One-to-one Edge Based Registration and Segmentation Based Validations in Hybrid Imaging

Technische Fakultät der
Universität Erlangen-Nürnberg

zur Erlangung des Grades

DOKTOR-INGENIEUR

vorgelegt von

by

Jingfeng Han

Erlangen — 2009

Als Dissertation genehmigt von der
Technische Fakultät der
Universität Erlangen-Nürnberg

Tag der Einreichung:	29. 07. 2009
Tag der Promotion:	23. 11. 2009
Dekan:	Prof. Dr.-Ing. R. German
Berichterstatter:	Prof. Dr.-Ing. J. Hornegger
	Prof. Dr. rer. nat. M. Rumpf

Abstract

During the past decade, image registration has become an essential tool for medical treatment in clinics, by finding the spatial mapping between two images, observing the changes of anatomical structure and merging the information from different modalities. On the other hand, the matching of appropriately selected features is becoming more and more important for the further improvement of registration methods, as well as for the qualitative validation of registration. The purpose of this thesis is to solve the following two problems: How to integrate feature detection into a non-rigid registration framework, so that a high quality spatial mapping can be achieved? How to systematically measure the quality of multi-modal registration by automatically segmenting the corresponding features?

For the first problem, we develop a general approach based on the Mumford-Shah model for simultaneously detecting the edge features of two images and jointly estimating a consistent set of transformations to match them. The entire variational model is realized in a multi-scale framework of the finite element approximation. The optimization process is guided by an EM type algorithm and an adaptive generalized gradient flow to guarantee a fast and smooth relaxation. This one-to-one edge matching is a general registration method, which has been successfully adapted to solve image registration problems in several medical applications, for example mapping inter-subject MR data, or alignment of retina images from different cameras.

For the second problem, we propose a new method validating the hybrid functional and morphological image fusion, especially for the SPECT/CT modality. It focuses on measuring the deviation between the corresponding anatomical structures. Two kinds of anatomical structures are investigated as validation markers: (1) the hot spot in a functional image and its counterpart in the morphological image (2) the kidneys in both modalities. A series of special methods are developed to segment these structures in both modalities with minimum user interaction. Accuracy of the validation methods have been confirmed by experiments with real clinical data-sets. The inaccuracies of hot spot based validation for neck regions are reported to be 0.7189 ± 0.6298 mm in X-direction, 0.9250 ± 0.4535 mm in Y-direction and 0.9544 ± 0.6981 mm in Z-direction. While the inaccuracies of kidneys based validation for abdomen regions are 1.3979 ± 0.8401 mm in X-direction, 1.9992 ± 1.3920 mm in Y-direction and 2.7823 ± 2.0672 mm in Z-direction. In the end, we also discuss a new interpolation based method to effectively improve the SPECT/CT fusion and present preliminary results.

Deutscher Titel:
Eins-zu-eins kantenbasierte Bildregistrierung und bildsegmentierungsbasierte Validierung des hybriden Scanners

Kurzfassung

Bildregistrierung wurde in den letzten Jahrzehnten für medizinische Anwendungen immer wichtiger, um Transformationen zwischen Bildern zu bestimmen, Veränderungen anatomischer Strukturen zu verfolgen oder verschiedene Bildmodalitäten zu vereinen. Weiterhin wird die Zuordnung von entsprechend ausgewählten Merkmalen immer wichtiger für die Verbesserung von Registrierungsverfahren sowie für eine qualitative Evaluierung. Ziel dieser Arbeit ist es Lösungsansätze zu folgenden Fragestellungen zu finden: Wie kann die Merkmalsdetektion in eine nicht-starre Registrierung eingebunden werden, so dass eine qualitativ gute örtliche Zuordnung erreicht werden kann? Wie kann die Qualität einer multi-modalen Registrierung durch automatische Segmentierung korrespondierender Merkmale evaluiert werden?

Für das erste Problem entwickelten wir einen allgemeinen Ansatz, der auf dem Mumford-Shah Modell aufbaut. Hierbei werden gleichzeitig Kantenmerkmale von zwei Bildern extrahiert und eine Menge von Transformationen geschätzt, die diese Merkmale aufeinander abbilden. Der Variationelle Ansatz ist durch ein Multi-Skalen-Rahmenwerk von Finite Elemente Schätzungen realisiert worden. Die Optimierung wurde durch einen EM-Algorithmus und einem adaptiven generalisierten Gradientenabstieg umgesetzt, um eine schnelle und glatte Relaxation zu garantieren. Die Eins-zu-Eins Zuordnung der Kanten ist ein allgemeiner Registrierungsansatz, der erfolgreich angewendet wird um Registrierungsprobleme bei vielen verschiedenen medizinischen Anwendungen zu lösen. Beispiele hierfür sind die MR Datensätze verschiedener Patienten oder die Ausrichtung von Retinaaufnahmen aufgrund unterschiedlicher Aufnahmegeräte.

Für die zweite Fragestellung stellen wir einen neuen Ansatz vor, um hybride funktionelle und morphologische Bildfusion, im speziellen SPECT/CT, zu evaluieren. Im Mittelpunkt steht hierbei die Vermessung der Abweichungen zwischen korrespondierenden anatomischen Strukturen. Zwei Kategorien von anatomischen Strukturen wurden untersucht: (1) der Hotspot in der funktionellen Bildgebung und sein Gegenstück in der morphologischen Aufnahme, (2) die Nieren in beiden Modalitäten. Eine Reihe von Methoden wurden entwickelt um diese Strukturen in beiden Modalitäten mit minimaler Benutzerinteraktion zu segmentieren. Die Experimente

mit echten medizinischen Daten bestätigen die Genauigkeit der Validierungsmethoden. Bei dem Hotspot-basierenden Verfahren beträgt die Ungenauigkeit 0.7189 ± 0.6298 mm in X-Richtung, 0.9250 ± 0.4535 mm in Y-Richtung und 0.9544 ± 0.6981 mm in Z-Richtung in Halsregionen. Die Ungenauigkeit bei der Nieren-basierten Evaluierung liegt bei 1.3979 ± 0.8401 mm in X-Richtung, 1.9992 ± 1.3920 mm in Y-Richtung und 2.7823 ± 2.0672 mm in Z-Richtung im Abdomenbereich. Wir stellen zusätzlich ein neues interpolation-basierendes Verfahren vor, um die SPECT/CT Fusion zu verbessern und zeigen erste vorläufige Ergebnisse.

Contents

List of Figures

List of Tables

List of Algorithms

Part I

Introduction

Chapter 1

Motivation

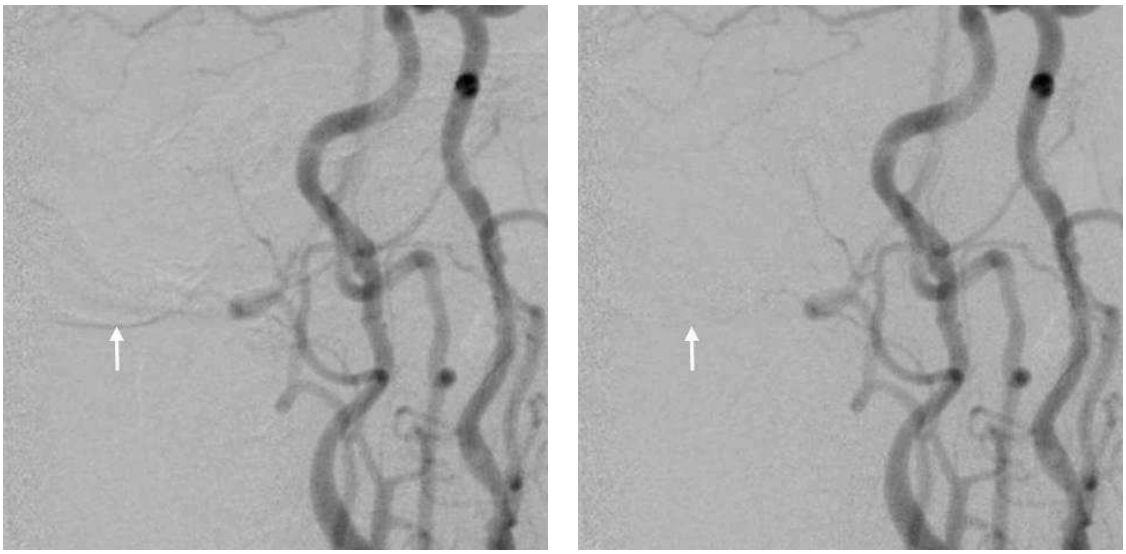
1.1 Medical Image Registration

The presented thesis deals with a specific problem of medical image analysis, namely image registration, also known as image fusion or image matching. Image registration is the process of finding an optimal geometric transformation, so that two given images are correctly aligned to each other. The concrete form of "optimal geometric transformation" varies a lot in different situations, but all these transformations define a point-to-point correspondence between the image pair.

Image registration has plenty of applications in the field of medical image processing. For instance, the typical requirements from physicians are to compare images acquired at different times, from different perspectives, of different patients or by different imaging modalities. Image registration is the fundamental and crucial processing step to determine the correspondence between the given images. In the following the effect of registration is illustrated on two examples. However, one should note that image registration has a much broader range of clinical applications.

Figures 1.1(a) and 1.1(b) show the motion compensation in Digital Subtraction Angiography (DSA). DSA is a standard way to visualize human vasculature by acquiring a pair of 2-D X-ray projection images and subtracting contrast-enhanced images (contrast images) from a contrast-free image (mask image). A fast and fully automatic image registration is usually required to remove the motion artifact between the contrast and mask image prior to subtraction.

Figures 1.2(a)-1.2(c) show the fusion of X-ray Computed Tomography (CT) and Photon Emission Tomography (PET) images of the same patient. These two modalities visualize different information: CT provides high-resolution images of density distribution of different tis-

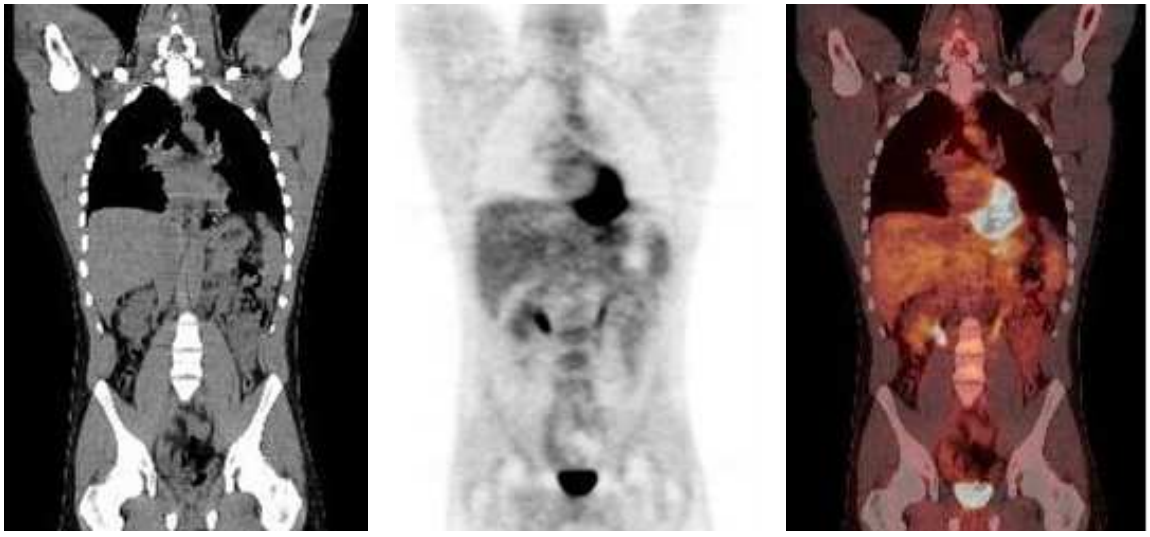


(a) (b)

Figure 1.1: Image registration for digital subtraction angiography. (a) Original substraction image without a registration process. (b) Substraction images processed by registration with subpixel precision. The images are courtesy of Mrs. Y. Deuerling-Zheng. Details on the applied registration method are described [DZLGH06].

sues, which can effectively show the anatomy of the patient, while PET modality records the alive functional and physiological activities of organs, for example, the glucose metabolism during the acquisition. The merge between these two modalities, e.g. CT as background and PET as foreground, is very useful for the diagnose, surgery plan as well as the observation of the follow-up. Nevertheless, if CT and PET scans are performed at different times or at different hospitals, rigid registration is obligated to transform two volumes into a common coordinate system. Then non-rigid registration is optionally used to compensate the non-rigid patient motion in the acquisitions.

The research on image registration has developed rapidly in the last twenty years. A substantial part of research on medical image processing deals with image registration. This trend was proven in a recent review study of image registration [PF03]. Image registration turned out to be more difficult than people expected. There are still several topics in the field of registration, for which many researchers are actively investigating more satisfactory solutions. In the following, we focus on these three challenges: multi-modal registration, non-rigid registration and validation of registration. Then we summarize the major contributions of this thesis, namely some new approaches proposed to solve these problems.



(a) CT (b) PET (c) Combined

Figure 1.2: Fusion of PET/CT volume data. The correct overlapping of two modalities noticeably shows physiological activities with the anatomical background. The images are courtesy of Dr. W. Römer (Nuclear Medicine Department of University of Erlangen).

Multi-modal Registration

In the past three decades, progress in medical imaging techniques and image processing methods has led to the fact that different imaging modalities with high resolution are available for medical treatment today. Currently, the most imaging modalities can be roughly classified into morphological and functional imaging modalities. For example, X-ray imaging, CT and Magnetic Resonance (MR) are considered morphological imaging modalities, whereas functional MR imaging (fMRI) and molecule imaging techniques, like Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT), are functional imaging modalities. These imaging modalities provide complementary information and the registration of these data brings significant clinical benefits for diagnosis and surgical planning. Even though a large number of methods have been invented in the past, the registration of different imaging modalities is still far away from being perfect. The fundamental reason is that the individual imaging modality cannot provide enough correlated information and sufficient contrast for a reliable registration. Simple intensity based similarity measures, typically computing of statistical dependencies, cannot reflect the correspondence of the same underlying anatomical structures in different modalities. The lack of knowledge of image contents is now more and more likely to be a bottleneck for further improvement of registration algorithms.

Non-rigid Registration

In contrast to rigid registration estimating the translation and rotation, non-rigid registration computes an elastic deformation field to align two images. Non-rigid registration is also known by many different names, such as “non-linear”, “elastic”, “non-parametric” or “deformable” registration. Non-rigid registration is a critical issue in many clinical applications. For instance, in computer assisted neurosurgery, the deformation of the brain between pre- and intraoperative MR data, referred to as the brain shift, needs to be corrected by non-rigid registration. A drawback of most current non-rigid registration algorithms is that they model all tissue as having the same degree of rigidity. However, physicians expect that the different tissues or different organs have different degree of rigidity, e.g. bone structures or instruments should be transformed rigidly. However, most algorithms uniformly compute the deformation, regardless of the underlining tissue classes. The second drawback is the inconsistency of the deformation field. Consistency of transformation means that if one computes the transformation from A to B and then switches the roles of A and B to compute the second transformation B from A, the two transformations should be inverse to each other. Consistent registration is not only more sound in the mathematical sense, but also very important for applications, where one is interested in determining the one-to-one correspondence of the same anatomical structures in different images, e.g. non-rigid registration for atlas construction [RFS03, MTT03] or historical biological images [STU05, CSM+06].

Validation

Besides of automatic fusion of different imaging modalities, physicians demand additional quantitative information on the reliability of the result. For example, they need information like, “The fusion of our new PET/CT machine has an average accuracy of 1 mm in the region of the head.” or “The registration error of the software is less than 5 mm in the region of the abdomen.” Validation is an essential part of the registration process and widely considered to be unsolved. It checks whether the anatomical structures in the first image is mapped to the corresponding ones in the second image. Although plenty of registration algorithms have been invented in the last three decades, very few researchers paid attention to validation methods for image registration. The most commonly used validation method is to compare the registration system against a gold standard, in which an optimal mapping has been pre-defined. For example, “Retrospective Image Registration Evaluation Project[1]” (RIRE) is a recently new possibility to compare various CT-

[1] http://www.insight-journal.org/rire/

MR and PET-MR registration techniques. People can download a group of datasets and perform registrations on them. Then the computed transforms are uploaded to compare with the "true" transforms, which are defined by a prospective, marker-based technique. The registration gold standard may be based on a computer simulation or a clinic phantom. However, the quantitative validation using a gold standard cannot always indicate the accuracy of the registration in a real clinical scenario, because real medical image datasets are much more complicated than simple simulations or phantom objects. In practice, visual assessment has often been used as a standard for validation. However, the reliability and reproducibility of human inspection are always questioned. In addition, most validation efforts have been concentrated primarily on rigid registration, but the systematic evaluation of non-rigid registration is still a great challenge.

1.2 Contributions of this Thesis

The author believes that image segmentation is the key to find better solutions for the registration and validation challenges outlined in the preceding sections. Generally speaking, image segmentation and image registration are two closely related problems. The goal of image segmentation is to simplify or to change the representation of an image into something that is more meaningful and easier to analyze [SS01], usually we call them "features". Image segmentation is typically used to locate objects or to find boundaries, i.e. lines, curves, among images. Whereas the task of image registration is to determine the correspondence between images. Ideally the same underlying anatomies are mapped to each other. Matching of features intuitively could be a natural criterion that drives as well as evaluates the image registration algorithms. Many researchers, including the author, believe that the integration of knowledge of image segmentation is a promising way to improve the registration method. From the point of methodology, this thesis has the following two contributions.

One-to-one Edge Based Registration

Since in practice neither the location of desired features, nor the mapping between images are known, it is intuitively a good idea to solve these two highly interdependent problems - image segmentation and image registration - in an iterative fashion in an uniform framework. In each iteration, the detected features guide the registration algorithm to find spatial correspondence, at the same time the estimated transformations constrain the search range of the feature detection. Many methods of joint segmentation and registration have already been proposed in the past ten years. Compared to most methods in literature [ZYK01, MC03, CW98, YL05, PFL+05],

one-to-one edge based registration method has three most distinguishing points:

- The method uses edge features as an input for registration. In this method, each image is approximated by two functions: a piecewise constant function that represents a reconstructed noise-free image and a phase field function that implicitly represents edge features. The experiments we carried out also proved that one-to-one edge based registration is more accurate than intensity based registration for the alignment of fine structures [HBR+06, HBD+07].

- The transformations are modeled as dense deformation fields. Most methods of joint segmentation and registration in literature are typically limited to rigid transformations [ZYK01, MC03, CW98]. This method uses non-rigid deformation fields to define transformation between images. This more flexible transformation model is usually required to compensate the non-rigid deformation, e.g. cardiac motion correction for 3-D reconstruction.

- The registration method is consistent. The method introduced in this thesis estimates forward and inverse transformation at the same time and explicitly constrains two transformations to be inverse to each other. Consistent registration estimates not only a smooth deformable transformation but also a one-to-one mapping of the images.

Automatic Accuracy Validation of SPECT/CT Fusion

Simple visual inspections with aid of a computer was a widely accepted way to measure the accuracy of multi-modal registration. It is essential for the physician to have a systematic method to measure the accuracy of multi-modal registration as well as to compare different registration methods in a clinical setting. This thesis presents accuracy validation methods for SPECT/CT registration using automatic segmentation of corresponding objects in two modalities. The degree of matching of segmented objects indicates the accuracy of registration. The validation can be roughly divided into two steps:

1. **Segmentation of corresponding structures**

   Two kinds of highly active objects in the SPECT modality are selected as the "validation markers": Mal-functional hot spots are used in the region of neck and kidneys are used in the region abdomen. The major challenge of this work was to find the best segmentation algorithms for different objects in different modalities. Several criteria are considered for the selection of algorithm: Whether the algorithm is robust with respect to noise, whether

it can maximize the degree of automation and reproducibility and whether the parameterizations need to adapt to different image data.

2. **Measurement of distance**

   The anatomic accuracy of SPECT/CT fusion is evaluated by measuring the distance between the centers of gravity of corresponding markers in X-, Y- and Z-directions. A small distance between two centers of gravity indicates a highly accurate fusion of SPECT/CT data-sets [HKC+08].

The experiments in this work use patient data generated from SPECT/CT hybrid scanners. Nevertheless, the validation method can also applied to other combinations of modalities, such as PET/CT.

1.3 Overview

This thesis is divided into three parts. Part I (Chapter 1 and 2) is devoted to the fundamental motivation of this work and a general introduction of image registration. In Part II (Chapter 3 and 4), the one-to-one edge based registration method and its successful applications are presented. Part III (Chapter 5 to 7) is devoted to introduce two validation methods of SPECT/CT hybrid imaging and the test results on the real clinic patient datasets. At the end, a brief summary of proposed methods is given.

PART I: Introduction

Chapter 1

This chapter gives the motivation as well as challenges of image registration in the field of medical image analysis. Two typical examples are selected to show the fact that image registration is an obligated process operation for clinic routines. Two major contributions of the thesis are described: (1) one-to-one edge based registration, a new registration method, and (2) a segmentation based validation method for SPECT/CT hybrid imaging.

Chapter 2

This chapter gives a general introduction of state of the art of registration methods. It is required for the presentation of our original contributions. Several fundamental topics of registration are

discussed: Examples are transformation models, interpolation, similarity measurement, variational framework, feature matching and so on.

PART II: One-to-one Edge based Registration

Chapter 3

This chapter presents the one-to-one edge based registration method. A clear relationship between the proposed and existing method is also defined. Functional definitions, computation of its first variations, finite element discretization, iterative algorithm as well as parameter study of the proposed method are described in detail.

Chapter 4

This chapter demonstrates the performance of the developed method. Four applications of one-to-one edge based registration are presented: (1) inter-subject registrations (2) alignment of retinal images from different sensors, (3) matching photographs of neurosurgery to MR Volume and (4) motion compensation for frame interpolation. The experiments prove that the proposed approach achieves a better match of fine structures respect to those existing techniques.

PART III: Segmentation based Validation for SPECT/CT Hybrid Imaging

Chapter 5

This chapter gives an introduction of hybrid SPECT/CT imaging. We focus on the benefit and the potential misalignment artifact as two modalities are combined into a hybrid imaging.

Chapter 6

This chapter is devoted to introduce the validation method of SPECT/CT hybrid imaging using markers of mal-functional hot spots. Hot spots in SPECT images are automatically segmented by a new “localized maximally stable extremal regions” method. The segmentation of corresponding structure in CT images is solved by a semi-automatic random walk method with minimal user interactions.

Chapter 7

This chapter describes the validation of SPECT/CT hybrid imaging using marker of kidneys. The target kidneys in CT image are detected automatically by a classic active shape model method and corresponding location in SPECT is computed by a shape based tracking method. Then, we also discuss the possibility to remove such misalignment artifact, which detected by this validation method.

Chapter 2

Image Registration

Image registration is the process of determining correspondences between points in two or more images. These images show completely or partially the same object or the same scenario, but the positions of image contents are different for the images. The images are not aligned or registered, that is, the direct spatial correspondence between them is not determined. Throughout this work, we consider only gray value images, typically modeled by spatial functions. This function assigns a gray value $u(\boldsymbol{x})$ to every spatial point $\boldsymbol{x} \in \mathbb{R}^d$, where $d = 2$ or 3 is the dimensionality of the images. Typically, two images are involved in the registration problem: a reference image R and a template image T. In this thesis, the reference and template images are in the same dimension, i.e. the 2D-3D registration is not discussed here. Ideally, we look for a transformation $\boldsymbol{\phi} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that the reference image R and the transformed template image $T_{\boldsymbol{\phi}}$ are as close as possible. The intensity function of reference and template images are denoted as u^R and u^T, respectively. The intensity function of $T_{\boldsymbol{\phi}}$ is often denoted as

$$u^T_{\boldsymbol{\phi}}(\boldsymbol{x}) = u^T \circ \boldsymbol{\phi}(\boldsymbol{x}) = u^T(\boldsymbol{\phi}(\boldsymbol{x})). \tag{2.1}$$

This transformation $\boldsymbol{\phi}(\boldsymbol{x})$ essentially determines the spatial correspondence between the reference and template image. Let us assume both images share the same coordinate system and the transformation $\boldsymbol{\phi}$ is invertible. The point $\boldsymbol{x}$ in the template image T corresponds to the point $\boldsymbol{\phi}(\boldsymbol{x})$ in the reference image R, or reversely, the point $\boldsymbol{x}$ in R corresponds to the point $\boldsymbol{\phi}^{-1}(\boldsymbol{x})$ in T.

In the rest of this chapter we review the existing approaches and discuss several important issues related with registration. For example, “how to define a suitable transformation?” “How to measure the similarity between the reference and template, based on image intensity or based

on image contents?" and "How to find the optimal transformation parameters?"

2.1 Transformation Models

In most clinical applications of image registration, the primary transformation model is rigid or non-rigid. Both of them are well-defined mappings of one image into another.

Rigid transformation is used for registration if the corresponding objects have no distortion. A simple rotation and translation are computed for the alignment:

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{R}\boldsymbol{x} + \boldsymbol{t}, \tag{2.2}$$

where $\boldsymbol{R} \in \mathbb{R}^{d\times d}$ is rotation matrix and $\boldsymbol{t} \in \mathbb{R}^d$ is the translation vector. However, spatial rotations in 3-D are more frequently represented by unit quaternions in practice. Quaternions are an extension of complex numbers with four components. Quaternions have many attractive properties: Compared to Euler angles they are simpler to compose and avoid the problem of gimbal lock. Compared to rotation matrices they are more efficient and more numerically stable. A concise definition of quaternions is given in [Alt86].

Non-rigid transformation is used for the registration problems where a deformation is expected to compensate the irregular distortions encountered in medical image analysis. Mathematically it is denoted by a continuous function

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{u}(\boldsymbol{x}), \tag{2.3}$$

where $\boldsymbol{u}(\boldsymbol{x}) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is the so-called displacement field. Figure 2.1 gives a 2-D example of a non-rigid transformation. Compared to the rigid transformation model, non-rigid transformations based on deformation are more flexible. On the other hand, the estimation of non-rigid transformation encounters more problems of parameter optimization and numerical implementation in practice. These topics will be further discussed in Section 2.5.

Besides of rigid and non-rigid transformation, there exist also other ways to model the transformation, for example, affine transformation, B-spline based transformation and so on. Because these transformations are not used in our work, we do not introduce them further. For the definitions and applications of these transformation, we refer to [CJ01, ZF03].

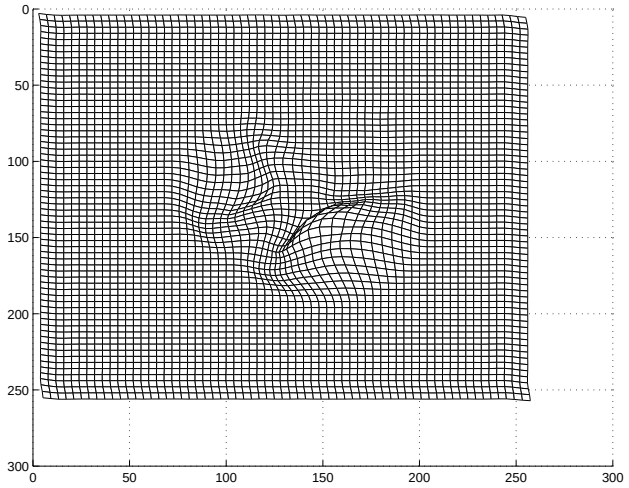


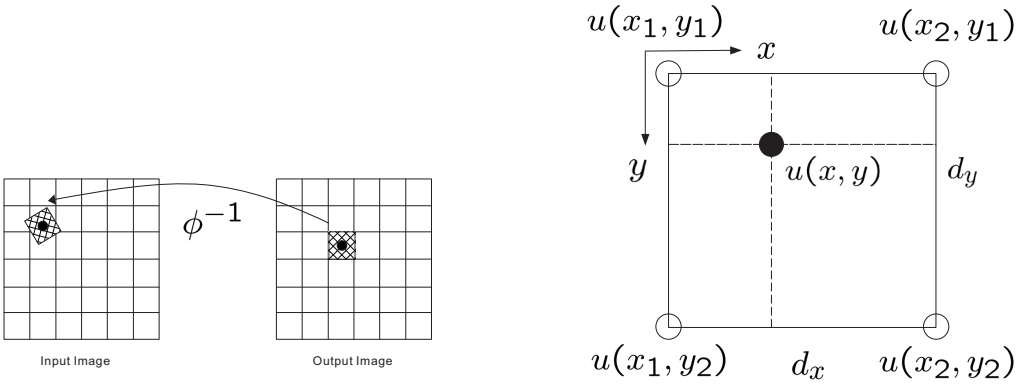
Figure 2.1: An example of 2-D non-rigid transformation.

2.2 Image Interpolation

Both reference and template images are digitalized images, which are basically the collections of samples on mesh points. The transformation between them, normally is defined by a continuous function, has to handle the problem that a mesh sample point on the one image could be transformed to a non-mesh point on the other image. Image interpolation is exactly the solution of this problem: it computes the image function value at a non-mesh point using the estimation based on the values of the neighboring mesh points. Figure 2.2(a) shows how to compute a transformed image. First we assume the transformation from the input to the output image to be invertible. For every voxel in the output image, we use the inverse transformation ϕ^{-1} to compute the corresponding position in the input image. Quite often, corresponding points are not located on the mesh and their intensity values need to be computed by interpolations.

Linear interpolation is often used in image registration, because of the trade-off between the effort of computation and the smoothness of estimation. In a 2-D case like Figure 2.2(b), the bilinear interpolation is computed by

$$\begin{aligned} u(x,y) = u(x_1,y_1)\frac{(x_2-x)(y_2-y)}{d_x d_y} + u(x_1,y_2)\frac{(x_2-x)(y-y_1)}{d_x d_y} \\ + u(x_2,y_1)\frac{(x-x_1)(y_2-y)}{d_x d_y} + u(x_2,y_2)\frac{(x-x_1)(y-y_1)}{d_x d_y}, \end{aligned} \tag{2.4}$$



(a) (b)

Figure 2.2: Image transformation and interpolation in 2D

where d_x and d_y are the width of voxel in x- and y-direction. A trilinear interpolation scheme for 3-D image data can be deduced in a similar way.

2.3 Voxel Similarity Measures

Generally speaking, voxel similarity measures are functions of transformation $\boldsymbol{\phi}$, that measure the agreement of matching. They typically compute the statistic dependency between images based directly on the voxel intensity. Registrations using voxel similarity measures have become more and more popular in the field of medical image analysis in the last decade. The major advantage is the fact that they are fully automatic and require less pre-processing of the images. This seems to be in high demand in a typical "just make it work" clinical environment. In the following, three groups of similarity measures that are the most commonly used in medical image registration will be introduced. For an extensive survey of voxel similarity measures, we refer to [ZF03, PMV03, Bro92].

Minimizing intensity difference between reference and template images is widely used for mono-modal registration. The most well-known intensity difference similarity measure is Sum of Squared Difference (SSD)

$$\mathcal{D}_{\text{SSD}}(\boldsymbol{\phi}) = \int_{\Omega} (u^R(\boldsymbol{x}) - u^T_{\boldsymbol{\phi}}(\boldsymbol{x}))^2 \, d\boldsymbol{x} \tag{2.5}$$

and the Sum of the Absolute Difference (SAD)

$$\mathcal{D}_{\text{SAD}}(\boldsymbol{\phi}) = \int_{\Omega} |u^R(\boldsymbol{x}) - u^T_{\boldsymbol{\phi}}(\boldsymbol{x})| \, d\boldsymbol{x} \tag{2.6}$$

Here Ω denotes the overlapping space of reference and transformed template images. Compared to the other similarity measures, they have really simple definitions and are easy to implement. The minimization of intensity difference is not suitable for multi-modal registration, but even for some mono-modal registrations, they do not always work as well. The reason is the fact that these similarity measures are more sensitive to the region with a large intensity difference (for instance due to bright artifacts) than to the rest with a small intensity difference, where the important structures or image contents may be present.

Correlation (COR) of signals has been widely used for feature detection. In the field of image registration, it is defined as

$$\mathcal{D}_{\text{COR}}(\boldsymbol{\phi}) = \int_{\Omega} u^R(\boldsymbol{x}) u^T_{\boldsymbol{\phi}}(\boldsymbol{x}) \, d\boldsymbol{x}. \tag{2.7}$$

In practice, the variant cross correlation (CC) is more commonly used as a similarity measure for image registration.

$$\mathcal{D}_{\text{CC}}(\boldsymbol{\phi}) = \frac{\int_{\Omega} (u^R(\boldsymbol{x}) - \overline{u^R})(u^T_{\boldsymbol{\phi}}(\boldsymbol{x}) - \overline{u^T_{\boldsymbol{\phi}}}) \, d\boldsymbol{x}}{\sqrt{\int_{\Omega} (u^R(\boldsymbol{x}) - \overline{u^R})^2 \, d\boldsymbol{x}} \cdot \sqrt{\int_{\Omega} (u^T_{\boldsymbol{\phi}}(\boldsymbol{x}) - \overline{u^T_{\boldsymbol{\phi}}})^2 \, d\boldsymbol{x}}}. \tag{2.8}$$

Compared to simple correlation (see Equation 2.7), cross correlation is invariant to global changes in intensity amplitude, because it accounts for the mean intensities $\overline{u^R}$ and $\overline{u^T_{\boldsymbol{\phi}}}$. Another important correlation measure is the correlation ratio (CR), which was first introduced by Roche et.al [RPA98] for image registration.

$$\mathcal{D}_{\text{CR}}(\boldsymbol{\phi}) = \frac{\int_{\Omega} u^R(\boldsymbol{x}) u^T_{\boldsymbol{\phi}}(\boldsymbol{x}) \, d\boldsymbol{x}}{\sqrt{\int_{\Omega} u^R(\boldsymbol{x})^2 \, d\boldsymbol{x}} \cdot \sqrt{\int_{\Omega} u^T_{\boldsymbol{\phi}}(\boldsymbol{x})^2 \, d\boldsymbol{x}}}. \tag{2.9}$$

An advantage of the correlation ratio is that it is normalized between 0 and 1.

Mutual Information is one of the most successful voxel similarity measures for multi-modal registration [PMV03]. Its definition is based on entropy and probability. Actually, all similarity measures can also be interpreted from the view point of probability of random variables. To differentiate random variables from the other variables, all random variables have a hat over the letter in the thesis, e.g. $\widehat{x}$.

When one random variable $\widehat{x}$ is a function of another $\widehat{y}$, i.e $\widehat{y} = f(\widehat{x})$, the measurement of $\widehat{x}$ can fully predict $\widehat{y}$. But in most cases, two variables are related, but not fully predictable from each other. Measuring $\widehat{x}$ tells something about $\widehat{y}$, but not everything. Predictability can

be examined by the joint probability density function (pdf) $p(\widehat{x}, \widehat{y})$, which tells us about the co-occurrence of events from two random variables [VW97].

In the context of registration, the two random variables are $\widehat{i}^R$ and $\widehat{i}^T_{\boldsymbol{\phi}}$, denoting respectively the intensities of reference and template image (transformed by $\boldsymbol{\phi}$) at the same spatial position. Let us define a 2-dimensional random variable $\widehat{\boldsymbol{i}} = (\widehat{i}^R, \widehat{i}^T_{\boldsymbol{\phi}})$. Figures 2.3(a)-2.3(d) illustrates the possible states of probability density function of $\widehat{\boldsymbol{i}}$, denoted by $p_{\boldsymbol{\phi}}(\widehat{\boldsymbol{i}})$, when the similarity measures are maximized or minimized. Minimizing intensity difference measures results in that an identical mapping $\widehat{i}^R = \widehat{i}^T_{\boldsymbol{\phi}}$ more likely occurs. The cross correlation seeks to measure of an affine dependency of two variables. The correlation ratio expresses the level of functional dependency. If $\widehat{i}^R$ and $\widehat{i}^T_{\boldsymbol{\phi}}$ are independent from each other, the correlation ratio of them is zero. If they are functional dependent i.e. $\widehat{i}^R = f(\widehat{i}^T_{\boldsymbol{\phi}})$, the correlation ratio is maximized to 1. Maximization of mutual information is the most "unconstrained" compared to the other approaches. It tends to shift the random variable $\widehat{i}^T_{\boldsymbol{\phi}}$ by registration to build up a number of clusters in $p_{\boldsymbol{\phi}}(\widehat{\boldsymbol{i}})$.

Mutual information is defined as

$$\mathcal{D}_{\text{MI}}(\boldsymbol{\phi}) = \int_{\mathbb{R}^2} p_{\boldsymbol{\phi}}(\widehat{\boldsymbol{i}}) \log \frac{p_{\boldsymbol{\phi}}(\widehat{\boldsymbol{i}})}{p(\widehat{i}^R) p_{\boldsymbol{\phi}}(\widehat{i}^T_{\boldsymbol{\phi}})} \, d\widehat{\boldsymbol{i}} \;\text{ where }\; \widehat{\boldsymbol{i}} = (\widehat{i}^R, \widehat{i}^T_{\boldsymbol{\phi}}). \tag{2.10}$$

Functions $p(\widehat{i}^R)$ and $p_{\boldsymbol{\phi}}(\widehat{i}^T_{\boldsymbol{\phi}})$ denote the marginal probability density functions of $\widehat{i}^R$ and $\widehat{i}^T_{\boldsymbol{\phi}}$. The definition of mutual information can be also interpreted as a kind of Kullback-Leibler divergence, which has the form of $\int p(\widehat{i}) \log(p(\widehat{i})/q(\widehat{i})) \; d\widehat{i}$ for two density functions p and q. Equation 2.10 is the Kullback-Leibler divergence between the joint probability density function $p_{\boldsymbol{\phi}}(\widehat{\boldsymbol{i}})$ and the one in the case of full independency $p(\widehat{i}^R) p_{\boldsymbol{\phi}}(\widehat{i}^T_{\boldsymbol{\phi}})$.

