Towards Robust and Accurate Image Registration by Incorporating Anatomical and Appearance Priors

LU YONGNING

(B. Eng., National University of Singapore)

A thesis submitted for the degree of

Doctor of Philosophy

NUS Graduate School for Integrative Sciences and Engineering

NATIONAL UNIVERSITY OF SINGAPORE

2014

Declaration

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis. This thesis has also not been submitted for any degree in any university previously.

Signed:

LU YONGNING

Date:

This thesis is dedicated to

all the people *– who never stopped having faith in me*

my parents *– who raised me and supported my*

education, for your love and sacrifices

my wife *– who is so understanding and supportive*

along the journey

my friends *– whom I am so grateful to have in my life*

Acknowledgements

Summary

Image registration is one of the fundamental computer vision problems, with applications ranging from motion modeling, image fusion, shape analysis, to medical image analysis. The process finds the spatial correspondences between different images that may be taken at different time or by modalities of acquisition. Recently, it has been shown that incorporating prior knowledge into the registration process has the potential to significantly improve the image registration results. Therefore, many researchers have been putting lots of effort in this field.

In this thesis, we investigate the possibility of improving the robustness and accuracy of image registration, by incorporating anatomical and appearance priors. We explored and formulated several methods to incorporate anatomical and appearance prior knowledge into image registration process explicitly and implicitly.

To incorporate the anatomical prior, we propose to utilize the segmentation information that is readily available. An intensity-based similarity measure named structural encoded mutual information is introduced by emphasizing the structural information. Then we use registration of the anatomical-meaningful point sets that are extracted from the surface/contour of the segmentation to generate an

anatomical meaningful deformation field. The two types of data-driven prior information are then combined in a hybrid manner to jointly guide the image registration process. The proposed method is fully validated in a pre-operative CT and non-contrast-enhanced C-arm CT registration framework for Trans-catheter Aortic Valve Implantation (TAVI) and other applications.

To incorporate the appearance prior, we proposed to describe the intensity matching information by using normalized pointwise mutual information which can be learnt from the training samples. The intensity matching information is then incorporated into the image registration framework by introducing two novel similarity measures, namely, weighted mutual information and weighted entropy. The proposed similarity measures have demonstrated their wide applicability ranging from natural image examples to medical images from different applications and modalities.

Lastly, we explored the feasibility of generating different image modalities from one source image based on prior image matching knowledge that is extracted from the database. The synthesized images based on prior knowledge can be then used for image registration. Using the synthesized images as the intermediate step in the multi-modality registration process explicitly simplifies the problem to a single modality image registration problem.

The methods and techniques we proposed in this thesis can be combined and/or tailored for any specific applications. We believe that

with more population databases made available, incorporating prior knowledge can become an essential component to improving the robustness and accuracy of image registration.

Contents

List of Figures

Chapter 1

Introduction

1.1 Image Registration: An Overview

In the field of image processing, it is often important to spatially align the images taken from different instants, from different devices, or different perspectives, so as to perform further qualitative and quantitative analysis of the images. The process of spatially aligning the images, is called *image registration*. More precisely, the goal of image registration is to find an optimal spatial transformation that maps the target image to the source image. From a mathematical perspective, given two input images, namely the source and target images, the image registration process is an optimization problem that finds the geometric transformation that brings the source image to be spatially aligned with the target image. The types of geometric transformation depends on the specific application. Generally, the transformation can be divided into two groups – *global* and *local*. The selection of the transformation model is highly dependent on the application.

As a fundamental computer vision problem, image registration has a wide range of applications, including motion modeling, image fusion, shape analysis, and medical image analysis. Detailed surveys and overviews on applications of image registration can be found in [1], [2], [2], [3], [4], [5], [6] and [7]. In this thesis, we will mainly focus on but not limited to deformable medical image registration, although the proposed methods can be straightforwardly applied to other applications, which we will also demonstrate in this thesis.

Image registration helps the clinicians to interpret the image information acquired from different modalities, different time points, or pre- and post- contrast-enhancement. Combining the image information from different time instants helps the clinicians to examine the disease progression over time. As the imagining technology develops, there are more and more imaging modalities that provide spatial co-localization of complementary information, including structural and functional information. These image modalities can be generally classified as either *anatomical* or *functional* [8, 9, 10]. Morphological information is explicitly depicted in the anatomical modalities, which include CT (computed tomography), MRI (magnetic resonance imaging), X-ray, US (ultrasound), etc. Metabolic information on the target anatomy is emphasized in the functional modalities, which include scintigraphy, PET (position emission tomography), SPECT (single photon emission computed tomography), and fMRI (functional MRI). Complementary information from different imaging modalities makes the assessment to be more convenient and accurate for the clinicians. With the rapid development of the clinical assessment technique and imaging techniques, medical applications increasingly rely more on the image registration; such applications range from examination of disease progression to the usage

of augmented reality in the minimal-invasive interventions. Therefore, image registration plays an essential role in medical image analysis.

Both mono- and multi- modality image registration play a very important role in medical applications. Applications of mono-modality image registration include treatment comparison between pre- and post- treatment images, registration of dynamic contrast enhanced (DCE) MRI for detecting abnormalities in myocardial perfusion with that have great potential for diagnosing cardiovascular diseases [11]. Multi-modality image registration also has a wide range of applications. In cardiology, for example, to support Trans-catheter Aortic Valve Implantation (TAVI) procedure, the 3D aortic model acquired from contrast-enhanced C-arm CT can be overlaid onto 2D fluoroscopy to provide anatomical details thus enabling more optimal valve deployment [12]. The procedure of extracting the 3D aortic model from contrast-enhanced C-arm CT requires extra radiation which may not be applicable for patients with kidney problems. To address this problem, a 3D/3D image registration between CT and non-contrast-enhanced C-arm CT is performed to obtained the 3D aortic model [13]. In neurosurgery, stereotaxy technology generally uses CT images, but for tumor identification MRI are typically preferred. Image registration allows the transfer of the tumor coordinates from the MR to the CT images. More discussion and analysis of the applications in neurosurgery can be found in [14]. Besides intra-subject registration, inter-subject image registration is playing a much more important role than ever before. Image registration has been extensively used in constructing statistical atlas [15] and atlas-based image segmentation [16].

Image registration algorithms consist of three major components. Firstly, a *transformation space* is needed to restrict the spatial transformation to a plau-

sible space. It is highly application-dependent. Rigid, affine, splines and non-parametric free-form are the typical spaces used for image registration. Secondly, a *similarity metric* is required to quantitatively measure the alignment between two images. Specifically, it quantifies the similarity between the source and target images using a mathematical expression. Similarity measures are generally classified into three groups, namely, intensity-based methods, feature-based methods, and hybrid methods. Thirdly, an *optimization method* is required to find the optimum parameters in the transformation space such that the defined similarity metric is optimized. This thesis will focus on designing adequate similarity metrics for more robust and accurate image registration.

Although numerous image registration techniques have been developed in the past few decades [4, 17, 18], ordinary image registration algorithms still fail to produce robust and accurate results due to different factors, for example, noise, occlusion, etc. Since medical images often contain significant amount of noise, contrast changes, occlusion and distortions due to lack of data acquisition protocols in some applications, image registration is particularly challenging for medical applications. In this thesis, we aim to develop image registration algorithms that increase the robustness and accuracy of image registration by incorporating anatomical and appearance priors.

1.2   Thesis Organization and Contributions

This thesis is organized as follows. Chapter 2 describes the image registration problem in more detail, and discusses existing image registration techniques.

In Chapter 3, we propose an algorithm that utilizes the segmentation in-

formation that is readily available. The anatomical prior is encoded into the registration framework by introducing a novel similarity measure, the structural encoded mutual information, and an anatomical meaningful deformation field to guide the image registration process. Feature-based image registration methods require highly accurate feature correspondence matching. statistically-constrained transformation model based methods usually demand for large size of training data which may not be practical in many applications. And intensity based methods only rely on the intensity information which often cost problems while optimizing the cost function. The proposed hybrid data-driven image registration framework draw upon the strength and avoids the shortcomings from the above mentioned methods, it benefits from the anatomical information which is extracted from the readily available segmentation, and the the prior anatomical prior deformation field does not require a large data set to train, thus providing a more robust and practical solution to the image registration problem.

In Chapter 4, we propose to describe the intensity matching information by using normalized pointwise mutual information. By learning the intensity matching information from the training images, the intensity matching prior is then incorporated into the image registration algorithm by designing two novel similarity measures: weighted mutual information and weighted entropy. The proposed normalized pointwise mutual information as an intensity matching prior is superior to the state-of-the-art methods where intensity joint histogram is learnt to guide the image registration process because NPMI is less sensitive to the change of field-of-view and size of the objects. Such a superior property is very important because now we can then obtain the intensity matching prior from a subset or even just a slice of the volume. NPMI better represents the

correlations between the intensities instead of being dominated by the number of co-occurrence, and thus brings the utilization of the prior intensity matching to a new level.

In Chapter 5, we explore the possibility of generating different image modalities from one source image based on prior image matching knowledge that is extracted from the database. Having the synthesized image, we essentially reduce the multi-modal image registration problem to a less challenging mono-modality registration problem. We propose to utilize the features such as intensity histogram and the Weber Local Descriptor for the matching process. The proposed matching framework provides much more robust and accurate matching results compared to the state-of-the-art method where SSD is used for the matching process. The more general and accurate matching scheme clearly shows its potential in helping image registration in the future.

Concluding remarks and discussion about future work are presented in Chapter 6.

Chapter 2

Background

This chapter aims to provide a comprehensive background on image registration. We first give a general introduction about the image registration problem. Then the major components of the image registration procedure are elaborated in details, with a literature review of state-of-the-art methods.

2.1 Introduction

Image registration is one of the fundamental computer vision problems, with applications ranging from motion modeling, image fusion, shape analysis, to medical image analysis. During the past decades, the rapid development of the image acquisition devices and more and more needs for image analysis invoked the research on image registration, targeting different applications. The process of image registration consists of establishing spatial correspondence between images acquired by different devices and/or at different time instances.

In general, image registration can be performed on a group of images [19,

20] or only two images. In this thesis, we focus on the image registration methods that involve only two images. Here, we give a more mathematical definition of the image registration problem. Given a source image, denoted by S, and a target image, denoted by T, the goal of image registration is to estimate the optimal transformation W^* such that the similarity metric $J(T, S \circ W)$ of the target image, and the transformed source image is optimized. Mathematically, image registration is to estimate the optimal transformation W^* such that the following objective function is optimized:

$$\arg \max_{W} J(T, S \circ W). \tag{2.1}$$

A image registration algorithm typically involves three main components: 1) a transformation model, 2) a matching criterion (similarity metric), and 3) an optimization method. In this thesis, we will mainly review on the transformation model and matching criterion. And in the methods we proposed in Chapter 3 and 4, we adapt the variational framework, and using gradient descent to solve the optimization problem.

2.2 Transformation Models

In this thesis, the definition of registration is based on geometrical transformations — we map the points from space X of the source image to space Y of the target image. The transformation W applied to a point $\boldsymbol{x}$ in space X produces a point in $\boldsymbol{x'}$,

$$\boldsymbol{x'} = W(\boldsymbol{x}). \tag{2.2}$$

We say that the registration is performed successfully if x' is matched or close to matched y in space Y, which is the exact correspondence of x. The set of possible transformations W can be divided into two groups: 1) global transformation models and 2) local transformation models. Each transformation group can be further classified into many subsets. Global transformation models make use of the information from the image for estimating a set of transformation parameters that is valid for the entire image. Global transformation is used to correct the misalignment of the images in a global scale, thus, it is usually a necessity as the first step of image registration. However, a global mapping is not able to handle images with local deformation, thus local mapping models are usually required after the global registration to further refine the registration process. Compared to global registration models, in which limited parameters are capable of specifying the transformation in 3D, local registration models are usually more application-dependent and require more parameters to be estimated.

2.2.1 Global Transformation Models

Linear models are the most frequently used for estimating global transformations. Although violations of the linearity assumption may require the use of higher order polynomial models, such as second or third-order, higher order polynomial models are rarely used in practical applications.

2.2.1.1 Rigid Transformations

Rigid transformations preserve all distances, and furthermore, they preserve the straightness of lines, the planarity of surfaces, and all angles between the straight

lines. The ubiquity of rigid objects in the real world makes rigid registration one of the most popular global transformation models. The rigid transformation model is very simple to specify, since it comprises only rotation and translation. In the 3D space, and under Cartesian coordinates, the translation vector $\mathbf{t}$ can be specified as a $3{\times}1$ matrix $[t_x, t_y, t_z]'$, where x, y, z are the Cartesian axes. It can also be specified in other coordinate systems, for example, spherical coordinates, however, we will consistently use Cartesian coordinate system to avoid confusion. Other coordinate systems can be easily derived from the Cartesian coordinate system. Specified using Euler angles, rotation can be parameterized in terms of three angles of rotation, $\theta_x, \theta_y, \theta_z$, with respect to the Cartesian axes. Here, we define the three basic rotations, using the right hand rule, as follows:

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix} \tag{2.3}$$

$$R_y(\theta) = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \tag{2.4}$$

$$R_z(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.5}$$

To generalize, other rotation matrices can be obtained by multiplying the three basic rotation matrices:

$$R = R_z(\alpha)R_y(\beta)R_x(\gamma) \tag{2.6}$$

We want to emphasize here that R is an orthogonal matrix, with $det(R) = +1$, where det is the determinant operator. Now, with the transformation W a rigid transformation, then

$$\boldsymbol{x}' = R\boldsymbol{x} + \mathbf{t}. \tag{2.7}$$

2.2.1.2 Affine Transformations

Another popularly used global transformation models are the affine transformations:

$$\boldsymbol{x}' = A\boldsymbol{x} + \mathbf{t}. \tag{2.8}$$

Affine transformations do not have any restriction on the elements a_{ij} in the matrix A. It preserves straight lines (and the planarity of surfaces) and straight line (and surface) parallelism. However, angles between the lines are allowed to change. Affine transformations are appropriate, because image acquisition may introduce a skew factor. Furthermore, it is widely used for multiview image registration, assuming that the distance from the camera to the scene is much larger than the scene area, with a pin-hole camera, a flat scene, and the geometric distortions do not contain local factors.

2.2.2 Local Transformation Models

The global transformations average out the geometric deformation over the entire image domain. Consequently, local deformation may not be properly handled. However, local deformation is a very important component in many applications, for example, medical applications where large organ deformation occurs. Therefore, local areas of the images should be taken care of with specific local transformation models.

Local transformation models are often referred to as non-rigid or deformable transformation models, we use them interchangeably in this thesis. It has been shown that local transformation models are superior to the global models when local geometric distortion is inherent in the images to be registered [4, 7, 21, 22, 23]. Moreover, the choice of local transformation models is important as it relates to the compromise between computational efficiency and richness of the description, as well as the relevance to the particular application. Here, we classify local transformation models into three main categories: 1) derived from physical models, 2) based on basis function expansions, and 3) knowledge-based transformation models.

2.2.2.1 Transformations derived from physical models

Following Modersitzki [21], we further divide the transformations derived from physical models into five categories: 1) linear elastic body models, 2) diffusion models, 3) viscous fluid flow models, 4) flows of diffeomorphisms and 5) curvature registration.

1) Linear Elastic Body Models

The linear elastic body models are described by the Navier-Cauchy Partial Differential Equation (PDE):

$$\mu\nabla^2\mathbf{u} + (\mu + \lambda)\nabla(\nabla \cdot \mathbf{u}) + \mathbf{F} = 0, \tag{2.9}$$

where $\mathbf{u}(\mathbf{x})$ is the transformation vector at location $\mathbf{x}$, $\mathbf{F}(\mathbf{x})$ is the force field that drives the registration process which is derived from maximizing the image matching criteria, λ is the Lamés first coefficient and μ specifies the stiffness of the material.

The Navier-Cauchy partial differential equation 2.9 is an optimization problem that balances the external force that comes from maximizing the matching criteria and the internal force that exhibits the elastic properties of the material. It was first proposed by Broit [24], in which the image grid was modeled as an elastic membrane. Subsequently, the models have been applied to range of applications.

2) Diffusion Models

The diffusion models can be described by the diffusion equation:

$$\triangle\mathbf{u} + \mathbf{F} = 0, \tag{2.10}$$

where $\triangle$ is the Laplace operator. Most of the algorithms based on the diffusion transformation model do not state 2.10 in their formulation or objective function. Nevertheless, in the regularization step, the transformation is convolved with a Gaussian kernel. This is based on the fact that Gaussian kernel is the Green's function of 2.10, thus applying a convolution with the Gaussian kernel is an

effective yet theoretically supported regularization step.

Inspired by Maxwell's Demons, Thirion [25] proposed to model image registration as a diffusion process. The idea is to consider the demons in the target image as semi-permeable membranes and to let the source image diffuse through the demons. The algorithm is an iterative process between: 1) estimating the forces for every demons (based on optical flow), and 2) updating the transformation based on the calculated forces in 1). The iterative process ends till it converges. In the course of medical image registration, it is common to treat all image elements as demons. Furthermore, a Gaussian filter can be applied after each iteration for regularization purpose. The publication of [25] has inspired many methods that share the iterative approach between estimating the forces and then regularizing the deformation field.

3) Viscous Fluid Flow Models

In this case, the transformation is modeled as a viscous fluid. Assuming there is only a small spatial variation in the hydrostatic pressure, and thus a low Reynold's number, the viscous fluid flow is described by the Navier-Stokes equation:

$$\mu_f \nabla^2 \mathbf{v} + (\mu_f + \lambda_f)\nabla(\nabla \cdot \mathbf{v}) + \mathbf{F} = 0. \tag{2.11}$$

The $\mu_f \nabla^2 \mathbf{v}$ term is related to the constant volume or incompressibility of the viscous flow. The expansion or contraction of the fluid is controlled by $(\mu_f + \lambda_f)\nabla(\nabla \cdot \mathbf{v})$. Different from the linear elastic body models, no assumption is made on the small transformations, therefore, the models are capable of recovering large deformations [26]. Multi-modal image registration using viscous fluid models is made possible in [27]. And an inverse consistent variant of viscous

fluid models is proposed in [28].

4) Flows of Diffeomorphisms

Local transformations can be also modeled by flows of diffeomorphisms. The velocity field over time is constrained to be smooth under the regularization term:

$$R = \int_0^1 \|\mathbf{v}_t\|_V^2 dt, \tag{2.12}$$

where $\| \cdot \|_V$ is a norm on the smooth velocity vector space V. Different types of spatial regularization can be specified through changing the kernel associated with V [29]. The choice of kernel may be either a single Gaussian kernel [30] or adaptive Gaussian kernel selections on the entire image domain [30, 31].

5) Curvature Registration

Under curvature registration, the constraint of the model is formulated by the equilibrium equation:

$$\triangle^2 \mathbf{u} + \mathbf{F} = 0. \tag{2.13}$$

Fischer and Modersitzki [32] show that curvature based image registration not only provides accurate and smooth solutions for the image registration task, but also avoids a pre-registration step. This is because the regularization scheme based on 2.13 does not penalize affine transformations. Equation 2.13 is solved by a finite difference scheme, which imposes the Neumann boundary conditions. Imposing the Neumann boundary conditions may have the effect of penalizing the affine transformations, Henn [33] proposed a full curvature based image registration method which includes second-order terms as boundary conditions to solve the problem.

2.2.2.2 Transformations based on basis function expansions

Another category of local transformation are modeled based on a set of basis functions. The coefficients of the basis functions are adjusted such that the resulted transformation maximizes some similarity metric that measures the alignment of the source and target images. The fundamental mathematical framework behind these set of transformation models are mainly from the theory of function interpolation [34] and approximation theory [35, 36]. Here, we only review five of the most important models that are based on basis function expansions, namely, 1) radial basis functions, 2) elastic body splines, 3) B-splines, 4) Fourier and wavelets, and 5) locally affine models.

1) Radial Basis Functions

Radial basis functions are ones of the most important interpolation strategies [37, 38, 39]. The value of the interpolation point **x** is calculated as a function of its distances to the landmark positions. It is defined as:

$$\mathbf{u}(x) = \sum_{i=0}^{N} \alpha_i R(\|\mathbf{x} - \mathbf{x}_i\|), \tag{2.14}$$

where i is the index for the landmarks, and α_i assigns different weights to the corresponding landmarks. The common choice of function R can be Gaussian [40] or inverse multiquadric [41]. A remarkable strength of the radial basis function to point out is its global support. The radial basis function are positive defined functions in which a closed-form solution of finding the optimal set of coefficients exists. Having said that, as the displacement at any landmark point would introduce influence to the whole image domain, sufficient or more dense landmarks are required at the area where local deformation is more dominant.

For medical image registration, landmark-based methods using the radial basis functions were extensively studied in [42] and [43]. A comparative evaluation study on the use of the radial basis functions in non-rigid image registration can be found in [44].

2) Elastic Body Splines

Davis et al. [45] introduced the Elastic Body Splines (EBS) which is a physics-based coordinate transformation. The Elastic Body Splines are the solutions to the Navier-Cauchy PDE 2.9. The equation can be solved analytically when the force field is given as a radial symmetric function of the distance from the landmark. The work was extended by Kohlrausch et al. [46] using Gaussian EBS, thus the local transformation can be better represented by the transformation model. Gaussian EBS was further developed by Wörz and Rohr [47], by taking into account the errors in the landmark displacement field using an approximated strategy instead of exact interpolation. Localization uncertainty was considered in [47], and an analytic solution was provided.

3) B-splines

Back to 1940s, Schoenberg [48] first introduced the B-splines for interpolation. Since then, the applications of B-splines have been widely developed, and it has become a popular tool for solving the interpolation problems in the field of signal processing [49, 50, 51]. In the course of image registration, free-form deformations (FFDs) is of the most commonly used local transformation models that belong to B-splines family. More specifically, coupling FFDs with cubic-B splines has been widely accepted in the medical image analysis community [52, 53, 54, 55]. In this section, we will mainly focus on reviewing the FFDs.

Given a 3D image size of $N_x \times N_y \times N_z$, a rectangular grid of $K_x \times K_y \times K_z$

is superimposed on the image. The transformation is modeled by the using the translations of the control points $\mathbf{d}_{i,j,k}$ on the grid. The displacement field is defined as:

$$\mathbf{u}(x) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(u_x) B_m(u_y) B_n(u_z) \mathbf{d}_{i+l,j+m,k+n}, \qquad (2.15)$$

where B_l is the lth basis function of the B-spline.

Despite the superiority of the FFDs in providing simple and efficient smooth transformations and requiring few degrees of freedom for describing the transformations, topology preservation is not guaranteed. Rueckert et al. [56] produced diffeomorphic deformation fields with some hard constraints imposed. The extensions of the original FFDs have been carried out widely. Different methods for placing the control points non-uniformly have been proposed [57, 58, 59]. Symmetric and inverse consistency of the transformation field is studied in [60, 61, 62].

4) Fourier and Wavelets

Fourier and wavelets are used to model the local transformations in many applications. An important, or probably the most important reason is that, Fourier and wavelets methods naturally decompose the transformation field in a multi-resolution manner, which is desirable in the image registration applications — less computational demand, larger capture range, harder for being trapped in the local optimum during the optimization process, etc.

Amit [63] considered image registration as a nonlinear variational problem, and presented two approaches for image matching — one based on Fourier basis and one based on wavelets. It was reported that, wavelets-based method

was more capable of capturing the local deformation than the Fourier-based method. Deformation field modeled by Fourier or wavelets are also reported in [64, 65, 66]. These methods all emphasized on the multi-resolution and/or multi-band decomposition for accelerating the calculation and increasing the computational efficiency. Furthermore, the topology of the image is preserved by imposing additional constraints. For example, hard constraints on the Jacobian were imposed in [67].

5) Locally Affine Models

As the name suggests, in this case, the local transformations are defined by locally affine models. The main strength of locally affine models is its computational efficiency. These family of local transformation models can be further divided into two categories: piecewise affine models and poly-affine models. The basic idea is to mosaic the image by a set of regions whose nodes parameterize the transformation. While the piecewise affine models usually define the regions quite evidently, the poly-affine models use fuzzy regions to avoid the lack of smoothness at the boundary regions.

Piecewise affine models: Hellier et al. [68] first introduced a multi-grid and multi-resolution approach based on a piecewise affine model which was regularized with an optical flow model. Not globally invertible is the main drawback of applying piecewise affine transformation in each region independently. This issue is partly addressed and tackled in [69].

Poly-affine models: To overcome the drawback of the piecewise affine models, poly-affine models use fuzzy regions. Arsigny et al. [70] introduced poly-affine transformations, which was later extended by Arsigny et al. [71]. A set of anchor points were defined and assigned with different weights (importance).

Fuzzy regions are defined according to the influences of the anchor points to each position **x**. The approach does not have a closed-form solution, thus it is computationally expensive.

2.2.2.3 Knowledge-based transformation models

Specific knowledge on the transformation can be incorporated into the transformation models to result in favorable transformation. The motivation behind incorporating knowledge into the transformation models is to increase the accuracy, robustness and plausibility of the transformation. Some general knowledge including topology preservation and volume preservation have been well studied, as briefly mentioned in the previous sections. In this section, we will focus more on knowledge that are derived from statistics and biomechanical/biophysical models.

1) Statistically-Constrained

Statistically-Constrained transformation models, or statistical deformations models (SDMs) are models that utilize the statistical information from the transformation fields. The statistical information is usually collected from a population of subjects. Due to the prior knowledge introduced, the degree of freedom of the transformation is constrained and thus reduced, which directly leads to less demand on computational power.

