medical applications with regards to the dimensionality, subject, object and modalities involved. The method may be automatic, interactive and semi-automatic, but they

can all be classified based on the basis of registration, nature and domain of transformation and optimization procedure according to [24].

The basis of medical image registration methods can be either image-based or non-image based. Non-image based methods are seldom used because they use calibration to directly align two coordinate systems, thus requiring the patient to remain motionless between both acquisitions. Most existing methods are image-based and they can be further divided to either extrinsic or intrinsic methods.

Extrinsic methods rely on artificial objects attached to the patient, which are designed to be visible and accurately detectable in all of the pertinent modalities, while intrinsic methods rely on patient generated image content only. Though extrinsic methods can make the registration comparatively easy, fast and usually automated, there is a need for intrinsic methods because of their noninvasive characteristic and improvement in patient comfort.

Intrinsic registration methods can be further divided into the following three categories based on their choice of feature: (1) landmark-based registration, land markers are used to obtain accurate registration result; (2) voxel property-based registration, no segmentation is needed before registration and usually it takes longer time in registration process; (3) Feature-based registration, segmentation is needed before registration.

### 2.3.1 Landmark-based Registration

This approach requires the segmentation procedure to identify points at the locus of the optimum of some geometric property [25, 26] or anatomical landmarks [27, 28].

By constraining the search space according to anatomical landmarks, mismatches are unlikely to occur, and the search procedure can be sped up significantly. However, due to the difficulties in computer recognition of landmarks, this kind of registration usually requires user-interaction.

### 2.3.2 Voxel Property-based Registration

This method uses image intensity for registration. There are two common approaches in this area. One approach attempts to reduce the image gray value content to representative scalars and orientations [29, 30], while the other uses the full image content throughout [31, 32].

### 2.3.3 Registration Based on Image Segmentation

This method needs to first extract anatomically the same structures (mostly surfaces) from the images to be registered. These structures are the sole input for the alignment procedure. Surface-based registration is commonly used for the following reasons: (1) it is less computationally intensive compared to volume-based registration since there are fewer data points; (2) it can be used to perform multimodality registration provided the surfaces can be accurately extracted from different image modalities, which is typically not easy; and (3) the surface is relatively invariant over time, which is useful, for example, in monitoring progression of bone disease. Popular methods of rigid model-based approaches are the “head-hat” method [33] and the fast chamfer matching technique [34]. Since rigid model based methods are always easy to perform and the computational complexity is relatively low, they are used extensively in the clinical field. With deformable models, however, a template model that is defined in one image is

required. The template may be deformed to match the segmented structure in the second image [35, 36] or the second image may be used unsegmented [37, 38, 39]. Deformable curves appear in the literature as snakes, active contours or nets. Deformable model based methods are best suited to find local curved transformations between images, and less so for finding (global) rigid or affine transformations. A drawback of the segmentation- based method is that the registration accuracy is limited to the accuracy of the segmentation step. The registration step is commonly performed automatically while the segmentation step is performed semi-automatically most of the time.

The transformation to be employed defines the nature of relationships between the coordinates of each point in one image (which is called the original image) and coordinates of the corresponding point in the other image (the reference image). It also decides the parameters to be found in the registration procedure. The nature of transformation can be rigid, affine, projective or elastic [24]. Only translations and rotations are allowed in rigid transformation. If the transformation maps parallel lines onto parallel lines, it is called affine. If it maps lines onto lines, it is called projective. Finally, if it maps lines onto curves, it is called curved or elastic. Figure 2.1 illustrates different 2D transformations.

The domain of the transformation is called global if it applies to the entire image, and local if regions of the image each have their own transformations defined. Local transformations are seldom used directly; the term is reserved for transformations that are composites of at least two transformations determined on sub-images that cannot be generally described as a global transformation. The most frequently used

transformation in registration applications is the global rigid transformation, because the rigid body constraint is a good approximation in many common medical images.

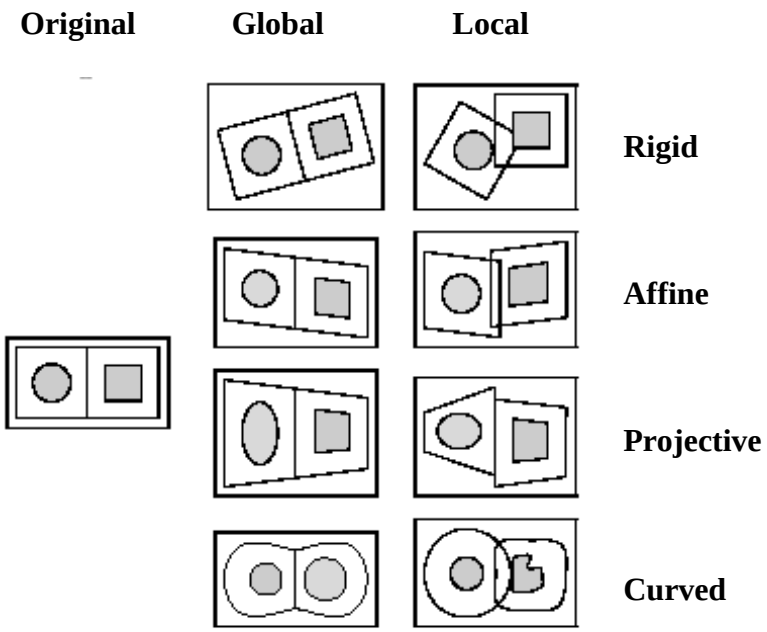



Figure 2.1. Examples of 2D transformations (adapted from [24]).

In the optimization procedure used in existing registration methods, transformation parameters can be either computed or search for. If the parameters can be determined in an explicit fashion, then the parameters can be computed directly. Otherwise the parameters need to be determined by finding an optimum of some function defined on the parameter space, i.e., searched for. In the former case, the manner of computation is completely determined by the paradigm. In the case of searching optimization methods, most registration methods are able to formulate the paradigm in a standard mathematical function of the transformation parameters to be optimized. If the similarity function is well behaved (quasi-convex), one of many standard and well-documented optimization techniques [40] can be used. Many applications use more than one optimization technique, frequently a fast but coarse technique followed by an accurate yet slow one. In addition, multi-resolution and

multi-scale approaches can be used to speed up convergence or to reduce the number of transformations to be examined and to avoid local minima.

### 2.3.4 CT Bone Registration

Here we are interested in *bone registration* based on CT segmentation. The transformations found in bone images are all rigid, as they concern mainly the displacement of bones. CT modality is used since it has better contrast for bone structures compared to other modalities.

Some special methods for bone registration were proposed by Münch [41], Jacq and Roux [42] and van den Elsen [43]. Münch performed an automatic registration by optimizing the cross-correlation of femural images; Jacq and Roux performed curved automatic registration on images of the humerus by minimization of the local grey value differences, and van den Elsen performed 3D rigid automatic registration in a full image content based way by optimizing the cross-correlation between a CT and MR image, where the CT gray values are first remapped using localized linear transforms.

However, most registration methods are surface-based, since anatomical surfaces are usually explicitly identified with tomograhic data such as MRI and CT, and are often closed. In the case of rigid models, these methods are always easy to perform and the computational complexity is relatively low. Those surface-based methods differ in elaboration of surface representation, similarity criterion, matching and global optimization. Besl and McKay propose the iterative closest point method [44] to determine the closest point pairs followed by computing the transformation from these pairs with a quaternion technique. This method is also a common basis of

many other methods that followed. Hemler, Naper, and Sumanaweerea propose a 3D registration system on an automatically extracted, user corrected surface, on CT calcaneus images [45] and on CT and MR spinal images in [46, 47]. In this system, the corresponding surface to be registered is first identified in each image set, and a set of 2D polygon points is used to represent the surface in the other image set. A least-squares minimization technique is then used to determine the rigid-body transformation which minimizes a cost function related to the sum-square perpendicular distance between the two surfaces. Bainville [48] found a local *curved* spline deformation using the local closest point of the surfaces combined with a regularization term. However, these methods all incur heavy computational cost in searching for point correspondences. Though some methods, e.g. [49], have been proposed to accelerate the process, the speed is still a problem in real-time applications.

Burel [50] has proposed a method for estimating the orientation of 3D objects without point correspondence information. It performs 3D registration by decomposing each surface into its spherical harmonics. The optimization is then done by using their special geometrical invariances. This method does not need point matching, it uses some direct linear algebra computations without an iterative search, and it is computationally fast. A crucial drawback of this method is that it is suitable for transformation which only has rotation. And it produces noticeable rotational error when the translation estimation is not accurate.

## 2.4 Statistical-based Modeling

The potential use of finite element (FE) models to plan, evaluate and investigate surgical treatments has been recognized for a long time since the early 1980s [51, 52] for musculoskeletal, [53] for bone mechanics, and [54] for prosthetic design. These investigations mainly employed generic FE meshes based on average patient geometries. Nevertheless, anatomical structures have significant variations in geometrical shape and tissue properties among different individuals. The complex boundaries of anatomical structures, further complicated by pathologies such as scoliosis, is very difficult to be represented mathematically. Manual development of patient-specific FE models from medical images is therefore a laborious task. A focus of our research is on fast generation of patient specific lumbar spinal model for surgical simulation using FE methods.

There is currently a variety of methods available for constructing patient specific FE meshes [55, 56, 57, 58]. Generally, these methods involve image segmentation to define the boundaries of the organ, geometrical modeling to reconstruct the surface of the organ from the boundaries, and discretization of the volume enclosed by the surface. In some published work, the latter two processes are combined into a single process. The direct-voxel conversion method by Keyak *et al*. [59, 60] converted the voxels from segmented CT images into hexahedral elements directly. Based on the structure model, Keyak *et al*. [61] further developed an accurate and precise method of predicting proximal femoral strength and fracture location for research and clinical studies of hip fracture related to osteoporosis and metastatic disease. Nevertheless, there might be inaccurate results at the surface of the structure after

the discretization process. Luboz *et al*. [62] proposed a method aimed at correcting irregularities of 3D model meshes in order to perform FE computations. The methodology is based on a mesh-matching method and a regularization technique using the Jacobian matrix transform related to the FE reference element and the current element. The marching cube algorithm [63] is used to improve the discretization process in our earlier work [57] to achieve better conformance to the boundaries. However, the resultant FE model has a large number of tetrahedral elements for a clinically relevant computational analysis.

Constructing a practical FE model from medical images is clearly not trivial due to the significant inter-subject variability of anatomy and function. The template-based approach, more commonly known as the atlas-based approach in medical image computing, addresses this problem by defining a common reference space. Mapping data sets into this common reference space not only accounts for anatomical and functional variations of individual subjects, it also offers a powerful tool which facilitates comparison of anatomy and function over time, between subjects, and between groups of subjects. The mapping can be achieved using various methods such as mapping functions and the non-rigid deformation algorithm based on free-form deformation with hierarchical multi-resolution representation of a deformation spline [62, 64]. Rossa et al. [65] developed a deformation method using the thin plate spline model and the minutia point correspondences between pairs of fingerprint impressions. In order to obtain the template or “baseline” model that is representative of the population, probabilistic and statistical approaches which include information from a group of subjects were proposed in [66, 67, 68].

# 3 Segmentation

## 3.1 Introduction

Segmentation is the image analysis process to isolate the object of interest from the background. The objective of segmentation is to identify which part of the data array makes up an object in the real world. Segmentation supports tasks such as measurement, visualization, registration, reconstruction and content-based search, each of them with specific needs. In the research work described in this Chapter, the role of segmentation is to separate the bone of interest from its surroundings, such as soft tissues. The segmentation results are then used to identify regions containing the 3D surface of the bone, which is used for subsequent registration.

There are various segmentation techniques developed. However, no standard segmentation technique can produce satisfactory results for all imaging applications. Automatic processing is desirable, but sometimes unattainable due to limitations imposed by image acquisition, abnormalities in the scene, or both [69]. The choice of a segmentation method is strongly dependent on the type and characteristics of the image. Likewise there is no universal segmentation method for bone images.

## 3.2 Method

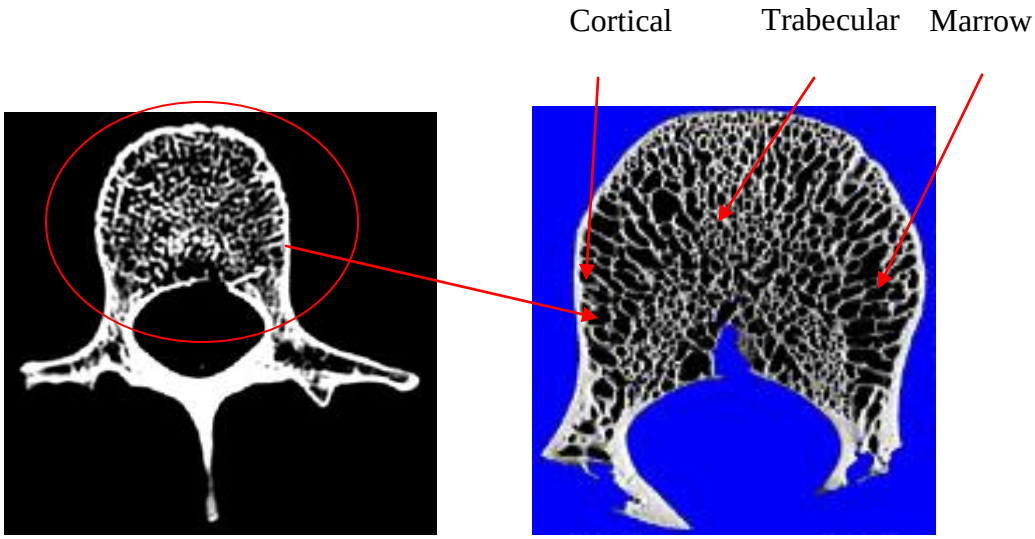
Haralick and Shapiro [70] have established the following qualitative guideline for a good image segmentation: “Regions of an image segmentation should be uniform

and homogeneous with respect to some characteristic such as gray tone or texture. Region interiors should be simple and without many small holes. Adjacent regions of segmentation area should have significantly different values with respect to the characteristic on which they are uniform. Boundaries of each segment should be simple, not ragged, and must be spatially accurate." Unfortunately, no quantitative image segmentation performance standard has been developed.

Kass *et al*. [71] developed the snake method which models a closed contour to the boundary of an object. The snake model is a controlled continuity closed contour that deforms under the influence of internal forces, image forces and external constraint forces.

Since bone structures are of high intensity levels in CT images, they can usually be separated from soft tissue using thresholding-based methods. However, simply employing global thresholding would fail due to the partial volume effect, beam hardening and intensity inhomogeneity of bone structures, and most segmentation methods are based on local (adaptive) thresholding. The local threshold can be selected based on local intensity distribution. Some methods use the mean plus standard deviation or mean of the maximum and minimum values [72, 73], while others use statistics based on local intensity gradient magnitude [74]. Nevertheless, those methods still do not perform well because of the partial volume effect and intensity inhomogeneity. In [75], a 2D iterative adaptive thresholding method, which is a variation of the ISODATA segmentation algorithm [76], is proposed for automatic and accurate segmentation of bone structures of CT images. However, it requires a lot of manual initialization work for volumetric images, and hence is not suitable for practical use. Here we have developed a 3D adaptive thresholding

method based on [75], which is near automatic, for the registration system. In this method, the 3D correlation of each object in the various slices is used to minimize the manual interactions.




(a) (b)

Figure 3.1. Spine structure. (a) A typical spine specimen. (b) Enlarged view of the vertebral body.

Segmentation by global thresholding will fail because of the partial volume effect (due to insufficient sampling and detector response), beam hardening (due to polychromaticity of the X-ray beam), intensity inhomogeneity of bone structures, and high gray level of surroundings. A typical spine specimen is shown in Figure 3.1 to illustrate bone structure. There are two major types of bone: cortical bone and trabecular bone. Cortical bone forms the outer shell and trabecular bone forms the inner portion. Cortical bone is 5% to 30% porous, with trabecular bone being 30% to 90% porous. The trabecular bone structures are of a branching pattern with marrow between them. The trabecular bone-marrow mixture is completely enclosed by a layer of cortical bone, which has a higher intensity.

In the following example, we are interested in obtaining a set of disjoint regions that correspond to individual bone and background. A CT image of the spine (Figure 3.2 (a)) is taken as an example to aid the explanation of the segmentation algorithm. A threshold that is too low is not sufficient to separate bone from the surroundings (Figure 3.2 (b)) and a threshold that is too high will misclassify bone regions that have gray level due to the partial volume effect (Figure 3.2 (c)). Our segmentation algorithm uses a local adaptive thresholding scheme that is capable of producing an accurate segmentation under these conditions (Figure 3.2 (d)).

The method we developed comprises two main steps: initial segmentation and iterative adaptive thresholding. Figure 3.3 illustrates the entire procedure. We manually select a region of soft tissue near the bone (Figure 3.3 (a)), and then perform initial thresholding using the threshold estimated from that region (Figure 3.3 (b)). A floodfilling procedure then gives us the result of initial segmentation (Figure 3.3 (c)). The final segmentation result is achieved after iterative adaptive thresholding (Figure 3.3 (d)).

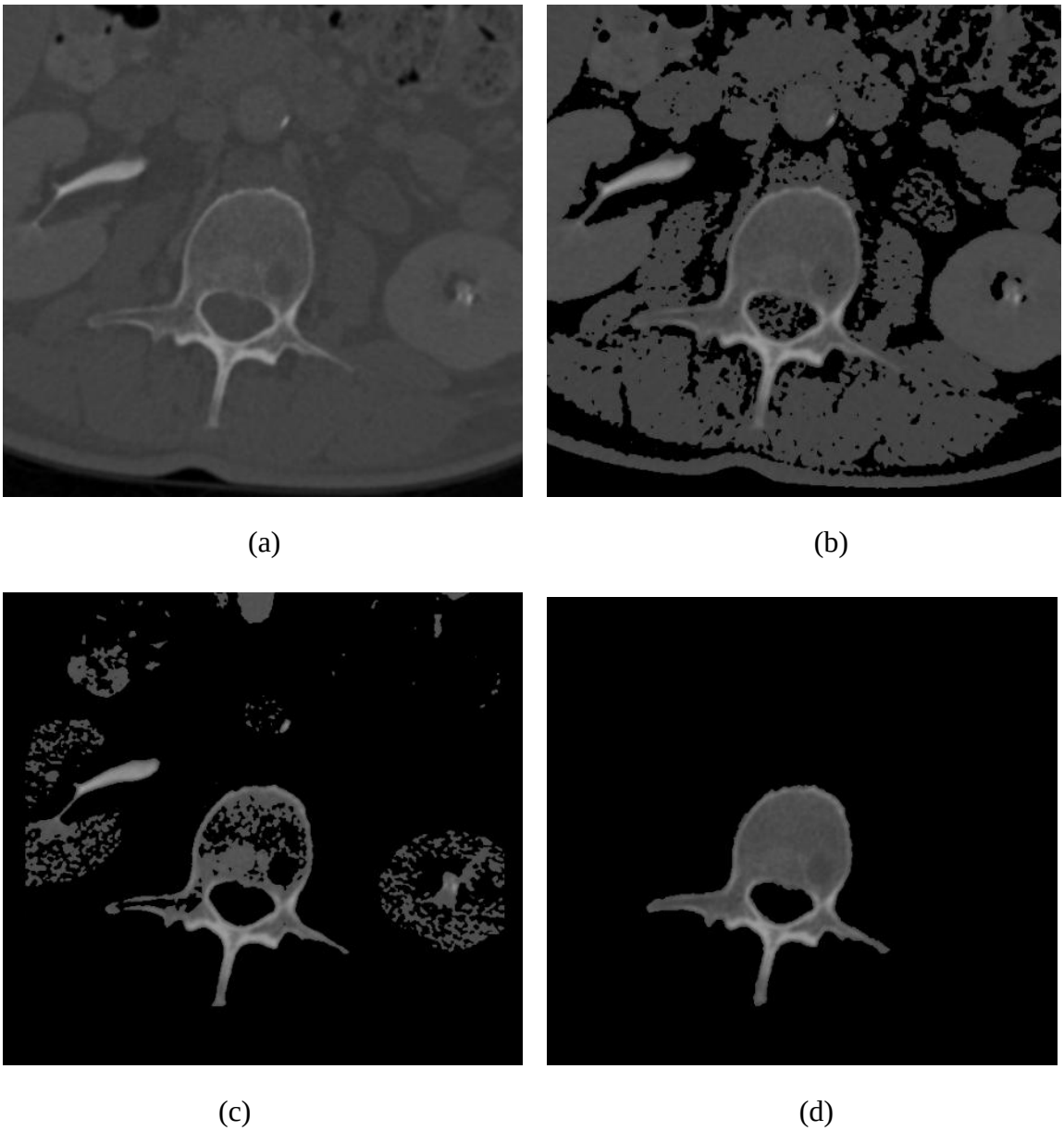



Figure 3.2. (a) CT image of spine. (b) Image produced by low threshold. (c) Image produced by high threshold. (d) Image produced by using our adaptive thresholding method.

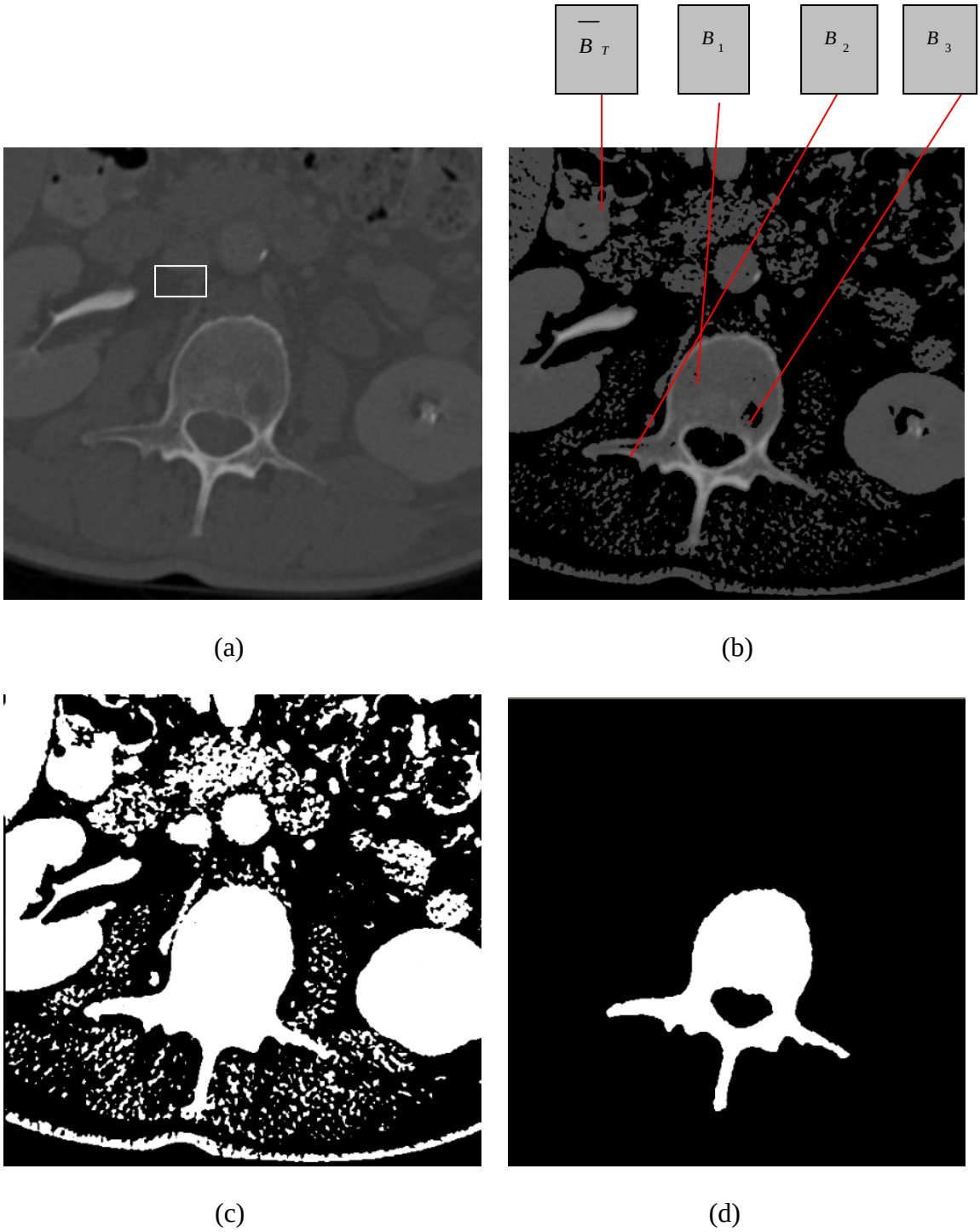



Figure 3.3. Illustration of segmentation procedure. (a) The pixels inside the white box are used to estimate the mean $\mu_f$ and the standard deviation $\sigma_f$ of soft tissue. (b) Image produced by thresholding the CT image with a threshold of $\mu_f + 2\sigma_f$. (c) Non-bone region extracted by floodfilling the thresholded image: the result of initial segmentation. (d) Bone region after iterative adaptive thresholding.