2.4 Estimation of Transformation Parameters

A rigid transformation for image registration can be uniquely defined by a number of parameters, as in Equation 2.2, the transformation has 7 parameters, 4 dual quaternion parameters and 3 translation parameters. The non-rigid transformation is defined by the d-dimensional function $\boldsymbol{u}$. For this reason, in some literatures of registration, non-rigid transformation is also called non-parametric transformation. In this section, we discuss the estimation of parameters for rigid registration. The estimation of non-rigid transformation will be investigated in the next section.

Gradient descent is the general method to solve the problem of optimization. It seeks to find a local extremum of a function in the parameter space by taking steps along the gradient direction of the function at the current position. Negative gradient for minimization and positive gradient

for maximization. Let us constrain the discussion to minimizing a given similarity measure $\mathcal{D}$ with respect to the unknown transformation parameter set $\boldsymbol{q}$. Vector $\boldsymbol{q}$ could be a concatenation of translations with rotation angles or with unit quaternions. The simple gradient descent is defined as in Algorithm 1, in which $\nabla_{\boldsymbol{q}}\mathcal{D}$ denotes the gradient of the similarity measure respect to the parameter vector $\boldsymbol{q}$. For a rigid transformation $\boldsymbol{\phi_t}$, the SSD similarity measure is

Algorithm 1 General Gradient Decent

while $\boldsymbol{t}$ has not yet converged **do**
$\boldsymbol{t}_{k+1} = \boldsymbol{t}_k - \tau \cdot \nabla_{\boldsymbol{q}}\mathcal{D}(\boldsymbol{t}_k)$
end while.

$$\mathcal{D}_{\text{SSD}}(\boldsymbol{\phi_q}) = \int_{\Omega} (u^R(\boldsymbol{x}) - u^T_{\boldsymbol{\phi_q}}(\boldsymbol{x}))^2 \, d\boldsymbol{x} \tag{2.11}$$

and the discretized formulation is

$$\mathcal{D}_{\text{SSD}}(\boldsymbol{\phi_q}) = \|\overrightarrow{U}^R - \overrightarrow{U}^T \circ \boldsymbol{\phi_q}\|^2 = \|\overrightarrow{U}^R - \overrightarrow{U}^T_{\boldsymbol{q}}\|^2,$$

where the image function u is represented by a vector $\overrightarrow{U}$ by stacking all the pixels in a given order. The gradient with respect to a single parameter q_j is computed by

$$\nabla_{q_j}\mathcal{D}_{\text{SSD}} = \langle \overrightarrow{U}^T_{\boldsymbol{q}} - \overrightarrow{U}^R, \partial_{q_j}\overrightarrow{U}^T_{\boldsymbol{q}} \rangle. \tag{2.12}$$

The computation of the gradient of various similarity measures is discussed in [Her02]. The positive real number τ in the gradient decent algorithm is the so-called step size, which determines the distance approaching along the descent direction. Intuitively, it is a good idea to adapt τ in each optimization iteration. Various strategies are proposed for the automatic adaption of the step size. An algorithm proposed in [Ven02] finds the optimal τ with respect to the current parameters $\boldsymbol{q}_k$, i.e $\tau_k = \arg\min_\tau \mathcal{D}(\boldsymbol{q}_k - \tau\nabla_{\boldsymbol{q}}\mathcal{D})$. The step size τ can also be computed to make subsequent gradients orthogonal to each other, i.e. $\langle \nabla\mathcal{D}(\boldsymbol{q}_k), \nabla\mathcal{D}(\boldsymbol{q}_{k-1}) \rangle = 0$. But the most reliable way to estimate τ is the so-called Armijo rule [Kos91]: Given an initial large step size $s > 0$, the factor of reduction $\beta \in (0, 1)$ and the fixed parameter of tolerance $\sigma \in (0, 1)$, the step size $\tau_k = s\beta^{k_0}$ is optimized in the sense of the Armijo rule when $k_0 \in \{0, 1, 2, ...\}$ is the maximal integer such that

$$\mathcal{D}(\boldsymbol{q}_k) - \mathcal{D}(\boldsymbol{q}_k - s\beta^{k_0}\nabla\mathcal{D}_k) \leq \sigma s\beta^{k_0}\|\nabla\mathcal{D}_k\|. \tag{2.13}$$

2.5 Non-rigid Registration

Regarding non-rigid registration, the optimization of a similarity measure is sufficient, because the flexibility of non-rigid transformation can still lead to undesired matchings, for example cracks and overlaps of the transformation field. In contrast to the rigid registrations, a new regularization measure $\mathcal{S}$ must be added to the object function as the remedy for the arbitrary irregularity of non-rigid transformation $\boldsymbol{\phi} = \mathbb{R}^d \rightarrow \mathbb{R}^d$. A non-rigid transformation is usually represented by a displacement field $\boldsymbol{u} = \boldsymbol{x} - \boldsymbol{\phi}(\boldsymbol{x})$. The problem is formulated as finding a function $\boldsymbol{u}$, such that

$$\mathcal{D}(\boldsymbol{u}) + \alpha\mathcal{S}(\boldsymbol{u}) \rightarrow \min . \tag{2.14}$$

The parameter α weights the regular property of transformation versus the similarity of the images. Different from the similarity measure $\mathcal{D}$, which is also a function of images (u^R, u^T), the regularization measure $\mathcal{S}$ is only dependent on the geometric properties of the transformation $\boldsymbol{\phi}$. In the rest of this section, we will review various regularizers (Section 2.5.1) and briefly introduce the variational framework for the estimation of transformation (Section 2.5.2).

2.5.1 Regularization

Diffusion regularization is often used to solve the optical flow problem in computer vision. In the context of image registration, the regularizer constrains the gradient magnitude of the transformation by minimizing

$$\mathcal{S}_{\text{diff}}(\boldsymbol{u}) = \frac{1}{2}\sum_{l=1}^{d}\int_{\Omega} \|\nabla u_l\|_2^2 \, d\boldsymbol{x}, \tag{2.15}$$

where $u_l : \mathbb{R}^d \rightarrow \mathbb{R}$ is the l-th component of $\boldsymbol{u}$ and $\|\cdot\|_2$ denotes a L_2-norm. A fast numerical implementation based on the additive operator splitting method is proposed in [FM99, MF02]. This numerical solution makes diffusion regularization very attractive for fast three dimensional non-rigid registrations.

Curvature regularization was firstly proposed in [FM04] to as an approach for non-rigid registration. Different from diffusion regularization, the curvature regularizer constrains the mag-

nitude of second order derivative of transformation by minimizing

$$\mathcal{S}_{\text{curv}}(\boldsymbol{u}) = \frac{1}{2}\sum_{l=1}^{d}\int_{\Omega}(\triangle u_l)^2\, d\boldsymbol{x}, \tag{2.16}$$

where $\triangle$ is the Laplace operator, i.e. $\triangle u_l = \sum_{m=1}^{d}\partial_{x_m,x_m}u_l$. The integral approximates the curvature in the l-th dimension of the displacement field and therefore penalizes oscillations of transformation. In theory, curvature regularization is invariant with respect to an affine transformation because $\mathcal{S}_{\text{curv}}(\boldsymbol{B}\boldsymbol{x}+\boldsymbol{c}) = 0$.

Elastic regularization for image registration was first investigated in [Bro81]. Elastic registration seeks to deform the image as an elastic object under external forces. The regularization forces the deformation subject to elasticity constraints, while the similarity measure works like the force resulting strain of the object. Elastic regularization based on the linearized elastic model is defined as

$$\mathcal{S}_{\text{elas}}(\boldsymbol{u}) = \frac{\mu}{4}\int_{\Omega}\sum_{l,m=1}^{d}(\partial_{x_l}u_m + \partial_{x_m}u_l)^2 + \frac{\lambda}{2}(\operatorname{div}\boldsymbol{u})^2\, d\boldsymbol{x}, \tag{2.17}$$

where λ and μ are the so-called Lamé-constants reflecting material properties of an elastic body.

Fluid regularization is another physically motivated regularization method for non-rigid registration introduced in [Chr94]. We can imagine that the deformation field $\boldsymbol{u}$ is not only a function of space but also of time, since $\boldsymbol{u}$ evolves from the current to the next iteration during the optimization, i.e $\boldsymbol{u}(\boldsymbol{x},t)$, where t is a time variable. In this way, a velocity field is defined by

$$\boldsymbol{v}(\boldsymbol{x},t) = \partial_t\boldsymbol{u}(\boldsymbol{x},t) + \nabla\boldsymbol{u}(\boldsymbol{x},t)\boldsymbol{v}(\boldsymbol{x},t). \tag{2.18}$$

The essential difference between elastic and fluid regularization is that a fluid regularizer does not directly constrain the displacement field $\boldsymbol{u}$ but the velocity field $\boldsymbol{v}$.

$$\mathcal{S}_{\text{flui}}(\boldsymbol{u}) = \mathcal{S}_{\text{elas}}(\boldsymbol{v}) \tag{2.19}$$

Compared with elastic registration, fluid registration is more deformable and allows large image deformations while preserving the topology of the object.

2.5.2 Variational Framework

One of the most popular numerical solutions for the minimization problem as in Equation 2.14 is to solve the correspondent Euler equation

$$\boldsymbol{f}[\boldsymbol{u}](\boldsymbol{x}) + \alpha \mathcal{A}[\boldsymbol{u}](\boldsymbol{x}) = \mathbf{0} \;\text{ for }\; \boldsymbol{x} \in \Omega, \tag{2.20}$$

where $\boldsymbol{f}$ and $\mathcal{A}$ are respectively the first variations of the similarity measure $\mathcal{D}$ and the regularization term $\mathcal{S}$. Assuming the objective function $\mathcal{I}(\boldsymbol{u})$ is sufficiently regularized in the function space $\mathcal{F}$, the first variation of $\mathcal{I}(\boldsymbol{u})$ with respect to $\boldsymbol{u} \in \mathcal{F}$ in the direction $\boldsymbol{k} \in \mathcal{F}$ is defined by

$$\delta\mathcal{I}(\boldsymbol{u}, \boldsymbol{k}) = \lim_{\epsilon \to 0} \frac{\mathcal{I}(\boldsymbol{u} + \epsilon \boldsymbol{k}) - \mathcal{I}(\boldsymbol{u})}{\epsilon}. \tag{2.21}$$

If $\boldsymbol{u}$ minimizes $\mathcal{I}$, then $\delta\mathcal{I}(\boldsymbol{u}, \boldsymbol{k}) = 0$ for every $\boldsymbol{k} \in \mathcal{F}$. Let us take the $\mathcal{S}_{\text{diff}}$ as an example to compute the variation:

$$\begin{aligned}
\delta\mathcal{S}_{\text{diff}}(\boldsymbol{u}, \boldsymbol{k}) &= \lim_{\epsilon \to 0} \frac{1}{\epsilon}(\mathcal{S}(\boldsymbol{u} + \epsilon \boldsymbol{k}) - \mathcal{S}(\boldsymbol{u})) \\
&= \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_{\Omega} \sum_{l=1}^{d} (\|\nabla(u_l + \epsilon k_l)\|_2^2 - \|\nabla u_l\|_2^2)\, d\boldsymbol{x} \\
&= \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_{\Omega} \sum_{l=1}^{d} (\|\nabla u_l\|^2 + \epsilon^2 \|\nabla k_l\|^2 + 2\epsilon \nabla u_l \nabla k_l - \|\nabla u_l\|^2)\, d\boldsymbol{x} \\
&= \int_{\Omega} \sum_{l=1}^{d} \langle \nabla u_l, \nabla k_l \rangle\, d\boldsymbol{x}
\end{aligned}$$

If Neumann boundary conditions are imposed, i.e.,

$$\langle \nabla u_l, \boldsymbol{n} \rangle = \langle \nabla k_l, \boldsymbol{n} \rangle = 0 \quad \text{for} \quad x \in \partial\Omega,$$

where $\boldsymbol{n}$ denotes the outer normal unit vector of $\partial\Omega$,

$$\begin{aligned}\int_\Omega \sum_{l=1}^{d} \langle \nabla u_l, \nabla k_l \rangle \, d\boldsymbol{x} &= \int_\Omega \sum_{l=1}^{d} \triangle u_l k_l \, d\boldsymbol{x} \\ &= \int_\Omega \langle \triangle \boldsymbol{u}, \boldsymbol{k} \rangle \, d\boldsymbol{x} \\ &= \int_\Omega \langle \mathcal{A}_{\text{diff}}[\boldsymbol{u}], \boldsymbol{k} \rangle \, d\boldsymbol{x}.\end{aligned}$$

Here $\triangle$ is a Laplace operator. In the similar way, the variation of regularization terms are computed as

$$\mathcal{A}_{\text{diff}}[\boldsymbol{u}] = \triangle \boldsymbol{u} \tag{2.22}$$

$$\mathcal{A}_{\text{curv}}[\boldsymbol{u}] = \triangle^2 \boldsymbol{u} \tag{2.23}$$

$$\mathcal{A}_{\text{elas}}[\boldsymbol{u}] = \mu \triangle \boldsymbol{u} + (\lambda + \mu) \nabla \, \text{div} \boldsymbol{u} \tag{2.24}$$

$$\mathcal{A}_{\text{flui}}[\boldsymbol{u}] = \mu \triangle \boldsymbol{v} + (\lambda + \mu) \nabla \, \text{div} \boldsymbol{v}. \tag{2.25}$$

All operators $\mathcal{A}$ in equations 2.22-2.25 are linear operators applying on the unknown displacement field $\boldsymbol{u}$, which can be formulated as matrix multiplications if the Euler equation is represented into a linear equation system.

The variation of similarity measures, the so-called force terms are computed by

$$\boldsymbol{f}_{\text{SSD}}[\boldsymbol{u}] = (u^R(\boldsymbol{x}) - u^T(\boldsymbol{x} - \boldsymbol{u})) \nabla u^T(\boldsymbol{x} - \boldsymbol{u}) \tag{2.26}$$

$$\boldsymbol{f}_{\text{CC}}[\boldsymbol{u}] = (G_\sigma * L_{\boldsymbol{u}}^{\text{CC}})(\widehat{\boldsymbol{i}}) \nabla u^T(\boldsymbol{x} - \boldsymbol{u}) \tag{2.27}$$

$$\boldsymbol{f}_{\text{CR}}[\boldsymbol{u}] = (G_\sigma * L_{\boldsymbol{u}}^{\text{CR}})(\widehat{\boldsymbol{i}}) \nabla u^T(\boldsymbol{x} - \boldsymbol{u}) \tag{2.28}$$

$$\boldsymbol{f}_{\text{MI}}[\boldsymbol{u}] = (G_\sigma * L_{\boldsymbol{u}}^{\text{MI}})(\widehat{\boldsymbol{i}}) \nabla u^T(\boldsymbol{x} - \boldsymbol{u}). \tag{2.29}$$

In above equations, $L^{\text{CC}}(\widehat{\boldsymbol{i}})$, $L^{\text{CR}}(\widehat{\boldsymbol{i}})$ and $L^{\text{MI}}(\widehat{\boldsymbol{i}})$ are the functions of the probability density functions of the 2-dimensional random variable of image intensities, see the discussion on page 17. They are convoluted with a Gaussian filter G_σ, because the joint probability density function $p_{\boldsymbol{u}}$ must be estimated by Parzen kernel with width σ

$$p_{\boldsymbol{u}}(\widehat{\boldsymbol{i}}) = \frac{1}{\Omega} \int_\Omega G_\sigma(\boldsymbol{I}_{\boldsymbol{u}}(\boldsymbol{x}) - \widehat{\boldsymbol{i}}) \, d\boldsymbol{x}, \tag{2.30}$$

where $\boldsymbol{I}_{\boldsymbol{u}}(\boldsymbol{x}) = (u^R(\boldsymbol{x}), u_\phi^T(\boldsymbol{x}))$ is a 2-dimensional intensity function. The exact mathematical

formulations of $L^{\mathrm{CC}}(\boldsymbol{i})$, $L^{\mathrm{CR}}(\boldsymbol{i})$ and $L^{\mathrm{MI}}(\boldsymbol{i})$ are out of range of this work. We refer to [Her02] for the further reading.

Regardless of the chosen similarity, the associated force term $\boldsymbol{f}$ is always non-linear with respect to the unknown $\boldsymbol{u}$ because of the gradient field of the deformed template $\nabla u^T(\boldsymbol{x}-\boldsymbol{u})$. Consequently we cannot analytically solve Equation 2.20 without an iterative method. Assuming the vector function $\boldsymbol{u}$ is sampled and the sample values are ordered into a discrete vector $\overrightarrow{U}$, the linear operator can be formulated as a matrix multiplication, i.e $\mathcal{A}[\boldsymbol{u}] \rightarrow \boldsymbol{A}\overrightarrow{U}$, typically a finite difference approximation is used. A simple fix-point iteration scheme like

$$\overrightarrow{U}^{(k+1)} = -\frac{1}{\alpha}\boldsymbol{A}^{-1}\overrightarrow{F}^{(k)}, \text{with } \overrightarrow{U}^{(0)} = \overrightarrow{0} \tag{2.31}$$

can be used to solve Equation 2.20. In each iteration, the force term $\boldsymbol{f}$ based on the $\boldsymbol{u}$ in the last step is the right-hand vector. More often, a time-stepping iteration scheme is employed

$$\partial_t \boldsymbol{u}(\boldsymbol{x},t) = \boldsymbol{f}[\boldsymbol{u}]((\boldsymbol{x},t)) + \alpha\mathcal{A}[\boldsymbol{u}](\boldsymbol{x},t). \tag{2.32}$$

Now the unknown $\boldsymbol{u}$ becomes a not only spatial but also time dependent function, which evolves during the iteration. The basic idea is that the solution of Equation 2.20 is found, as the solution converges, i.e. $\partial_t \boldsymbol{u}(\boldsymbol{x},t) = 0$. Equation 2.32 is often solved in a semi-implicit iterative scheme

$$\partial_t \boldsymbol{u}^{(k+1)}(\boldsymbol{x},t) - \alpha\mathcal{A}[\boldsymbol{u}^{(k+1)}](\boldsymbol{x},t) = \boldsymbol{f}[(\boldsymbol{u}^{(k)}](\boldsymbol{x},t)) \tag{2.33}$$

A time-step $\tau > 0$ is introduced for the time discretization and $\partial_t \boldsymbol{u}^{(k+1)}(\boldsymbol{x},t)$ is approximated by $(\boldsymbol{u}^{(k+1)} - \boldsymbol{u}^{(k)})/\tau$. In the end, we obtain the iteration scheme like

$$\overrightarrow{U}^{(k+1)} = (\boldsymbol{I} - \tau\alpha\boldsymbol{A})^{-1}(\overrightarrow{U}^{(k)} + \tau\overrightarrow{F}^{(k)}), \text{with } \overrightarrow{U}^{(0)} = \overrightarrow{0}. \tag{2.34}$$

No matter if the fix-point or the time-stepping iteration scheme is used, the most expensive computation is always the inversion of the matrix. Many efficient numerical solvers, e.g the multigrid method, have been developed or adapted to this problem. The parameter τ does not have to be pre-fixed, it can be dynamically determined based on the principles discussed in Section 2.4.

2.6 Feature based Registration

Feature based registration methods rely on the limited set of features generated from image content, such as identified points, segmented binary structures or object surfaces to register images. Compared to intensity based registration, feature based methods more likely emphasize matching of image contents, not of intensity patterns. Automatic or interactive selection of the desired features is a pre-condition of the registration, thus, the error of feature location may influence the quality of registration. In this section we will briefly discuss various methods of feature based registration.

2.6.1 Point based Match

Point based registration seeks to find a transformation that approximately aligns two groups of given points in two spaces. The points can be anatomical landmarks interactively selected by the user e.g. fiducial landmarks, or automatically identified by algorithms e.g. salient points. They can also be feature points with some given geometrical properties, e.g, corners or local curvature extrema.

If the correspondences between two groups of points are given, the transformations usually can be determined analytically. For rigid transformations, it can be formulated as a least squares fitting problem: given two groups of points $\boldsymbol{t} = (\boldsymbol{t}_1, \boldsymbol{t}_2, ..., \boldsymbol{t}_{n_s})^T$ and $\boldsymbol{r} = (\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_{n_s})^T$ with one-to-one correspondence, search for the transformation $\boldsymbol{\phi}$ minimizing the squared Procrustes distance described by

$$\sum_{s=1}^{n_s} \|\boldsymbol{t}_s \circ \boldsymbol{\phi} - \boldsymbol{r}_s\|^2. \tag{2.35}$$

The translation vector can be computed from the difference between the centroids $\bar{\boldsymbol{t}}$ and $\bar{\boldsymbol{s}}$, subsequently the rotation matrix can be determined by singular value decomposition (SVD) of the correlation matrix

$$\boldsymbol{K} \;=\; \boldsymbol{P}^T\boldsymbol{Q} = \left[\; \boldsymbol{t}_1 - \bar{\boldsymbol{t}}, \boldsymbol{t}_2 - \bar{\boldsymbol{t}}, ..., \boldsymbol{t}_{n_s} - \bar{\boldsymbol{t}} \;\right]^T \left[\; \boldsymbol{r}_1 - \bar{\boldsymbol{r}}, \boldsymbol{r}_2 - \bar{\boldsymbol{r}}, ..., \boldsymbol{r}_{n_s} - \bar{\boldsymbol{r}} \;\right], \tag{2.36}$$

i.e. $\boldsymbol{K} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T$ and the rotation matrix $\boldsymbol{R} = \boldsymbol{V}\boldsymbol{U}^T$.

Landmark based interpolation

The transformations can also be deformable. The problem can be treated as a point based interpolation where landmark correspondences are given sparse data in a multi-dimensional scalar field. We know the deformation $\boldsymbol{\phi}(\boldsymbol{t}_i) = \boldsymbol{r}_i$ and wish to estimate the value of the deformation field at the other spatial locations. There are the kriging [MKGL96], thin-plate spline [Boo89] and radial basis function approaches [Mic86]. Each method uses different constraints and assumptions to estimate the deformation outside the landmarks. Here we briefly summarize the general solution of the thin-plate spline method. An extensive discussion can be found in [Roh01].

For thin-plate spline method, the deformation is defined by a set of landmarks. Usually, the landmarks are pairs of points that are selected from the reference and template image. The deformation needs to exactly match these landmarks and keep smooth elsewhere. Because we handle SPECT and CT volume data here, let us constrain the discussion in a 3-D case, where the coordinate of a three-dimensional point is $\boldsymbol{x} = (x_1, x_2, x_3)^T$. Assume that n_s points $\boldsymbol{t}_s = (t_1^s, t_2^s, t_3^s)^T$ need to be transformed to new positions $\boldsymbol{r}_s = (r_1^s, r_2^s, r_3^s)^T$, respectively, the desired transformation with the manner $\boldsymbol{\phi}(\boldsymbol{x}) = (\phi_1(\boldsymbol{x}), \phi_2(\boldsymbol{x}), \phi_3(\boldsymbol{x}))^T$ is constrained by

$$\phi_1(\boldsymbol{t}_s) = r_1^s, \phi_2(\boldsymbol{t}_s) = r_2^s, \phi_3(\boldsymbol{t}_s) = r_3^s. \tag{2.37}$$

According to the theory of landmark based interpolation [Roh01], point $\boldsymbol{x}$ that is transformed by such function $\boldsymbol{\phi}$ that

$$\phi_l(\boldsymbol{x}) = \underbrace{\sum_{j=1}^{M} a_l^j g_j(\boldsymbol{x})}_{\text{linear part}} + \underbrace{\sum_{r=1}^{n_s} b_l^s \sigma(\boldsymbol{x}, \boldsymbol{t}_s)}_{\text{radius base part}}, l = 1, 2 \text{ or } 3. \tag{2.38}$$

In the linear part, g_j's are the linear basis function given by $\boldsymbol{g}(\boldsymbol{x}) \in \mathbb{R}^M$, where $M = 4$ and

$$g_1(\boldsymbol{x}) = 1, g_2(\boldsymbol{x}) = x_1, g_3(\boldsymbol{x}) = x_2, g_4(\boldsymbol{x}) = x_3. \tag{2.39}$$

In the radius base part, $\sigma(\boldsymbol{x}, \boldsymbol{t}_s)$ is usually a function of the radius $\boldsymbol{h}_s = \boldsymbol{x} - \boldsymbol{t}_s$. The coefficients in Equation 2.38 are the solution of the equation system

$$\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{G} \\ \boldsymbol{G}^T & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{A} \end{pmatrix} = \begin{pmatrix} \boldsymbol{U} \\ \boldsymbol{0} \end{pmatrix}, \tag{2.40}$$

where

$$
\begin{aligned}
&\boldsymbol{\Sigma}_{ij} = \sigma(\boldsymbol{t}_i, \boldsymbol{t}_j),\\
&\boldsymbol{G}_{ij} = g_j(\boldsymbol{t}_i), i = 1, ..., n_s, j = 1, ..., M, M = 4\\
&\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3), \boldsymbol{a}_l = (a_l^1, ..., a_l^M)^T, l = 1, 2 \text{ or } 3,\\
&\boldsymbol{B} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3), \boldsymbol{b}_l = (b_l^1, ..., b_l^{n_s})^T, l = 1, 2 \text{ or } 3,\\
&\boldsymbol{U} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3), \boldsymbol{u}_l = (u_l^1, ..., u_l^{n_s})^T, l = 1, 2 \text{ or } 3.
\end{aligned}
$$

In order to ensure that $\boldsymbol{\Sigma}$ is invertible, we choose the basis function of a thin-plate spline interpolation, $\sigma(\boldsymbol{x}, \boldsymbol{t}_s) = \|\boldsymbol{h}_s\|_2^2 \log \|\boldsymbol{h}_s\|_2$ in two dimensions and $\sigma(\boldsymbol{x}, \boldsymbol{t}_s) = \|\boldsymbol{h}_s\|$ in three dimensions. There are many functions which also ensure a nonsingular and invertible $\boldsymbol{\Sigma}$ [Buh03].

Iterative closest point (ICP) algorithm

For matching the point sets without one-to-one correspondence, for example, matching the sets of salient points detected from two images, the most widely used methods are based on the iterative closest point (ICP) algorithm. In the following, the basic concept of the ICP algorithm is introduced. We refer to the survey [RL01] for its variants.

Let say that we have a fixed point cloud $\boldsymbol{t} = (\boldsymbol{t}_1, \boldsymbol{t}_2, ..., \boldsymbol{t}_{n_s})^T$ and a floating point cloud $\boldsymbol{r} = (\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_{n_s'})^T$, where numbers of two point groups are normally different, i.e. $n_s \neq n_s'$. We want to estimate a rigid transformation $\boldsymbol{\phi}$ to match them. In each iteration, the ICP algorithm actually selects the closest points as correspondences. For instance, in the k-th iteration, the floating point cloud is moved by current transformation and is denoted as $\boldsymbol{r}^{(k)}$. The searching operation can be denoted as

$$\boldsymbol{t}^{(k)} = \mathcal{C}(\boldsymbol{r}^{(k)}, \boldsymbol{t}). \tag{2.41}$$

Among the fixed points cloud $\boldsymbol{t}$, the operation $\mathcal{C}$ looks for the closest point $\boldsymbol{t}_s$ with respect to the the point $\boldsymbol{r}_s \in \boldsymbol{r}^{(k)}$ and finally generates the set of correspondent points $\boldsymbol{t}^{(k)} = \{\boldsymbol{t}_s\}_{s=1,...,n_s}$. Then, the ICP algorithm calculates the transformation $\boldsymbol{\phi}$ for minimizing the distance between point set $\boldsymbol{t}^{(k)}$ and $\boldsymbol{r}^{(k)}$:

$$d(\boldsymbol{t}^{(k)}, \boldsymbol{r}^{(k)}) = \frac{1}{n_s} \sum_{s=1}^{n_s} \|\boldsymbol{t}_s^{(k)} - \boldsymbol{r}_s^{(k)} \circ \boldsymbol{\phi}^{-1}\|^2.$$

The transformation can be calculated based on any of these four methods: A SVD based method

by Arun et al. [AHB87], a quaternion method by Horn [Hor87], an algorithm using orthogonal matrices by Horn et al. [HHN88] and a calculation based on dual quaternions by Walker et al. [WSV91]. These algorithms show similar performance and stability concerning noisy data [LEF]. Similar to Equation 2.41, the least squares registration can be denoted as:

$$(\boldsymbol{\phi}, d) = \mathcal{Q}(\boldsymbol{t}_s^{(k)}, \boldsymbol{r}_s^{(k)}) \tag{2.42}$$

The ICP algorithm can now be stated as follows:

Algorithm 2 Iterative Closest Point (ICP)

given two meshes $(\boldsymbol{t}, \boldsymbol{r})$ and maximal iteration number $N_{\max}$
set $\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{x}$ and $\boldsymbol{t}^{(0)} = \boldsymbol{t}$
for $k = 0, ..., N_{\max}$ **do**
  Compute the correspondence: $\boldsymbol{t}^{(k)} = \mathcal{C}(\boldsymbol{r}^{(k)}, \boldsymbol{t})$
  Registration: $(\boldsymbol{\phi}^{(k)}, d) = \mathcal{Q}(\boldsymbol{t}_s^{(k)}, \boldsymbol{r}_s^{(k)})$
  Transform the floating points: $\boldsymbol{r}^{(k+1)} = \boldsymbol{\phi}^{(k)}(\boldsymbol{r})$
  if convergence **then**
    break
  end if
end for

2.6.2 Structure based Registration

Structure based methods register images by aligning the common structures found in both images. In [MV97, MCOS+02, ZF03] the most important structure based registration methods in the last decade have been summarized. The most common structures, such as surfaces, edges or contours, are represented by a set of feature points, which can easily be handled with point based matching methods, for instances, the ICP based method presented above.

A kind of "Head and Hat" algorithm is a simple but also widely used method in clinical practice to register 3-D CT, MR and PET images. First the skin surfaces in both images are segmented with less computational complexity. Then the algorithm models the contours from the higher resolution image as a surface ("header") and the contours of the lower resolution image a series of points ("hat"). The optimized rigid transformation is determined such that the mean squared deviation between "hat points" and "head contour" is minimized.

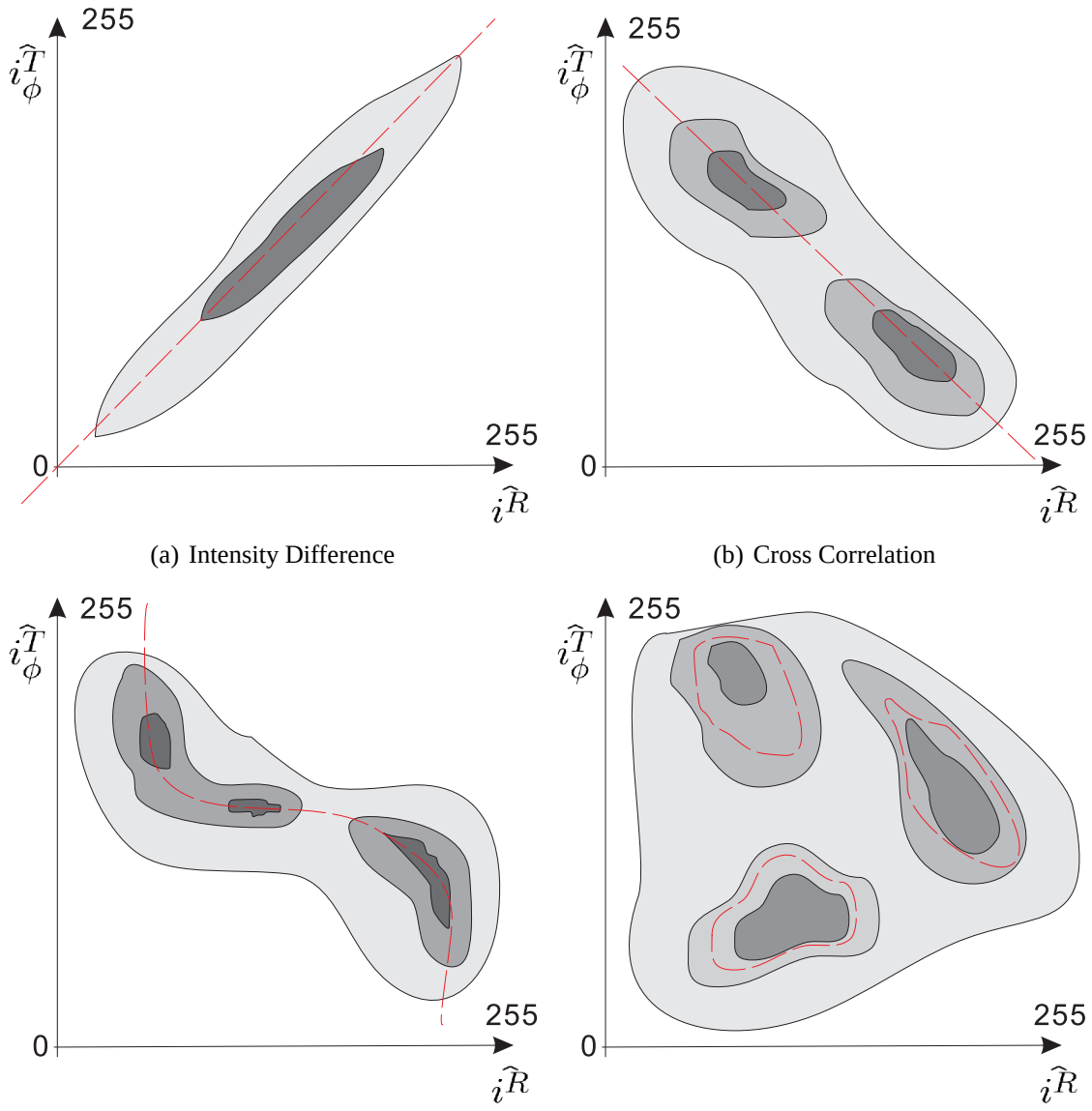
The matching of distance mapping is another popular way for the alignment of segmented binary structures. The basic idea is that the surface or contour feature is represented by a distance function. An comparative survey of Euclidean distance transform algorithms can be

found in [FBTC07]. Then the distance functions of two images are rigidly [Bor88] or non-rigidly [Par03] registered to each other. The drawback of segmentation based registration is that the accuracy of registration is influenced by the error in the segmentation step. Although the registration part is commonly automatic, the segmentations of common structures in reference and template images quite often requires user interaction.

2.6.3 Simultaneous Segmentation and Registration

Several attempts have been published in literature to develop methods for detecting the features and aligning images simultaneously. In 2001, the first hybrid segmentation and registration framework was proposed in [ZYK01]. The method utilizes multi-channel Chan-Vese active contour to segment the desired edge features and find the optimal Euclidean transformation between images. In 2002, Moelich improved this framework by substituting the Chan-Vese active contour with logic models that allow better control of the segmentation and a richer context information about dissimilarity of images [MC03]. In 2004, Chen presented a joint framework of classification and registration for MR data [CW98]. This was achieved by a maximizing a posteriori (MAP) model. In 2005, Young introduced a method that combines partial differential equations based on morphing active contours with Yezzi and Zollei's algorithms for joint segmentation and registration [YL05]. In the same year, a statistical framework using an Expectation Maximization based algorithm appeared in [PFL$^+$05]. The approach simultaneously estimates image inhomogeneities, anatomical label-map and a mapping from the atlas to the image space.

Due to our knowledge, most the existing approaches are restricted to lower dimensional rigid transformations for image registration. Recently, in [DR06, DRR09] a novel approach for non–rigid registration by edge alignment was presented. The key idea of this work is to modify the Ambrosio–Tortorelli approximation of the Mumford–Shah model, which is traditionally used for image segmentation, so that the new functional can also estimate the spatial transformation between images. This method is actually the theoretical fundament of the one-to-one edge matching method introduced in the next chapter.



(a) Intensity Difference

(b) Cross Correlation

(c) Correlation Ratio

(d) Mutual Information

Figure 2.3: Sketch of possible states of the joint probability function by minimizing intensity difference, cross correlation, correlation ratio and mutual information. Axes $\widehat{i^R}$ and $\widehat{i_\phi^T}$ denote the random variables of the gray levels (0-255) of reference and template images, respectively. The gray levels shown in sketches tell us the degrees of concentration of joint probability functions.

Part II

One-to-one Edge based Registration

Chapter 3

One-to-one Edge based Registration

Edge features, usually determined by intensity discontinuities, reflect the underlying structure information among the images. Detecting and matching edge features are two important and challenging image processing problems in the fields of computer vision and medical image analysis. Typically, solutions are developed for each of these two problems mutually independent. However, in various applications, the solutions of these problems depend on each other. Tackling each task would benefit from prior knowledge of the solution of the other task. This advantage has already been pointed out in [KYZ01]. In this work, edges in different images are segmented by an active edge model, similar to the one proposed in [CV01], and images are simultaneously matched to each other with an affine transformation.

In 2007, Mumford–Shah model [MS89] was expanded in [DR06, DRR09] with the capability of matching the edge features of two images. The edge features are represented by two different cartoon approximations of the images. A smooth dense warping function defines the mapping between the edge features. The modified Mumford–Shah model seeks to simultaneously tackle two highly interdependent tasks: edge segmentation and non-rigid registration. However, the non-symmetrical functional definition and the transformation model constrain the applicability of this model. This drawback will be further discussed in Section 3.2.

In this work we introduce a new symmetric model for edge matching based on the Mumford–Shah model as well. We use two relatively separated discontinuity sets to explicitly represent the edge sets of the associated images. For the ambiguity problem of the correspondence, we apply the idea of consistent registration [CJ01, JC02] to simultaneously estimate the forward and reverse transformations and to constrain one transformation to be the inverse of the other one. In this way, the edge sets of the images have equal influence on the edge registration. Thus, the proposed method is one-to-one in the sense, that it allows to determine one-to-one correspondences

between the edge features of two images. Symmetric one-to-one edge matching is not only more sound in mathematical sense, but also very important in a broad range of applications, where one is interested in determining the correspondence of the same structure in different images. For example, non-rigid registration for atlas construction [RFS03, MTT03], historical biological images [STU05, CSM+06] or motion estimation.