Principal component analysis (PCA) has be widely used while learning the statistical models. Tang et al. [72] accelerate the SDM learning process by applying PCA. Rueckert et al. [15] studied the movement of each of the FFD control points, and then used the statistical result to constrain on the transformation of the FFD control points. The stiffness of different structures influences the

deformability of the objects/tissues, statistical prior knowledge on this was introduced by Commowick et al. [73]. One common drawback of statistical learning model was the high dimensionality of the prior knowledge, Xue et al. [74] tackled the problem of high dimensional SDMs using wavelet based decompositions. Nevertheless, SDMs still relies on the representability of its training samples for a good transformation model.

2) Biomechanical/Biophysical inspired

Biomechanical/biophysical properties of the tissues can be explicitly encoded, thus the complex transformation field can be more easily obtained with the reduced degree of freedom. The searching space of the transformation is thus much reduced, and therefore the increase of computation efficiency. Embedding the biomechanical/biophysical properties into the image registration process also helps to produce more anatomically plausible transformation. These models are closely related to anatomy and physiology. Therefore, they should be very carefully examined. Failing to represent the anatomy/physiology and assign the parameters correctly may produce undesirable results.

These models are usually applicable to specific tasks. For example, [75], [16] and [76] focus on modeling the tumor growth. [77], [78] and [79] work on biomechanical models of the breast. And biomechanical models of the prostate are also reported in [80, 81].

2.3 Matching Criterion

The alignment of the target and source images is described by the matching criterion, or the similarity metric. There are mainly three approaches in design-

ing similarity measures, namely, 1) feature-based methods, 2) intensity-based methods, and 3) hybrid methods. The feature-based methods aim to establish an optimum correspondences for the landmark points, e.g., anatomical locations, salient points, etc. It involves feature points detection, and then solving for the correspondence problem. Intensity-based methods measures the images alignment based on the information extracted from the intensity relationships between the images. While hybrid methods combines the information provided by landmarks and intensity.

2.3.1 Feature-Based

2.3.1.1 Feature Points Detection

Detecting the features from the source and target images is the first step of feature-based registration methods. Depending on the application, features can be salient points/regions, lines, corners, line intersections, anatomical meaningful points/regions, etc. An extensive review on point-detectors and descriptors and be found in [82]. The detection and matching of the landmark points are highly dependent on the richness of the description. The descriptors should be discriminant enough for distinguishing between the potential matching candidates. Furthermore, it is desirable for the descriptors to be invariant to factors, such as, rotation, translation, deformation, and intensity changes, thus a robust matching can be obtained.

2.3.1.2   Transformation Estimation based on Feature Points

In the field of image registration, we hope to estimate a plausible transformation such that the two landmark point sets are aligned. In the case where the landmark correspondences are made known, the transformation can be calculated by adapting interpolation strategy, e.g., radial basis functions or thin-plate splines straightforwardly [42, 47], or incorporating a regularization energy based on correspondence constraint.

When the landmark correspondences are unknown, the landmark points can be represented as probability distributions. The transformation is thus estimated by minimizing the distance between two distributions. Kernel correlation [83] and kernel density correlation [84] were used as the distance measure. More recently, Gaussian mixture models (GMMs) have been more actively used in this research area. In [85], each point set is modeled as GMM, the distance between the point set is measured by L_2 distance. Each feature of each shape is modeled as GMM in [86]. Geometric information can also be adopted to estimate the transformation without exact landmark correpondences. Signed distance functions were used for shape representation [87, 88]. Euclidean distance transform was deployed in [89].

2.3.2   Intensity-Based

Intensity-based methods measures the registration accuracy based on the information extracted from the intensity relationships between the images. The intensity-based methods do not require landmarks detection, therefore, the procedure for intensity-based methods is more straightforward. Nevertheless, com-

pared to landmarks based methods where only a subset of the pixels are considered as landmarks, intensity-based methods make use of the information from every pixel. This comes at the cost of consuming more computational power. Based on the modalities of two images, we further classify the intensity-based methods into two categories: 1) for mono-modal image registration applications, and 2) for multi-modal image registration applications.

2.3.2.1   Mono-modal Image Registration

In the mono-modal setup, both the source and target images come from the same modality, and thus share the same intensity properties. The most straightforward matching criterion is the sum of squared difference:

$$SSD(T, S) = \sum_{i=0}^{N} (T(\mathbf{x}_i) - S(\mathbf{x}_i))^2, \qquad (2.16)$$

where the squared difference is summed up over the whole image domain. This similarity metric assumes that the same structures have the same intensity values in the two images. Similar assumption has been made in sum of absolute difference (SAD):

$$SSD(T, S) = \sum_{i=0}^{N} |T(\mathbf{x}_i) - S(\mathbf{x}_i)|. \qquad (2.17)$$

The SSD and SAD as matching criteria are reported to be sensitive to the noise and outliers. Later, the intensity matching relationship between the source and

target images is relaxed to linear — cross coefficient:

$$CC(T, S) = \frac{Cov(T, S)^2}{Var(T)Var(S)}, \tag{2.18}$$

where Cov and Var are the covariance and variance operators respectively. Or functional — correlation ratio [90]:

$$CR(T, S) = \frac{Var(E(S|T))}{Var(S)}, \tag{2.19}$$

where E is the expectation.

Besides purely relying on the pixel-based intensity values, attribute-based methods were also proposed to provide richer information, thus some ambiguous matching resulted from purely relying on pixel-based intensity values as stated above can be avoided. Shen and Davatzikos [91] proposed a hierarchical attribute matching mechanism for elastic registration, where the attribute vector includes geometric moment invariants. Local histograms were utilized in [92]. Gabor [93] and alpha stable filters [94] were applied in different image registration tasks. Myronenko and Song [95] analyzed the complexity of the residual image, by minimizing the basis functions of the residual image, image registration can be thus obtained.

2.3.2.2 Multi-modal Image Registration

Multi-modal image registration is naturally more difficult to solve compared with mono-modal registration problems, because the intensity matching information is not as predictable. Furthermore, the structural appearance may be

significantly different in the images from different modalities.

For the cases, where the intensity matching information is still linear or functional, or at least can be assumed as linear or functional in the local region, cross correlation and correlation ratio can be still applicable with the location constraints [96]. Nevertheless, the most widely used approaches for multi-modal image registration is based on information theory. Among all information theoretic approaches, mutual information (MI) is the most popular approach which has been extensively investigated [97, 98]. Mutual information is defined as:

$$MI(T, S) = H(T) + H(S) - H(T, S), \qquad (2.20)$$

where H is the differential entropy. MI removes the assumption on functional relationship, and uses the statistical dependency between the intensity values from the corresponding pixel to evaluate the registration results. Many following studies came out to solve the shortcomings of MI. Studholme et al. [99] proposed the normalized mutual information (NMI) to solve the overlap invariant problem. Higher order of mutual information, which takes into account of the spatial information were also developed rapidly [100, 101, 102, 103].

The idea of incorporating prior knowledge of the underlying registration problem has shown significant improvement in registration robustness and accuracy. The effort of intensity standardizing of different modalities makes incorporating intensity matching prior more practical. More specifically, using divergence measures to comparing the joint intensity distributions from training and testing cases has attracted much attention. Chung et al. [104], Guetter et al. [105] and Cremers et al. [106] proposed to use Kullback-Leibler divergence

(KLD) as the similarity measure to register multimodal images. Liao et al. [107] used Jensen-Shannon divergence (JSD) to compare learnt joint distribution with the observed joint distribution. JSD is more robust compared to KLD due to its symmetry, being theoretically upper-bounded, and well-defined with histogram non-continuity. Despite the reported success of using leaned joint intensity distributions to increase the registration robustness and accuracy, a major drawback of using the leaned joint intensity distributions is that the joint intensity distributions may deviate a lot from the training to testing images which degenerate the effectiveness of the intensity matching prior. To solve the above mentioned problem, in this thesis, we propose to learn the intensity matching relationship from pre-registered image training pairs. Instead of using the learnt joint distributions, we proposed to learn the intensity matching relationship from the joint histogram through the normalized pointwise mutual information (NPMI), and apply the learnt intensity matching as a more general, flexible and robust prior.

Another group of researchers try to solve the multi-modal image registration problem by reducing the multi-modal setup to mono-modal, and hence only mono-modal image registration needs to be performed subsequently. In particular, ultrasound images are simulated from MR [108] and CT [109] images respectively. Andronache et al. [110] and Maintz et al. [111] try to map the source and target images to a common pseudo modality. With the introduction of modality synthesis techniques [112, 113] that are based on constructed dictionaries, databases, we believe that there is a huge potential to improve the image registration accuracy by reducing the multi-modal images to mono-modal images.

2.3.3   Hybrid

Hybrid methods aim to combine the strengths from feature-based and intensity-based methods, and capitalize them in a complementary way. Feature-based and intensity-based methods are taking independently in two steps in [114, 115, 116]. In most of these methods, one type of information (either feature or intensity) dominates in coarse registration, while the other dominates at the finer level. Another group of researchers utilize the additional information as constraints in the registration process. In particular, landmarks/surfaces were used as soft constraints in [117, 118, 119]. And Joshi et al. [120] imposed a hard constraint of geometric correspondences in brain image registration. In most of the above mentioned methods, only one type of information benefits from another. However, these two types of information can be beneficial to each other, so the solution of each problem can take advantage of its counterpart. Such approach can be found in [121, 122, 123].

2.3.4   Group-wise

Although many techniques have been proposed for image registration, most of them belong to the category of pairwise image registration, where only two images, naming the source and target images are involved. Registration of a group of images has traditionally been tackled by repeatedly applying a pairwise registration [124, 125]. Recently, group-wise registration has been proposed for simultaneous and consistent registration of all images in a group [19, 126, 127, 128, 129]. Group-wise image registration is well studied for atlas construction [130, 131, 132]. More specifically, Seghers et al. performed

pairwise registration between all possible pairs of images in the group, and constructed the atlas by voxelwise averaging of all images after mapping them to their mean morphological images [133]. Park et al. [134] defined an image closest to the population mean geometry as a tentative template and generated the atlas by iteratively registering all images onto the template and replacing the template with the mean of the aligned images. More efficient group-wise registration methods are then subsequently proposed. Joshi et al. proposed a method for atlas estimation in a large deformation diffeomorphic setting [135]. A gradient-based stochastic optimizer proposed by Zollei et al. is employed to minimize an information-theoretic objective function, and an affine congealing mechanism is used to drive each image to the center of the group simultaneously [126]. [126] is further extended to a nonrigid group-wise registration algorithm by incorporating free-form B-Splines to represent nonrigid deformations [136]. More recently, it has been pointed out that a single mode is not sufficient to account for the variation of all images in a population, and thus multi-class approaches are proposed [137, 138, 139].

2.4 Conclusions

To this end, we have carefully introduced the background of image registration, including two of the most important components: the transformation model and similarity measure. Despite the efforts from the researchers all over the world, image registration is still a very challenging problem and it is still open for the researchers to improve and solve. In particular, we work from the angle of incorporating additional knowledge to improve the accuracy and robustness

of the image registration process, including anatomical and appearance priors, which will be presented in the subsequent chapters of this thesis.

Chapter 3

Image Registration: Utilizing Anatomical Priors

In this chapter, we present a hybrid multi-modal deformable registration framework using a data-driven deformation prior. The deformation prior is generated by utilizing the anatomical information contained the images, for example, the segmentation of certain organs or objects. The proposed approach belongs to the hybrid approach as described in Chapter 2. We will present the detailed algorithm in the following sections.

3.1 Introduction

Image registration helps the clinicians to combine the image information acquired from different modalities, different time points, or pre- and post- contrast-enhancement for better evaluation. For some cases, rigid/affine registration may be sufficient; however, in many cases, deformable registration is needed to com-
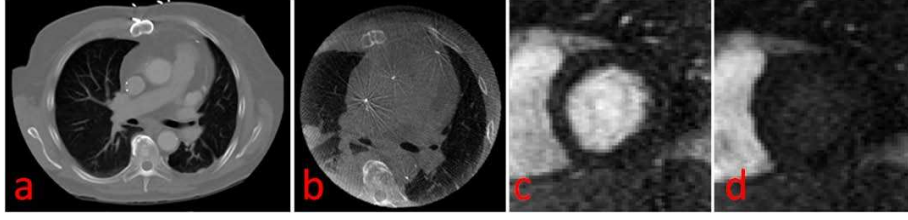
Figure 3.1: Structure appearance may be largely different due to different levels of contrast-enhancement. (a) and (b) is a pair of images from pre-operative contrast-enhanced CT and intra-operative non-contrast-enhanced C-arm CT for TAVI procedure. (c) and (d) is a pair of images from a perfusion cardiac sequence at different phases.

pensate for local movements.

Deformable registration is inherently ill-posed and under-constrained from the mathematical point of view. It becomes more challenging when dealing with different structural appearances due to different levels of contrast-enhancement between two images. This problem widely exists in the field of medical image registration, e.g., registration of the perfusion cardiac images at the wash in/out phases, and 3D/3D registration of pre-operative contrast-enhanced CT and intra-operative non-contrast-enhanced C-arm CT images (See Fig 3.1). In these cases, purely relying on the intensity information produces anatomically implausible deformation. Integrating *a priori* information about the deformation is thus highly desirable. Landmark constraint was proposed to increase the registration accuracy and robustness [140, 141, 142]. These methods added a penalty term to constrain the correspondence pairs from moving too far apart, and thus exact correspondence matching is very crucial. They optimize the energy function using the thin-plate spline (TPS) and B-spline models. TPS are based on the bending energy of a thin plate, thus it only represents a relatively coarse deformation. B-spline models usually require numerical solutions at each

iteration during the optimization process, which is computationally expensive. There is another group of methods incorporating *a priori* information from the statistical point of view. Statistical analysis on shape and displacement field variability is incorporated into the image registration process. Xue et al. [74] tackled the problem of high dimensional statistical deformation models (SDMs) using wavelet based decompositions. Despite the promising results, training the SDMs suffers from the curse of dimensionality, and how to select the training data to represent the population remains unclear. Among the aforementioned methods, one important and potentially readily available information is missing and may be utilized — the segmentation of some dominant and common objects in the images. The motion of these segmented objects could be modeled and may greatly improve registration accuracy. In addition, from the clinical workflow perspective, this segmentation may be needed for diagnosis and guidance purpose alone, and as a result, utilization of the available segmentation results does not impose additional requirement for the purpose of image registration.

3.2 Dense Matching and The Variational Framework

In this thesis, we use the variational framework for image registration [96]. We shall provide a brief introduction here as it is extensively used in Chapter 2 and 3.

Given the image domain Ω, we want to find a function $\boldsymbol{h}$ to transform each point $\boldsymbol{x}$ with a displacement vector $\boldsymbol{h}(\boldsymbol{x})$. We briefly review the variational frame-

work in this section. The function $\boldsymbol{h}$ is optimized such that an energy functional $J(\boldsymbol{h})$ is minimized. The energy functional $J(\boldsymbol{h})$ is of the form

$$J(\boldsymbol{h}) = J_1(\boldsymbol{h}) + R(\boldsymbol{h}), \tag{3.1}$$

where $J_1(\boldsymbol{h})$ measures the "dissimilarity" between the source and target images, and $R(\boldsymbol{h})$ is designed for smoothness regularization. Thus, image registration is to find the $\boldsymbol{h}$ such that $J_1(\boldsymbol{h}) + R(\boldsymbol{h})$ is minimized.

The first variation of $J(\boldsymbol{h})$ at $\boldsymbol{h}$ in the direction of $\boldsymbol{k}$ is defined by

$$\delta_{\boldsymbol{k}} J(\boldsymbol{h}) = \frac{\delta J(\boldsymbol{h} + \epsilon \boldsymbol{k})}{\delta \epsilon}\big|_{\epsilon=0}. \tag{3.2}$$

The gradient $\nabla J(\boldsymbol{h})$ of J is defined by requiring the equality

$$\delta_{\boldsymbol{k}} J(\boldsymbol{h}) = (\nabla_{\boldsymbol{k}} J(\boldsymbol{h}), \boldsymbol{k}) \tag{3.3}$$

to hold for every $\boldsymbol{k}$. $\delta_{\boldsymbol{k}} J(\hat{\boldsymbol{h}})$ equals to zero for every $\boldsymbol{k}$ for the minimizer $\hat{\boldsymbol{h}}$ to be existed. This is equivalent to $\delta_{\boldsymbol{k}} J(\hat{\boldsymbol{h}}) = 0$. The above functions are the Euler-Lagrange equations associated with the energy functional J which is usually impossible to find the close solution. Therefore, a gradient decent strategy is usually deployed to find the minimizer $\hat{\boldsymbol{h}}$. A time-dependent, differentiable function $\boldsymbol{h}$, given an initial estimate $\boldsymbol{h}_0$ is computed as the following initial value problem:

$$\begin{cases} \dfrac{d\boldsymbol{h}}{dt} = -\nabla J(\boldsymbol{h}) \\ \boldsymbol{h}(0)(.) = \boldsymbol{h}_0(.) \end{cases} \tag{3.4}$$

The solution of the matching problem is chosen at $\boldsymbol{h}(t)$ where $t \to \infty$.

3.3   Method

3.3.1   Anatomical Knowledge-based Deformation Field Prior

A TPS can be used to represent an image deformation field [143, 144]. It maps a point $\boldsymbol{x}$ from source image I_2 to the corresponding point $\boldsymbol{v}(\boldsymbol{x})$ in the target image I_1. Given a set of control points $\boldsymbol{c}_k \in R^3$ and the associated coefficients a_{ij}, w_{ki} $\in R$ with $i = 1, 2, 3$, $j = 1, 2, 3, 4$, and $k = 1, 2, ..., K$, the TPS transformation $\boldsymbol{v}$ can be written as:

$$\boldsymbol{v}_i(\boldsymbol{x}) = a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + a_{i4} + \sum_{k=1}^{K} w_{ki}\phi(\|\boldsymbol{x} - \boldsymbol{c}_k\|), \qquad (3.5)$$

with the condition

$$\sum_{k=1}^{K} w_{ki} = 0, \text{ and } \sum_{k=1}^{K} \boldsymbol{c}_{ku}w_{ki} = 0, \; i, u = 1, 2, 3, \qquad (3.6)$$

where ϕ is a radial-basis function which is defined as

$$\phi(r) = r^2 \log(r^2). \qquad (3.7)$$

Note that parameters include 12 global affine parameters a_{ij} and $3K$ local coefficients w_{ki} for the control points. In our method, the two additional constraints in (3.6) ensure that the plate would not move or rotate under the imposition of the loads and remain stationary [44]. The control points are placed in a uniform grid to capture the deformations.

Landmark-based methods are commonly used for non-rigid medical image registration [140, 141, 142, 145]. However, in many applications, it is very dif-

ficult to find exact/accurate landmark correspondences from the images. It may be due to poor image quality and/or lack of feature points in the object, e.g., heart surface. Instead of assuming a one-to-one correspondence based on the nearest neighbor criterion, one-to-many relaxations have been proposed to allow for fuzzy correspondences. These approaches do not establish the explicit point correspondence, and thus are less sensitive to the missing correspondences and outliers. In particular, the registration problem is expressed as a joint optimization over the transformation parameter and correspondence matrix in [146]. Tsin and Kanade [83] proposed a kernel correlation based point set registration approach where the cost function is proportional to the correlation of two kernel density estimates. Myronenko et al. [147] proposed another robust nonrigid point set registration algorithm, where they maintain the same Gaussian affinity matrix and also adopt a similar alternating update strategy interpreted in an expectation maximization framework. For the group of methods mentioned above, where fuzzy correspondences are used, they all can be viewed as special cases in Jian and Vemuri's framework using Gaussian mixture models [85]. Because of its generality [85], in this thesis, we use mixture of Gaussians to represent the point sets of interest, which can be sampled from the segmentation. Then we efficiently and robustly register the point sets using [85]'s method. More specifically, the Gaussian mixture model from the given point set in a simplified setting is as follows: 1) The number of Gaussian components is the number of the points in the point set and all components are weighted equally, 2) for each component, the mean vector is given by the spatial location of each point, and 3) all components share the same spherical covariance matrix. Jian and Vemuri's framework using Gaussian mixture models for robust point set registra-

tion is well suited for our application because of its simplicity, and insensitivity to missing correspondences and outliers which commonly exist in our application. Here, we generate an anatomical knowledge-based deformation field prior using TPS model by optimizing the following cost function:

$$E_{\text{TPS}}(\boldsymbol{v}) = \int (f_{\boldsymbol{v}} - g)^2 d\boldsymbol{x} + \lambda E_{\text{bending}}(\boldsymbol{v}), \tag{3.8}$$

where $f_{\boldsymbol{v}}$ is the distribution representing the transformed point set warped by $\boldsymbol{v}$, g is the distribution of the target point set, and $\boldsymbol{x}$ is the location. A small λ ensures that the TPS approximates local deformations well [42]. In our work, we experimentally set λ to 0.001. E_{bending} is the bending energy of the TPS, which can be written as:

$$E_{\text{bending}}(\boldsymbol{v}) = \int \int \int L(\boldsymbol{v}_x) + L(\boldsymbol{v}_y) + L(\boldsymbol{v}_z) dx dy dz, \tag{3.9}$$

$$L(.) = (\tfrac{\partial^2}{\partial x^2})^2 + (\tfrac{\partial^2}{\partial y^2})^2 + (\tfrac{\partial^2}{\partial z^2})^2 + 2(\tfrac{\partial^2}{\partial x \partial y}) + 2(\tfrac{\partial^2}{\partial x \partial z}) + 2(\tfrac{\partial^2}{\partial y \partial z}), \tag{3.10}$$

where $\boldsymbol{v}_x$, $\boldsymbol{v}_y$ and $\boldsymbol{v}_z$ are components for $\boldsymbol{v}$ at different directions respectively.

TPS is chosen to represent the underlying transformation model due to its nice properties, including its smoothness, no free parameters to tune manually, closed-form solutions for both warping and parameter estimation, and physical explanation for its energy function [142, 148]. Compared to other transformation models that have been successfully applied in the course of heart surface registration, for instance, B-spline models [149], TPS is more suited to our application mainly in two aspects: 1) the deformation is more global: the generated deformation prior is sufficient to provide a high-level knowledge of the plausible

deformation field to guide the deformable registration process; and 2) The computational complexity is much lower than other transformation models where the iterative calculation in the optimization process causes long running time. In addition, the distribution of the point sets are modeled as mixture of Gaussians for the purpose of efficient and robust registration [85]. It was demonstrated in [85] that even with outliers and missing parts in the point sets, the algorithm is still able to register the point sets robustly and correctly. Registration using mixtures of Gaussians may not be highly accurate at the edges, compared to other computationally-expensive landmark-based registration methods that focus on point-to-point matching. However, the deformation prior generated from the point sets registration results is sufficient to provide high-level knowledge of the plausible deformation field. Note that, unlike many other spline-based optimization schemes where iterative volume intensity interpolation is required, we only use TPS to generate a deformation prior based on the location distributions of the point sets, which leads to a much higher computational efficiency.

3.3.1.1 Penalty from Prior Deformation Field

Optimizing Equation 3.11 provides a data-driven prior deformation field $\boldsymbol{v}$, and we want the prior deformation field $\boldsymbol{v}$ to guide the deformable registration process. The penalty term is thus defined as:

$$E_{\text{prior}}(\boldsymbol{h}) = - \int_{\Omega} w(\boldsymbol{x}) ||\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{v}(\boldsymbol{x})||^2 d\boldsymbol{x}. \tag{3.11}$$

A local weight term $w(\boldsymbol{x})$ is included in the penalty term. $w(\boldsymbol{x})$ should be large at the structure mismatching area where the deformation should rely more on the

prior deformation knowledge, and small at the area where intensity information is rich in both images.

3.3.2   Similarity Measure for Deformable Registration

To further refine the registration results, deformable registration is performed after rigid-body registration. MI [124] is widely used in the field of multi-modal image registration. Readers can refer to Chapter 2 of the thesis for the review of MI-based image registration techniques. However, in the MI setup, every pixel is treated equally, regardless of the importance in geometric location or structural information. We propose a novel intensity-based similarity metric — structure-encoded mutual information, which assigns different weights to the pixels according some anatomical prior knowledge. Here, we combine the prior deformation field as described in the previous section and the novel structure-encoded mutual information (described in the following section) for the following energy functional:

$$E(\boldsymbol{h}) = E_{\mathrm{SMI}}(\boldsymbol{h}) + \alpha E_{\mathrm{Prior}}(\boldsymbol{h}), \tag{3.12}$$

where E_{SMI} denotes the structure-encoded mutual information of the observed and target data, and E_{Prior} denotes the similarity measure of the current deformation field $\boldsymbol{u}$ and the prior deformation field $\boldsymbol{v}$ which is generated by optimizing Equation 3.11. A deformation field that maximizes the SMI is influenced by the prior deformation field to achieve a more clinically meaningful alignment. The factor α controls the amount of guidance by the prior deformation field. It can be seen that our similarity measure considers both the low-level

information from the image content and the high-level prior on the deformation field. The combination of the two soft constraints integrates the anatomical information which largely helps the registration process. Compared to other registration methods that use some image salient features or boundary features as additional constraints, our proposed method does not require numerical solutions to optimize the energy function. The optimum of the similarity measure can be found by means of variational calculus. We derive the gradient of the combined functional with respect to the displacement field, thus the computational time is largely reduced which is very important for clinical usage. The details of structural encoded mutual information and regularization are given in the following subsection.

