

**ADAPTIVE REGISTRATION AND ATLAS BASED SEGMENTATION**

**by**

**Hyunjin Park**

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Biomedical Engineering)  
in The University of Michigan  
2003

Doctoral Committee:

Professor Charles R. Meyer, Co-Chair  
Associate Professor Jeffrey A. Fessler, Co-Chair  
Professor Paul L. Carson  
Professor Alfred O. Hero III  
Associate Professor Douglas C. Noll

## **ABSTRACT**

### **ADAPTIVE REGISTRATION AND ATLAS BASED SEGMENTATION**

by

Hyunjin Park

Co-Chairs : Charles R. Meyer and Jeffrey A. Fessler

With the rapid developments in image registration techniques, registrations are applied not only as linear transforms but also as warping transforms with increasing frequency. The latter is especially suitable for soft tissue organs in the human body. When using Thin Plate Spline (TPS) as the warping transform of choice, a high degree of freedom (DOF) warping can be either manually specified by the placement of control points or implemented by using a dense grid of control points. The former leads to problems of operator bias, whereas the latter is very computationally expensive. Instead, we propose to automate the determination of DOF by locally increasing the density of control points in regions where they are needed rather than globally increasing the density of control points. Local estimates of Mutual Information (MI) and entropy are used to identify local regions requiring higher DOF.

There have been significant efforts to build a probabilistic atlas of the brain and to use it for many common applications like segmentation and registration. Though the work related to brain atlases can be applied to non-brain organs, less attention has been

paid to actually building an atlas for organs other than the brain. We present a method to construct a probabilistic atlas of an abdomen consisting of 4 organs (i.e., liver, kidneys and spinal cord). Using 32 non-contrast abdominal CT scans, 31 are mapped onto one individual scan using TPS as the warping transform and MI as the similarity measure. Except for an initial coarse placement of 4 control points by the operators, the MI based registration is automatic. Additionally, the four organs in each of the 32 CT data sets are manually segmented. The manual segmentations are warped onto the “standard” patient space using the same transform computed from their gray scale CT data set and a probabilistic atlas is calculated. Then the atlas is used to aid the segmentation of low contrast organs in additional 20 CT data sets not included in the atlas. By incorporating the atlas information into the Bayesian framework, segmentation results clearly showed improvements over a standard unsupervised segmentation method.

© Hyunjin Park 2003  
All Rights Reserved

To my wife Hyeun and my parents.

## ACKNOWLEDGMENTS

I wish to thank my advisor, Prof. Charles Meyer, for the guidance and friendship he has provided during my graduate student life in University of Michigan. I would never have made it this far without his helpful insights and suggestions. Secondly I wish to thank my co-chair, Prof. Jeff Fessler, who provided great advice every time I struggled with my research problems.

I also wish to thank rest of my committee member, Prof. Paul Carson, Prof. Alfred Hero and Prof. Doug Noll for all the expertise they have contributed to my dissertation.

To my parents, thank you for your support and belief in me. Also to my only sibling, Namjoon, thank you for your support.

I would also like to thank all my DIPL labmates, especially Ram, for all the feedback and comments. Also I would like to acknowledge my friends in Ann Arbor, Jongbum, Sangwoo and Hyunseok for letting me keep my life balanced outside my school work.

And finally and most importantly, I would like to express my deepest gratitude to my wife, Hyeun, whose love and support made this possible.

## TABLE OF CONTENTS

<b>DEDICATION</b> .....	<b>ii</b>
<b>ACKNOWLEDGMENTS</b> .....	<b>iii</b>
<b>LIST OF TABLES</b> .....	<b>vi</b>
<b>LIST OF FIGURES</b> .....	<b>vii</b>
<b>CHAPTER 1</b>	
<b>INTRODUCTION</b> .....	<b>1</b>
1.1. Dissertation Overview .....	1
1.2. Adaptive Registration .....	2
1.3. Abdominal Atlas .....	4
1.4. Bayesian Segmentation based on Atlas .....	5
<b>CHAPTER 2</b>	
<b>REGISTRATION OVERVIEW</b> .....	<b>7</b>
2.1. Formal Statement of Registration .....	7
2.2. Similarity Measures .....	7
2.3. Geometric Transforms .....	11
2.4. Local Effects of Control points in TPS.....	18
2.5. MIAMI Fuse .....	19
<b>CHAPTER 3</b>	
<b>ADAPTIVE REGISTRATION</b> .....	<b>20</b>
3.1. Shortcoming of high DOF non-linear registrations .....	20
3.2. General Approach for Adaptive Registration .....	21
3.3. Local Mismatch Measure M .....	21
3.4. Triad of Control Points .....	23
3.5. Iterative algorithm.....	24