This chapter is organized as follows: In Section 3.1, we introduce some basic knowledge about the classic Mumford–Shah model, the approximation proposed by Ambrosio and Tortorelli and the Finite Element approximation as a preparation for the discussion of the proposed method. In Section 3.2, we present the non-symmetrical Mumford–Shah model for edge matching and discuss the potential drawbacks of this model. Then, in Section 3.3, the symmetrical model is introduced, including functional definitions, variational formulations, numerical implementations and algorithm. In Section 3.4, we study the parameter setting of the algorithm and show experimental results. Finally, we summarize the method of one-to-one edge based registration in Section 3.5.

3.1 Fundamentals

3.1.1 Mumford-Shah Model

In their pioneering work [MS89], Mumford and Shah proposed to model an image as follows: For an image function $u : \Omega \to \mathbb{R}$ on an image domain $\Omega \subset \mathbb{R}^d$ with $d = 2$ or 3 and non-negative constants α, β and ν, the Mumford–Shah (MS) functional is given by

$$\begin{aligned} E_{\text{MS}}(w, K) =& \frac{\alpha}{2} \int_{\Omega} (w - u)^2 \, \mathrm{d}\boldsymbol{x} \\ &+ \frac{\beta}{2} \int_{\Omega \setminus K} |\nabla w|^2 \, \mathrm{d}\boldsymbol{x} + \frac{\nu}{2} \mathcal{H}^{d-1}(K). \end{aligned} \tag{3.1}$$

By minimizing this functional, the noisy image function u, has been represented by a cartoon set (w, K), where $w : \Omega \to \mathbb{R}$ is a reconstructed noise-free piecewise constant function, while K is a set of discontinuity in Ω. The first term measures the degree of fidelity of the approximation w with respect to the input data u. The second term acts as a kind of "edge-preserving smoother", which penalizes large gradients of w in regions except of K, i.e. $\Omega \setminus K$, while not smoothing the image in the edge set. The last term $\mathcal{H}^{d-1}$ denotes the $d-1$ dimensional Hausdorff measure, which is used to regularize the reconstruction of w and control the length of the edge set. Existence theory for Mumford-Shah model established in [CDR02] proposed to consider

the minimization of an equivalent energy depending on w only. The Mumford-Shah model has turned out to be versatile and has been applied widely in image segmentation, denosing, shape modeling, image inpaint and data-reconstruction. See [MS95, CSV00, CV01, Fri09] and the references therein.

3.1.2 Ambrosio–Tortorelli Approximation

It is difficult to minimize the original Mumford–Shah functional in Equation 3.1 because of its implicit definition of the discontinuity set K. Various approximations have been proposed during the last two decades. In this work we focus on the Ambrosio–Tortorelli approximation with elliptic functionals [AT90].

In the Ambrosio–Tortorelli (AT) approximation the discontinuity set K is expressed by a phase field function v. This scalar function v approximates the characteristic function of the complement of K, i.e., $v(\boldsymbol{x}) \approx 0$ if $\boldsymbol{x} \in K$ and $v(\boldsymbol{x}) \approx 1$ otherwise. The approximation functional is defined as follows:

$$\begin{aligned} E^{\epsilon}_{\text{AT}}[w, v] =& \frac{\alpha}{2} \int_{\Omega} (w-u)^2 \,\mathrm{d}\boldsymbol{x} + \frac{\beta}{2} \int_{\Omega} v^2 \|\nabla w\|^2 \,\mathrm{d}\boldsymbol{x} \\ &+ \frac{\nu}{2} \int_{\Omega} (\epsilon \|\nabla v\|^2 \; + \frac{1}{4\epsilon}(v-1)^2) \,\mathrm{d}\boldsymbol{x}, \end{aligned} \tag{3.2}$$

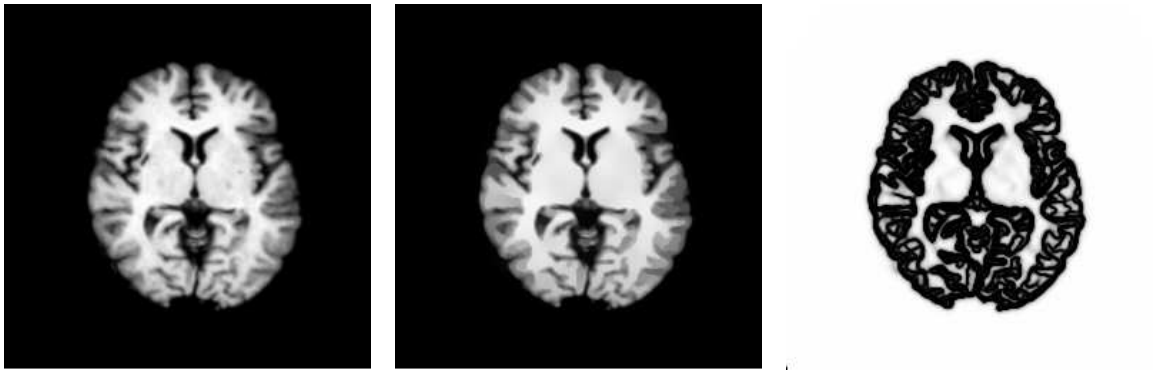
The second term, still working as an "edge-preserving smoother", couples zero-regions of v with regions where the gradient of w is large. The following "coupling" between w and v is energetically preferable:

$$v(x) \begin{cases} \approx 0 & \text{where } \|\nabla w\| \gg 0, \\ \approx 1 & \text{where } \|\nabla w\| \approx 0. \end{cases} \tag{3.3}$$

The last term approximates the edge length, i. e. the $d-1$ dimensional measure $\mathcal{H}^{d-1}(K)$ of the edge set K. The parameter ϵ controls the "width" of the diffusive edge set. Mathematically speaking, the sequence of functionals E^{ϵ}_{AT} $\Gamma-$converges to the Mumford–Shah functional, i.e.

$$\Gamma - \lim_{\epsilon \to 0} E^{\epsilon}_{\text{AT}} = E_{\text{MS}}.$$

For a rigorous proof and further explanation we refer to [Mar92].



(a) u (b) w (c) v

Figure 3.1: A 2-D example of Ambrosio–Tortorelli approximation. (a) The original image u. (b) The piecewise constant function w. (c) The phase field function v.

3.1.3 Finite Element Discretization

Finite Element (FE) methods are used in this work to discretize equations. The whole image domain Ω is covered by a uniform rectangular mesh $\mathcal{C}$, on which a standard multi-linear Lagrange finite element space is defined. We consider all images as sets of voxels, where each voxel corresponds to a grid node of $\mathcal{C}$. Let $\mathcal{N} = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\}$ denote the nodes of $\mathcal{C}$. The FE basis function of node x_i is defined as the piecewise multi-linear function that fulfills

$$\varphi_i(\boldsymbol{x}_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Figures 3.2(a) and 3.2(b) show the basis functions in 1-D and 2-D space. The FE-space $\mathcal{V}$ is the linear hull of φ_i, i.e.

$$\mathcal{V} := \operatorname{span}(\varphi_1, ..., \varphi_n).$$

The FE-space of vector valued functions is $\mathcal{V}^d$, the canonical basis of this space, is

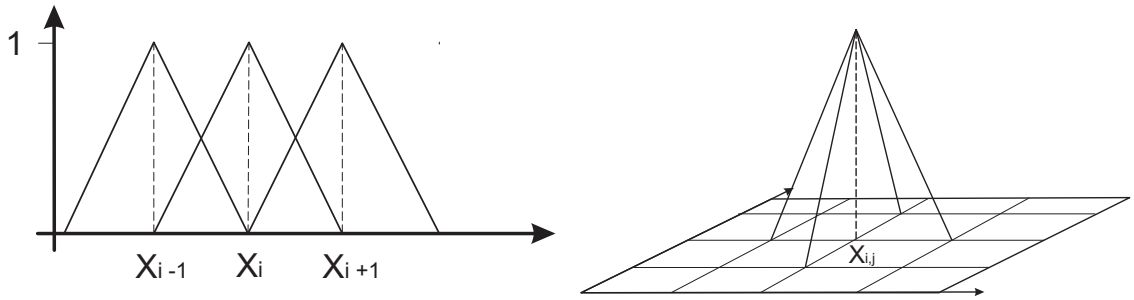
$$\varphi_1 e_1, ..., \varphi_n e_1, ..., \varphi_1 e_d, ..., \varphi_n e_d,$$

where e_i is the i-th canonical basis vector of $\mathbb{R}^d$. In FE-space scalar and vector valued functions, e.g. u and $\boldsymbol{\phi}$, are approximated by

$$u \approx \overline{U} := \sum_{i=1}^{n} u(\boldsymbol{x}_i)\varphi_i(\boldsymbol{x}) \quad \text{and}$$

$$\boldsymbol{\phi} = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_d \end{bmatrix} \approx \overline{\Phi} := \begin{bmatrix} \sum_{i=1}^{n} \phi_1(\boldsymbol{x}_i)\varphi_i(\boldsymbol{x}) \\ \vdots \\ \sum_{i=1}^{n} \phi_d(\boldsymbol{x}_i)\varphi_i(\boldsymbol{x}) \end{bmatrix}.$$

The FE approximation of a function can also be represented by a vector that collects the function values on the nodes, e.g. $\overrightarrow{U} := (u(\boldsymbol{x}_i), \cdots, u(\boldsymbol{x}_n))^T$ and $\overrightarrow{\Phi} := (\overrightarrow{\Phi_1}, \cdots, \overrightarrow{\Phi_d})^T$ where $\overrightarrow{\Phi_l} = (\phi_l(\boldsymbol{x}_1), \cdots, \phi_l(\boldsymbol{x}_n))^T$. In this work we denote continuous functions by lowercase letters (e.g. u or $\boldsymbol{\phi}$), their FE representation by "over-lined" uppercase letters (e.g. $\overline{U}$ or $\overline{\Phi}$) and their vector representation by "over-arrowed" uppercase letters (e.g. $\overrightarrow{U}$ or $\overrightarrow{\Phi}$).



(a) 1-D basis functions. (b) 2-D basis function in the rectangular mesh.

Figure 3.2: Basis functions of the finite element method.

3.2 Mumford-Shah Model for Edge Matching

A promising way to extend the Mumford-Shah model for deformable edge matching is introduced in [DR06, DRR09]. Given the reference u^R and the template image u^T, the functional is

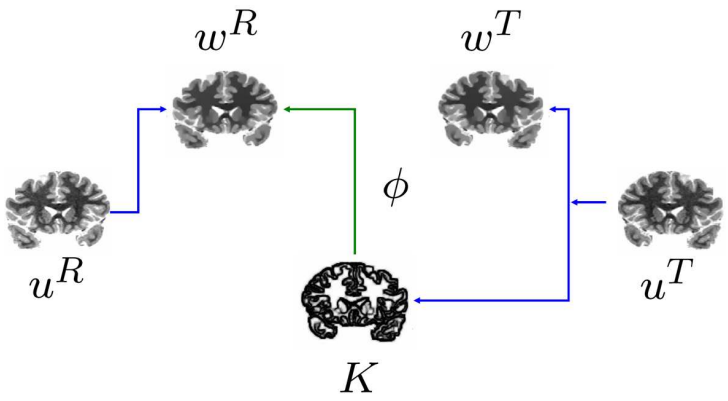


Figure 3.3: Non-symmetric Mumford–Shah model for edge matching. u^R and u^T are the given reference and template images. w^R and w^T are the restored, piecewise smooth functions of image R and image T. K is the combined discontinuity set of both images. Function ϕ represents the spatial transformation from image T to image R.

defined as

$$\begin{aligned} E_{\text{MSreg}}(w^R, w^T, K, \boldsymbol{\phi}) = &\frac{\alpha}{2}\int_\Omega (w^R - u^R)^2 \,\mathrm{d}x + \frac{\alpha}{2}\int_\Omega (w^T - u^T)^2 \,\mathrm{d}x \\ &+ \frac{\beta}{2}\int_{\Omega\setminus K} |\nabla w^R|^2 \,\mathrm{d}x + \frac{\beta}{2}\int_{\Omega\setminus K_{\boldsymbol{\phi}}} |\nabla w^T|^2 \,\mathrm{d}x \\ &+ \frac{\nu}{2}\mathcal{H}^{d-1}(K) + C_{\text{REG}}[\boldsymbol{\phi}]. \end{aligned} \tag{3.4}$$

Here w^R and w^T are the reconstructed piecewise constant functions of the reference and template images. $\boldsymbol{\phi}$ is the non-rigid transformation that K denotes the edge-set of reference image u^R, while $K_{\boldsymbol{\phi}}$ denotes the deformed edge-set K under the transformation $\boldsymbol{\phi}$. $C_{\text{REG}}[\boldsymbol{\phi}]$ is the regularization functional of the transformation $\boldsymbol{\phi}$, which we will further discuss in the next section. Matching the edge between two image data-sets is reflected by the fourth functional term in Equation 3.4. That is, in this model the edge-set K energetically not only prefers to couple the gradient of reference image ∇w^R but also to be transformed to couple the gradient of template image ∇w^T.

A major drawback of the above Mumford-Shah based matching is its asymmetry with respect to edge features and spatial mapping between them. The scheme of the model is shown in Figure 3.3. The definition of the similarity measure is not symmetrical: a joint discontinuity set K is used to estimate the edges of the restored template image T and the deformed edges of the restored reference image R. The model of the spatial mapping between the two images is not symmetrical: the transformation $\boldsymbol{\phi}$ in Figure 3.3 is only defined in one direction, from the

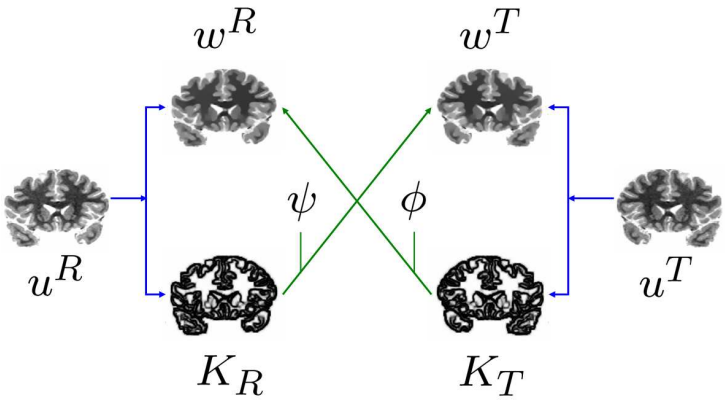


Figure 3.4: Symmetric Mumford–Shah model for one-to-one edge matching. u^R and u^T are the given images. w^R and w^T are the restored, piecewise smooth functions of image R and image T. K_R and K_T are the discontinuity sets of the images R and T, respectively. Function ϕ represents the transformation from image T to image R and function ψ represents the transformation from image R to image T.

image T to the image R. The asymmetry of the similarity measure and the single directional transformation, as pointed out in [RK06], cannot ensure that the method is consistent. That is, if one computes the transformation $\boldsymbol{\phi}$ from T to R and then switches the roles of T and R to compute the transformation $\boldsymbol{\psi}$ from R to T, it is uncertain whether these transformations are inverse to each other.

3.3 One-to-one Deformable Edge Matching

In this section we propose a new symmetric model for edge matching based on the Mumford–Shah model. Figure 3.4 shows the scheme of this symmetric model. We use two separated discontinuity sets (K_R and K_T in Figure 3.4) to explicitly represent the edge sets of the associated images. For the ambiguity problem of the correspondence, we apply the idea of consistent registration [CJ01, JC02] to simultaneously estimate the forward and reverse transformations and to constrain one transformation to be the inverse of the other one. In this way, the edge sets K_R and K_T of the images R and T, respectively, have equal influence on the edge registration. Thus, the proposed method is one-to-one in the sense that it allows to determine one-to-one correspondences between the edge features of two images.

The major task of one-to-one edge based registration is stated as follows: Find an appropriate transformation $\boldsymbol{\phi}$ such that the transformed template image $u^T \circ \boldsymbol{\phi}$ becomes similar to the

reference image u^R [Mod04]. The degree of similarity (or dissimilarity) is evaluated using the gray values u^R and u^T or certain features such as edges. We consider an edge based matching method that seeks to register two images based on joint edge extraction and registration. Thus, the algorithm simultaneously has to fulfill the two following tasks:

- Detection of the edge features from two noisy images.
- Registration of two images using these detected edge features.

The first task is more related to image denoising and edge detection, for which we simply employ the Mumford–Shah model as the feature representation. In practice, the discontinuity sets are approximated by phase field functions as in Equation 3.2 of Ambrosio–Tortorelli approximation. In this algorithm, the four unknowns $\{w^R, w^T, v^R, v^T\}$ are estimated, where (w^R, v^R) and (w^T, v^T) are the feature representations of R and T, respectively.

The second task is more related to image registration. The non-rigid transformation from image R to image T is mostly different from the inverse function of the transformation from T to R. In order to overcome such correspondence ambiguities, we follow the method of consistent registration [CJ01] to jointly estimate the transformations in both forward and reverse directions. We denote the transformation from T to R as $\boldsymbol{\phi}$ and the transformation from R to T as $\boldsymbol{\psi}$. Functions $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ are estimated to match the two feature representations (w^R, v^R) and (w^T, v^T) to each other. Additionally, $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ are required to be smooth and approximately inverse to each other. For the desired spatial properties a regularization functional and a consistency functional are proposed to constrain the transformations to satisfy these requirements.

3.3.1 Functional Definitions

The six unknowns - the restored reference image w^R, the restored template image w^T, the edge describing phase-fields v^R and v^T of the reference and the template image, respectively, and the deformations $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ from the template to the reference domain and vice versa - are estimated by minimizing a joint functional with the following structure:

$$E_{\text{SYM}} = E_{\text{AC}} + \mu E_{\text{CC}} + \lambda E_{\text{REG}} + \kappa E_{\text{CON}}, \tag{3.5}$$

where μ, λ and κ are nonnegative constants which control the contributions of the associated functionals. The detailed definitions of these functionals are given in the following.

Auto-coupling Functional

$$\begin{aligned} E_{\text{AC}} =& C_{\text{AC}}[w^R, v^R] + C_{\text{AC}}[w^T, v^T] \\ :=& E^{\epsilon}_{\text{AT}}[w^R, v^R] + E^{\epsilon}_{\text{AT}}[w^T, v^T]. \end{aligned} \tag{3.6}$$

Here E^{ϵ}_{AT} denotes the functional of the Ambrosio–Tortorelli approximation whose definition has been given in Equation 3.2, where u is replaced by u^R or u^T respectively. The single auto-coupling cost function, e.g. $C_{\text{AC}}[w^R, v^R]$, essentially makes use of the mechanisms of the Mumford–Shah model and its Ambrosio–Tortorelli approximation to estimate the feature representation (w^R, v^R) of the image R, such that the piecewise smooth function w^R optimally couples with the phase field function v^R in a manner similar to Equation 3.3. Roughly speaking, this auto-coupling functional is responsible for detecting the edge features of each image and for defining the internal relation between the phase field function v^R (and v^T respectively) and the piecewise smooth function w^R (and w^T respectively). In this functional the segmented edge features of the two images, i.e. (w^R, v^R) and (w^T, v^T), are totally independent of each other.

Cross-coupling Functional

$$\begin{aligned} E_{\text{CC}} =& C_{\text{CC}}[w^R, v^R, \boldsymbol{\phi}] + C_{\text{CC}}[w^T, v^T, \boldsymbol{\psi}] \\ :=& \frac{1}{2} \int_{\Omega} (v^T \circ \boldsymbol{\phi})^2 \left\| \nabla w^R \right\|^2 \, \mathrm{d}\boldsymbol{x} \\ &+ \frac{1}{2} \int_{\Omega} (v^R \circ \boldsymbol{\psi})^2 \left\| \nabla w^T \right\|^2 \, \mathrm{d}\boldsymbol{x}. \end{aligned} \tag{3.7}$$

This functional is responsible for matching the edge features of the two images. It favors spatial transformations $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ which optimally couple the feature representations (w^R, v^R) and (w^T, v^T) in the following way:

$$v^T \circ \boldsymbol{\phi} \approx \begin{cases} 0 & \text{where } \|\nabla w^R\| \gg 0, \\ 1 & \text{where } \|\nabla w^R\| \approx 0. \end{cases}$$

$$v^R \circ \boldsymbol{\psi} \approx \begin{cases} 0 & \text{where } \|\nabla w^T\| \gg 0, \\ 1 & \text{where } \|\nabla w^T\| \approx 0. \end{cases}$$

By definition, this functional jointly treats segmentation and registration: Regarding registration, the functional acts the similarity measure based on the intermediately segmented edge features. Instead of directly matching the phase-fields functions ($v^R \leftrightarrow v^T$) and the smooth functions

($w^R \leftrightarrow w^T$), the functional seeks to match the gradient field of the smooth function of one image to the phase field function of the other image ($v^R \leftrightarrow \nabla w^T, v^T \leftrightarrow \nabla w^R$). Regarding segmentation, this functional also imposes the influence of the edge features segmented in the other image. In the following subsection we will see that both spatial transformations are controlled by regularization. The regularized spatial transformations lead to local edge feature correspondence.

Regularization Functional

$$\begin{aligned} E_{\text{REG}} =& C_{\text{REG}}[\boldsymbol{\phi}] + C_{\text{REG}}[\boldsymbol{\psi}] \\ :=& \frac{1}{2}\int_{\Omega} \|J(\boldsymbol{\phi} - \mathbb{1})\|^2 \, \mathrm{d}\boldsymbol{x} \\ &+ \frac{1}{2}\int_{\Omega} \|J(\boldsymbol{\psi} - \mathbb{1})\|^2 \, \mathrm{d}\boldsymbol{x}, \end{aligned} \tag{3.8}$$

where $J(\cdot)$ denotes Jacobian of vector-value function. Here $\mathbb{1} : \boldsymbol{x} \mapsto \boldsymbol{x}$ denotes the identity mapping and $\boldsymbol{\phi} - \mathbb{1}$, $\boldsymbol{\psi} - \mathbb{1}$ the displacement fields corresponding to $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$. Generally speaking, the regularization functional is used to rule out singular transformations which may lead to cracks, foldings, or other undesired properties. In this work the regularization constraint is the sum of the norm of the Jacobian of both displacement fields. See [AAF99] for further explanations of regularization based on the Jacobians of transformations.

Other candidates for regularization constraints are linear elastic [Bro81, CJM97] and viscous fluid [BNG96, CJM97] regularizations. These two constraints make use of a continuous mechanical model to regularize the transformations [Gur81]. Another alternative, which already ensures a homeomorphism property, is the nonlinear elastic regularization. It separately deals with length, area and volume deformation and in particular penalizes volume shrinkage [DR04].

Consistency Functional

$$\begin{aligned} E_{\text{CON}} =& C_{\text{CON}}[\boldsymbol{\phi}, \boldsymbol{\psi}] + C_{\text{CON}}[\boldsymbol{\psi}, \boldsymbol{\phi}] \\ :=& \frac{1}{2}\int_{\Omega} \|\boldsymbol{\phi} \circ \boldsymbol{\psi}(\boldsymbol{x}) - \boldsymbol{x}\|^2 \, \mathrm{d}\boldsymbol{x} \\ &+ \frac{1}{2}\int_{\Omega} \|\boldsymbol{\psi} \circ \boldsymbol{\phi}(\boldsymbol{x}) - \boldsymbol{x}\|^2 \, \mathrm{d}\boldsymbol{x}. \end{aligned} \tag{3.9}$$

The forward and reverse transformations $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ are purely independent of each other in E_{AC} and E_{REG} and are implicitly correlated in E_{CC} via the matching of the two image / phase-fields

pairs, i.e. $(w^R, v^T \circ \boldsymbol{\phi}) \leftrightarrow (w^T, v^R \circ \boldsymbol{\psi})$. The consistency functional E_{CON} in Equation 3.9 explicitly specifies the relationship between forward and reverse transformations. E_{CC} is minimal if and only if $\boldsymbol{\phi} \circ \boldsymbol{\psi}(\boldsymbol{x}) = \boldsymbol{x} = \boldsymbol{\psi} \circ \boldsymbol{\phi}(\boldsymbol{x})$, i.e., $\boldsymbol{\phi} \approx \boldsymbol{\psi}^{-1}$ and $\boldsymbol{\psi} \approx \boldsymbol{\phi}^{-1}$. The transformation in one direction has to be the inverse function of the transformation in the other direction. For registration, this consistency constraint favors an invertible and bijective correspondence of the segmented edge features.

3.3.2 Variational Formulation

We assume that the minimum of the entire energy E_{SYM} is the zero crossing of its variation with respect to all the unknowns $\{w^R, w^T, v^R, v^T, \boldsymbol{\phi}, \boldsymbol{\psi}\}$. The definition of the entire functional E_{SYM}, as well as of each individual functional E_{AC}, E_{CC}, E_{REG} and E_{CON}, is symmetric with respect to the two groups of unknowns: $\{w^R, v^R, \boldsymbol{\phi}\}$ and $\{w^T, v^T, \boldsymbol{\psi}\}$. Thus, we restrict ourselves to the description of the computation of variations with respect to $\{w^R, v^R, \boldsymbol{\phi}\}$. The variational formulas of the other group can be deduced in a complementary way.

Given an arbitrary scalar test function $\vartheta \in C_0^\infty(\Omega)$, we obtain the variations with respect to w^R and v^R:

$$\begin{aligned}
\langle \partial_{w^R} E_{\text{SYM}}, \vartheta \rangle &= \langle \partial_{w^R} E_{\text{AC}}, \vartheta \rangle + \mu \langle \partial_{w^R} E_{\text{CC}}, \vartheta \rangle \\
&= \int_\Omega \alpha(w^R - u^R)\vartheta + \beta(v^R)^2 \nabla w^R \cdot \nabla \vartheta \, \mathrm{d}\boldsymbol{x} \\
&\quad + \int_\Omega \mu(v^T \circ \boldsymbol{\phi})^2 \nabla w^R \cdot \nabla \vartheta \; \mathrm{d}\boldsymbol{x},
\end{aligned} \tag{3.10}$$

$$\begin{aligned}
\langle \partial_{v^R} E_{\text{SYM}}, \vartheta \rangle &= \langle \partial_{v^R} E_{\text{AC}}, \vartheta \rangle + \mu \langle \partial_{v^R} E_{\text{CC}}, \vartheta \rangle \\
&= \int_\Omega \beta \left\| \nabla w^R \right\|^2 v^R \vartheta \; + \frac{\nu}{4\epsilon}(v^R - 1)\vartheta \; \mathrm{d}\boldsymbol{x} \\
&\quad + \int_\Omega \nu\epsilon \nabla v^R \cdot \nabla \vartheta \, \mathrm{d}\boldsymbol{x} \\
&\quad + \int_\Omega \mu \| \nabla w^T \circ \boldsymbol{\psi}^{-1} \|^2 v^R \vartheta | \det J(\boldsymbol{\psi}^{-1}) | \, \mathrm{d}\boldsymbol{x}.
\end{aligned} \tag{3.11}$$

Here we have used the transformation rule

$$\begin{aligned}
&\int_\Omega \frac{\mu}{2} \| \nabla w^T \|^2 (v^R)^2 \circ \boldsymbol{\psi} \, \mathrm{d}\boldsymbol{x} \\
&\qquad = \int_{\boldsymbol{\psi}(\Omega)} \frac{\mu}{2} \| \nabla w^T \circ \boldsymbol{\psi}^{-1} \|^2 (v^R)^2 | \det J(\boldsymbol{\psi}) |^{-1} \, \mathrm{d}x
\end{aligned}$$

and $\boldsymbol{\psi}(\Omega) = \Omega$. Given an arbitrary vector-valued test function $\zeta \in C_0^\infty(\Omega, \mathbb{R}^d)$, we obtain the variation with respect to $\boldsymbol{\phi}$:

$$\begin{aligned}
\langle \partial_{\boldsymbol{\phi}} E_{\text{SYM}}, \zeta \rangle &= \mu \langle \partial_{\boldsymbol{\phi}} E_{\text{CC}}, \zeta \rangle + \lambda \langle \partial_{\boldsymbol{\phi}} E_{\text{REG}}, \zeta \rangle + \kappa \langle \partial_{\boldsymbol{\phi}} E_{\text{CON}}, \zeta \rangle \\
&= \int_\Omega \mu \left\| \nabla w^R \right\|^2 (v^T \circ \boldsymbol{\phi}) \nabla (v^T \circ \boldsymbol{\phi}) \cdot \zeta \, \mathrm{d}\boldsymbol{x} \\
&+ \int_\Omega \lambda J(\boldsymbol{\phi}) : J(\zeta) \, \mathrm{d}\boldsymbol{x} \\
&+ \int_\Omega \kappa([\boldsymbol{\phi} \circ \boldsymbol{\psi}](\boldsymbol{x}) - \boldsymbol{x}) \cdot [\zeta \circ \boldsymbol{\psi}](\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \\
&+ \int_\Omega \kappa([\boldsymbol{\psi} \circ \boldsymbol{\phi}](\boldsymbol{x}) - \boldsymbol{x}) J(\boldsymbol{\psi}(\boldsymbol{\phi}(\boldsymbol{x}))) \cdot \zeta(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}.
\end{aligned} \tag{3.12}$$

Due to the high complexity of the minimization problem (four scalar functions and two vector-valued functions) the unknowns are estimated in an Estimation-Minimization type procedure:

Algorithm 3 Estimation-Minimization type Procedure

Let $f_{i=1,\ldots,m}$ denote the unknown functions and $E := E[f_1, ..., f_m]$ denote the functional.
while $f_{i=1,\ldots,m}$ has not yet converged **do**
  for $i = 1$ to m **do**
    $f_i = \arg\min_f E[f_1, ..., f_{i-1}, f, f_{i+1}, ..., f_m]$.
  end for
end while

3.3.3 Solution of the Linear Part

First we introduce generalized mass and stiffness matrices, which play the key roles in the discretization of Equations 3.10 and 3.11 using FE approximation.

Given a function $f(\boldsymbol{x}) : \Omega \mapsto R$, the generalized mass $\boldsymbol{M}[f]$ and stiffness matrices $\boldsymbol{L}[f]$ are defined as follows:

$$\boldsymbol{M}[f] = \left(\int_\Omega f(\boldsymbol{x}) \varphi_i(\boldsymbol{x}) \varphi_j(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right)_{i,j} \tag{3.13}$$

$$\boldsymbol{L}[f] = \left(\int_\Omega f(\boldsymbol{x}) \nabla \varphi_i(\boldsymbol{x}) \cdot \nabla \varphi_j(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right)_{i,j} \tag{3.14}$$

Both matrices are $n \times n$-dimensional, where n is the number of nodes in the FE space. Both matrices are sparse, i.e. most entries are zero. An entry is non-zero, if and only if $i = j$

or node i and j are adjacent in the mesh. To compute the integrals in these non-zero entries we use a numerical Gaussian quadrature scheme of order three (cf. [SW92]). Obviously, the common mass matrix $\boldsymbol{M}$ and stiffness matrix $\boldsymbol{L}$ are just special cases of the generalized ones, i.e. $\boldsymbol{M} := \boldsymbol{M}[1]$ and $\boldsymbol{L} := \boldsymbol{L}[1]$.

The variations in Equation 3.10 and 3.11 are linear with respect to the unknowns w^R and v^R respectively. In each iteration of the Estimation-Minimization procedure, the zero-crossings are simply calculated by solving the corresponding linear systems. Replacing the continuous functions u^R and w^R with their FE approximations $\overline{U^R}(\boldsymbol{x}) = \sum_{i=1}^n \overrightarrow{U_i^R}\varphi_i(\boldsymbol{x})$ and $\overline{W^R}(\boldsymbol{x}) = \sum_{i=1}^n \overrightarrow{W_i^R}\varphi_i(\boldsymbol{x})$ and considering basis functions φ_i and φ_j of the FE space as test functions, the equation for zero crossings (see Equation 3.10) is equivalent to:

$$\begin{aligned}
&\alpha \sum_{i=1}^{n}\sum_{j=1}^{n} \overrightarrow{W_i^R} \int_\Omega \varphi_i(\boldsymbol{x})\varphi_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&+ \beta \sum_{i=1}^{n}\sum_{j=1}^{n} \overrightarrow{W_i^R} \int_\Omega (v^R(\boldsymbol{x}))^2 \nabla\varphi_i(\boldsymbol{x}) \cdot \nabla\varphi_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&+ \mu \sum_{i=1}^{n}\sum_{j=1}^{n} \overrightarrow{W_i^R} \int_\Omega (v^T \circ \boldsymbol{\phi})^2(\boldsymbol{x}) \nabla\varphi_i(\boldsymbol{x}) \cdot \nabla\varphi_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \alpha \sum_{i=1}^{n}\sum_{j=1}^{n} \overrightarrow{U_i^R} \int_\Omega \varphi_i(\boldsymbol{x})\varphi_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}
\end{aligned} \tag{3.15}$$

Using the notations of generalized mass (3.13) and stiffness matrices (3.14), Equation 3.15 can be rewritten as

$$\left(\alpha\boldsymbol{M} + \beta\boldsymbol{L}\left[(v^R)^2\right] + \mu\boldsymbol{L}\left[(v^T \circ \boldsymbol{\phi})^2\right]\right)\overrightarrow{W^R} = \alpha\boldsymbol{M}\overrightarrow{U^R}. \tag{3.16}$$

Similarly Equation 3.11 leads to

$$\begin{aligned}
&\Big(\mu\boldsymbol{M}\left[\|\nabla w^T \circ \boldsymbol{\psi}^{-1}\|^2 |\det J(\boldsymbol{\psi})|^{-1}\right] \\
&+\beta\boldsymbol{M}\left[\left\|\nabla w^R\right\|^2\right] + \frac{\nu}{4\epsilon}\boldsymbol{M} + \nu\epsilon\boldsymbol{L}\Big)\overrightarrow{V^R} = \frac{\nu}{4\epsilon}\boldsymbol{M}\overrightarrow{1}.
\end{aligned} \tag{3.17}$$

Here $\overrightarrow{1}$ denotes the one-vector, i.e. $(1, \cdots, 1)^T$. Analogously, we get the linear systems for $\overrightarrow{W^T}$ and $\overrightarrow{V^T}$:

$$\left(\alpha\boldsymbol{M} + \beta\boldsymbol{L}\left[(v^T)^2\right] + \mu\boldsymbol{L}\left[(v^R \circ \boldsymbol{\psi})^2\right]\right)\overrightarrow{W^T} = \alpha\boldsymbol{M}\overrightarrow{U^T}. \tag{3.18}$$

$$\Big(\mu \boldsymbol{M}\left[\|\nabla w^R \circ \boldsymbol{\phi}^{-1}\|^2 |\det J(\boldsymbol{\phi})|^{-1}\right] + \beta \boldsymbol{M}\left[\left\|\nabla w^T\right\|^2\right] + \frac{\nu}{4\epsilon}\boldsymbol{M} + \nu\epsilon\boldsymbol{L}\Big)\overrightarrow{V^T} = \frac{\nu}{4\epsilon}\boldsymbol{M}\overrightarrow{1}. \tag{3.19}$$

The linear equation systems 3.16 - 3.19 are solved with a preconditioned Conjugate-Gradient (CG) method. For the theory of CG method, we refer to [HS52].

3.3.4 Solution of the Nonlinear part

Equation 3.12 shows that the variation of energy is nonlinear with respect to one of the transformations. Thus, the unknown transformation cannot be estimated by solving a linear system. Instead we employ a regularized gradient descent method to iteratively find the zero-crossing:

$$\overline{\Phi}^{(k+1)} = \overline{\Phi}^{(k)} - \tau^{(k)} \cdot \text{grad}^{\varpi_\sigma}_{\overline{\Phi}} E[\overline{\Phi}^{(k)}], \tag{3.20}$$

where $\text{grad}^{\varpi_\sigma}_{\overline{\Phi}} E[\overline{\Phi}^{(k)}]$ is the regularized gradient with respect to the unknown $\overline{\Phi}$ and a metric ϖ_σ, and $\tau^{(k)}$ is the step size. In the following, we introduce the computation of $\text{grad}^{\varpi_\sigma}_{\overline{\Phi}} E[\overline{\Phi}^{(k)}]$ and the estimation of τ^k.

Regularized Gradient flow - $\text{grad}^{\varpi_\sigma}_{\overline{\Phi}} E[\overline{\Phi}^{(k)}]$

Optimization of the functional is usually updated in the direction of the gradient descent. However, due to the highly complex functional, the common gradient descent could be easily trapped by the large amount of local minima in practice. For this reason, the gradient flow is regularized in the sense that the target minimum can be reached by a smooth path from the initial guess, while irrelevant local minima are ruled out by a regularized metric.

Various regularized metric used for image registration have been summarized in [CHR02, Dro05]. In this work we choose the Helmholtz type operator $\mathcal{A} = \mathbb{1} - \frac{\sigma^2}{2}\triangle$ for $\sigma \in \mathbb{R}^+$. The metric representing $\mathcal{A}$ is

$$\varpi_\sigma(\overline{\Phi_1}, \overline{\Phi_2}) = (\overline{\Phi_1}, \overline{\Phi_2})_{L^2} + \frac{\sigma^2}{2}(J\overline{\Phi_1}, J\overline{\Phi_2})_{L^2}.$$

Here $(\cdot,\cdot)_{L^2}$ denotes the intrinsic scalar product of L^2. This regularized gradient combined with the time discretization is closely related to the iterative Tikhonov regularization, which leads to smooth paths from the initial deformations towards the set of minimizers of the matching

energy. For theoretical details we refer to [SW00, CHR02, CDR02]. In our implementation, the regularized gradient $\mathrm{grad}^{\varpi^\sigma}_{\overline{\Phi}} E[\overline{\Phi}^{(k)}]$ is computed in two steps:

1. Compute the variation
$$\partial_{\overline{\Phi}} E[\overline{\Phi}^{(k)}] = \left\langle \partial_{\overline{\Phi}} E_{\mathrm{SYM}}[\overline{\Phi}^{(k)}], \zeta \right\rangle$$
according to Equation 3.12, where the integrals are computed with a Gaussian quadrature scheme of order three and the test functions are the canonical basis functions of $\mathcal{V}^d$, see Section 3.1.3.