### 3.2.1 Initial Segmentation

We manually select a region (e.g., see Figure 3.2 (a)) of soft tissue near the bone to obtain estimates for the mean gray level $\mu_f$ and standard deviation $\sigma_f$ of the soft tissue. We then produce an image by thresholding the CT image with a threshold $\mu_f + 2\sigma_f$ (Appendix A). This image is used to classify each pixel of the CT image into two classes: $B$ (bone) and $\overline{B}$ (non-bone), that is, for a pixel $x$ with gray level $I(x)$,

$$x \in \begin{cases} B & \text{if } I(x) \geq \mu_f + 2\sigma_f \\ \overline{B} & \text{otherwise} \end{cases} \tag{3.1}$$

From Figure 3.2 (b), we note that the interior of each bone, which is a mixture of trabecular bone and marrow, has a gray level below the threshold and is misclassified as non-bone. The trabecular bone-marrow mixture is completely enclosed by a layer of cortical bone, which has a very high gray level and is always classified correctly as $B$. This means that regions of trabecular bone-marrow mixture are not connected to regions of true non-bone (soft tissue, fat, and air). Hence $\overline{B}$ can be written as

$$\overline{B} = \overline{B}_T \cup B_1 \cup B_2 \cup B_3 \cup \cdots \cup B_n \tag{3.2}$$

where $\overline{B}_T$ is the true non-bone region and the $B_i$'s are the trabecular-bone-marrow mixture regions and $n$ is the number of regions.

To extract $\overline{B}_T$, we first locate any pixel $u \in \overline{B}_T$. A pixel $v \in \overline{B}$ is connected to $u$ if at least one of the following conditions is true:

- $v$ is one of the 8-connected neighbors of $u$.

- $v$ is connected to $w \in \overline{B}$ and $w$ is connected to $u$.

Thus $v$ is connected to $u$ if and only if $v \in \overline{B}_T$. Hence, we can segment out $\overline{B}_T$ by identifying all pixels in $\overline{B}$ that are connected to $u$. We can then re-classify the trabecular bone-marrow regions to $B$ and rename $\overline{B}_T$ as $\overline{B}$ (Figure 3.2 (c)):

$$B \leftarrow B \cup B_1 \cup B_2 \cup B_3 \cup \cdots \cup B_n = B \cup \overline{B} / \overline{B}_T \qquad (3.3)$$

$$\overline{B} \leftarrow \overline{B}_T$$

The segmentation result is used as the initial input for an iterative adaptive thresholding scheme that is described.

### 3.2.2 Iterative Adaptive Thresholding Algorithm

The regions after the initial segmentation are $B$ and $\overline{B}$. A pixel is said to be on the boundary of $B$ and $\overline{B}$ if at least one of its connected neighbors does not belong to the same category ($B$ or $\overline{B}$) as the pixel. We gather all the boundary pixels in $B$ to form a set $E_B$. Next, we define $W(x)$, a window centered on pixel $x$. The iterative adaptive thresholding algorithm, which is a variation of the ISODATA segmentation algorithm, is described by the following steps:

- Compute $E_B$ from the current segmentation ($B$ and $\overline{B}$).
- For each pixel $x$ in $E_B$
  - assume that the CT data in $W(x)$ come from a mixture of two Gaussian distributions (bone and non-bone) having respective means and variances $(\mu_b, \sigma_b^2)$ and $(\mu_{nb}, \sigma_{nb}^2)$,
  - classify $W(x)$ using the *Bayes decision rule* (Appendix A),
  - add $x$ to the error class $R$ if $x$ is classified as non-bone.

- Update the current segmentation: $B \leftarrow B / R$, $\overline{B} \leftarrow \overline{B} \cup R$
- Iterate until convergence.

By using the iterative adaptive thresholding algorithm with the initial segmentation, the final segmentation result that is shown in Figure 3.3 (d) is achieved.

In practice, if we simply apply the above algorithm in a 2D case, a lot of manual work is needed to get the proper initial segmentation (as explained in section 3.3.1) for volumetric images. Hence, we require an automatic segmentation method, 3D adaptive thresholding, for the volumetric images.

### 3.2.3 3D Adaptive Thresholding

In medical applications, 2D images are stacked up to form a 3D dataset. This dataset can be treated as a digital representation of the region of interest. To maximize the 3D correlation of each object in the various slices, we implement a 3D adaptive thresholding procedure. This procedure requires minimal manual interaction.

In 3D adaptive thresholding, we use a 10-voxel neighborhood (Figure 3.4) to gather all the boundary pixels in $B$ to form a set $E_B$, which stores all the boundary pixels. To be specific, for each pixel $x$ belonging to $B$, if one of its 10-connected neighbors ($N(x)_i \; i = 1,2,3,\cdots 10$) does not belong to $B$, we put $x$ into the set $E_B$.

For each boundary pixel in $E_B$, we define the local window $W(x)$ for each boundary pixels to be a cylindrical region, as shown in Figure 3.5. We then compute the respective means and variances ($\mu_b, \sigma_b^2$) and ($\mu_{nb}, \sigma_{nb}^2$) for bone and non-bone inside the corresponding region and similarly reclassify all the voxels as

described in section 3.3.1 and section 3.3.2. The entire volumetric image is processed in each iteration.

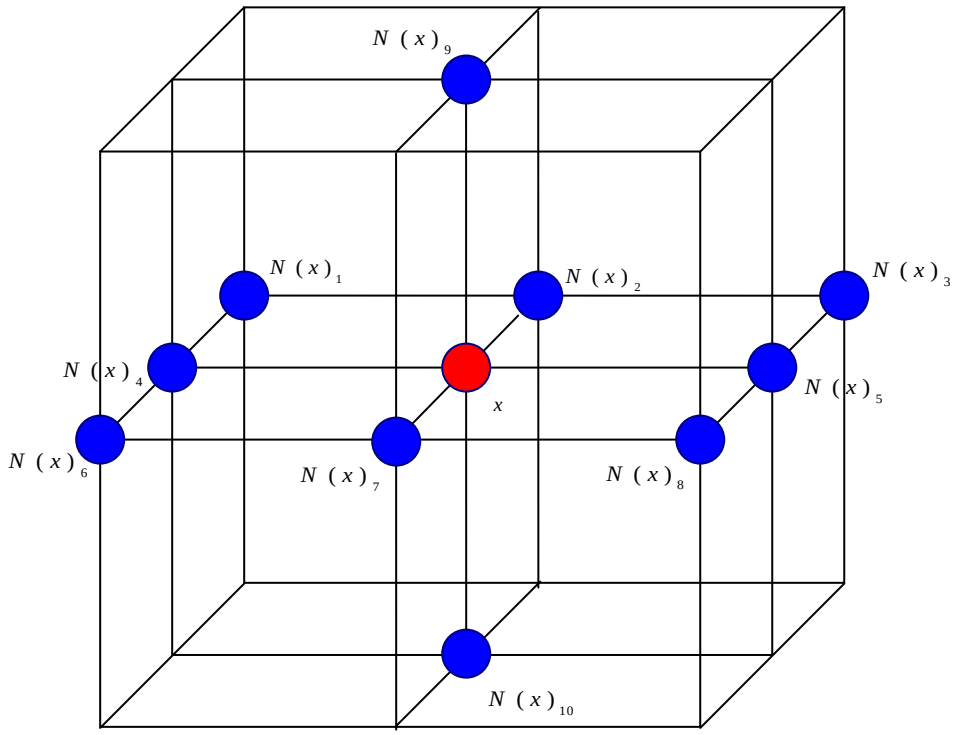



Figure 3.4. 3D neighborhood definitions.

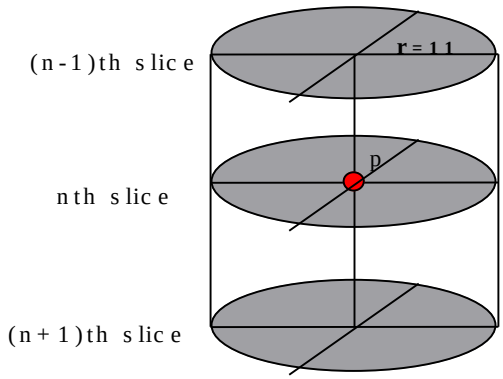



Figure 3.5. 3D window definitions.

## 3.3 Experiments

### 3.3.1 Dataset

*(a) Calcaneus*

CT scans using GE HiSpeed CT/i system from NASA Ames Research Center. Image volume contains 144 slices. Every slice has 512×512 voxels with voxel dimensions 0.3mm×0.3mm×0.5mm (the slice thickness), at 12 bits.

*(b) Spine*

Dataset 1:

CT scans using Toshiba high-resolution multislice CT machine located at Johns Hopkins University, Dept of radiology. Image volume contains 295 slices. Every slice has an in-plane resolution of 512x512 voxels with *voxel* dimensions 0.8mm× 0.8mm×0.7mm (the slice thickness),at 16 bits.

Dataset 2:

CT scans using Siemens system, located at the National University Hospital of Singapore. Image volume contains 59 slices. Every slice has 512×512 voxels with voxel dimensions 0.488mm×0.488mm×0.4mm (the slice thickness), at 12 bits.

### 3.3.2 Experimental Design

The segmentation procedure may be broken down into three steps for implementation:

(a) Initial thresholding

(b) Automatic floodfilling

(c) Iterative adaptive thresholding

Figure 3.6 shows the flowchart of this process.

*(a) Initial Thresholding*

Initial thresholding is the first step to get the initial contour of the 3D surface. The threshold is selected based on the region that is of lower gray level and uniform distribution. The mean $\mu_f$ and standard deviation $\sigma_f$ of this region are used to calculate the threshold. Since the gray level is similar for all two-dimensional image slices, we can manually select a region just outside the bone (e.g. the soft tissues) of interest from any one slice. Then we perform global thresholding on the entire data set.

*(b) Automatic Floodfilling*

In order to obtain the initial contour for 3D surface, we need to do floodfilling. In section 3.3, we mentioned that if we simply applied the above algorithm, much manual work would be needed to obtain the proper initial segmentation. The 3D adaptive thresholding method was developed to minimize manual interaction for the volumetric images. We can find the reason here by comparing it with the 2D adaptive thresholding method.

Figure 3.7 shows two consecutive slices after initial thresholding. When we select a seed at the top left corner of the image and perform floodfilling, we can get good results using 3D adaptive thresholding, while the boundary pixels inside the bone region could not be detected using 2D adaptive thresholding. This is because we can use information from the $(N+1)th$ slice in the 3D case, while in the 2D case we cannot. Figure 3.8 and Figure 3.9 show the results using different seed selections for

the 2D case. Figure 3.10 shows the typical process and the final results for the 3D case.

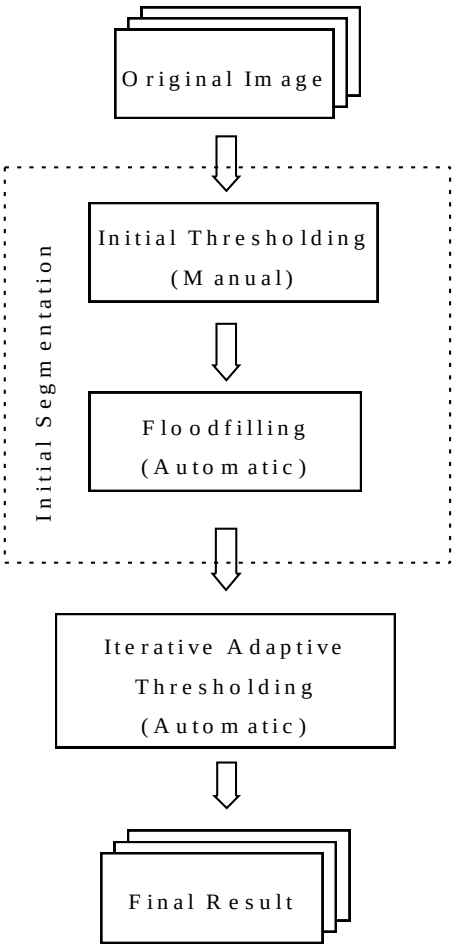
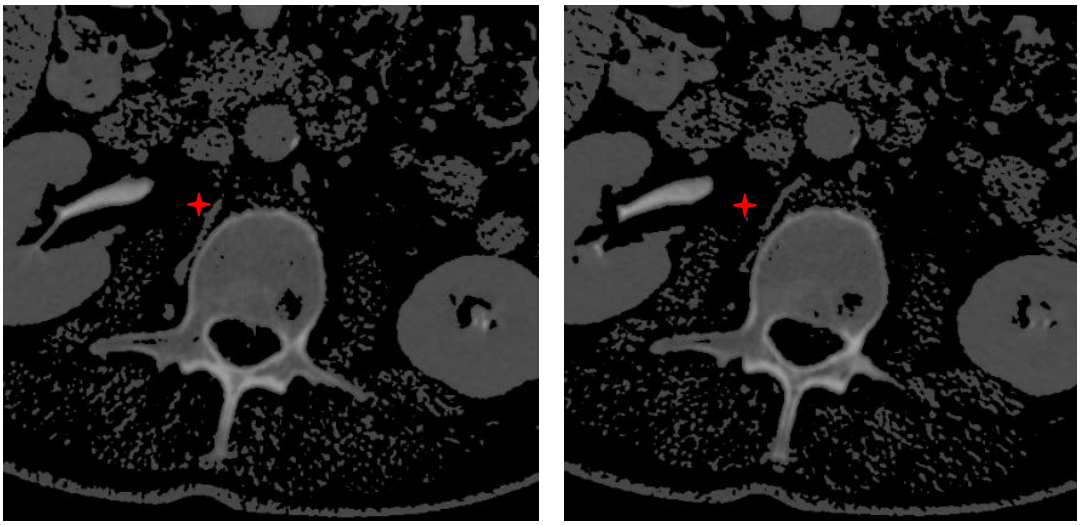



Figure 3.6. Implementation procedure.

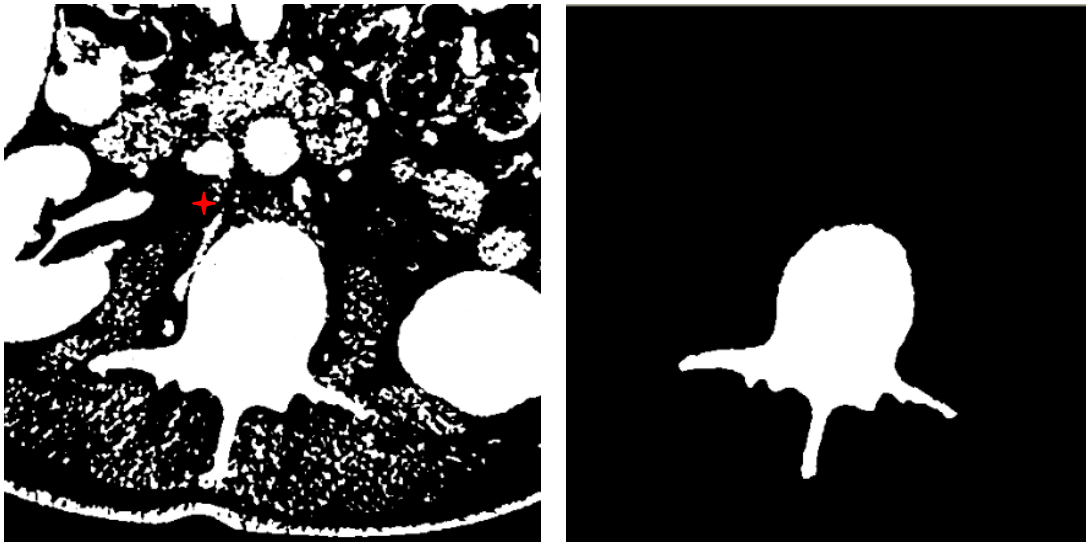


(a) (b)

Figure 3.7. Original initial thresholded images. (a) *N*th slice. (b) (*N*+1)th slice.

(a) (b)

Figure 3.8. 2D adaptive thresholding result of *N*th slice using automatic seed selection at the top left corner of image. (a) Initial contour, *N*th slice, automatic seed selection. (b) Final result, *N*th slice.



(a) (b)

Figure 3.9. 2D adaptive thresholding result of *N*th slice using manual seed selection. (a) Initial contour, *N*th slice, manual seed selection. (b) Final result, *N*th slice.

We can see from Figure 3.8 and Figure 3.9 that in the 2D case, manual selection of seed for floodfilling is needed for good results. However in the 3D case (Figure 3.10), we only need to automatically select one seed from the background (for all CT images, the pixel located at (1,1) belongs to the background). And using the information from $(N+1)th$ slice together with neighborhood definition described in Figure 3.4, the boundary pixels inside the bone region could be easily figured out for further processing.

*(c) Iterative adaptive thresholding*

At each iteration, we similarly reclassify all the pixels on the boundary using the 3-D definition of the window (cylinder). The algorithm will automatically process the entire volumetric data iteratively until convergence.

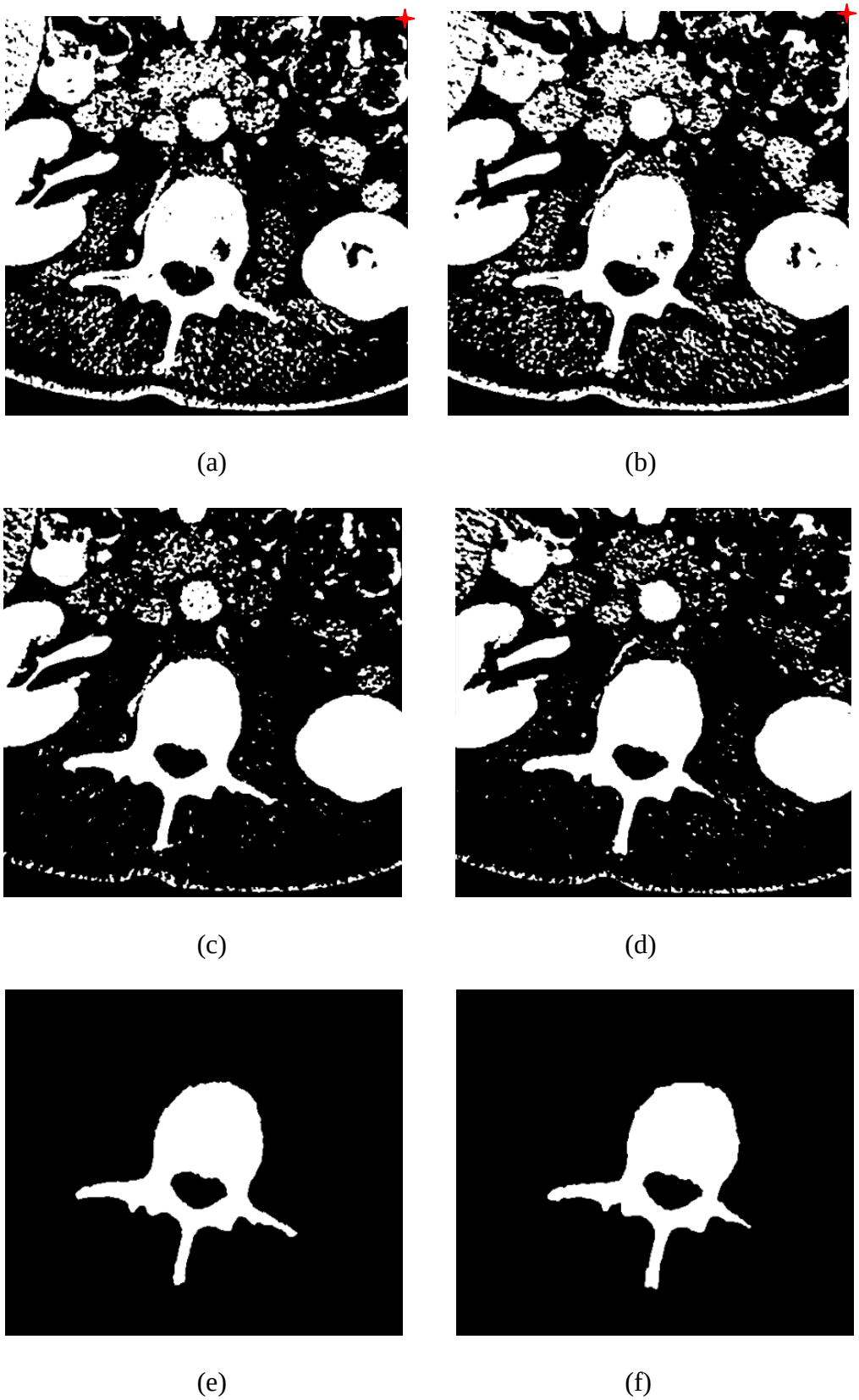



Figure 3.10. 3D adaptive thresholding result of *N*th slice. (a) Initial contour, *N*th slice. (b) Initial contour, (*N*+1)th slice. (c) 1st iteration, *N*th slice. (d)1st iteration, (*N*+1)th slice. (e) Final result, *N*th slice. (f) Final result, (*N*+1)th slice.

## 3.4 Results and Discussion

We have developed this 3D adaptive thresholding method for CT images of bone structures and applied to CT scans of the calcaneus and spine.

*(a) Accuracy*

Figure 3.11 shows the results of the calcaneus. Figure 3.12 and Figure 3.13 show those of the spine.

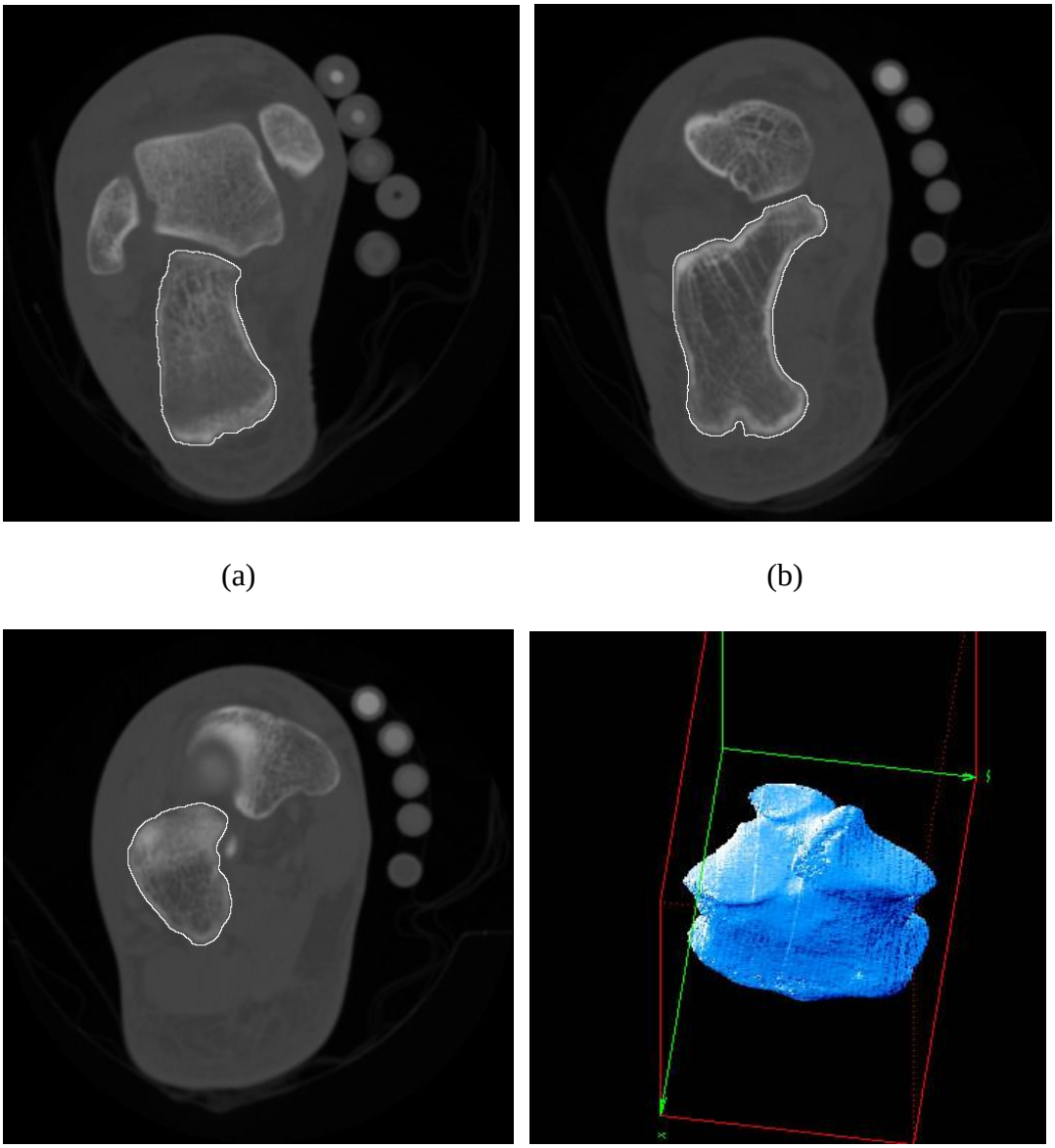
We assess the accuracy of the segmentation by comparing the segmentation result with the manual segmentation performed by an experienced radiologist. The volumetric overlap of these two segmentations is measured by the Hausdorff distance (HD) and mean distance shown in Table 3.1. The HD is calculated from

$$\mathrm{HD}(\mathbf{P}, \mathbf{Q}) = \max_{\mathbf{p} \in \mathbf{P}} \{ \min_{\mathbf{q} \in \mathbf{Q}} \{ D(\mathbf{p}, \mathbf{q}) \} \}, \tag{3.4}$$

where $\mathbf{P}$ and $\mathbf{Q}$ represent the surface point clouds of two datasets; $\mathbf{p}$ and $\mathbf{q}$, respectively, are points on the two surfaces, and $D$ is the distance between any two points.

Table 3.1. Segmentation accuracy measurements.