<b>CHAPTER 4</b>	
<b>2D FEASIBILITY DEMONSTRATION OF ADAPTIVE     REGISTRATION .....</b>	<b>28</b>
4.1. Accuracy of Triad Position .....	28
4.2. Issue of Scale .....	32
4.3. Results using 2D MR Scans with Known Deformations.....	34
4.4. Summary and Discussion.....	43
<b>CHAPTER 5</b>	
<b>COMPARISON WITH OTHER ADAPTIVE REGISTRATION     ALGORITHMS.....</b>	<b>46</b>
5.1. Overview of Adaptive Registration Algorithms .....	46
5.2. Common Framework in Adaptive registration .....	47
5.3. Other Adaptive Registration Algorithms .....	48
5.4. Experiments .....	49
5.5. Summary and Discussion.....	54
<b>CHAPTER 6</b>	
<b>ABDOMEN ATLAS .....</b>	<b>55</b>
6.1. Introduction.....	55
6.2. Probabilistic Atlas on Abdomen .....	56
6.3. Methods.....	57
6.4. Results.....	61
<b>CHAPTER 7</b>	
<b>ATLAS BASED SEGMENTATION ON ABDOMEN.....</b>	<b>65</b>
7.1. Segmentation Overview.....	65
7.2. Methods.....	67
7.3. Segmentation Results.....	72
7.4. Summary and Discussions .....	77
<b>CHAPTER 8</b>	
<b>SUMMARY AND DISCUSSIONS.....</b>	<b>79</b>
8.1. Summary and Discussion for Adaptive Registration.....	79
8.2. Summary and Discussion for Atlas Based Segmentation.....	80
8.3. Future work.....	81
<b>APPENDIX.....</b>	<b>83</b>
<b>BIBLIOGRAPHY.....</b>	<b>87</b>



## LIST OF TABLES

### Tables

5.1. Comparison of global MI and local MI .....	52
6.1. Variances of the probabilistic map in some slices .....	64
7.1. Performance measures for the sample case .....	77
7.2. Performance measures for 20 data sets .....	77

## LIST OF FIGURES

### Figure

2.1. Example of control points .....	12
2.2. Control point realization .....	14
2.3. B-spline realization .....	15
2.4. Sample Wu's radial basis function based non-linear transform .....	17
2.5. Nine control point pairs placed uniformly across the $[0\ 1] \times [0\ 1]$ squares.....	19
3.1. Flow chart of the iterative algorithm (fixed scale).....	26
4.1. Illustration of mismatched measure .....	29
4.2. Various dilation deformations in homologous image.....	31
4.3. Error vs. subblock size of mismatch metric.....	33
4.4. Images for simulation .....	35
4.5. Experiment 1 results .....	37
4.6. Intermediate stage results in experiment 1.....	38
4.7. Experiment 2 results .....	39
4.8. Intermediate stage results in experiment 2.....	40
4.9. Experiment 3 results .....	41
4.10. Intermediate stage results in experiment 3.....	42
5.1. General adaptive non-rigid registration algorithm.....	48
5.2. Data sets used.....	50

5.3. Global/Local MI vs. degree of freedom.....	52
5.4. Local mismatch measure M vs. gradient of global MI for large and small deformations.....	54
6.1. Distribution of control points for organs of interests.....	59
6.2. Comparison of probabilistic atlases.....	63
7.1.a. One slice of the data set .....	73
7.1.b. Atlas information .....	73
7.1.c. Segmentation results without atlas information.....	74
7.1.d. Segmentation results with atlas information.....	74
7.1.e. Comparison between grayscale data set and segmentation with atlas infomation.....	74
7.1.f. Difference between segmentation with and without atlas information.....	74
7.1.g. Gaussian modeling for different tissue types.....	75
7.1.h. Subsection of the ROC curve.....	75
7.1.i. Actual histogram and modeling of the histogram.....	75
7.1. Sample segmentation results.....	75
A.1. 4-point example for invertibility of TPS.....	84
A.2. Error with respect to number of control points used for inversion.....	85
A.3. Schematic for invertibility .....	86

# CHAPTER 1

## INTRODUCTION

Advances in medical imaging technologies, such as magnetic resonance imaging (MRI) and computed tomography (CT), have brought high quality data sets into reality. This in turn poses new challenges to image analysis techniques (i.e., registration and segmentation). This dissertation introduces novel approaches to improving registration and segmentation in medical image data sets.