3.3.2.1   Structure-Encoded Mutual Information

Encoding location/structural information into the MI similarity measure is investigated by Suh et al. [150] where they only consider the statistics in the region of interest. Different from [150], we emphasize the structure information contained in each voxel. In particular, we propose to assign different weights to the pixels according to some anatomical prior knowledge, e.g., structures that appear in both images should be given higher weights, such as organ boundaries which can be easily extracted if segmentation is readily available. Lower weights are given to the pixels in the homogeneous region where not much information was contained or in the area with mismatching structures, which may lead to significant registration errors. Specifically, our density estimator is based

on a normalized Gaussian kernel of variance β, noted $G_\beta(\boldsymbol{i})$:

$$P(\boldsymbol{i}, \boldsymbol{u}) = \frac{1}{|\sum_\Omega D(\boldsymbol{x})|} \int_\Omega D(\boldsymbol{x}) G_\beta(\boldsymbol{I}_u(\boldsymbol{x}) - \boldsymbol{i}) d\boldsymbol{x}, \qquad (3.13)$$

where $\boldsymbol{i}$ represents the co-occurrence of an intensity pair, $D(\boldsymbol{x})$ is a weighting factor assigned to every pixel. For each pair of intensities $\boldsymbol{i}$, the value of the estimated joint pdf is a nonlinear function of $\boldsymbol{u}$. Note that in the conventional way of calculating the joint density, $D(\boldsymbol{x})$ is assigned to 1. However, we introduce a weighting factor so that we can incorporate the structural information into the calculation of the joint pdf, e.g., the pixels that are closer to the object borders will be given higher weights because the intensity-based SMI term is more capable of matching such discriminant areas.

The definition of the SMI is the same as MI which takes the following form:

$$E_{\text{SMI}} = \int_{R^2} P(\boldsymbol{i}, \boldsymbol{h}) \log \frac{P(\boldsymbol{i}, \boldsymbol{h})}{p(i_1)p(i_2)} d\boldsymbol{i}, \qquad (3.14)$$

where $p(i_1)$ and $p(i_2)$ are the marginal probability density function of the target and source images respectively. Compared to the conventional MI, we encode the anatomical segmentation information into the calculation of joint density function, where pixels that are around the segmentation borders are given higher weights because of its reliability. Such a weighting scheme implicitly makes the SMI content-aware — the similarity measure is dominated by the pixels that are more anatomically reliable. The deformation field therefore is driven more by the areas that are regarded as more reliable, making the optimization less prone to noise and local optima, and therefore leading to more accurate registration

results.

3.3.3 Optimization

For the optimization step, we follow the variational framework proposed by Hermosillo et al. [96] that has demonstrated nice properties in terms of accuracy and capture range compared to the parametric deformable registration. In particular, we perform variational minimization of the proposed energy functional by gradient descent:

$$\frac{\partial \boldsymbol{h}}{\partial t} = -\frac{\partial E(\boldsymbol{h})}{\partial \boldsymbol{h}} = -\frac{\partial E_{\text{SMI}}(\boldsymbol{h})}{\partial \boldsymbol{h}} - \alpha \frac{\partial E_{\text{Prior}}(\boldsymbol{h})}{\partial \boldsymbol{h}}. \tag{3.15}$$

Based on (3.15), the deformation field $\boldsymbol{h}$ is updated at every iteration. The gradient of SMI is similar to the gradient of MI [96] with an additional term on the structure weight term $D(\boldsymbol{x})$, and can be shown to be:

$$\frac{\partial E_{\text{SMI}}}{\partial \boldsymbol{h}(\boldsymbol{x})} = \frac{D(\boldsymbol{x})}{|\sum_{\Omega} D(\boldsymbol{x})|} G_{\beta} * \left(\frac{\partial_2 P(\boldsymbol{i},\boldsymbol{h}(\boldsymbol{x}))}{P(\boldsymbol{i},\boldsymbol{h}(\boldsymbol{x}))} - \frac{p'(i_2,\boldsymbol{h}(\boldsymbol{x}))}{p(i_2,\boldsymbol{h}(\boldsymbol{x}))} \right) \nabla I_2^{\sigma}(\boldsymbol{x} + \boldsymbol{h}(\boldsymbol{x})), \tag{3.16}$$

where $\partial_2 P$ denotes the partial derivative of P with respect to its second variable, p' is the derivative of the marginal pdf, and $*$ is the convolution operator.

Following the notation in [96], the gradient of $E_{\text{Prior}}(\boldsymbol{h})$ is derived from:

$$\delta_{\boldsymbol{k}} E_{\text{Prior}}(\boldsymbol{h}) = \int_{\Omega} \frac{\partial}{\partial \epsilon} w(\boldsymbol{x}) ||\boldsymbol{h}(\boldsymbol{x}) + \varepsilon \boldsymbol{k}(\boldsymbol{x}) - v(\boldsymbol{x})||^2 |_{\epsilon=0} d\boldsymbol{x}$$

$$= \int_{\Omega} w(\boldsymbol{x})(2(\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{v}(\boldsymbol{x})) \cdot \boldsymbol{k}(\boldsymbol{x})) d\boldsymbol{x}, \tag{3.17}$$

where $\delta_{\boldsymbol{k}} E_{\text{Prior}}(\boldsymbol{h})$ is the first variation of $E_{\text{Prior}}(\boldsymbol{h})$ at $\boldsymbol{h}$ in the direction of $\boldsymbol{k}$ (see

[151]).

Thus we have

$$\frac{\partial E_{\text{Prior}}(\boldsymbol{h})}{\partial \boldsymbol{h}(\boldsymbol{x})} = 2w(\boldsymbol{x})(\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{v}(\boldsymbol{x})). \tag{3.18}$$

From (3.16) and (3.18), we immediately notice the term $\left(\frac{\partial_2 P(\boldsymbol{i},\boldsymbol{u}(\boldsymbol{x}))}{P(\boldsymbol{i},\boldsymbol{h}(\boldsymbol{x}))} - \frac{p'(i_2,\boldsymbol{h}(\boldsymbol{x}))}{p(i_2,\boldsymbol{h}(\boldsymbol{x}))}\right)$ and $(\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{v}(\boldsymbol{x}))$ as the comparison functions of our registration method. $\frac{\partial_2 P(\boldsymbol{i},\boldsymbol{h}(\boldsymbol{x}))}{P(\boldsymbol{i},\boldsymbol{h}(\boldsymbol{x}))}$ tends to cluster the joint histogram $P(\boldsymbol{i},\boldsymbol{h})$, while $-\frac{p'(i_2,\boldsymbol{h}(\boldsymbol{x}))}{p(i_2,\boldsymbol{h}(\boldsymbol{x}))}$ prevents the marginal distribution of the deformed source image from becoming too clustered. For $(\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{v}(\boldsymbol{x}))$, the deformation field $\boldsymbol{h}$ is lead by $\boldsymbol{v}$, thus $\boldsymbol{h}$ will not deviate significantly from the prior deformation field $\boldsymbol{v}$. In addition, the use of the weight terms $D(\boldsymbol{x})$ and $w(\boldsymbol{x})$ leads to desirable properties while updating the deformation field at each iteration. Specifically, in the locations where $D(\boldsymbol{x})$ is large, the deformation field will be dominated by the SMI term, because in these regions, there is clear structure information in both source and target images. Thus the registration benefits from the intensity information. While in the regions with larger $w(\boldsymbol{x})$, these regions typically contain nonradiopaque structures in the target images and mismatched structures between the source and target images occur. A larger $w(\boldsymbol{x})$ imposes a stronger prior deformation knowledge at location $\boldsymbol{x}$ to avoid anatomically implausible deformation. The combination effect from $D(\boldsymbol{x})$ and $w(\boldsymbol{x})$ ensures that our deformable registration is context-aware and well regularized.

3.4 Experiments

To validate the proposed framework, here, we present three sets of experiments, on 1) pre-operative CT and non-contrast-enhanced C-arm CT, 2) myocardial perfusion MRI, and 3) simulated pre- and post- liver tumor resection MRI.

3.4.1 Pre-operative CT and Non-contrast-enhanced C-arm CT

Registration of pre-operative contrast-enhanced CT and non-contrast-enhanced C-arm CT eliminates the need for acquiring contrast-enhanced C-arm CT, which is harmful to trans-catheter aortic valve implantation (TAVI) patients with kidney impairments [152]. We validated our proposed method on 20 TAVI patients who had undergone both CT and contrast-enhanced C-arm CT scans. The 20 C-arm CT images are with standard quality. The size of each slice in a volume is 256×256 or 512×512 pixels. A volume contains around 100 - 300 slices. The image resolution is isotropic and varies from 0.49 to 0.92 mm^3. We artificially removed the contrast in the aorta area for C-arm CT by replacing intensities corresponding to the contrast agent with intensities generated from a Gaussian distribution with the mean equal to the heart area of the C-arm CT volume. The generated data are visually inspected to be non-distinguishable from the real non-contrast enhanced C-arm CT volumes (Fig. 3.2). Thus in the experiments, we are essentially matching the CT volume with non-contrast-enhanced C-arm CT volume. The experimental setup takes advantage of the known ground truth. Although artificially generated non-contrast enhanced C-arm CT image may not be an ideal experiment setup, it is so far the best way to make quantitative evaluation on the performance of the proposed registration algorithm for real

45
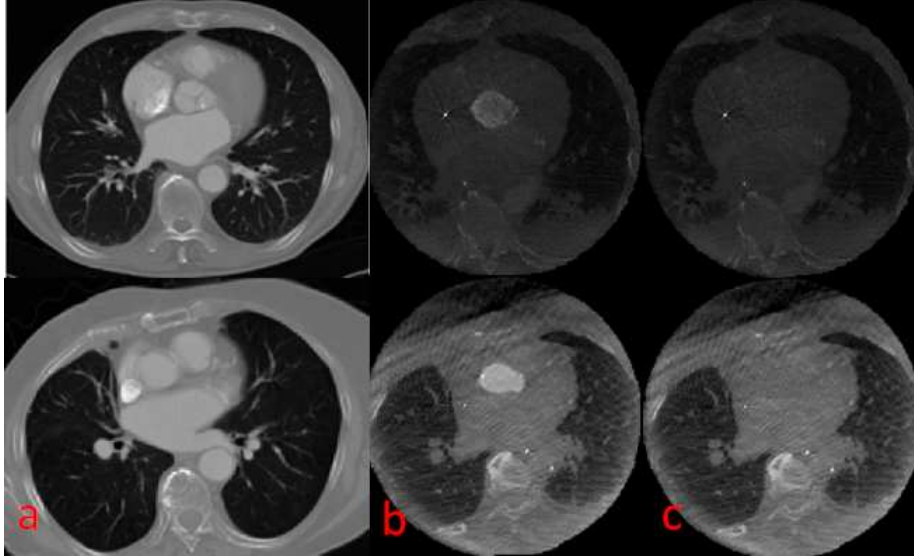
Figure 3.2: Pre-operative CT, intra-operative contrast-enhanced C-arm CT and Simulated non-contrast-enhanced C-arm CT image examples from two patients. Column (a): Pre-operative CT. Column (b): Intra-operative contrast-enhanced C-arm CT. Column (c): Simulated non-contrast-enhanced C-arm CT.

patient data. Compared to possibly alternative validation methods on animal / cadaver data, the deformation property of the heart is more realistic for the target (live human being) application. In addition, using metallic markers to generate ground truth positions raises the issue of the interference of the markers on registration performance due to their high gradients in the images. For all the experiments, the parameters are fixed and without tuning for the best performance for individual data. For quantitative evaluation purpose, a cardiologist manually delineated the aortic root of all the CT and C-arm CT data sets.

3.4.1.1 Qualitative Evaluation on Artificial Non-Contrast Enhanced C-arm CT

To evaluate the deformable registration results, we further compared the results obtained from rigid-body registration initialized using the heart center (for the three cases where both initialization methods failed, we used the manual rigid registration results), conventional MI-based deformable registration under variational framework [96], purely relying on our proposed deformation prior, Lu et al.'s method [152], and the proposed method. Our earlier method in [152] only incorporates the rigid nature of the spine while generating the deformation field prior. More specifically, the point sets are sampled from the spine area, whereas in this thesis, we extend the sample points to the heart surface area to better model the heart motion, thus giving a more effective deformation field prior. For evaluation purposes, we extract surface meshes of the aortic roots from the deformed CT and the C-arm CT images (see Fig. 3.4). In our experiment, the segmentation of the aortic roots is done by a cardiologist, however, it can also be achieved automatically by using the method in [153]. The mesh-to-mesh distance is calculated by the average distance from the points on the aortic root surface mesh root of the deformed CT to the closest point on the aortic root surface mesh of the C-arm CT. The mesh-to-mesh errors are 3.22 ± 1.14 mm, 3.26 ± 1.60 mm, 2.05 ± 0.67 mm, 2.16 ± 0.64 mm, and 1.76 ± 0.43 mm, respectively. The comprehensive comparison can be found in Fig. 3.3.

The mesh-to-mesh error from the rigid registration is larger than 3 mm on average, which exceeds the tolerance in the practical requirement. The result from the rigid registration indicates that the residual motion of the heart still ex-

ists after rigid registration. Therefore, deformable registration should be applied to compensate for the residual motion.

However, using pure intensity-based conventional MI does not reduce the amount of registration error on average, adversely, it introduces a much larger variance. A closer look at every individual cases reveals that pure intensity-based MI reduces the registration error in majority of the cases comparing to rigid registration. But unfortunately, pure intensity-based MI produces significant errors for the cases in which C-arm CT has bad quality. Without incorporating any anatomical information, pure intensity-based MI generates deformation fields that are anatomically implausible, which leads to significant registration errors (See Fig. 3.5 column (b)). As robustness is extremely critical for medical applications, and there is no guarantee to acquire C-arm CT data with good image quality, pure intensity-based conventional MI cannot be applied in the practical situation due to its unreliability.

On the other hand, we also tested the reliability of our proposed deformation field prior. Without incorporating any of the intensity information, we directly apply our proposed prior deformation field on the pre-operative CT data. It can
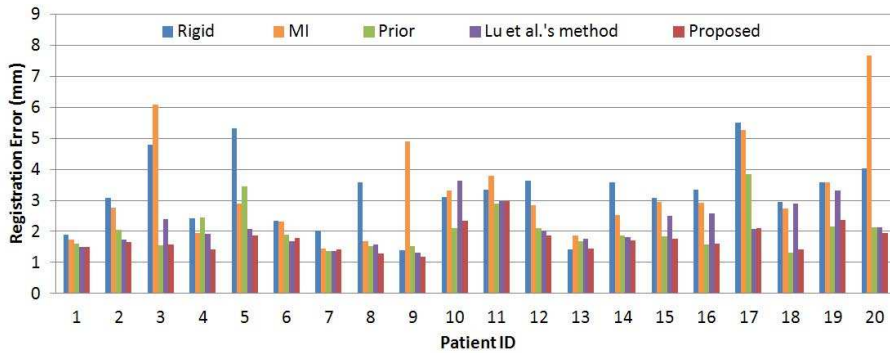
Figure 3.3: Registration performance of 20 patients measured using mesh-to-mesh error.
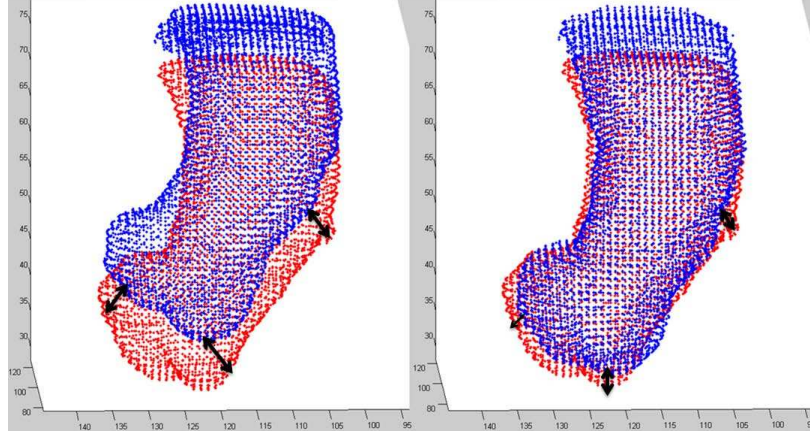
Figure 3.4: Point sets extracted from the aortic root surface, before (left) and after (right) deformable registration. Red point set is the ground truth, and blue point sets are extracted from pre-operative CT. The black arrows demoindicate the errors calculated at the three corresponding points.

be seen that our proposed prior deformation field produces robust and reasonable registration results. This is because the proposed deformation field prior is derived based on the global motion of the heart, thus the result obtained using the prior directly is quite robust. A comparison is made on the results obtained from solely the deformation field prior and the proposed method where intensity information is further incorporated through SMI. Paired t-test between these two methods results a two-tailed P-value of 0.0293. By conventional criteria, this difference is considered to be statistically significant, indicating that the proposed method is statistically significantly better than applying the prior deformation field alone. The improvement can be explained as the followings. The deformation field prior is derived using the method of point set registration, which is less sensitive to points at the area with sharp changes. Therefore, the deformation at those areas may not be fully captured by the deformation prior. To demonstrate this, one registration example from Patient 5 is shown in Fig.
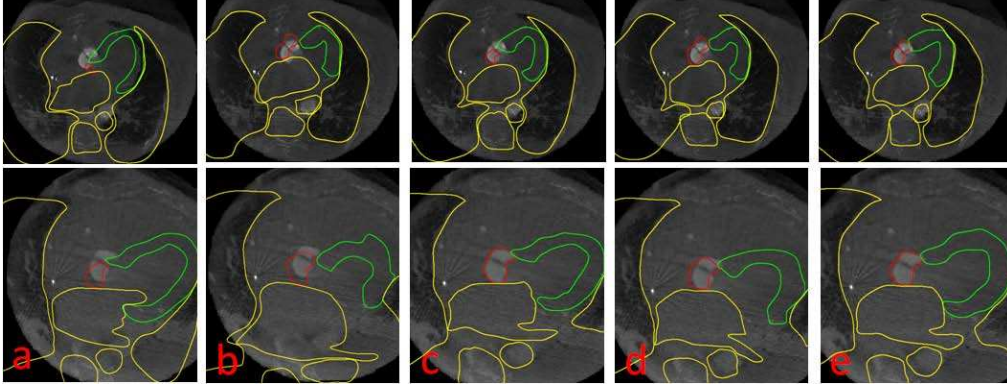
Figure 3.5: The registration results from Patients 5 (Row 1) and 9 (Row 2). (a) Rigid. (b) Deformable using MI. (c) Directly applying prior deformation field. (d) Lu's method (e) The proposed method. The red lines delineate the aortic root, the green lines delineate the myocardium and the yellow lines delineate the other visible structures from the CT images.

3.5, Row 1. In this case, the deformation field prior is not well generated mainly because point sets registration using TPS model does not do well in the region with sharp structures, e.g., the sharp thoracic cavity (because the image was acquired at the inspiratory phase). Furthermore, although the prior deformation field provides a high level prior knowledge about how the heart deforms, inaccurate estimation may occur due to the errors from the lung segmentation. One such case is shown in Fig. 3.5, Row 2. In this case, because of the inaccurate segmentation on the right lung, solely applying the deformation prior would cause the epicardium area to be misaligned and thus leads to inaccurate registration result in the aortic root area. Despite the above mentioned shortcomings of the deformation field prior, with further incorporating the SMI term, the misaligned area can be well corrected which leads to a more desirable registration result. Therefore, the aortic root is registered to a more optimum position.

Our proposed method is closely related to Lu et al.'s method [152] which combines strength from intensity-based method and the anatomical knowledge

into the deformable registration framework. Compared to [152]'s method where only rigid nature of the spine is incorporated, our proposed method further constrains the deformation field by using a TPS model to estimate the heart motion. With the least registration error and standard deviation, the result demonstrates the superiority of the proposed similarity measure which combines soft constraints derived from the anatomy, as well as the numerical stability of our gradient descent based optimization scheme. Clinically, a registration error below 2.5 mm is deemed acceptable. Compared to Lu et al.'s method, we improve the results for Patient 3, 15 and 16 from borderline acceptable to very accurate, and furthermore, the results for Patient 10, 18 and 19 are improved from clinically not acceptable to acceptable. We further perform a paired t-test between these two methods, and the two-tailed P value equals to 0.0012, showing that the proposed method is statistically significantly better than Lu et al.'s method. This is largely attributed to the proposed deformation prior, which is able to model the deformable heart motion, instead of simple rigid-body motion in the spine area as proposed in [152]. It is noted that in certain cases, the improvement of the proposed method is minor compared to [152]. This is because in the cases where heart motion is not significant, modeling the spine motion is sufficient for the registration task, while in the cases where heart motion becomes apparent, the proposed method provides significantly improved registration results.

To qualitatively compare our method with Lu et al.'s method, we show another registration example from Patient 9 in Fig. 3.5, Row 2. In this case, the C-arm CT image is very noisy especially at the area near the spine. The proposed method produces the most accurate registration result at the targeted area – the aortic root (red contours). Furthermore, the anatomical structure at the

51

heart area is nicely preserved, due to the incorporated deformation prior. We can see that MI-based method fails completely because of the large area of mismatched structures, while Lu et al.'s method performs well around the spine and heart surface (yellow contours). However, because the aortic root is quite far away from both the heart surface and the spine area, the constraints imposed by [152] is not strong enough to produce a good registration result. Furthermore, due to the significant noise, the registration result at the heart area is not clinically meaningful, e.g., the myocardium (green contours) is badly distorted.

A nice property of our proposed similarity measure is the balance between E_{SMI} and E_{prior}. At the initial stages of the registration process, E_{prior} is the main driving force because initially, the deformation field is very different from the deformation prior, thus the magnitude of E_{prior} is large. As the registration process continues, the E_{SMI} term acts to fine tune the registration results by utilizing the intensity information. And E_{prior} constrains the registration process by penalizing significant deviation from the prior. Note that these two terms are both soft constraints, and the combination of these constraints nicely present the best registration results.

	Aortic Ostia		
	Mean	STD	Median
After Rigid-body Registration	6.78	7.19	7.32
After Deformable Registration	2.86	0.77	2.73

Figure 3.6: The left and right coronary ostia at the aortic valve of two example data: (a) C-arm CT image (b) Pre-operative CT image. The table on the right shows the landmark registration error between the registered coronary ostia in the CT image to the corresponding points in the C-arm CT image. The mean, standard deviation (STD), and median of the errors are reported (measured in millimeters).

Registration accuracy in the aortic valve landmark detection is measured to further validate the clinical applicability of the proposed registration framework. During the TAVI procedure, the left and right coronary ostia are overlaid onto 2-D fluoroscopic images to provide guidance to physicians to avoid blocking the ostia after valve deployment. Thus accurate registration of left and right coronary ostia is crucial for the TAVI procedure. The landmark registration accuracy of the coronary ostia is measured using the Euclidean distance from the coronary ostia in the C-arm CT image to the corresponding coronary ostia in the registered CT image. We can see that the landmark registration accuracy of the coronary ostia is significantly improved from 6.78 ± 7.32 mm to 2.86 ± 0.77 mm. The average landmark registration error of below 3 mm indicates the clinical feasibility of our proposed registration framework. Figure 3.6 shows the coronary ostia from C-arm CT and pre-operative CT respectively.

3.4.1.2 Qualitative Evaluation on Real Non-Contrast Enhanced C-arm CT

We also perform our proposed registration framework on three sets of CT/real non-contrast enhanced C-arm CT data. The specs of the three data sets are similar to those in Section 3.1.1. The non-contrast C-arm CT images do not have any visible structure at the targeted area of this application — aorta. Furthermore, as mentioned in Section 3.1.1, the only way to determine positions on non-contrast-enhanced C-arm CT image is to place metallic markers to obtain the ground truth. However, doing so raises the issue of the interference of the markers on registration performance due to their high gradients in the images. Therefore, we are not able to provide quantitative comparison, and only quali-

tative evaluation is performed.

It is found that the experiment results are very similar to the results obtained by using artificially generated non-contrast enhanced C-arm CT. This validates that the quantitative evaluation in the Section 3.1. We qualitatively evaluate our image registration framework on CT and real non-contrast enhanced C-arm CT. An example is shown in Figure 3.7. Firstly, rigid-body registration brings the pre-operative CT to the same coordinate system as C-arm CT image. However, major misalignment still exits mainly due to the cardiac and respiratory motion (see the red arrows in Figure 3.7 Row 2). After performing deformable registration, the cardiac and respiratory motion is corrected, and the aortic root is registered to a more optimum location (Figure 3.7 Row 3). In this example, we show that the proposed image registration framework successfully registers the CT image to the non-contrast enhanced C-arm CT image. In particular, the spine is correctly registered with its rigid nature embedded in the deformable registration process. Furthermore, the heart structure is nicely preserved and the aortic root is registered to a optimum position, which is clinically applicable, thanks to the anatomic knowledge based deformation field prior.

3.4.2 Myocardial Perfusion MRI

We perform our second set of experiment on 8 myocardial perfusion MRI sequences. The data was acquired by Siemens Sonata/Avanto MR scanners following bolus injection of gadolinium-DTPA contrast agent. Due to the intensity change caused by the contrast enhancement, registration of myocardial perfusion MRI is treated as multi-modal image registration.
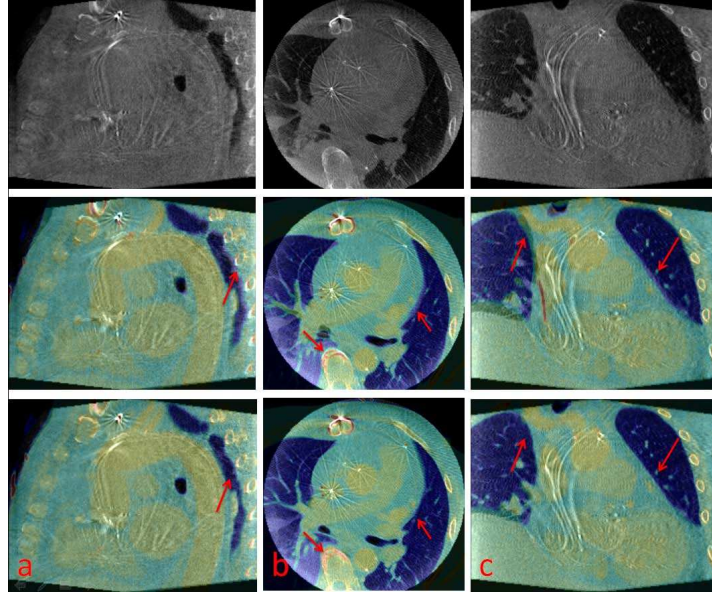
Figure 3.7: Qualitative evaluation on image registration of CT and real non-contrast enhanced C-arm CT. Row 1: Non contrast-enhanced C-arm CT. Row 2: After rigid-body registration. Row 3: After deformable registration.