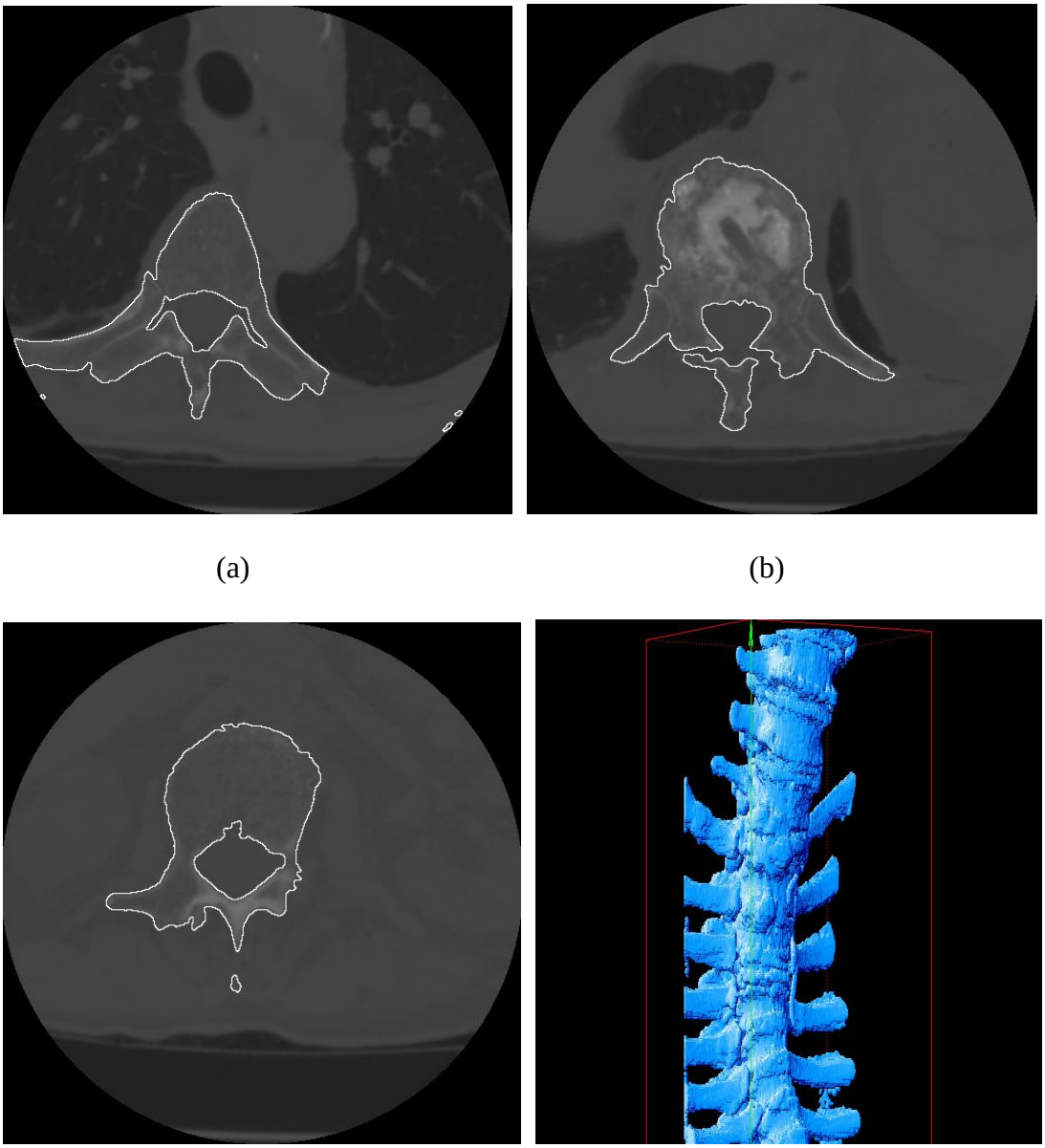
| Datasets | Minimum Distance (mm) | Mean Distance (mm) |
|---|---|---|
| Calcaneus | 0.21 | 0.08 |
| Spine dataset 1 | 0.64 | 0.26 |
| Spine dataset 2 | 0.43 | 0.15 |



(a) (b)

(c) (d)

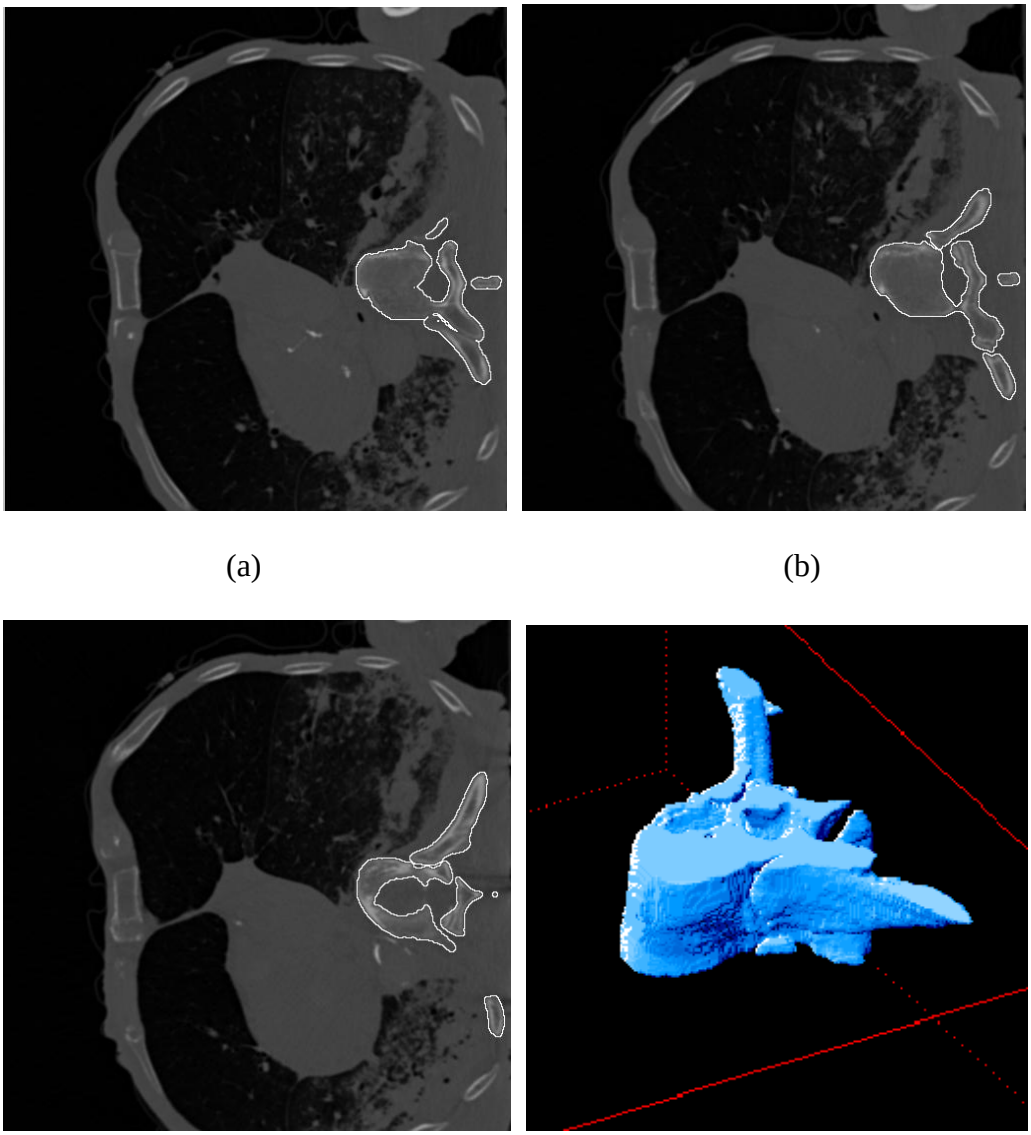
Figure 3.11.Calcaneus segmentation results. (a)-(c) An overlay of the detected surface results at different locations of calcaneus. (d) Reconstructed 3D image based on segmentation results.



(a) (b)

(c) (d)

Figure 3.12. Spine segmentation results, dataset 1. (a)-(c) An overlay of the detected surface results at different locations of spine. (d) Reconstructed 3D image based on segmentation results.



(a) (b)

(c) (d)

Figure 3.13. Spine segmentation results, dataset 2. (a)-(c) An overlay of the detected surface results at different locations of spine. (d) Reconstructed 3D image based on segmentation results.

*(b) Processing time*

The segmentation time of each dataset are listed in Table 3.2. All simulations were running using Dell workstation with Pentium IV 2.6GHz, 2GB memory. Visual C++ (version 6.0), OpenCV library and IPPI (Intel® Integrated Performance Primitives: Image and Video Processing) were used to implement the above segmentation procedures for fast execution.

Table 3.2. Processing time.

| | Calcaneus | Spine, data set 1 | Spine, data set 2 |
|---|---|---|---|
| Total time | 1 hour | 4 hours | 25min |
| Average time | 25s | 48s | 25s |

We see that the 3D adaptive thresholding method is good for detection of the outer contour of these two kinds of bone structures. We have identified some limitations of this segmentation algorithm, e.g., narrow gaps between bones in Figure 3.14. This could be improved by the use of prior knowledge of bone structure to adjust the window definition to obtain sufficient statistical information for threshold selection. The processing time is applicable for the segmentation of preoperative scans, but directly applying the method is still time-consuming for intraoperative segmentation. Multi-resolution methods could help to achieve higher execution speeds in practice.

## 3.5 Conclusion

Due to the partial volume effect, beam hardening, intensity homogeneity of bone structures and high gray level of surroundings, simple thresholding techniques are

not able to extract the bone from normal CT images accurately and automatically. We have developed a semi-automatic 3D adaptive thresholding segmentation algorithm to extract bone structures from clinical CT data. The fairly good results can be achieved within a short period.

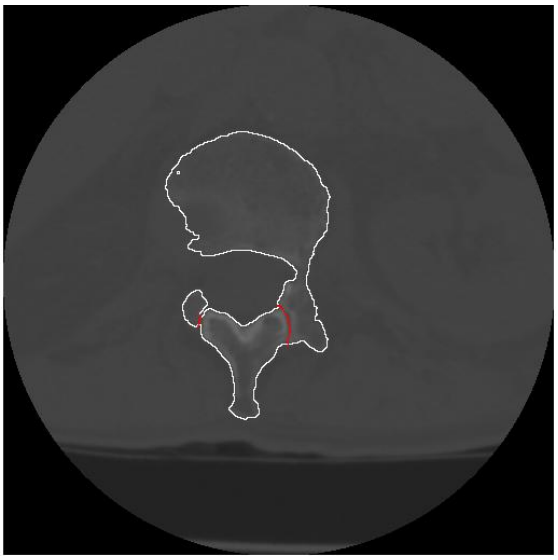


Figure 3.14. Red line highlights the narrow gaps that were not detected.

# 4 Surface Based Registration

## 4.1 Overview of Registration System

Figure 4.1 shows the architecture of the proposed registration system and its interfaces with external entities. The registration system comprises two main processes: semi-automatic segmentation for both preoperative and intra-operative scans, and automatic real-time registration for intra-operative scans.

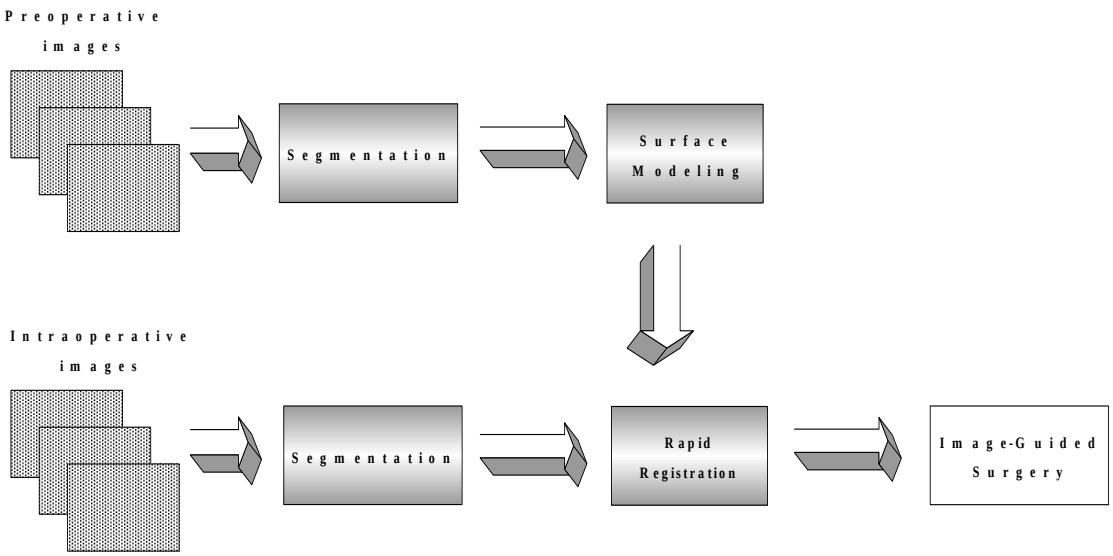
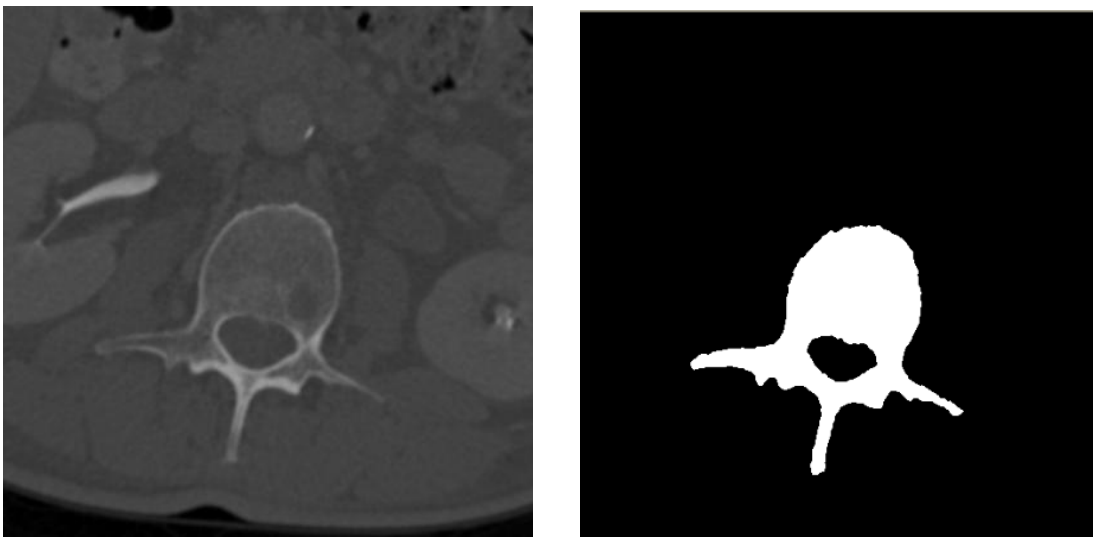



Figure 4.1. A registration system for image-guided surgery.

Segmentation is first performed to separate the bone of interest from its surroundings. From this, we identify the bone surface that will be used in the registration procedure. Since CT imaging is a high-resolution modality, the set of 2D contours extracted slice by slice constitutes the 3D bone surface model. The preoperative and intra-operative image volumes are then aligned or registered into the same geometric space. The registration algorithm comprises two steps: 3D

surface modeling with a NN, and an optimization procedure to determine the transformation that best aligns the bone surfaces from the preoperative and intra-operative scans. The suitably registered datasets can then be employed for image-guided surgery [77, 78].

We use the local adaptive thresholding scheme described in chapter 3 to segment the bone structures. An example of the final segmentation result is shown in Figure 4.2.



(a) (b)

Figure 4.2. Segmentation results. (a) Original CT image. (b) Bone region after iterative adaptive thresholding.

The registration algorithm consists of two steps: coarse registration to obtain an initial estimate of the transformation followed by fine registration to achieve sub-pixel accuracy. The fine registration step achieves fast computation by reducing the computational requirement of the cost function. Using the transformed surface coordinates of the intra-operative dataset as input, the cost is the sum of the output of the forward-feed NN. Current registration techniques focus mainly on fast optimization routines to reduce the overall time for surface-based registration. However, the cost function used in surface-based registration is highly

computationally intensive. We employ a novel NN-based technique to achieve sub-pixel accuracy within a few minutes. This is sufficiently fast for intra-operative registration. The segmentation, coarse and fine registration procedures are described in the next section.

## 4.2 Methods

### 4.2.1 CT Image Segmentation

Using the method mentioned in Chapter 3, segmentation is first performed to separate the bone of interest from its surroundings. From this, we identify the bone surface that will be used in the registration procedure. The set of 2D contours extracted slice by slice constitutes the 3D bone surface model.

### 4.2.2 Coarse Registration and Neural-Network-based Registration

It is assumed that the reference dataset and the current dataset are related by a rigid body transformation. We denote the reference data point by $\mathbf{p}$ and the intra-operative data point by $\mathbf{q}$, where $\mathbf{p}$ and $\mathbf{q}$ are 3D vectors. They are related by $\mathbf{p} = \mathbf{R}\mathbf{q} + \mathbf{t}$, where $\mathbf{R}$ is a $3 \times 3$ rotation matrix and $\mathbf{t}$ a 3D translation vector. The aim of our registration algorithm is to obtain a quick and accurate estimate for $\mathbf{R}$ and $\mathbf{t}$.

*(a) Coarse Registration*

The main function of coarse registration is to provide an initial approximate and robust estimate of the transformation. The registration should be fast and yet be able

to avoid local minima so that the estimated transformation parameters can be used as a good starting point for subsequent fine registration.

If $\mathbf{p}$ denotes a point from the reference image data and $\mathbf{q}$ the corresponding point from the intra-operative image data, $\mathbf{p}$ and $\mathbf{q}$ are related by

$$\mathbf{p} = \mathbf{R}\mathbf{q} + \mathbf{t} \tag{4.1}$$

With $N$ corresponding point pairs, $(\mathbf{p}_1, \mathbf{q}_1) \cdots (\mathbf{p}_N, \mathbf{q}_N)$, equation (4.1) can be extended to

$$\mathbf{P} = \mathbf{R}\mathbf{Q} + \mathbf{T}, \tag{4.2}$$

where $\mathbf{P}, \mathbf{Q}$ and $\mathbf{T}$ are $3 \times N$ matrices defined by

$$\mathbf{P} = (\mathbf{p}_1 \mathbf{p}_2 \cdots \mathbf{p}_N), \qquad \mathbf{Q} = (\mathbf{q}_1 \mathbf{q}_2 \cdots \mathbf{q}_N), \qquad \textit{and} \qquad \mathbf{T} = (\overbrace{\mathbf{t} \quad \mathbf{t} \quad \cdots \quad \mathbf{t}}^{N}). \tag{4.3}$$

By averaging over all the measured points, equation (4.2) can be written as

$$\overline{\mathbf{p}} = \mathbf{R}\,\overline{\mathbf{q}} + \mathbf{t} \tag{4.4}$$

with

$$\overline{\mathbf{p}} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{p}_i, \qquad \overline{\mathbf{q}} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{q}_i. \tag{4.5}$$

Equations (4.2) and (4.4) can be combined to give

$$\mathbf{P} - \overline{\mathbf{P}} = \mathbf{R}(\mathbf{Q} - \overline{\mathbf{Q}}) \tag{4.6}$$

where $\overline{\mathbf{P}}$ and $\overline{\mathbf{Q}}$ are $3 \times N$ matrices

$$\overline{\mathbf{P}} = (\overbrace{\overline{\mathbf{p}} \quad \overline{\mathbf{p}} \quad \cdots \quad \overline{\mathbf{p}}}^{N}) \quad \text{and} \quad \overline{\mathbf{Q}} = (\overbrace{\overline{\mathbf{q}} \quad \overline{\mathbf{q}} \quad \cdots \quad \overline{\mathbf{q}}}^{N}). \tag{4.7}$$

Denoting $\mathbf{P} - \overline{\mathbf{P}}$ and $\mathbf{Q} - \overline{\mathbf{Q}}$ by $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{Q}}$, respectively, the covariance of $\tilde{\mathbf{P}}$ can be written as

$$\tilde{\mathbf{P}}\tilde{\mathbf{P}}^t = \mathbf{R}\tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^t\mathbf{R}^t \tag{4.8}$$

with the superscript $t$ denoting the transpose. In equation (4.6), $\mathbf{p}_i$ and $\mathbf{q}_i$ have to be corresponding point pairs, whereas in equation (4.8), $\tilde{\mathbf{P}}\tilde{\mathbf{P}}^t$ and $\tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^t$ are the covariance of the surface points in the reference and current image coordinates, respectively, and there is no need to establish corresponding point pairs. In other words, $\tilde{\mathbf{P}}\tilde{\mathbf{P}}^t$ and $\tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^t$ can be constructed from independent sets of surface points from the reference and current surfaces. By expressing the SVD of $\tilde{\mathbf{P}}\tilde{\mathbf{P}}^t$ and $\tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^t$ as $\mathbf{U}\mathbf{\Lambda}_{\mathbf{p}}\mathbf{U}^t$ and $\mathbf{V}\mathbf{\Lambda}_{\mathbf{q}}\mathbf{V}^t$, respectively, the rotation matrix $\mathbf{R}$ can be estimated by minimizing the expression $\left\|\mathbf{U}\mathbf{\Lambda}_{\mathbf{p}}^{1/2} - \mathbf{R}\mathbf{V}\mathbf{\Lambda}_{\mathbf{q}}^{1/2}\right\|_F$ subject to $\mathbf{R}^t\mathbf{R} = \mathbf{I}$. The minimum is attained by setting $\mathbf{R} = \mathbf{U}\mathbf{V}^t$ [79], after which, the translation $\mathbf{t}$ can be estimated using equation (4.4).

The above technique is based upon principal-axes alignment method which is not commonly used in image registration. In summary, two independent sets of surface points ($\mathbf{P}$, $\mathbf{Q}$) are extracted from the reference and current surfaces. We next compute the centroids ($\bar{\mathbf{p}}$, $\bar{\mathbf{q}}$) and covariance matrices ($\tilde{\mathbf{P}}\tilde{\mathbf{P}}^t$, $\tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^t$). The rotation matrix is estimated by the SVD computation of the covariance matrices and the translation matrix by using equation (4.8). As the number of extracted surface points is finite, there will inevitably be errors in determining the centroids and covariance matrices. However, this technique is computationally inexpensive and is capable of producing a coarse but robust estimate that can be used as a good starting point for the more accurate rigid registration method described in the following section.

*(b) Surface Representation*

The cost function used in most surface-based registration techniques is a nonlinear function of the transformation parameters, the reference surface and the current surface. Denoting the transformation matrix by $\mathbf{T}$, the cost of a transformation function may be written as

$$C(\mathbf{T}) = f(\mathbf{D}_r, \mathbf{T}(\mathbf{D}_c)), \tag{4.9}$$

where $\mathbf{D}_r$ and $\mathbf{D}_c$ are the reference and current surfaces, respectively. The derivation of the nonlinear representation of the surface model is explained in the following section. The function achieves its minimum value when the transformed current surface is the closest, in terms of Euclidean distance, to the reference surface. This non-linear function is used as the criterion for selecting the best transformation parameters. Since, in typical image-guided clinical procedures, $\mathbf{D}_r$ is acquired prior to the operation, it may be used to create a computationally efficient function $f_r$ for the cost calculation. We can write equation (4.9) as

$$C(\mathbf{T}) = f(\mathbf{D}_r, \mathbf{T}(\mathbf{D}_c)) = f_r(\mathbf{T}(\mathbf{D}_c)). \tag{4.10}$$

Neural networks [80] have been commonly used in image processing but not for surface representation. The advantages of using NNs for this purpose include: (1) the ability to perform nonlinear modeling (suitable for complex surfaces such as vertebrae); (2) low computational requirements in cost calculation; (3) ease of implementation on hardware field-programmable gate array (FPGA); and (4) the ability to acquire $\mathbf{D}_r$ for extensive network training prior to the operation.

A neural network models the distance map of a point $\mathbf{p}$ from the reference surface. Letting $N_r$ denote the NN model derived from the reference surface and $\mathbf{p}_c$ the point from the current surface, the cost function can be computed by

$$C(\mathbf{T}) = \sum_{p_c} N_r(\mathbf{T}(\mathbf{p}_c)) \ . \tag{4.11}$$

This surface modeling procedure is explained in detail below.

*1) Creation of training data*: A function $d(x, y, z)$ is defined as the signed distance from a point $(x, y, z)$ to the reference surface. Points inside the reference surface have a negative distance while points outside the surface have a positive distance. Since a sample point on the surface satisfies $d(x, y, z) = 0$, the reference surface is defined implicitly as the zero set of this function. Within a spherical volume that encloses the entire reference surface, points are sampled uniformly for preparing the training data. The points satisfying $|d(x, y, z)| \leq 5$ will be used for training. This range was obtained empirically. From our observation, the registration error from coarse registration was less than 5 pixels. We only need to consider the 5-pixel surrounding region of any point. Hence, the distance function is defined as less than 5.

*2) Surface modeling using MLP*: The multilayer perceptron (MLP) is used to model the bone surface. This NN is computationally efficient and can be implemented efficiently using hardware [81]. Let the coordinates $(x, y, z)$ of a point $\mathbf{p}$ be the three input neurons of the MLP, and the distance $d$ of $\mathbf{p}$ to the reference surface be the output of the neuron in the last layer of the network. A multilayer network is constructed to map the relationship from the 3D Euclidean input space,

input point $(x, y, z)$, to a 1D Euclidean output space, the shortest signed distance to the reference surface. The activation function φ(·) is a hyperbolic tangent. The structure of the NN is shown in Figure 4.3.