### 1.1. Dissertation Overview

This dissertation primarily consists of three parts: adaptive registration, atlas construction of abdominal organs, and segmentation based on atlas priors. For adaptive registration, I will start with a general review of registration, including registrations with both linear (i.e., affine) and non-linear (i.e., warping) transforms paired with various similarity measures (i.e., correlation and mutual information). Next, I will explore the shortcomings of traditional non-linear high degrees of freedom (DOF) warping registrations, and suggest a way to overcome those shortcomings by an adaptive registration method. Using a new adaptive registration method, we will demonstrate its feasibility in 2D MRI scans. I also compare our adaptive registration method with other adaptive registration methods and conclude that our method is better. For the second part of the dissertation, atlas construction of abdominal organs, I first focus on a general overview of atlas construction and next focus on probabilistic atlas construction on four

abdominal organs (i.e., liver, both kidney and spinal cord) from CT scans of 32 anonymous patients. Finally, for the third part, the atlas information is used as priors to a Bayesian segmentation problem on non-contrast CT scans.

## **1.2. Adaptive Registration**

Thanks to recent advances in medical imaging systems, there are many data sets of the same patient/object from various modalities of imaging [1-14, 28]. Registration is a process that combines data sets of different or the same modalities into one pose/orientation so that the user can view them in one spatial frame. As a result, better knowledge of the patient/object is gained. Much of the past work in registration assumed rigidity between two data sets and the rigid transform that best aligns two data sets is being sought. The rigid transform is further generalized to an affine transform, which includes scaling and shearing. The rigidity assumption is true for data sets with incompressible organs like brains, bones etc.; thus, most of the past work typically concentrated on registering brain scans of several modalities [12, 28]. For organs like the liver, stomach etc., we cannot achieve a satisfactory registration with rigid transforms between data sets. As a result many non-linear transforms have been proposed. Among the notables are Thin Plate Spline (TPS), B-splines and other radial basis functions [18, 22-25]. TPS is a non-linear transform that is endorsed by a rich literature in shape statistics and morphometrics [18, 29, 22, 23]. Other non-linear transforms like B-splines and Wu's radial basis function are commonly used in the registration literature for their local characteristics and computation efficiencies [26, 27]. When computing the best transform that aligns two data sets of interest, we need an objective function to measure the "goodness of alignment," namely an objective function for similarity between two data sets under a particular transformation. There have been many similarity measures for

different registration problems but mutual information (MI) has gained considerable support over the past years for inter-modality registrations [28].

Degrees of Freedom (DOF) related to a non-linear transform can vary greatly. The more DOF we have, the more flexible the transform is. For full affine transforms, the needed DOF are 6 and 12 for 2D and 3D respectively. In the case of non-linear transforms, the needed DOF may increase tremendously depending on the geometric complexity between the data sets. For example, abdominal data sets with many soft organs, like liver and kidneys, will require 100 or more DOF. In case of the TPS, the DOF are specified by the location and number of control points. Thus the operator is required to manually identify many control points in the data sets. In the case of abdominal data sets in 3D, users are required to identify 30 or more control points (>90 DOF). This manual specification is laborious, biased and frequently inconsistent. Another way to specify high DOF is to lay a dense grid of control points, which is common in B-spline based, non-linear registrations [26]. This avoids the manual specification of control points, but the DOF associated are very high, typically in the thousands or more, thus rendering this approach very computationally expensive.

Our work tries to automatically place control points without increasing the density of control points globally. In our *mutual information-based automatic multimodality image fusion software* (i.e., MIAMI Fuse) [9, 11, 15], the manual selection of *initial* control points is the only user intervention, and in many cases, the registration process can be automated. Compared to laying a dense grid of control points, this approach has far fewer control points; thus it is computationally more efficient.