3.4.2.1 Experimental Setup.

We select a source frame which has the best contrast in the sequence, and the selected floating frame is registered to every frame of the sequence. In this experiment, we can obtain the epicardium segmentation using [154]. The point sets are sampled from the epicardium outline. $D(\boldsymbol{x})$ is larger at the locations near edges, and $w(\boldsymbol{x})$ is larger at the locations near the segmented epicardium. The information of epicardium segmentation is thus implicitly embedded into the registration process.

3.4.2.2 Experiment Results

For our data set, myocardial contours (epicardium and endocardium) of all the slices are drawn by a cardiologist. These contours serve as the ground truth.
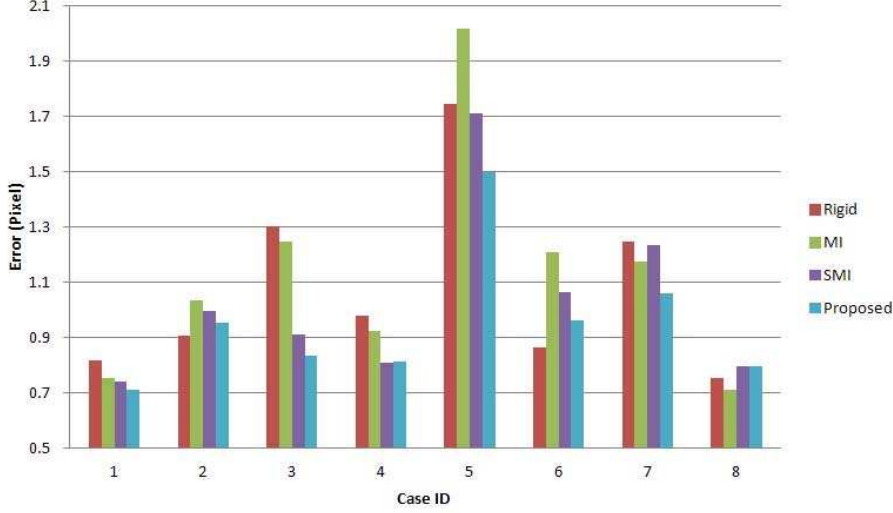
Figure 3.8: Quantitative comparison of the registration errors (in *pixel*) obtained by rigid registration, MI, SMI and the proposed method.

We calculated the root mean square distance from the ground truth to the propagated contours (yellow contours in Fig. 3.9). The paired t-test indicates that our hybrid method is statistically significantly better than the intensity-based methods with P values equaling to 0.039 and 0.0263 when compared to MI and SMI respectively. In particular, for the example shown in Fig. 3.9), simple warping using the segmentation information results in noticeable registration errors at the structure-rich areas as the intensity information is ignored. Intensity-based registration does not perform well in the homogeneous area because of the lack of structure information. However, combining the strength of both intensity-based and segmentation-based methods, our hybrid method produces the best result. It is clearly shown that our proposed method provides the most robust registration result, especially when the initial alignment is unsatisfied, e.g. Case 5. Thanks to the guidance from the high-level knowledge prior deformation field, the registration problem is much better constrained and a more optimal solution can
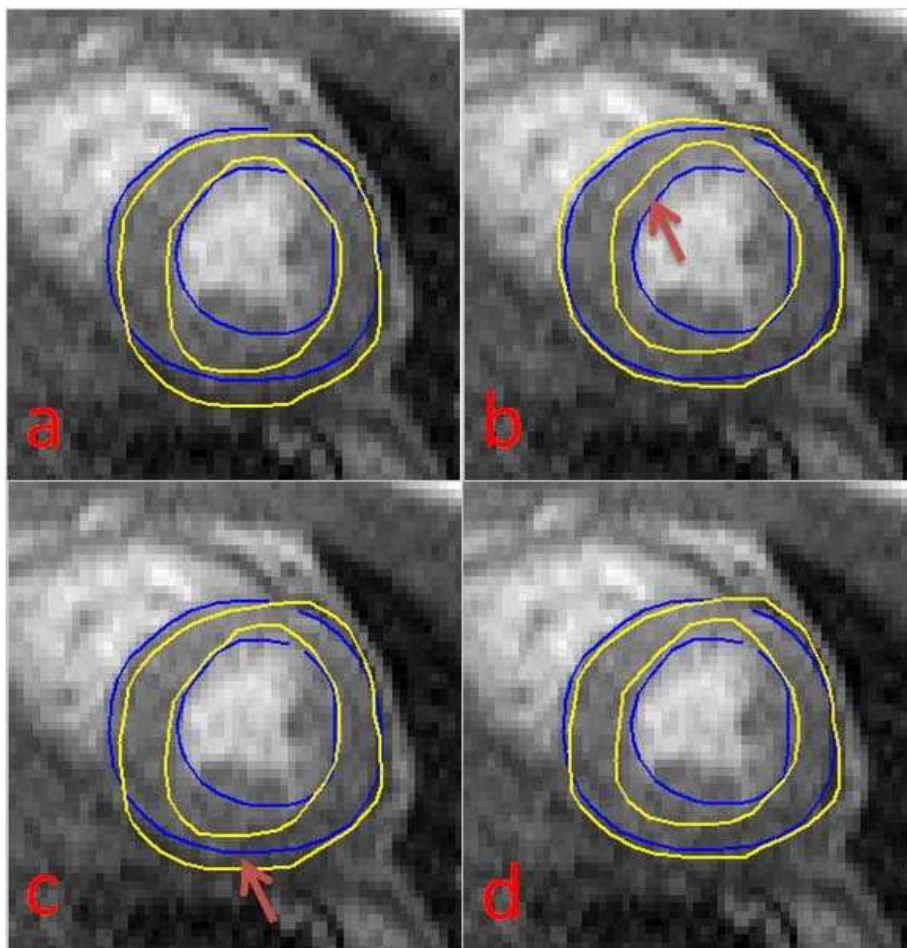
Figure 3.9: Registration results (a) Rigid (b) Simple warping using segmentation information (c) SMI (d) Proposed method. Yellow and blue lines are the propogated and the ground truth contour.

be found. Note that in certain cases, e.g., Case 4 and Case 8, the difference is insignificant with and without the proposed deformation field prior, this is because in these cases, less motion of at the myocardium area is observed, thus the intensity based SMI is sufficient to register these image sequences even without incorporating the deformation field prior.

3.4.3   Simulated Pre- and Post- Liver Tumor Resection MRI

The proposed hybrid method could be potentially applied to another category of registration problems with mismatching structures, i.e. registration between volumes of pre- and post- tumor resection. In this experiment, the registration is performed on pre-operative MRI and simulated post-operative MRI.

3.4.3.1   Experimental Setup

We simulated a post tumor resection image by manually segmenting the tumor and the surrounding tissues, and then replacing the segmented area with 0 intensities. Then we artificially deform the pre-operative MRI, and the registration is performed between the deformed pre-opeartive MRI and the simulated post-operative MRI. $D(\boldsymbol{x})$ is set to one except for the resected area and $w(\boldsymbol{x})$ is one at the resected area and zero otherwise. We assume the liver segmentation is available, the point set is extracted from the liver surface.

3.4.3.2   Results

Here we get the preliminary results using one data set. Qualitatively, intensity-based registration does not perform well in the resected area, and simple warping using the liver segmentation does not preserve the detailed structures well. The proposed hybrid method guides the registration using the deformation prior at the resected area, while at the rest of the area, intensity-based method dominates. By combining the strength of both, the hybrid method achieves the best registration result as demonstrated in Fig. 3.10.
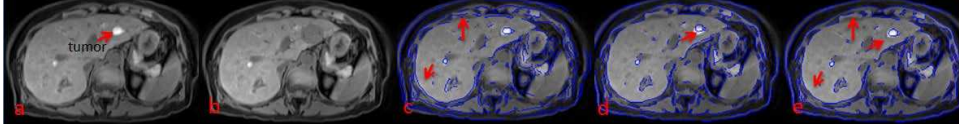
Figure 3.10: (a) Pre-operative MRI. (b) Simulated post-operative MRI. (c), (d) and (e) are the registration results using simple warping, SMI and our method respectively.

3.5 Conclusion

In this chapter, we presented a hybrid multimodal deformable registration framework with incorporates two sources of anatomical prior information. The proposed method addresses the image registration issue while dealing with images with different structure appearance due to different levels of contrast medium, and was validated on both TAVI, perfusion MR data. In addition, preliminary results show that the proposed method can also be applied to registration of pre- and post- tumor resection images. The experimental results are encouraging which demonstrates its superiority compared to intensity-based method and simple warping using segmentation. Furthermore, we descend the gradient of the combined functional with respect to the deformation field under the variational framework which is more computationally efficient compared to spline-based optimization schemes.

Chapter 4

Multi-modal Image Registration: Utilizing Appearance Priors

Incorporating context-specific prior knowledge of the intensity mapping has shown promising registration results. In this chapter, we propose to learn the intensity matching information through normalized pointwise mutual information, from existing perfectly aligned training images or image pairs that are roughly registered. Then novel similarity measures — weighted mutual information and weighted entropy of intensity mapping confidence map are proposed. The details of the algorithm will be presented in the following sections.

4.1 Introduction

As mentioned in Chapter 1, mutual information is widely used in the field of multi-modal image registration. Nevertheless, incorporating prior knowledge of the underlying registration problem has shown significant improvement in

registration robustness and accuracy. Hence, it is common to incorporate prior information into the energy optimization scheme by the following formula:

$$J(T, S \circ W) = J_{MI}(T, S \circ W) + \alpha J_{Prior}(T, S \circ W), \qquad (4.1)$$

where the similarity measure J consists of the mutual information based similarity J_{MI} and a prior similarity term J_{Prior}.

There is much prior information that can be incorporated into the image registration process. The effort of intensity standardizing over different modalities makes incorporating intensity matching prior more practical. More specifically, using divergence measures to compare the joint intensity joint distributions from training and testing cases has attracted much attention. [104], [105] and [106] proposed to use Kullback-Leibler divergence (KLD) as the similarity measure to register multi-modal images. Liao et al. used Jensen-Shannon divergence (JSD) to compare learnt joint distribution with the observed joint distribution [107]. The major drawback of using the leant joint intensity distributions is that these joint intensity distributions may deviate a lot between the training and testing images, which reduces the effectiveness of the intensity matching prior. This is because the joint intensity distribution depends not only on the intensity matching relationship, but also the quantity of the intensity matching pairs. The number of the intensity matching pairs may vary significantly from training cases to testing cases. In this chapter, we propose to learn the intensity matching relationship from pre-registered image training pairs. Instead of using the learnt joint distributions, we propose to learn the intensity matching relationship from the joint histogram through the normalized pointwise mutual information

(NPMI), and apply the learnt intensity matching as a more general, flexible and robust prior. The motivation comes from the observation that the joint distributions not only account for how the intensities are matched, but also the amount of each intensity pair. The acquired intensity matching prior is thus sensitive to field-of-view and size of the objects, while the statistically significant intensity mapping relationship remains unchanged. For example, in Fig. 4.2, the joint intensity distributions change a lot from slice to slice for brain images, while the proposed learnt intensity matching relationship based on NPMI is largely the same.

Compared to direct application of the joint distributions, the proposed method further utilizes the intensity mapping information. Using NPMI as the measure of intensity matching relationship makes the proposed algorithm more robust to changes in both the object and the background; the latter may even become the dominant force in the learnt joint distributions due to its dominant size in many applications. More importantly, we can obtain our intensity matching prior from a subset or even just a slice of the volume, provided that the slice is representative of the object to be registered. This makes the acquisition of the prior much simpler. In this chapter, two novel similarity metrics based on weighted mutual information and weighted entropy are proposed.
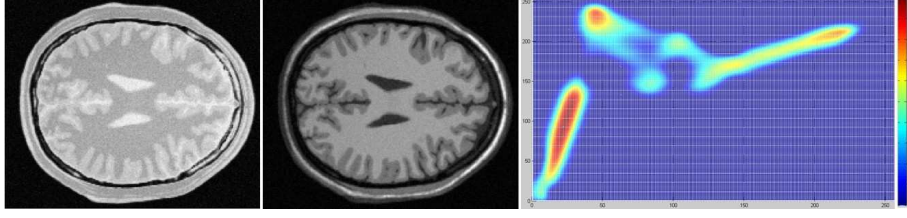
Figure 4.1: Two corresponding PD/T1 brain MRI slices and the computed NPMI. The red value shown in the NPMI map shows high correlation between the intensity pairs.

4.2 Method

4.2.1 Normalized Pointwise Mutual Information

Normalized pointwise mutual information (NPMI) is widely used in the text categorization field [155]. It ranges from -1 to 1 where a positive value indicates the trend of appearing together and a negative value indicates the trend of not appearing together, 0 indicates statistical independence. NPMI is defined as:

$$\text{NPMI}(i_1, i_2) = \frac{\log \frac{p(i_1, i_2)}{p(i_1)p(i_2)}}{-\log(\max(p(i_1), p(i_2)))}, \tag{4.2}$$

where in the application of image registration, i_1 and i_2 are the intensity values from the image pair, $p(i_1)$ and $p(i_2)$ are the marginal distributions of the training image pair and $p(i_1, i_2)$ is their joint distribution. Figure 4.1 shows an example the learnt NPMI based on two corresponding PD/T1 brain MR images.

NPMI reflects the correlation between the intensity values from source and target images. The low frequencies bias is much reduced by normalizing with the factor $\log(\max(p(i_1), p(i_2)))$. Therefore, it provides more insight regarding the correlation between the values from two sources. The properties of NPMI nicely fit into the application of image registration, especially multi-modality
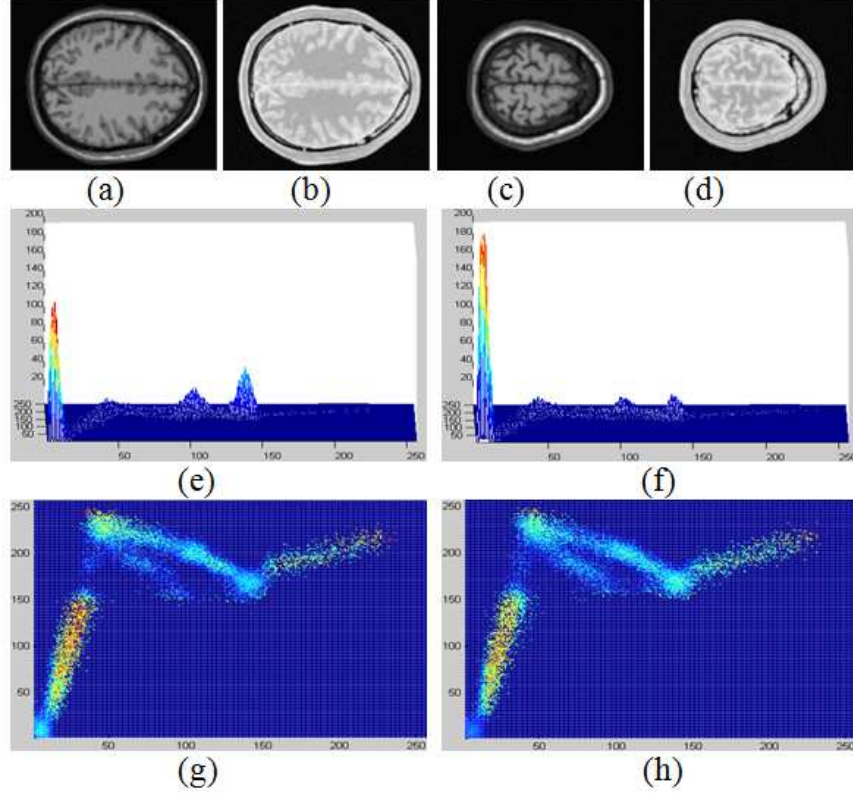
Figure 4.2: Different training slices may result in different joint histograms but similar intensity matching relationship. (a) (b) A training pair of brain image (T1/PD). (c) (d) another training pair of brain image. (e) (f) the resulting learnt joint histograms from pair (a) (b) and (c) (d) respectively. (g) (h) the resulting learnt intensity matching prior from pair (a) (b) and (c) (d) respectively.

image registration where one to many and many to one intensity matchings are possible. To be more specific, in the field of medical image analysis, two regions may have the same mean intensity in one modality but totally two different intensities in another modality. The NPMI mainly reflects the intensity matching information, because the bias that would have been introduced from the quantity of the intensity values has been normalized (see Fig. 4.2). To capture the regional intensity mappings more accurately, we calculate the NPMI at every patch of the training images (the patch size may depend on the application), and

then obtain the final NPMI by choosing the maximum value among all the patch pairs. In the context of image registration, when two intensity values always appear at the same time, NPMI will be given a value of 1, the less correlation the two intensities have, the smaller value NPMI would be given. We do not consider the intensities that have negative correlation due to their rarity.. The training images may be obtained from the prior image pairs that have already been carefully aligned. Alternatively, as in this chapter, the training images are obtained by MI-based registration rather than from prior image pairs.

The NPMI obtained from the training process has the capability to model the intensity relationship from the source and target images. Compared to the joint intensity histogram as the intensity matching prior, the influence from the number of intensity matching co-occurrence has been minimized, thanks to the normalization factor. Furthermore, as NPMI is much less affected by the size of the object and the proportion of the background, the training process for NPMI is more straightforward. In many cases, we can easily extract the NPMI information from only few slides of the image volume as long as the intensity matching in these slices are representative for the entire volume. NPMI is a more elegant intensity matching prior compared to the joint intensity histogram, because of its simplicity in the training process, and robustness and accurateness as an intensity matching prior. In the next subsection, we will propose two novel similarity measures that use the learnt NPMI for image registration.

4.2.2 Weighted Mutual Information

4.2.2.1 Formation of WMI

In the traditional formulation of mutual information:

$$J_{\text{MI}} = \int_{R^2} p_{\boldsymbol{h}}(i_1, i_2) \log \frac{p_{\boldsymbol{h}}(i_1, i_2)}{p_{f_1}(i_1)p_{f_2}(i_2)} di_1 di_2 \qquad (4.3)$$

every intensity pair (i_1, i_2) is weighted by the corresponding co-occurrence probability $p(i_1, i_2)$. This assumes that apart from the probability of occurrence, the intensity pairs are treated equally. However, such an assumption is not completely true in the course of image registration. Some intensity pairs are semantically more important than the others, because certain intensities from source and target images are inherently related to each other, regardless of the probability of occurrence, e.g., the thin structures in the images, although the probability of occurrence is low, the correlation of the intensity matching pair is very high.

To address this issue, we propose to incorporate the learnt intensity matching prior into the MI framework. The weighted mutual information (WMI) is thus defined as:

$$J_{\text{WMI}} = \int_{R^2} p_{\boldsymbol{h}}(i_1, i_2) \log \frac{p_{\boldsymbol{h}}(i_1, i_2)}{p_{f_1}(i_1)p_{f_2}(i_2)} w(i_1, i_2) di_1 di_2, \qquad (4.4)$$

where $w(i_1, i_2)$ is the weight learnt from the training data for each intensity pair. When $w(i_1, i_2)$ is assigned to 1 for every intensity pair, WMI degenerates to the conventional MI where every intensity pair is treated equally. However, it can be easily understood that, certain intensity pairs should play a more important part in the image registration process, e.g., intensities that are highly correlated. A

higher weight is given to the intensity pairs that are more highly correlated while calculating the weighted mutual information. The correlation of the intensity pairs can be learnt from the training data sets. In particular,

$$
w(i_1, i_2) = \begin{cases} 1 + \text{NPMI}(i_1, i_2) & \text{NPMI}(i_1, i_2) > 0 \\ 1 & \text{NPMI}(i_1, i_2) \leq 0 \end{cases} \tag{4.5}
$$

The weight $w(i_1, i_2)$ indicates the matching relationship between the two intensities i_1 and i_2. Intensity pairs that are positively correlated will result in higher NPMI values and thus the higher weights. On the other hand, if the intensities are negatively correlated or independent, a low weight will be given. By giving weights to the intensity pairs, weighted mutual information explicitly encodes the intensity matching information into the formation of the original mutual information. Therefore, unlike mutual information, which is an unsupervised similarity measure, the proposed weighted mutual information makes use of the prior information obtained from the training data set to facilitate the image registration process. As the intensity matching information has been incorporated by assigning different weights to different intensity pairs, the weighted mutual information is theoretically superior in describing the image alignments compared to mutual information. Naturally, it may lead to a better image registration result. Note that, the weight $w(i_1, i_2)$ is always greater than or equal to 1, so the weighted mutual information in fact combines the conventional mutual information and the contribution from the learnt intensity matching prior.

4.2.2.2 Probabilistic Interpretation Using Bayesian Inference

In probability terms, minimization of our energy function in Equation (4.8) can be interpreted as maximizing the posterior distribution of the diffeomorphic warping $\boldsymbol{h}$, given the target image T, the source image S, and the learnt intensity mapping M:

$$
\begin{aligned}
\boldsymbol{h} &= \arg\sup_{\boldsymbol{h}} \log p(\boldsymbol{h}|T, S, M) \\
&= \arg\sup_{\boldsymbol{h}} \log(p(T, S, M|\boldsymbol{h})p(\boldsymbol{h})).
\end{aligned}
\tag{4.6}
$$

The second term, $\log p(\boldsymbol{h})$, can be recognized as the geometric prior on the displacement field, and in our model, it is considered as generic regularization. In our implementation, we apply a fast filtering technique to the deformation field at each iteration to regularize the deformation. Hence we focus on the first term $\log p(T, S, M|\boldsymbol{h})$. It can be shown from Bayesian inference theory that:

$$
\begin{aligned}
p(T, S, M|\boldsymbol{h}) &\propto \int_{R^2} p(T, S, M|\boldsymbol{h}, \boldsymbol{i})p_{\boldsymbol{h}}(\boldsymbol{i})d\boldsymbol{i} \\
&\propto \int_{R^2} p(T, S|M, \boldsymbol{h}, \boldsymbol{i})p(M|\boldsymbol{h}, \boldsymbol{i})p_{\boldsymbol{h}}(\boldsymbol{i})d\boldsymbol{i},
\end{aligned}
\tag{4.7}
$$

where $\boldsymbol{i} = (i_1, i_2)$ is a given intensity mapping pair in the R^2 domain. Proportionality in the above equation means that only those factors that do not depend on the deformation field $\boldsymbol{h}$ and thus do not affect the maximization are neglected. Thus the probability calculation separates into three parts and can be interpreted as follows: the first term, $p(T, S|M, \boldsymbol{h}, \boldsymbol{i})$, measures the similarity of the two images given one intensity mapping pair, and in our model can be recognized as the term $E^{\mathrm{MI}}(\boldsymbol{i}, \boldsymbol{h})$ in Equation (4.10); the second term, $p(M|\boldsymbol{h}, \boldsymbol{i})$, denotes the

learnt prior probability for a given intensity mapping pair, and in our method is modeled using the weight $w(\boldsymbol{i})$ defined in Equation (4.5). When the prior on the intensity mapping is not available and hence $p(M|\boldsymbol{h},\boldsymbol{i})$ is the same (i.e. $w(\boldsymbol{i})=1$) for all the intensity pairs, the above equation boils down to the conventional mutual information similarity measure. The conditional distribution is integrated along the dimension of possible intensity mapping pairs via the third term $p_{\boldsymbol{h}}(\boldsymbol{i})$.