2. The representation of the metric in FE-terms is
$$\varpi_\sigma(\overline{\Phi_1}, \overline{\Phi_2}) = \left(\boldsymbol{M}_{\text{bl.}} + \tfrac{\sigma^2}{2} \boldsymbol{L}_{\text{bl.}} \right) \overrightarrow{\Phi_1} \cdot \overrightarrow{\Phi_2}$$
which leads to
$$\mathrm{grad}^{\varpi^\sigma}_{\overline{\Phi}} E[\overline{\Phi}^{(k)}] = \left(\boldsymbol{M}_{\text{bl.}} + \tfrac{\sigma^2}{2} \boldsymbol{L}_{\text{bl.}} \right)^{-1} \left(\partial_{\overrightarrow{\Phi_i}} E[\overline{\Phi}^{(k)}] \right)_i .$$
Here $\boldsymbol{M}_{\text{bl.}}$ and $\boldsymbol{L}_{\text{bl.}}$ denote $d \times d$ block matrices with the standard mass and stiffness matrices respectively on the diagonal positions, and zero matrices on the off diagonal positions. We use $\sigma = \sqrt{10h}$, where h is the mesh resolution. The solution of the linear system is computed by a single V-cycle of a multigrid solver.

At this point, we see that the principle difference to "classical" gradient descent methods is that the regularized method does not use the primitive variation but a regularized (smoothed) one as descent direction.

Armijo-rule - $\tau^{(k)}$

The step size of the gradient flow is determined by the Armijo-rule [Kos91], choosing the largest $\tau^{(k)}$ such that energy is minimized in a successive reduction rule.

The natural way in the Estimation-Minimization procedure is to estimate the step size for each transformation individually, i.e. estimate $\tau_{\overline{\Phi}}$ for the transformation $\overline{\Phi}$ then estimate $\tau_{\overline{\Psi}}$ for $\overline{\Psi}$. However, if $\tau_{\overline{\Phi}}$ and $\tau_{\overline{\Psi}}$ are estimated sequentially in each iteration, the consistency functional in Equation 3.7 prevents $\tau_{\overline{\Phi}}$ and $\tau_{\overline{\Psi}}$ from being large, because large individual step sizes will largely increase the consistency functional. Consequently, the regularized gradient descent requires a large number of iterations to approach the minimum. In order to solve this problem, we

simultaneously estimate both transformations and compute one step size for both of them:

$$\begin{bmatrix} \overline{\Phi}^{(k+1)} \\ \overline{\Psi}^{(k+1)} \end{bmatrix} = \begin{bmatrix} \overline{\Phi}^{(k)} \\ \overline{\Psi}^{(k)} \end{bmatrix} - \tau^{(k)} \begin{bmatrix} \text{grad}^{\varpi_\sigma}_{\overline{\Phi}} E[\overline{\Phi}^{(k)}, \overline{\Psi}^{(k)}] \\ \text{grad}^{\varpi_\sigma}_{\overline{\Psi}} E[\overline{\Phi}^{(k)}, \overline{\Psi}^{(k)}] \end{bmatrix}. \tag{3.21}$$

Since $\overline{\Phi}$ and $\overline{\Psi}$ are updated at the same time, the consistency energy does not penalize a large step size $\tau^{(k)}$ any more.

Let $\Theta := [\overline{\Phi}, \overline{\Psi}]^T$, $\text{grad}E[\Theta] := [\text{grad}^{\varpi_\sigma}_{\overline{\Phi}} E, \text{grad}^{\varpi_\sigma}_{\overline{\Psi}} E]^T$ and $E'[\Theta] := [E'[\overline{\Phi}], E'[\overline{\Psi}]]^T$. We define the condition for the Successive Reduction Rule (SRR) as:

$$\frac{E_{\text{REG}}[\Theta^{(k)} - \tau^{(k)} \cdot \text{grad}E[\Theta^{(k)}]] - E_{\text{REG}}[\Theta^{(k)}]}{\langle E'[\Theta^{(k)}], \text{grad}E[\Theta^{(k)}] \rangle} > \frac{1}{4}.$$

The step size $\tau^{(k)}$ is estimated as in Algorithm 4. The regularization of the gradient and the

Algorithm 4 Adaptive Step Size Estimation

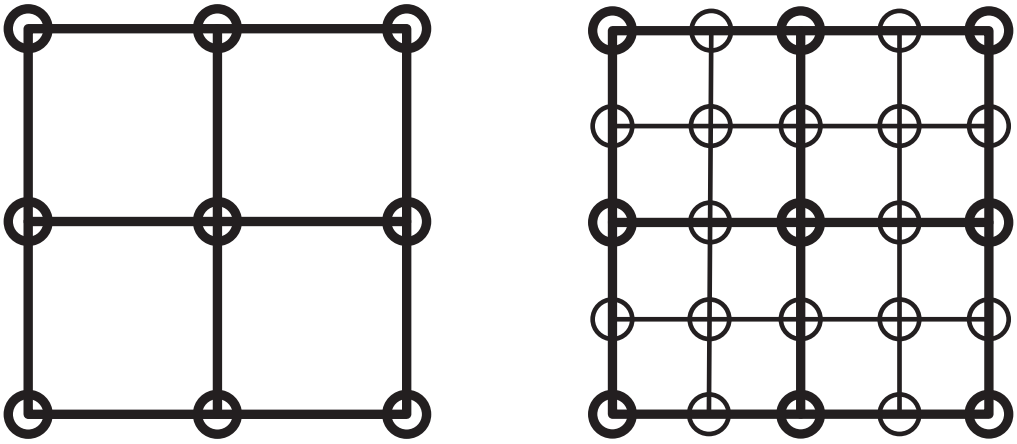
Initialize $\tau^{(k)}$ from previous iteration:
  if $k = 0$ **then** $\tau^{(k)} = 1.0$
  else $\tau^{(k)} = \tau^{(k-1)}$
Find the largest $\tau^{(k)}$ fulfilling SSR:
if SSR succeeds **then**
    do $\tau^{(k)} = 2\tau^{(k)}$ **until** SSR fails
else
    do $\tau^{(k)} = 0.5\tau^{(k)}$ **until** SSR succeeds
end if

adaptive estimation of the step size allow the regularized gradient descent method to perform more efficiently than the classical ones. In most cases we use five gradient descent steps to estimate the transformations in each iteration of the Estimation-Minimization procedure.

3.3.5 Multi-scale Algorithm

In order to avoid being trapped in local minima, the algorithm employs a spatial multi-scale scheme, in which global structures are segmented and matched before local ones.

The image domain $\Omega := [0, 1]^d$ is discretized by a rectangular mesh $\mathcal{C}_m$, which has $2^m + 1$ equidistant nodes in each axis, thus $n := (2^m + 1)^d$ nodes total. m is called the level of the mesh. A discrete function on the mesh $\mathcal{C}_m$ can also be called a function on level m. Figures 3.5(a) and 3.5(b) show a 2-D example of two nested meshes $\mathcal{C}_1$ and $\mathcal{C}_2$, in which the feature represen-



(a) $\mathcal{C}_1$ (b) $\mathcal{C}_2$

Figure 3.5: A simple 2-D example of nested mesh hierarchy. The nodes of the coarse mesh $\mathcal{C}_1$ are a subset of the nodes of the fine mesh $\mathcal{C}_2$. The prolongation of a function from the mesh $\mathcal{C}_1$ to the mesh $\mathcal{C}_2$ only requires the interpolation of the function values on the new nodes.

tations $\{w^R, v^R, w^T, v^T\}$ and the transformations $\{\boldsymbol{\phi}, \boldsymbol{\psi}\}$ are first computed on the coarse mesh $\mathcal{C}_1$. Then the results are prolongated to the next higher level on the finer mesh $\mathcal{C}_2$.

Although such a nested mesh hierarchy is not natural for finite difference methods, where commonly discrete images with 2^m voxels in each axis are used, it is common for the canonical hierarchy in the Finite Element context. This way the prolongation from one level to the next higher level is very convenient. Let $\mathcal{N}_m$ denote the set of nodes of the m-th mesh, as shown in Figure 3.5(a) and 3.5(b). The nested mesh hierarchy ensures $\mathcal{N}_{m-1} \subset \mathcal{N}_m$. During prolongation from level $m-1$ to m the function values stay the same on the nodes in $\mathcal{N}_{m-1}$ and the function values on the nodes in $\mathcal{N}_m \setminus \mathcal{N}_{m-1}$ are determined by multi-linear interpolation from the values on the neighboring nodes in $\mathcal{N}_{m-1}$. The entire multi-scale implementation is summarized in Algorithm 5.

3.4 Experiments

In this section, we will study the parameter setting of the method and illustrate the effects of different parameters by the registration of two MR images.

Two MR volumes are acquired from the same individual and with the same machine but with the different scan parameters (T1/T2). The original T1 weighted MR (reference image R) and T2 weighted MR (template image T) volumes are already nearly perfectly matched to each other. In order to demonstrate the effect of registration, the T2 weighted MR is artificially

Algorithm 5 One-to-one Edge based Registraiton Algorithm

given images u^R and u^T.
given starting level m_0 and ending level m_1.
given number of iterations on each level N_m.
initialize $[w^R, w^T, v^R, v^T]$ with 0.
intialize $[\boldsymbol{\phi}, \boldsymbol{\psi}]$ with $\mathbb{1}$.
for $m = m_0$ **to** m_1 **do**
  for $k = 1$ **to** N_m **do**
    update $\overline{W^R}$ by Equation 3.16
    update $\overline{V^R}$ by Equation 3.17
    update $\overline{W^T}$ by Equation 3.18
    update $\overline{V^T}$ by Equation 3.19
    update $[\overline{\Phi}, \overline{\Psi}]$ with 5 regularized gradient descent steps by Equation 3.21
  end for
  if $m \neq m_1$ **then**
    prolongation $[\overline{W^R}, \overline{V^R}, \overline{W^T}, \overline{V^T}, \overline{\Phi}, \overline{\Psi}]$ from level m to $m+1$
  end if
end for

deformed by a given elastic transformation. We specify the displacement vectors on eight points and computed the displacement vectors in the remaining part of the data using thin-plate spline interpolation. Both of the given volumes are of size $512 \times 512 \times 101$ and have been resampled to $129 \times 129 \times 129$ pixels to comply with the mesh hierarchy presented before. We performed 18 experiments with different parameter settings. For each experiment 10 iterations were run on the $129 \times 129 \times 129$ mesh. It took approximately two hours for the C++ implemented program to run an experiment on a standard PC with Intel Pentium 4 processor with 2.26 GHz and 2.0 GB RAM. It is expected that the computation time will decrease significantly by further optimization of the code. Although these parameters are only tested for registration of T1- and T2-weighted MR, they can also be used to determine the parameters for edge matching of the other modalities.

Experiments A1-A4 demonstrate how the parameters β, μ and λ balance edge detection and edge matching in the algorithm. The other parameters are fixed at $\alpha = 2550$, $\nu = 0.1$, $\kappa = 100$, $\epsilon = 0.01$. In this example, we denote the phase field functions of T1- and T2-weighted MR volumes as v^1 and v^2 respectively. Figure 3.6 shows how the two phase field functions varied in a local region with different parameters. In experiments A1-A3, the very large regularization weighting parameter λ ($= 1000$) prevents the algorithm from matching the edge features of the two images. Without consideration of the edge matching, the detection of edge features is controlled by the ratio between the auto-coupling weighting parameter β and the cross-coupling weighting parameter μ. In experiment A1, since β is much larger than μ, the auto-coupling

functional E_{AC} has more influence than the cross-coupling functional E_{CC}. The resulting phase field functions are more likely to describe its own edge feature. Experiment A2 is exactly the opposite case of A1. With small β and large μ the phase field function is more likely to represent the edge features of its counterpart. Namely, v^1 shows the edge features of image T2 and v^2 shows the edge features of image T1. The parameters β and μ need to be customized to specific applications. A general principle: β and μ need to be set in such a way that the resulting phase field functions v^1 and v^2 clearly describe the edge features of both images, as shown in experiment A3. For the T1-/T2-MR data in this experiment, it is reasonable to set β and μ equal. However, when the intensity patterns of images are largely different, like in neurosurgery photographs and the brain MR projection in Section 4.3, it is necessary to choose the parameters β and μ differently. In experiment A4, we activate the edge matching through a relatively small regularization weighting parameter λ ($= 10$). Each phase field function describes not only its own edge features, but also the transformed edge features of the other image. From Figure 3.6, one can observe that the phase field functions are merged with respect to experiment A3.

Experiments B1-B7 and C1-C7 were used to study the setting of the parameters λ and κ. We measured the cross-coupling cost C_{CC}, regularization cost C_{REG} and consistency cost C_{CON} for each experiment. The values of these costs are shown in Table 3.1 and 3.2 and have been scaled by 10000 for presentation purposes. The minimum and the inverse of the maximum of the determinant of the Jacobians of the forward and reverse transformations are computed to measure the degree of preservation of the topology. If a transformation is regular, these determinants should be close to 1.

Experiments B1-B7 demonstrate the effect of the regularization functional as the weight parameter λ is varied. In experiments B1 and B2, there are minor regularization constraints. A negative Jacobian of the transformations appeared. This means that the estimated transformation failed to preserve the topology of the images. As λ increases, the regularization constraints improve the transformations because the minimum Jacobian and the inverse of the maximum Jacobian are far from being singular. Experiments C1-C7 demonstrate the effect of the consistency functional as the weight parameter κ varies. In experiment C1, the consistency functional E_{CON} has no influence on the registration. The forward and reverse transformations are relatively independently estimated. The inconsistency of the two transformations are confirmed by the relatively large cost of the consistency functional. As κ increases, the cost of the consistency functional approaches zero. This means that one transformation is more likely to be the inverse function of the other one. Note that the cost of cross-coupling functional increases when the consistency constraints and regularization constraints become strong,

Table 3.1: Study of the weight of the regularization functional E_{REG}

Exp.	λ	C_{CC}		C_{REG}		C_{CON}		$\det J\boldsymbol{\phi}$		$\det J\boldsymbol{\psi}$	
		For	Rev	For	Rev	For	Rev	$1/\max$	$\min$	$1/\max$	$\min$
$B1$	0.01	1982	2939	780.9	128.5	3.700	3.667	0.4736	-0.057	0.6933	0.4197
$B2$	0.1	2221	2944	517.3	94.27	2.965	2.940	0.5671	0.088	0.7617	0.5195
$B3$	1	2709	2971	181.2	59.70	2.032	2.029	0.7348	0.4737	0.7358	0.4899
$B4$	5	3120	3050	44.24	27.27	1.149	1.146	0.8738	0.7296	0.8518	0.7464
$B5$	10	3328	3165	20.02	11.00	0.9419	0.9415	0.9253	0.8209	0.9070	0.8706
$B6$	20	3517	3243	6.180	3.031	0.7674	0.7699	0.9403	0.9000	0.9479	0.9301
$B7$	50	3550	3314	1.053	0.5344	0.1792	0.1833	0.9832	0.9599	0.9802	0.9724

C_{CC}: Cross-coupling functional, C_{REG}: Regularization functional, C_{CON}: Consistency functional. The other weight parameters were set as follows: $\alpha = 2550, \beta = 1, \nu = 0.1, \mu = 0.5, \kappa = 100$.

which indicates a worse matching of edge features between the two images. The optimal parameters should be chosen so as to achieve optimal feature matching, least amount of topological distortion and acceptable inconsistency of the transformations. According to our experience, it is safe to roughly fix five of the parameters in most 2-D and 3-D applications, i.e. $\lambda = 10, \kappa = 100, \alpha = 2550, \nu = 0.1 \sim 1, \epsilon = 0.01$ usually achieves good results.

3.5 Summary

This new edge matching method simultaneously performs the following three tasks: detecting the edge features from two images, computing two dense warping functions in both forward and reverse directions to match the detected features, and constraining each dense warping function to be the inverse of the other. An adaptive regularized gradient descent, in the framework of multi-resolution Finite Element approximation, enables the algorithm to efficiently find the pair of dense transformations.

Although the idea of simultaneous edge detection and edge matching has been successfully applied on various applications of medical image analysis (see Chapter 4), it is still very difficult to register functional and morphological imaging modalities, which is one of the most challenging multi-modal registration problems. Actually, it is very difficult to define the so-called edge feature among the functional images (e.g SPECT, PET). These imaging modalities record the

Table 3.2: Study of the weight of consistency functional E_{CON}

Exp.	κ	C_{CC}		C_{REG}		C_{CON}		$\det J\boldsymbol{\phi}$		$\det J\boldsymbol{\psi}$	
		For	Rev	For	Rev	For	Rev	$1/\max$	min	$1/\max$	min
$C1$	0	3044	3121	28.85	41.87	3.054	3.047	0.8824	0.7507	0.8475	0.7950
$C2$	50	3072	3137	27.19	45.13	0.7922	0.7891	0.8782	0.7251	0.8548	0.8136
$C3$	100	3088	3157	27.24	42.26	0.3255	0.3249	0.8751	0.7495	0.8569	0.8230
$C4$	200	3236	3115	32.69	25.19	0.1720	0.1720	0.8996	0.8032	0.8624	0.8246
$C5$	300	3279	3154	27.72	17.06	0.1430	0.1426	0.9061	0.8046	0.8971	0.8824
$C6$	400	3291	3169	26.82	17.50	0.1118	0.1165	0.9079	0.8086	0.8977	0.8758
$C7$	500	3334	3182	24.74	32.82	0.0803	0.0803	0.9115	0.8170	0.9917	0.8794

C_{CC}: Cross-coupling functional, C_{REG}: Regularization functional, C_{CON}: Consistency functional. The other weight parameters were set as follows: $\alpha = 2550, \beta = 1, \nu = 0.1, \mu = 0.5, \lambda = 10$.

accumulation of signals during acquisition time (typically more than 30 minutes) and the result is a kind of "average" image of functional activities of organs. Therefore, the borders between organs are blurry and edge features are hard to be located. The other drawback of this method is the difficulty of finding appropriate parameter settings for different applications, where seven constant parameters ($\alpha, \beta, \gamma, \epsilon, \kappa, \mu, \nu$) should be determined in an ad-hoc way. Although in the last section the study gives a good guideline to find an approximative parameter setting, it is still a tedious task in practice to find the optimal one for various registration applications.

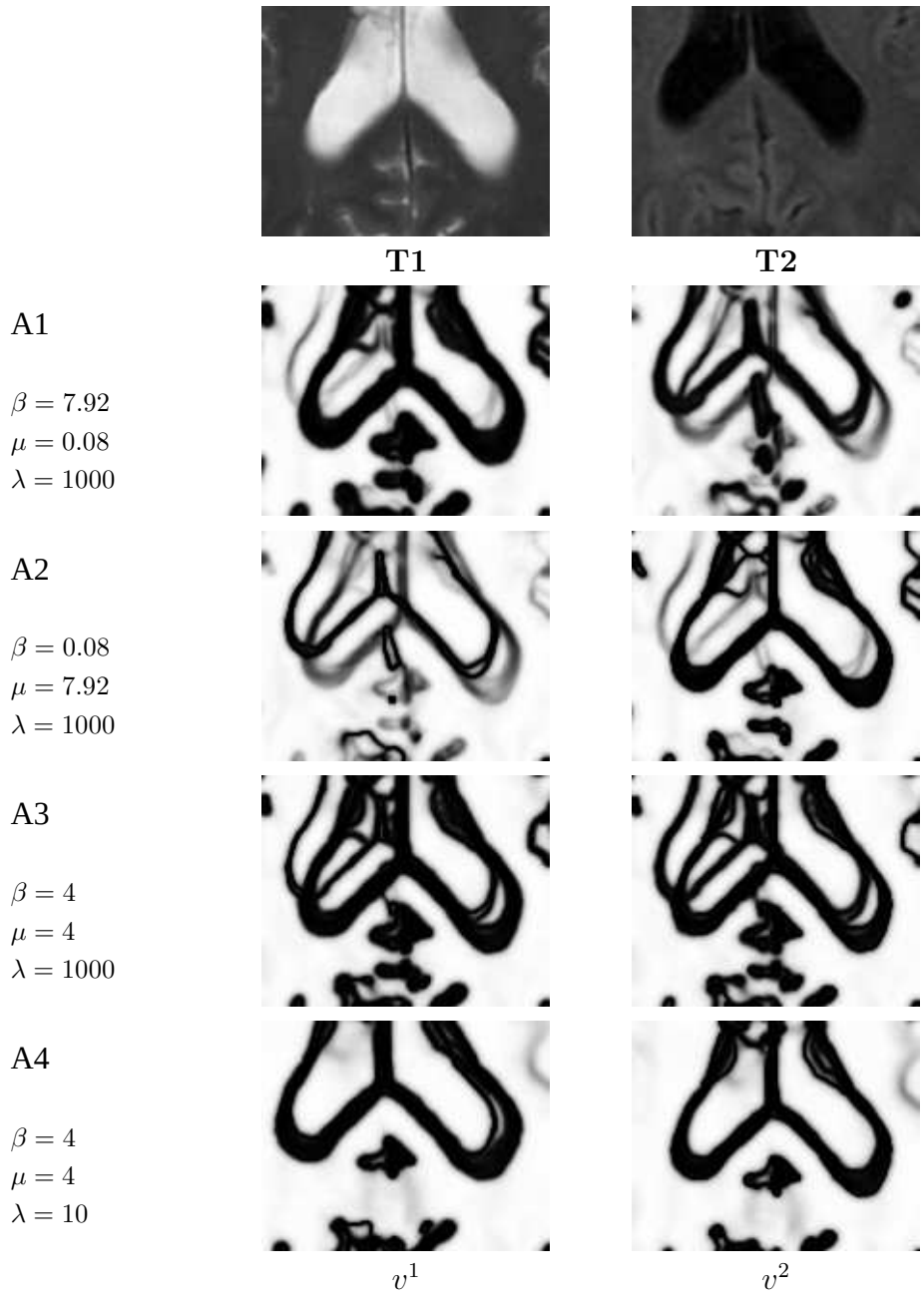


Figure 3.6: Experiments A1-A4 show the influence of the parameters β, μ and λ on the phase field functions. In experiments A1-A3, the very large λ disables the edge matching functionality and allows only edge detections. Furthermore, the ratio between β and μ determines whether the phase-fields represent edge features of its own image or the features of its counterpart. In experiment A4 edge matching as well as edge detection are enabled. Note that edge matching merged the phase-fields of both sides compared to experiment A3.

Chapter 4

Applications of One-to-one Edge based Registration

4.1 Registration of Inter-subject MR Datasets

In the following two experiments we use the one-to-one edge matching method to solve the inter-object mono-modal registration problem: registering two MR data sets (MR-to-MR) and two CT data sets (CT-to-CT). The two MR data show healthy brains of two individuals. The two CT data show two other patients, one normal and one abnormal. The data sets are collected by the same MR and CT scanners with the same scanning parameters. The MR data sets are preprocessed by segmenting the brain from the head using MRIcro[1].

The original sizes of the two CT data sets were $512 \times 512 \times 58$ and $512 \times 512 \times 61$ while the two MR data sets were $256 \times 256 \times 160$ and $256 \times 256 \times 170$. All of them have been resampled into a $257 \times 257 \times 257$ voxel lattice with the same resolution in all three directions. The experiments were performed with the previously described multi-scale scheme, with 10 iterations for each of the levels: $33 \times 33 \times 33$, $65 \times 65 \times 65$, $129 \times 129 \times 129$ and $257 \times 257 \times 257$. It took approximately 1 minute, 10 minutes, 90 minutes and 5 hours respectively for each level. The parameters of the algorithm were set as follows: $\alpha = 2550$, $\beta = 1$, $\nu = 0.1$, $\mu = 1$, $\lambda = 10$, $\kappa = 100$, $\epsilon = 0.01$.

The matching results of the data sets are visualized by a pattern of "interlace-stripe", showing the two data sets in turns within a single volume. For the MR-to-MR registration, sub-figures 4.1(a) and 4.1(c) show the interlace-stripe volumes of the original data sets R and T, while sub-figures 4.1(b) and 4.1(d) show the interlace-striped volumes of registered data sets in forward and reverse directions. For the CT-to-CT registration in Figures 4.2(a)-4.2(d), the layout

[1] http://www.sph.sc.edu/comd/rorden/mricro.html

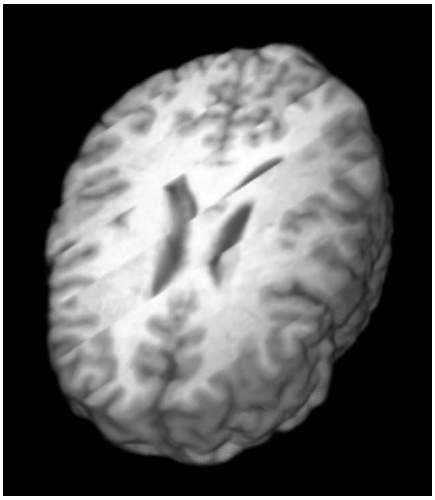
of sub-figures is same.

By visual inspection, the algorithm of one-to-one edge matching successfully registers MR-to-MR and CT-to-CT volume data sets of different individuals in both directions. Figures 4.1(a)-4.1(d) show precise alignments of the edges such as the brain's volume shape, hemispheric gap and ventricular system for inter-object MR-to-MR registration. In the inter-object CT-to-CT registration the main interest is to obtain the fitting shape of the bone. In Figures 4.2(a)-4.2(d) axial cuts of the 3-D CT data set are shown. Figures 4.3(a) and 4.3(c) shows that the initial mismatch of the data sets, visible by the discontinued bone edges in the top row, is dissolved with the computed transformation, as is evident from the continuous bone edges in Figures 4.3(b) and 4.3(d).

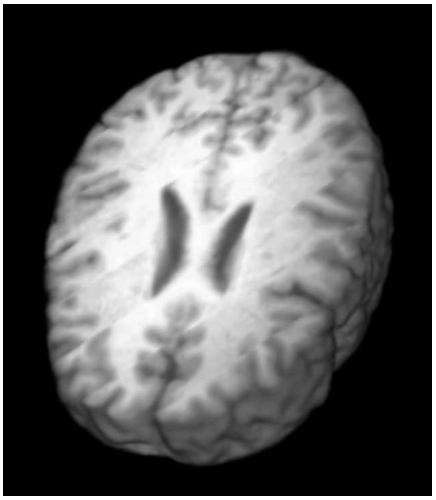
4.2 Registration of Multi-modal Retinal Images

A concurrent representation of the optic nerve head and the neuroretinal rim in various retina image modalities is significant for a definite diagnosis of glaucoma. Several modalities of retina images have been used in the ophthalmic clinic: the reflection-free photographs with an electronic flash illumination and the depth/reflectance retina images acquired by scanning-laser-tomograph. By acquisition, the depth and reflectance images normally have been perfectly matched to each other. Thus, the task of this application is the registration of multi-modal retina images, i.e. to match the reflectance and depth images with the photograph. For the registration of mono-modal retina images we refer to [CSRT02a, CSRT02b]. Figures 4.4(a)-4.4(c) show an example of multi-modal retina images of the same patient. In a recent paper [KJ04], an affine transformation model and an extended mutual information similarity are applied for registration of bi-modal retina images. However, as shown in Figures 4.5(a) and 4.5(d), this method (using the software described in [KJ04]) still cannot recover the minor deviations in the domain of vessels and neuroretinal rims. In this experiment we employ our one-to-one edge matching algorithm as a post-registration to compensate such small deviations of fine vessels.

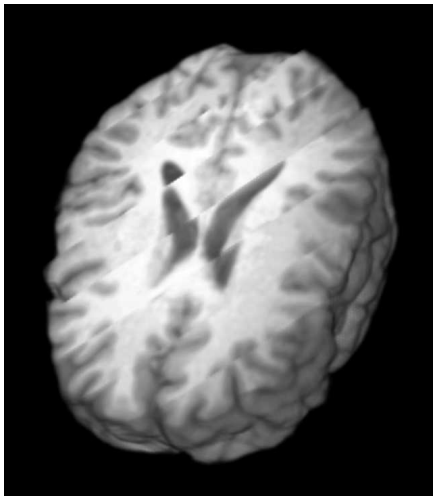
The images are pre-processed in the following way: first extracting the green channel of the photograph as the input for the registration. Then the photograph is affinely pre-registering to reflectance and depth images using the automatic software described in [KJ04]. In the last step, the pre-registered images are sampled in a mesh of 257×257. The algorithm is run for 10 iterations in three levels, which takes less than three minutes altogether. The parameters of the algorithm are set as follows: $\alpha = 2550$, $\beta = 1$, $\nu = 0.1$, $\mu = 0.5$, $\lambda = 10$, $\kappa = 100$, $\epsilon = 0.01$. From Figures 4.5(a)-4.5(f), one can observe that most minor deviations in the domain of vessels



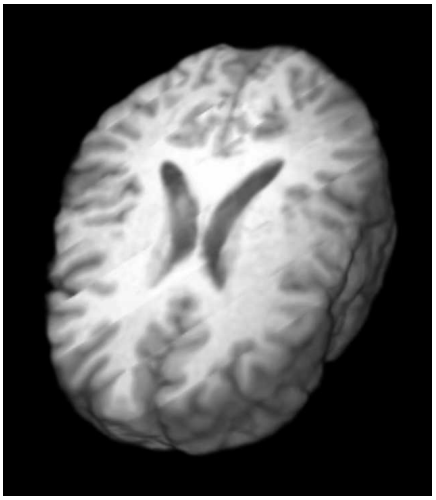
(a) $u^R \| u^T$: Interlace-stripe volume of between original data-sets R and T.



(b) $u^R \| u^T \circ \boldsymbol{\phi}$: Interlace-stripe volume of transformed T and original R.

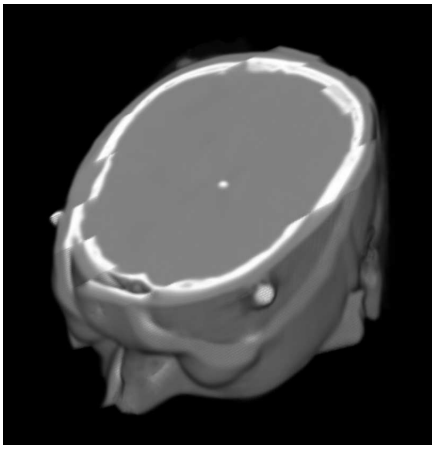


(c) $u^T \| u^R$: Interlace-stripe volume of between original data-sets T and R.

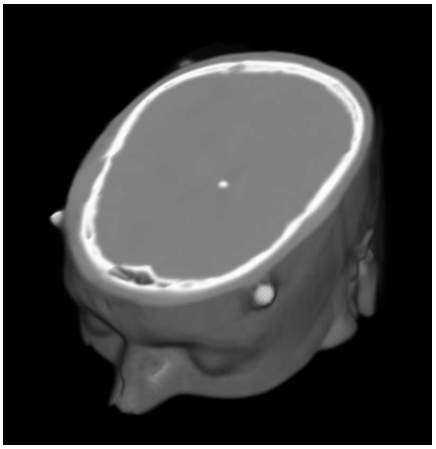


(d) $u^T \| u^R \circ \boldsymbol{\psi}$: Interlace-stripe volume of transformed R and original T.

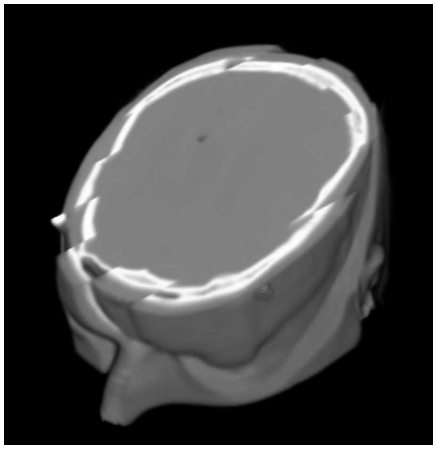
Figure 4.1: Inter-object MR-to-MR registration using one-to-one edge matching. The sub-figures (a) and (c) show the interlace-stripe volumes of the original data sets R and T, while the sub-figures (b) and (d) show the interlace-striped volumes of registered data sets in forward and reverse directions.



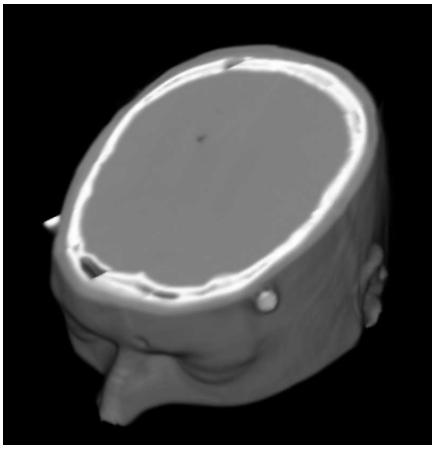
(a) $u^R \| u^T$: Interlace-stripe volume of between original data-sets R and T.



(b) $u^R \| u^T \circ \boldsymbol{\phi}$: Interlace-stripe volume of transformed T and original R.

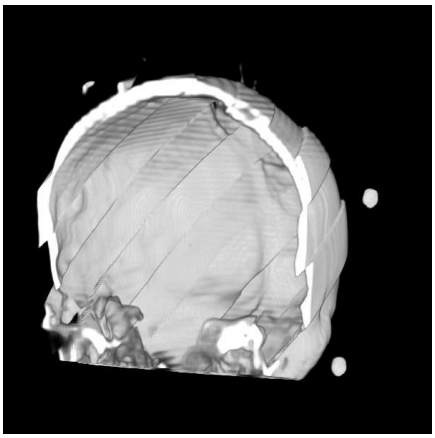


(c) $u^T \| u^R$: Interlace-stripe volume of between original data-sets T and R.

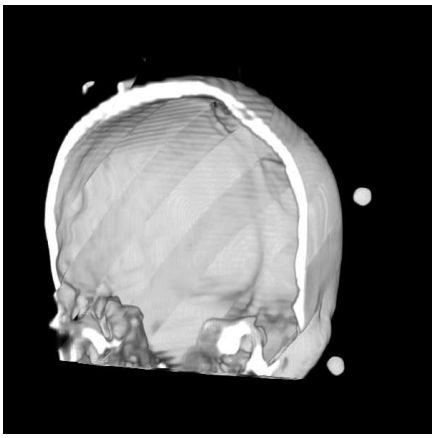


(d) $u^T \| u^R \circ \boldsymbol{\psi}$: Interlace-stripe volume of transformed R and original T.

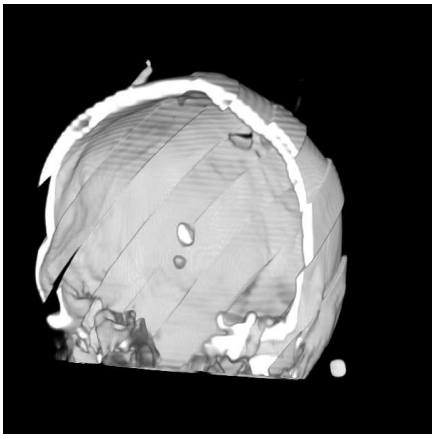
Figure 4.2: Inter-object CT-to-CT registration using one-to-one edge matching. The sub-figures (a) and (c) show the interlace-stripe volumes of the original data sets R and T, while the sub-figures (b) and (d) show the interlace-striped volumes of registered data sets in forward and reverse directions.



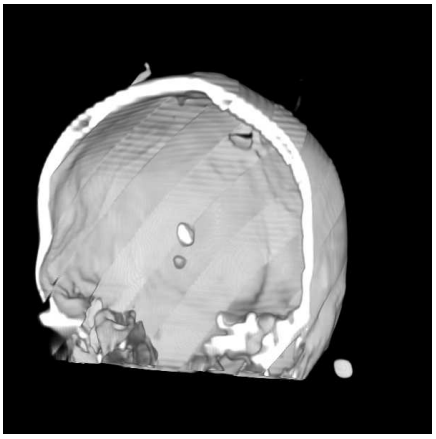
(a) $u^R \| u^T$: Interlace-stripe skulls of between original data-sets R and T.



(b) $u^R \| u^T \circ \boldsymbol{\phi}$: Interlace-stripe skulls of transformed T and original R.

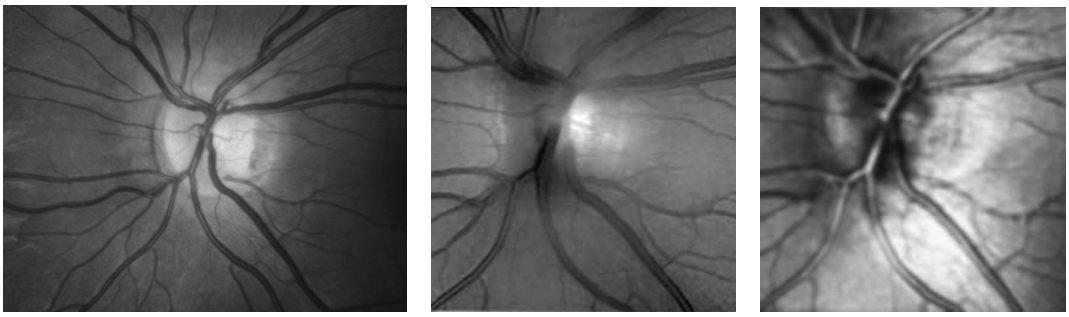


(c) $u^T \| u^R$: Interlace-stripe skulls of between original data-sets T and R.



(d) $u^T \| u^R \circ \boldsymbol{\psi}$: Interlace-stripe skulls of transformed R and original T.

Figure 4.3: The matching of skulls in CT-to-CT registration. (a) and (c): Interlace-stripe volumes of skulls of original data sets. (b) and (d): Interlace-stripe volumes of matched skulls.