### 1.3. Abdominal Atlas

Recently, significant effort has been directed towards the development of deformable templates typically for segmentation of the human brain [31-41]. Only one recent publication addresses segmentation of the liver and kidneys using a deformable surface model [42]. Rich literature has been compiled on probabilistic atlases of the brain [43-53]; while most of the literature is extendable to abdominal organs, these extensions exist only in theory and have not been actually demonstrated on abdominal organs. Thus, we have pursued the simultaneous construction of a probabilistic atlas that includes a set of abdominal organs similar to that described by Evans [57]. The term “probabilistic atlas” does not simply mean the average boundary of an organ, but rather the complete spatial distribution of probabilities that a voxel belongs to one or more organs, i.e., each voxel is an  $n$ -vector, where  $n$  depends on the number of organs in the modeling system. We continue to focus on probabilistic atlases in spite of the remarkable, even spectacular results of deformable templates, because of the belief that atlases, if well formed, bring more prior information to the process of defining complex organs in low signal-to-noise settings. After mapping a new patient onto the atlas using the same methods used to construct the atlas, the computed inverse transform maps the probability distributions of the atlas back onto the patient. Then the atlas can be used to find the most probable loci of edges of the patient’s organ. Furthermore, the derivative of the atlas defines a Bayesian weighting (i.e., probability density function), which provides very rich information for identifying the organ’s boundaries.

In chapter 6, we present the results of a full probabilistic atlas construction for 32 patients, where 31 of the patients have been mapped onto one individual using thin plate splines (TPS) as the non-linear transform and mutual information (MI) as the similarity

measure. Registering 31 patients onto one arbitrarily chosen reference patient can introduce bias towards that specific reference patient, especially if the reference patient is far from the average population for which the atlas is intended. Thus, the reference patient selected best represents the population to our knowledge aided by an expert inspection. By iterating this construction process; i.e., using the resulting atlas as the reference target of another atlas construction phase, other groups have shown that the variance of the atlas and the dependence of the reference patient can be reduced further [44, 51]. While in the past, many groups focused only on the mean surface, the Bayesian value of the probabilistic atlas in the segmentation task lies not only in the mean but also in the variance of the atlas.

#### **1.4. Bayesian Segmentation based on Atlas**

Segmentation can be thought as the process of assigning labels to individual voxels in the data set. Extensive review of the segmentation methods and deformable surface models can be found in these papers [67, 84]. Most segmentation methods can be divided into two groups. The first one is supervised segmentation where a certain Volume of Interest (VOI) has to be manually specified for training different tissue types. The other is unsupervised segmentation where the discovery of different tissue types is automated. Both segmentation algorithms include regularization in their labeled fields. A common way to regularize is to incorporate a Markov Random Field (MRF) priors. Basically, the Markov Random Field model encourages adjacent voxels having similar labels. Our contribution is to extend the unsupervised segmentation with the additional information provided by the probabilistic atlas. Improvements in segmentation quality compared to cases with no atlas information are readily noticeable, especially where there is little contrast between organs of interest. Specifically, we can eliminate most of the soft



tissue false positives around liver/kidneys/spinal cord and distinguish liver, spinal cord, left kidney and right kidney. To apply the atlas information, a CT data set should be mapped onto the atlas space by the same method that is originally used to construct the atlas.

## CHAPTER 2

### REGISTRATION OVERVIEW

#### 2.1. Formal Statement of Registration

The goal of registration is to find the best transform  $T$  such that both data sets are best aligned by a particular similarity measure. In this paper the following notations are assumed.  $A(x)$  is the reference data set and  $B(x)$  is the homologous or floating data set.  $T(x)$  is the transform between two data sets, where  $x$  is the coordinates in 2D or 3D.

$$\hat{T} = \arg \max_{T \in F} SIM(A(\bullet), B(T(\bullet))) \quad (2.1)$$

$\hat{T}$ ; estimate of the transform

$F$ ; family of feasible transforms

$SIM$ ; similarity measure

Note that the homologous data set is mapped onto the reference data set before calculating the similarity measure. Once  $T$  is found, all the coordinates are assumed to in the reference coordinate frame since the homologous coordinate frame can always be found by applying the transform  $T$ .

#### 2.2. Similarity Measures

When computing the best transform that aligns two data sets of interest, we need a cost function to measure the “goodness of alignment”, namely a cost function for similarity between two data sets under a particular geometric transformation. There have

been many similarity measures for different registration problems ranging from a simple correlation to mutual information (MI). The MI similarity measure has gained considerable support over the past years [28].