4.2.2.3 Optimization of Variational Formulation

Based on the proposed WMI, our learning-based registration scheme is defined as the minimization of the following energy functional with respect to the deformation field $\boldsymbol{h}$:

$$J(\boldsymbol{h}) = -J_{WMI} \tag{4.8}$$

For the optimization step, we follow the variational framework proposed by Hermosillo et al. [96]. In particular, we adopt the variational minimization of the energy functional by gradient descent:

$$\frac{\partial \boldsymbol{h}}{\partial t} = -\frac{\partial J(\boldsymbol{h})}{\partial \boldsymbol{h}} = \frac{\partial J_{\mathrm{WMI}}(\boldsymbol{h})}{\partial \boldsymbol{h}} \tag{4.9}$$

The remaining challenge is to calculate the gradient of the data term J_{WMI}. For simplicity, we use the notation $\boldsymbol{i} = (i_1, i_2)$, $\boldsymbol{I_h}(\boldsymbol{x}) = (I_1^\sigma(\boldsymbol{x}), I_2^\sigma(\boldsymbol{x} + \boldsymbol{h}(\boldsymbol{x})))$, and define:

$$J_{\mathrm{MI}}(\boldsymbol{i}, \boldsymbol{h}) = -\log \frac{p(\boldsymbol{i}, \boldsymbol{h})}{p(i_1)p(i_2, \boldsymbol{h})} \tag{4.10}$$

We estimate the joint density estimation by:

$$p(\boldsymbol{i}, \boldsymbol{h}) = \frac{1}{|\Omega|} \int_\Omega G_\beta(\boldsymbol{I_h}(\boldsymbol{x}) - \boldsymbol{i}) d\boldsymbol{x}, \qquad (4.11)$$

where G_β is a normalized Gaussian kernel of variance β, and

$$\frac{\partial p(\boldsymbol{i}, \boldsymbol{h})}{\partial \boldsymbol{h}} = \frac{1}{|\Omega|} \int_\Omega \partial_2 G_\beta(\boldsymbol{I_h}(\boldsymbol{x}) - \boldsymbol{i}) \nabla I_2^\sigma(\boldsymbol{x} + \boldsymbol{h}(\boldsymbol{x})) d\boldsymbol{x}, \qquad (4.12)$$

where ∂_2 denotes the partial derivative of $G_\beta(\boldsymbol{I_h}(\boldsymbol{x}) - \boldsymbol{i})$ with respect to its second variable. It can be shown that the gradient of the data term E_{WMI} can then be simplified to:

$$\frac{\partial J_{\text{WMI}}}{\partial \boldsymbol{h}} = -\int_{R^2} \frac{\partial}{\partial \boldsymbol{h}}(p_{\boldsymbol{h}}(\boldsymbol{i}))(J_{\text{MI}}(\boldsymbol{i}, \boldsymbol{h}) w(\boldsymbol{i})) d\boldsymbol{i} \qquad (4.13)$$

By substituting (8-9) into (10) with some algebraic manipulation, we obtain

$$\frac{\partial J_{\text{WMI}}}{\partial \boldsymbol{h}} = -\{G_\beta * [(\partial_2 w(\boldsymbol{i})) J_{\text{MI}}(\boldsymbol{i}, \boldsymbol{h})] + G_\beta *$$
$$[(\frac{\partial_2 p(\boldsymbol{i}, \boldsymbol{h})}{p(\boldsymbol{i}, \boldsymbol{h})} - \frac{p'(i_2, \boldsymbol{h})}{p(i_2, \boldsymbol{h})}) w(\boldsymbol{i})]\} \nabla I_2^\sigma(\boldsymbol{x} + \boldsymbol{h}(\boldsymbol{x})), \quad (4.14)$$

where $p(i_2, \boldsymbol{h})$ is the marginal distribution of the transformed source image and $p'(i_2, \boldsymbol{h})$ is the derivative of $p(i_2, \boldsymbol{h})$, * is the convolution operator. We applied a fast filtering technique [156] to the deformation field at each iteration to regularize the deformation.

4.2.3   Weighted Entropy of Intensity Mapping Confidence Map

We now use the Weighted Entropy of Intensity Mapping Confidence Map to utilize the learnt NPMI. It consists of two steps: first, obtaining an intensity matching confidence map; second, calculating the weighted entropy of the intensity matching confidence map.

4.2.3.1   Intensity Matching Confidence Map

For any given pair of images, there is a corresponding intensity matching confidence map. The intensity matching confidence map represents the NPMI value of the corresponding pixels between the target and (transformed) source images (Fig. 4.3). Let IMCM denote the intensity matching confidence map. IMCM is defined as:

$$\text{IMCM}(\mathbf{x}) = \text{NPMI}(T(\mathbf{x}), (S \circ W)(\mathbf{x})) \tag{4.15}$$

where we assign each point $\mathbf{x} \in \Omega$ an intensity matching value based on NPMI, and thus form an intensity matching confidence map. If we make an analogy with mono-modal images, intensity matching confidence map can be viewed as the difference map in the multi-modal set-up. In the mono-modal setup, we can easily create the difference map by subtracting the source image from the target image. However, it is not as straightforward for the multi-modal setup. We obtain the intensity matching confidence map by incorporating the information from the learnt NPMI. At every location $\mathbf{x} \in \Omega$, the intensity matching confidence map reflects the matching score of the corresponding intensity pair $(T(\mathbf{x}), S(\mathbf{x}))$, and thus is capable of describing the accuracy of the image alignment in the multi-modal setup. The intensity matching confidence map

71

naturally incorporates intensity mapping where a higher value in the confidence map indicates higher probability of correct registration. Furthermore, a smooth confidence map indicates a better registration, because abrupt changes should only occur at the places where tissue type has changed and thus the significant change of the intensity.

To gain a deeper understanding of the intensity matching confidence map, we can make an analogy to the intensity difference map in the mono-modal setup. In the mono-modal setup, the intensity difference map is obtained by simply subtracting the source image from the target image. Thus each pixel of the intensity difference map reflects the image registration accuracy of the source and target images — the larger the difference, the worse the matching result. SSD is designed based on the assumption that the sum of the squared difference map should be minimized to achieve the best registration accuracy. SSD provides decent image registration results when such an assumption is not violated. However, in the real situation, the intensity differences of pixels cannot be exactly zero, especially in the situation where the lighting condition has changed, and the assumption of applying SSD does not hold thus resulting in unsatisfactory image registration. To increase the robustness and accuracy when utilizing the image difference map for registration purposes, [157] proposed to use Shannon entropy of the image intensity difference map as a similarity measure, which is shown to be more accurate and robust compared to SSD, MI and NMI in the mono-modality situation [157].

The intensity matching confidence map is closely related to the intensity difference map, but it works even in the multi-modal setup (see Fig. 4.3). Each individual pixel of the intensity matching confidence map reflects the matching

confidence at that particular location. It is therefore natural to design a similarity measure based on the sum of the intensity matching confidence map values. However, it is easy to realize that summing up the intensity matching confidence map values as the similarity measure may be affected by the training accuracy of the NPMI, as well as the noise level of the source and target images. These are the same reasons that cause SSD to be defective in the mono-modal setup. Inspired by the entropy-based similarity measure [157], we here propose to utilize intensity matching confidence map using weighted entropy, and thus make the entropy-based similarity measure applicable in the multi-modal setup.
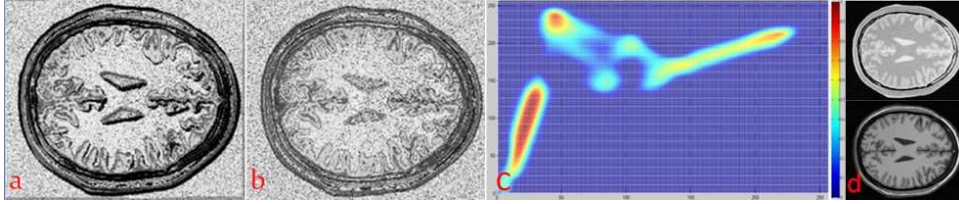
Figure 4.3: (a)Intensity matching confidence map before image registration, the black area indicates low matching confidence which is a sign of mis-alignment. (b)Intensity matching confidence map after registration where high matching confidence value is across the map. (c) NPMI obtained from the training data set. (d) Training images.

4.2.3.2 Weighted Entropy

To utilize the information provided by the intensity matching confidence map, we propose to use weighted entropy of the intensity matching confidence map to measure the accuracy of the image registration results. Weighted entropy is the measure of information supplied by a probabilistic experiment whose elementary events are characterized both by their objective probabilities and by some

quantitative weights [158]. It is defined as:

$$J_{\text{Prior}} = -\int_R w(y) P_{\text{IMCM}}(y) \log P_{\text{IMCM}}(y) dy. \qquad (4.16)$$

In our formulation, y is the value from every pixel of the intensity matching confidence map; the weighted term $w(y)$ is larger for larger y, $P_{\text{IMCM}}(y)$ is the histogram of the intensity matching confidence map. Specifically, we estimate the histogram based on a normalized Gaussian kernel of variance β, denoted by $G_\beta(y)$:

$$P_{\text{IMCM}}(y, \boldsymbol{h}) = \frac{1}{|\Omega|} \int_\Omega G_\beta(\text{NPMI}(T(\boldsymbol{x}), (S \circ W)(\boldsymbol{x})) - y) d\boldsymbol{x}. \qquad (4.17)$$

The formulation is very similar to the well-known Shannon entropy, except for the newly introduced weighting factor $w(y)$. Shannon entropy is not suited to our application, because by minimizing the Shannon entropy of the intensity matching confidence map, the value of the confidence map would shrink to 0, which leads to a total mis-alignment. Therefore, we define the weighting factor $w(y)$ as:

$$w(y) = \tanh((i - a)/b), \qquad (4.18)$$

where $\tanh$ is the hyperbolic tangent, b defines the smooth transition region, and a defines the switch point of the IMCM. For value i that is greater than a in the IMCM, we consider it as a good matching. The proposed weighting factor $w(y)$ differentiates the intensity matchings. It gives positive values to intensity matchings that are likely to appear according to the training data set, while penalizing the intensity matchings that are not likely to appear in the training data set by

giving them negative weights. Unlike the Heavyside step function, Equation 4.18 is differentiable, and the transition region controlled by b provides a buffer region, in case the switch point a is not accurately defined. Furthermore, $w(y)$ allows weighted entropy to be biased towards certain events that carry higher weight, in our application, the intensity matching confidence map has higher values; in contrast, the heavyside step function gives only two flat values for the entire domain. For our image registration problem, a good image registration produces an intensity matching confidence map that is smooth and with high values. From the mathematical perspective, this requires the weighted entropy of the intensity matching confidence map to be small given that the weight for the higher values are larger. Therefore, weighted entropy is well suited for the problem.

4.2.3.3 Optimization of Variational Formulation

Our deformable registration scheme is defined as minimizing the sum of the mutual information and the weighted entropy of the intensity matching confidence map with respect to the deformation field $\boldsymbol{h}$:

$$J(\boldsymbol{h}) = J_{\mathrm{MI}}(\boldsymbol{h}) + J_{\mathrm{Prior}}(\boldsymbol{h}). \tag{4.19}$$

The gradient of $J_{\mathrm{MI}}(\boldsymbol{h})$ at each location $\boldsymbol{x}$ is calculated as

$$\nabla J_{\mathrm{MI}}(\boldsymbol{h}, \boldsymbol{x}) = L_{J_{\mathrm{MI}}}(I_1(\boldsymbol{x}), I_2(\boldsymbol{x}, \boldsymbol{h})) \nabla I_2^{\sigma}(\boldsymbol{x} + \boldsymbol{h}(\boldsymbol{x})), \tag{4.20}$$

where

$$L_{J_{\mathrm{MI}}}(I_1(\boldsymbol{x}), I_2(\boldsymbol{x}, \boldsymbol{h})) = \frac{1}{|\Omega|} G_\beta * \left(\frac{\partial_2 P(\mathbf{i}, \boldsymbol{h})}{P(\mathbf{i}, \boldsymbol{h})} - \frac{p'(i_2, \boldsymbol{h})}{p(i_2, \boldsymbol{h})}\right), \tag{4.21}$$

with p being the marginal distribution of the respective image. Here, we further derive the first variation of the J_{Prior} at $\boldsymbol{h}$:

$$\delta_{\boldsymbol{k}} J_{\mathrm{Prior}}(\boldsymbol{h}) = w(y)(1 + \log P(y))\delta_{\boldsymbol{k}} P(y), \tag{4.22}$$

where

$$\begin{aligned}
\delta_{\boldsymbol{k}} P(y) = \int_\Omega \partial G_\beta(\mathrm{NPMI}(I_1(\boldsymbol{x}), I_2(\boldsymbol{x}, \mathbf{h})) - y) \frac{\mathrm{NPMI}(I_1(\boldsymbol{x}), I_2(\boldsymbol{x}, \boldsymbol{h})) - y}{\beta} \\
\times \partial_2 \mathrm{NPMI}(I_1(\boldsymbol{x}), I_2(\boldsymbol{x}, \boldsymbol{h})) \nabla I_2^\sigma(\boldsymbol{x} + \boldsymbol{h}(\boldsymbol{x})) \cdot \boldsymbol{k}(\boldsymbol{x}) d\boldsymbol{x},
\end{aligned} \tag{4.23}$$

$\partial_2 \mathrm{NPMI}$ denotes the partial derivative of NPMI with respect to its second variable. The gradient of $J_{\mathrm{Prior}}(\boldsymbol{h})$ at each location $\boldsymbol{x}$ is calculated as:

$$\nabla J_{\mathrm{Prior}}(\boldsymbol{h}, \boldsymbol{x}) = L_{J_{\mathrm{Prior}}}(I_1(\boldsymbol{x}), I_2(\boldsymbol{x}, \boldsymbol{h})) \nabla I_2^\sigma(\boldsymbol{x} + \boldsymbol{h}(\boldsymbol{x})), \tag{4.24}$$

where

$$\begin{aligned}
L_J(I_1(\boldsymbol{x}), I_2(\boldsymbol{x}, \boldsymbol{h})) = \\
\frac{1}{|\Omega|} G_\beta * (w(y(\mathbf{x}))(1 + \log P(y(\mathbf{x})))) \partial_2 \mathrm{NPMI}(I_1(\boldsymbol{x}), I_2(\boldsymbol{x} + \boldsymbol{h}(\boldsymbol{x}))),
\end{aligned} \tag{4.25}$$

4.3 Experiments

We evaluate the proposed registration method through five sets of experiments. In Section 4.3.1, a phantom study is performed to demonstrate the flexibility

and robustness of utilizing the normalized pointwise mutual information as the intensity matching prior. Section 4.3.2 shows that even with partial occlusions and background change, the proposed method still provides the flexibility of object-specific image registration. In Section 4.3.3, a quantitative and qualitative comparison is performed in synthetic brain MR images. We use the knowledge of the ground truth of the deformation field and the correspondence of the land marks to perform the evaluation.

4.3.1 Synthetic Image Study

Utilizing the intensity matching prior can guide the image registration process to converge to the desired optimum point. To demonstrate, we have an ambiguous setup (Figure 4.4) such that there are two optimum alignments. In Figure 4.4, (a) (b) and (c) are the synthetic images of the target image, the source image, and the contour of the source image overlaid onto the target image before image registration. Without incorporating any prior knowledge, the circle in the source image will match to the outer circle of the target image, by applying mutual information as the similarity measure (Figure 4.4(d)). On the contrary, by changing the prior intensity matching information, both weighted mutual information and the proposed method with weighted entropy of the IMCM as the prior can flexibly match the circle in the source image to either the outer or inner circle of the target image (see Figure 4.4(e) and (f)). This experiment shows the importance of incorporating context-specific prior information, which helps to lead to the desire alignment.

A comparison between our proposed method and [105]'s method is further

Figure 4.4: (a) Target image. (b) Source image. (c) Contour of the source image overlaid onto the reference image before registration. (d) Registration result using MI. (e) (f) Registration result using the proposed method with different matching profiles. For (d) (e) (f), green line indicates the contour of the source image after registration.

performed, to test the robustness of the training prior with respect to the change in the size of the object. We increase the radius of the circles in both the target and source images, while the prior remains. Due to the change of the radius, the joint histogram of the target and source images is no longer the same as the training prior joint histogram; thus using KL divergence to measure the deviation of the test and training joint histogram is inappropriate. The experiment result also validates the issue we raised above, the method in [105] converges to the undesired global optima, where the circle in the source image is matched to the outer circle of the target image. This is because the prior learnt from the image with a certain object size cannot be generalized to help register the
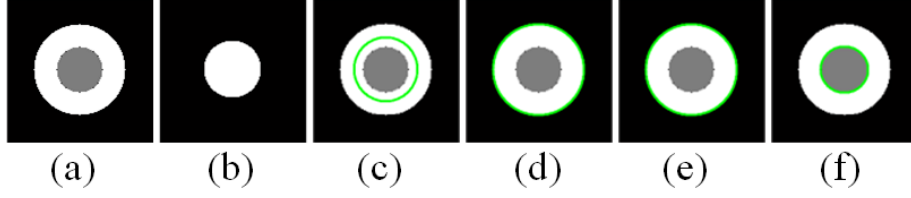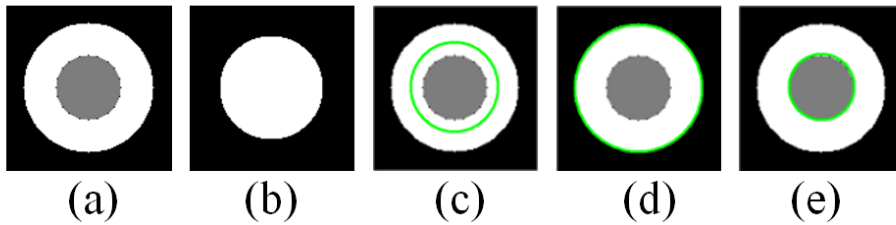

Figure 4.5: (a) Target image. (b) Source image. (c) Contour of the source image overlaid onto the target image before registration. (d) Registration result using the method in [105]. (e) Registration result using the proposed method. In (d) (e), green line indicates the contour of the source image after registration.

object with a slightly different size. On the contrary, our proposed method still correctly registers the circle in the source image to the inner circle of the target image (Figure 4.5(e)).

4.3.2 Face Images with Occlusion and Background Changes

In this experiment, we show that the proposed method which incorporates the intensity matching prior improves the registration result significantly in the presence of occlusion, change of lighting condition and background. We aim to show that the proposed method can be applied to general applications, rather than limited to the field of medical image where most of the images in this thesis are from. We use the data sets as been employed in [105], where the source image is taken under different lighting conditions with the person wearing sun glasses. The first two images of Figure 4.6 show a pair of manually registered training data used to learn the intensity matching prior. The two training images were captured under different lighting conditions, and with the person in the source image wearing a sunglasses. To demonstrate the power of context-specific capability of our proposed algorithm, only the area with the person occupied is used for training, without including the white background. In the registration process, we change the background of both the source and the target images, and the entire image with the added background is used in registration. We compared the registration results using the proposed methods (both WMI and weighted entropy) with conventional MI, and the method proposed in [105] which incorporates the trained joint histogram as the prior information. The registration results are shown in Figure 4.7, where we superimpose the edge maps

79
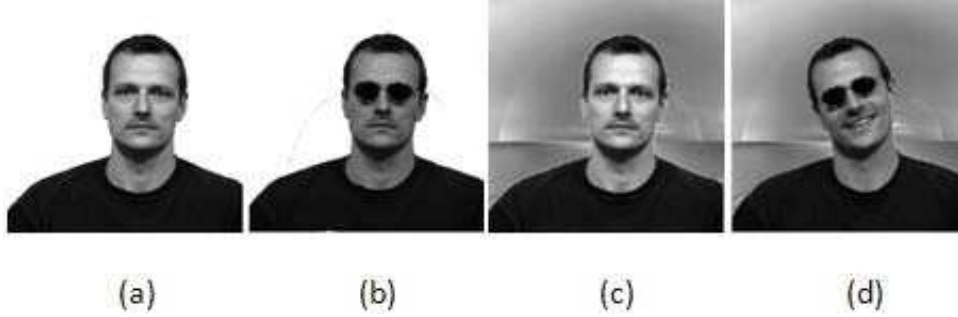
Figure 4.6: Face images used for training and registration. (a) (b), training images. (c) (d) target and source images used for registration, with the addition of the different backgrounds.

of the registered source images on the target images for better visualization of the registration results. The image registration results show that conventional MI fails for all of the three cases possibly because the identical background in the source and target images makes the mutual information less sensitive in the foreground change, in these cases, the human. In the second example, although MI is able to match the outline of the person correctly, the details in the human face, especially region around the sun glasses, is largely misaligned, due to lack of prior matching information. Incorporating the intensity matching prior using learnt joint histogram [105] helps to align the outline of the person correctly in the first and third cases, but fails to align the sun glasses with the eye region. The drop in performance of the method [105] in this case is expected and simple to interpret. Firstly, the intensity matching knowledge of the changing background is not in the prior, and thus using the global intensity joint histogram as the prior is inaccurate. Secondly, global intensity joint histogram is less sensitive to the local changes, in this case, the sunglasses, where it only occupies a small portion of the image. The joint histogram prior is not be able to guide the misalignment on the sunglass area due to the fact that it only contributes a very small amount

in the calculation of the joint histogram. The above mentioned two factors make the effectiveness of the prior weaker. On the other hand, our proposed methods are able to detect the region that we want to register and thus in all the three cases, our proposed method reaches a perfect alignment for the human region.

This example demonstrates the advantage of using intensity mapping over joint histogram. Using the joint histogram directly as the prior fails when the background changes significantly and/or if the background is not in the training pool, while our proposed method enables the registration process to focus on the region of interest and correctly performs the image alignment in the target area.

Figure 4.7: Three different backgrounds are tested during registration. (a) (b) (c) overlay the edge of the source image to the target image before registration. (d) (e) (f) show the result obtained by conventional mutual information. (g) (h) (i) show the result obtained by the method proposed in [105]. (j) (k) (l) show the result obtained by using WMI. And (m) (n) (o) show the result obtained by using weighted entropy

4.3.3 Simulated MRIs

In Section 4.3.1 and Section 4.3.3.1, we have shown qualitatively that incorporating the appearance prior can significantly increase the image registration robustness and accuracy. Furthermore, in the scenario where size of the objects/backgrounds change significantly, we demonstrated that incorporating NPMI as the intensity matching prior information, is superior than using the joint histogram as the prior [105]. Qualitatively, the two similarity measures we proposed achieve similar registration results in Section 4.3.1 and Section 4.3.3.1. To provide more insights of the proposed similarity measures, in this experiment, we focus on examining the strengths and weakness of the proposed similarity measures.

In this study, we used the simulated MR brain images generated using the BrainWeb MR simulator [159]. A pair of perfectly aligned T1/PD brain data with resolution $1 \times 1 \times 1$ mm^3, noise level 3% and 20% non-uniformity are created. With this data set, we perform two sets of experiment: 1) to examine the similarity measure with respect to translation and rotation; and 2) to study on the deformable registration application.

4.3.3.1 Similarity Measure Comparison

To quantitatively evaluate the effectiveness of our proposed method, we examine the similarity measures, including MI, WMI and weighted entropy, with respect to the horizontal (from -20 mm to 20 mm), and rotational shift (from -30 degree to 30 degree) of the brain MR image T1 over PD.

1) NPMI Learnt From The Exact Alignment
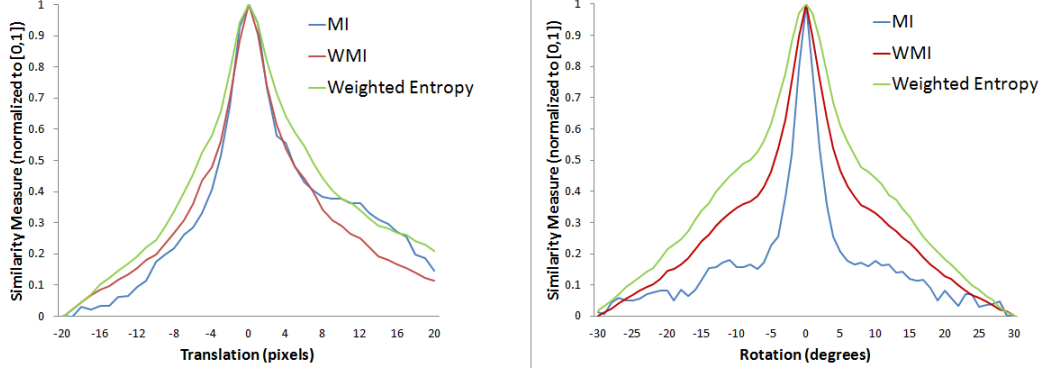
Figure 4.8: Plot of three similarity measures (MI, WMI and weighted entropy with an accurate NPMI) with respect to the translational and rotational shift. Zero translation and rotation corresponds to the perfect alignment.

In this section, the NPMI is learnt from the exact alignment, thus the intensity matching information is well presented. It is shown in Fig. 4.8 that, the three similarity measures all achieve optimum at the correct image alignment. However, the similarity measure curve generated from MI is not smooth while the image is moving away from the optimum alignment. In contrast, for the two proposed similarity measures, the similarity score decays smoothly while the image shifts away. A smooth similarity measure curve indicates that the proposed similarity measures are good measurements of the alignment, because the similarity scores are highly correlated to the amount of misalignment. Furthermore, as we can see from Fig. 4.8, weighted mutual information shows a narrower, but sharper attraction basin to the optimum; this implies that WMI may have a better discrimination power for local misalignments. On the contrary, weighted entropy shows a wider attraction basin to the optimum; this is a strong indication that it has a wider capture range, which is a desirable property especially when the deformation is large.

2) A Shifted NPMI
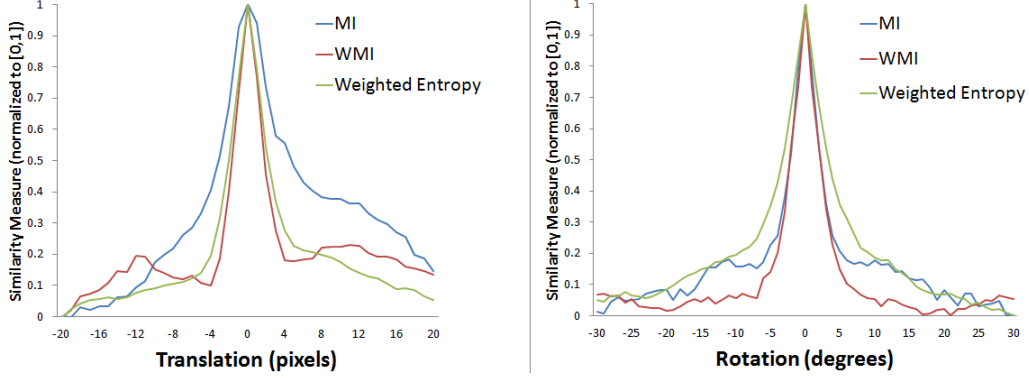
Figure 4.9: Plot of three similarity measures (MI, WMI and weighted entropy with less accurate NPMI) with respect to the translational and rotational shift. Zero translation and rotation corresponds to the perfect alignment.