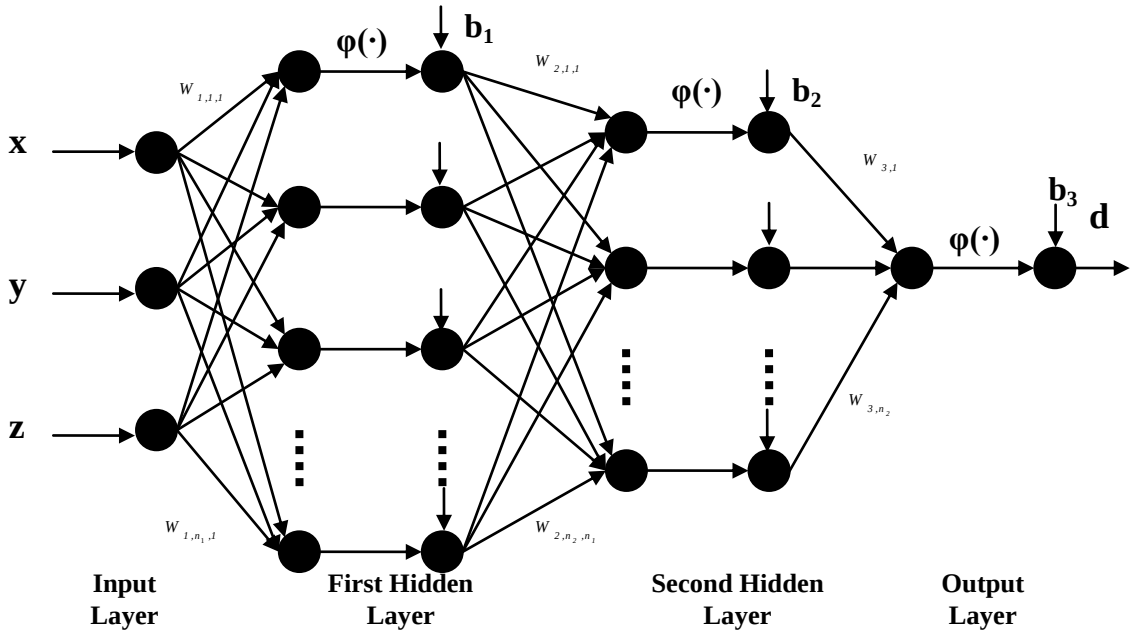



Figure 4.3. Network structure for surface function approximation. $i$ denotes the number of nodes in the first hidden layer; $j$ denotes the number of nodes in the second hidden layer.

Our network structure employs two hidden layers for the following reasons. Though a single hidden layer is adequate for surface function approximation, the number of neurons of the hidden layer will be very large due to the complexity of the surface [82]. The problem is that the neurons therein will tend to interact with each other globally, which makes it difficult to improve the approximation at one point without worsening it at some other point. Consequently the calculation time of the cost function will be greatly increased and it will eventually affect the registration time. On the other hand, with two hidden layers, the approximation (curve-fitting) process becomes more manageable [83]. To achieve similar training accuracy, the learning time of the network constructed from two hidden layers is usually much shorter than

that from a single hidden layer. Furthermore, two hidden layers are usually sufficient in practice [84].

It has to be noted that there is no ideal number of neurons for every problem, and unless prior information is available of the problem, the fine-tuning of the number of neurons involves a rather qualitative than analytical approach, and is to be expected from system identification problems [85], such as the variation in bone structures. A feasible method is to search for the optimal number in each layer by experimenting on test datasets. Since the number of layers and the number of nodes represent the complexity of a surface function, these can be determined using a test dataset such as the Visible Human Dataset (VHD). For example, if it is obtained empirically that a 2-layer NN with 20 neurons for the first layer and 10 neurons for the second is optimum for the lumbar spine of the VHD, this setup can be assumed as optimum for the lumbar spine of the incoming patient. One common rule is that the number of neurons of the second layer cannot be more than that of the first layer because local features modeled by the first layer are usually more complex than the global features modeled by the second layer.

*3) Derivation of surface representation function*: From the above network structure definition, the surface representation function can be derived as:

$$\sum_{j=1}^{n_2} (W_{3,j} \tanh( \sum_{i=1}^{n_1} (W_{2,j,i} \times \tanh( W_{1,i,1} x + W_{1,i,2} y + W_{1,i,3} z + b_{1,j})) + b_{2,j})) + b_3 = 0 . \quad (4.12)$$

In the equation, $n_1$ and $n_2$ are the respective number of neurons of the first hidden layer and the second hidden layer, $W_{1,i,j}$ the weights between the $jth$ node of the input layer and the $ith$ neuron of the first hidden layer, $W_{2,j,i}$ the weights between the $ith$ neuron of the first hidden layer and the $jth$ neuron of the second hidden

layer, $w_{3,j}$ the weights between the $jth$ neuron of the second hidden layer and the output layer, and $\mathbf{b}_1$, $\mathbf{b}_2$, $b_3$ the respective biases for the first hidden layer, the second hidden layer and the output layer.

*(c) Optimization*

Using the start point obtained by coarse registration, fine registration is obtained by minimizing the cost function. A standard optimization method, the downhill simplex method [40], is used for this optimization to calculate the final rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$.

## 4.3 Experiments

In this section, we describe the experiments and the datasets used to evaluate our registration algorithm.

### 4.3.1 Datasets

Two medical datasets that of a spine and a calcaneus, are used for testing the algorithm.

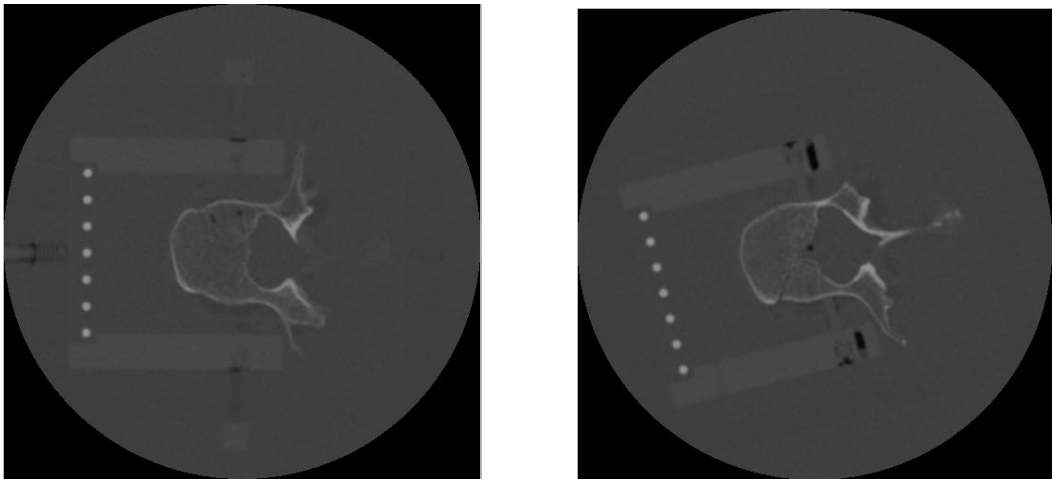
*1) Spine dataset*: This dataset is obtained from the NASA Ames Research Center. A section of two vertebral bodies are firmly enclosed in a cage with linear structures. The entire structure is submerged in a cylindrical water bath and scanned using a GE HiSpeed CT/i system at three different orientations to give the datasets SA, SB and SC. The scanning parameters are: helical scan, pitch 1, detector width 1mm, FOV 170mm, reconstructed at 1mm interval, pixel spacing 0.332mm/0.332mm. Figure 4.4 shows a slice from both SA and SB.

*2) Calcaneus dataset*: This dataset is also from NASA Ames Research Center. A calcaneus bone is firmly attached to a cage with linear structures and submerged in a water bin. All scans were taken at the following CT scanner settings: 120 kVp, 190 mA, reconstructed slice thickness 0.5 mm, pixel spacing 0.352mm/0.352mm. Scanning was done using a GE HiSpeed CT/i system at five different orientations to give the datasets denoted by CA, CB, CC, CD and CE. Figure 4.5 shows two corresponding slices from CA and CB.

### 4.3.2 Experiment Design

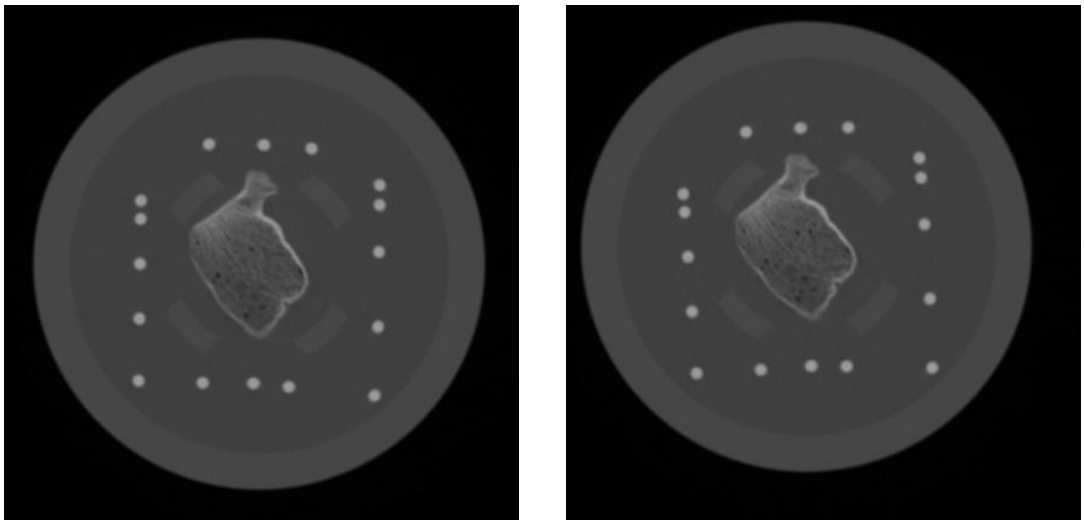
*1) Segmentation*: In the experiment, we first extract the surfaces of the calcaneus in CA, CB, CC, CD and CE, and the surfaces of the spine in SA, SB and SC using the segmentation algorithm described in Chapter 3. Figures 4.6(a), (c) and (e) show the extracted surfaces of CA, SA and SB, respectively.

*2) Neural network modeling*: The extracted surfaces from CA, SA (two surfaces) and SB (two surfaces) are modeled using neural networks. A 2-layer NN (first hidden layer: 20 nodes, second hidden layer: 10 nodes) is use to model the vertebral surface. Training process stops when the cumulative mean square error between the NN output and distance $d(x, y, z)$ is less than 0.0001 of the total number of training points or the error does not decrease for 1000 successive iterations. This network is also used to model the calcaneus dataset. The NN modeled surface is used to register the corresponding surfaces from other datasets using the algorithm described in section 4.3. The extracted surfaces from SB and SC are registered to SA, and CB, CC, CD and CE are registered to CA.



(a) (b)

Figure 4.4. Original images from different spine datasets. (a) 38th slice of SA. (b) 38th slice of SB.



(a) (b)

Figure 4.5. Original images from different calcaneus datasets. (a) 90th slice of CA. (b) 90th slice of CB.

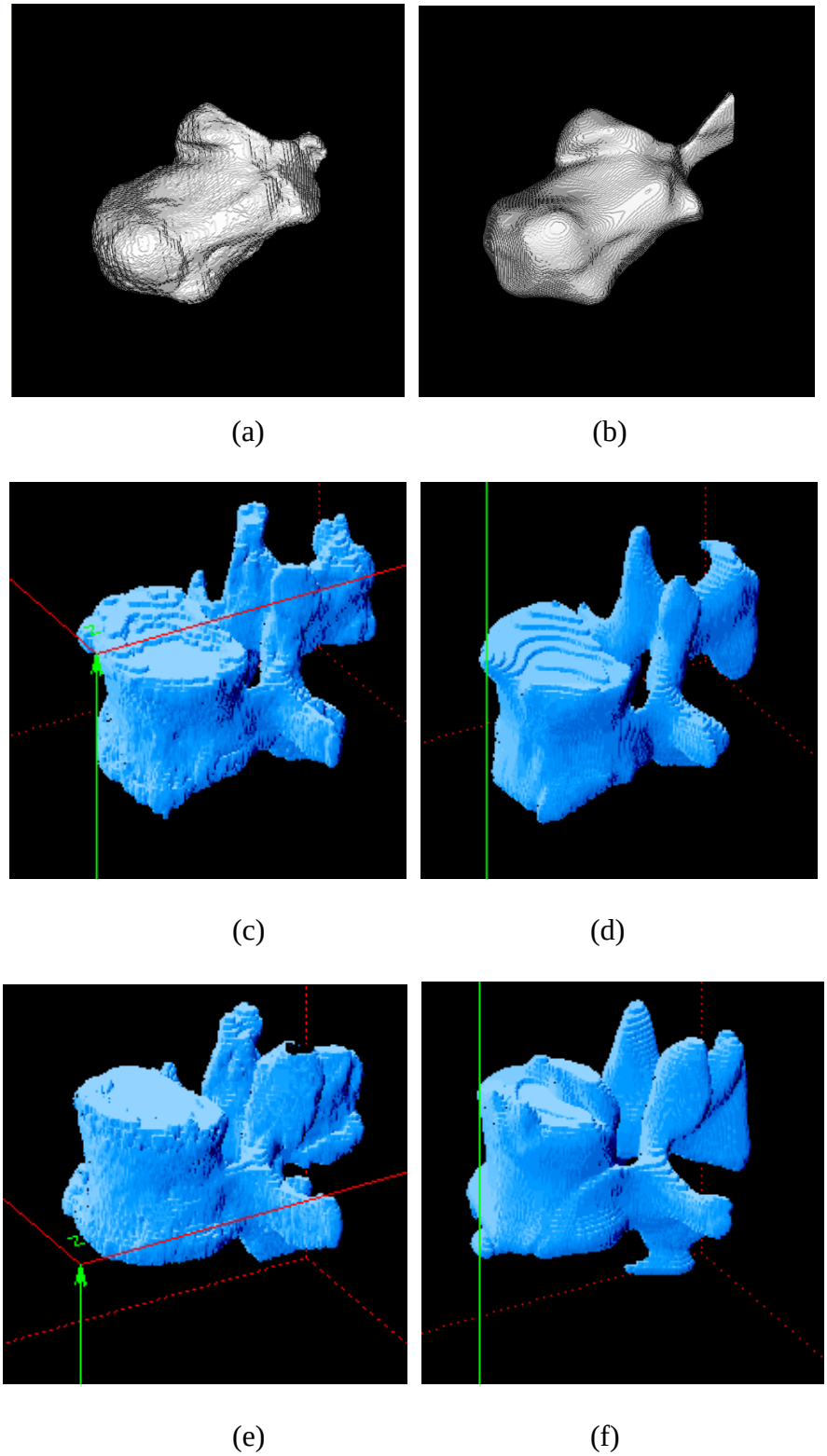



Figure 4.6. Surface modeling results. (a) CA (c) SA-V1 (e) SB-V1: Extracted surface. (b) CA (d) SA-V1 (f) SB-V1: NN surface model.

*3) Registration accuracy*: Execution time and registration accuracy are the two important specifications for a registration system. To assess the accuracy, we compare our results with that of a frame-based registration method [86] and the commonly used ICP surface registration technique [87]. We implement the frame-based registration algorithm [85] and use it to register CB, CC, CD and CE to CA. Similarly, we use our registration module to register CB, CC, CD and CE to CA. The results from the frame-based registrations are extremely accurate [85] and hence can be used as the ground truth to assess the accuracy of our registration system. SB and SC are registered to SA using our system and ICP methods. Similarly, we use frame-based registration to register SB and SC to SA and use the results to assess the accuracy of our registration system and ICP algorithm.

In [85], with a reference rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$, for any transformation ($\hat{\mathbf{R}}$, $\hat{\mathbf{t}}$), the largest absolute eigenvalue of $(\hat{\mathbf{R}} - \mathbf{R})$ is denoted by $\varepsilon_R$ and $\|\hat{\mathbf{t}} - \mathbf{t}\|_2$ is denoted by $\varepsilon_T$. The translation error has a constant value of $\varepsilon_T$ throughout the volume of interest (VOI) while the rotational error varies with the distance from the center of rotation and its value is bounded by $\varepsilon_R \|\mathbf{P}\|_2$, where $\|\mathbf{P}\|_2$ is the distance between the point of assessment and the center of the frame. With the availability of $\varepsilon_R$ and $\varepsilon_T$, we can calculate the upper bound of the registration error for each point in any desired VOI. We can further determine the upper bound of the average or maximum registration error over the VOI. Thus $\varepsilon_R$ and $\varepsilon_T$ are used as the error measures to evaluate our registration system and the ICP algorithm.

*4) Speed*: To compare the registration speed of our NN method and the ICP method, we record the execution times. All the procedures, including cost function

calculation, surface modeling and optimization, are implemented in C++ and executed on a dual XEON 3.06 GHz Pentium computer with a memory size of 2GB. Intel IPP libraries [88] are used whenever possible to shorten the execution time.

## 4.4 Results and Discussion

In this section, we present the registration results of our proposed algorithm on the CT spine and calcaneus datasets.

Figure 4.6 shows the surface modeling results of CA, SA and SB. The registration results shown below are based on these models. We can see that the NN representations of the surface models are fairly similar to the extracted surfaces but with additional smoothing. The surface modeling time is listed in Table 4.1. It takes about 2 hours to train each dataset's surface using about 40,000 training points. The accuracy of the surface model is evaluated using the following method. The real surface points are used as input of the NN. We use the average output of NN to evaluate the NN model. Ideally, the output should be zero since all the surface points have zero output in the training dataset. We note that the average cost for each dataset is less than 0.1, which means the NN model is less than 1 voxel away from the real CT surface.

Table 4.1. Surface modelling results.

| Dataset | CA | SA-V1 | SA-V2 | SB-V1 | SB-V2 |
|---|---|---|---|---|---|
| Modeling time (s) | 8743 | 7200 | 7213 | 7210 | 7209 |
| Average cost (pixel) | 0.015 | 0.053 | 0.034 | 0.042 | 0.065 |

Table 4.2 shows that in the calcaneus registration experiment, the results of the NN-based registration are accurate when compared to the frame based method. The

maximum VOI error between the two methods is less than 0.5mm. This is smaller than the CT slice thickness, which means that our registration method achieves sub-voxel accuracy. Table 4.3 compares the results obtained with NN-based and ICP registration methods. The results of the two are similar, with the difference less than 0.1mm. The registration accuracy map for a particular slice in SB for the SB to SA registration experiment is shown in Figure 4.7.

From the error map (Figure 4.7), we notice that the center region has the smallest registration error, below 0.15 mm. This is because the displacement error due to the rotation error will increase with the distance from the rotation center. From Table 4.3, we further note that the execution time of the NN-based method (1 min) is much shorter than that from ICP (15 mins), due mainly to the fact that the former requires no searching of point pair correspondence. Since the ICP method is a popular surface-based method that has been proven to be very accurate, we can conclude that the NN-based method is not only very accurate but is also extremely fast.

In short, the experimental results demonstrate that the proposed NN-based method can register intra-operative data with pre-operative data efficiently and accurately.

## 4.5 Conclusion

We have described a 3-D surface-based rigid registration system for image-guided surgery on bone structures. This system includes near-automatic segmentation for both preoperative and intra-operative scans and automatic real-time registration for intra-operative scans. The segmentation algorithm is used to extract the 3-D bone surfaces for both preoperative and intra-operative scans. A novel automatic

surface-based method using a neural network is then used to perform intra-operative registration. We use the NN to construct an invariant descriptor for human bone to speed up the registration process. Significantly improved computational efficiency is apparent with the NN-based approach.

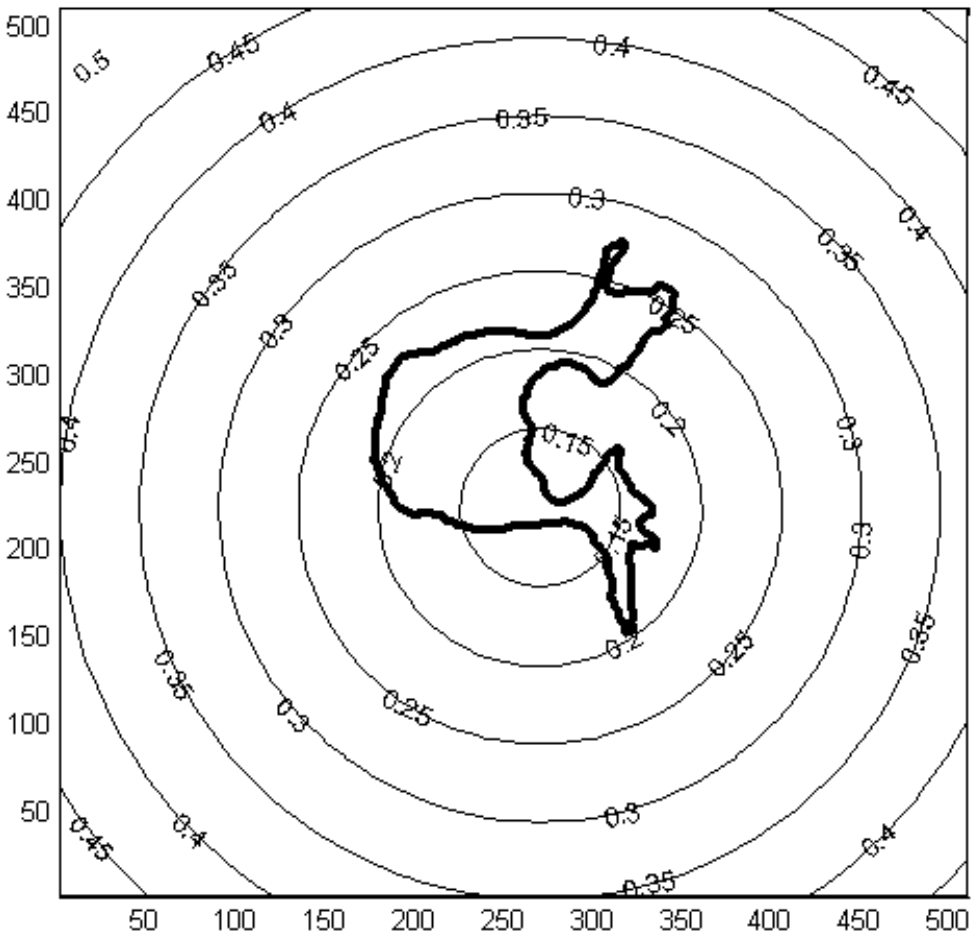



Figure 4.7. Registration error map of one slice from SB in registering SB to SA using V1.

Table 4.2. Calcaneus comparison results with frame-based registration (reference dataset is CA).

| Dataset | $\varepsilon_R$ (radian) | $\varepsilon_T$ (mm) | Maximum VOI Error (mm) |
|---|---|---|---|
| CA- CB | 0.0033 | 0.2301 | 0.3806 |
| CA- CC | 0.0040 | 0.2437 | 0.4270 |
| CA- CD | 0.0019 | 0.1193 | 0.2045 |
| CA- CE | 0.0032 | 0.3434 | 0.4898 |

Table 4.3. Full surface registration accuracy results and execution time of spine datasets (reference dataset is SA, V1 is the first vertebra and V2 the second vertebra).

| Datasets | Neural network based | | | | ICP | | | |
|---|---|---|---|---|---|---|---|---|
| | Accuracy | | | Time | Accuracy | | | Time |
| | $\varepsilon_R$ (radian) | $\varepsilon_T$ (mm) | Max VOI Error (mm) | (s) | $\varepsilon_R$ (radian) | $\varepsilon_T$ (mm) | Max VOI Error (mm) | (s) |
| SA- SB, V1 | 0.0034 | 0.0956 | 0.2542 | 47.17 | 0.0041 | 0.0550 | 0.2438 | 856.45 |
| SA- SB, V2 | 0.0031 | 0.0816 | 0.2376 | 36.66 | 0.0029 | 0.1446 | 0.2923 | 745.13 |
| SA- SC, V1 | 0.0062 | 0.1413 | 0.4320 | 61.20 | 0.0031 | 0.2195 | 0.3648 | 814.70 |
| SA- SC, V2 | 0.0017 | 0.5231 | 0.6113 | 56.26 | 0.0043 | 0.3215 | 0.5390 | 987.32 |
| SB- SC, V1 | 0.0012 | 0.3666 | 0.4247 | 54.48 | 0.0028 | 0.2177 | 0.3497 | 830.02 |
| SB-SC, V2 | 0.0013 | 0.2776 | 0.3451 | 59.53 | 0.0033 | 0.2682 | 0.4371 | 845.13 |

We applied the methods to perform several registrations on CT/CT calcaneus and spine images. We have shown that our registration algorithm is as accurate as the commonly used ICP technique. Both are able to achieve sub-pixel registration accuracy. However, our registration process is about 15 times faster than the ICP technique. The execution time for our registration process is about 1min, which is much shorter than those using standard techniques. A partial volume registration method [89] could be used to further shorten execution time of the whole registration procedure. Also, the methods presented are well placed to be implemented in hardware (FPGA). These may provide a good solution for intra-operative rigid registration in image-guided surgery systems.

# 5 Iterative Weighted CT/MR Image Registration

## 5.1 Introduction

X-ray and CT provide well-contrasted images of high-density biological objects such as bones and tumors but are usually not preferred for detailed soft tissue examination. MR imaging, with its moderate resolution and good signal-to-noise ratio is the modality of choice for soft tissues. Fusing CT and MR images will help overcome the limitation of relying on a single modality for image guided surgery. A typical fusion procedure comprises segmentation of the CT and MR images, followed by registration and spatial alignment/fusion. The region of interest in CT images (e.g., bone) or MR images (e.g., kidney and liver) of a patient is first segmented. After spatial registration, the segmented CT and MR images are aligned to give a model comprising well-contrasted bone structure and the surrounding soft tissues. Such a composite model is important for surgical planning and education. For example, a vertebra, which is a hard tissue, may have to be examined with the intervertebral disc, a soft tissue, for effective spinal surgery planning. An important motivation of this work was the development of a patient-specific hybrid model of the spine for image guided spinal surgery although the techniques described here may also be employed for different anatomies, e.g., the ankle.