### 2.2.1. Correlation

One of the simplest similarity measures is the sum of squared differences (SSD) between data sets.

$$SSD = \frac{1}{N} \sum |A(x) - B(T(x))|^2, \quad (2.2)$$

where  $N$  is total number of voxels in the reference data set. It can be shown that this is the optimal similarity measure when two data sets are different only by additive isotropic Gaussian noise [12]. It is quite obvious this cannot be applied to inter-modality registrations since the gray scale difference can no longer be modeled as isotropic Gaussian. Even for intra-modality registrations, this strict requirement is seldom true since noise in medical images is not always additive isotropic Gaussian. The SSD measure is very sensitive to outlier voxels that exhibit large value differences between data sets. This is quite possible when one data set is a non-contrast scan and the other is a contrast scan. Effects of these outlier voxels can be reduced by using the sum of absolute differences (SAD) instead of the SSD [28].

$$SAD = \frac{1}{N} \sum |A(x) - B(T(x))| \quad (2.3)$$

Under the strict Gaussian assumption implicit in SSD, there are very few actual applications for registration. A slightly relaxed assumption would be linear relationship

between voxel values between data sets and additive Gaussian isotropic noise. This assumption leads to the correlation coefficient (CC) similarity measure.

$$CC = \frac{\sum (A(x) - \bar{A})(B(T(x)) - \bar{B})}{\{\sum (A(x) - \bar{A})^2 \sum (B(T(x)) - \bar{B})^2\}^{1/2}}, \quad (2.4)$$

where  $\bar{A}, \bar{B}$  are mean voxel values for A(x) and B(x). This similarity measure is popular for intra-modality registrations [85].

### 2.2.2. Mutual Information (MI)

With inter-modality registrations, there is usually no simple linear relationship between voxel values of data sets. There are some similarity measures where this linear relationship is not required [28]. Among those, MI has gained considerable popularity [12-15, 28]. The MI used here is the classical Shannon mutual information. Basically, two co-registered data sets yield a joint probability density function (PDF) with tight clusters, whereas un-registered data sets yield a joint PDF with dispersed clusters. Tighter clusters (i.e., more correlation) translate into higher MI values than dispersed ones (i.e., less correlation).

$$MI(A, B) = \sum \sum p(a, b) \log_2(p(a, b) / p(a)p(b)) \quad (2.5)$$

$p(a), p(b)$ : marginal densities of A and B  
 $p(a, b)$ : joint density of A and B

While other papers used Parzen windowing for the estimation of PDFs [13, 14] in our implementation all PDFs (i.e., both marginal and joint) are estimated by non-interpolated histograms with fixed bin width. Bin width of the histogram is calculated from Freedman and Diaconis [20], where the optimal bin width,  $h_n$ , is chosen to minimize integrated

mean squared error over all squared-integrable PDFs. Optimal bin width assumes independent and identically distributed (I.I.D) voxel intensity pairs in both data sets.

$$h_n = 2(IQR)n^{-1/3} \quad (2.6)$$

$IQR$ ; inter quartile range  
 $n$ ; number of samples

### 2.2.3. Normalized MI

Note that calculation of MI or entropy requires estimating a joint PDF, which in turn is calculated from the overlap of  $A(x)$  and  $B(T(x))$  under the hypothetical transform. There are cases where the amount of overlap between the two data sets during the optimization of MI affect the final registration result [29]. Specifically, excessive overlap of background between data sets can wrongfully contribute to MI disproportionately, thus leads to errors in registration. To overcome this overlap dependency problem, Normalized MI (NMI) has been proposed to be insensitive to amount of overlap [29]. The formula for NMI is the following.

$$NMI = \frac{H(A) + H(B)}{H(A,B)}, \quad (2.7)$$

where  $H(A)$ ,  $H(B)$  are marginal entropies for  $A(x)$  and  $B(T(x))$  and  $H(A,B)$  is the joint entropy. Note that entropies are defined in traditional Shannon fashion as follows

$$H(A) = -\sum p(a) \log_2(p(a)) \quad (2.8)$$

$p(a)$ ; PDF for  $A(x)$

## 2.3. Geometric Transforms

As stated in section 2.1, there are numerous evaluations of MI or the equivalent similarity measure under the hypothetical geometric transform. Geometric transforms can be divided into two categories, linear and non-linear transforms. I will first review linear transforms and introduce commonly used non-linear transforms.