It is often concerned that the intensity matching learnt from the training images may be different from the testing images. To test how the deviation may affect the similarity measure, in this section, the NPMI is learnt from the exact alignment, but we purposely increase the intensity values of the training images by 10 units, thus the trained NPMI is slightly different from the intensity matching information of the testing images.

It is shown in Fig. 4.9 that the similarity measures of both WMI and weighted entropy are not as smooth as they were shown in Fig. 4.8, mainly due to imperfection of the learnt NPMI. Nevertheless, we can see that the similarity measure of weighted entropy is less affected by the change, it is still monotonically decreasing as the image shifted away from the optimum alignment. In contrast, fluctuation of the similarity measure of WMI can be seen as the image moved away.

3) Discussion

The set of experiments shown in Section 4.3.3.1 demonstrates that, incorporating NPMI as the appearance prior can help the image registration process in

general. However, the choice of WMI and weighted entropy may be application-dependent. WMI is more suitable to the applications that have significant local changes, given that the NPMI is accurately learnt. While weighted entropy is more robust, shown by its wide capture range and less sensitive to the change of NPMI.

4.3.3.2   Deformable Registration Evaluation

To evaluate the deformable registration results, we randomly generate ten artificial deformations using thin-plate splines with a maximum displacement of 10 mm. The artificial deformations are applied to the T1 image for evaluation purpose. For quantitative assessment, we chose 433 brain landmark points at cortex, ventricles etc, with the assistance from an expert. We evaluate the performance using the root mean square error (RMSE) between the transformed position and the corresponding ground truth position. And to examine how an inaccurate NPMI could affect the registration result, we perform the image registration based on: 1) NPMI that was learnt from the exact alignment, and 2) NPMI that has been artificially shifted.

1) NPMI Learnt From The Exact Alignment

The prior is learnt from the intrinsically registered volume pair. Only 10 slices out of a total 181 slices are used for training purpose, while the whole brain volume are tested for registration accuracy.

The quantitative results can be found in Fig. 4.10. The root mean squared errors for MI, WMI and weighted entropy are 2.67 ± 0.75 mm, 1.68 ± 0.16 mm and 1.72 ± 0.21 mm respectively. It is clear to see that the proposed two similarity measures achieve significantly more accurate results and better robustness

Figure 4.10: Quantitative comparison of the registration results obtained by conventional MI, WMI and the proposed weighted entropy by applying ten randomly created deformation fields using TPS. Accurate intensity matching prior information is used.

(i.e., smaller root mean squared error values on average and variance) than the conventional mutual information. To reconfirm, we perform the paired t-tests with the MI. The paired t-test between MI and WMI shows a P-value of 0.0026 which indicates that results obtained from WMI are statistical significantly better than MI. Similarly, the paired t-test between MI and weighted entropy gives a P-value of 0.0020, which reconfirms the improvement of our proposed methods. We further perform the paired t-test between WMI and weighted entropy, although in this experiment, WMI gives smaller root mean squared error and variance, the P-value of 0.19 indicates that the difference between WMI and weighted entropy is considered to be not statistically significant.

An qualitative study can be found in Fig. 4.11. The purple arrows show that the proposed methods are still able to perform well at the area where image

Figure 4.11: Qualitative comparison of the registration results of the MR brain images obtained by (a) conventional MI, (b) WMI and (c) the proposed weighted entropy. Accurate intensity matching prior information is used. The major differences of the registration results are indicated by the arrows.

registration using MI fails significantly. We further compare the result obtained from WMI and weighted entropy, it can be seen that, WMI performs better at the edge area as pointed by the black arrows. Such observation reconfirms the hypothesis we made in Section 4.3.3.1 that WMI has a better discriminant power at the local region, given that the NPMI is accurately estimated.

2) A Shifted NPMI

This experiment aims to investigate the robustness of the proposed similarity measures to the accuracy of NPMI. The prior is still learnt from the intrinsically registered volume pair. Same as the previous section, only 10 slices out of a total 181 slices are used for training purpose, while the whole brain volume are tested for registration accuracy. However, in this time, we artificially increase the intensity value of the training data by 5 units. This creates an less accurate appearance prior compared to the previous section.

The quantitative result is shown in Fig. 4.12. The root mean squared errors for MI, WMI and weighted entropy are 2.67 ± 0.75 mm, 1.93 ± 0.29 mm and 1.78 ± 0.26 mm respectively. The proposed WMI and weighted entropy still outperform the conventional MI. However, it is worth noticing that, the image
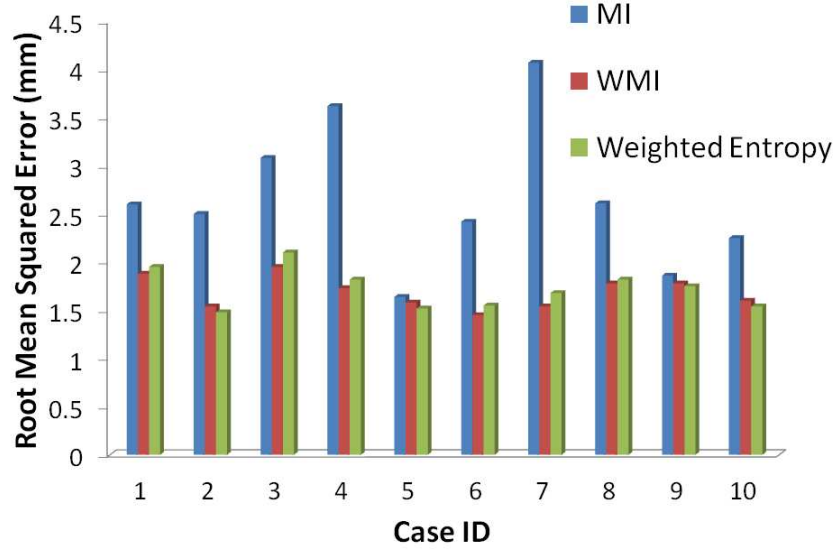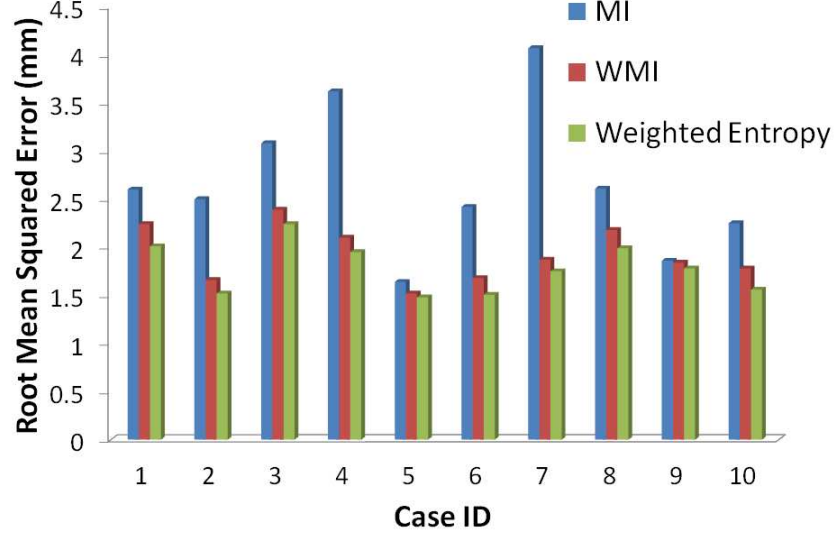
Figure 4.12: Quantitative comparison of the registration results obtained by conventional MI, WMI and the proposed weighted entropy by applying ten randomly created deformation fields using TPS. Shifted intensity matching prior information is used.

registration results produced by both proposed similarity measures deteriorate. In particular, the root mean squared error using WMI increases significantly from 1.68 mm to 1.93 mm, while the change of using weighted is much smaller, from 1.72 mm to 1.78 mm. Again, we perform a formal paired t-test, comparing the proposed methods with MI. The P-values of 0.0065 and 0.0024 show that the proposed methods are still statistically significantly better than the conventional MI. The paired t-test between WMI and weighted entropy indicates that using the weighted entropy as the similarity measure significantly outperforms WMI in this experiment setup.

To qualitatively access the experiment result, an example is provided in Fig. 4.13. We used a purple eclipse to highlight the area where significant area of mis-registration occurs. The mis-registration occurs mainly due to the existence
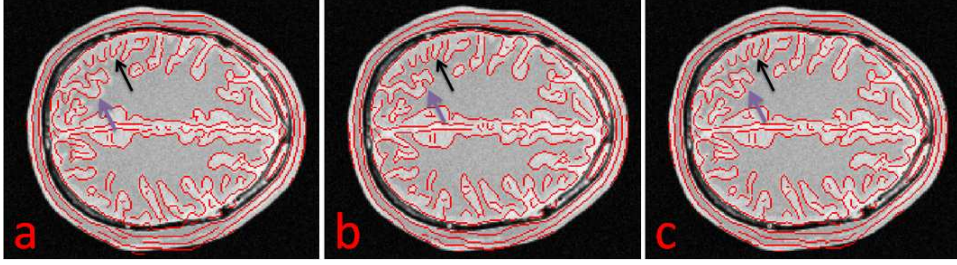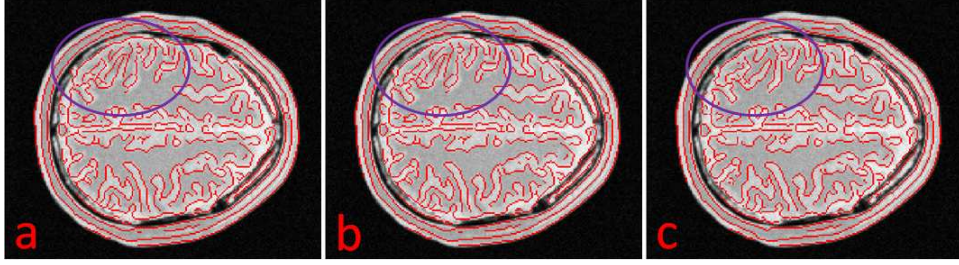
Figure 4.13: Qualitative comparison of the registration results of the MR brain images obtained by (a) conventional MI, (b) WMI and (c) the proposed weighted entropy. Shifted intensity matching prior information is used. The purple circle indicates the area where large misalignment occurs for MI and WMI.

of intensity non-uniformity in this area. MI fails to register at the area where the intensity is faded and becomes more similar to the intensity from the surrounding structure. WMI is not able to discriminate the faded structures, mainly because of the inaccurate appearance prior information. This coincides with the findings in Section 4.3.3.1, where the inaccurate appearance prior brings a rough similarity curve for WMI which has many local optima. In the process of deformable image registration, the local optima leads to premature termination of the optimization process, which leads to the inaccurate image registration results. In this case, the proposed weighted entropy is still capable of registering the misaligned structures that could not be recovered using MI and WMI, thanks to the wider attraction basin as shown in Section 4.3.3.1, which indicates the wide capture range, as well as the robustness of the proposed weighted entropy with respect to NPMI.

3) Discussions

The experiment results shown in Section 4.3.3.2 coincide with the findings from Section 4.3.3.1. Still, the choice of WMI and weighted entropy may be application-dependent. WMI has a more discriminant power on local changes

given that the NPMI is accurate. In the case where the intensity matching relationship obtained from the training sets deviates from the test images, it is best to use weighted entropy which is more robust to the deviation of NPMI.

4.4   Conclusion

In this chapter, we propose to incorporate the appearance prior into the image registration framework, by utilizing the normalized pointwise mutual information. Two similarity metrics based on NPMI are proposed. The problem is extensively studied in the deformable registration framework. We showed that the accuracy and robustness of image registration have been improved significantly when incorporating such prior information. Furthermore, compared to the state-of-the-art methods, the proposed method provides more flexibility and robustness with respect to the change in the global profile of the learnt joint histogram. Although the advantages of the proposed method are mainly demonstrated in the non-rigid image registration setup, it can be easily extended to rigid-body registration.

Chapter 5

Modality Synthesis: From Multi-modality to Mono-modality

Compared to mono-modality image registration, multi-modality image registration is a much harder problem due to the complexity of modeling the intensity matching relationship. Recently, image modality synthesis has attracted much attention. It has been shown that the synthetic images have the potential to reduce the amount of image acquisition to perform certain analysis, e.g., tumor growth. Furthermore, image registration benefits from image modality synthesis as it can reduce the multi-modal image registration problem to a simpler mono-modality registration problem. In this chapter, we propose a general framework for modality synthesis, utilizing the features such as intensity histogram and the Weber Local Descriptor. We show that the proposed modality synthesis method is superior to the state-of-the-art methods.

5.1 Introduction

In the field of image analysis, images from different modalities reflect different characteristics of the underlying anatomy. For example, local bone densities can be shown clearly on CT images while MRI is more suitable for examining organs and soft tissues. Although acquiring images of the same patient from different modalities help the physicians/doctors for anatomical analysis, performing image registration for images from different modalities is not simple. It has been shown that in the case of mono-modal image registration, the optimal similarity criterion exist, which are cross correlation and correlation coefficient. With this in mind, it is natural to think of finding a way to transform the multi-modal image registration problem to a simpler mono-modal image problem for which there exists an optimal similarity criterion. Recently, the interest of finding methods for performing subject-specific synthesis from a given modality to some target modality has drawn a lot of attention.

In this chapter, we aim to develop a modality synthesis technique that will benefit image registration applications. Modality synthesis generally refers to generating different appearances of the same anatomy based on prior knowledge, instead of actual acquisition. It has the potential to solve the dilemma in choosing multiple image acquisitions and time/money costs. With such techniques, we can also transform images from any modalities to the same modality. The image analysis difficulties will become much simpler. Modality synthesis has shown its wide range of applications including multi-modal image registration [112, 160, 161] and segmentation [112, 162]. Most of the above mentioned methods are application-specific, for example, [109] focuses on trans-

forming CT to ultrasound images for image-guided intervention purpose. Ou and Chefd'Hotel [163] assume a polynomial intensity matching relationship between different modalities, but the assumption is not flexible and not always true. To provide a more general framework for modality synthesis, database driven, or exemplar-based methods has be considered [112, 113]. The underlying principle of these methods is to utilize a training database from different subjects where images from different modalities are well aligned. Given any source images, these methods try to synthesize the target images of another modality based on the matching information provided by the database. These methods are particularly of interest because they avoid explicit modeling of intensity matching information and naturally incorporates the spatial context. Therefore, they are very general and produce visually impressive results even with limited size of image databases. Having said that, the exemplar-based methods often suffer from having a large database because computational efficiency increases significantly with database size. This problem is not mentioned in [112, 113], which however is very important, because it relates to the applicability of the proposed methods.

Our proposed approach for modality synthesis belongs to the exemplar-based category. For each point in the source image (modality A), we perform a local patch-based search in the database to estimate the target value for the point in modality B. The proposed method is very closely related to methods in [112, 113]. However, in [112], N most similar candidates were selected and averaged to produce the final estimation which is very time consuming and ineffective. Ye et al. [113] introduced a data-driven regularization technique to provide a more coherent synthesis of different modalities, which is more advanced

compared to averaging the N most similar candidate patches as proposed in [112]. Both [112] and [113] use Sum of Squared Distances (SSD) to measure the distance between the source patch and the patches in the database. SSD is not general enough as an image similarity measure, especially when the lighting condition differs, or the parameters of the medical image changes. And as we will demonstrate later, using SSD as the similarity measure limits the possibility of reducing the size of the database. Therefore, we propose a framework for modality synthesis that uses patch based intensity histogram and Weber local descriptor features, which is more general and discriminative than the existing methods. Compared to SSD, intensity histogram and Weber local descriptor are more suitable for finding the corresponding patch from the database, which leads to a much robust and accurate image synthesis result. Furthermore, we introduce a weighting factor based on normalized pointwise mutual information which can be learnt from the database. Thanks to the robustness of our proposed distance measure, we can apply a technique for database reduction to further decrease the computational time. Our framework potentially outperforms the state-of-the-art modality synthesis methods in terms of robustness, accuracy and computational efficiency.

5.2 Method

The process of modality synthesis refers to the following task: Given a source image S from modality M_a, we aim to generate a corresponding target image T for modality M_b. The target image T is constructed based on the source image S and the population database. In the population database, there are N exemplar

image pairs $(M_{a_n}, M_{b_n})|_{n=1,...,N}$, where every pair is spatially aligned. The idea behind our proposed modality synthesis method is that, by searching for the most similar patch M_{a_m} to P_S, we can automatically find the corresponding patch in modality M_b as M_{b_m}. Therefore, the N exemplar image pairs act as a black box to map any image from modality M_a to modality M_b.

For each image point $\boldsymbol{x}$, $T(\boldsymbol{x})$ is estimated by first finding the patch which is most similar to $P_S(\boldsymbol{x})$ in the database, say $P_{a_m}(\boldsymbol{y})$, then $P_{b_m}(\boldsymbol{y})$ is selected to fill up the corresponding position in the target image $T(\boldsymbol{x})$.

The process of modality synthesis is inspired by [112] and [113], however, we speed up and improve the synthesis quality by introducing: 1) database reduction, and 2) a novel distance measure for patch matching.

5.2.1 Database Reduction

Although a large database provides rich information for the synthesis task, redundant information will however decrease the computational efficiency. To remove the redundant information from the database, we aim to select the most representative exemplars that well represent the data set. By doing that, the searching task will only be performed within a much reduced database, yet the information is still well preserved. Elhamifar et al. [164] proposed to find the representative data points from the data set by finding a subset of the data set that minimizes the reconstruction error of each data point in the data set as a linear combination of the representative data. It is formulated as minimizing the expression:

$$\Sigma_{i=1}^{N} \parallel y_i - Y c_i \parallel_2^2 = \parallel Y - YC \parallel_F^2, \tag{5.1}$$

with respect to the coefficient matrix $C = [c_1, c_2, ..., c_N] \in R^{N \times N}$. Then the nonzeros of C is constrained as

$$\| C \|_{0,q} \leq k. \tag{5.2}$$

$\| C \|_{0,q}$ counts the number of nonzero rows of C, and the indices of the nonzero rows of C correspond to the indices of the columns of Y which are chosen as the data representatives. To reduce the computational complexity, the constraint on C is relaxed to $L1$ norm, thus the optimization is formulated as:

$$\min \| Y - YC \|_F^2 \ s.t. \ \| C \|_{1,q} \leq \tau. \tag{5.3}$$

τ is used instead of k since for the k optimal representatives, $\| C \|_{1,q}$ is not necessarily bounded by k. In our application, for every patch, we stack the intensity and WLD histograms (see detailed description in the next section) to form Y_i. By reducing the size of the data set, the computational efficiency is significantly improved, because they are linearly related.

5.2.2 Modality Synthesis

5.2.2.1 Locality Search Constraint

We further reduce the search space by imposing a locality search constraint. Let's now assume that all the subjects within the database and the source images are linearly registered. To achieve the spatial restriction, for any position $\boldsymbol{x}$, we define a small search window $\mathbf{W}_{\boldsymbol{x}}$ centered at $\boldsymbol{x}$. We only consider the candidates from the database that are within the searching window. Restricting the search

space significantly increases the computational efficiency, especially when the database is large. Although the search space is drastically reduced, the accuracy will not be affected because it is reasonable to assume that the most similar patch is spatially located near $\boldsymbol{x}$.

5.2.2.2 Modality Synthesis Using a Novel Distance Measure

To synthesize $T(\boldsymbol{x})$, we first extract a patch $P_S(\boldsymbol{x})$ centered at $\boldsymbol{x}$ from the source image S. Based on the locality search constraint defined above, we search within the constrained database to find a patch P_{a_n} from the database such that the dissimilarity between $P_S(\boldsymbol{x})$ and P_a is minimized. It can be formulated as:

$$(n^*, \boldsymbol{y}^*) = \arg\min_{n,\boldsymbol{y}} w(P_S(\boldsymbol{x}), P_{a_n}(\boldsymbol{y})) * d(P_S(\boldsymbol{x}), P_{a_n}(\boldsymbol{y})). \tag{5.4}$$

Recent papers on modality synthesis [112, 113] use sum of squared error (SSD) as the dissimilarity measure. However, under different lighting conditions, or in the case of MRI, the changes in imaging parameters may affect the intensity value, therefore resulting in the failure of SSD. Here, we propose a novel distance measure based on the intensity histogram (H) and Weber local descriptor (WLD) features [165]. The intensity histogram reflects the number of pixels in the patch at each different intensity value. While WLD reflects not only the intensity information, but also the change of intensity, using the WLD histogram is discriminative yet robust in many applications, including facial recognition under different illumination [166], etc. Intuitively, we formulate our dissimilarity measure as the dissimilarity of the intensity histogram and WLD histogram between $P_S(\boldsymbol{x})$ and $P_{a_n}(\boldsymbol{y})$. For histogram comparison,

we use Jensen-Shannon divergence (JSD) which is symmetrical, theoretically upper-bounded, and well-defined with histogram non-continuity [107].

We then define our novel dissimilarity measure as the following:

$$d(P_S(\boldsymbol{x}), P_{an}(\boldsymbol{y})) = \text{JSD}(\text{H}(P_S(\boldsymbol{x})) \parallel \text{H}(P_{an}(\boldsymbol{y}))) + \text{JSD}(\text{WLD}(P_S(\boldsymbol{x})) \parallel \text{WLD}(P_{an}(\boldsymbol{y})))$$

$$(5.5)$$

where

$$\text{JSD}(p_1 \parallel p_2) = \frac{1}{2}\text{KL}(p_1 \parallel M) + \frac{1}{2}\text{KL}(p_2 \parallel M), \tag{5.6}$$

$$M = \frac{1}{2}(p_1 + p_2), \tag{5.7}$$

$$\text{KL}(p_1 \parallel p_2) = \int p_1(i)\log\frac{p_1(i)}{p_2(i)}. \tag{5.8}$$

Based on Equation (5.4) and the proposed distance measure (5.5), we can find the best matching patch $P_{a_{\hat{n}}}(\boldsymbol{y}))$ that best matches the histogram of intensity and the WLD features. However, intensity and WLD histogram are robust global measurements. For our application, we aim to find the intensity mapping pixel by pixel, therefore, it is highly possible that the features are nicely matched while the particular pixel is not. We thus introduce a weighting factor $w(i_1, i_2)$ in our formulation.

The global intensity matching information can be learnt using NPMI [167], which we have introduced earlier in this thesis. Aiming to minimize Equation (5.4), we define $w(i_1, i_2)$ as follows:

$$w(i_1, i_2) = 1 - \text{NPMI}(i_1, i_2), \tag{5.9}$$

In the field of image matching, NPMI indicates the correlation relationship be-
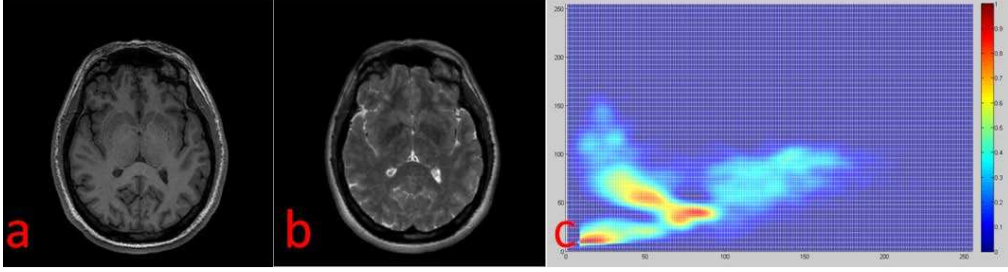
Figure 5.1: NPMI training example, using a pair of T1/T2 brain MR images. (a) T1 image (b) Corresponding aligned T2 image (c) Obtained NPMI.

tween different intensities. Figure 5.1 shows the NPMI map using a pair of perfectly aligned T1/T2 images. When intensity pair (i_1, i_2) tends to appear together, NPMI(i_1, i_2) is positive. The larger the NPMI value, the higher the correlation between the intensities. When two intensities are independent of each other, NPMI equals zero. In our application, to find a similar patch in the database, $\mathrm{NPMI}(P_S(\boldsymbol{x}), P_{a_n}(\boldsymbol{y}))$ has to be a positive value, which is an indication of correlation of these two intensity values.

Combining the information from $w(P_S(\boldsymbol{x}), P_{a_n}(\boldsymbol{y}))$ and $d(P_S(\boldsymbol{x}), P_{a_n}(\boldsymbol{y}))$, we can now find the best matching patch $M_{a_{\hat{n}}}(\boldsymbol{y})$ thus the corresponding patch in Modality b — $P_{b_{\hat{n}}}(\boldsymbol{y}))$. $T(\boldsymbol{x})$ will be filled in with the intensity value of $P_{b_{\hat{n}}}(\boldsymbol{y}))$.

5.2.2.3 Search in Multi-Resolution

A multi-resolution search scheme is employed. For every location $\boldsymbol{x}$, only the corresponding patches that are well matched with the source patches under different resolutions will be selected as the corresponding patch / point. The multi-resolution setup helps to increase the accuracy and robustness of the matching process, because the matching information in both local and global extends are

all considered.

5.3   Experiments

The proposed modality synthesis framework is generally applicable to any source-target modality pair provided a database with well aligned examples exists. To validate the proposed framework, in this section, we provide two sets of experiments. We first use a simple synthetic toy problem to demonstrate the flexibility, effectiveness and robustness of our proposed method. Then, we synthesize brain T2 images from T1. Compared with the state-of-the-art solution [113], the proposed method is more accurate and robust.

For the parameter settings, the patch size is fixed at 7×7 pixels ($7 \times 7 \times 7$ voxels in the 3D cases), the local search window W is set to 9×9 ($9 \times 9 \times 9$ in the 3D cases). Three resolutions are used in our experiments.