(a) Photograph (b) Depth image (c) Reflectance image

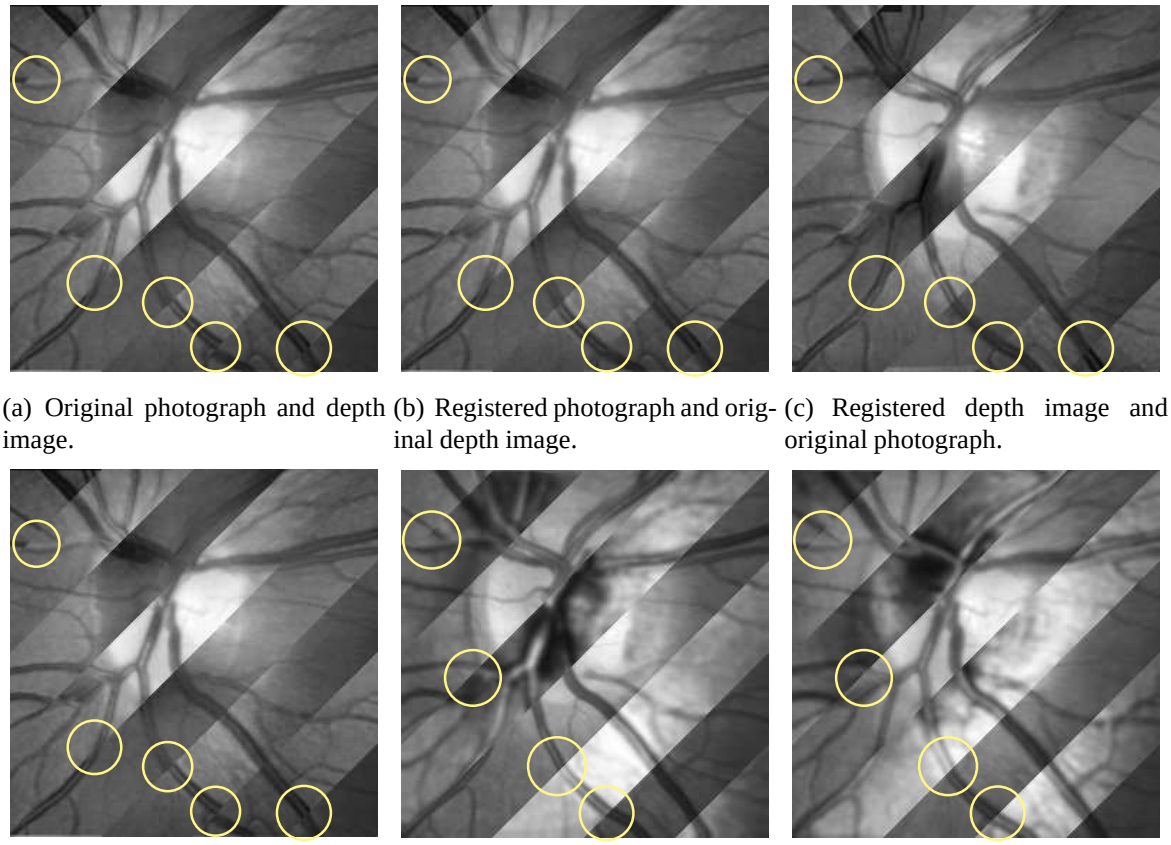
Figure 4.4: Multi-modal retina images of the same patient.

are compensated by the computed non-rigid transformations. Note that in this example with fine elongated structures, different from more volumetric image structures in the other applications, an affine pre-registration is used to compensate the large initial mismatch and to avoid getting stuck in a local minimum.

4.3 Matching Photographs of Neurosurgery to MRI Volume

In neocortical epilepsy surgery, the tumor may be located adjacent to, or partly within, the so-called eloquent (functionally very relevant) cortical brain regions. For a safe neurosurgical planning, the physician needs to map the appearance of the exposed brain to the underlying functionality. Usually, an electrode is placed on the surface of the brain in the first operation for electrophysiological examination of the underlying brain functionalities, then the photograph within the tested anatomical boundaries is colored according to the function of electrode contacts. On the other hand, the pre-operative 3-D MR data set contains the information of the underlying tumor and healthy tissue as well. In the second procedure, the registered photograph and MRI volume are used together to perform the cutting without touching eloquent areas. Currently, a neocortical expert needs to manually rotate the 3-D MR to find the best 2-D projection matching to the photographs. However, due to the different acquisitions and the brain shift during surgery, the photograph and MR projection cannot be accurately aligned. In this experiment, we make use of our one-to-one edge matching algorithm to refine the matching between a 2-D digital photograph of epilepsy surgery to the projection of 3-D MR data of the same patient.

The digital photographs of the exposed cortex are taken with a handheld Agfa e1280 digital camera (Agfa, Cologne, Germany) from the common perspective of the neurosurgeon's view. The high-resolution 3-D data set is acquired according to the T1-weighted MR imaging proto-



(a) Original photograph and depth image.

(b) Registered photograph and original depth image.

(c) Registered depth image and original photograph.

(d) Original photograph and reflectance image.

(e) Registered photograph and original reflectance image.

(f) Registered reflectance image and original photograph.

Figure 4.5: The example of post-registration of bi-modal retina images using one-to-one edge matching. The photograph is registered with the depth image (b)-(c) and the reflectance image (e)-(f). A registration method published in [KJ04] for bi-modal retina images cannot fully recover the minor deviations of fine structures in (a) and (d). The forward and reverse transformations estimated by the one-to-one edge matching successfully remove such minor mismatching.

col (TR 20, TE 3.6, flip angle 30°, 150 slices, slice thickness 1mm) using 1.5 Tesla Gyroscan ACS-NT scanner (Philips Medical Systems). The brain is automatically extracted from the MRI volume using the SISCOM module of the Analyze software (Mayo Foundation, Rochester, MN). For both the photograph and the MR projection, the regions of interest are manually selected by a physician.

Figures 4.6(a)-4.6(d) show the input images, preprocessed images, interlace-stripe registered and unregistered images. In Figure 4.6(a), the digital photograph shows the exposed left hemisphere from an intraoperative viewpoint, the frontal lobe on the upper left, the parietal lobe on the upper right and parts of the temporal lobe on the bottom. The surface with the gyri and sulci and

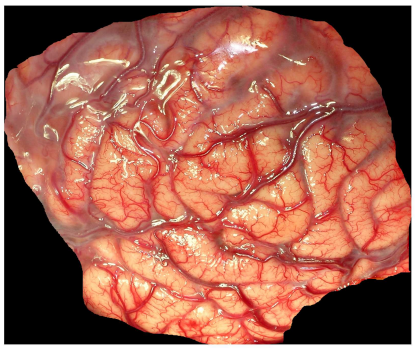
the overlying vessels are clearly visible. Alongside, Figure 4.6(b) displays the left-sided view of the rendered MR volume in the corresponding parts. Comparing Figures 4.6(a) and 4.6(b), one can notice that the undesired surface vessels and reflectance flash are strongly presented in the digital photograph, while the MR projection images clearly display the desired edge features. The photographic image and the projection image were preprocessed by appropriate GIMP filter chains for edge enhancement. The preprocessed images are displayed in Figures 4.6(c) and 4.6(d), respectively. Both images were resampled to 2049×2049 pixels. The algorithm was run from level 3 to level 11. We note that the values of the parameters β and μ are quite different from the other examples. The reason is that the image modalities of the photograph and the MR projection differ largely from each other. The two parameters are set to $\beta = 100$ and $\mu = 0.1$, so that both phase field functions clearly represent the edge features on the brain and have comparable influence on the registration. In Figures 4.7(a) and 4.7(c), the interlace-stripe images illustrate the mismatch of photograph and MR projection. Figures 4.7(b) and 4.7(d) show that the method greatly refines the matching of the desired edge features. Especially the brain sulci and gyri, which are significant for neurosurgery, are nearly perfectly aligned to each other. We have implemented a mutual information algorithm in the same Finite Element framework (including the step sized controlled, regularized, multi-scale descent) for a comparison. Overall, our method gives comparable results in most cases, especially when dealing with coarse structures. However, in this example that contains a large number of fine structures, the edge based matching gives better alignment. The zoom views of local regions in Figures 4.8(a)-4.8(d) show that the edge-matching method can achieve a better alignment of fine structures than the mutual information based registration.

4.4 Motion Compensation for Frame Interpolation

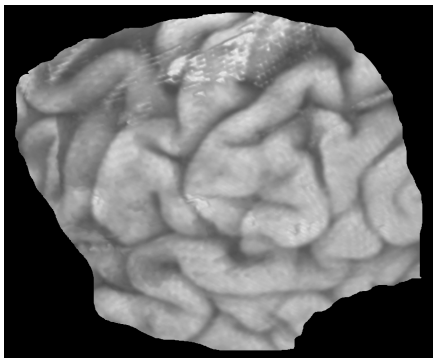
Temporal interpolation of video frames in order to increase the frame rate requires the estimation of a motion field (transformation). Then pixels in the intermediate frame are interpolated along the path of the motion vector. In this section, we give a proof of concept that the one-to-one edge matching method can be used for this application. For a review of the frame interpolation method, we refer to [KWM99, KBS04].

We perform our test on the Susie sequence[2] and interpolate frame 58 in Figures 4.9(a)-4.9(f). We use a 257×257 cropped version for the experiment. Frames 57, 58 and 59 are denoted as F_{57}, F_{58} and F_{59} respectively. The forward transformation $\phi : F_{57} \rightarrow F_{59}$ and reverse transformation

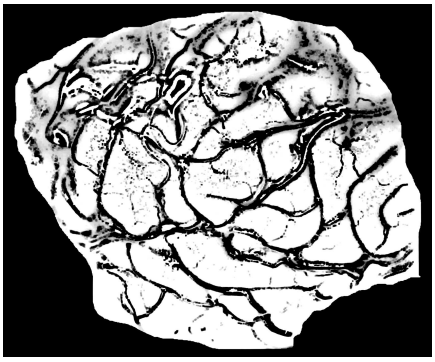
[2]Susie sequence from http://image.cse.nsysu.edu.tw/testimage/sequence/gray/susie.rar



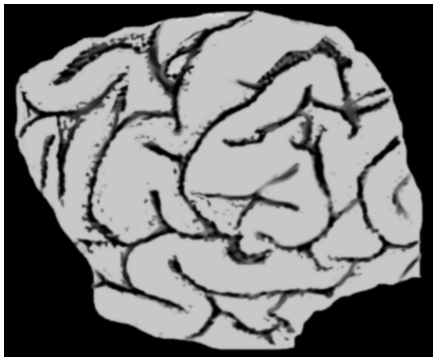
(a) The original photograph of the exposed left hemisphere from an intraoperative view point.



(b) A section of the projection of the MR volume, whose orientation is specified by physicians.



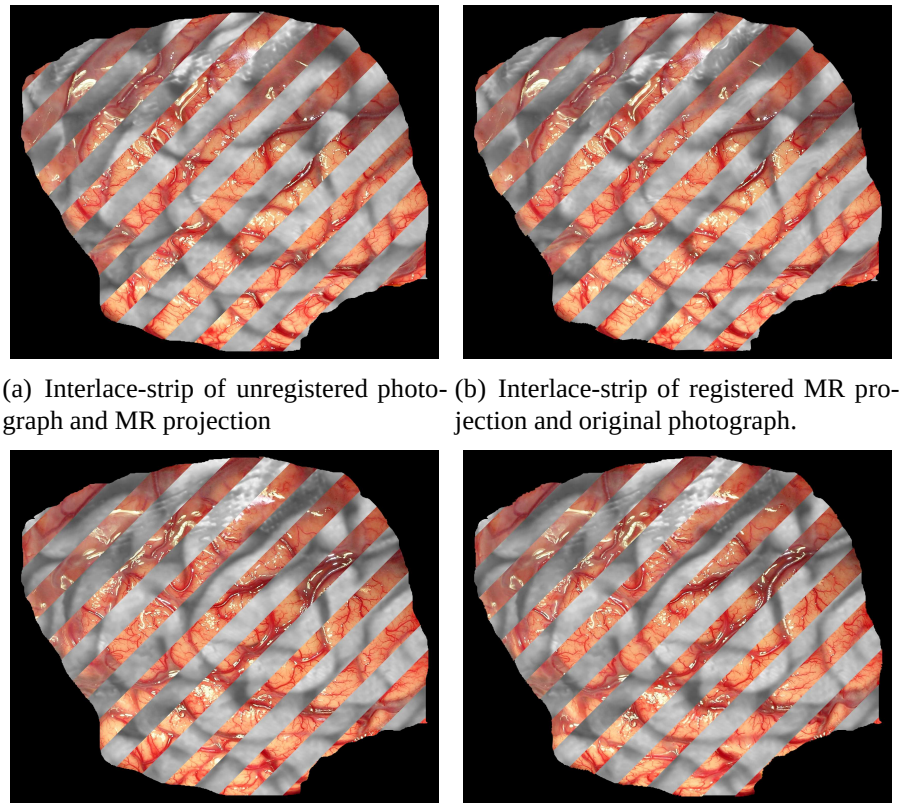
(c) Preprocessed photograph.



(d) Preprocessed MR projection.

Figure 4.6: A neurosurgery photograph of a section of the brain and its MR projection. All the sub-figures only display the region of interest: the exposed cortex. Image (a) and (b) are courtesy of Dr. J. Scorzin (Department of Neurosurgery, Bonn University Hospital).

$\boldsymbol{\psi} : F_{59} \rightarrow F_{57}$ are estimated by the one-to-one edge matching with the parameter setting: $\alpha = 2550$, $\beta = 1$, $\nu = 0.1$, $\mu = 1$, $\lambda = 10$, $\kappa = 100$, $\epsilon = 0.01$. Frame 58 is interpolated as: $F_{58} = 0.5 \times (F_{57} \circ 0.5\boldsymbol{\phi} + F_{59} \circ 0.5\boldsymbol{\psi})$. It is compared with a standard block matching algorithm using an adaptive rood pattern search [NM02], 16×16 blocks and a search range of $[-16, 16]$ in the horizontal and vertical directions. The experimental results show that the block matching algorithm produces blocky and noisy motion fields, while the one-to-one edge matching based motion estimation gives an excellent visual quality of frame interpolation.



(a) Interlace-strip of unregistered photograph and MR projection

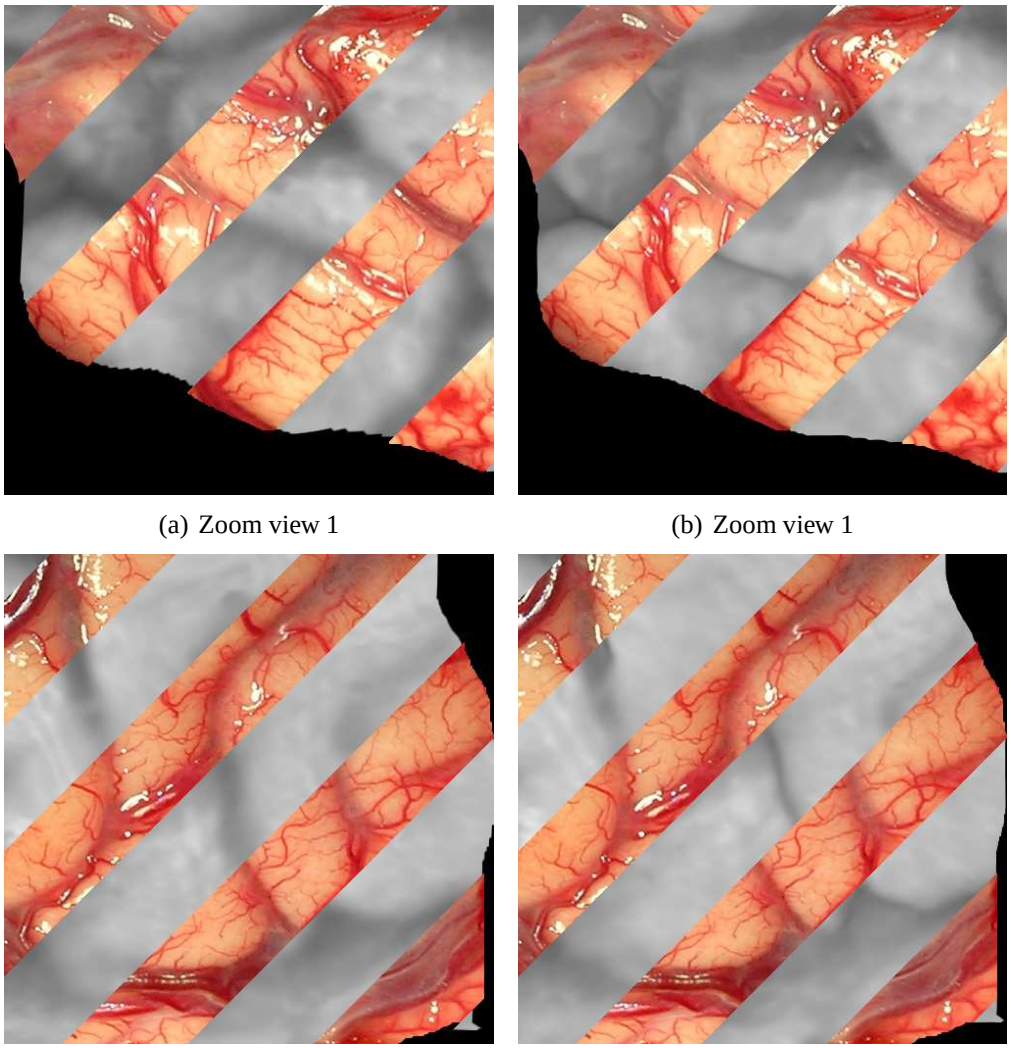
(b) Interlace-strip of registered MR projection and original photograph.

(c) Interlace-strip of unregistered MR projection and photograph

(d) Interlace-strip of registered photograph and original MR projection.

Figure 4.7: (a) and (c) are interlace-stripe image of unregistered image. It illustrate the mismatch of photograph and MR projection. (b) and (d) are interlace-stripe image of registered image. They show that the method greatly refines the matching of the desired edge features. Especially the brain sulci and gyri, which are significant for neurosurgery, are nearly perfectly aligned to each other.

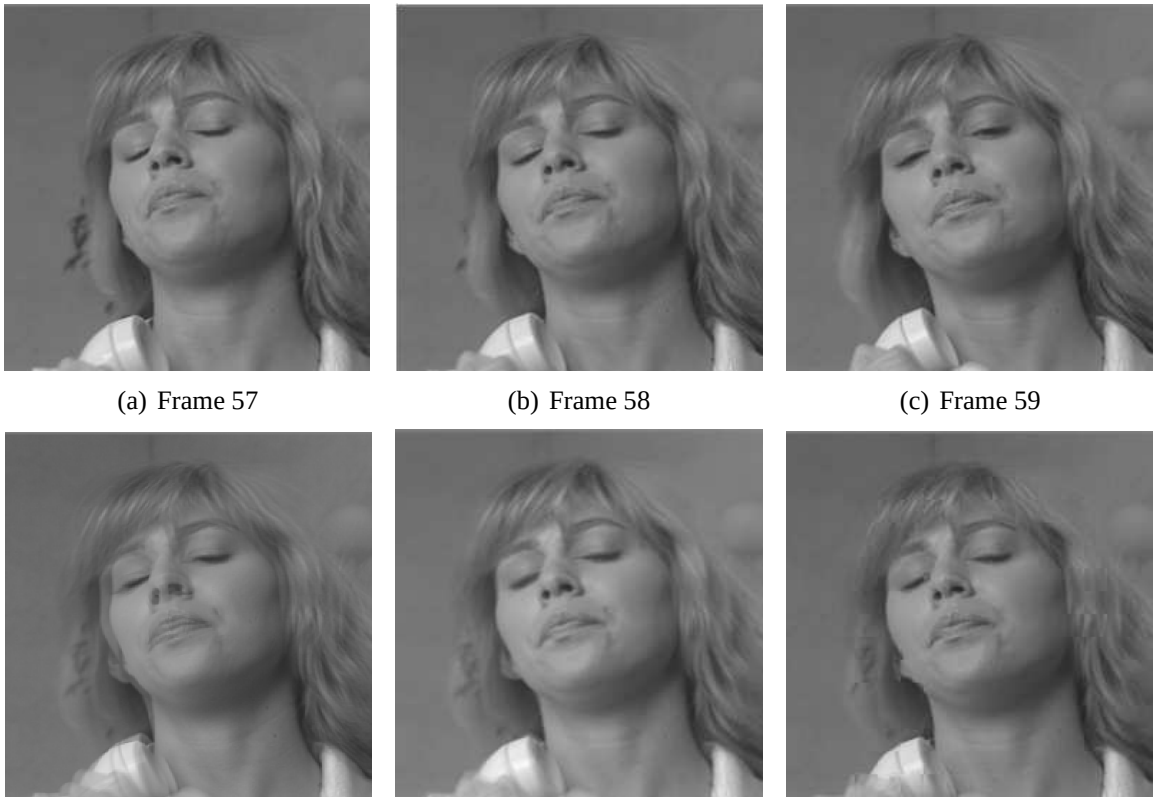
One-to-one edge based registration Mutual information based registration



(a) Zoom view 1 (b) Zoom view 1

(c) Zoom view 2 (d) Zoom view 2

Figure 4.8: Comparison of one-to-one edge matching (a) and (c), and the mutual information based matching (b) and (d). The two algorithms are implemented in a same Finite Element framework including the step size controlled, regularized multi-scale descent. The first row shows how the pre-processed images are registered by the two methods. The zoomed views of local regions in the registered images show that one-to-one edge matching performs a better registration of fine structures.



(a) Frame 57 (b) Frame 58 (c) Frame 59

(d) Averaging interpolation (e) One-to-one edge based registration (f) Standard block matching

Figure 4.9: Top: Original frame 57, 58 and 59 of Susie sequence. Bottom: the interpolated frame 58 using simply averaging, one-to-one edge matching motion estimation and standard block matching motion estimation. The experiment shows that one-to-one edge registration based motion estimation gives an excellent visual quality of frame interpolation, respect to the other two methods.

Part III

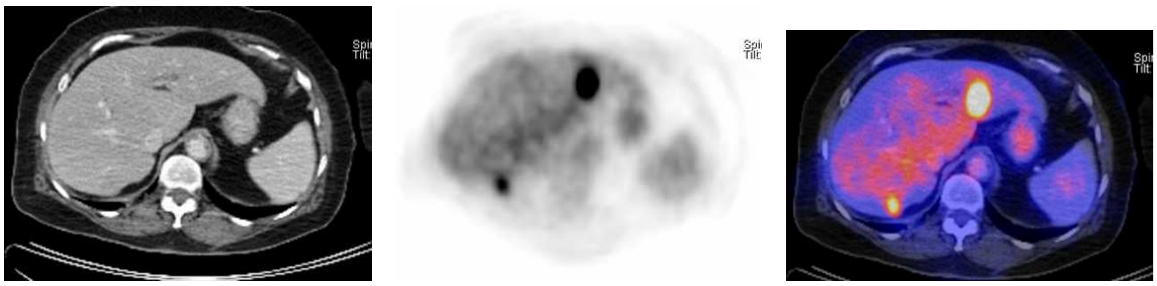
Segmentation based Validation for SPECT/CT Hybrid Imaging

Chapter 5

Accuracy Evaluation of SPECT/CT Imaging

5.1 Fusion of SPECT and CT

Hybrid scanners, which enable the performance of Single Photon Emission Computed Tomography (SPECT) and X-ray Computed Tomography (CT) in one imaging session, have been one of the greatest advancement in the field of medical imaging in the last decade. See the example in Figures 5.1(a)-5.1(c). The coupling of SPECT and CT has been proved to have considerable diagnostic potential. It offers physicians the opportunity to acquire spatially correlated physiological and morphological information in a single session. These hybrid systems have greatly improved the diagnostic accuracy and have therefore been widely accepted clinically [HR06, RNU+06, USM+06, SRCG04, KO07, KRH07, SLRG07, BAO06].



(a) CT (b) SPECT (c) SPECT/CT

Figure 5.1: Combining SPECT/CT imaging. Image source: nuclear medicine department of university of Erlangen.

With respect to traditional SPECT imaging, the largest contribution of hybrid SPECT/CT

Figure 5.2: Siemens Symbia SPECT/CT system combines variable angle dual detector SPECT with 6-slice CT and allows accurate CT based attenuation correction. The graphics courtesy of Prof. T. Kuwert from nuclear medicine department of university of Erlangen.

imaging is that the metabolism activities of patients are not only clearly visualized but also correctly mapped to the anatomical position. "Historically, nuclear medicine has focused on radiopharmaceuticals trapped in organ structures and the presence of disease hallmarked by the absence of activity. A conventional nuclear medical exam is usually followed with additional procedures, like a biopsy, to determine the particular disease process." explains David Rollo, chief medical officer of Philips Medical Systems. Today, hybrid imaging allows newly developed SPECT tracers to be target-specific, concentrating in the designated tissues or organs. Therefore the physicians can observe the particular disease processes more precisely.

Moreover, hybrid imaging eases the correction of attenuation effects of SPECT imaging, using anatomic maps derived from CT. The measurement principle of SPECT shows that attenuation of SPECT is the loss of these useful photons, either by photoelectric absorption or by scattering in an angle sufficiently large that they can no longer be detected. Assuming point A and point B are respectively close to and far away from the surface. These two points have the same degree of radioactivity. Due to the attenuation effect, the detection of activity from point A is easier than from point B and in the reconstructed image the pixel intensity of A is higher than of point B. In order to eliminate this attenuation artifact, tissue density maps (distribution of attenuation coefficient) have to be taken into account. On stand-alone SPECT scanners, transmission scans are often performed to determine them. On hybrid scanners, the attenuation coefficients can be computed directly from the CT data (Hounsefield units), according to a

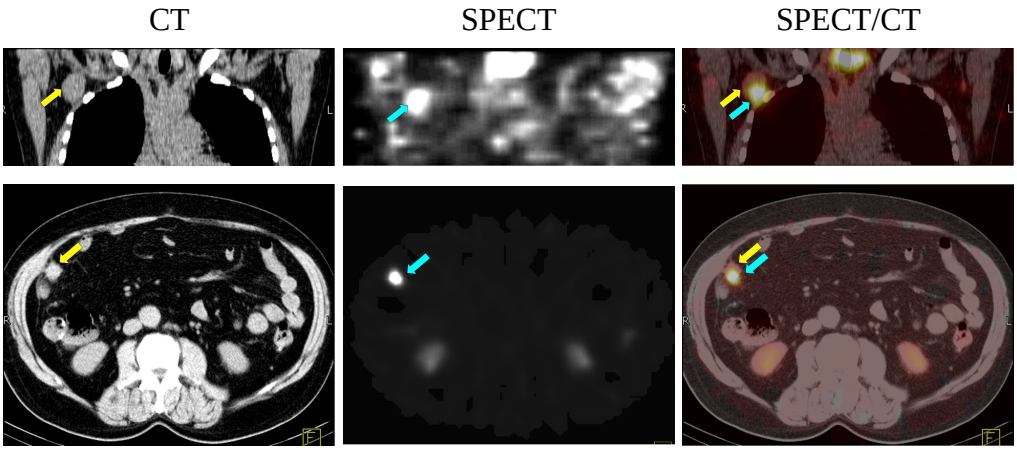


Figure 5.3: Two examples of mismatching of hybrid scanners. The yellow arrows point to the lesions in CT and blue arrows point to the corresponding hot spot in SPECT. Upper row: Coronal views of somatostatin receptor SPECT and CT in a patient with a neuroendocrine carcinoma (NEC). Lower row: Transversal views of somatostatin receptor SPECT and CT in another NEC patient.

general model developed in [BSDSZ03].

5.2 Evaluation of Accuracy

However, combined SPECT/CT imaging is still far from being perfect. A major problem is that SPECT and CT scanning cannot be performed at the same time, although they are done within one examination. For example, the Siemens Symbia SPECT/CT system in Figure 5.2, combining a dual-detector variable angle gamma camera with a spiral CT scanner, requires about one hour for the SPECT exam and only a few seconds for the CT exam of a lung region. It seems that the CT scanner takes a “snapshot” of the objects while the SPECT image accumulatively records a blurry “averaging” functional information. Between the acquisition of the CT and SPECT, the respiratory or cardiac motion as well as patient movements result in the mis-registration of hybrid imaging. The misalignment artifact does not only deteriorate the imaging quality, but also negatively influences the diagnoses based on the SPECT/CT. Figure 5.3 shows two examples of misalignment between CT and SPECT datasets.

Evaluating the anatomical accuracy of image fusion inherent to these systems remains a challenge. Phantom studies are not suitable for the validation of the hybrid scanners because it is technically difficult to simulate such complicated deformations. Due to our knowledge,

Table 5.1: Segmentation algorithms for two validation markers of SPECT/CT imaging

	kidneys	mal-functional hot spots
CT	Active Shape Model [CTCG95, HWWM05]	Random Walk Segmentation [Gra06]
SPECT	Shape Template Matching	Localized Maximally Stable Extremal Regions [MCUP02]

the anatomical accuracy of SPECT/CT scanners has not been sufficiently validated. In a recent study [NWD+06], the accuracy of a SPECT/CT system has been preliminarily evaluated by measuring the distance between the centers of gravity of corresponding lesions in two modalities. However, reproducibility and accuracy of the validation method were not guaranteed, since the centers of gravity were interactively selected by the users.

The next two chapters present more reliable and more accurate methods for evaluating the fusion quality of SPECT/CT hybrid scanners with minimum user interaction. The basic idea is to segment the corresponding objects in both modalities and then measure the distance between the centers of gravity of the segmented objects. This distance measurement quantitatively describes the fusion quality of SPECT/CT datasets. A small distance between two centers of gravity indicates a high accuracy fusion. Actually, such kinds of distance measurements have been used to compare various registration techniques in several clinical studies. Compared to previous studies, the work reported in this chapter makes the following three major contributions:

- Two kinds of anatomical objects have been selected as evaluation markers for two different examining regions, i.e. hot spots for the neck region and kidneys for the abdominal region.
- Four different full- or semi-automatic segmentation methods have been found as well as adapted, in order to correctly extract hot spots and kidneys in both modalities, respectively. The selected segmentation algorithms are listed in Table 5.1.
- The proposed evaluation methods have been validated using real clinical datasets and the accuracy and reproducibility of the measurements have been proven by the experiments.

The validation tool has been successfully integrated into a commercial software application for medical image analysis (InSpace, Siemens Medical Solutions).

Chapter 6

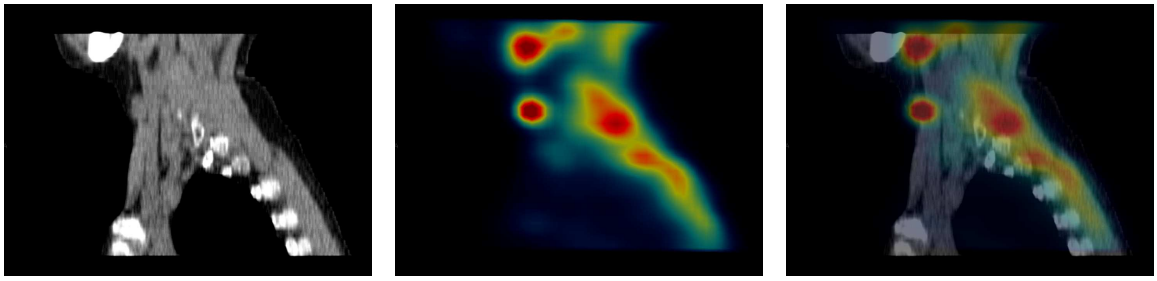
Evaluation using Hot Spots

In this chapter we will introduce the accuracy validation of SPECT/CT fusion using hot spot segmentation and present the evaluation of the validation method on patient datasets.

6.1 Segmentation for SPECT Hot Spots

The determination of the surface that separates different physiological features in functional images is difficult because of the low spatial resolution, the blurring of the edges and the high noise characteristics of functional images. Thresholding is one of the most widely used techniques to segment the volume of interest in functional image data. The threshold can have a fixed value, for example in [GDMK76] $25\%, 40\%$, or 50% of the maximal gray level is used. The threshold can also be automatically computed for each individual image. A classic adaptive thresholding method, histogram based thresholding [Ots79], has been widely used to segment the object in the SPECT volume [MJG$^+$86]. It determines the threshold value by maximizing the variance between the population of background voxels and object voxels. But the distinction between the background and the object class in the histogram is often too weak to find the optimal threshold. The study [KLB91, EWLE95] shows that the threshold of the SPECT lesions can be correctly calculated with knowledge of the size of lesions and the activity of the background. However, the prior knowledge about lesions is often not available in routine circumstances. In this work, we present a fully automatic thresholding method for segmentation of SPECT hot spots. The method is based on a localized version of MSER algorithm [DB06, MCUP02], which does not need any prior knowledge about the object.

A discrete three-dimensional image is defined as a mapping $I : \Omega_z \in \mathbb{Z}^3 \rightarrow G_z$, where Ω_z is the set of voxels and G_z is the set of gray levels, e.g. $\{0, 1, ..., 255\}$ is a typical set for



(a) CT (b) SPECT (c) Fusion

Figure 6.1: Sagittal views of a hot spot in the neck region. It is a physiological accumulation of submandibular gland.

SPECT images. **Region** $\mathcal{Q}$ is a continuous subset of Ω_z, for which each pair $(\boldsymbol{x}_1, \boldsymbol{x}_2) \in \mathcal{Q}$ of pixels is connected by a path fully contained in $\mathcal{Q}$. The algorithm starts from a seed $\boldsymbol{x}_0$, typically user-specified, which is located within the hot spot. In this work, we define a **localized extremal region** $\mathcal{Q}_g : g \in G_z$ as a maximal region that contains the seed $\boldsymbol{x}_0$ and $I(\boldsymbol{x}) \geq g$ for all $\boldsymbol{x} \in \mathcal{Q}_g$. We say g is the **base-level** of the localized extremal region $\mathcal{Q}_g$. It is clear that the sequence of localized extremal regions $\mathcal{Q}_0, ..., \mathcal{Q}_g, \mathcal{Q}_{g+1}, ..., \mathcal{Q}_{I(\boldsymbol{x}_0)}$ is nested, i.e. $\mathcal{Q}_g \supset \mathcal{Q}_{g+1}$ and $\boldsymbol{x}_0 \in \mathcal{Q}_g$. For the purpose of segmentation, we are more interested in the extremal region $\mathcal{Q}_g^*$ that is maximally stable if the area variation

$$\rho^{\Delta}(g) = \frac{\#(\mathcal{Q}_{g-\Delta}) - \#(\mathcal{Q}_{g+\Delta})}{\#(\mathcal{Q}_g)} \tag{6.1}$$

has a minimum at g^*. Here $\#(.)$ denotes cardinality and $\Delta \in \mathbb{Z}$ is a given parameter.

A straightforward way is to compute the area of all the extremal regions $\mathcal{Q}_0, ..., \mathcal{Q}_{I(\boldsymbol{x}_0)}$, then to choose the region with minimum area variation. This method can successfully segment the "strong" hot spots, which have relatively cool backgrounds and have no other compatible hot objects in the neighborhood. The submandibular gland hot spot in the left of Figure 6.2 belongs to this case.