### 2.3.1. Linear – rigid and affine

The simplest geometric transform between two data sets is the rigid transform, where objects to be registered are rigid (e.g. bones), thus non-scalable. And thus, they are only allowed to translate and rotate. Clearly rigid transforms are not applicable to registration between different objects. Such a transform is strictly confined to the same object where the transformation is known to be rigid. The formulation for rigid transform is the following.

$$T(x) = A^{rot}x + b, \quad (2.9)$$

where  $A^{rot}$  is the rotation matrix and  $b$  is the translation vector in 2D or 3D. For 2D, DOF are 1 rotation parameter and 2 translation parameters, thus 3 all combined. For 3D, DOF are 3 rotations and 3 translations.

A rigid transform can be generalized to an affine transform where the object can be scaled and sheared in addition to being translated and rotated.

$$T(x) = A^{affine}x + b \quad (2.10)$$

Here  $A^{affine}$  is a full rank matrix, thus 4 DOF for 2D and 9 DOF for 3D. An affine transform can be applied to register objects that differ by rigid motion (i.e., translation

and rotation), scale and shear. There is indeed more flexibility to register with the affine transform than the rigid transform. It still does not provide enough flexibility to register common soft organs (e.g. liver) where organs can deform quite freely within a patient or across patients.

### 2.3.2. Nonlinear – Thin Plate Spline (TPS)

Non-linear transforms can supply more flexibility to register soft organs in intra-patient or inter-patient registrations. There are many ways to specify a non-linear transform. It is not feasible to cover all the non-linear transforms. Hence, I will only introduce the most commonly and widely used ones here.

The first is thin plate spline (TPS). In TPS, control points are needed to formulate a non-linear transform, which are placed in pairs on the corresponding loci of both data sets. The loci can be anatomically or mathematically identified [22, 23]. Fig. 2.1 shows a sample application of control points with anatomically identifiable control points.

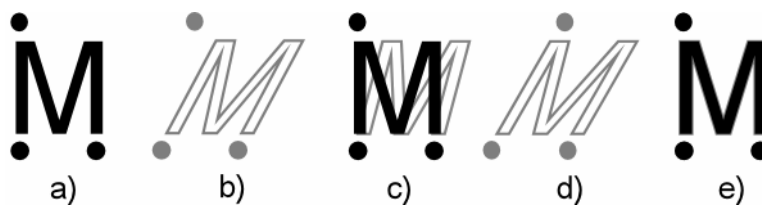


Fig 2.1 : Example of control points.

Fig. 2.1.a shows the reference image, the letter M in this case, and an associated set of 3 control points. Fig. 2.1.b shows the "floating" image to be registered with the reference and the initial loci of its 3 control points. Fig. 2.1.b is referred to as "floating" because as the geometric mapping is changed, the pose of the letter M changes relative to the fixed reference in Fig. 2.1.a. Fig. 2.1.c demonstrates the geometric mapping of Fig. 2.1.b onto a using the registration prescribed by the control points, i.e., the control points in Fig.

2.1.b map directly onto the corresponding control points in Fig. 2.1.a and map the rest of the pixels of Fig. 2.1.b into Fig. 2.1.a. Note that because the initial location of the upper left control point in Fig. 2.1.b is not homologous with its pair in Fig. 2.1.a, the resulting registration shown in Fig. 2.1.c is inaccurate. After the optimization of MI the position of the control points are automatically adjusted to obtain the registration shown in Fig. 2.1.e.

Assuming that  $x$  is the set of reference coordinates,  $x'$  is the set of homologous coordinates and that there are  $N$  control point pairs  $(x_1, \dots, x_N)$  and  $(x'_1, \dots, x'_N)$  that are manually specified as in Fig. 2.1, the formulation of TPS follows.