5.3.1   Synthetic Image Study

We illustrate the power of the proposed framework through a simple synthetic image problem. To build the database, two perfectly aligned circles are placed in two images. To reflect the multi-modal setup, the intensity values of these two circles are different. Gaussian noise is added to the images. If you can always find an exact matching patch in the database, it is quite obvious that any matching algorithm for modality synthesis will work. However, it is not usually the case in real application where some patches from the source image S is not available in the database, for example, some structures have been deformed. A good modality synthesis algorithm should be able to overcome such problem to
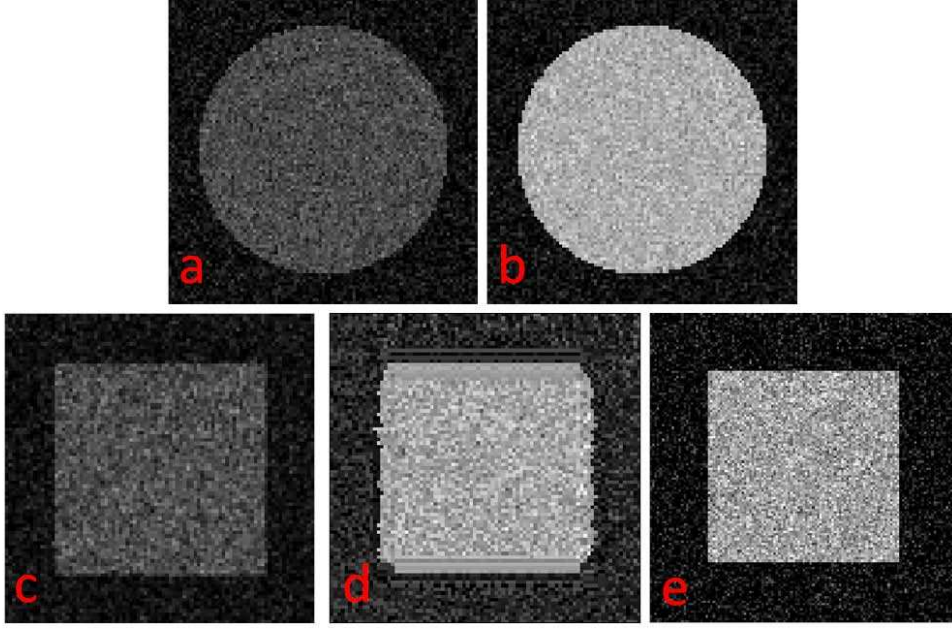
Figure 5.2: (a) Training Image Modality A (b) Training Image Modality B (c) Source Image Modality A (d) Synthesized Target Image using [113]'s method (e) Synthesized Target Image using the proposed method.

preserve the quality of the synthesized image. To examine the situation when there may not be a perfect match from the database, we create a source image S with a squared shape at the center. Such a setup creates a problem at the edge of the square, as there's no way to find a matched patch in the database, and it is interesting to see what the synthesized image would be with this setup.

To examine the robustness and the accuracy of the respective framework, we applied our proposed method in the setup mentioned above, and then make a direct comparison with the method in [113]. The result is shown in Figure 5.2. It is shown clearly that, our proposed method nicely preserve the square shape. On the other hand, the method in [113] fails at the edge of the square, because there's no straight edge in the training database.

This experiment clearly shows the advantages of the proposed framework. Firstly, the learnt weight $w(i_1, i_2)$ based on NPMI constrains the matching process in a global intensity point of view, so intensity matchings that are not correlated according to the database will be excluded. Secondly, using the histogram matching from intensity and the WLD features allows greater tolerance in the discrepancy between the patch from the source image and the database. Although the circle has deformed to a square, the edge and the intensity distribution information still remains. The proposed method is still able to produce a reasonable result. On the other hand, using SSD as the matching criteria [113] loses the generality of the information learnt from the database, especially near the edge area, thus resulting in poorer performance when no perfect match is found from the database. In other words, using SSD as the matching criteria may require a larger database to achieve a good result. The flexibility and the robustness of our proposed method make the database reduction possible, which leads to reduction of the searching space and thus the improvement of the computational efficiency.

5.3.2 Synthesis of T2 from T1 MRI

The second experiment is performed on synthesizing T2 from the corresponding T1-weighted MR brain images. Such modality transformation is useful for multi-modality registration [161], segmentation and abnormality detection [113]. In this experiment, we use the NAMIC database (http://hdl.handle.net/1926/1687) where T2 and T1 images are perfectly aligned. We only use the 13 brain image pairs where both T2 and T1 modalities exist in the database. In the pre-
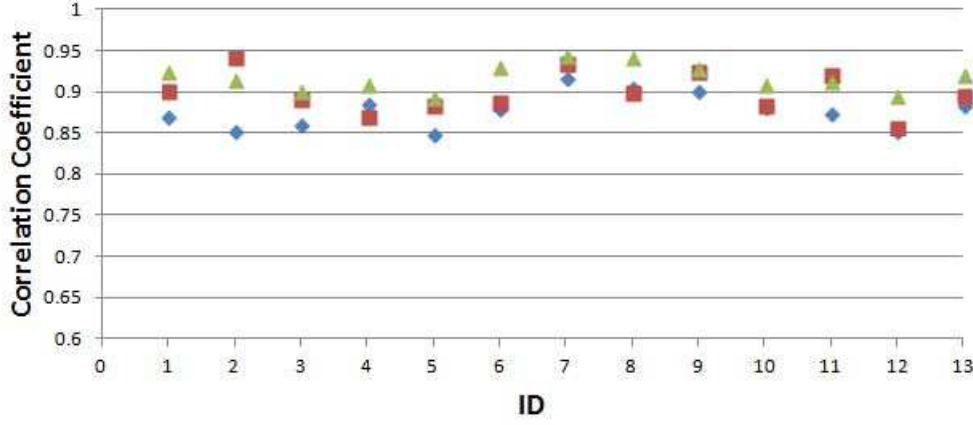
Figure 5.3: Correlation coefficients between synthesis T2 and the ground truth T2 computed by proposed method (with full database) (green), proposed method (with the reduced database) (red) and [113]'s method (with full database) (blue).

processing step, we linearly register the images, the brain images are skull-stripped, the inhomogeneity has been corrected and the intensity histogram has matched in each modality.

To quantitatively assess the experiment, we perform a leave-one-out cross-validation. Therefore, every synthesized brain image is based on the rest of the 12 subjects. Such setup allows us to compute the similarity between the synthesized and ground truth T2 images. Similar to the previous experiment, we compare our proposed method with method in [113]. Furthermore, we show the the effectiveness of the database reduction technique we introduced in Section 2.1, by using the full database, and the reduced database when synthesizing T2 using the proposed method. To measure the similarity between the synthesized and ground truth T2 images, we used correlation coefficient(CC) [168, 169] because it is the optimal intensity-based matching criterion when linear relation is assumed between the signal intensities [18].

Graphs in Figure 5.3 show the similarity measure (CC) calculated from the

synthesized T2 images and the ground truth. The symbols shown in green, red and blue indicate the CC value computed by proposed method (with full database), proposed method (with the reduced database) and [113]'s method (with full database) respectively. The size of the reduced database is only 1/10 of the full database. We found that reducing the database will cause [113]'s method to deteriorate quite significantly, thus we did not include [113]'s method with the reduced database in this paper. It is clearly shown in Figure 5.3 that, in all of the 13 cases, the proposed method outperforms [113]'s method. This is because the proposed matching criteria using patch based intensity histogram and Weber local descriptor features is more reliable compared to SSD. The proposed distance measure utilizes the histogram of the features, and the overall intensity matching relationship learnt from NPMI. In the modality synthesis application, the proposed distance measure is more reliable and demands less in terms of database size. It is because the intensity and WLD distribution within a patch is very limited (a patch only contains few tissue types), even the structure within the patch deforms, the histogram of intensity and WLD may still remain the same or with minimum changes. Supported by this, on our proposed distance measure can still provide a reasonable estimation even when the the source patch can only find a deformed match in the database. On the other hand, [113]'s method could not produce reliable results when there is not a good match in the database. It is also observed that the proposed method does not deviate a lot when using the reduced database compared to the full database. Such result also demonstrates the robustness of the proposed method.

A qualitative comparison is shown in Figure 5.4. Qualitatively, the proposed method performs better than [113]'s method, especially at the regions where
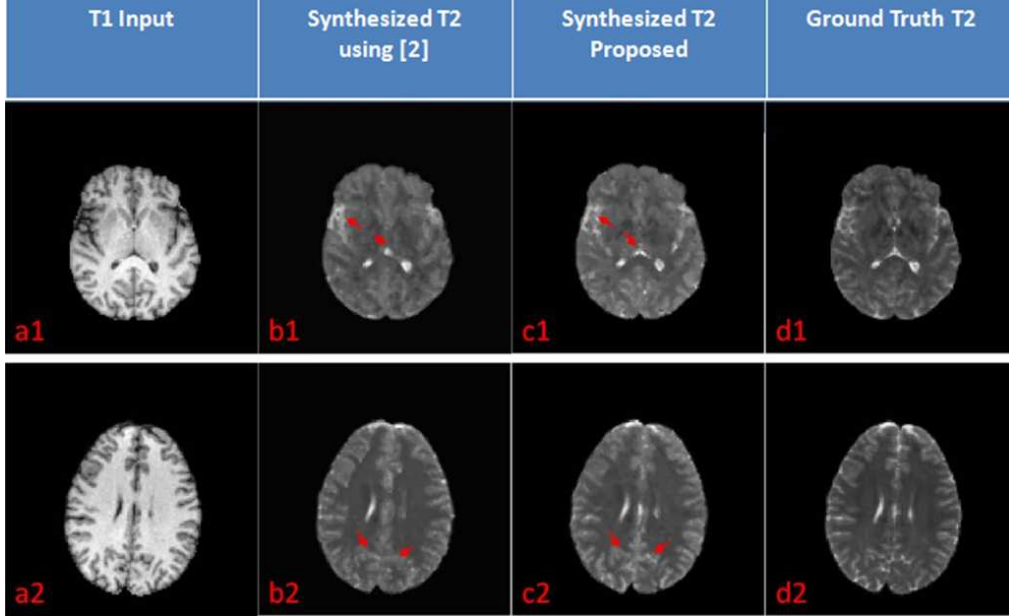
Figure 5.4: Visual results for synthesis of T2 from different data sets. Col (a) Input Images from T1 (b) Synthesis of T2 using [113](c) Synthesis of T2 using the proposed method (d) Ground truth T2 images.

rich structure exist. To demonstrate, let's focus on the regions where the red arrows point to. In the four regions pointed by the red arrow, synthesis of T2 using [113]'s method lose the fine detail of the ventricle, because in such structure rich region, using SSD to search the matched patch in the database is very difficult (it is too hard or impossible to find exact ventricle structures within the database), which leads to false matchings. In comparison, enforcing by the intensity matching information provided by NPMI, and with a more flexible and robust matching histogram-based distance measure, the proposed method preserve the fine structure even at the regions where exact similar patch is difficult to find in the database.

5.4 Conclusion

In this chapter, we proposed a general framework for modality synthesis. The proposed framework combined the strength from local intensity and Weber local descriptor features, together with the intensity matching information from the normalized pointwise mutual information, to produce a robust and impressive synthesis results. Furthermore, we adopt a database reduction technique that significantly reduces the size of the database and thus increases the computational efficiency. We evaluated the proposed method on a synthetic image study and through synthesis of T2 from T1 brain MR images. We show that the proposed method outperforms the state-of-the-art method in terms of accuracy and robustness. Furthermore, we demonstrate and explain that the proposed method is much less dependent to the size of the database, thus opening up the possibility of further reducing the size of the database. With more and more population databases made available, modality synthesis is becoming more important, how to fully utilize the information from the database is still an open question. Although the work and experiments presented in this chapter are still preliminary, we believe our work can become essential component for developing cross modality analysis tools. Such technique is also capable of bringing the multi-modal image analysis to a much more straightforward mono-modal image analysis. And image registration will be one of the biggest beneficiaries of such development.

Chapter 6

Conclusion and Future Work

In this chapter, we conclude the thesis with an overview of achievements and future work directions.

6.1 Incorporating Anatomical Prior

We have presented to incorporate the anatomical prior into the image registration framework. Firstly, a novel intensity-based similarity measure — structural encoded mutual information is proposed. The similarity metric weigh the pixels differently according to the anatomical significance. Secondly, an anatomical knowledge-based deformation field prior is derived. We use mixture of Gaussians to represent the point sets of interest that are extracted from the segmented organs. Point sets registration using a TPS model is then performed, which derives our data-driven deformation prior. Lastly, we proposed a novel similarity measure that combines the structural encoded mutual information and the knowledge from the prior deformation field. The similarity measure combines the low level intensity information from SMI and high level prior information on the deformation field.

We validate the proposed image registration work on three sets of experiment. In particular, it is shown that the proposed registration framework works well on registering pre-operative CT and non-contrast-enhanced C-arm CT, which is an essential step in the TAVI surgery. The significant improvement of the proposed method makes TAVI more applicable to many patients. The proposed method also show its strengths in registering perfusion cardiac MRIs and tumor resection MRIs.

The application of 3D/3D registration of pre-operative CT and non-contrast-enhanced C-arm CT is still in the research stage, and we plan to apply our proposed framework to the clinical TAVI surgery in the near future. Furthermore, we seek to further generalize the way of acquiring the anatomical information, such that it is not limited to the segmentation information. Possible options include biomechanical information of the objects, analysis of the respiratory and cardiac cycles, etc. We will also look to improve the effectiveness of the deformation field prior. Further studies include more effective sampling of the point sets and generating the prior deformation field in a more effective and efficient manner.

6.2 Incorporating Appearance Prior

We have proposed the use of normalized pointwise mutual information to model the complex intensity matching relationship in the multi-modality setup. Two novel similarities based on weighted mutual information and weighted entropy of the intensity matching confidence map were suggested. We find the optimum of the proposed similarity measures using the variational calculus, where gra-

dient of the similarity measures are calculated and gradient descent is used for finding the optimum solution.

The proposed similarities are validated on three sets of experiment. From the synthetic image example, and registration of human face with occlusions and background changes, we demonstrate that the NPMI is a more flexible and robust intensity matching prior compared to the state-of-the-art methods which use joint histogram as the appearance prior. We further compared the two similarity measures using the simulated brain MRIs. It is found that WMI may have more discriminant power on the local misalignments, given that the NPMI is very accurate. While weighted entropy has a larger capture range, and a less sensitive to the deviation of the intensity matching information. The choice of WMI and weighted entropy is thus very application-dependent.

NPMI has shown its capability of modeling complex intensity matching information, however, it should be noted that a good registration result depends on an accurate NPMI. In the future, more insight into NPMI is worth investigating. Ideally, NPMI should be trained by a small yet representative data set. We will look for the optimum solution for training the NPMI. More rigorous investigation on how noise, e.g., with the addition of white noise, biased field etc, would affect the image registration result of the proposed methods is required in our future work. Furthermore, the applications of our proposed method shown in this thesis is limited, and we intend to test the proposed similarity measures (both WMI and weighted entropy) on more applications.

6.3 Modality Synthesis

We presented a general framework for modality synthesis, utilizing the features such as intensity histogram and the Weber Local Descriptor. Such modality synthesis technique has the potential to transform the multi-modal image registration problem to a much simpler mono-modal image registration problem.

In the experiments, we validated that the proposed modality synthesis method outperforms the state-of-the-art methods. The experiment results are still preliminary. As the technique is still in the premature research stage, we thus did not provide the experiments with image registration applications. However, we believe that modality synthesis has the potential for developing cross modality analysis tools. And image registration will soon benefit from it.

It is highly possible that the modality synthesis result will be much improved if we take the underlying "physics" of different modalities into consideration in generating the synthetic images, e.g., the imaging parameters of the MR images. In the future, we will work towards the direction, so that our framework can be better suited to specific applications.

Chapter 7

Publication List

The publication list during the PhD candidature can be found below:

Journal Publication:

Lu, Y., Sun, Y., Liao, R., Ong, S. H. (2014). A Pre-operative CT and Non-contrast-enhanced C-arm CT Registration Framework for Trans-catheter Aortic Valve Implantation. Computerized Medical Imaging and Graphics, accepted.

Conference Publication:

Lu, Y., Sun, Y., Liao, R., Ong, S. H. (2014, April).Learning Based Deformable Image Registration Using Weighted Entropy of Intensity Mapping Confidence Map. In Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium. IEEE.

Lu, Y., Sun, Y., Liao, R., Ong, S. H. (2013, April). A new similarity measure for deformable image registration based on intensity matching. In Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on (pp. 234-237). IEEE.

Lu, Y., Sun, Y., Liao, R., Ong, S. H. (2013). Hybrid Multimodal Deformable

Registration with a Data-Driven Deformation Prior. In Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions (pp. 45-54). Springer Berlin Heidelberg.

Lu, Y., Sun, Y., Liao, R., Ong, S. H. (2013). Registration of pre-operative CT and non-contrast-enhanced C-arm CT: An application to trans-catheter aortic valve implantation (TAVI). In Computer Vision ACCV 2012 (pp. 268-280). Springer Berlin Heidelberg.

Lu, Y., Liao, R., Zhang, L., Sun, Y., Chefd'Hotel, C., Ong, S. H. (2012, November). Learning-based deformable registration using weighted mutual information. In Pattern Recognition (ICPR), 2012 21st International Conference on (pp. 2626-2629). IEEE.

US pending patents:

Liao, R., Lu, Y., Sun, Y., Ong, S. H. Modality Synthesis: Using Patch Based Intensity Histogram and Weber Local Descriptor Features, 2014P13356US, 2014.

Liao, R., Lu, Y., Sun, Y., Ong, S. H. Learning-based Deformable Image Registration Using Weighted Entropy of Intensity Mapping Confidence Map, 2014E09702US 2014.

Liao, R., Lu, Y., Sun, Y., Ong, S. H. A New Similarity Measure for Deformable Registration Based on Intensity Mapping, 2013P02148US, 2013.

Liao, R., Lu, Y., Sun, Y., Ong, S. H. Hybrid Multimodal Deformable Registration With A Data-Driven Deformation Prior, 2013E05548US, 2013.

Liao, R., Lu, Y., Sun, Y., Ong, S. H. Deformable Registration of Pre-operative CT and non-contrast-enhanced C-arm CT: An application to Trans-catheter Aortic Valve Implantation (TAVI), 2012P05569US, 2012.

Liao, R., Lu, Y., Sun, Y., Ong, S. H. A System and Method for Organ-

Specific Deformable Registration Using Conditional Mutual Information, 2011P27535US, 2011.

Liao, R., Lu, Y., Zhang, L., Sun, Y., Chefd'Hotel, C., Ong, S. H. Learning-Based Deformable Registration Using Weighted Mutual Information, 2011P26584US, 2011.

References

[1] L. G. Brown, "A survey of image registration techniques," *ACM computing surveys (CSUR)*, vol. 24, no. 4, pp. 325–376, 1992.

[2] J. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical image analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[3] D. L. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Physics in medicine and biology*, vol. 46, no. 3, p. R1, 2001.

[4] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[5] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: a survey," *Medical image analysis*, vol. 1, no. 2, pp. 91–108, 1996.

[6] H. Lester and S. R. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern recognition*, vol. 32, no. 1, pp. 129–149, 1999.

[7] A. A. Goshtasby, *2-D and 3-D image registration: for medical, remote sensing, and industrial applications*. John Wiley & Sons, 2005.

[8] J. Beutel, J. M. Fitzpatrick, S. C. Horii, Y. Kim, H. L. Kundel, M. Sonka, and R. L. Van Metter, *Handbook of Medical Imaging, Volume 3. Display and PACS*. Washington, DC: SPIE Press, 2002.

[9] C. L. Epstein, *Introduction to the mathematics of medical imaging*. Siam, 2008.

[10] J. T. Bushberg and J. M. Boone, *The essential physics of medical imaging*. Lippincott Williams & Wilkins, 2011.

[11] C. Li and Y. Sun, "Nonrigid registration of myocardial perfusion mri using pseudo ground truth," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*. Springer, 2009, pp. 165–172.

[12] M. John, R. Liao, Y. Zheng, A. Nöttling, J. Boese, U. Kirschstein, J. Kempfert, and T. Walther, "System to guide transcatheter aortic valve implantations based on interventional c-arm ct imaging," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*. Springer, 2010, pp. 375–382.

[13] Y. Lu, Y. Sun, R. Liao, and S. H. Ong, "A pre-operative ct and non-contrast-enhanced c-arm ct registration framework for trans-catheter aortic valve implantation," *Computerized Medical Imaging and Graphics*, 2014.

[14] T. Peters, B. Davey, P. Munger, R. Comeau, A. Evans, and A. Olivier, "Three-dimensional multimodal image-guidance for neurosurgery," *Medical Imaging, IEEE Transactions on*, vol. 15, no. 2, pp. 121–128, 1996.

[15] D. Rueckert, A. F. Frangi, and J. A. Schnabel, "Automatic construction of 3-d statistical deformation models of the brain using nonrigid registration," *Medical Imaging, IEEE Transactions on*, vol. 22, no. 8, pp. 1014–1025, 2003.

[16] M. Bach Cuadra, M. De Craene, V. Duay, B. Macq, C. Pollo, and J.-P. Thiran, "Dense deformation field estimation for atlas-based segmentation of pathological mr brain images," *Computer methods and programs in biomedicine*, vol. 84, no. 2, pp. 66–75, 2006.

[17] T. Makela, P. Clarysse, O. Sipila, N. Pauna, Q. C. Pham, T. Katila, and I. E. Magnin, "A review of cardiac image registration methods," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 9, pp. 1011–1021, 2002.

[18] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 7, pp. 1153–1190, 2013.

[19] S. Marsland, C. J. Twining, and C. J. Taylor, "A minimum description length objective function for groupwise non-rigid image registration," *Image and Vision Computing*, vol. 26, no. 3, pp. 333–346, 2008.

[20] L. Zollei, "A unified information theoretic framework for pair-and group-wise registration of medical images," 2006.

[21] J. Modersitzki, *Numerical methods for image registration*. OUP Oxford, 2003.

[22] B. Fischer and J. Modersitzki, "Ill-posed medicine—an introduction to image registration," *Inverse Problems*, vol. 24, no. 3, p. 034008, 2008.

[23] J. Modersitzki, *FAIR: flexible algorithms for image registration*. SIAM, 2009, vol. 6.

[24] C. Broit, "Optimal registration of deformed images," 1981.

[25] J.-P. Thirion, "Image matching as a diffusion process: an analogy with maxwell's demons," *Medical image analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[26] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *Image Processing, IEEE Transactions on*, vol. 5, no. 10, pp. 1435–1447, 1996.

[27] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "A viscous fluid model for multimodal non-rigid image registration using mutual information," *Medical image analysis*, vol. 7, no. 4, pp. 565–575, 2003.

[28] M.-C. Chiang, A. D. Leow, A. D. Klunder, R. A. Dutton, M. Barysheva, S. E. Rose, K. L. McMahon, G. I. De Zubicaray, A. W. Toga, and P. M. Thompson, "Fluid registration of diffusion tensor images using information theory," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 4, pp. 442–456, 2008.

[29] P. Dupuis, U. Grenander, and M. I. Miller, "Variational problems on flows of diffeomorphisms for image matching," *Quarterly of applied mathematics*, vol. 56, no. 3, p. 587, 1998.

[30] L. Risser, F. Vialard, R. Wolz, M. Murgasova, D. D. Holm, and D. Rueckert, "Simultaneous multi-scale registration using large deformation diffeomorphic metric mapping," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 10, pp. 1746–1759, 2011.

[31] S. Sommer, M. Nielsen, F. Lauze, and X. Pennec, "A multi-scale kernel bundle for lddmm: towards sparse deformation description across space and scales," in *Information Processing in Medical Imaging*. Springer, 2011, pp. 624–635.

[32] B. Fischer and J. Modersitzki, "Curvature based image registration," *Journal of Mathematical Imaging and Vision*, vol. 18, no. 1, pp. 81–85, 2003.

[33] S. Henn, "A full curvature based algorithm for image registration," *Journal of Mathematical Imaging and Vision*, vol. 24, no. 2, pp. 195–208, 2006.

[34] E. Meijering, "A chronology of interpolation: from ancient astronomy to modern signal and image processing," *Proceedings of the IEEE*, vol. 90, no. 3, pp. 319–342, 2002.

[35] M. J. D. Powell, *Approximation theory and methods*. Cambridge university press, 1981.

[36] L. L. Schumaker, *Spline functions: basic theory*. Cambridge University Press, 1981.

[37] M. J. Powell, "Radial basis functions for multivariable interpolation: a

review," in *Algorithms for approximation*. Clarendon Press, 1987, pp. 143–167.

[38] ——, "The theory of radial basis function approximation in 1990," *Advances in numerical analysis*, vol. 2, pp. 105–210, 1992.

[39] M. D. Buhmann, M. D. Buhmann, and M. D. Buhmann, *Radial basis functions: theory and implementations*. Cambridge university press Cambridge, 2003, vol. 5.

[40] N. Arad, N. Dyn, D. Reisfeld, and Y. Yeshurun, "Image warping by radial basis functions: Application to facial expressions," *CVGIP: Graphical models and image processing*, vol. 56, no. 2, pp. 161–172, 1994.

[41] D. Ruprecht and H. Müller, *Free form deformation with scattered data interpolation methods*. Springer, 1993.

[42] K. Rohr, H. S. Stiehl, R. Sprengel, T. M. Buzug, J. Weese, and M. Kuhn, "Landmark-based elastic registration using approximating thin-plate splines," *Medical Imaging, IEEE Transactions on*, vol. 20, no. 6, pp. 526–534, 2001.

[43] M. Fornefett, K. Rohr, and H. S. Stiehl, "Radial basis functions with compact support for elastic registration of medical images," *Image and Vision Computing*, vol. 19, no. 1, pp. 87–96, 2001.