Prevailing methods to perform registration/fusion of different modalities incur heavy computational cost in searching for optimal transformation parameters (mutual information (MI) method) or require the input of extracted object surface (surface

based registration). Further details of surface-based registration can be found in [88]. However, it is not easy to extract object surface from MR images.

It is good to solve the CT/MR registration problem and MR image segmentation problem simultaneously. This Chapter describes our approach for simultaneously solving the problems of CT/MR registration and MR image segmentation. The algorithm can perform fast and accurate CT/MR feature-based registration, accurate extraction of the bone surface from MR images, and fast fusion of the two modalities. Since the bone surface in CT images can be extracted accurately [88], the segmented CT image is used as the reference for MR image segmentation. Our novel segmentation approach employs a shape-based adaptive level set to handle the fuzzy boundaries of the MR images. The iterative process starts with a coarse extraction of the bone surface from MR images, which is then registered to the accurate bone surface extracted from CT images. The CT bone surface, after spatial transformation and re-sampling, is used as the initial estimate for MR image segmentation. The new segmented MR image is subsequently registered to the CT bone surface. MR image segmentation is improved after each iterative step using the results of the registered CT segmentation. This iterative process converges when the MR and CT image segmentation results agree within a specified tolerance. In this iterative registration/segmentation process, only fine adjustments are needed since the target boundaries in MR images is close to the initial estimate, thus reducing computational time. Inaccurate segmentation due to poor scans or complex anatomies such as the vertebrae can be prevented.

## 5.2 Methods

### 5.2.1 Iterative Segmentation/Registration System

Figure 5.1 shows the flowchart of the proposed CT/MR registration system. It comprises the following components: initial segmentation of CT and MR images, iterative CT/MR registration and refinement of MR image segmentation. Initial segmentation is first performed on CT images to separate the region of interest (bone) from its surroundings [90]. The bone surface is then identified and used in registration. It is clear that MR images, with their inherent low signal-to-noise ratio, poor contrast and fuzzy boundaries are unlikely to be segmented accurately in a single step. The first segmentation step captures the general shape of the target object (the vertebrae). A coarse registration result is obtained by registering the MR and CT surfaces with a weighted surface-based registration algorithm. With the registered CT surface model as the reference, we use the intermediate results of MR image segmentation and registration to iteratively refine the suboptimal MR image segmentation. This iterative process is carried out until the segmented CT and MR surfaces match within a specified tolerance.

### 5.2.2 MR Image Segmentation

We propose the double-front level set for fast segmentation of MR datasets. The level set is a time evolving function, and is the so-called “zero level curve” corresponding to a propagating front. It is a simple and versatile method of computing and analyzing the motion of an interface $\Gamma$ in two or three dimensions.

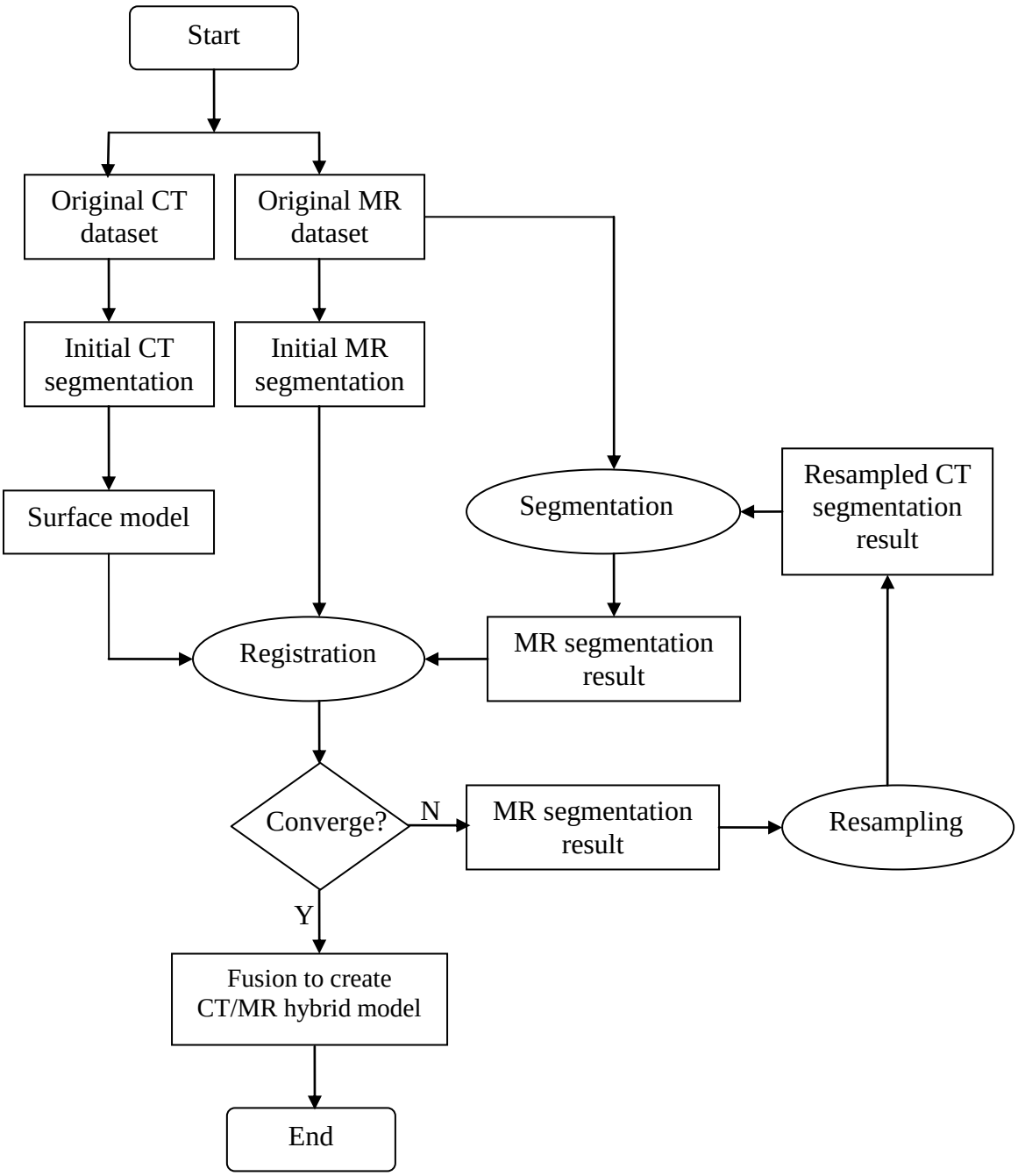



Figure 5.1. Flowchart of feedback segmentation-registration.

This method can deal with convoluted shapes and sharp corners but is not capable of bi-directional growing, i.e., when the expanding front exits the target boundary, it may not be able to “shrink back”. The single-front level set is thus prone to leak into the background at a fuzzy boundary [23]. The idea of gradient vector flow was proposed to help overcome the problem [91] but it does not always lead to a satisfactory solution.

In our method, the level sets are bi-directional since they can either expand or shrink. In bi-directional propagation, a “balloon” force, together with the velocity field derived from image intensity, prevents the front from being trapped in local minima. The vertebral boundaries in MR images are often fuzzy. With single-direction propagation, the front is likely to leak beyond the target boundary. It is difficult to determine the correct magnitude of the balloon force for the bi-directional level set; a small force may lead to trapping in local minima, while a large force could give rise to leakage. Our proposed solution is to extend the single-front level set to a double-front level set (DFLS), which comprises a pair of one-directional propagating single-front level sets with one shrinking and the other expanding. The two fronts prevent each other from intersecting, hence minimizing the leakage.

DFLS can be regarded as a pair of level sets, one propagating forwards and the other backwards. Ideally, the final boundary will be the inter section of both level sets. Leakages are prevented since the back-propagating level set $\varphi_B$ and the forward-propagating level set $\varphi_F$ only meet at the final boundary. In other words, when the single-front level set cannot find the boundary position at a fuzzy edge, the boundary position defined by DFLS is the location where the back-propagating front meets the forward-propagating front.

With a velocity field $\mathbf{v}$, the level set will vary with time as [23]

$$\frac{\partial \varphi}{\partial t} + \mathbf{v} \cdot \nabla \varphi = 0 . \tag{5.1}$$

For the double-front level set, we have

$$\frac{\partial \varphi_F}{\partial t} + v_N \, | \nabla \varphi_F | = 0 \tag{5.2}$$

and

$$\frac{\partial \varphi_B}{\partial t} - v_N \, | \nabla \varphi_B | = 0 , \tag{5.3}$$

where the level set functions $\varphi_B$ and $\varphi_F$ are time-variant $N$ dimensional hyper surfaces and $v_N = \mathbf{v} \cdot \frac{\nabla \varphi}{| \nabla \varphi |}$ is a function of the direction of the unit normal. The two level sets share the same velocity field with opposite signs. The solution to MR image segmentation is the cross section of the two hyper surfaces.

In our MR image segmentation method, the stopping criterion is determined via the estimation of the distribution of the gradient values at edge points between hard and soft tissues. A small number of regions are first selected at the bone/tissue interface. The gradient at an edge point depends on the type of tissues on either side of the interface. For example, the gradient between bone and muscle is different from that between bone and ligament. We would like the level set function to stop propagating once it reaches the edge of the bone; thus, it is important that we capture the statistics of the edge gradient at various different bone/tissue interfaces. We observe that there are four main types of interfaces. We compute the gradient values in four representative regions and use K-means clustering to classify the points into two groups, edge points and non-edge points, according to these values. From the mean

$\mu_e$ and standard deviation $\delta_e$ of the edge points, we compute the threshold $\tau = \mu_e + \delta_e$. The level set stops growing at a point whose gradient value exceeds $\tau$.

Assume that we have an initial estimated surface. Several subsequent iterations of dilation and erosion will result in the creation of a pair of distance iso-surfaces. The number of dilation and erosion iterations is determined by the average distance between the segmented contours from CT images and the desired boundary of the MR scan. To ensure a reasonable estimation of the boundary in fuzzy areas, the initial contours need to be approximately the same distance to the goal boundary. Dilation and erosion of the binary CT contours can yield the initial fronts. The same number of iterations of dilation and erosion are typically performed. By adjusting the velocity field, the back-propagating level set $\varphi_B$ and the forward-propagating level set $\varphi_F$ will meet at the original CT surface.

The numerical approximation of the level set equation with curvature dependant velocity is

$$\max^2(\max(D_{ij}^{-x}\varphi_F), -\min(D_{ij}^{+x}\varphi_F)) + \max^2(\max(D_{ij}^{-y}\varphi_F), -\min(D_{ij}^{+y}\varphi_F)) = \frac{1}{v_{ij}^2}$$

From the numerical approximations derived in [92], our iterative solution for the forward-propagating and back-propagating double-front level sets are

$$\max^2(\max(D_{ij}^{-x}\varphi_F), -\min(D_{ij}^{+x}\varphi_F)) + \max^2(\max(D_{ij}^{-y}\varphi_F), -\min(D_{ij}^{+y}\varphi_F)) = \frac{1}{v_{ij}^2} \quad (5.4)$$

and

$$\max^2(\max(D_{ij}^{-x}\varphi_B), -\min(D_{ij}^{+x}\varphi_B)) + \max^2(\max(D_{ij}^{-y}\varphi_B), -\min(D_{ij}^{+y}\varphi_B)) = \frac{1}{v_{ij}^2}, \quad (5.5)$$

where

$$D_{ij}^{+x}\varphi = (\varphi_{i+1,j} - \varphi_{i,j}) / \Delta x ,$$

$$D_{ij}^{-x}\varphi = (\varphi_{i,j} - \varphi_{i-1,j}) / \Delta x ,$$

$$D_{ij}^{+y}\varphi = (\varphi_{i,j+1} - \varphi_{i,j}) / \Delta y ,$$

$$D_{ij}^{-y}\varphi = (\varphi_{i,j} - \varphi_{i,j-1}) / \Delta y . \qquad (5.6)$$

In the equations 5.4 and 5.5, $v_{ij}$ is velocity derived from the two-space dimension discrete approximation to the Hamilton-Jacobi equations and $\Delta x$ and $\Delta y$ respectively are the changes over x and y axis [91]. Only those image points between the two initial contours are processed. The computing cost is significantly reduced. Since clinical MR images have large inter-slice distances, 3D MR image segmentation [93] has the drawback of heavy computational cost and produces similar results as 2D segmentation. Thus, MR image segmentation is performed in 2D in our implementation.

### 5.2.3 Weighted Surface Registration

In the initial MR image segmentation with the single level set, some edge points are more reliable than others. The registration process can be guided by assigning more priority to the reliable edges and less to unreliable ones. DFLS yields a weighted segmentation result in which weights are derived from both intensity gradient and position.

Maurer *et al*. [94] used weighted geometrical features to register CT images of the head to physical spaces. In our method, the weight of each edge point is determined

according to its gradient value and spatial location. We assign discrete weights of 1, 2 or 3 to edge points according to their gradient values. The weights of the inner front points are given greater importance relative to the outer front points by an additional increment of 3. This helps to ensure accurate registration since the outer front may stop incorrectly at the soft tissue around the bone instead of at the bone surface.

The purpose of CT/MR registration is to align the vertebral bodies in MR and CT images. Letting $D$ denote the distance of a point in MR images to the reference surface in CT images, $w_c$ the weight of $\mathbf{p}_c$ the point of the current vertebral edge in MR images, and $\mathbf{p}_d$ the closest corresponding points in CT images, the cost function can be represented as

$$f(\mathbf{p}_c) = \sum_{\mathbf{p}_c} w_c D(\mathbf{p}_c, \mathbf{p}_d) . \tag{5.7}$$

We implemented a neural network based approach [88] to solve this surface-based registration problem. A downhill simplex optimization technique is used to locate the minimum of this derived cost function:

$$f(\mathbf{p}_c) = \sum_{\mathbf{p}_c} w_c D(\mathbf{p}_c) . \tag{5.8}$$

### 5.2.4 Iterative Segmentation/Registration

The iterative process is shown in Figure 5.2. Two initial fronts or contours of MR image segmentation are obtained from the dilation and erosion of the binary image obtained after CT segmentation. With the weighted surface obtained from the DFLS segmentation, a weighted surface-based registration is performed with the CT surface model. The CT segmentation result is then transformed and re-sampled to fit

the MR image specification. The contours of the re-sampled CT segmentation are used as the initial reference to redo the segmentation of the original MR image. This iterative process of segmentation and registration is repeated until the registration the cost function converges to less than a predefined small value $\delta$ , i.e., $f(\mathbf{p}_c) < \delta$ .

The registered CT/MR images are used to construct the CT/MR hybrid model. Knowing the correspondence between CT and MR images obtained by image registration, the CT segmentation result can be transformed to the MR image. The vertebral volume delineated in the above process is replaced by the transformed CT segmentation result. This model comprises the bone structure from CT images and soft tissues from MR images. The fused images provide detailed information of both soft and hard tissues, unlike the image from a single modality.

## 5.3 Experiments

In this section, we describe the datasets and experiments for evaluating the registration/segmentation system.

### 5.3.1 Dataset

Six pairs of MR/CT images are used in the experiments: three human spines, one human ankle and two pig spines. The details of the datasets are shown in Tables 5.1 and 5.2. The MR datasets were imaged with the T2 sequence.

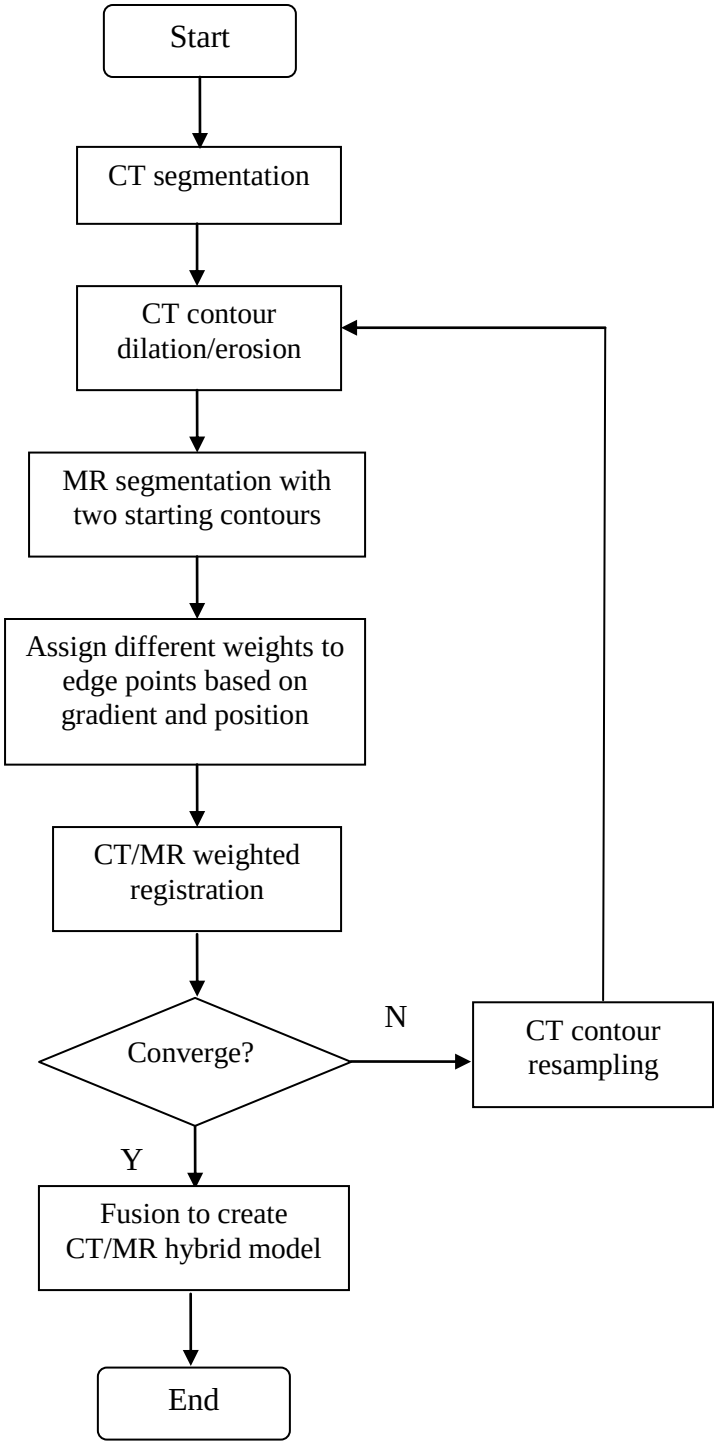



Figure 5.2. Flowchart of iterative segmentation/registration.

Two human spine MR images are shown in Figures 5.3(a) and 5.3(b). They comprise three lumbar vertebrae, L2, L3 and L4, from mid-L2 to mid-L4. The first human specimen is an 80 year old man with damaged vertebrae. There are 90 CT images and 12 MR images with slice gaps of 1mm and 7mm, respectively. The second patient has a curved spine. The three human spine MR datasets are of low contrast. In general, there is an intensity difference between the bone and surrounding tissues, but the boundaries are very fuzzy and poorly defined.

Table 5.1. Datasets used in the experiments.

| **Dataset** | **CT** | **MR** |
|---|---|---|
| Human spine 1 | HS1_CT | HS1_MR |
| Human spine 2 | HS2_CT | HS2_MR |
| Human spine 3 | HS3_CT | HS3_MR |
| Human ankle | HA_CT | HA_MR |
| Pig spine 1 | PS1_CT | PS1_MR |
| Pig spine 2 | PS2_CT | PS2_MR |

The MR dataset HA_MR is a scan of a human ankle Figure 5.3(c). This dataset is of much better visual quality than the previous two, but there is a very large rotation between the CT and MR images.

### 5.3.2 Experimental Design

1) *Segmentation and modeling*: In this experiment, we use the segmentation algorithm described in Chapter 3 to first extract the surface of the ankle bone in HA_CT and the surfaces of the vertebrae from the CT images of the three human spine and two pig spines. The extracted surfaces from these datasets are modeled

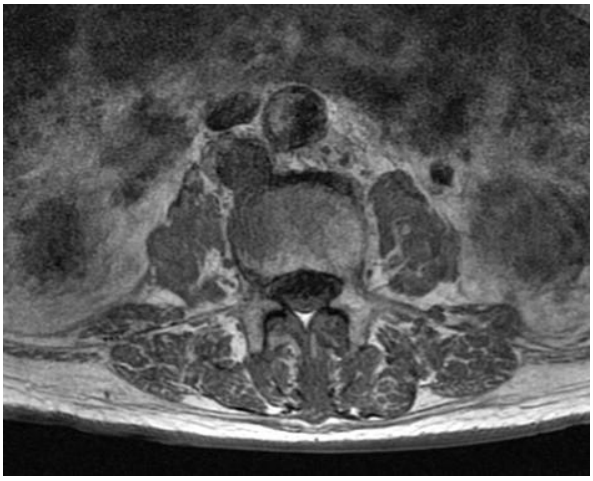
using neural networks. Since two spine images cannot be matched by rigid registration, one vertebra is chosen from each of the CT and MR datasets for segmentation, registration and fusion. For example, L3 of the third patient is used in the experiment.

2) *Weighted registration and resampling*: MR images are segmented using DFLS. The extracted surfaces from HS1_MR, HS2_MR, HS1_MR, PS1_MR, PS2_MR and HA_MR are registered to HS1_CT, HS2_CT, HS1_CT, PS1_CT, PS2_CT and HA_CT, respectively, using the weighted registration method described in Section 3. After registration, the CT contour is used to generate the initial contours for MR image segmentation with DFLS. This process is iterated until convergence. The stopping criteria are: translation in each direction is less than 0.01mm, and rotation of the x, y, and z axes is less than $0.01^{\circ}$.

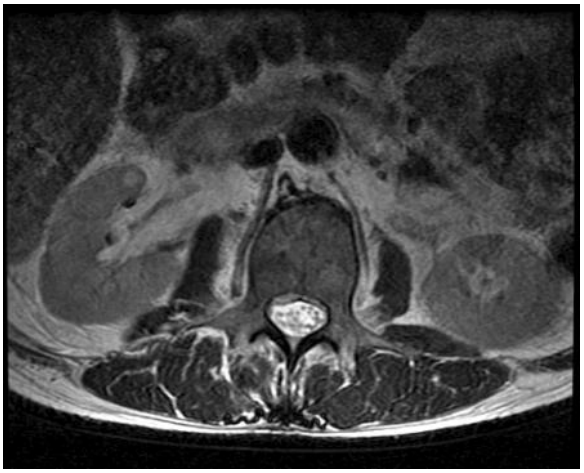
*3) Segmentation and registration accuracy:* We assess the accuracy of the segmentation in two ways. First, we compare the segmentation result with the manual segmentation performed by an experienced radiologist. Second, the normalized error for the converged segmentation result registered to the CT model is evaluated. The cost of registering the CT surface to the CT model is used as the reference, where cost is defined as the output of the NN from the input transformed point coordinates. This error is intrinsic.

Table 5.2. Dataset specifications.

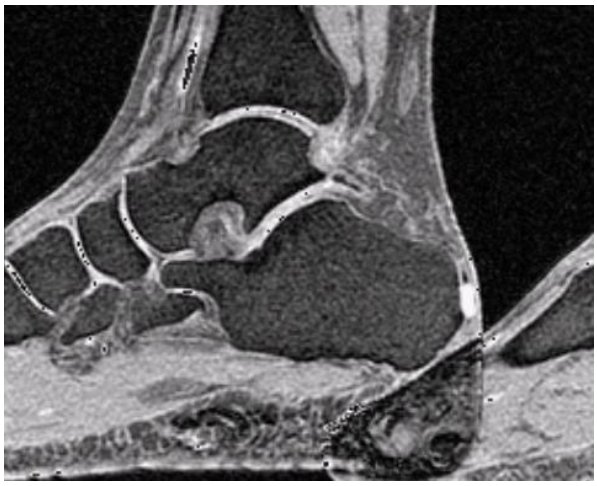
| Dataset | Object | Slices | X-resolution (mm) | Y-resolution (mm) | Z-resolution (mm) |
|---|---|---|---|---|---|
| HS1_CT | Human spine | 90 | 0.313 | 0.313 | 1 |
| HS1_MR | Human spine | 12 | 0.39 | 0.39 | 7 |
| HS2_CT | Human spine | 96 | 0.652 | 0.652 | 1 |
| HS2_MR | Human spine | 20 | 0.39 | 0.39 | 5 |
| HS3_CT | Human spine | 57 | 0.273 | 0.273 | 4 |
| HS3_MR | Human spine | 16 | 0.39 | 0.39 | 5 |
| HA_CT | Human ankle | 100 | 0.41 | 0.41 | 0.5 |
| HA_MR | Human ankle | 76 | 0.703 | 0.703 | 0.8 |
| PS1_CT | Pig spine | 100 | 0.293 | 0.293 | 1 |
| PS1_MR | Pig spine | 30 | 0.547 | 0.547 | 3 |
| PS2_CT | Pig spine | 100 | 0.391 | 0.391 | 0.8 |
| PS2_MR | Pig spine | 50 | 0.391 | 0.391 | 0.7 |



(a)



(b)



(c)

Figure 5.3. Experimental datasets: (a) HS1_MR, (b) HS2_MR, (c) HA_MR.