$$T(x) = a_0 + a_1x + \sum_{i=1}^N w_i U(r_i) \quad (2.11)$$

$a_0$ ; translation parameter  $a_1$ ; affine parameter  $w_i$ ; warp coefficients,

$$U(r) = \begin{cases} r^2 \log(r^2) & \text{in 2D} \\ |r| & \text{in 3D} \end{cases}$$

where  $x'=T(x)$  is the transformation,  $U(r)$  is the basis function,  $r_i$  is the Euclidean distance between  $x_i$  and  $x$ , (i.e.,  $|x - x_i|$ ). A detailed description of TPS and control points can be found in [22, 23]. One important property of TPS is that TPS minimizes a value called bending energy as in equation 2.12, which represents smoothness of the transform. So with TPS, one can produce a smooth non-linear transform with few control points (i.e., DOF) compared to other non-linear transforms. TPS is a non-linear transform that is endorsed by a rich literature in shape statistics and Morphometrics [18, 19, 22, 23]. Recently it has been found that TPS is the Maximum Likelihood Estimator (MLE) for unspecified landmarks not identified as control points [19]. DOF in TPS are  $2N$  or  $3N$  in 2D and 3D respectively, where  $N$  is the number of control points. If the number of control point pairs is less than 4 in 2D and 5 in 3D, the TPS automatically reduces to an affine solution, thus it is possible to specify affine transforms through the use of control points instead of explicit parameters, like translation and rotation. We use control points for both affine and non-linear transforms to maintain coherency of the user interface.



$$BE = \iint_{R^2} \left( \left( \frac{\delta^2 z}{\delta x^2} \right)^2 + 2 \left( \frac{\delta^2 z}{\delta x \delta y} \right)^2 + \left( \frac{\delta^2 z}{\delta y^2} \right)^2 \right) dx dy \quad (2.12)$$

BE; bending energy in 2D

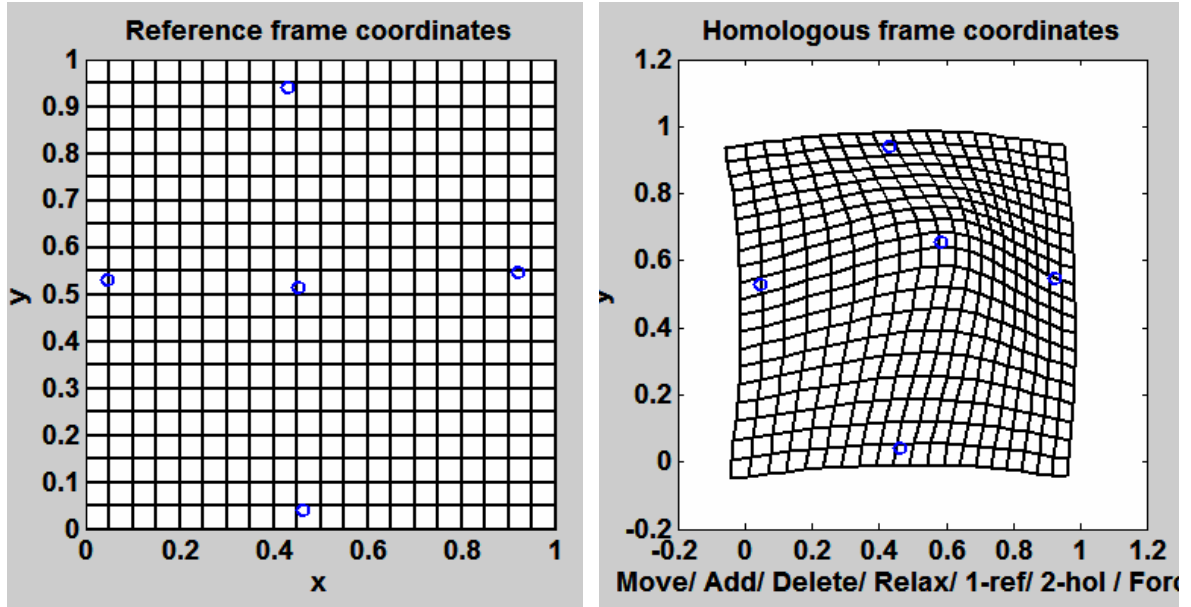


Fig. 2.2. Control point realization.

Five control point pairs are placed in “+” shape across the  $[0 \ 1] \times [0 \ 1]$  square. Control point in the middle in the homologous side is moved towards northeast direction. This figure show that a smooth transform is formed by specifying only 5 control points (i.e., 5 knots of 10 DOF).