However, in practice lesions frequently appear as "weak" hot spots, whose intensities are not so distinctive from the background, or some other hot objects are close to the hot spots lesion. These global maximally stable regions cannot always correctly segment this class of weak hot spots. The liver lesion hot spot in the right of Figure 6.2 is an example of a weak hot spot. The plot of area variation with respect to various base-levels clearly shows the problem. The minimum of area variation lies in the low base-level, which is far away from the true hot spot base-level. In order to avoid this problem, the algorithm should not search for the maximally

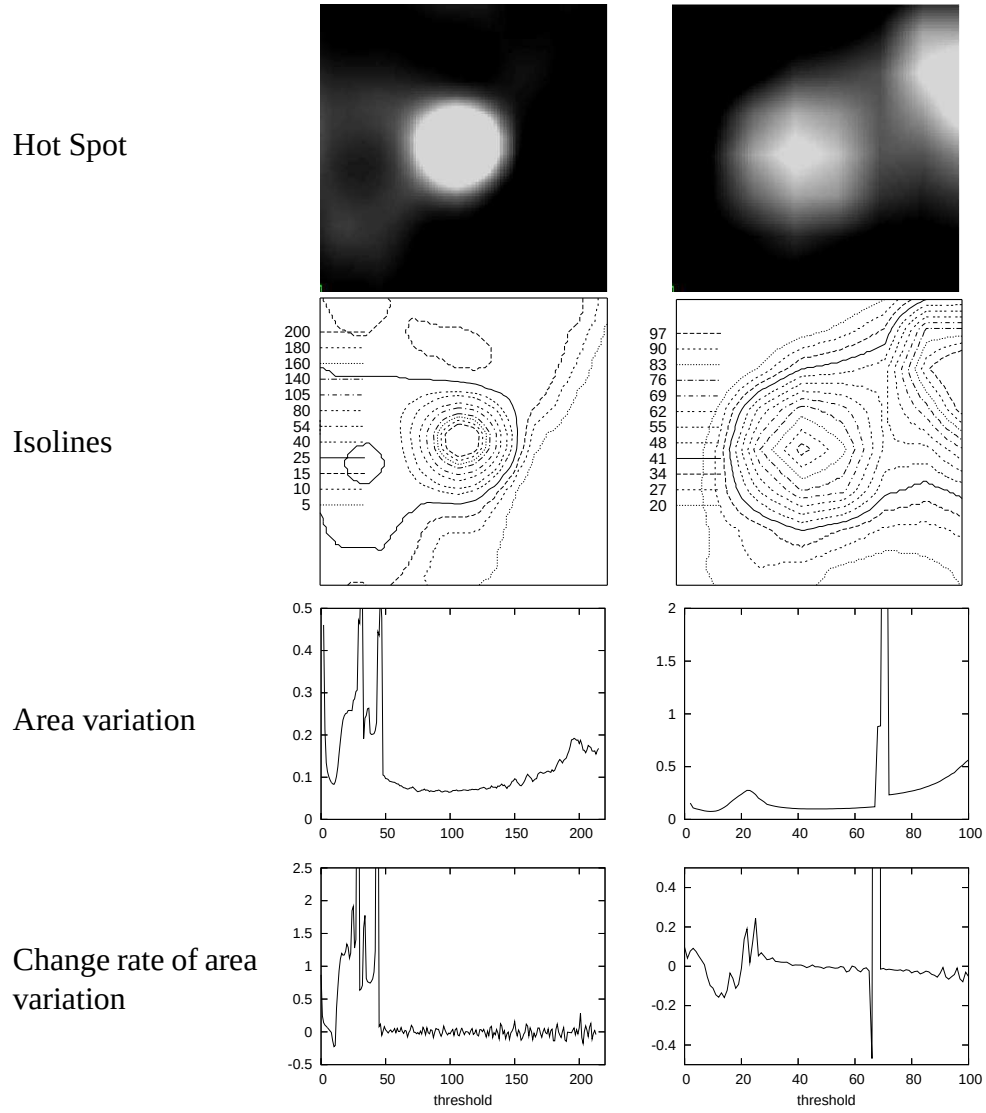


Figure 6.2: Two examples of segmentation of SPECT hot spot. Left: A physiological accumulation of submandibular gland. Right: A tumor lesion of liver. $\Delta = 1$ for both examples.

stable region among all the extremal regions but only among a set of localized extremal regions that always contain the seed $\boldsymbol{x}_0$ but have not yet merged with surrounding regions. The merge of the localized extremal region with surrounding regions usually results in a dramatic increase of region area. It can be easily detected by the change rate of area variation (see the last row of Figure 6.2). Because the algorithm proceeds from relatively high base-level $\mathcal{Q}_{I(\boldsymbol{x}_0)}$ to the lowest base-level, the change rate of area variation is defined as

$$\varrho^{\Delta}(g) = \frac{\rho^{\Delta}(g) - \rho^{\Delta}(g+1)}{\rho^{\Delta}(g+1)}. \tag{6.2}$$

The implementation of the method is summarized as follows:

Algorithm 6 Localized MSER for SPECT Hot Spot

```
Manually select a point x_0 within the SPECT hot spot as a seed
for i = I(x_0), ..., 0 do
   compute the area of extremal region |Q_g|
   if g > 2Δ then
      compute the area variation of region ρ^Δ(g − Δ)
      if ϱ^Δ(g − Δ) > ξ, set g' = g − Δ then break
   end if
end for
among the regions: Q_{I(x_0)}, ..., Q_{g'},
find the localized MSER Q_{g*} that has minimal area variation.
```

Relationship with standard MSER. The localized MSER is actually a simplified version of the standard MSER. In the standard MSER method, the stable analysis is performed on a region tree, which is efficiently calculated by a partition method [CLG05]. The merge of two regions has been implicitly encoded in this rooted tree structure. The additional merge detections are not necessary any more. While the localized MSER actually starts the connected thresholding from the given seed and seeks a range of thresholds that leaves the peak of hot spot effectively unchanged. Although both methods theoretically can achieve almost the same segmentation of the hot spot, the localized MSER appears to be more efficient than the standard MSER method. The partition algorithm of standard MSER method has a computational complexity of $\mathcal{O}(n \log\log n)$, where n is the number of voxels. The computation time of the localized MSER method is dependent on the size of the desired object. For the segmentation of a small hot spot among a large volume data, the localized MSER method becomes more efficient than the standard MSER method.

6.2 Random Walk Segmentation in CT Datasets

In this application, the largest challenge of segmentation of corresponding structures in CT is the unclear boundary between the desired object and irrelevant surrounding tissues. This is demonstrated in Figure 6.3(a).

Most fully automatic segmentation methods that do not use prior knowledge of the shape and size of the object cannot reliably locate such almost invisible boundaries. For example, the standard fast marching level set method iteratively evolves the level set interface (boundary) based on a speed function, which is typically a function of the gradient magnitude. As shown in Figure 6.3(b), the speed function cannot stop the level set interface crossing the invisible boundary. Given a perfect parameterization, some more advanced segmentation methods can satisfactorily segment the lesion in this example. However, it is very tricky for users to determine the optimal setting of multiple algorithm parameters for each segmentation task. For example, the well-known gradient vector flow method [XJ. L. Prince98] requires the user to specify four parameters at the same time. Moreover, it is difficult to define objective stopping criteria for many fully automatic iterative methods. In practice, the user gives a fixed number of iterations or interactively stops the iterations. Due to the tedious parameterization and subjective stopping criterion of fully automatic methods, many medical researchers favor interactive segmentation, with the help of a mouse or a light pen. However, the manual selection of a three-dimensional boundary is not only time-consuming, but also non-reproducible.

Random walk segmentation [Gra06] is chosen to detect structures in this application. This intuitive semi-automatic segmentation allows the user to select seeds to mark the inside and outside regions. Then the algorithm determines the optimal partition based on the image intensity and pre-selected seeds. Figure 6.3(c) shows the selected seeds of the neck lesion and Figure 6.3(d) shows the lesion segmented by random walk method.

The two-label (object and background) random walk is described here with a simple 3×3 image in Figure 6.4. For the description of the general multi-label case, we refer to [Gra06]. In the random walk segmentation, an image is modeled as a graph of nodes and edges, where each node is a voxel and an edge connects the adjacent nodes. Assuming that the user selects node 2 and 7 as a seed of background and object respectively, for each unlabeled node i, the algorithm would determine the following: given a random walker starting from this node, what is the probability p_i that it first reaches the object seed? Obviously the probabilities of object seed and background seed is 1 and 0. When the probabilities of unlabeled nodes are known, the boundary of the object (the curve in Figure 6.4) can be approximated by interpolation between the adjacent object nodes ($p_i > 0.5$) and background nodes ($p_i < 0.5$).

In the random walk algorithm, the image structure is represented by the edge weights:

$$w_{ij} = \exp(-\beta(I(\boldsymbol{x}_i) - I(\boldsymbol{x}_j))^2), \tag{6.3}$$

where $I(\boldsymbol{x}_i)$ and $I(\boldsymbol{x}_j)$ are the intensities of nodes i and j, $\beta \in \mathbb{R}$ is the only free parameter in the algorithm. Intuitively, if two adjacent voxels have a large variation of intensity, the random walker has a relatively low probability of crossing the edge between them. The probabilities of unlabeled nodes are determined by a combinatorial Dirichlet problem, in which the discrete Laplacian matrix L is defined as:

$$\boldsymbol{L}_{ij} = \begin{cases} -w_{ij} & \text{if node } i \text{ and } j \text{ are adjacent,} \\ d_i = \sum_i w_{ij} & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

The given seeds serve as known values ($p_2 = 0, p_7 = 1$), which can be moved into the right hand side vector. The corresponding linear system is

$$\begin{bmatrix} d_0 & -w_{01} & -w_{03} & & & & & \\ -w_{01} & d_1 & & -w_{14} & & & & \\ -w_{03} & & d_3 & -w_{34} & & -w_{36} & & \\ & -w_{14} & -w_{34} & d_4 & -w_{45} & & & \\ & & & -w_{45} & d_5 & & -w_{58} & \\ & & -w_{36} & & & d_6 & & \\ & & & & -w_{58} & & d_8 & \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ p_8 \end{bmatrix} = \begin{bmatrix} 0 \\ p_2 w_{12} \\ 0 \\ p_7 w_{14} \\ p_2 w_{25} \\ p_7 w_{16} \\ p_7 w_{78} \end{bmatrix}. \tag{6.4}$$

Generally, the system equation of the random walk segmentation is denoted as

$$\boldsymbol{L}\boldsymbol{p} = \boldsymbol{f}. \tag{6.5}$$

With respect to the retrospective methods, random walk segmentation has five advantages.

1. If sufficient seeds can be correctly selected, small objects with partially unclear boundaries can be satisfactorily segmented.

2. Compared to a fully manual selection, the user interaction is minimal. From our experience, drawing the seeds on the middle slice in the axial, sagittal and coronal views is sufficient.

3. The method does not require any stopping criteria. The algorithm only requires the solution of a large, sparse, symmetric positive-definite system of linear equations.

4. The segmentation is almost real-time. A fast multigrid solver that has complexity $\mathcal{O}(n)$ is implemented to solve the linear system.

5. The method has only one algorithm parameter that is kept constant in all experiments.

Next we briefly introduce a fast node based multigrid (MG) method to solve the linear system of Equation 6.5. For the definitions of mathematical notations and the detail of this multigrid solver, we refer to [TOS01, BHM00] and our previous works [KKR07, KSR07]. Multigrid solvers are based on the assumption that high frequency errors can be treated efficiently by an appropriate smoother on a fine scale and low frequency errors are approximated on coarser scales. Therefore we build up an image pyramid and construct Equation 6.5 on each level. The whole algorithm of MG solver can be recursively defined by one multigrid iteration, computing $\boldsymbol{p}_m^{(k+1)} = \mathcal{M}_m(\boldsymbol{p}_m^{(k)}, \boldsymbol{L}_m, \boldsymbol{f}_m, \nu_1, \nu_2)$ by Algorithm 7.

Algorithm 7 Multigrid Correction Scheme

given an initial guess $\boldsymbol{p}_m^{(0)} = \mathbf{0}$
pre-smoothing $\widetilde{\boldsymbol{p}}_m^{(k)} = S_m^{\nu_1}(\boldsymbol{p}_m^{(k)}, \boldsymbol{L}_m, \boldsymbol{f}_m)$
compute residual $\boldsymbol{r}_m = \boldsymbol{f}_m - \boldsymbol{L}_m \widetilde{\boldsymbol{p}}_m^{(k)}$
restrict residual $\boldsymbol{r}_{m'} = \mathcal{I}_m^{m'} \boldsymbol{r}_m$
if number of coarse grid points $< \epsilon_{min}$ **then**
  Solve $\boldsymbol{L}_{m'} \boldsymbol{e}_{m'} = \boldsymbol{r}_{m'}$ exactly
else
  $\boldsymbol{e}_{m'} = \mathcal{M}_{m'}(0, \boldsymbol{L}_{m'}, \boldsymbol{r}_{m'}, \nu_1, \nu_2)$
end if
interpolate error $\boldsymbol{e}_m = \mathcal{I}_{m'}^{m} \boldsymbol{e}^{m'}$
coarse grid correction $\widetilde{\boldsymbol{p}}_m^{(k)} = \widetilde{\boldsymbol{p}}_m^{(k)} + \boldsymbol{e}_m$
post-smoothing $\boldsymbol{p}_m^{(k+1)} = S_m^{\nu_2}(\widetilde{\boldsymbol{p}}_m^{(k)}, \boldsymbol{L}_m, \boldsymbol{f}_m)$

As smoother we apply a line-wise red-black Gauss-Seidel method denoted by S_m^{ν}, where the parameter ν specifies the number of performed Gauss-Seidel iterations and m is the current scale or level. Then we compute the residual $\boldsymbol{r}_m$ and restrict it to the next coarser level m' by a full weighting restriction operator $\mathcal{I}_m^{m'}$. Afterwards the so-called error equation $\boldsymbol{L}_{m'} \boldsymbol{e}_{m'} = \boldsymbol{r}_{m'}$ is solved on level m' recursively by multigrid. The coarse matrix $\boldsymbol{L}_{m'} = \mathcal{I}_m^{m'} \boldsymbol{L}_m \mathcal{I}_{m'}^{m}$ is computed by Galerkin coarsening. Next the error $\boldsymbol{e}_{m'}$ is interpolated by a trilinear interpolation operator $\mathcal{I}_{m'}^{m}$ to the fine level m and used there as a correction to the current solution $\boldsymbol{p}_m$. Finally

we apply again some post-smoothing Gauss-Seidel steps at level m. As stopping criterion we check, if the norm of the residual $\|\boldsymbol{r}_m\|$ drops below a given threshold that we choose as 10^{-8} in our implementation. To reduce the number of multigrid iterations we additionally use iterant recombination [BM95, CKR05] that is similar to preconditioning.

The solution of the system for a typical $40\times40\times40$ sub-volume, using the described multigrid solver with 2 pre- and 2 post-smoothing steps and 5 V-cycles, requires less than 2 seconds on an AMD Athlon 3200+ computer (2.20 GHz, 2.00 GB RAM).

6.3 Experiments

6.3.1 Patient Dataset

To evaluate the validation tool, 21 patients, 13 females and 8 males between 10-80 years old with the average age at 59.22, were examined by a SPECT/spiral CT scanner (Siemens Symbia system) between November 2006 and March 2007. Datasets were selected where both the hot spot on SPECT and the corresponding structure on CT were clearly visible in the neck region. We chose adenomas of the parathyroid glands on 8 patients, the physiological accumulations of the submandibular gland on 10 patients, thyroid nodule on 1 patient, neuroendocrine tumor on 1 patient, thyroid carcinoma on 1 patient for this study, where Tc-99m MIBI (18 patients), I-131-NaI (2 patients) and In-111-SMS (1 patients) were used as tracers.

6.3.2 SPECT Hot Spot Segmentation

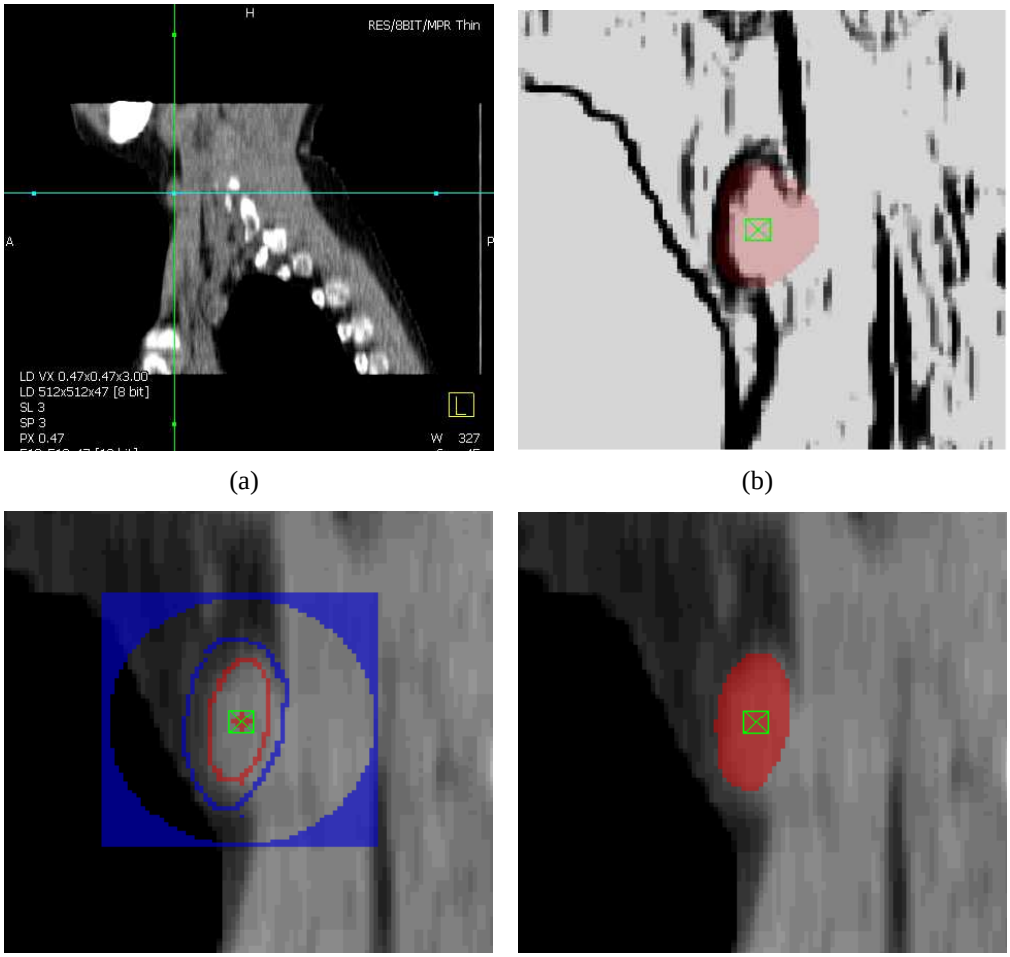
First we evaluated the localized MSER method for the segmentation of SPECT hot spots. The proposed method was compared with the histogram based thresholding [Ots79], a widely used segmentation method for functional images [MJG+86]. Two radiologists were asked to evaluate the segmented hot spots by the two methods. To avoid any bias in the ratings, the evaluation was carried out independently by the two radiologists. The segmentation results were scored between 0-2 by the radiologists: (1) If one segmentation method was obviously better than the other, this method would be scored 2 and the other method would be scored 0. (2) If both segmentation methods had comparable results, both methods would be scored 1. For the hot spots in the selected data-sets, the evaluation showed that the radiologists were much more satisfied with localized MSER segmentation: The average scores of the localized MSER method were 1.667 and 1.619, while the average scores of the histogram based thresholding were only 0.333 and 0.381, respectively.

6.3.3 Accuracy Test

The accuracy of the validation tool was evaluated as follows: Two operators perform the validations independently. One operator directly used the validation tool to measure the distances in X-, Y- and Z-direction (t_x, t_y, t_z) between the hot spot on SPECT and the structure on CT. In the same way, the second operator validated the SPECT/CT volumes, where the SPECT volume had been artificially shifted in X-, Y- and Z-directions. The shift parameters (s_x, s_y, s_z) were randomly generated between 5 mm and 10 mm or between -5 mm and -10 mm. We denote the distances measured by the second operator as $(\widetilde{t}_x, \widetilde{t}_y, \widetilde{t}_z)$. The extent to which the ground truth shift (s_x, s_y, s_z) and the measured shift $(d_x, d_y, d_z) := (\widetilde{t}_x - t_x, \widetilde{t}_y - t_y, \widetilde{t}_z - t_z)$ match, indicates the accuracy of the validation. As shown in Figures 6.5(a)-6.5(c), the experiment yielded a clear linear association between the ground truth and the measurement: The correlation coefficients are 0.9927, 0.9909 and 0.9853 in X-, Y- and Z-directions, respectively. The anatomical inaccuracies, measured by the mean $\pm$ standard deviation of the absolute error, were reported to be 0.7189 ± 0.6298 mm in X-direction, 0.9250 ± 0.4535 mm in Y-direction and 0.9544 ± 0.6981 mm in Z-direction, respectively.

6.3.4 Reproducibility Test and Time Measurement

To evaluate the intra-observer reproducibility, the distances between the SPECT hot spot and CT structure were measured 20 times in five different patients, yielding a mean standard deviation of 0.2177 mm in the X-direction, 0.3039 mm in the Y-direction and 0.3350 mm in the Z-direction respectively. This indicated a high intra-observer reproducibility of the measurements of the X-, Y- and Z-distances. The mean time for a full validation process, including data loading and user operations, was less than 2 minutes on an AMD Athlon 3200+ computer (2.20 GHz, 2.00 GB RAM).



(a) (b)

(c) (d)

Figure 6.3: (a): The sagittal view of a patient with a lesion in the neck. (b): Gray levels represent the normalized speed function of the fast marching level set method. The red region represents the segmentation result after 100 iterations. (c): The view of seed selection for the random walk method. The blue frame defines the region of interest. The blue lines (seeds), drawn by users, define the outside of the lesion, while the red ones define the inside of the lesion. (d): The region segmented by the random walk method.

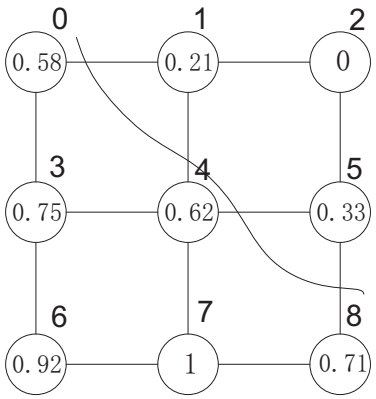
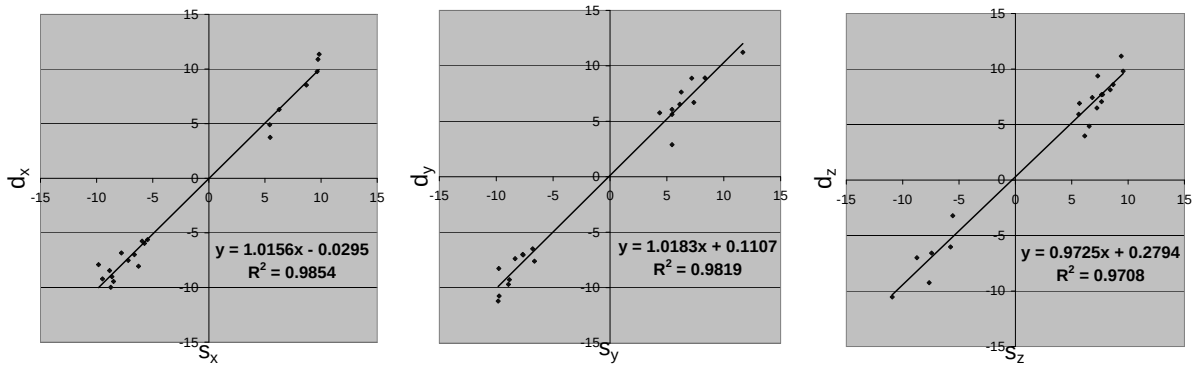


Figure 6.4: The graph of a simple 3×3 image, where node 2 and node 7 are the seed of background and object. The values inside each node indicate the probabilities that the random walker starting from this location reaches the object seed.



(a) s_x-d_x (b) s_y-d_y (c) s_z-d_z

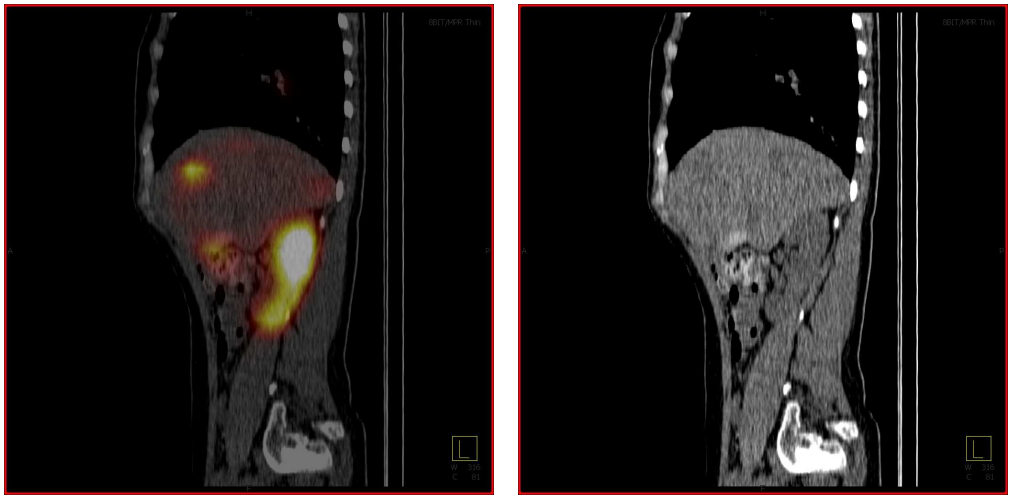
Figure 6.5: Comparison of the ground truth shift (s_x, s_y, s_z) and the measured shift (d_x, d_y, d_z). Two shift parameters in X-, Y- and Z-directions are close to lines of identity.

Chapter 7

Evaluation using Kidneys

Although the method introduced in the previous chapter has been successfully used in some clinical studies and its accuracy and reproducibility have been proven in some specific regions, like regions of neck and head, it is still far away from a general evaluation scheme for multi-modal registration. The choice of a hot spot as the evaluation marker limits the applicability of this evaluation method, because the patient data sometimes has no hot spots in both modalities or the hot spots in either modality (quite often in the low-dose CT) are not clear enough to be correctly segmented. Figures 7.1(a) and 7.1(b) shows an example where the lesion is almost invisible in the low-dose CT, so that the automatic segmentation of such a almost invisible object is nearly impossible. The snapshot in Figures 7.1(a) and 7.1(b) also gives the clue to solving the problem in this case: the kidney can be a good alternative as the evaluation marker.

More generally, the accuracy of matching between morphological and functional data can be measured by the degree of matching of the organs that can be reliably segmented from the both modalities. The candidates can be kidney, liver or bladder, not only because all these organs are normally highly active in the functional modalities but also the statistic shape models of organs can be created to effectively guide the segmentation of the organ among the noisy images. In this chapter, we mainly discuss the kidneys based evaluation method for SPECT/CT fusion. One should note that the idea can be easily extended to other organ markers and other combinations of modalities. Section 7.1 gives an introduction on how to construct the active shape model of kidneys and how to customize the model to segment kidneys in CT volumes. In Section 7.2, a template matching method is used to localize the kidneys among SPECT modality. In Section 7.3 we introduce a fast method to automatically correct the misalignment of kidneys between SPECT and CT images. Section 7.4 presents the validation test of the method.



(a) SPECT/CT (b) CT

Figure 7.1: Sagittal views of a patient abdomen region. In this example, the hot spot in the liver is not suitable to evaluate the accuracy of SPECT/CT fusion, because the corresponding structure in the low-dose CT dataset is almost invisible. A possible alternative of the evaluation marker could be the kidney, which is clearly visible in both modalities.

7.1 Active Shape Model Segmentation for CT Kidneys

Active Shape Models (ASMs), first proposed in [CTCG95], are statistical models of the shape of objects which are built by learning patterns of variability from a training set of correctly labeled images. Active shape models are an elegant way to represent the inherent inter- and intra-personal shape variability inherent in biological structures, where shapes of structures are similar but not identical. The shape models can iteratively deform to fit the new objects in the images, however, the models should only be able to deform in ways such that the characteristic of the class of objects is still represented. The constrained deformation or model based fitting enables ASM to more reliably segment the desired biological structure, even though the image may be corrupted by the noise or parts of the boundaries are missing. The active shape model based segmentation can be roughly decomposed into two parts:

- Construction of the point model of the object of interest. It requires the alignment of training data, determining the point correspondence between different data sets and building the point distribution models using Principal Component Anaysis (PCA).

- Instantiation of the derived shape model for the segmentation of the object among the current image data. It updates not only the pose parameters, i.e. translation and rotation, but also the shape parameters defined in the shape models.

In the rest of this section, the active shape model used for segmentation of a single kidney will be briefly introduced. The implementation details of the model can be found in [Spi06, SHV+09]. Because the Point Distribution Model (PDM) plays the key role in the ASM for modeling the training data, parameterization of shape as well as segmentation of the object, PDM will be defined in Subsection 7.1.1. Subsequently, the method of pre-registration of training data and the method for point correspondence will be discussed in Subsection 7.1.2. Finally, we show in Subsection 7.1.3 how to adapt ASM to segment the object among the image.

7.1.1 Point Distribution Model

In this subsection we focus on how to apply the concept of PCA to build up the statistic model. Let us assume that we have a set of n_r shapes as training data and these shapes have been correctly aligned to each other. Each input training set is represented by a set of n_s landmarks stored in a single vector $\boldsymbol{\varsigma} \in \mathbb{R}^{3n_s}$. Let $\boldsymbol{\varsigma}_r$ be the vector collecting the coordinates of the n_s points of the r-th shape in such way:

$$\boldsymbol{\varsigma}_r = (x_{r1}, x_{r2}, ..., x_{rn_s}, y_{r1}, y_{r2}, ..., y_{rn_s}, z_{r1}, z_{r2}, ..., z_{rn_s})^T, \tag{7.1}$$

where the $(x_{r2}, y_{r2}, z_{r2})^T$ is, for instance, the coordinate of the second point in the r-th shape. The vectors of all the training data, defined in Equation 7.1, form the columns of the landmark configuration matrix:

$$\boldsymbol{L}_c := (\boldsymbol{\varsigma}_1, \boldsymbol{\varsigma}_2, ..., \boldsymbol{\varsigma}_{n_r}). \tag{7.2}$$

We will discuss the method of determination of these consistent landmarks among the training data in the Subsection 7.1.2.

One can imagine that these $3n_s$-D points cluster within some region of the space, the so-called "Allowable Shape Domain" [CTCG95]. The center of this point cloud, presenting the mean shape $\overline{\boldsymbol{\varsigma}}$ of n_r different training examples, is calculated by

$$\overline{\boldsymbol{\varsigma}} = \frac{1}{n_r} \sum_{r=1}^{n_r} \boldsymbol{\varsigma}_r. \tag{7.3}$$

PCA attempts to use a $3n_s$-D ellipsoid to fit this point cloud and each axis of this ellipsoid defines a mode of variation, a way in which the landmarks tend to move as the shape varies. We calculate

the $3n_s \times 3n_s$ covariance matrix:

$$\boldsymbol{S}_c = \frac{1}{n_r}\sum_{r=1}^{n_r}(\boldsymbol{\varsigma}_r - \overline{\boldsymbol{\varsigma}})(\boldsymbol{\varsigma}_r - \overline{\boldsymbol{\varsigma}})^T. \tag{7.4}$$

The principal axes of the ellipsoid are described by $\boldsymbol{p}_i (i = 1, ..., 3n_s)$, the unit eigenvectors of $\boldsymbol{S}_c$ such that

$$\boldsymbol{S}_c \boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i \tag{7.5}$$

where λ_i is the i-th eigenvalue of $\boldsymbol{S}_c$, $\lambda_i \geq \lambda_{i+1}$ and $\boldsymbol{p}_i^T \boldsymbol{p}_i = 1$. The eigenvector corresponding to the largest eigenvalues describes the longest axes of the ellipsoid, and also the most significant variation mode among the training population. The eigenvalue indicates the significance of the corresponding variation mode. In practice, most of the variation can usually be described by a small number of modes (let us say t) that have the largest eigenvalue. In this work, we choose the smallest t, such that more than 90% variation is included, i.e.,

$$\sum_{i=1}^{t} \lambda_i \geq 0.9 \sum_{i=1}^{3n_s} \lambda_i \tag{7.6}$$

Let $\boldsymbol{P}_t = [\boldsymbol{p}_1, ..., \boldsymbol{p}_t]$. The training examples can be approximated by the mean shape and a linear combination of principal axes, i.e.

$$\boldsymbol{\varsigma} = \overline{\boldsymbol{\varsigma}} + \boldsymbol{P}_t \boldsymbol{b} \tag{7.7}$$

The weight vector $\boldsymbol{b} \in \mathbb{R}^t$ can be regarded as the parameter that defines the allowable deformation of shape. The Figure 7.1.1 shows the effects of varying the first and second parameters of the right kidney model.

7.1.2 Point Correspondence

In order to build statistic shape models of right and left kidneys, a set of training shapes needs to be generated. Therefore, 40 different kidney-pairs have been manually segmented by clinical experts. The shape of every single kidney is represented by a set of mesh points of the surface. In order to automatically generate ASM from this training dataset, first of all, the training dataset need to be registered to each other. In other words we need to filter out effects of scale, translation and rotation. One should note that the correspondence of the mesh points are unknown at this

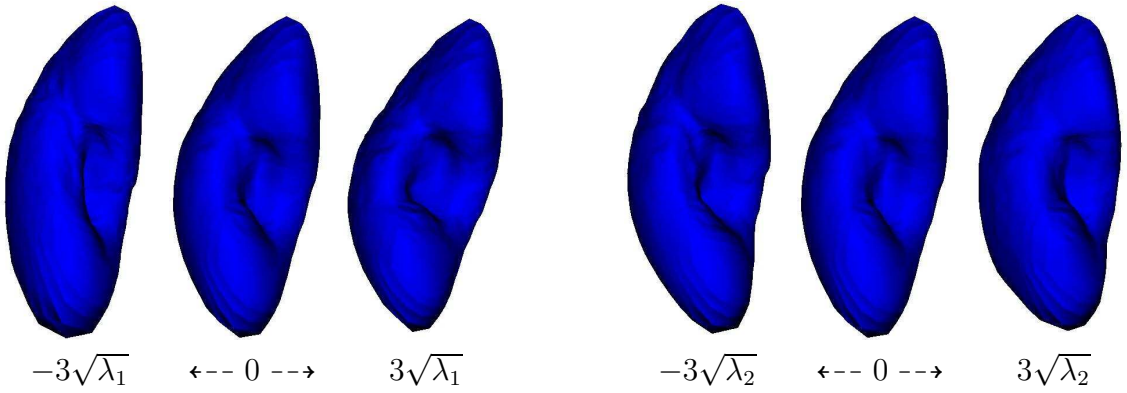


Figure 7.2: Effects of varying the first (left) and the second (right) parameters of the variation model. Graphics courtesy of Martin Spiegel [SHV+09].

point. Thus, for such an alignment problem without given point correspondences, algorithms based on the concept of Iterative Closest Point (ICP) [BM92, CM92] need to be used to register these training surfaces. The details of the ICP algorithm have been introduced on page 27. One notes that the transformation is defined by the parameter vector $\boldsymbol{q} = [\boldsymbol{q}_R|\boldsymbol{q}_T]^t$, where $\boldsymbol{q}_R = [q_1, q_2, q_3, q_4]$ is the unit quaternion and $\boldsymbol{q}_T = [q_5, q_6, q_7]$ is the translation vector.

The next task is to automatically define landmarks, such that the landmarks on different training examples are located at corresponding positions. The problem of establishing a dense correspondence over a set of training surfaces can be posed as that of defining a new parameterization for each training data, leading to a dense correspondence between equivalently parameterized boundary points [DTC+02]. Various attempts have been made to determine the dense correspondence from sets of training surfaces. An overview and comparison of different algorithms for point correspondence can be found in [SRN+03]. The Minimum Description Length (MDL) approach for ASM, which was first proposed in [DCT01, DTC+02] and extended in [HWWM05, HWM06], is implemented to solve the point correspondence problem of the shape model of kidneys. This method minimizes a cost function based on MDL of the resulting statistical shape model and shows a superior performance in comparison with the other approaches.

Mesh Parameterization

Parameterization of a surface mesh is important for the initialization of the correspondence between the training surfaces as well as for an efficient manipulation for a better correspondence in the optimization process afterwards. Roughly speaking, a normal kidney has a shape of genus 0 [LL06], which is topologically equivalent to a sphere. Therefore, a one-to-one mapping is

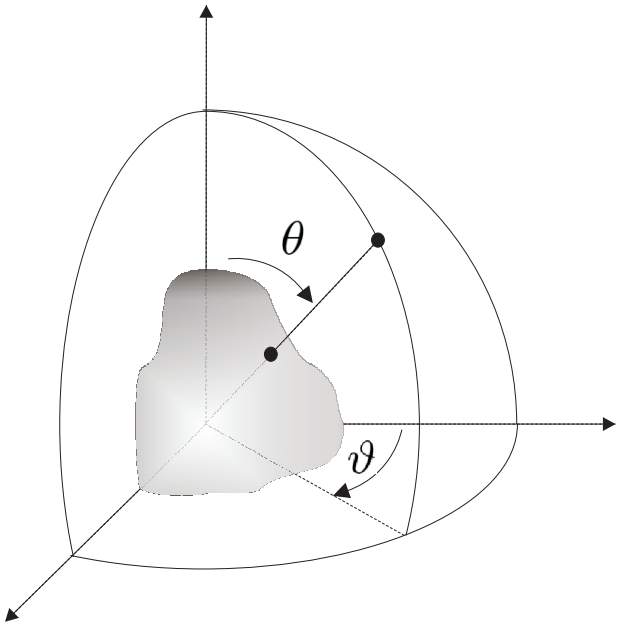


Figure 7.3: Example of barycentric coordinate transformation. One-to-one mapping between the points of shape and unit sphere.

assigned from every points in the mesh to an unique position on the unit sphere, which can be specified by the longitude angle $\theta \in [0, 2\pi]$ and the latitude angle $\vartheta \in [0, \pi]$, as shown in Figure 7.3. There exist a number of approaches to generate the surface parameterization and minimize the distortion error. A comprehensive overview on this topic can be found in [FH05]. In this work, we choose a method that preserves the angles, i.e., if the point on the unit sphere moves in one direction, the corresponding point in the shape will move in the coherent direction.

Cost Function: Minimum Description Length

For automatic landmark generation, a certain cost function needs to be defined to quantify the current parameter setting, such that the chosen landmarks on all the training data are correctly located on the corresponding position. Davies et al. [DCT01, DTC+02] first used the minimum description length (MDL) of the resulting shape model as the cost function to measure the correspondence of landmarks. The key insight is that the "best" model is the one that describes the entire training set as "efficiently" as possible. In practice, less complex variants based on the eigenvalues have been proposed in [Tho03, KT98]. In this work, we choose the MDL cost

function described in [Tho03] to measure the landmark correspondence. It is defined as follows:

$$E_{\text{MDL}} := \sum L_i \quad \text{with} \quad L_i = \begin{cases} 1 + \log(\lambda_i/\lambda_c) & \text{for }, \lambda_i \geq \lambda_c \\ \lambda_i/\lambda_c & \text{for } \lambda_i < \lambda_c, \end{cases} \quad \text{and} \quad \lambda_c = (\frac{\sigma}{\overline{r}})^2, \tag{7.8}$$

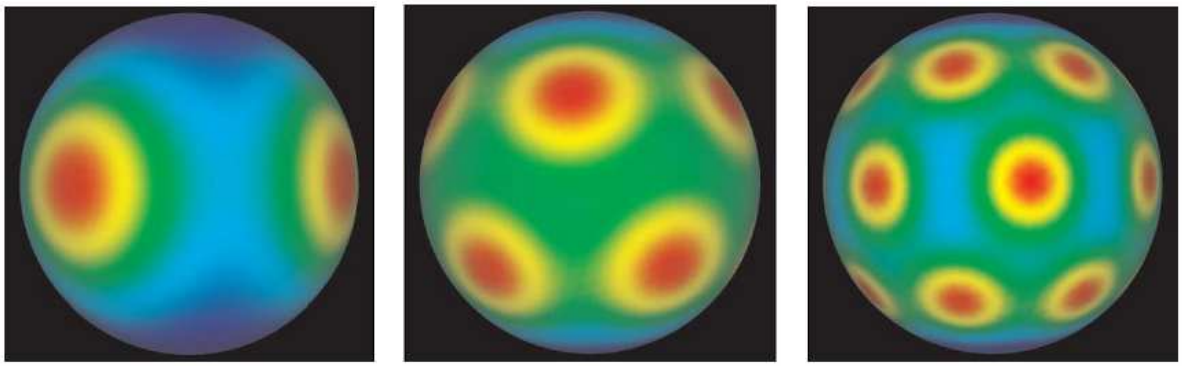
where λ_i is the eigenvalue defined in Equation 7.5, σ is the standard deviation of noise in the training data and $\overline{r}$ is the average radius of the training shapes. Parameter λ_c is the threshold controlling the noise in the training data by separating the eigenvalues into a high lambda region $(\lambda_i \geq \lambda_c)$ and a low lambda region $(\lambda_i < \lambda_c)$. The attractive property of this cost function is that it tends to zero when all the eigenvalues tend to zero, and at the same time both L_i and its derivative are continuous at the cut-off λ_c [Tho03].