[44] L. Zagorchev and A. Goshtasby, "A comparative study of transformation functions for nonrigid image registration," *Image Processing, IEEE Transactions on*, vol. 15, no. 3, pp. 529–538, 2006.

[45] M. H. Davis, A. Khotanzad, D. P. Flamig, and S. E. Harms, "A physics-based coordinate transformation for 3-d image matching," *Medical Imaging, IEEE Transactions on*, vol. 16, no. 3, pp. 317–328, 1997.

[46] J. Kohlrausch, K. Rohr, and H. S. Stiehl, "A new class of elastic body splines for nonrigid registration of medical images," *Journal of Mathematical Imaging and Vision*, vol. 23, no. 3, pp. 253–280, 2005.

[47] S. Wörz and K. Rohr, "Physics-based elastic registration using non-radial basis functions and including landmark localization uncertainties," *Computer Vision and Image Understanding*, vol. 111, no. 3, pp. 263–274, 2008.

[48] I. J. Schoenberg, "Contributions to the problem of approximation of equidistant data by analytic functions," *Quart. Appl. Math*, vol. 4, no. 2, pp. 45–99, 1946.

[49] M. Unser, A. Aldroubi, and M. Eden, "Fast b-spline transforms for continuous image representation and interpolation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 3, pp. 277–285, 1991.

[50] ——, "B-spline signal processing. ii. efficiency design and applications," *Signal Processing, IEEE Transactions on*, vol. 41, no. 2, pp. 834–848, 1993.

[51] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited [medical images application]," *Medical Imaging, IEEE Transactions on*, vol. 19, no. 7, pp. 739–758, 2000.

[52] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images," *Medical Imaging, IEEE Transactions on*, vol. 18, no. 8, pp. 712–721, 1999.

[53] J. Declerck, J. Feldmar, M. L. Goris, and F. Betting, "Automatic registration and alignment on a template of cardiac stress and rest reoriented spect images," *Medical Imaging, IEEE Transactions on*, vol. 16, no. 6, pp. 727–737, 1997.

[54] M. Sdika, "A fast nonrigid image registration with constraints on the jacobian using large scale constrained optimization," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 2, pp. 271–281, 2008.

[55] D. Perperidis, R. H. Mohiaddin, and D. Rueckert, "Spatio-temporal free-form registration of cardiac mr image sequences," *Medical image analysis*, vol. 9, no. 5, pp. 441–456, 2005.

[56] D. Rueckert, P. Aljabar, R. A. Heckemann, J. V. Hajnal, and A. Hammers, "Diffeomorphic registration using b-splines," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*. Springer, 2006, pp. 702–709.

[57] J. A. Schnabel, D. Rueckert, M. Quist, J. M. Blackall, A. D. Castellano-Smith, T. Hartkens, G. P. Penney, W. A. Hall, H. Liu, C. L. Truwit *et al.*, "A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations," in *Medical Image Computing and*

Computer-Assisted Intervention–MICCAI 2001. Springer, 2001, pp. 573–581.

[58] W. Shi, X. Zhuang, L. Pizarro, W. Bai, H. Wang, K.-P. Tung, P. Edwards, and D. Rueckert, "Registration using sparse free-form deformations," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012.* Springer, 2012, pp. 659–666.

[59] J. Wang and T. Jiang, "Nonrigid registration of brain mri using nurbs," *Pattern Recognition Letters*, vol. 28, no. 2, pp. 214–223, 2007.

[60] W. Feng, S. J. Reeves, T. S. Denney, S. Lloyd, L. Dell'Italia, and H. Gupta, "A new consistent image registration formulation with a b-spline deformation model," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on.* IEEE, 2009, pp. 979–982.

[61] V. Noblet, C. Heinrich, F. Heitz, and J.-P. Armspach, "Symmetric non-rigid image registration: application to average brain templates construction," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008.* Springer, 2008, pp. 897–904.

[62] A. Sotiras and N. Paragios, "Discrete symmetric image registration," in *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on.* IEEE, 2012, pp. 342–345.

[63] Y. Amit, "A nonlinear variational problem for image matching," *SIAM Journal on Scientific Computing*, vol. 15, no. 1, pp. 207–224, 1994.

[64] G. E. Christensen and H. J. Johnson, "Consistent image registration," *Medical Imaging, IEEE Transactions on*, vol. 20, no. 7, pp. 568–582, 2001.

[65] J. Ashburner and K. J. Friston, "Nonlinear spatial normalization using basis functions," *Human brain mapping*, vol. 7, no. 4, pp. 254–266, 1999.

[66] Y.-T. Wu, T. Kanade, C.-C. Li, and J. Cohn, "Image registration using wavelet-based motion model," *International Journal of Computer Vision*, vol. 38, no. 2, pp. 129–152, 2000.

[67] O. Musse, F. Heitz, and J.-P. Armspach, "Topology preserving deformable image matching using constrained hierarchical parametric models," *Image Processing, IEEE Transactions on*, vol. 10, no. 7, pp. 1081–1093, 2001.

[68] P. Hellier, C. Barillot, É. Mémin, and P. Pérez, "Hierarchical estimation of a dense deformation field for 3-d robust registration," *Medical Imaging, IEEE Transactions on*, vol. 20, no. 5, pp. 388–402, 2001.

[69] R. Narayanan, J. A. Fessler, H. Park, and C. R. Meyer, "Diffeomorphic nonlinear transformations: A local parametric approach for image registration," in *Information Processing in Medical Imaging*. Springer, 2005, pp. 174–185.

[70] V. Arsigny, X. Pennec, and N. Ayache, "Polyrigid and polyaffine transformations: a novel geometrical tool to deal with non-rigid deformations–application to the registration of histological slices," *Medical image analysis*, vol. 9, no. 6, pp. 507–523, 2005.

[71] V. Arsigny, O. Commowick, N. Ayache, and X. Pennec, "A fast and log-euclidean polyaffine framework for locally linear registration," *Journal of Mathematical Imaging and Vision*, vol. 33, no. 2, pp. 222–238, 2009.

[72] S. Tang, Y. Fan, G. Wu, M. Kim, and D. Shen, "Rabbit: rapid alignment of brains by building intermediate templates," *NeuroImage*, vol. 47, no. 4, pp. 1277–1287, 2009.

[73] O. Commowick, R. Stefanescu, P. Fillard, V. Arsigny, N. Ayache, X. Pennec, and G. Malandain, "Incorporating statistical measures of anatomical variability in atlas-to-subject registration for conformal brain radiotherapy," in *Medical Image Computing and Computer-Assisted Intervention– MICCAI 2005*. Springer, 2005, pp. 927–934.

[74] Z. Xue, D. Shen, and C. Davatzikos, "Statistical representation of high-dimensional deformation fields with application to statistically constrained 3d warping," *Medical Image Analysis*, vol. 10, no. 5, pp. 740–751, 2006.

[75] E. I. Zacharaki, D. Shen, S.-K. Lee, and C. Davatzikos, "Orbit: a multiresolution framework for deformable registration of brain tumor images," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 8, pp. 1003–1017, 2008.

[76] E. I. Zacharaki, C. S. Hogea, D. Shen, G. Biros, and C. Davatzikos, "Non-diffeomorphic registration of brain tumor images by simulating tissue loss and tumor growth," *Neuroimage*, vol. 46, no. 3, pp. 762–774, 2009.

[77] F. S. Azar, D. N. Metaxas, and M. D. Schnall, "A deformable finite element model of the breast for predicting mechanical deformations under external perturbations," *Academic Radiology*, vol. 8, no. 10, pp. 965–975, 2001.

[78] A. P. Del Palomar, B. Calvo, J. Herrero, J. Lopez, and M. Doblaré, "A finite element model to accurately predict real deformations of the breast," *Medical Engineering & Physics*, vol. 30, no. 9, pp. 1089–1097, 2008.

[79] L. Han, J. Hipwell, T. Mertzanidou, T. Carter, M. Modat, S. Ourselin, and D. Hawkes, "A hybrid fem-based method for aligning prone and supine images for image guided breast surgery," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*. IEEE, 2011, pp. 1239–1242.

[80] A. Mohamed, C. Davatzikos, and R. Taylor, "A combined statistical and biomechanical model for estimation of intra-operative prostate deformation," in *Medical Image Computing and Computer-Assisted InterventionMICCAI 2002*. Springer, 2002, pp. 452–460.

[81] R. Alterovitz, K. Goldberg, J. Pouliot, I.-C. J. Hsu, Y. Kim, S. M. Noworolski, and J. Kurhanewicz, "Registration of mr prostate images with biomechanical modeling and nonlinear parameter estimation," *Medical physics*, vol. 33, no. 2, pp. 446–454, 2006.

[82] A. Goshtasby and J. Le Moign, *Image registration*. Springer, 2012.

[83] Y. Tsin and T. Kanade, "A correlation-based approach to robust point

set registration," in *Computer Vision-ECCV 2004*.   Springer, 2004, pp. 558–569.

[84] M. Singh, H. Arora, and N. Ahuja, "Robust registration and tracking using kernel density correlation," in *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04. Conference on*.   IEEE, 2004, pp. 174–174.

[85] B. Jian and B. C. Vemuri, "Robust point set registration using gaussian mixture models," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 8, pp. 1633–1645, 2011.

[86] A. S. Roy, A. Gopinath, and A. Rangarajan, "Deformable density matching for 3d non-rigid registration of shapes," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*.   Springer, 2007, pp. 942–949.

[87] N. Paragios, M. Rousson, and V. Ramesh, "Non-rigid registration using distance functions," *Computer Vision and Image Understanding*, vol. 89, no. 2, pp. 142–165, 2003.

[88] X. Huang, N. Paragios, and D. N. Metaxas, "Shape registration in implicit spaces using information theory and free form deformations," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 8, pp. 1303–1318, 2006.

[89] M. Savinaud, N. Paragios, and S. Maitrejean, "Motion-based enhancement of optical imaging," in *Biomedical Imaging: From Nano to Macro,*

2009. ISBI'09. IEEE International Symposium on. IEEE, 2009, pp. 738–741.

[90] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in *Medical Image Computing and Computer-Assisted InterventationłMIC-CAI98.* Springer, 1998, pp. 1115–1124.

[91] D. Shen and C. Davatzikos, "Hammer: hierarchical attribute matching mechanism for elastic registration," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 11, pp. 1421–1439, 2002.

[92] D. Shen, "Image registration by local histogram matching," *Pattern Recognition*, vol. 40, no. 4, pp. 1161–1172, 2007.

[93] Y. Ou, A. Sotiras, N. Paragios, and C. Davatzikos, "Dramms: Deformable registration via attribute matching and mutual-saliency weighting," *Medical image analysis*, vol. 15, no. 4, pp. 622–639, 2011.

[94] S. Liao and A. C. Chung, "Feature based nonrigid brain mr image registration with symmetric alpha stable filters," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 1, pp. 106–119, 2010.

[95] A. Myronenko and X. Song, "Intensity-based image registration by minimizing residual complexity," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 11, pp. 1882–1891, 2010.

[96] G. Hermosillo, C. Chefd'Hotel, and O. Faugeras, "Variational methods

for multimodal image matching," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 329–343, 2002.

[97] W. M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical image analysis*, vol. 1, no. 1, pp. 35–51, 1996.

[98] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multi-modality image registration based on information theory," in *Information processing in medical imaging*, vol. 3, no. 6, 1995, pp. 263–274.

[99] C. Studholme, D. L. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3d medical image alignment," *Pattern recognition*, vol. 32, no. 1, pp. 71–86, 1999.

[100] D. Rueckert, M. Clarkson, D. Hill, and D. J. Hawkes, "Non-rigid registration using higher-order mutual information," in *Medical Imaging 2000*. International Society for Optics and Photonics, 2000, pp. 438–447.

[101] C. Studholme, C. Drapaca, B. Iordanova, and V. Cardenas, "Deformation-based mapping of volume change from serial brain mri in the presence of local tissue contrast change," *Medical Imaging, IEEE Transactions on*, vol. 25, no. 5, pp. 626–639, 2006.

[102] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens, "Nonrigid image registration using conditional mutual information," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 1, pp. 19–29, 2010.

[103] X. Zhuang, S. Arridge, D. J. Hawkes, and S. Ourselin, "A nonrigid registration framework using spatially encoded mutual information and free-form deformations," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 10, pp. 1819–1828, 2011.

[104] A. C. Chung, R. Gan, and W. Wells III, "Robust multi-modal image registration based on prior joint intensity distributions and minimization of kullback-leibler distance," *Report CS07-01, Hong Kong Univ. of Science & Technology*, 2007.

[105] C. Guetter, C. Xu, F. Sauer, and J. Hornegger, "Learning based non-rigid multi-modal image registration using kullback-leibler divergence," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*. Springer, 2005, pp. 255–262.

[106] D. Cremers, C. Guetter, and C. Xu, "Nonparametric priors on the space of joint intensity distributions for non-rigid multi-modal image registration," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 2. IEEE, 2006, pp. 1777–1783.

[107] R. Liao, C. Guetter, C. Xu, Y. Sun, A. Khamene, and F. Sauer, "Learning-based 2d/3d rigid registration using jensen-shannon divergence for image-guided surgery," in *Medical Imaging and Augmented Reality*. Springer, 2006, pp. 228–235.

[108] A. Roche, X. Pennec, G. Malandain, and N. Ayache, "Rigid registration of 3-d ultrasound with mr images: a new approach combining inten-

sity and gradient information," *Medical Imaging, IEEE Transactions on*, vol. 20, no. 10, pp. 1038–1049, 2001.

[109] W. Wein, S. Brunke, A. Khamene, M. R. Callstrom, and N. Navab, "Automatic ct-ultrasound registration for diagnostic imaging and image-guided intervention," *Medical image analysis*, vol. 12, no. 5, pp. 577–585, 2008.

[110] A. Andronache, M. von Siebenthal, G. Székely, and P. Cattin, "Non-rigid registration of multi-modal images using both mutual information and cross-correlation," *Medical image analysis*, vol. 12, no. 1, pp. 3–15, 2008.

[111] J. A. Maintz, P. A. van den Elsen, and M. A. Viergever, "3d multimodality medical image registration using morphological tools," *Image and vision computing*, vol. 19, no. 1, pp. 53–62, 2001.

[112] J. E. Iglesias, E. Konukoglu, D. Zikic, B. Glocker, K. Van Leemput, and B. Fischl, "Is synthesizing mri contrast useful for inter-modality analysis?" in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*.   Springer, 2013, pp. 631–638.

[113] D. H. Ye, D. Zikic, B. Glocker, A. Criminisi, and E. Konukoglu, "Modality propagation: Coherent synthesis of subject-specific scans with data-driven regularization," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*.   Springer, 2013, pp. 606–613.

[114] H. J. Johnson and G. E. Christensen, "Consistent landmark and intensity-based image registration," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 5, pp. 450–461, 2002.

[115] D. Paquin, D. Levy, and L. Xing, "Hybrid multiscale landmark and deformable image registration," DTIC Document, Tech. Rep., 2007.

[116] E. Gibson, A. R. Khan, and M. F. Beg, "A combined surface and volumetric registration (savor) framework to study cortical biomarkers and volumetric imaging data," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*. Springer, 2009, pp. 713–720.

[117] K. Rohr, P. Cathier, and S. Wörz, "Elastic registration of electrophoresis images using intensity information and point landmarks," *Pattern recognition*, vol. 37, no. 5, pp. 1035–1048, 2004.

[118] T. Hartkens, D. L. Hill, A. D. Castellano-Smith, D. J. Hawkes, C. R. Maurer Jr, A. J. Martin, W. A. Hall, H. Liu, and C. L. Truwit, "Using points and surfaces to improve voxel-based non-rigid registration," in *Medical Image Computing and Computer-Assisted InterventionłMICCAI 2002*. Springer, 2002, pp. 565–572.

[119] P. Hellier and C. Barillot, "Coupling dense and landmark-based approaches for nonrigid registration," *Medical Imaging, IEEE Transactions on*, vol. 22, no. 2, pp. 217–227, 2003.

[120] A. A. Joshi, D. W. Shattuck, P. M. Thompson, and R. M. Leahy, "Surface-constrained volumetric brain registration using harmonic mappings," *Medical Imaging, IEEE Transactions on*, vol. 26, no. 12, pp. 1657–1669, 2007.

[121] P. Cachier, J.-F. Mangin, X. Pennec, D. Rivière, D. Papadopoulos-Orfanos, J. Régis, and N. Ayache, "Multisubject non-rigid registration of

brain mri using intensity and geometric features," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*.   Springer, 2001, pp. 734–742.

[122] A. Joshi, R. Leahy, A. W. Toga, and D. Shattuck, "A framework for brain registration via simultaneous surface and volume flow," in *Information Processing in Medical Imaging*.   Springer, 2009, pp. 576–588.

[123] A. Cifor, L. Risser, D. Chung, E. M. Anderson, and J. A. Schnabel, "Hybrid feature-based log-demons registration for tumour tracking in 2-d liver ultrasound images," in *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*.   IEEE, 2012, pp. 724–727.

[124] J. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE TMI*, vol. 22, no. 8, pp. 986–1004, 2003.

[125] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp. 825–841, 2002.

[126] L. Zöllei, E. Learned-Miller, E. Grimson, and W. Wells, "Efficient population registration of 3d data," in *Computer Vision for Biomedical Image Applications*.   Springer, 2005, pp. 291–301.

[127] M. De Craene, B. Macq, F. Marques, P. Salembier, and S. Warfield, "Unbiased group-wise alignment by iterative central tendency estimations," *Mathematical Modelling of Natural Phenomena*, vol. 3, no. 06, pp. 2–32, 2008.

[128] W. R. Crum, T. Hartkens, and D. Hill, "Non-rigid image registration: theory and practice," 2014.

[129] A. W. Toga and P. M. Thompson, "The role of image registration in brain mapping," *Image and vision computing*, vol. 19, no. 1, pp. 3–24, 2001.

[130] K. K. Bhatia, J. V. Hajnal, B. K. Puri, A. D. Edwards, and D. Rueckert, "Consistent groupwise non-rigid registration for atlas construction," in *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*. IEEE, 2004, pp. 908–911.

[131] N. Kovacevic, J. Chen, J. G. Sled, J. Henderson, and M. Henkelman, "Deformation based representation of groupwise average and variability," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004*. Springer, 2004, pp. 615–622.

[132] N. Kovačević, J. Henderson, E. Chan, N. Lifshitz, J. Bishop, A. Evans, R. Henkelman, and X. Chen, "A three-dimensional mri atlas of the mouse brain with estimates of the average and variability," *Cerebral Cortex*, vol. 15, no. 5, pp. 639–645, 2005.

[133] D. Seghers, E. DAgostino, F. Maes, D. Vandermeulen, and P. Suetens, "Construction of a brain template from mr images using state-of-the-art registration and segmentation techniques," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004*. Springer, 2004, pp. 696–703.

[134] H. Park, P. H. Bland, A. O. Hero III, and C. R. Meyer, "Least biased target selection in probabilistic atlas construction," in *Medical Image Comput-*

ing and Computer-Assisted Intervention–MICCAI 2005. Springer, 2005, pp. 419–426.

[135] S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, pp. S151–S160, 2004.

[136] S. K. Balci, P. Golland, and W. Wells, "Non-rigid groupwise registration using b-spline deformation model," *Open source and open data for MICCAI*, pp. 105–121, 2007.

[137] S. Allassonnière, Y. Amit, and A. Trouvé, "Towards a coherent statistical framework for dense deformable template estimation," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, no. 1, pp. 3–29, 2007.

[138] D. J. Blezek and J. V. Miller, "Atlas stratification," *Medical Image Analysis*, vol. 11, no. 5, pp. 443–457, 2007.

[139] M. R. Sabuncu, S. K. Balci, and P. Golland, "Discovering modes of an image population through mixture modeling," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008.* Springer, 2008, pp. 381–389.

[140] C. O. Sorzano, P. Thevenaz, and M. Unser, "Elastic registration of biological images using vector-spline regularization," *Biomedical Engineering, IEEE Transactions on*, vol. 52, no. 4, pp. 652–663, 2005.

[141] X. Papademetris, A. P. Jackowski, R. T. Schultz, L. H. Staib, and J. S.

Duncan, "Integrated intensity and point-feature nonrigid registration," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004*. Springer, 2004, pp. 763–770.

[142] J. Mitra, Z. Kato, R. Martí, A. Oliver, X. Lladó, D. Sidibé, S. Ghose, J. C. Vilanova, J. Comet, and F. Meriaudeau, "A spline-based non-linear diffeomorphism for multimodal prostate registration," *Medical Image Analysis*, vol. 16, no. 6, pp. 1259–1279, 2012.

[143] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 11, no. 6, pp. 567–585, 1989.

[144] A. Goshtasby, "Registration of images with geometric distortions," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 26, no. 1, pp. 60–64, 1988.

[145] H. J. Johnson and G. E. Christensen, "Landmark and intensity-based, consistent thin-plate spline image registration," in *Information Processing in Medical Imaging*. Springer, 2001, pp. 329–343.

[146] H. Chui and A. Rangarajan, "A new algorithm for non-rigid point matching," in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, vol. 2. IEEE, 2000, pp. 44–51.

[147] A. Myronenko, X. Song, and M. A. Carreira-Perpinán, "Non-rigid point set registration: Coherent point drift," in *Advances in Neural Information Processing Systems*, 2006, pp. 1009–1016.

[148] R. Richa, P. Poignet, and C. Liu, "Efficient 3d tracking for motion compensation in beating heart surgery," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*. Springer, 2008, pp. 684–691.

[149] W. W. Lau, N. A. Ramey, J. J. Corso, N. V. Thakor, and G. D. Hager, "Stereo-based endoscopic tracking of cardiac surface deformation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004*. Springer, 2004, pp. 494–501.

[150] J. Suh, D. Scheinost, X. Qian, A. Sinusas, C. Breuer, and X. Papademetris, "Serial nonrigid vascular registration using weighted normalized mutual information," in *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*. IEEE, 2010, pp. 25–28.

[151] R. Courant and J. Moser, *Calculus of Variations 1945-1946*, 1957.

[152] Y. Lu, Y. Sun, R. Liao, and S. H. Ong, "Registration of pre-operative ct and non-contrast-enhanced c-arm ct: An application to trans-catheter aortic valve implantation (tavi)," in *Computer Vision–ACCV 2012*. Springer, 2013, pp. 268–280.

[153] Y. Zheng, M. John, R. Liao, J. Boese, U. Kirschstein, B. Georgescu, S. K. Zhou, J. Kempfert, T. Walther, G. Brockmann *et al.*, "Automatic aorta segmentation and valve landmark detection in c-arm ct: application to aortic valve implantation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*. Springer, 2010, pp. 476–483.

[154] C. Li, X. Jia, and Y. Sun, "Improved semi-automated segmentation of cardiac ct and mr images," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*. IEEE, 2009, pp. 25–28.

[155] G. Bouma, "Normalized (pointwise) mutual information in collocation extraction," *Proceedings of GSCL*, pp. 31–40, 2009.

[156] C. Chefd'hotel, G. Hermosillo, and O. Faugeras, "Flows of diffeomorphisms for multimodal image registration," in *Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposium on*. IEEE, 2002, pp. 753–756.

[157] S.-K. S. Fan and Y.-C. Chuang, "An entropy-based image registration method using image intensity difference on overlapped region," *Machine Vision and Applications*, vol. 23, no. 4, pp. 791–804, 2012.

[158] S. Guiaşu, "Weighted entropy," *Reports on Mathematical Physics*, vol. 2, no. 3, pp. 165–179, 1971.

[159] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans, "Brainweb: Online interface to a 3d mri simulated brain database," in *NeuroImage*. Citeseer, 1997.

[160] T. Cao, C. Zach, S. Modla, D. Powell, K. Czymmek, and M. Niethammer, "Registration for correlative microscopy using image analogies," in *Biomedical Image Registration*. Springer, 2012, pp. 296–306.

[161] D.-J. Kroon and C. H. Slump, "Mri modalitiy transformation in demon registration," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*. IEEE, 2009, pp. 963–966.

[162] Z. Tu, K. L. Narr, P. Dollár, I. Dinov, P. M. Thompson, and A. W. Toga, "Brain anatomical structure segmentation by hybrid discriminative/generative models," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 4, pp. 495–508, 2008.

[163] W. Ou and C. Chefd'Hotel, "Polynomial intensity correction for multi-modal image registration," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*. IEEE, 2009, pp. 939–942.

[164] E. Elhamifar, G. Sapiro, and R. Vidal, "See all by looking at a few: Sparse modeling for finding representative objects," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 1600–1607.

[165] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikainen, X. Chen, and W. Gao, "WLD: A robust local image descriptor," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 9, pp. 1705–1720, 2010.

[166] B. Wang, W. Li, W. Yang, and Q. Liao, "Illumination normalization based on weber's law with application to face recognition," *Signal Processing Letters, IEEE*, vol. 18, no. 8, pp. 462–465, 2011.

[167] Y. Lu, R. Liao, L. Zhang, Y. Sun, C. Chefd'Hotel, and S. H. Ong, "Learning-based deformable registration using weighted mutual informa-

tion," in *Pattern Recognition (ICPR), 2012 21st International Conference on*.   IEEE, 2012, pp. 2626–2629.

[168] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain," *Medical image analysis*, vol. 12, no. 1, pp. 26–41, 2008.

[169] J. Kim and J. A. Fessler, "Intensity-based image registration using robust correlation coefficients," *Medical Imaging, IEEE Transactions on*, vol. 23, no. 11, pp. 1430–1444, 2004.