Execution time and registration accuracy are two important performance indicators for a registration system. We compare our results with those of the commonly used ICP surface registration technique [44] and normalized MI (NMI) [95]. With the manual segmentation performed by an experienced radiologist, we obtain the segmented MR image and registered it with pre-segmented CT image using ICP. The NMI method implemented in our previous work [94] is used for CT/MR registration. The volumetric overlap of these three registrations are measured by three metrics, the HD, mean distance between surfaces and DICE similarity coefficient [96]. The DICE similarity coefficient, $s$ , is given by

$$S = \frac{2\left|H(A) \cap H(B)\right|}{\left|H(A)\right| + \left|H(B)\right|}, \tag{5.9}$$

where $H(A)$ represents the entropy of the CT image dataset and $H(B)$ represents the entropy of the MR image dataset.

## 5.4 Results and Discussion

In this section, we present and discuss the segmentation and registration results. The algorithm was implemented using Microsoft Visual C++ on a Pentium 4 (3.2 GHz) workstation with 2GB memory. Table 5.3 shows the computational time for CT/MR registration and the MR image segmentation of each dataset. Since HA_MR has a larger rotation angle than that of the other datasets, more iterations are required for convergence.

*Dataset HA_MR* – In this experiment, there are eight iterations of segmentation and registration prior to convergence. This is because a large rotation (nearly $90^\circ$) is required to align the MR and CT images. We observed from Figures 5.4(a)-(f) that the segmentation result helps the registration process to match the two datasets in the first three iterations. When the MR and CT images are nearly matched, the registration step helps to improve the MR image segmentation.

*Dataset HS1_MR* – In this experiment, there are four iterations of segmentation and registration prior to convergence. The final result is shown in Figure 5.5(a). The CT contour matches the resampled MR image perfectly.

*Dataset PS1_MR* – Only two iterations of segmentation and registration are required before convergence. The result is shown in Figure 5.5(b). Smoother boundaries are obtained after the second iteration.

From our experiments with the MR datasets, segmentation accuracy clearly depends on the quality of the original images. When we compare automated segmentation and the manual segmentation performed by an experienced radiologist, automated segmentation of HA_MR gives the best result with a highly creditable maximum error of only 1 voxel (0.7 mm). The other five pairs of datasets have maximum error of 2 voxels (0.8 mm). The segmented MR datasets of HS1_MR, HS2_MR and HS3_MR were resampled and fused with the original CT dataset with an error below 1 mm. Figure 5.6(a) shows the axial, sagittal and coronal views of the fused CT/MR hybrid model of HS1_CT and HS1_MR and Figure 5.6(b) the axial view of the hybrid model

of HS2_CT and HS2_MR. The segmentation result of dataset PS1_MR is better than that of dataset PS2_MR since the boundaries in the former dataset are smoother.

Table 5.3. Registration/Segmentation time.

| **Dataset** | **Number of slices** | **Total time (s)** | **MR segmentation time (s)** | **CT/MR registration time (s)** | **Number of iterations** |
|---|---|---|---|---|---|
| HS1_MR | 12 | 116 | 98 | 18 | 6 |
| HS1_MR | 20 | 124 | 104 | 20 | 4 |
| HS1_MR | 16 | 119 | 100 | 19 | 5 |
| HA_MR | 76 | 364 | 253 | 110 | 8 |
| PS1_MR | 30 | 42 | 33 | 9 | 2 |
| PS2_MR | 50 | 225 | 190 | 35 | 3 |

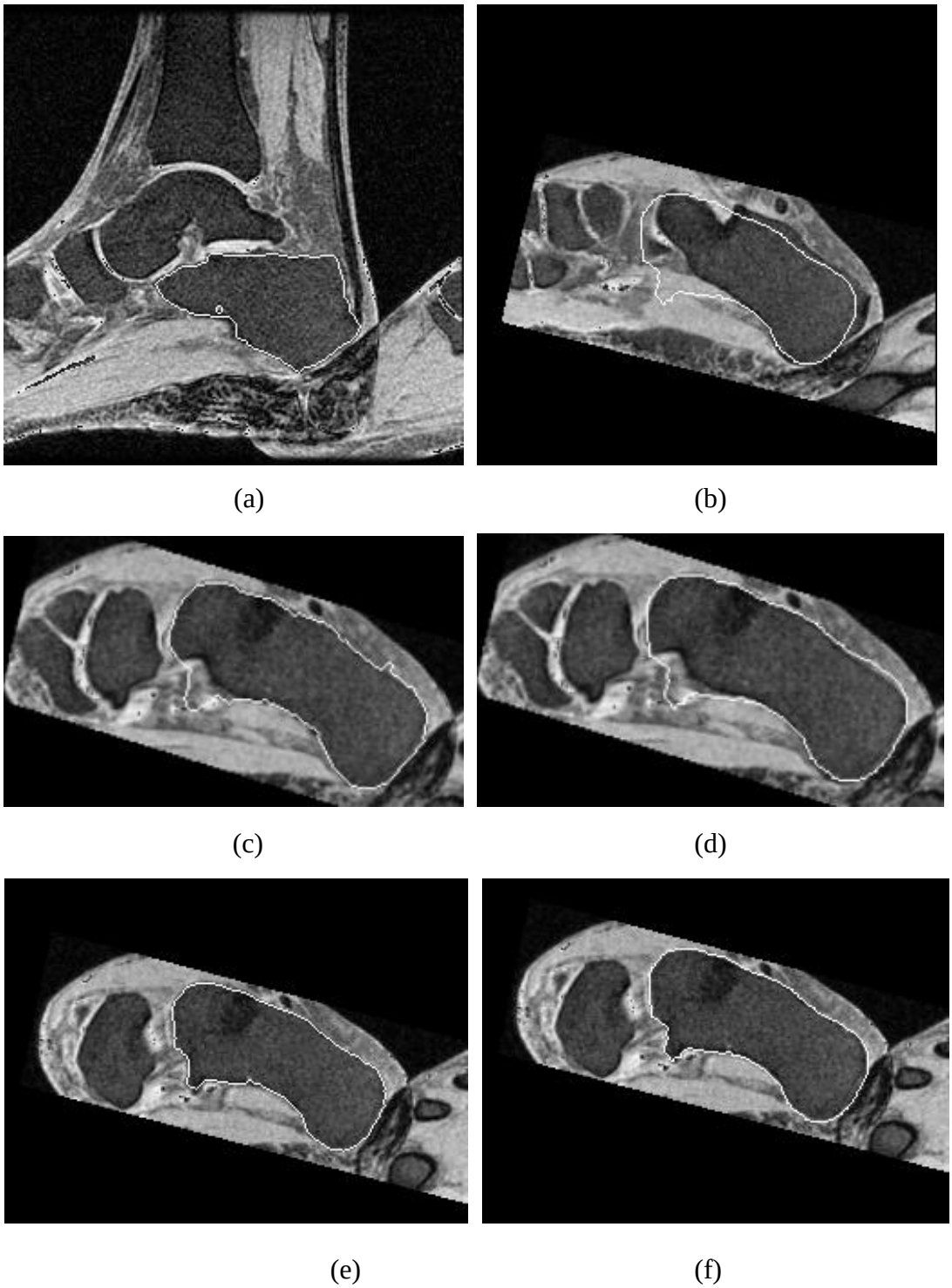
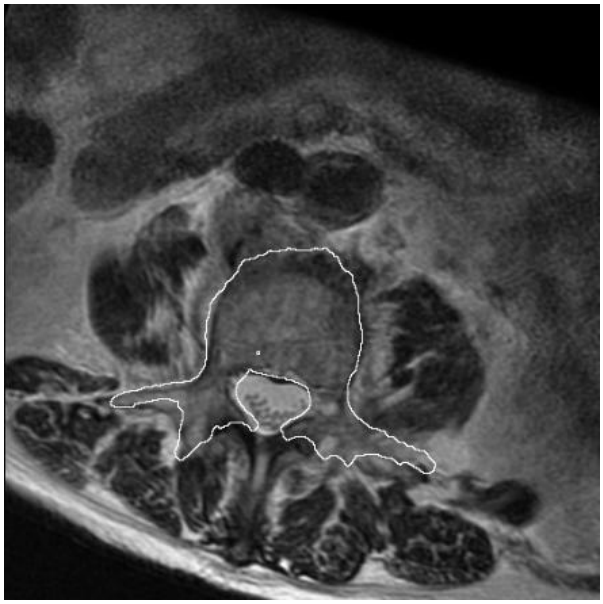
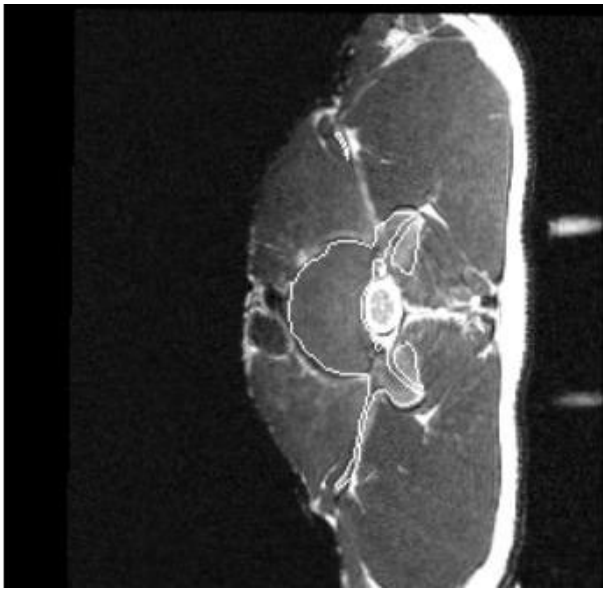



Figure 5.4. Experiment segmentation/registration results of dataset HA_MR: (a) 1st segmentation (b) 1st registration, (c) 3rd segmentation, (d) 3rd registration, (e) 8th segmentation, (f) 8th registration - converged.

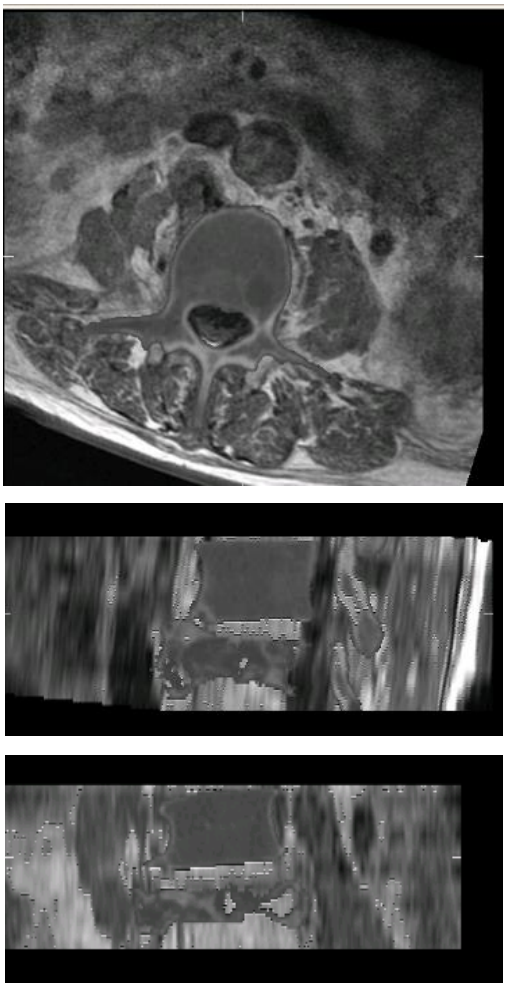


(a)

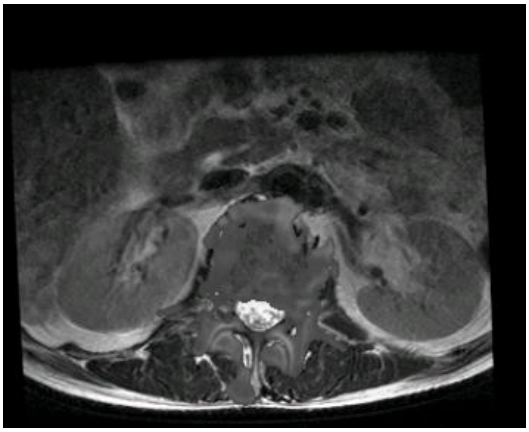


(b)

Figure 5.5. Converged registration results: (a) dataset HS1_MR, (b) dataset PS1_MR.



(a)



(b)

Figure 5.6. (a) Axial, sagittal and coronal views of the fused CT/MR hybrid model of a patient with cracked vertebrae. (b) Axial view of the fused CT/MR hybrid model of a patient with curved spine.

The normalized error for a converged segmentation result registered to the CT model is presented in Table 5.4. We note that the normalized error of the MR surface registered to the CT model ranges from 0.871 to 1.617 times the intrinsic error of the registration model. The NN model has an average error of 0.5 voxel, i.e., less than 1 voxel.

Table 5.4. Average cost after converging.

| | **PS1_CT** | **PS1_MR** | **PS2_CT** | **PS2_MR** | **HA_CT** | **HA_MR** |
|---|---|---|---|---|---|---|
| Number of points | 40834 | 25139 | 53676 | 44875 | 61465 | 14121 |
| Total cost | 2461.7 | 1319.9 | 2013.3 | 2997.1 | 933.4 | 347.1 |
| Average cost | 0.0603 | 0.0525 | 0.0375 | 0.0668 | 0.0152 | 0.0246 |
| Normalized error | 1 | 0.871 | 1 | 1.781 | 1 | 1.617 |

Table 5.5 compares the results obtained with the ICP registration method and NMI with our weighted registration method. The results are similar, with HD values within 0.2mm. The worst case of matching is 92%. We also note that the execution time of our method ($< 6$ min) is much shorter than that of traditional ICP (12 min) and NMI (1 hour). This would be mainly due to the fact that the former does not require any searching of point-pair correspondences. Since traditional ICP and NMI are popular registration methods that have been proven to be very accurate, we conclude that our weighted registration approach is not only accurate but also extremely fast.

Table 5.5. Execution time and volumetric overlap results.

| Dataset | Weighted Registration | | | | ICP | | | | NMI | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Total time (s)** | **Hausdorff distance (mm)** | **Mean distance (mm)** | **DICE** | **Total time (s)** | **Hausdorff distance (mm)** | **Mean distance (mm)** | **DICE** | **Total time (s)** | **Hausdorff distance (mm)** | **Mean distance (mm)** | **DICE** |
| HS1_MR | 116 | 0.25 | 0.07 | 97.7% | 809 | 0.29 | 0.09 | 96.7% | 3728 | 0.23 | 0.07 | 97.3% |
| HS1_MR | 124 | 0.49 | 0.14 | 98.9% | 864 | 0.41 | 0.13 | 99.0% | 4205 | 0.45 | 0.13 | 98.6% |
| HS1_MR | 119 | 0.33 | 0.11 | 95.8% | 821 | 0.36 | 0.12 | 95.3% | 3985 | 0.39 | 0.12 | 95.5% |
| HA_MR | 364 | 0.70 | 0.21 | 97.3% | 1445 | 0.75 | 0.23 | 96.6% | 6524 | 0.77 | 0.23 | 96.8% |
| PS1_MR | 42 | 0.51 | 0.16 | 94.5% | 882 | 0.62 | 0.17 | 94.7% | 4353 | 0.48 | 0.15 | 94.2% |
| PS2_MR | 225 | 0.78 | 0.22 | 93.6% | 979 | 0.83 | 0.23 | 93.2% | 5744 | 0.91 | 0.24 | 93.0% |

DFLS requires the input of two initial contours, the locations and positions of which will significantly influence the segmentation results. Our algorithm has error correction ability in the $x$, $y$ and $z$ directions. If there is some translation and rotation of the CT initialization from the MR dataset, the segmentation method will still work if the user sets a larger error tolerance but this will be at the expense of increased computing time. Due to the vertebral shape, it would be easier for the algorithm to correct the initialization error horizontally (in the $x$ and $y$ directions) than vertically (in the $z$ direction), especially in the segmentation of the spinal processes.

## 5.5 Conclusion

In this Chapter, we have described a new iterative methodology for fast and accurate multimodal CT/MR registration and segmentation of MR dataset executed in a concurrent manner. In MR image segmentation, we extend the ordinary single-front level set to the double-front level set. This effectively reduces computational time by limiting the search area around the target and enhances segmentation accuracy by avoiding leakage and distraction by other objects. The iterative segmentation/registration method helps to refine the segmentation of MR images and the registration of MR to CT. We have tested the algorithm on six pairs of image datasets. The technique is fully automatic but is still able to give results that are comparable to manual segmentation. The proposed segmentation/registration approach will aid the development of image based pre-surgery planning, image guided surgery and post-surgery examination.

We have demonstrated our method in the creation of a hybrid model – a fused CT/MR image dataset – with patient data. The hybrid model of a patient with damaged vertebrae is particularly useful for surgical planning of procedures such as vertebroplasty [97]. With both the spinal cord and damaged vertebrae displayed, the surgeon could plan a needle trajectory that could reach the center of the collapse vertebra without inflicting damages to the spinal cord.

The iterative registration/segmentation process is currently applicable to the CT and MR scans of the same subject [98]. In future, we can set up a standard spine model, which can be used as a standard initial template or reference model for the segmentation of MR images of different patients. We could extract prominent features of the MR image segmentation (e.g., the spinal cord axis and positions of inter-vertebral discs) and re-shape the reference model according to these features by registering the segmented MR images to the standard model.

# 6 Statistical Modeling of Vertebrae

## 6.1 Introduction

Medical images such as CT images can be used to analyze the mechanical properties of vertebrae [99]. The vertebrae in different sections of the spine have distinctively different geometrical shapes and hence biomechanical behaviors. The lumbar section of the human spine, shown in Figure 6.1, have been under the focus of intensive research because it is the main load-bearing region of the entire vertebral column and its abnormality contributes to the development of an array of the pathological symptoms, such as low back pain.

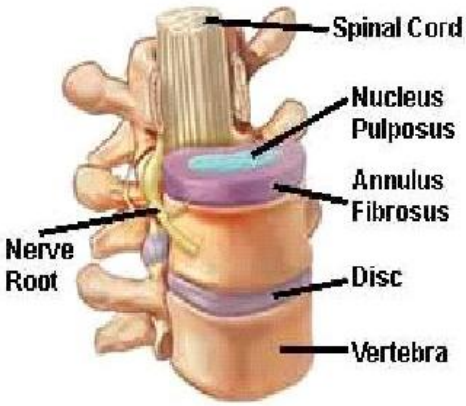



Figure 6.1. A segment of lumbar spine revealing the internal structure of an intervertebral disc and spinal nerve system (adapted from [100]).

FE models of human vertebrae can be used to assess strain and stress fields. The interpretation of medical images is a very challenging task considering the significant inter-subject variability of anatomy and function. Atlas-based approaches address this problem by defining a common reference space. Mapping data sets into this common reference space not only accounts for anatomical and functional variations of

individual subjects, it also offers a powerful tool which facilitates comparison of anatomy and function over time, between subjects, between groups of subjects. Different elastic [101] and fluid [102] warping techniques have been developed for this purpose.

We propose a statistical model-based framework to rapidly create FE meshes with patient-specific geometry. A center firing searching method was implemented to find the corresponding control points for training statistical shape model. The proposed framework may be used to generate FE models of complex geometrical structure such as human vertebrae from medical images.

## 6.2 Methods

Traditionally, landmarks are anatomically characteristic points which can be uniquely identified across a set of individuals. When constructing a statistical model, it is necessary to choose control points properly. The control points should represent the vertebral shape well and has small number at the same time. Given a set of such dense correspondences, one can build a statistical model of the deformation field [103].

In our proposed method, we first construct a statistical shape model database by extracting surface points from existing FE models. Taking into consideration the tubular vertebral body shape, a center firing method with rays directed outwards from the center is used to determine the control points to construct the statistical shape model. Shape parameters derived from the statistical variation among the surface

points are used to represent the reference object model. A Bayesian formulation, based on this prior knowledge and surface information of the input image set, is used to find the most suitable reference model from the database for mapping. The selected model, which acts as a template, is then deformed elastically to match the input image geometry. Figure 6.2 shows the structure of the proposed system.

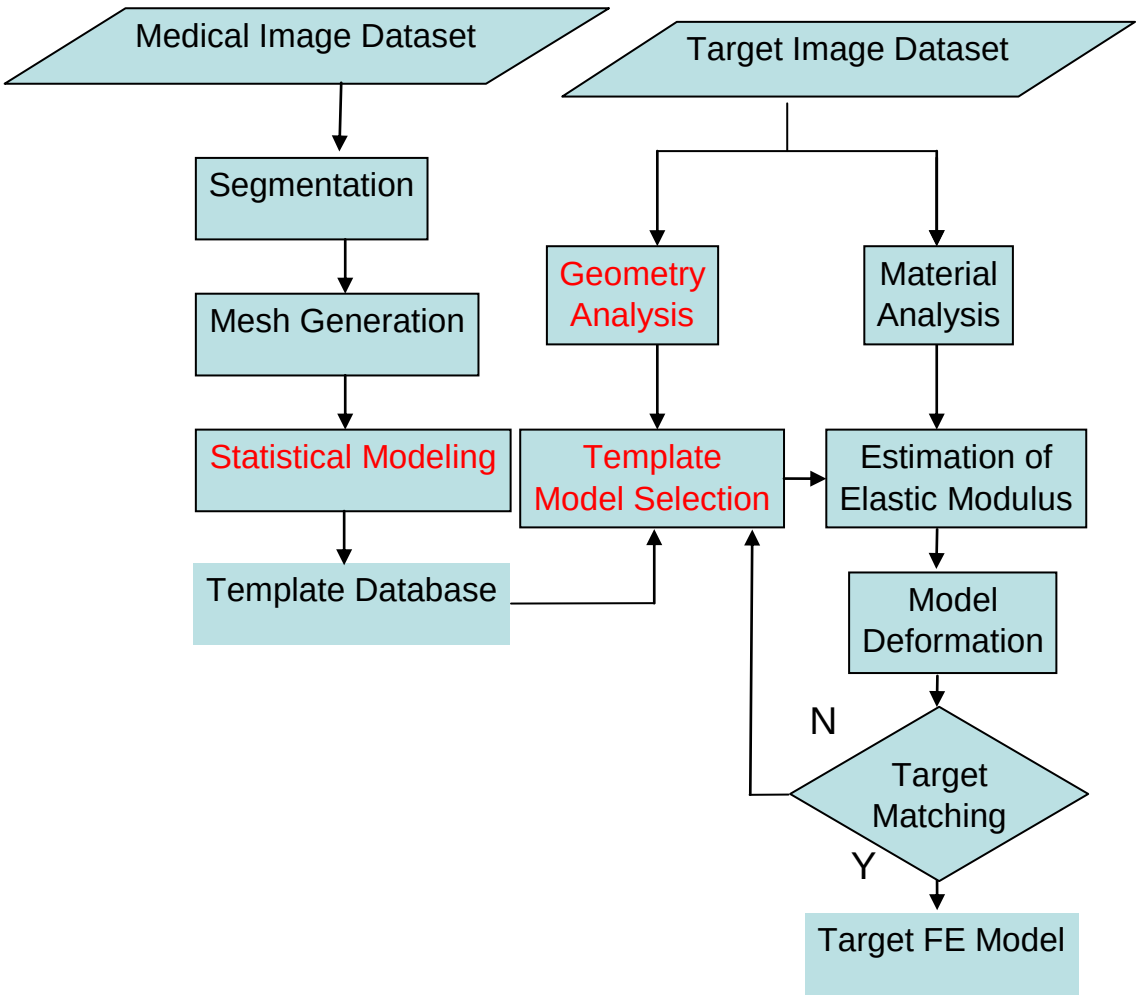



Figure 6.2. System structure.

After processing $N$ sets of sample vertebrae images, we now have a full set of $N$ surface meshes. The next step is to align these surfaces into a common reference frame for analysis. The meshes are processed to remove view and size dependent

considerations to isolate variations in shape alone, which is the goal of our analysis. We align the samples with the following steps:

- Translate centroids of samples to the same point
- Scale the samples to the same size
- Rotate the samples to the same orientation

To train a statistical shape model, we need to use the locations of the corresponding control points in the training dataset. Instead of the manual selection or fixed distance interpolation from surface points, we use a center firing searching technique to automatically locate the corresponding points for each surface nodes of the template model.

From the center of gravity of a 3D model, we fire out rays at different directions. We record the position of the intersections of the rays and the model surface. The 2D illustration is shown in Figure 6.3. Under the assumption that the training data sets have similar shape, these corresponding points are used to train the vertebrae statistical shape model.

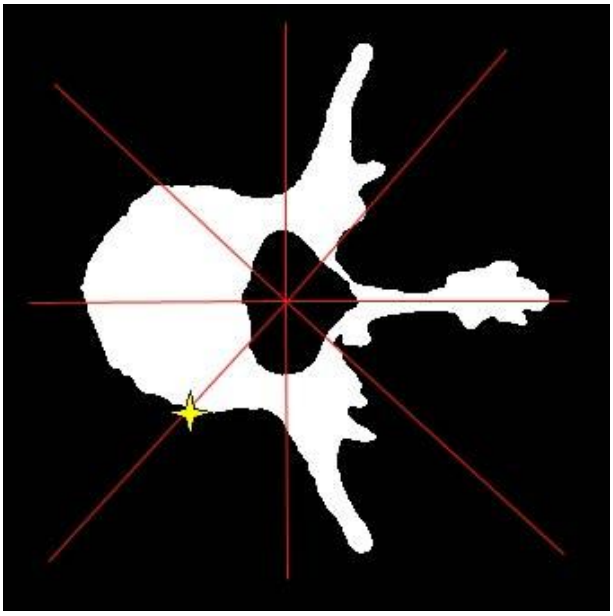


Figure 6.3. CG firing searching.