Re-parameterization

Assume that we have a known principal direction $(\Delta\theta, \Delta\vartheta)$, in which the vertices in the neighborhood of (θ, ϑ) should move to improve the landmark correspondence. A Gaussian envelope function defined as:

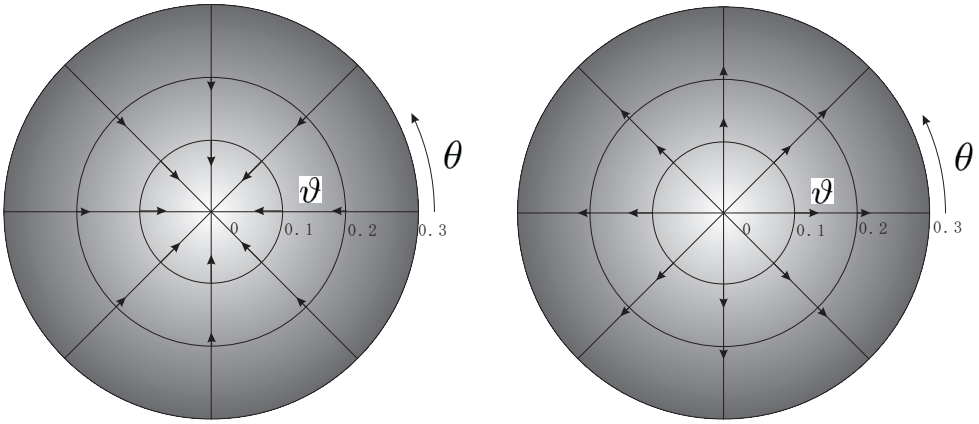
$$g(d, \sigma) = \begin{cases} \exp(\frac{-d^2}{2\sigma^2}) - \exp(\frac{-9d^2}{2\sigma^2}) & \text{for } \ d < 3\sigma \\ 0 & \text{for } \ d \geq 3\sigma \end{cases} \tag{7.9}$$

restricts the re-parameterization within a local neighborhood by the multiplication: $g(d, \sigma) \cdot (\Delta\theta, \Delta\vartheta)$, where d denotes the Euclidian distance between the current vertex to the center of the kernel, while σ defines the size of the kernel. In this way, the distant movements, outside of 3σ, are ignored. At the beginning large kernels with large σ are applied to optimize the large region. During the course of optimization, σ is decreased to optimize the details. Figures 7.4(a)-7.4(c) show three examples of kernels with different σ. One should note that this method works in the most regions of the (θ, ϑ) space, except for the poles ($\vartheta = 0$ or $\vartheta = \pi$). The reason is that all the vertices within the polar region move toward or away from the polar points, depending on the given $\Delta\vartheta$. Figures 7.5(a) and 7.5(b) present two examples of such phenomenon nearby the north polar point $(\vartheta \approx 0)$. To solve this problem, a random matrix is used to change the relative kernels. The details of this method are described in [Arv92].



(a) $\sigma = 0.4$ (b) $\sigma = 0.3$ (c) $\sigma = 0.2$

Figure 7.4: Kernel configuration for σ values of $0.4, 0.3$ and 0.2. Red colors mark regions allowing large movements. The graphics are taken from [Tho03].



(a) $\Delta\theta = 0, \Delta\vartheta < 0$ (b) $\Delta\theta = 0, \Delta\vartheta > 0$

Figure 7.5: Examples of movement of vertices in the pole region $\vartheta \approx 0$.

Calculating MDL Gradients

Now we need to compute the gradient $(\Delta\theta, \Delta\vartheta)$ in a certain position to minimize MDL cost function of the resulting shape model. The cost function in Equation 7.8 is actually defined with the eigenvalues of the covariance matrix $\boldsymbol{S}_c$ in Equation 7.4. The eigenvalues and eigenvector of $\boldsymbol{S}_c$ can be computed by a singular value decomposition (SVD) of a $3n_s \times n_r$ matrix

$$\boldsymbol{A} := \frac{1}{\sqrt{n_r - 1}}(\boldsymbol{L}_c - \overline{\boldsymbol{L}_c}), \tag{7.10}$$

where $\boldsymbol{L}_c$ is the landmark configuration matrix defined in Equation 7.2 and $\overline{\boldsymbol{L}_c}$ is a matrix with all columns set to $\overline{\boldsymbol{\varsigma}}$. The matrix $\boldsymbol{A}$ is an unbiased landmark configuration matrix, with respect to

the biased landmark configuration matrix $\boldsymbol{L}_c$. It is decomposed as:

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T, \tag{7.11}$$

where $\boldsymbol{U} \in \mathbb{R}^{3n_s \times n_r}$ and $\boldsymbol{V} \in \mathbb{R}^{n_r \times n_r}$ are two column-orthogonal matrices and the matrix $\boldsymbol{D} := \text{diag}(d_r) \in \mathbb{R}^{n_r \times n_r}$ is a diagonal matrix. According to the theory of SVD, the matrices $\boldsymbol{U}$ and $\boldsymbol{D}^2$ contain the corresponding eigenvectors and eigenvalues of the matrix $\boldsymbol{A}\boldsymbol{A}^T$. One should notice that an advantage of this method is that PCA of training data can be obtained without an explicit computation of covariance matrix.

The derivative of i-th singular value d_i with respect to the landmark matrix $\boldsymbol{A}$ is computed by:

$$\frac{\partial d_i}{\partial a_{jk}} = u_{ji} v_{ki} \tag{7.12}$$

Here a_{jk}, u_{ji} and v_{ki} are the single entries of the matrix $\boldsymbol{A}$, $\boldsymbol{U}$ and $\boldsymbol{V}$ respectively, while d_i is the i-th diagonal entry of the matrix $\boldsymbol{D}$. Since $\lambda_i = d_i^2$, the derivative of the i-th eigenvalue is computed as

$$\frac{\partial \lambda_i}{\partial a_{jk}} = \frac{\partial \lambda_i}{\partial d_i} \cdot \frac{\partial d_i}{\partial a_{jk}} = 2 d_i u_{ji} v_{ki} \tag{7.13}$$

Thus, the gradient of the cost function E_{MDL} in Equation 7.8 for every landmark can be calculated by

$$\frac{\partial E_{\text{MDL}}}{\partial a_{jk}} = \sum_i \frac{\partial L_i}{\partial a_{jk}} \text{ with } \frac{\partial L_i}{\partial a_{jk}} = \begin{cases} 2 u_{ji} v_{ki} / d_i & \text{for } \lambda_i \geq \lambda_c \\ 2 d_i u_{ji} v_{ki} / \lambda_c & \text{for } \lambda_i < \lambda_c. \end{cases} \tag{7.14}$$

The last step is the computation of the movement direction $(\Delta\theta, \Delta\vartheta)$ within the parameterization mesh. For $\boldsymbol{t} := (\theta, \vartheta)$, the gradient with respect to $\boldsymbol{t}$ is calculated by

$$\frac{\partial E_{\text{MDL}}}{\partial \boldsymbol{t}} = \frac{\partial E_{\text{MDL}}}{\partial a_{ij}} \cdot \frac{\partial a_{ij}}{\partial \boldsymbol{t}} \tag{7.15}$$

where the surface gradient $\partial a_{ij} / \partial \boldsymbol{t}$ can be approximated by the finite difference method. The resulting optimization algorithm works in a gradient decent fashion, i.e., in each iteration, the gradient is computed by Equation 7.15 to guide the re-parameterizations for every training surfaces. The step lengths are adapted, until there are no parameters left to tune. For the implementation

of the algorithm, we refer to [HWWM05].

7.1.3 Model Customization

In this subsection, we discuss how to customize a shape model of kidneys, according to the image data, in order to correctly segment a new kidney within the image. ASM based segmentation was sometimes called image search in literature. Given an instance of model $\boldsymbol{X} = \{\boldsymbol{\varsigma}_s\}_{s=1,...,n_s}$, this iterative algorithm proceeds as follows: First locate the current model in the image and examine the region around each model point $\boldsymbol{\varsigma}_s$ to find the best nearby matching point $\boldsymbol{\varsigma}_s^{'}$, then update the pose and shape parameters to best fit the set of matching points $\boldsymbol{X}^{'} = \{\boldsymbol{\varsigma}_s^{'}\}_{s=1,...,n_s}$. The procedure continues until no parameters change any more. In practice, the initial approximation does not need to be very close to the final solution, thus, the method can automatically segment the object on its own in most cases. Compared with conventional segmentation methods, a distinguishing feature of this method is that the models attempt to deform to better fit the data, but only in ways which are consistent with the shapes found in the training set [CTCG95].

Detection of Matching Points

Usually, the mean shape of a training set is placed in the image space as an initial estimate of the position of the model. For each model point, a matching point needs to be searched along the line normal to the surface of the model. A straightforward way to detect the matching point is to simply locate the strongest edge along the profile. However, the desired matching point is not always located on a strong edge. Especially for medical images, it happens quite often that weak edges represent the boundary of structures. A method of modelling the local structure has been developed in [CT99] to effectively solve this problem.

Based on a given model point, we sample the derivative along a profile with l pixel normals to the boundary on both sides among each training data. A vector $\boldsymbol{g}_r = (g_{r0}, ..., g_{r,2l})^T$ collects the $2l + 1$ derivative samples for this model point in the r-th training data. This vector is further normalized as follows

$$\boldsymbol{g}_r \leftarrow \frac{1}{\sum_{j=0}^{2l} |g_{rj}|} \boldsymbol{g}_r. \tag{7.16}$$

Let $\{\boldsymbol{g}_r\}_{r=1,...,n_r}$ denote the normalized samples vectors of this model points among n_r training dataset. Assume that the $\boldsymbol{g}_r$'s are samples of n_r-dimensional Gaussian distribution $\boldsymbol{g}$, whose mean and covariance are denoted by $\overline{\boldsymbol{g}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{g}}$. This Gaussian distribution gives a statistical

model of intensity patterns at these model points. Given a new profile vector $\boldsymbol{g}^{'}$, the quality of fit with the intensity pattern is measured by the Mahalanobis distance:

$$d_M(\boldsymbol{g}^{'}) = (\boldsymbol{g}^{'} - \overline{\boldsymbol{g}})^T \boldsymbol{\Sigma}_{\boldsymbol{g}}^{-1} (\boldsymbol{g}^{'} - \overline{\boldsymbol{g}}) \tag{7.17}$$

During the fitting of the model, a profile of $\widetilde{l}$ pixels on either side is sampled $(\widetilde{l} > l)$, then we search a position on the profiles that best fits the given intensity pattern, minimizing the distance in Equation 7.17. This process is repeated for every model point and a set of matching points is generated in the end.

In order to improve the efficiency of the algorithm, as well as to avoid being trapped in local minima, the searching of matching points is implemented in a multi-resolution framework [CT99]. A three-level volume pyramid is build for either a training example or a test volume. The base volume in the level 0 is the original one, while the volume on level 1 is smoothed and down-sampled with half the number of pixels in each dimension. The volume on level 2 is built in the same fashion based on the level 1 volume. The profiles of the model point have the same number of pixels in all levels. Obviously, a profile with a coarse resolution has a larger range in the physical space than the one with a fine resolution. Assume the search is performed in a pyramid of 3 levels. The search begins with level 2, where large movement is allowed and local minima are rejected. The searching procedure enters the next level when 90% of the matching points are stable. From the current matching points, the procedure restarts in the next level and continues searching on a finer resolution.

Computation of Pose Parameters and Shape Parameter

For given model points and corresponding points, first a rigid transformation needs to be estimated to match them. A function $\boldsymbol{T}_{\boldsymbol{R},\boldsymbol{t}}$ defines the transformation with a rotation matrix $\boldsymbol{R}$ and translation vector $\boldsymbol{t}$. The scaling effects are not considered, because both the model points and matching points are explicitly defined in the physical coordinates. The alignment involves the following three steps:

1. Compute the centers of gravity of model points and matching points, which are denoted as $\boldsymbol{c}_m$ and $\boldsymbol{c}_f$ respectively.

2. Translate all the model points, so that the centers of gravity of two sets of points overlap. The translation vector is $\boldsymbol{t} = \boldsymbol{c}_m - \boldsymbol{c}_f$.

3. Singular value decomposition is applied on the correlation matrix for obtaining the rotation

matrix [Kan94]. See the formulation in Equation 2.36.

Given the transformation $\boldsymbol{T_{R,t}}$, the matching points $\boldsymbol{X}^{'}$ are projected into the frame of the model coordinates by the inverse transformation $\boldsymbol{T_{R,t}^{-1}}$, i.e

$$\boldsymbol{X}^{'}_{\boldsymbol{T}^{-1}} = \boldsymbol{T_{R,t}^{-1}}(\boldsymbol{X}^{'}) \tag{7.18}$$

The shape parameters $\boldsymbol{b}$, defined in Equation 7.7, are computed to fit the resulting model points $\boldsymbol{X}$ and $\boldsymbol{X}^{'}_{\boldsymbol{T}^{-1}}$, i.e. minimizing the sum of square distance $\|\boldsymbol{X} - \boldsymbol{X}^{'}_{\boldsymbol{T}^{-1}}\|^2$. According to the Equation 7.7, the shape parameters are updated by

$$\boldsymbol{b} = \boldsymbol{P}_t^T(\boldsymbol{X}^{'}_{\boldsymbol{T}^{-1}} - \overline{\boldsymbol{\varsigma}}). \tag{7.19}$$

The estimation of post parameters and shape parameters is summarized in Algorithm 8.

Algorithm 8 Computation of post and shape parameters

given matching points $\boldsymbol{X}^{'}$
set shape parameter $\boldsymbol{b}$ to zero.
repeat
  Compute the current model instance: $\boldsymbol{\varsigma} = \overline{\boldsymbol{\varsigma}} + \boldsymbol{P}_t\boldsymbol{b}$
  Estimate the transform $\boldsymbol{T_{R,t}}$ from the model points to matching points.
  Transform the matching points into the model coordinates: $\boldsymbol{X}^{'}_{\boldsymbol{T}^{-1}} = \boldsymbol{T_{R,t}^{-1}}(\boldsymbol{X}^{'})$
  Compute the shape parameters: $\boldsymbol{b} = \boldsymbol{P}_t^T(\boldsymbol{X}^{'}_{\boldsymbol{T}^{-1}} - \overline{\boldsymbol{\varsigma}})$.
until $\boldsymbol{b}$ converges

7.2 Shape based Tracking for Kidneys in SPECT

This section introduces a method to localize the kidneys in a SPECT image based on a segmented kidneys in a corresponding CT image. As introduced in Chapter 5, the interest of nuclear medicine is the biodistribution of a radiolabeled substance – the radioactive tracer – that is determined by the body's physiological and biochemical functioning. Compared to morphological imaging, it is more difficult to determine the location of surfaces in functional imaging, because of the blurring of the edges, the sampling and the presence of noise [KLB91]. Currently, the boundary of an organ is determined by either thresholding methods or the maximum in the local gradient. See [PvAL$^+$93] and the reference therein. The segmentation of the kidneys in a CT dataset, which was introduced in the previous section,can be very useful to localize the kidneys in

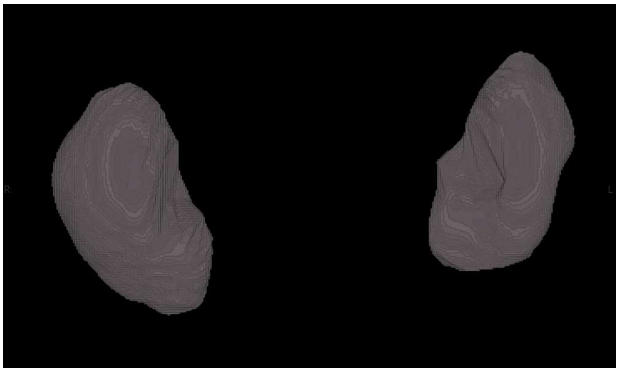


Figure 7.6: An example of a shape template of the kidney-pair.

the corresponding SPECT dataset. In other words, the task may be defined as tracking an object in the SPECT image, which has a similar shape as the CT kidney template. The prior knowledge of the size, shape and contrast of the organ has been used to determine the threshold for organ segmentation in SPECT images [LKS92, EWLE95]. But the method presented in the following makes use of the framework of rigid registration to estimate the pose parameters, so that the shape template is optimally aligned with SPECT image dataset. Analogously, segmentation of kidneys in CT, which includes the construction of the model and the customization of the model, the segmentation of the kidneys in SPECT consists of two sequential steps: the construction of the shape template and the template tracking of the object. They will be introduced in the rest of this section.

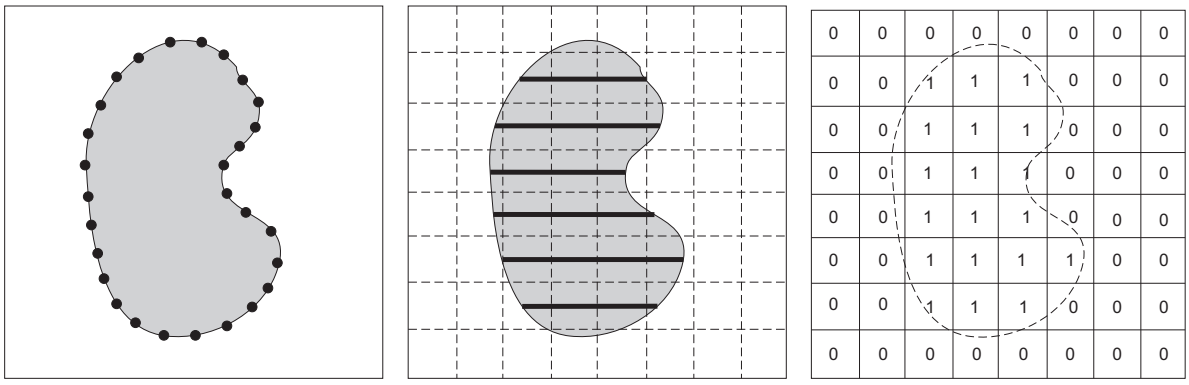
7.2.1 Construction of Shape Template

In Section 7.1, kidneys in CT segmented by the ASM method are represented by a set of continuous geometric primitives, which cannot be directly registered to SPECT volume data. An alternative to conventional geometric representation, the volumetric representation is an uniform, simple and robust description of measured objects and is very convenient to serve as the shape template in the registration frame. Figure 7.6 shows an example of such a shape template volume of a kidney-pair. The reformulation process that accomplishes the conversion from a set of continuous geometric primitives to an array of voxels in the 3-D discrete space is called voxelization. Many methods have been proposed to conduct voxelization. A short summary of voxelization can be found in [ZCB+04].

In this work, we use the implementation of from open source library ITK[1] to perform vox-

[1]Insight Segmentation and Registration Toolkit. http://www.itk.org/

elization of kidneys. The basic idea of voxelization is illustrated by the 2-D example in Figures 7.7(a)-7.7(c).



(a) (b) (c)

Figure 7.7: A 2-D example of solid voxelization. (a): The shape of an object is represented by a set of vertices and lines (geometric primitives).(b): Given the grid, the algorithm casts rays in either direction and computes the fragment that intersects with the segmented kidney. (c): The output image matrix is generated by assigning different values (e.g 1 and 0) to the voxels of inside and outside of intersections.

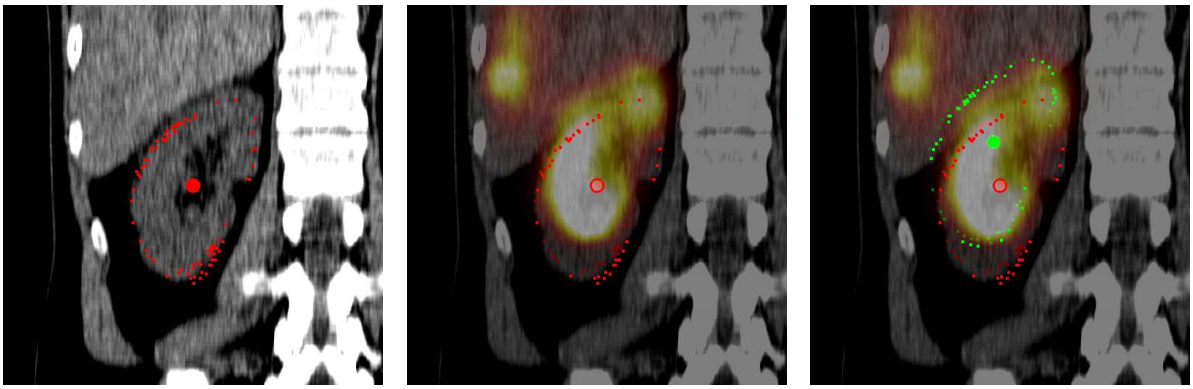
For the voxelization of the 3-D shape model of kidneys, some additional aspects of the implementation are considered. The elementary geometric primitive of the 3-D shape is the polygon or triangular. For reasons of efficiency, the algorithm does not really cast a ray for every element of the yz-matrix (assume the casting direction parallel to the x-axis), but traverses every polygon or triangular and records its intersection with the casting ray in the direction of x-axis. After all the pieces are processed, the surface intersections of every yz-element are known. For the complex shape, a casting ray may go through the object more than once, but the valid yz-element always has an even number of intersections. The output shape template volume has the same grid as the original CT volume. A high resolution template volume could reduce the error of voxelization. However, an extremely high resolution has little positive effect in practice, because the generated template volume will later be registered with the SPECT volume of a much lower resolution, normally $5 \sim 10$ times lower than the CT resolution.

7.2.2 Shape based Tracking

Conventional image based segmentation approaches perform poorly when segmenting organs in molecular images, because little contrast is present along boundaries and different parts of the same organ sometimes have different gray levels. Thus, the a-priori knowledge of the

shape of the underlying structures can be very useful. The basic idea of so-called "shape based segmentation", actually very similar to image registration, is to estimate a spatial transformation such that the transformed shape model is mapped to the underlying structure in the image. Various shape representations and transformation models have been investigated for different segmentation tasks. For more details on this topic, we refer to the recent representative works [TYW+03, Par03, PFK+05] and the references therein.

Since the SPECT/CT pair belongs to the same patient, there should be no great shape variance between kidneys in the two modalities. Therefore, if the kidneys have been correctly segmented in the CT volume, a rigid transformation can map the template with the SPECT kidneys. Due to this reason, we prefer to call the method "shape based tracking". A binary volume (introduced in Section 7.2.1) is used to represent the a-priori knowledge about the shape of the kidneys. Although the level-set function is possibly a more accurate shape representation, which has appeared in some works [Par03], we find out that it does not greatly improve the shape based segmentation of SPECT kidneys in this work.



(a) (b) (c)

Figure 7.8: An example of tracking the kidney from SPECT. (a) The kidney in CT is segmented by ASM. (b) An obvious misalignment exists between hybrid SPECT/CT. (c) The kidney in SPECT is detected by shape based tracking method.

Normalized cross correlation is chosen as the similarity measure of the registration. Let $u^R(\boldsymbol{x})$ and $u^T_{\boldsymbol{\phi}}(\boldsymbol{x})$ denote the fixed SPECT volume and the floating template volume transformed by $\boldsymbol{\phi}$, respectively. The normalized cross correlation between two volumes is defined by

$$\mathcal{S}_{\text{NCC}} = \frac{\int_\Omega u^R(\boldsymbol{x}) u^T_{\boldsymbol{\phi}}(\boldsymbol{x})\, d\boldsymbol{x}}{\sqrt{\int_\Omega (u^R(\boldsymbol{x}))^2\, d\boldsymbol{x} \int_\Omega (u^T_{\boldsymbol{\phi}}(\boldsymbol{x}))^2\, d\boldsymbol{x}}}. \tag{7.20}$$

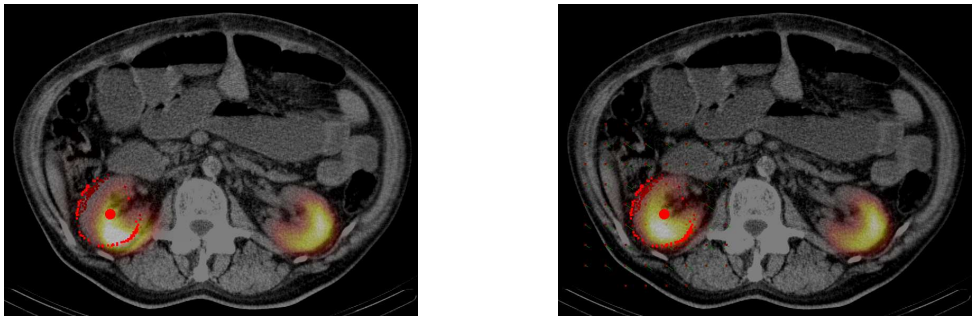
Gradient descent approach is applied for optimization. The step size of each iteration is auto-

matically computed according to Armijo's rule and the optimization stops when the estimated step size is too small (see the introduction on page 47). Like many registration systems, this registration work is also implemented in a multi-level fashion, i.e. at first images are registered in a low resolution then in a fine resolution, and the transformation estimated in a low resolution is the initial point of the optimization in finer resolutions. As defined in Section 7.1.1, the transformation is parameterized by the vector $\boldsymbol{q} = [\boldsymbol{q}_R|\boldsymbol{q}_T]^t$, where $\boldsymbol{q}_R$ is the unit quaternion and $\boldsymbol{q}_T$ is the translation vector. These seven parameters do not always have to be estimated at the same time in practice. Because the shape template is usually close to the kidneys in SPECT and the misalignment is dominated by shifting in most situations of validation, the search for parameters can be divided into two parts: First the quaternion vector $\boldsymbol{q}_R$ is fixed, estimating the translation vector $\boldsymbol{q}_T$ until convergence. Then search for optimal $\boldsymbol{q}_R$ with fixed $\boldsymbol{q}_T$. The optimization converges after several rounds, when both $\boldsymbol{q}_T$ and $\boldsymbol{q}_R$ are stable.

7.3 Fast Correction of Misalignment

In this section we present a method that can quickly correct the organ misalignments of the SPECT/CT hybrid imaging, see the example in Figure 7.9, in which the segmented kidney in the SPECT image is transformed as a rigid structure, while surrounding tissues are deformed elastically. In [LHH97] an efficient algorithm is proposed to solve this problem. The deformation with and without rigid constraints in Figures 7.10(a)-7.10(d) shows clearly the distinguishing properties of the method. It seems that the fast correction incorporating rigid structures can achieve more reasonable deformations than the pure non-rigid registration. Roughly speaking, the method is a kind of extension of landmark based deformable registration (introduced on page 26). The basis functions combined with a weight function are formulated to handle the rigid constraints. This method computes the transformation not iteratively but analytically. In other words, once the validation result is at hand, the transformation can be computed immediately.

In the following, the mathematical framework and implementation details of the fast correction method will be presented. First, we integrate single rigid transformation constraint into this framework in Subsection 7.3.1. Then, the framework is extended in Subsection 7.3.2 to allow multiple rigid objects movement. The mathematical notations in this section are consistent with those of point based registration framework introduced in page 26.



(a) (b)

Figure 7.9: Automatic correction of misalignment of organs: (a) Surface of right kidney has been marked by red points. The corresponding kidney in SPECT is measured by the method introduced in Chapter 7 about 8 mm deviated from its true position. (b) The kidney in SPECT is transformed by a deformation field (marked by the green arrows) to achieve a better fusion in the local region. The deformation defines the kidney to be transformed as a rigid object.

7.3.1 Deformation with Single Rigid Constraint

The difficulty lies in the integration of a rigid transform into the framework of point based deformation. The desired deformation should transform a part of the image with a pre-defined rigid transformation, while in the rest of the image, the deformation is controlled by landmarks. Actually the point based deformation in Equation 2.38 is naturally divided into a linear part and a non-linear (radius base) part. The coefficient matrix $\boldsymbol{A}$ in Equation 2.40 controls the linear parts of the transformation. The basic idea is to use the matrix $\boldsymbol{A}$ to control the linear transformation in the pre-selected region and modify the radius basis function of the nonlinear part, such that in the pre-selected regions the radius basis function tends towards zero.

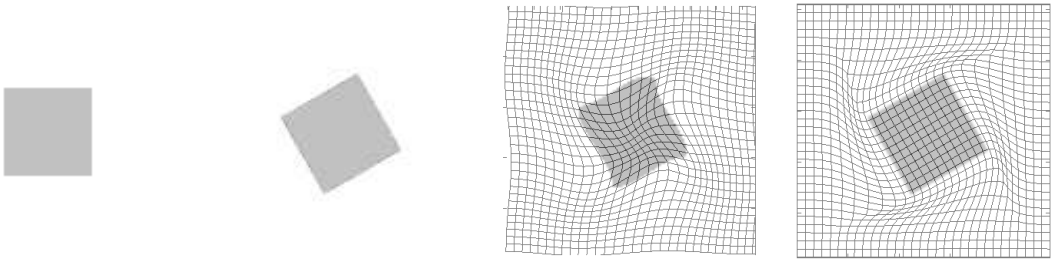
Assume that we have only one rigid object called $\mathcal{G}$, whose transformation can be represented using the linear basis function in Equation 2.39 as

$$\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{G}(\boldsymbol{x})\boldsymbol{A}. \tag{7.21}$$

Let a distance function $d_{\mathcal{G}}$ be defined as

$$d_{\mathcal{G}}(\boldsymbol{x}) = \begin{cases} 0 & \text{if } \boldsymbol{x} \in \mathcal{G} \\ \text{distance from } \boldsymbol{x} \text{ to } \mathcal{G} & \text{elsewhere.} \end{cases} \tag{7.22}$$

In this work, a linear-time Algorithm [FH04] is employed to compute the distance. The radius



(a) Template (b) Reference (c) Non-Rigid (d) Rigid Constrained

Figure 7.10: Comparison of deformations with and without rigid constraints. The template image, a gray square in the center of the image, is registered to the reference image, the gray square rotated in 30° with respect to the template square. The curvature regularized non-rigid registration, although it nearly perfectly registers them, ignores the desired underlying rigid rotation. The deformation with rigid constraints treats the objects and surrounding separately and the rigid movement is seamlessly merged into the deformable neighborhood.

basis function is weighted with the distance function: For one point $\boldsymbol{x}$ and one landmark $\boldsymbol{t}_s$, it is defined by

$$\sigma_{\mathcal{G}}(\boldsymbol{x}, \boldsymbol{t}_s) = d_{\mathcal{G}}(\boldsymbol{x}) d_{\mathcal{G}}(\boldsymbol{t}_s) \sigma(\boldsymbol{x}, \boldsymbol{t}_s). \tag{7.23}$$

The new radius basis function $\sigma_{\mathcal{G}}(\boldsymbol{x}, \boldsymbol{t}_s)$ tends to zero as the point $\boldsymbol{x}$ tends to the rigid object $\mathcal{G}$. The resulting matrix $\boldsymbol{\Sigma}^{\mathcal{G}}$ where $\boldsymbol{\Sigma}^{\mathcal{G}}_{ij} = \sigma_{\mathcal{G}}(\boldsymbol{t}_i, \boldsymbol{t}_j)$ can be decomposed as

$$\boldsymbol{\Sigma}^{\mathcal{G}} = \boldsymbol{D}^{\mathcal{G}} \boldsymbol{\Sigma} \boldsymbol{D}^{\mathcal{G}}, \tag{7.24}$$

and

$$\boldsymbol{D}^{\mathcal{G}}_{ij} = \begin{cases} d_{\mathcal{G}}(\boldsymbol{t}_i) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{7.25}$$

Matrix $\boldsymbol{D}^{\mathcal{G}}$ is diagonal and invertible if no landmarks are defined on the rigid object $\mathcal{G}$. The top equation of 2.40 is rewritten as

$$\boldsymbol{\Sigma}^{\mathcal{G}} \boldsymbol{B}^{\mathcal{G}} + \boldsymbol{G}\boldsymbol{A} = \boldsymbol{U} \tag{7.26}$$

Rearranging the equation, the unknown parameter matrix is computed by

$$\boldsymbol{B}^{\mathcal{G}} = (\boldsymbol{\Sigma}^{\mathcal{G}})^{-1}(\boldsymbol{U} - \boldsymbol{G}\boldsymbol{A}) \tag{7.27}$$

Thus the interpolation solution with one rigid object for $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ is defined in the following manners,

$$\boldsymbol{x} \rightarrow \boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{A}^T \boldsymbol{g}(\boldsymbol{x}) + \sum_{s=1}^{n_s} (\boldsymbol{B}^{\mathcal{G}})_s^T \sigma^{\mathcal{G}}(\boldsymbol{x}, \boldsymbol{t}_s) \tag{7.28}$$

7.3.2 Deformation with Multiple Rigid Constraints

Now we consider a more general situation, where the desired deformation transforms several objects with their pre-defined rigid transformations, while in the rest of the image, the deformation is still controlled by landmark based interpolation. With respect to the deformation with single rigid constraints, the major difference is the definition of the linear part of the transformation, i.e. the matrix $\boldsymbol{A}^T \boldsymbol{g}(\boldsymbol{x})$ in Equation 7.28, respectively.

In the following, we use o as index of rigid objects. Assume that we have n_o objects, named as $\mathcal{G}_o, o = 1, ..., n_o$. The objects may have any shape, but cannot be overlapping. If $\mathcal{G}$ denotes a union of all objects, i.e $\mathcal{G} = \mathcal{G}_1 \cup \mathcal{G}_2 \cup ... \cup \mathcal{G}_{n_o}$, the definition of the distance map is the same as in Equation 7.22. Let $\boldsymbol{l}_o = (l_1^o, l_2^o, l_3^o, l_4^o)^T, o = 1, ..., n_o$ denote the coefficient vector of the linear transformations associated with object $\mathcal{G}_o$, i.e. $\boldsymbol{\phi}_o(\boldsymbol{x}) = l_1^o g_1(\boldsymbol{x}) + l_2^o g_2(\boldsymbol{x}) + l_3^o g_3(\boldsymbol{x}) + l_4^o g_4(\boldsymbol{x})$.

In order to represent n_o linear transformations in a single matrix, as $\boldsymbol{GA}$ in Equation 7.26, a kind of inverse distance weighted interpolation [She68] is used. The linear term is defined by a weighted sum of each linear transformation. The weight is defined by

$$w_i(\boldsymbol{x}) = \frac{q_i(\boldsymbol{x})}{\sum_{o=1}^{n_o} q_o(\boldsymbol{x})} \text{ where } q_o(\boldsymbol{x}) = \frac{1}{d_o^{\mathcal{G}}(\boldsymbol{x})^\mu} \text{ and } i = 1, ..., n_o. \tag{7.29}$$

Here $d_o^{\mathcal{G}}(\boldsymbol{x})$ denotes the distance map (defined by Equation 7.22) with respect to the object $\mathcal{G}_o$. The weight is normalized, i.e $\sum_{i=1}^{n_o} w_i(\boldsymbol{x}) = 1$ for all $\boldsymbol{x}$. The weight works only among the associated object, i.e

$$w_i(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \boldsymbol{x} \in \mathcal{G}_i \\ 0 & \text{if } \boldsymbol{x} \in \mathcal{G}_j, j = 1, ..., n_o, j \neq i. \end{cases} \tag{7.30}$$

Smoothness of the interpolation is determined by the parameter $\mu = 1.5$ in this work. Thus the overall linear transformation is defined as

$$\boldsymbol{L}_w(\boldsymbol{t}) = \sum_{i=1}^{n_o} w_i(\boldsymbol{t}) \boldsymbol{l}_i. \tag{7.31}$$

The matrix of linear transformation $\boldsymbol{GA}$ in Equation 7.26 is replaced by a new matrix $\boldsymbol{T}$, which is defined by

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{g}(\boldsymbol{t}_1)^T \boldsymbol{L}_w(\boldsymbol{t}_1) \\ \boldsymbol{g}(\boldsymbol{t}_2)^T \boldsymbol{L}_w(\boldsymbol{t}_2) \\ \vdots \\ \boldsymbol{g}(\boldsymbol{t}_n)^T \boldsymbol{L}_w(\boldsymbol{t}_{n_s}) \end{pmatrix}. \tag{7.32}$$

and

$$\boldsymbol{B}^{\mathcal{G}} = (\boldsymbol{\Sigma}^{\mathcal{G}})^{-1}(\boldsymbol{U} - \boldsymbol{T}). \tag{7.33}$$

The interpolation solution with multiple rigid objects for $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ is defined in the following manners,

$$\boldsymbol{x} \rightarrow \boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{L}_w(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{x}) + \sum_{s=1}^{n_s} (\boldsymbol{B}^{\mathcal{G}})_s^T \sigma^{\mathcal{G}}(\boldsymbol{x}, \boldsymbol{t}_s). \tag{7.34}$$

7.4 Experiments

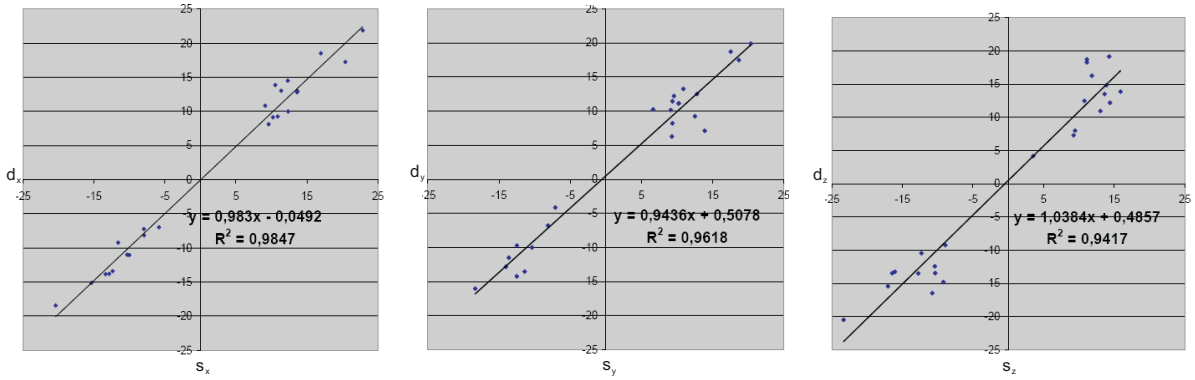
7.4.1 Patient Datasets

To evaluate the kidney based validation, we selected 24 patient datasets, where the kidneys are clearly visible on both SPECT and CT images. Among these 24 patients, there are 10 female and 14 male between 15-78 years old and the average age is 55.21. All of them are abdomen studies generated by a SPECT/Spiral CT scanner of the University of Erlangen, in which 9 patients were examined between November 2006 and March 2007, the rest were examined between August and November 2008.