With the preprocessed data, we are ready to perform the statistical analysis as in [104]. Vertex locations are represented as a vector $\mathbf{m}_i = (x_1; y_1; z_1; x_2; y_2; z_2; :::; x_n; y_n; z_n)$. The training vectors create a cloud of points in 3*n* dimensions, which are modeled as a multivariate Gaussian distribution. We can therefore compute the mean of the distribution (centroid of the point cloud)

$$\overline{\mathbf{m}} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{m}_i, \qquad (6.1)$$

and the covariance matrix

$$\boldsymbol{\omega} = \frac{1}{N}\sum_{i=1}^{N} (\mathbf{m}_i - \overline{\mathbf{m}})(\mathbf{m}_i - \overline{\mathbf{m}})^t. \qquad (6.2)$$

Our goal at this point is to construct a method for approximating each instance of this shape using only a small number of parameters. In order to accomplish this, we determine the principal components of the Gaussian distribution, which represent the axes that contain the most variation in the set. This principal component analysis

(PCA) produces modes of variation given by the eigenvectors of the covariance matrix. These are eigenvectors $\mathbf{u}_k$ such that for some eigenvalue $\lambda_k$, we have

$$\omega \mathbf{u}_k = \lambda_k \mathbf{u}_k . \tag{6.3}$$

By creating a matrix $\mathbf{U}$ with the $n$ eigenvectors corresponding to the largest eigenvalues as the matrix columns, we can now compactly approximate each training instance using the mean of these $n$ modes of variation as:

$$\mathbf{m} \approx \overline{\mathbf{m}} + \mathbf{U}\boldsymbol{\alpha}, \tag{6.4}$$

for some set of mode coefficients $\boldsymbol{\alpha}$.

In the prediction process, the target image surface points will be aligned as described above. With the statistical shape model, the similarity between the target image set and training dataset can be computed. Then the most similar mesh model is chosen for elastic deformation.

While elastic models are useful in non-rigid registration, they are limited by themselves because they are too generic [102]. Considering the statistical information, the elastic model has stronger constrain to deform. Statistical models can be powerful tools to directly capture the character of the variability of the individuals being modeled. Instead of only relying on an elastic model to guide the deformation in a roughly plausible way, the statistics of a sample of images can be used to guide the deformation in a way governed by the measured variation of individuals.

## 6.3 Statistical Model Based Deformation Results

The dataset shown in Figure 4.4 is used in this experiment. This dataset includes four set of CT images of different vertebrae. After performing random deformation to these four set of images, we obtain a total of 32 vertebral images, which are used as the training dataset.

We first extracted the surfaces of the vertebrae, and then used the resultant surface points to align the vertebral body with the template. The center firing method was used to determine the initial set of control points. These control points were marked on the original slices, as shown in Figure 6.4.

With these control points from 32 datasets, we constructed a statistical shape model. Statistical models of shape variation [105] have been shown to be powerful tools for image interpretation. By changing the shape parameter $\alpha$ , different deformed shapes can be created. This is shown in Figure 6.5. The shape parameter is able to represent the various deformed shapes. In this example, the appropriate values are 1, 5 and 10 for $\alpha_1$ , $\alpha_2$ , and $\alpha_3$ , respectively.

Shape parameters derived from the statistical variation among the surface points have been used to represent the reference object model. A Bayesian formulation based on this prior knowledge and surface information of the target image set is used to select the most suitable reference model from the database for mapping. From the literature, the elastic modulus of cortical bone is between 10 GPa and 30 GPa. Thus, an analysis

on the target bone material provides an estimated elastic modulus within this range. The selected model, which acts as a template, was deformed elastically using FE method to match the target image geometry with the given elastic modulus. When the resultant model of deformation did not match the targeted image dataset well, elastic modulus was adjusted to achieve a better matching. The resulted model was then compared with the target dataset. This iteration ended when an optimal matching was found. Figure 6.6 illustrates the effect of different elastic modulus on deformation. A material with large elastic modulus is difficult to deform, while one with small elastic modulus is easier to deform.

Figure 6.7 shows a generated finite element model of a patient specific vertebra. The resultant finite element model and reconstructed patient vertebral model using volume rendering are visually similar. The elastic modulus used in the deformation range from 10 to 30 GPa assuming a fixed Poisson's ratio of 0.3. This is consistent with the estimation of bone mechanical properties reported in the literature. The final estimated value depends on the tolerance and numerical convergence of the finite element methods in the deformation loop. Nevertheless, the iterations have served well in fine tuning the estimation based on our experience to-date. A match quantitatively defined by average Euclidean distance may not represent a good visual geometrical match.

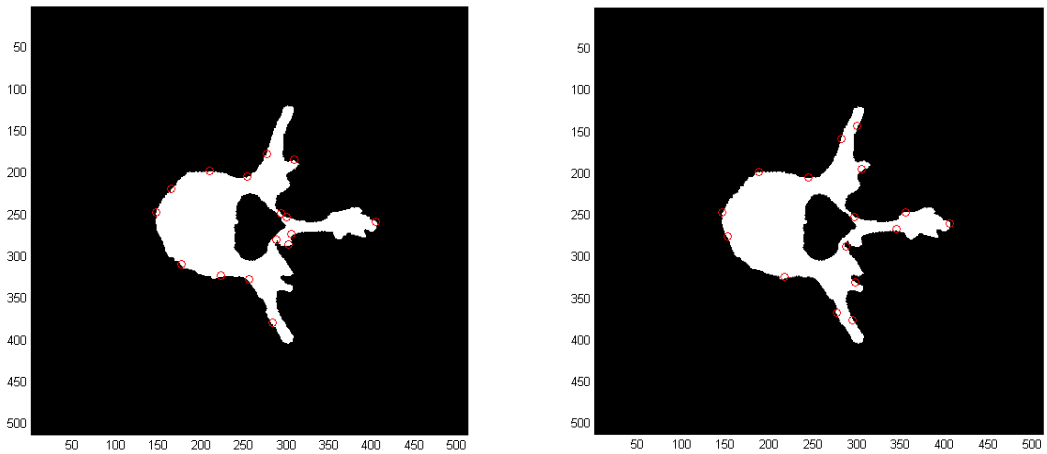
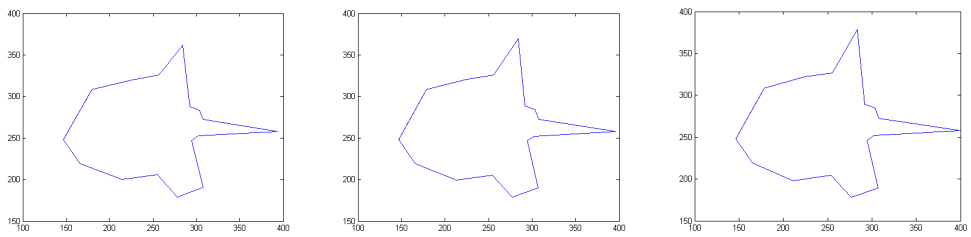
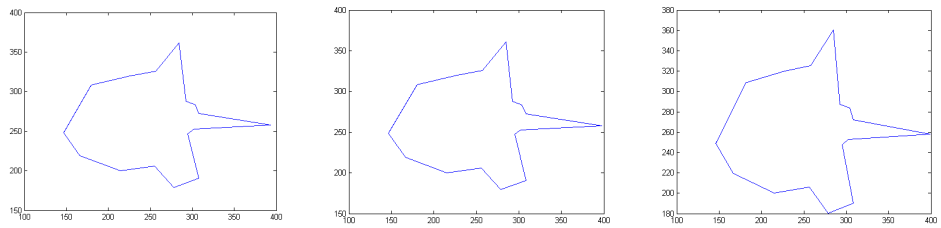


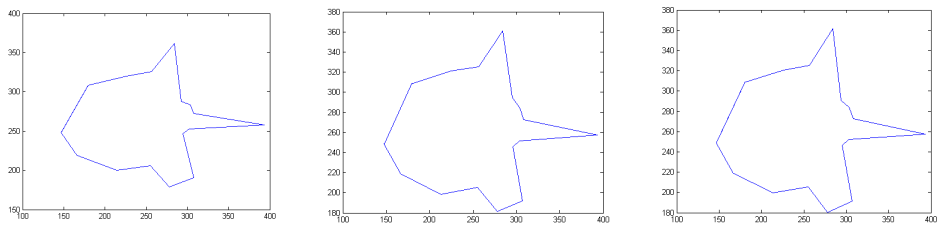
Figure 6.4. Control points marked by center firing method.



(a)

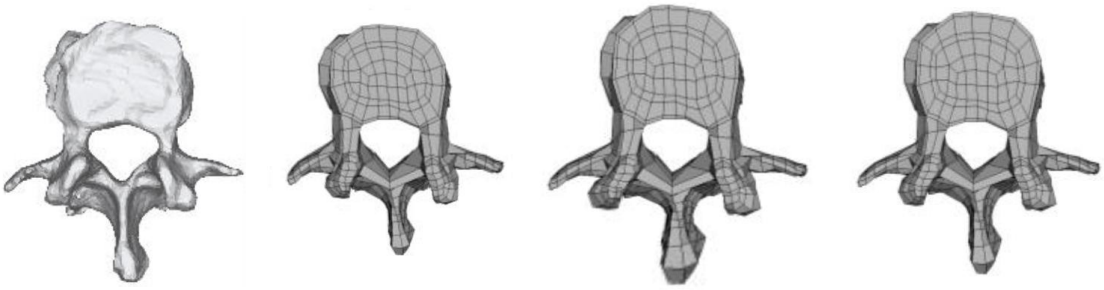


(b)



(c)

Figure 6.5. Deformed shape by changing the shape parameter. Varying (a) first shape parameter $\alpha_1$, (b) second shape parameter $\alpha_2$, (c) third shape parameter $\alpha_3$.



(a) (b) (c) (d)

Figure 6.6. Deformation results of different elastic modulus. (a) Target image. Results of (b) small elastic modulus, (c) large elastic modulus, (d) optimal elastic modulus.

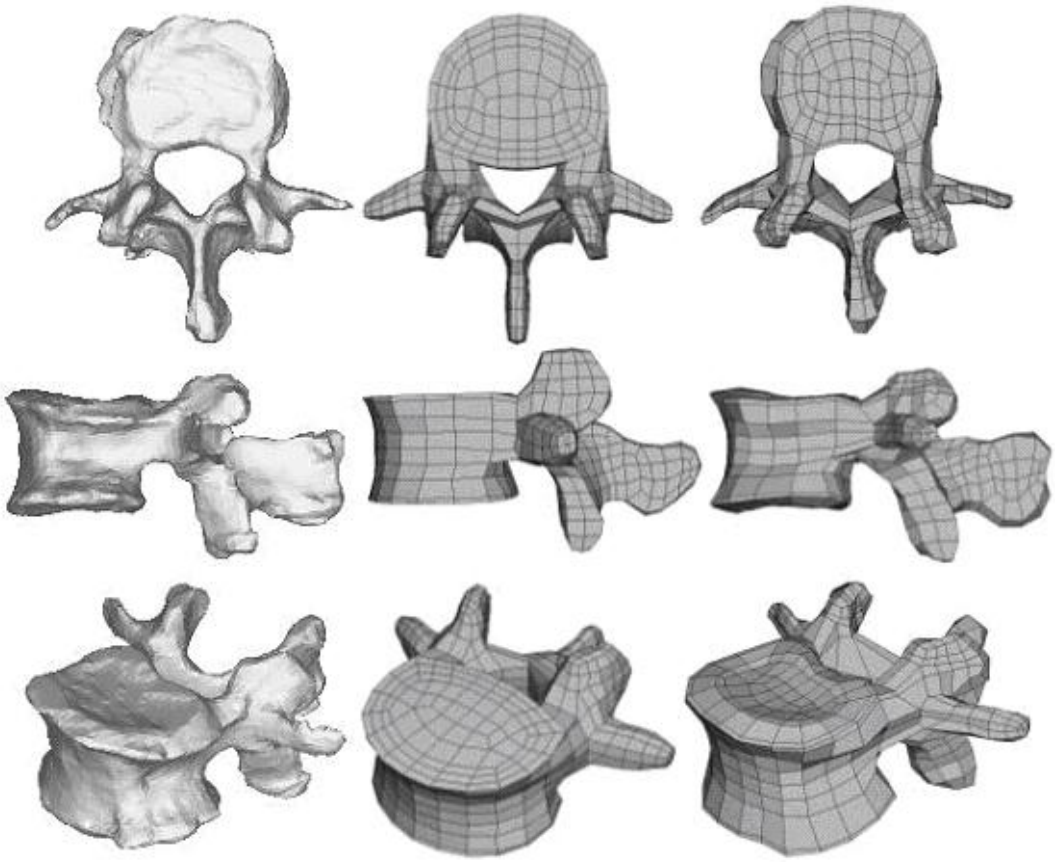


Figure 6.7. Patient specific finite element model. Left, central and right column are the top, side and perspective view of the target vertebrae geometry, template mesh and the transformed mesh, respectively.

## 6.4 Conclusion

Our method has the following advantages over conventional template-based mesh-generation methods. First, high mapping quality is ensured. We select proper vertebral templates using statistical analysis of a pre-trained database instead of using a single template, which reduces the possibility of mapping error for a complex structure such as vertebra. Secondly, minimum preprocessing, e.g., pre-adjustment, is required. Hence, we can generate the FE mesh faster and more accurately than those methods with no or minimal manual intervention. The method also has the benefits of conventional template-based mesh-generation methods. The generated mesh model has a relatively small number of elements that are shaped and organized efficiently to represent the essential geometrical features of the structures.

In preliminary experiments, we applied the proposed method to model human lumbar spine with an initial database comprising six datasets. Preliminary results show that the statistical shape information has significantly augmented and improved template-based mesh generation.

# 7 Conclusion and Future Work

## 7.1 Conclusion

Registration helps the surgeon to match the information from preoperative scan images with that of the intra-operative patient data during image guided surgery. There is a need for an accurate registration system that improves surgical outcomes and patient comfort via elimination of invasive implants. The work described has involved the research and development of a new registration system based on computational model. Preoperative images of patient are segmented using an adaptive thresholding method. The adaptive thresholding method takes into consideration the inhomogeneity of bone structure. A patient-specific surface model is then constructed and used in the registration process.

We proposed and developed a new automatic surface-based rigid registration system using the NN techniques for CT/CT registration. We use a MLP NN to construct the bone surface model. A surface representation function is derived from the resultant NN model and is adopted in intra-operative registration. An optimization process is used to search for optimal transformation parameters together with the NN model. Since no point correspondence is required in our NN based model, the intra-operative registration process becomes significantly faster compared to standard techniques. These advantages are demonstrated in our applications to several medical datasets.

We achieve sub-voxel accuracy using our registration method which is comparable to that of conventional approaches.

In order to produce a complete image volume with clearly visible hard and soft tissues in high resolution from both CT and MR modalities, we propose a weighted method for CT/MRI registration. This is an iterative methodology that can achieve accurate MRI segmentation and CT/MRI registration simultaneously. A semi-automatic segmentation is performed for CT dataset. After a suboptimal MRI segmentation, the segmented MRI dataset is registered with the segmented CT dataset. The registered CT contour is then used as prior knowledge (or initial condition) for the MRI segmentation. This iterative process is carried out until the segmented CT surface matches the segmented MRI perfectly. The experimental results of six pair of images showed the feasibility of this system and the advantages compared to other conventional methods. We have also investigated a statistical model-based framework to rapidly create FE meshes with patient-specific geometry.

The above CT/CT and CT/MRI registration methods were integrated into a generic software toolkit. The software toolkit has already been used in segmentation of various human and animal images. It has also been applied to register human bone structures for image-guided surgery. The successful completion of the weighted registration method greatly enhances the state-of-art for CT/MRI registration.

## 7.2 Image-based Bone Material Estimation

Newer techniques such as quantitative ultrasound (QUS) have been introduced recently for measuring bone density. QUS sends non-ionizing sound waves to detect mineral density. Piezoelectric transducers transmit ultrasound energy that travels through the bone to the receiving transducer. Reductions in ultrasound signal are attributed to attenuation by bone and tissue. QUS is also an averaged area method and cannot distinguish cortical from trabecular bone. It is therefore used mainly in thin cortex regions and is not able to measure sites at risk of osteoporotic fracture such as the hip or spine [106]. Studies have shown that adding an ultrasound measurement to a Dual-Energy X-Ray Absorptiometry (DXA) does not improve the prediction of fractures [107, 108]. Although some have said that ultrasound measures the "quality" of bone, more careful studies suggest that it mainly measures the bone mass. Newer systems incorporate imaging techniques to aid in positioning and improve precision [109, 110]. The advantages of QUS include no radiation exposure, low cost, portability and rapidity of scanning. Assessment of fracture risk in elderly women by QUS has been proven [107, 111, 112], and studies indicate that in the elderly, QUS is as good a predictor of hip fracture as DXA [110]. However, QUS is not suitable for assessing the spine. A primary disadvantage of QUS is the lack of sensitivity, making it inappropriate for long term monitoring of osteoporosis or response to drug therapy. Significant false negative rate has been detected in discriminating healthy from osteoporotic groups [98].

Since the mechanical properties of the bone depend largely on both the density and structure of the trabecular bone, imaging techniques with direct measures of trabecular bone structure may improve the analysis of bone biomechanical properties [113] compared to those that only measure the average area Bone Mineral Density (BMD). The biomechanical properties of bone can be integrated to our hybrid model for further study.

## 7.3 Clinical Applications

There is a recent study showing that in spinal surgery, robotically assisted needle insertion is feasible and enhances placement accuracy, especially in complicated cases [8, 9]. This application performed pre-operation planning with three projections of vertebrae reconstructed from CT images. During operation, a geometric relation between the coordinate systems of the patient's anatomy and the preoperative plan is established automatically by matching the preoperative reconstructed CT images to intra-operative fluoroscopic images of the patient and targeting device in place.

The ability to support minimally invasive procedures is one of the most attractive features of this technology. Deformity and revision cases are also compelling indications of this system. Surgeons can find the correct entry point and trajectory by using unusual or absent anatomical landmarks. Clinically acceptable placement reached high accuracy rate as 99% in all the cases. It verifies the system's accuracy and supports its use in minimally invasive and open spine surgery for pedicle screws, as well as for translaminar facet screw techniques.

# Bibliography


[1] R. Boellaard, Obtaining cardiac images from positron emission tomography, computed tomography, and magnetic resonance imaging: physical principles, Heart and Metabolism, 34: 33-37, 2007.

[2] http://www.magnet.fsu.edu/education/tutorials/magnetacademy/mri/index.html

[3] R.H. Hashemi, and W.G. Bradley, Jr., MRI: The Basics, Lippincott, Williams and Wilkins, Baltimore, 1997.

[4] C. Scharver, R. Evenhouse, A. Johnson and J. Leigh, Designing cranial implant in a haptic augmented reality environment, Communications of the ACM, 47(8): 32-38, 2004.

[5] M. Agus, A. Giachetti, E. Gobbetti, G. Zanetti, and A. Zorcolo, Antonio Zorcolo, A multiprocessor decoupled system for the simulation of temporal bone surgery, Computer Visual Science, 5: 35-43, 2002.

[6] M. Meehan, D. Morris, C.R. Maurer, A.K. Antony, F. Barbagli, K. Salisbury and S. Girod, Virtual 3D planning and guidance of mandibular distraction osteogenesis, Computer Aided Surgery, 2(11): 51-62, 2006.

[7] F.A. Jolesz, I.F. Talos, R.B. Schwartz, H. Mamata, D.F. Kacher, Hynynen K, N. McDannold, P. Saivironporn and L. Zao, Intraoperative magnetic resonance imaging and magnetic resonance imaging-guided therapy for brain tumors, Neuroimaging clinics of North America, 12(4): 665-683, 2002.

[8] G.S. Fischer, A. Deguet, C. Csoma, R.H. Taylor, L. Fayad, J.A. Carrino, S.J. Zinreich and G. Fichtinger, MRI image overlay: Application to arthrography needle insertion, Computer Aided Surgery, 12(1): 2-14, 2007.

[9] M. Shoham, I.H. Lieberman, E.C. Benzel, D. Togawa, E. Zehavi, B. Zilberstein, M. Roffman, A. Bruskin, A. Fridlander, L. Joskowicz, S. Brink-Danan and N. Knoller, Robotic assisted spinal surgery - from concept to clinical practice, Computer Aided Surgery, 12(2): 105-115, 2007.

[10] S.P. DiMaio, S. Pieper, K. Chinzei, N. Hata, E. Balogh, G. Fichtinger, C.M. Tempany and R. Kikinis, Robot-assisted needle placement in open MRI: System architecture, integration and validation, Computer Aided Surgery, 12(1): 15-24, 2007.

[11] G. Schulte-Altedorneburg, M. Gebhard, W.A. Wohlgemuth, W. Fischer, J. Zentner, R. Wegener, T. Balzer and K. Bohndorf, MR arthrography: pharmacology, efficacy and safety in clinical trials. Skeletal Radiology, 32(1): 1-12, 2003.

[12] C.A. Binkert, F.R. Verdun, M. Zanetti, C.W. Pfirrmann and J. Hodler, CT arthrography of the glenohumeral joint: CT fluoroscopy versus conventional CT and fluoroscopy--comparison of image-guidance techniques, Radiology, 229(1): 153-158, 2003.

[13] C.V. Thakar, O. Liangos, J.P. Yared, D.A. Nelson, S. Hariachar and E.P. Paganini, Predicting Acute Renal Failure after Cardiac Surgery: Validation and

Re-definition of a Risk-Stratification Algorithm, Hemodialysis International, 7(2): 143-147, 2003.

[14] D. L. Pham, C. Xu, and J. L. Prince, Current methods in medical image segmentation, Annual Review of Biomedical Engineering, 2(1): 315-337, 2000.

[15] S.S. Jasijit, S.S. Kamaledin and S. Sameer(Eds), Advanced Algorithmic Approaches to Medical Image Segmentation, State-of-the-Art Applications in Cardiology, Neurology, Mammography and Pathology, Springer-Verlag, pp 238, 2002.

[16] L. G. Nyul, and J. K. Udupa, On standardizing the MR image intensity scale, Magnetic Resonance in Medicine, 42(6): 1072-1081, 1999.

[17] C. Pluempitiwiriyawej, J. M. F. Moura, Y. J. L. Wu and C. Ho, STACS: new active contour scheme for cardiac MR image segmentation, IEEE Transactions on Medical Imaging, 24(5): 593-603, 2005.

[18] L. K. Van, F. Maes, D. Vandermeulen and P. Suetens, A unifying framework for partial volume segmentation of brain MR images, IEEE Transactions on Medical Imaging, 22(1): 105-119, 2003.

[19] M. E. Algorri, F. F. Mangas, Classification of anatomical structures in MR brain images using fuzzy parameters, IEEE Transactions on Biomedical Engineering, 51(9): 1599-1608, 2004.

[20] D. Selvathi, A. Arulmurgan, S. S. Thamarai and S. Alagappan, MRI image segmentation using unsupervised clustering techniques, Six International

Conference on Computational Intelligence and Multimedia Applications, pp. 105-110, 16-18 August 2005.

[21] S. Shan; W. Sandham, M. Granat and A. Sterr, MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization, IEEE Transactions on Information Technology in Biomedicine, 9(3): 459-467, 2005.

[22] J. A. Sethian, An analysis of flame propagation, Ph.D. Thesis, Department of Mathematics, University of California, Berkeley, CA, 1982.

[23] J. A. Sethian, A fast marching level set method for monotonically advancing fronts, Applied Mathematics, 93: 1591-1595, 1996.

[24] J.B.A. Maintz and M.A. Viergever, A survey of medical image registration, Medical Image Analysis, 2(1): 1-36, 1998.

[25] Y. Amit and A. Kong, Graphical templates for model registration, IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(3): 225-236, 1996.

[26] J. Thirion, New feature points based on geometric invariants for 3D image registration, International Journal of Computer Vision, 18(2): 121-137, 1996.

[27] S. Fang, R. Raghavan and J.T. Richtsmeier, Volume morphing methods for landmark based 3D image deformation. In M.H. Loew and K.M. Hanson (eds), Medical Imaging: Image processing, 404-415, Newport Beach, CA. 1996.

[28] T. Peters, B. Davey, P. Munger, R. Comeau, A. Evans and A. Olivier, Three-dimensional multimodal imageguidance for neurosurgery, IEEE Transactions on Medical Imaging, 15(2): 121-128,1996.

[29] P.J. Slomka, G.A. Hurwitz, J. Stephenson and T. Cradduck, Automated alignment and sizing of myocardial stress and rest scans to three-dimensional normal templates using an image registration algorithm, Journal of Nuclear Medicine, 36: 1115-1122, 1995.

[30] L. Dong and A.L. Boyer, A portal image alignment and patient setup verification procedure using moments and correlation techniques, Physics in Medicine and Biology, 41, 697-723, 1996.

[31] J.B.A. Maintz, P.A. van den Elsen and M. A. Viergever, Comparison of edge-based and ridge-based registration of CT and MR brain images, Medical Image Analysis, 1(2): 151-161,1996.

[32] T. Buzug and J. Weese, Improving DSA images with an automatic algorithm based on template matching and an entropy measure, In H.U. Lemke, M.W. Vannier, K. Inamura and A.G. Farman (eds), Computer Assisted Radiology, Vol. 1124 of Excerpta Medica - International Congress Series, 145-150, Elsevier. 1996.