7.4.2 Accuracy Test

The accuracy of the validation method is measured in the same way as in Section 6.3.3. The right kidneys in the data-sets were used as validation objects and the shift parameters (s_x, s_y, s_z) were randomly generated between 7 mm and 23 mm or -7 mm and -23 mm. As shown in Figures 7.11(a)-7.11(c), the experiment yielded a clear linear dependency between the ground truth shift and the measured shift: The correlation coefficients are 0.9923, 0.9807 and 0.9704 in X-, Y- and Z-directions, respectively. The anatomical inaccuracies, measured by mean $\pm$

standard deviation of the absolute error, were 1.3979 ± 0.8401 mm in X-direction, 1.9992 ± 1.3920 mm in Y-direction and 2.7823 ± 2.0672 mm in Z-direction, respectively.



(a) s_x-d_x (b) s_y-d_y (c) s_z-d_z

Figure 7.11: Comparison of the ground truth shift (s_x, s_y, s_z) and the measured shift (d_x, d_y, d_z). Two shift parameters in X-, Y- and Z-directions are close to lines of identity.

Chapter 8

Conclusion

In this thesis, medical image registration and quantitative validation methods have been explored. Firstly, reviews of state-of-art of registration methods were given. Then, two novel methods were introduced, namely, an original non-rigid edge based registration method and a new validation method for hybrid imaging by automatic segmentations of markers. In the end of this thesis, we summarize the methods, present their contributions and state the the future works.

One-to-one Edge Based Registration

This new method makes use of the Mumford–Shah model to simultaneously detect the edge features of two images and jointly estimate a consistent set of transformations to match them. Compared to the current asymmetric methods in the literature, this fully symmetric method allows one to determine one-to-one correspondences between the edge features of two images. The entire variational model is realized in a multi-scale framework of the Finite Element approximation. The optimization process is guided by an Estimation-Minimization type algorithm and an adaptive generalized gradient flow to guarantee a fast and smooth relaxation.

The algorithm is tested on T1- and T2-weighted MR datasets to study the parameter setting. We also present promising results of four applications of the proposed algorithm: inter-object mono-modal registration, retina image registration, matching digital photographs of neurosurgery with its volume data and motion estimation for frame interpolation.

One-to-one edge based registration is a general framework, which provides enough flexibilities to adapt to different registration applications. At the moment, users have to spend a lot of effects to find the optimal parameter setting or adapt the implementation for each application. In front of us, there is still a lot of work to customize, or in some sense to simplify, this general method to specific application. We also hope this method can be widely accepted by the others,

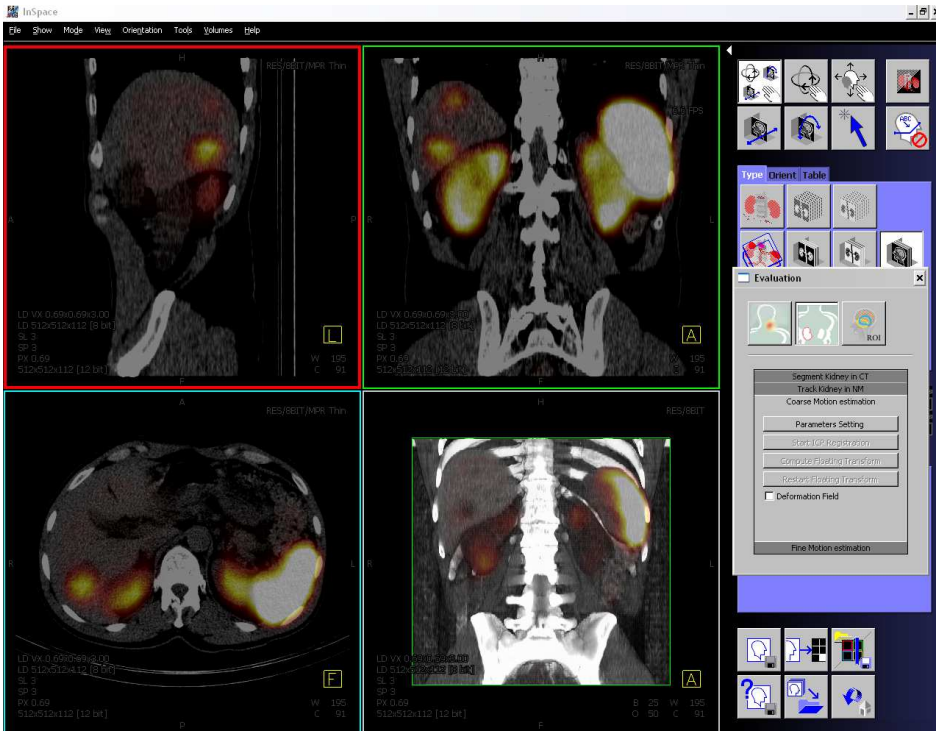


Figure 8.1: User interface of validation software.

especially for the non-rigid registration of medical image datasets.

Automatic Segmentations for Validation of Hybrid Imaging

This validation method measures the anatomical accuracy of the hybrid imaging by the computation of the distance between the segmented corresponding markers in both modalities. Two kinds of objects are used as markers in this work for validation: hot spots and kidneys. Both of them can be segmented in hybrid modalities with minimum user interaction.

The experimental results on the clinical data sets - 21 patients for hot-spot-marker and 24 patients for kidney-marker - show that the measurement of this validation tool is sufficiently accurate and reproducible for clinical datasets. According to the testing results, the inaccuracies of hot spot based validation for neck regions are 0.7189 ± 0.6298 mm in X-direction, 0.9250 ± 0.4535 mm in Y-direction and 0.9544 ± 0.6981 mm in Z-direction. The inaccuracies of kidneys based validation for abdomen regions are 1.3979 ± 0.8401 mm in X-direction, 1.9992 ± 1.3920 mm in Y-direction and 2.7823 ± 2.0672 mm in Z-direction. The tool implemented in a plugin of InSpace software platform has been used by Clinic of Nuclear Medicine, Friedrich-Alexander

University of Erlangen, for the purpose of research. See the user interface of the software in Figure 8.1. Although we only tested the method with the data generated by the Siemens Symbia SPECT/spiral CT system in this work, the tool and the same principle can be also applied on the other SPECT/CT systems.

There is still a long way to go, and this thesis has only solved a small part of the validation problems. More potential validation markers, e.g. liver and bladder, need to investigated. Appropriate segmentation methods need to be developed for new markers. The concept of segmentation based validation can be easily extended to PET/CT hybrid imaging. In our future work, we plan to apply this validation tool to analyze the variation of the accuracy of hybrid scanners with respect to different patient positions, tracers or acquisition protocols. Since the tool has been installed and used in clinics, we expect that more and more medical studies using this tool would appear.

In Section 7.3, we introduce a fast method to remove the misalignments of SPECT/CT images. It constraints specified misaligned objects to be rigidly transformed, while in the meantime in the other regions non-rigid deformations are computed. In this thesis, the mathematical definition of the method was presented and its effects need to be further validated by testing on more patient datasets. However, This method presents a new segmentation based registration method and could have a huge mount of applications in future.

Bibliography

[AAF99] J. Ashburner, J. Andersson, and K. J. Friston. High-dimensional nonlinear image registration using symmetric priors. *NeuroImage*, 9(6):619–628, 1999.

[AHB87] K. S. Arun, T. S. Huang, and S.D. Blostein. Least square fitting of two 3-d point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):698 – 700, 1987.

[Alt86] S. L. Altmann. *Rotations, Quaternions, and Double Groups*. Oxford Science Publications, Oxford, England, 1986.

[Arv92] J. Arvo. *Fast random rotation matrices*, pages 117–120. Academic Press Professional, Inc., New York, USA, 1992.

[AT90] L. Ambrosio and V. M. Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Communications on Pure and Applied Mathematics*, 43:999–1036, 1990.

[BAO06] T. Bunyaviroch, A. Aggarwal, and M. E. Oates. Optimized scintigraphic evaluation of infection and inflammation: Role of single-photon emission computed tomography/computed tomography fusion imaging. *Seminars in Nuclear Medicine*, 36(4):295–311, 2006.

[BHM00] W. Briggs, V. E. Henson, and S. F. McCormick. *A Multigrid Tutorial*. Society for Industrial and Applied Math, Philadelphia, USA, 2nd edition, 2000.

[BM92] P. J. Besl and N. D. Mckay. A method for registration of 3-d shapes. *IEEE Transation on Pattern Analysis and Machine Intellegence*, 14(2):239–256, 1992.

[BM95] A. Brandt and V. Mikulinsky. On recombining iterants in multigrid algorithms and problems with small islands. *SIAM Journal on Scientific Computing*, 16(1):20–28, 1995.

[BNG96] M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. In *Lecture Notes In Computer Science; Vol. 1131, Proceedings of the 4th International Conference on Visualization in Biomedical Computing*, pages 267–276, Londom, UK, 1996. Springer-Verlag.

[Boo89] F. Bookstein. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[Bor88] G. Borgefors. Hierarchical chamfer matching: a parametric edge matching algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6):849–865, 1988.

[Bro81] C. Broit. *Optimal registration of deformed images*. Ph.d thesis, University of Pensylvania, 1981.

[Bro92] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.

[BSDSZ03] C. Bai, L. Shao, A. J. Da-Silva, and Z. Zhao. A generalized model for the conversion from ct numbers to linear attenuation coefficients. *IEEE Transaction on Nuclear Science*, 50(5):1510–1514, 2003.

[Buh03] M. D Buhmann. *Radial basis functions theory and implementations*. Cambridge University Press, July 2003.

[CDR02] U. Clarenz, M. Droske, and M. Rumpf. Towards fast non–rigid registration. In *Inverse Problems, Image Analysis and Medical Imaging, AMS Special Session Interaction of Inverse Problems and Image Analysis*, volume 313, pages 67–84. AMS, 2002.

[Chr94] G. E. Christensen. *Deformable shape models for anatomy*. Ph.d thesis, Sever Institue of Technology,Washington University, 1994.

[CHR02] U. Clarenz, S. Henn, and K. Rumpf, M. Witsch. Relations between optimization and gradient flow methods with applications to image registration. In *Proceedings of the 18th GAMM Seminar Leipzig on Multigrid and Related Methods for Optimisation Problems*, pages 11–30, 2002.

[CJ01] G. E. Christensen and H. J. Johnson. Consistent image registration. *IEEE Transactions on Medical Imaging*, 20(7):568–582, 2001.

[CJM97] G. E. Christensen, S. C. Joshi, and M. I. Miller. Volumetric transformation of brain anatomy. *IEEE Transactions on Medical Imaging*, 16(6):864–877, 1997.

[CKR05] I. Christadler, H. Köstler, and U. Rüde. Robust and efficient multigrid techniques for the optical flow problem using different regularizers. In *Proceedings of 18th symposium simulationstechnique ASIM 2005*, volume 15, pages 341–346, Erlangen, Germany, September 2005.

[CLG05] M. Couprie, L. Najman, and G. Bertrand. Quasi-linear algorithms for the topological watershed. *Journal of Mathematical Imaging and Vision*, 22(2-3):231–249, 2005.

[CM92] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, 3(10):145–155, 1992.

[CSM+06] I. A. Carreras, C. Sorzano, R. Marabini, J. Carazo, C. O. De Solorzano, and J. Kybic. Consistent and elastic registration of histological sections using vector-spline regularization. In *Computer Vision Approaches to Medical Image Analysis*, pages 85–95. 2006.

[CSRT02a] A. Can, C. V. Stewart, B. Roysam, and H. L. Tanenbaum. A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):347–364, 2002.

[CSRT02b] A. Can, C. V. Stewart, B. Roysam, and H. L. Tanenbaum. A feature-based technique for joint, linear estimation of high-order image-to-mosaic transformations: mosaicing the curved human retina. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):412–419, 2002.

[CSV00] T. F. Chan, B. Y. Sandberg, and L. A. Vese. Active contours without edges for vector–valued images. *Journal of Visual Communication and Image Representation*, 11:130–141, 2000.

[CT99] T. F. Cootes and C. Taylor. Statistical models of appearance for computer vision. Technical report, University of Manchester, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, Manchester M13 9PT, UK, 1999.

[CTCG95] 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.

[CV01] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[CW98] Y. Chen and L. Wu. *Second order elliptic equations and elliptic systems*, volume 174 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1998. Translated from the 1991 Chinese original by Bei Hu.

[DB06] Michael Donoser and Horst Bischof. 3d segmentation by maximally stable volumes (MSVs). In *Proceeding of international conference on pattern recogition (ICPR)*, pages 63–66, Hong Kong, China, August 2006.

[DCT01] R. H. Davies, T. F. Cootes, and C. J. Taylor. A minimum description length approach to statistical shape modelling. In *Proceedings of the 17th International Conference on Information Processing in Medical Imaging*, pages 50 – 63, London, UK, June 2001. Spring-Verlag.

[DR04] M. Droske and M. Rumpf. A variational approach to non-rigid morphological registration. *SIAM Journal on Scientific Computing*, 64(2):668–687, 2004.

[DR06] M. Droske and W. Ring. A Mumford-Shah level-set approach for geometric image registration. *SIAM Journal on Applied Mathematics*, 66(6):2127–2148, 2006.

[Dro05] M. Droske. *On Variational Problems and Gradient Flows in Image Processing*. Ph.d thesis, University Duisburg, 2005.

[DRR09] M Droske, W. Ring, and M. Rumpf. Mumford-Shah based registration: a comparison of a level set and a phase field approach. *Computing and Visualization in Science*, 12(3):101–114, 2009.

[DTC+02] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor. A minimum description length approach to statistical shape modeling. *IEEE Transactions on Medical Imaging*, 21(5):525 – 537, 2002.

[DZLGH06] Y. Deuerling-Zheng, M. Lell, A. Galant, and J. Hornegger. Motion compensation in digital subtraction angiography using graphics hardware. *Computerized medical imaging and graphics*, 30(5):279–289, 2006.

[EWLE95] Y. E. Erdi, B. W. Wessels, M. H. Loew, and A. K. Erdi. Threshold estimation in single photon emission computed tomography (SPECT) and planar imaging for clinical radioimmunotherapy. *Cancer Research*, 55(23 Suppl.):5823–5826, 1995.

[FBTC07] R. Fabbri, O. M. Bruno, J. C. Torelli, and L. F. Costa. 2d Euclidean distance transform algorithms: A comparative survey. *ACM Computing Surveys*, 40(1):1–44, 2007.

[FH04] P. F. Felzenszwalb and D. P. Huttenlocher. Distance transforms of sampled functions. Technical report, Cornell Computing and Information Science Technical Report, 2004.

[FH05] M. S. Floater and K. Hormann. Surface parameterization: a tutorial and survey. In N. A. Dodgson, M. S. Floater, and M. A. Sabin, editors, *Advances in multiresolution for geometric modelling*, pages 157–186. Springer Verlag, 2005.

[FM99] B. Fischer and J. Modersitzki. Fast inversion of matrices arising in image processing. *Numerical Algorithms*, 22(1):1–11, 1999.

[FM04] B. Fischer and J. Modersizki. A unified approach to fast image registration and a new curvature based registration technique. *Linear Algebara and its Applications*, 380:107–124, 2004.

[Fri09] M. Fried. Multichannel image segmentation using adaptive finite elements. *Computing and Visualization in Science*, 12(3):125–135, 2009.

[GDMK76] M. L. Goris, S. G. Daspit, P. McLaughlin, and J. P. Kriss. Interpolative background subtraction. *Journal of Nuclear Medicine*, 17(8):744–747, 1976.

[Gra06] L. Grady. Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783, 2006.

[Gur81] M. E. Gurtin. *An Introduction to Continuum Mechanics*. Academic Press, Orlando, Florida, 1981.

[HBD+07] J. Han, B. Berkels, M. Droske, J. Hornegger, M. Rumpf, C. Schaller, J. Scorzin, and H. Urbach. Mumford-shah model for one-to-one edge matching. *IEEE Transation on Image Processing*, 16(11), 2007.

[HBR+06] J. Han, B. Berkels, M. Rumpf, J. Hornegger, M. Droske, M. Fried, J. Scorzin, and C. Schaller. A Variational Framework for Joint Image Registration, Denoising and Edge Detection . In H. Handels, J. Ehrhardt, A. Horsch, H. Meinzer, and T. Tolxdorff, editors, *Bildverarbeitung für die Medizin 2006*, pages 246–250, Hamburg, 2006. Springer.

[Her02] G. Hermosillo. *Variational Methods for Multi-modal Image Matching*. PhD thesis, Université de Nice, France, 2002.

[HHN88] B. K. P. Horn, H. M. Hilden, and Sh. Negahdaripour. Closed-form solution of absolute orientation using orthonormal matrices. *Journal of the Optical Society of America*, 5(7):1127 – 1135, 1988.

[HKC+08] J. Han, H. Köstler, Bennewitz C., Kuwert T., and Hornegger J. Computer-Aided Evaluation of Anatomical Accuracy of Image Fusion between X-Ray CT and SPECT. *Computerized Medical Imaging and Graphics*, 32(5):388–395, 2008.

[Hor87] B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America*, 4(4):629 – 642, 1987.

[HR06] M. Horger and R. Bares. The role of single-photon emission computed tomography/computed tomography in benign and malignant bone disease. *Seminars in Nuclear Medicine*, 36(4):286–294, 2006.

[HS52] Magnus R. Hestenes and Eduard Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6):409–436, 1952.

[HWM06] T. Heimann, I. Wolf, and H. P. Meinzer. Optimal landmark distributions for statistical shape model construction. In *The Proceedings of SPIE, Medical Imaging 2006, Image Processing*, volume 6144, San Diego, CA, USA, March 2006.

[HWWM05] T. Heimann, I. Wolf, T. G. Williams, and H. P. Meinzer. 3d active shape models using gradient descent optimization of description length. In *The Proceedings of Information Processing in Medical Imaging*, pages 566–577, Colorado, USA, July 2005.

[JC02] H. J. Johnson and G. E. Christensen. Consistent landmark and intensity-based image registration. *IEEE Transactions on Medical Imaging*, 21(5):450–461, 2002.

[Kan94] K. Kanatani. Analysis of 3-d rotation fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(5):543–549, 1994.

[KBS04] H. A. Karim, M. Bister, and M. U. Siddiqi. Multiresolution motion estimation for low-rate video frame interpolation. *Journal on Applied Signal Processing*, 11:1708–1720, 2004.

[KJ04] L. Kubecka and J. Jan. Registration of bimodal retinal images - improving modifications. In *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2004)*, volume 3, pages 1695– 1698, 2004.

[KKR07] E. M. Kalmoun, H. Köstler and U. Rüde. 3D optical flow computation using a parallel variational multigrid scheme with application to cardiac C-arm CT motion. *Image and vision computing*, 25(9):1482–1494, 2007.

[KLB91] M. A. King, D. T. Long, and A. B. Brill. SPECT volume quantitation: influence of spatial resolution, source size and shape, and voxel size. *Medical Physics*, 18(5):1016–1024, 1991.

[KO07] Y. Krausz and O. Israel. Single-photon emission computed tomography/computed tomography in endocrinology. *Seminars in Nuclear Medicine*, 36(4):267–274, 2007.

[Kos91] P. Kosmol. *Optimierung und Approximation*. de Gruyter Lehrbuch, Berlin, Germany, 1991.

[KRH07] T. Kuwert, W. Römer, and J. Hornegger. Correlative imaging in cardiology. principles and clinical application. *Herz*, 32(2):122–128, 2007.

[KSR07] H. Köstler, M. Stürmer, and U. Rüde. A fast full multigrid solver for applications in image processing. *Numerical Linear Algebra with Applications*, 15(2-3):187 – 200, 2007.

[KT98] A. C. W. Kotcheff and C. J. Taylor. Automatic construction of eigenshape models by direct optimization. *Medical Image Analysis*, 2(4):303–314, 1998.

[KWM99] R. Krishnamurthy, J. W. Woods, and P. Moulin. Frame interpolation and bidirectional prediction of video usingcompactly encoded optical-flow fields and label fields. *IEEE Transactions on Circuits and Systems for Video Technology*, 9:713–726, 1999.

[KYZ01] T. Kapur, L. Yezzi, and L. Zöllei. A variational framework for joint segmentation and registration. *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (IEEE CVPR - MMBIA)*, pages 44–51, 2001.

[LEF] A. Lorusso, D. Eggert, and R. Fisher. A comparison of four algorithms for estimating 3-d rigid transformations. In *Proceedings of the 4th British Machine Vision Conference (BMVC '95)*, pages 237 – 246.

[LHH97] J. A. Little, D. L. G. Hill, and D. J. Hawkes. Deformations incorporating rigid structures. *Computer vision and image understanding*, 66(2):223–232, 1997.

[LKS92] D. T. Long, M. A. King, and J. Sheehan. Comparative evaluation of image segmentation methods for volume quantitation in spect. *Medical Physics*, 19(2):483–489, 1992.

[LL06] S.J. Liu and J. Li. Genus-zero shape classification using spherical normal image. In *ICPR 2006. 18th International Conference on Pattern Recognition*, volume 2, pages 126–129, 2006.

[Mar92] R. March. Visual reconstructions with discontinuities using variational methods. *Image and Vision Computing*, 10(1):30–38, 1992.

[MC03] M. Moelich and T. Chan. Joint segmentation and registration using logic models. Technical report, Y.-N. YOUNG AND D. LEVY, Technical Report, 2003.

[MCOS+02] T. Makela, P. Clarysse, N. Pauna O. Sipilä, Q. C. Pham, T. Katila, and I. E. Magnin. A review of cardiac image registration methods. *IEEE Transactions on Medical Imaging*, 21(9):1011–1021, 2002.

[MCUP02] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *Proceedings of the british machine vision conference (BMVC)*, volume 1, pages 384–393, Cardiff, UK, September 2002.

[MF02] J. Modersitzki and B. Fischer. Fast diffusion registration. In M. Z. Nashed and O. Scherzer, editors, *AMS Special Session on Interactions of Inverse Problems and Image Analysis: Inverse Problems, Image Analysis and Medical Imaging*, pages 117–127. AMS, 2002.

[Mic86] C. A. Micchelli. Interpolation of scattered data: Distance matrices and conditionally positive definite functions. *Constructive Approximation*, 2(1):11–22, 1986.

[MJG+86] L. Mortelmans, J. Nuyts, G. Van Pamel, V. van den Maegdenbergh, M. De Roo, and P. Suetens. A new thresholding method for volume determination by SPECT. *European Journal of Nuclear Medicine and Molecular Imaging*, 12(5-6):284–290, 1986.

[MKGL96] K.V Mardia, J. T. Kent, C. R. Goodall, and J. A. Little. Kriging and splines with derivative information. *Biometrika*, 85(2):207–221, 1996.

[Mod04] J. Modersitzki. *Numerical Methods for Image Registration*. Oxford University Press, Oxford, 2004.

[MS89] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42:577–685, 1989.

[MS95] J. M. Morel and S. Solimini. *Variational methods in image segmentation*. Birkhauser Boston Inc., Cambridge, MA, USA, 1995.

[MTT03] S. Marsland, C.J. Twining, and C.J. Taylor. Groupwise non-rigid registration using polyharmonic clamped-plate splines. In Randy E. Ellis and Terry M. Peters, editors, *Sixth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCIA'03)*, LNCS, pages 246–250, Montreal, 2003. Springer Verlag.

[MV97] J. B. A Maintz and M. A. Viergever. An overview of medical image registration methods. *Symposium of the Belgian hospital physicists association(SBPH/BVXF)*, 12(5):1–12, 1996/1997.

[NM02] Y. Nie and K.-K. Ma. Adaptive rood pattern search for fast block-matching motion estimation. *IEEE Transactions on Image Processing*, 11(12):1442 – 1449, 2002.

[NWD+06] A. Neomayr, W. Römer, D. Strobel, W. Bautz, and T. Kuwert. Anatomical accuracy of hybrid SPECT/spiral CT in the lower spine. *Nuclear Medicine Communications*, 27(6):521–528, 2006.

[Ots79] N. Otsu. A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man, and Cybernetic*, 9(1):62–66, 1979.

[Par03] N. Paragios. A level set approach for shape-driven segmentation and tracking of the left ventricle. *IEEE Transactions on Medical Imaging*, 22(6):773 – 776, 2003.

[PF03] J. P. W. Pluim and J. M. Fitzpatrick. Image registration. *IEEE Transaction on Medical Imaging*, 22(11):1341–1343, 2003.

[PFK+05] K. M. Pohl, J. Fisher, R. Kikinis, W. E. L. Grimson, and W. M. Wells, III. Shape based segmentation of anatomical structures in magnetic resonance images. In *Computer Vision for Biomedical Image Applications 2005*, pages 489–498, 2005.

[PFL+05] K. M. Pohl, J. Fisher, J. J. Levitt, M. E. Shenton, R. Kikinis, W. E. L. Grimson, and W. M. Wells. A unifying approach to registration, segmentation, and intensity correction. In *8th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pages 310–318, 2005.

[PMV03] J. Pluim, J. Maintz, and M. Viergever. Mutual information based registration of medical images: A survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004., 2003.

[PvAL+93] P. H. Pretorius, A. van Aswegen, M. G. Lötter, M. G. Nel C. P. Herbst, and A. C. Otto. Verification of a varying threshold edge detection spect technique for spleen volume: A comparison with computed tomography volumes. *The Journal of Nuclear Medicine*, 34(6):963–967, 1993.

[RFS03] D. Rueckert, A. F. Frangi, and J. A. Schnabel. Automatic construction of 3-d statistical deformation models of the brain using nonrigid registration. *IEEE Transactions on Medical Imaging*, 22(8):1014–1025, 2003.

[RK06] P. Rogelj and S. Kovačič. Symmetric image registration. *Medical Image Analysis*, 10(3):484–493, 2006.

[RL01] S. Rusinkiewicz and M. Levoy. Efficient variants of the icp algorithm. In *Proceedings of Third International Conference on 3-D Digital Imaging and Modeling (3DIM 2001)*, pages 145–152, 2001.

[RNU+06] W. Reomer, A. Nömayr, M. Uder, W. Bautz, and T. Kuwert. SPECT-guided CT for evaluating foci of increased bone metabolism classified as indeterminate on SPECT in cancer patients. *Journal of Nuclear Medicine*, 47(7):1102–1106, 2006.

[Roh01] K. Rohr. *Landmark-Based Image Analysis: Using Geometric and Intensity Models*. Kluwer Academic Publishers, Norwell, MA, USA, 2001.

[RPA98] A. Roche, X. Pennec, and N. Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 98*, October 1998.

[She68] D. Shepard. A two-dimensional interpolation function for irregularly space data. In *Proceedings 23rd National Conference of the ACM*, pages 517–524, 1968.

[SHV+09] M. Spiegel, D. Hahn, Daum V., J. Wasza, and Hornegger J. Segmentation of Kidneys using a New Active Shape Model Generation Technique based on Non-rigid Image Registration. *Computerized Medical Imaging and Graphics*, 33(1):29–39, 2009.

[SLRG07] O. Schillaci, L. Filippi, R. Danieli, and G. Simonetti. Single-photon emission computed tomography/computed tomography in abdominal diseases. *Seminars in Nuclear Medicine*, 37(1):48–61, 2007.

[Spi06] M. Spiegel. Generation of active shape models for segmentation and registration. Prediploma thesis in Computer Science, Chair for Pattern Recognition, University of Erlangen-Nuremberg, July 2006.

[SRCG04] O. Schillaci, R. Danieli, C. Manni, and G. Simonetti. Is SPECT/CT with a hybrid camera useful to improve scintigraphic imaging interpretation? *Nuclear Medicine Communications*, 25(7):705–710, 2004.

[SRN+03] M. Styner, K. T. Rajamani, L. P. Nolte, G. Zsemlye, G. Székely, C. J. Taylor, and R. H. Davies. Evaluation of 3d correspondence methods for model building. In *Information Processing in Medical Imaging*, pages 63–75, July 2003.

[SS01] L. G. Shapiro and G. C. Stockman. *Computer Vision*, pages 279–325. Prentice-Hall, New Jersey, USA, 2001.

[STU05] C. O. S. Sorzano, P. Thevenaz, and M. Unser. Elastic registration of biological images using vector-spline regularization. *IEEE Transactions on Biomedical Engineering*, 52(4):652–663, April 2005.

[SW92] R. Schaback and H. Werner. *Numerische Mathematik*. Springer-Verlag, Berlin, 4th edition, 1992.

[SW00] O. Scherzer and J. Weickert. Relations between regularization and diffusion filtering. *Journal of Mathematical Imaging and Vision*, 12(1):43–63, 2000.

[Tho03] H. H. Thodberg. Minimum description length shape and appearance models. In *Information Processing in Medical Imaging*, pages 51–62, 2003.

[TOS01] U. Trottenberg, C.W. Oosterlee, and A. Schüller. *Multigrid*. Academic Press, London and San Diego, 2001.

[TYW+03] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Meical Imaging*, 22(2):137–152, February 2003.

[USM+06] D. Utsunomiya, S. Shiraishi, M. Imuta, S. Tomiguchi, K. Kawanaka, S. Morishita, K. Awai, and Y. Yamashita. Added value of SPECT/CT fusion in assessing suspected bone metastasis: comparison with scintigraphy alone and nonfused scintigraphy and CT. *Radiology*, 238(1):264–271, 2006.

[Ven02] P. Venkataraman. *Applied Optimization with MATLAB Programming*. John Wiley and Sons Inc., Hoboken, USA, 2002.

[VW97] P. Viola and W. M. Wells. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, 1997.

[WSV91] M. W. Walker, L. Shao, and R. A. Volz. Estimating 3-d location parameters using dual number quaternions. *Image Understanding*, 54(3):358 – 367, 1991.

[XJ. L. Prince98] C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7(3):359–369, 1998.

[YL05] Y. Young and D. Levy. Registration-based morphing of active contours for segmentation of ct scans. *Mathematical Biosciences and Engineering*, 2(1):79–96, 2005.

[ZCB+04] D. Zhao, W. Chen, H. Bao, H. Zhang, and Q. Peng. Real-time voxelization for complex polygonal models. In *Proceedings of Pacific Graphics 2004*, pages 73–78, October 2004.

[ZF03] B. Zitová and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.

[ZYK01] L. Zöllei, A. Yezzi, and T. Kapur. A variational framework for joint segmentation and registration. In *MMBIA'01: Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 44–51, Washington, DC, USA, 2001. IEEE Computer Society.

List of Publications

Refereed Journal Articles

- J. Han, B. Berkels, M. Droske, J. Hornegger, M. Rumpf, C. Schaller, J. Scorzin and H. Urbach, *Mumford-Shah Model for One-to-One Edge Matching*, *IEEE Transaction on Image Processing*, 16 (2007) No. 11 pp. 2720-2732.

- J. Han, H. Köstler, C. Bennewitz, T. Kuwert, J. Hornegger, *Computer-Aided Evaluation of Anatomical Accuracy of Image Fusion between X-Ray CT and SPECT*, *Computerized Medical Imaging and Graphic*, 32 (2008) No. 5 pp. 388-395

Conference Proceeding Articles

- J. Han, C. Bennewitz, J. Hornegger, T. Kuwert, *Semi-automatical Validation of SPECT/CT Scanners*, *3rd Russian-Bavarian Conference on Biomedical Engineering*, Erlangen, Germany, 2007 vol.1, pp. 93-100.

- J. Han, B. Berkels, M. Rumpf, J. Hornegger, M. Droske, M. Fried, J. Scorzin, C. Schaller, *A Variational Framework for Joint Image Registration, Denoising and Edge Detection*, *Bildverarbeitung für die Medizin 2006* Hamburg, Germany, 2006, pp. 246-250.

- J. Han, J. Hornegger, T. Kuwert, W. Bautz, W. Römer, *Feature Constrained Non-rigid Image Registration*, *Frontiers in Simulation*, Erlangen, Germany, 2006, pp. 638-643.

- J. Han, M. Qiao, J. Hornegger, T. Kuwert, W. Bautz, W. Römer, *Automatic sub-volume registration by probabilistic random search*, *Procceeding SPIE Medical Imaging 2006*, San Diego, USA, 2006, vol. 6144, pp. 799-807.

- F. Jäger, J. Han, J. Hornegger, T. Kuwert, *A Variational Approach to Spatially Dependent Non-Rigid Registration*, *Procceeding SPIE Medical Imaging 2006*, San Diego, USA, 2006, vol. 6144, pp. 860-869.

- F. Jäger, J. Han, J. Hornegger, T. Kuwert, *Wissensbasierte Nicht-Starre Registrierung von SPECT/CT Datensätzen*, *Bildverarbeitung für die Medizin 2006* Hamburg, Germany, 2006, pp. 236-240.
- J. Han, M. Qiao, J. Hornegger, T. Kuwert, W. Bautz, W. Römer, *Partial Image Data Registration using Stochastic Optimization*, *Frontiers in Simulation*, Erlangen, Germany, 2006, pp. 644-645.
- M. Prümmer, J. Han, J. Hornegger, *2D-3D Non-rigid Registration using Iterative Reconstruction*, *Vision Modeling and Visualization*, Erlangen, Germany, 2005, pp. 187-194.
- W. Römer, J. Hornegger, J. Han, W. Bautz, T. Kuwert, *Non-rigid Fusion of Morphological and Functional Images Using Anatomical Fix Points and Contours - A New Approach to Overcome the Current Drawbacks of Retrospective Image Fusion The Radiological Society of North America (RSNA)* Chicago, USA, 2005

Awards in Conferences

- The work *A Variational Framework for Joint Image Registration, Denoising and Edge Detection* was awarded with "BVM-Preis für ein Herausregendes Poster" in *Bildverarbeitung für die Medizin Konferenz*, Hamburg, Germany, 2006
- The work *Automatic sub-volume registration by probabilistic random search* was awarded with "Honorable Mention Poster Award" in *SPIE Medical Imaging 2006 Symposium*, San Diego, USA, 2006

Supervised Student Theses

- S. Gaffling *Non-rigid Registration for Interpolation of Defect Slices in Sequences of Microscopic Photographs* Master thesis, 2007.
- Y. Zhou, *Automatic segmentation of liver lesions in CT data sets*, Master thesis, 2006.
- Z. Mou, *Multigrid methods for non-rigid image registration*, Master thesis, 2006.
- M. Qiao, *Mono-modal Image Registration for Partial Volume Data*, Master thesis, 2005.
- C. Siller, *Vergleich von Ähnlichkeitsmassen für eine multimodale Bildregistrierung unter Berücksichtigung anatomischer Merkmale*, Prediploma thesis, 2004.

Acknowledgements

First and foremost, I would like to thank my advisor Prof. Dr. Joachim Hornegger for giving me the opportunity to participate in the project C10 in Sonderforschungsbereich (SFB) 603, funded by the German Research Foundation (DFG). I acknowledge his inspirative advice, constant encouragement and especially his trust and support during some difficult periods of my Ph.D study. Likewise, my appreciations go to Prof. Dr. Martin Rumpf who invited me to work in his team for four months. This wonderful cooperation did not only widen my perspective of the methods of mathematical analysis in the field of image processing but also inspired me to formulate one of the most important theoretical foundations of this thesis. I am also thankful to Prof. Dr. Torsten Kuwert for my fruitful inclusion at the nuclear medicine department of university of Erlangen. His expertise and valuable comments helped me very much to discover new aspects of registration techniques in the clinical field.

I was very lucky to work with the colleagues at the chair for pattern recognition for more than three years. First of all, I want to thank Eva Rothgang and Michael Balda for the time and effort to improve this manuscript. I received so much help from my office mate Marcus Prümmer who always patiently answered all my questions on teaching, research and our backup system. Eva Kollorz was so nice to endure my poor German for such a long time and patiently review my papers, presentations and reports. Also I want to thank Dieter Hahn and Volker Daum, two very talented programmers, for their quick and sharp answers to my questions on implementation topics. I would also like to thank Martin Spiegel and Andreas Wimmer for passing on their knowledge on computer graphics and the InSpace framework to me. Although I cannot mention the names of all those involved here, I would nevertheless like to express my gratitude for the support they gave me and also the fun we had together.

In addition to the colleagues at our chair, I am very glad that I was able to have wonderful cooperations with Dr. Wolfgang Römer, Dr. Harald Köstler, Benjamin Berkels and Christian Bennewitz during my Ph.D.

Last, but not least I would like to thank my family and friends. Especially important are my wife Yu and my son Jitong, who constantly give me joy, strength and confidence in my life. Without their understanding and encouragement, I would not have been able to finish this thesis.

Jingfeng Han