[33] C. Chen, C.A. Pelizzari, G.T.Y. Chen, M.D. Cooper and D.N. Levin, Image analysis of PET data with the aid of CT and MR images, Information Processing in Medical Imaging, 601-611, Plenum Press. 1988.

[34] G. Borgefors, Hierarchical chamfer matching: a parametric edge matching algorithm, IEEE Transactions on Pattern Analysis and Machine Intelligence, 10: 849-865, 1988.

[35] C. Davatzikos, Nonlinear registration of brain images using deformable models, Proceedings of Mathematical Methods in Biomedical Image Analysis, 94-103, San Francisco, CA. 1996.

[36] J. Thirion, Non-rigid matching using demons, Proceedings of Computer Vision and Pattern Recognition, 245-251, San Francisco, CA. 1996.

[37] R. Bajcsy, R. Lieberson and M. Reivich, A computerized system for the elastic matching of deformed radiographic images to idealized atlas images, Journal of Computer Assisted Tomography, 7(4): 618-625. 1983.

[38] A. Guéziec, Large deformable splines, crest lines and matching, Proceedings of International Conference on Computer Vision, 650-657, Berlin, Germany. 1993.

[39] D. MacDonald, D. Avis and A.C. Evans, Multiple surface identification and matching in magnetic resonance images, Proceedings of SPIE Visualization in Biomedical Computing, 160-169, Rochester, MN. 1994.

[40] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, Numerical Recipes in C, Cambridge University Press, 2nd edition. 1992.

[41] B. Münch and Rüegsegger, 3-D repositioning and differential images of volumetric CT measurements, IEEE Transactions on Medical Imaging, 12(3): 509-514, 1993.

[42] J. Jacq and C. Roux, Registration of non-segmented images using a genetic algorithm, In N. Ayache (ed.), Proceedings of Computer Vision, Virtual Reality, and Robotics in Medicine, Nice, France, 905: 205-211, 1995.

[43] P.A. Van den Elsen, E.J.D. Pol, T.S. Sumanaweera, P.F. Hemler, S. Napel and J. Adler, Grey value correlation techniques used for automatic matching of CT and MR brain and spine images, Proceedings of SPIE Visualization in Biomedical Computing, Rochester, MN. 2359: 227-237, 1994.

[44] P.J. Besl and N.D. McKay, A method for registration of 3D shapes, IEEE Transactions on Pattern Analysis and Machine Intelligence, 14: 239-256, 1992.

[45] P.F. Hemler, T.S. Sumanaweera, P.A. van den Elsen, S. Napel and J. Adler, A versatile system for multimodality image fusion, Journal of Image Guided Surgery, 1(1): 35-45, 1995.

[46] P.F. Hemler, T. Sumanaweera, R. Pichumani, P.A. van den Elsen, S. Napel and J. Adler, A system for multimodality image fusion, Proceedings of IEEE Symposium on Computer-based Medical Systems, Winston-Salem, NC. 335-340, 1994.

[47] P.F. Hemler, T. Sumanaweera, R. Pichumani, P.A. van den Elsen, S. Napel, J. Drace, and J. Adler, A system for multimodality image fusion of the spine, In AAAI Spring Symposium Series: Applications of Computer Vision in Medical Images Processing, Stanford University, 42-45, 1994.

[48] E. Bainville, G. Champledoux, P. Cinquin, V. Dessenne, A. Hamadeh, J. Troccaz, S. Lavall ́ee, O. P ́eria, P. Sautot and R. Szeliski, Anatomical surfaces based 3D/3D and 3D/2D registration for computer assisted medical interventions, In L. Beolchi and M.H. Kuhn (eds), Medical Imaging: Analysis of Multimodality 2D/3D Images, IOS Press. 19: 53-69, 1995.

[49] C.A. Kapoutsis, C.P. Vavoulidis and I. Pitas, Morphological iterative closest point algorithm, IEEE Transactions on Image Processing, 8(11): 1644-1646, 1999.

[50] G. Burel, H. Henocq, and J. Catros, Registration of 3D objects using linear algebra, In N. Ayache (ed.), Proceedings of Computer Vision, Virtual Reality, and Robotics in Medicine, Springer-Verlag, 905: 252-256, 1995.

[51] A. Shirazi-adl, A.M. Ahmed and S.C. Shrivastava, Mechanical response of a lumbar motion segment in axial torque alone and combined with compression, Spine, 11(9): 914-927, 1986.

[52] A. Shirazi-adl, S.C. Shrivastava and A.M. Ahmed, Stress analysis of the lumbar disc-body unit in compression, Spine, 9: 120-134, 1984.

[53] R. A. McLaughlin, J. Hipwell, D. J. Hawkes, J. A. Noble, J. V. Byrne, and T. C. Cox, A comparison of a similarity-based and a feature-based 2-D-3D registration method for neurointerventional Use, IEEE Transactions on Medical Imaging, 24(8): 1058-1066, 2005.

[54] M.L. Jan, K.S. Chuang, G.W. Chen *et al*., A three-dimensional registration method for automated fusion of micro PET-CT-SPECT whole-body images, IEEE Transactions on Medical Imaging, 24: 886-893, 2005.

[55] J.R. Cebral and R. Lohner, From medical images to anatomically accurate finite element grids, International Journal For Numerical Methods In Engineering, 51: 985-1008, 2001.

[56] A.E. Anderson, C.L. Peters, B.D. Tuttle and J.A. Weiss, Subject-specific finite element model of the pelvis: development, validation and sensitivity studies, ASME Journal of Biomechanical Engineering, 127: 364-373, 2005.

[57] Z. Wang, J.C.M. Teo, C.K. Chui, S.H. Ong, C.H. Yan, S.C. Wang, H.K. Wong, S.H. Teoh, Computational biomechanical modeling of the lumbar spine using marching-cubes surface smoothened finite element voxel meshing, Computer Methods and Programs in Biomedicine, 80: 25-35, 2005.

[58] L. Baghdadi, D.A. Steinman and H.M. Ladak, Template-based finite-element mesh generation from medical images, Computer Methods and Programs in Biomedicine, 77: 11-21, 2005.

[59] J.H. Keyak, J.M. Meagher, H.B. Skinner and C.D. Mote, Automated three-dimensional finite element modeling of bone: a new method, Journal of Biomedical Engineering, 12: 389-397, 1990.

[60] J.H. Keyak, M.G. Fourkas, J.M. Meagher, H.B. Skinner, Validation of an automated method of three-dimensional finite element modeling of bone, Journal of Biomedical Engineering, 15: 505-509, 1993.

[61] J.H. Keyak, T.S. Kaneko, J. Tehranzadeh and H.B. Skinner, Predicting proximal femoral strength using structural engineering models, Clinical Orthopaedics and Related Research 2005, 437: 219-228, 2005

[62] V. Luboz, M. Chabanas, P. Swider and Y. Payan, Orbital and maxillofacial computer aided surgery: patient-specific finite element models to predict surgical

outcomes, Computer Methods in Biomechanics and Biomedical Engineering, 8(4): 259-265, 2005.

[63] G.Y. Zheng and L.P. Nolte, Surface reconstruction of bone from X-ray images and point distribution model incorporating a novel method for 2D-3D correspondence, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2: 2237-2244, 2006.

[64] B.W. Fei, J.L. Duerk, D.B. Sodee and D.L. Wilson, Semiautomatic nonrigid registration for the prostate and pelvic MR volumes, Academic Radiology, 12(7): 815-824, 2005.

[65] A. Rossa, S. Dassb and A. Jainc, A deformable model for fingerprint matching, Pattern Recognition, 38: 95-103, 2005.

[66] D.C. Mathieu, B.A. Aloys, M. Benoit and K.W. Simon, Multi-subject registration for unbiased statistical atlas construction, Proceedings MICCAI 2004, 655-662, 2004.

[67] F. Destrempes and M. Mignotte, A statistical model for contours in images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(5): 626-638, 2004.

[68] R.H. Davies, C.J. Twining, T.F. Cootes, J.C. Waterton and C.J. Taylor, 3D statistical shape models using direct optimisation of description length, Proceedings of the 7th European Conference on Computer Vision-Part III, 3-20, 2002.

[69] C.R. Brice and C.L. Fenema, Scene analysis using regions, Artificial Intelligence, 1: 205-226, 1970.

[70] R.M. Haralick and L.G. Shapiro, Computer and Robot Vision (Volume II) (Paperback), Prentice Hall, 2002.

[71] M. Kass, A. Witkin, and D. Terzopoulos, Snakes: Active contour models, International Journal of Computer Vision, 1(4): 321-331, 1987.

[72] K.R. Castleman, Digital Image Processing. Prentice Hall. 1996.

[73] E.R. Davies, Machine Vision. Academic Press. 1997.

[74] J. Kittler, J. Illingworth and J. Foglein, Threshold based on image statistics, Computer Vision, Graphics and Image Processing, 30: 125-147, 1985.

[75] C.H. Yan, Measuring Changes In Local Volumetric Bone Density: New Approaches To Quantitative Computed Tomography, Ph.D Thesis, Stanford University, 1998.

[76] R.O. Duda, and P.E. Hart, Pattern Classification and Scene Analysis. John Wiley & Sons. 1973.

[77] A. Santos, J. Pascau, M. Desco, J.A. Santos, F. Calvo, *et al*, Multimodality image integration for radiotherapy treatment, an easy approach, Proceedings of SPIE, 4319: 715-723, 2001.

[78] C.A. Pelizzari, G.T.Y. Chen, D.R. Spelbring, R.R. Weichselbraum and C. Chen, Accurate three dimensional registration of CT, PET and/or MR images of the brain, Journal of Computer Assisted Tomography, 13: 20-26, 1989.

[79] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., The Johns Hopkins University Press, 1996.

[80] S. Haykin, Neural Networks: A Comprehensive Foundation, second ed., Prentice Hall, 1999.

[81] S. Hezel, A. Kugel, R. Manner and D.M. Gavrila, FPGA-based template matching using distance transforms, Proceedings 10th Annual IEEE Symposium on Field-Programmable Custom Computing Machines, 89-97, 2002.

[82] E.J. Teoh, K.C. Tan, and C. Xiang, Estimating the number of hidden neurons in a feedforward network using the singular value decomposition, IEEE Transactions on neural networks, 17; 1623-1629, 2006.

[83] E.D. Sontag, Feedback stabilization using two-hidden-layer nets, IEEE Transactions on Neural Networks, 3: 981-999, 1992.

[84] L. Ljung, System Identification: Theory for the User. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[85] C.H. Yan, R.T. Whalem, G.S. Beaupre, G.S. Sumanaweera, S.Y. Yen and S. Nape, A new frame registration algorithm, Medical Physics, 25: 121-128, 1998.

[86] http://www.intel.com/technology/computing/opencv/index.htm, Open Source Computer Vision Library.

[87] H.M. Chen and P.K. Varshney, Mutual information-based CT-MR brain image registration using generalized partial volume joint histogram estimation, IEEE Transactions on Medical Imaging, 22: 1111-1119, 2003.

[88] J. Zhang, Y. Ge, S.H. Ong, C.K. Chui, S.H. Teoh, C.H. Yan, Rapid surface registration of 3D volumes using a neural network approach, Image and Vision Computing, 26(2): 201-210, 2008.

[89] C.H. Yan, S.H. Ong, Y. Ge, J. Zhang, S.H. Teoh, B.H. Okker, A neural network approach for 3D surface modeling and registration, IEEE International Workshop on BioMedical Circuits & Systems, Singapore, 1-3 December 2004.

[90] C. Xu and J.L. Prince, Snakes, shapes and gradient vector flow, IEEE Transactions on Image Processing, 7(3): 359-369, 1998.

[91] E. Rouy and A. Tourin, A viscosity solutions approach to shape-from-shading, SIAM Journal on Numerical Analysis, 29: 867-884, 1992.

[92] A. Yezzi, L. Zollei and T. Kapur, A variational framework for integrating segmentation and registration through active contours, Medical Image Analysis, 7(2): 171-185, 2001.

[93] C. R. Maurer, Jr., R. J. Maciunas and J. M. Fitzpatrick, Registration of head CT images to physical space using a weighted combination of points and surfaces, IEEE Transactions on Medical Imaging, 17(5): 753-761, 1998.

[94] B. H. Okker, C. H. Yan, J. Zhang, S. H. Ong and S. H. Teoh, Accurate and fully automatic 3D registration of spinal images using normalized mutual information, IEEE International Workshop on BioMedical Circuits & Systems, Singapore, 1-3 December 2004.

[95] J. Zhang, C.H. Yan, C.K. Chui, S.C. Wang, S.H. Ong, Multimodality Image Registration System for Image Guided Orthopaedic Surgeries, IEEE Transactions on Medical Imaging, revised and resubmitted.

[96] P. Slagmolen, D. Loeckx, S. Roels, X. Geets, F. Maes, K. Haustermans and P. Suetens, Nonrigid registration of multitemporal CT and MR images for radiotherapy treatment planning, Biomedical Image Registration, Lecture Notes in Computer Science (Heidelberg: Springer-Verlag), 297-305, 2006.

[97] S. T. Wang, H. L. Ma, C. L. Liu, W. K. Yu, M. C. Chang and T. H. Chen, Is fusion necessary for surgically treated burst fractures of the thoracolumbar and lumbar spine?: A prospective, randomized study, Spine, 31(22): 2646-2652, 2006.

[98] B.L. Riggs, L.J. Melton, 3rd. Bone turnover matters: the raloxifene treatment paradox of dramatic decrease in vertebral fractures without commensurate increases in bone density, Journal of Bone and Mineral Research, 17: 11-14, 2002.

[99] http://www.espine.com/normal_anatomy.html

[100] J. O'Brien, E. Burton, I. McKeith, N. Fox, *et al*., Alzheimer's disease and dementia with Lewy bodies show different regional patterns of evolving atrophy using fluid registration of serial MRI scans, Neurobiology of Aging, 298, 2004.

[101] W.R. Crum, R.I. Scahill, N.C. Fox, Automated hippocampal segmentation by regional fluid registration of serial MRI: Validation and application in Alzheimer's disease, Neuroimage, 13(5): 847-855, 2001.

[102] D. Rueckert, A.F. Frangi, and J.A. Schnabel, Automatic construction of 3D statistical deformation models using non-rigid registration, In MICCAI, 77-84, 2001.

[103] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham, Active shape models-their training and application, Computer Vision and Image Understanding, 61(1): 38-59, 1995.

[104] T.F. Cootes, C.J. Twining and C.J. Taylor, Diffeomorphic statistical shape models, Proceedings of British Machine Vision Conference, 1: 447-456, 2004.

[105] E.M. Lochmüller, R. Müller, V. Kuhn, C.A. Lill and F. Eckstein, Can novel clinical densitometric techniques replace or improve DXA in predicting bone strength in osteoporosis at the hip and other skeletal sites?, Journal of Bone and Mineral Research, 18: 906-912, 2003.

[106] X.G. Cheng, P.H.F. Nicholson, S. Boonen, G. Lowet, P. Brys, J. Aerssens, G. Van De Perre and J. Dequeker, Prediction of vertebral strength in vitro by spinal bone densitometry and calcaneal ultrasound, Journal of Bone and Mineral Research, 12: 1721-1728, 1997.

[107] A. Stewart and D.M. Reid, Precision of quantitative ultrasound: comparison of three commercial scanners, Bone, 27: 139-143, 2000.

[108] G. Falgarone, R. Rorcher, A. Duche, S. Kolta, M. Dougados and C. Roux, Discrimination of osteoporotic patients with quantitative ultrasound using imaging or non-imaging device, Joint Bone Spine, 71: 419-423, 2004.

[109] D.C. Bauer, C.C. Gluer, J.A. Cauley *et al*, Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women, A prospective study, Study of Osteoporotic Fractures Research Group. Archives of Internal Medicine, 157: 629-634, 1997.

[110] C.C. Gluer, Quantitative ultrasound techniques for the assessment of osteoporosis: expert agreement on current status, The International Quantitative Ultrasound Consensus Group, Journal of Bone and Mineral Research, 12: 1280-1288, 1997.

[111] M.A. Moyad, Osteoporosis: a rapid review of risk factors and screening methods, Urologic Oncology, 21: 375-379, 2003.

[112] C.H. Chesnut 3rd and C.J. Rosen, Reconsidering the effects of antiresorptive therapies in reducing osteoporotic fracture, Journal of Bone and Mineral Research, 16: 2163-2172, 2001.

# Appendix A

## Bayes Decision Theory

Let $i$ be a gray level whose distribution depends on the corresponding material (bone or non-bone). Let $p(i \mid B)$ be the conditional probability density function for $i$ given that the corresponding material is bone related. Similarly $p(i \mid \overline{B})$ represents the conditional probability density function for $i$ given that the corresponding material is non-bone. Then the difference between $p(i \mid B)$ and $p(i \mid \overline{B})$ describes the difference in attenuation between bone region and non-bone region, see Figure 1. Suppose that the priori probabilities, $P(B)$ and $P(\overline{B})$, are known, the posteriori probability for each material given a gray level $i$ can be computed using the *Bayes Rule*:

$$P(B \mid i) = \frac{p(i \mid B)P(B)}{p(i)}$$

$$P(\overline{B} \mid i) = \frac{p(i \mid \overline{B})P(\overline{B})}{p(i)} \tag{A.1}$$

where

$$p(i) = p(i \mid B)P(B) + p(i \mid \overline{B})P(\overline{B}) \tag{A.2}$$

For classification of the underlying material based on the observed gray level, the *Bayes decision rule* minimizes the probability of classification error: Decide $B$ if $p(B \mid i) \geq P(\overline{B} \mid i)$; otherwise decide $\overline{B}$. This rule can be re-written as

$$\text{Decide } B \text{ if } p(i \mid B)P(B) \geq p(i \mid \overline{B})P(\overline{B}); \text{ otherwise decide } \overline{B}. \tag{A.3}$$

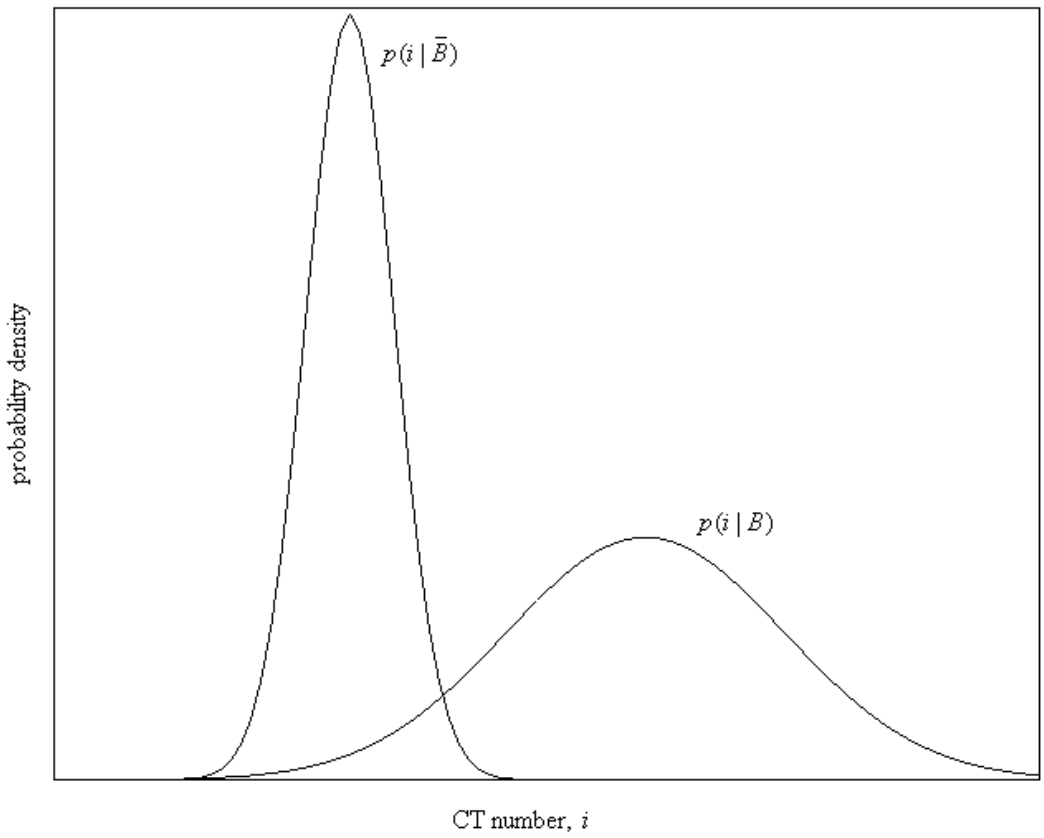



Figure A.1. Class conditional probability density function.

Consider in a neighborhood $N(x)$ of a pixel $x$ and we assume that the gray level from this region come from a mixture of two Gaussian distributions ($B$ and $\bar{B}$) having respective means, variances and priori probability ($\mu_b, \sigma_b^2, P(B)$) and ($\mu_{nb}, \sigma_{nb}^2, P(\bar{B})$). Based on the current segmentation of the image ($B$ and $\bar{B}$), these parameters can be estimated by the following equations:

$$\mu_b = \frac{1}{|N(x) \cap B|} \sum_{y \in N(x) \cap B} I(y),$$

$$\sigma_b^2 = \frac{1}{|N(x) \cap B| - 1} \sum_{y \in N(x) \cap B} (I(y) - \mu_b)^2, \qquad \text{(A.4)}$$

$$P(B) = \frac{|N(x) \cap B|}{|N(x)|},$$

$$\mu_{nb} = \frac{1}{\left|N(x) \cap \overline{B}\right|} \sum_{y \in N(x) \cap \overline{B}} I(y),$$

$$\sigma_{nb}^2 = \frac{1}{\left|N(x) \cap \overline{B}\right| - 1} \sum_{y \in N(x) \cap \overline{B}} (I(y) - \mu_{nb})^2, \qquad \text{(A.5)}$$

$$P(\overline{B}) = 1 - P(B).$$

Thus Bayes Decision Rule can be re-formulated as

Decide $B$ if

$$\frac{\left|N(x) \cap B\right|}{\sigma_b} \exp\left(\frac{-(i - \mu_b)^2}{2\sigma_b^2}\right) \geq \frac{\left|N(x) \cap \overline{B}\right|}{\sigma_{nb}} \exp\left(\frac{-(i - \mu_{nb})^2}{2\sigma_{nb}^2}\right) \qquad \text{(A.6)}$$

otherwise decide $\overline{B}$

# Appendix B

## Publication List:

**Journal Papers:**

J. Zhang, C.H. Yan, C.K. Chui, S.C. Wang, S.H. Ong, Multimodality Image Registration System for Image Guided Orthopaedic Surgeries, IEEE Transactions on Medical Imaging, revised and resubmitted.

C.K. Chui, L Bian, Z Wang, J. Teo, J Zhang, J. Ong, C.H. Yan, S.H. Ong and S.H. Teoh: A component oriented software toolkit for patient specific mesh generation and its applications in spinal simulation, Advances in Engineering Software, revised and resubmitted.

J. Zhang, Y. Ge, S.H. Ong, C.K. Chui, S.H. Teoh, C.H. Yan, Rapid Surface Registration of 3D Volumes Using a Neural Network Approach, Image and Vision Computing, to be published, 26(2): 201-210, 2008.

C.K. Chui, Jackson S.K. Ong, Z.Y. Lian, Z.L. Wang, J. Teo, J. Zhang, C.H. Yan, S.H. Ong, S.S. Wang, H.K. Wong, C.L. Teo and S.H. Teoh, Haptic Interaction in Computer Game-like Surgical Simulation for Vertebroplasty Training, Simulation and Gaming, 2006; 37(4): 438-451.

**Book Series:**

C-K Chui, JCM Teo, Z Wang, J Zhang, JSK Ong, K-M Si-hoe, S-H Ong, C-H Yan, S-C Wang, H-K Wong, JH Anderson and S-H Teoh (2006): Integrative haptic and visual interaction for simulation of PMMA injection during vertebroplasty, Health Technology and Informatics 119 - Medicine Meets Virtual Reality 14, JD Westwood et al (Eds), IOS Press. pp:96-98.

**Conference Papers:**

J. Zhang, C.W Yeo, C.K. Chui, C.H. Yan, S.H. Ong, S.H. Teoh, Bone Material Estimation in Statistical Model Aided Elastic Deformation, 2nd MRS-S Conference on Advanced Materials, 18-20 January 2006, Singapore.

J. Zhang, C.K. Chui, C.H. Yan, S.H. Ong, Z.L. Wang, J.C.M. Teo, S.H. Teoh, Fast Template-Based Finite-Element Mesh Generation using Statistical Shape Information, The 12$^{th}$ International Conference On Biomedical Engineering, 2005, Singapore.

J. Ong, Z.Y. Lian, C.K. Chui, Z. Wang, J. Zhang, J. Teo, C.H. Yan, S.H. Ong, C.L. Teo and S.H. Teoh: Biomechanical modeling of bone for needle insertion, 2nd MRS-S Conference on Advanced Materials, 18-20 January 2006, Singapore.

C.H. Yan, S.H. Ong, Y. Ge, J. Zhang, S.H. Teoh, B.H. Okker, A Neural Network Approach for 3D Surface Modeling and Registration, IEEE International Workshop on BioMedical Circuits & Systems, 1-3 December 2004, Singapore.

B.H. Okker, C.H. Yan, J. Zhang, S.H. Ong, S.H. Teoh, Accurate and Fully Automatic 3D Registration of Spinal Images using Normalized Mutual Information, IEEE International Workshop on BioMedical Circuits & Systems, 1-3 December 2004, Singapore.