### 2.3.3. Nonlinear – B-splines

The most widely used non-linear transform for image registration is the B-spline representation [24, 26]. In B-spline representation, a non-linear transform is represented by a sum of third order B-spline polynomials at a uniformly spaced grid. Assuming the  $x = (x_1, x_2, x_3)$ , a coordinate in 3D, following is the formulation of B-spline geometric transform.

$$T(x_1, x_2, x_3) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 B_l(u) B_m(v) B_n(w) \phi_{l+j} \phi_{j+m} \phi_{k+n} \quad (2.13)$$

$$B_0(u) = (1-u)^3 / 6$$

$$B_1(u) = (3u^3 - 6u^2 + 4) / 6$$

$$B_2(u) = (-3u^3 + 3u^2 + 3u + 1) / 6$$

$$B_3(u) = u^3 / 6$$

where  $\Phi$  is the uniformly spaced grid of B-splines, and  $B_0$  to  $B_3$  are the third order B-spline polynomials [26]. DOF in B-splines are  $2N$  or  $3N$  in 2D and 3D respectively, where  $N$  is the number of B-splines used. For B-splines, the support of the transform is local; i.e., if a point in a given grid moves locally, then the transformation changes in that local area only. Thus for a local change of a grid, the transform needs to be updated only locally, and it is thus computationally efficient. Fig 2.3 is the actual instance of B-spline transformation.

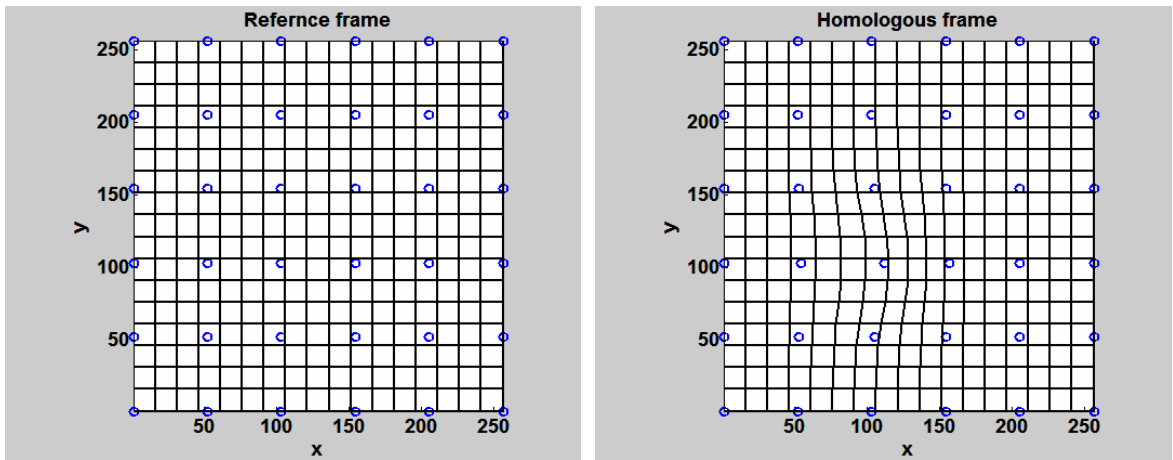


Fig. 2.3. B-spline realization.  $6 \times 6 = 36$  B-splines are used. One spline around the middle is translated to right on the homologous side. Notice the local support property of the transform.

### 2.3.4. Nonlinear – Wu’s radial basis function

Another notable non-linear transform is formed by using Wu’s radial basis function [61, 62]. One problem with TPS is that the support of the transform is not strictly local as with in B-splines. Thus, it takes more computation power to update a TPS transform than a B-spline transform given that the change in geometric transform is local. This is indeed a big problem because in the course of optimization of the similarity measure, geometric transforms are constantly evaluated. Wu’s radial basis function has finite support built in, thus it has the efficient strict local property. The formulation is exactly the same as TPS except that the basis function is now replaced by Wu’s radial basis function. With this approach, the achieved non-linear transform is less smooth than TPS, thus it does not minimize the bending energy.

$$T(x) = a_0 + a_1x + \sum_{i=1}^N w_i U(r_i) \quad (2.14)$$

$a_0$ ; translation parameter  $a_1$ ; affine parameter  $w_i$ ; warp coefficients

$$U(r) = R(r/s)$$

$$R(r) = (1-r)_+^4 (3r^3 + 12r^2 + 16r + 4)$$

$$(1-r)_+ = \max(1-r, 0)$$

where  $r_i$  is the Euclidean distance between  $x_i$  and  $x$  and  $s$  is the radius support of the basis function. Fig 2.4 is the sample realization of Wu’s radial basis function based non-linear transform.

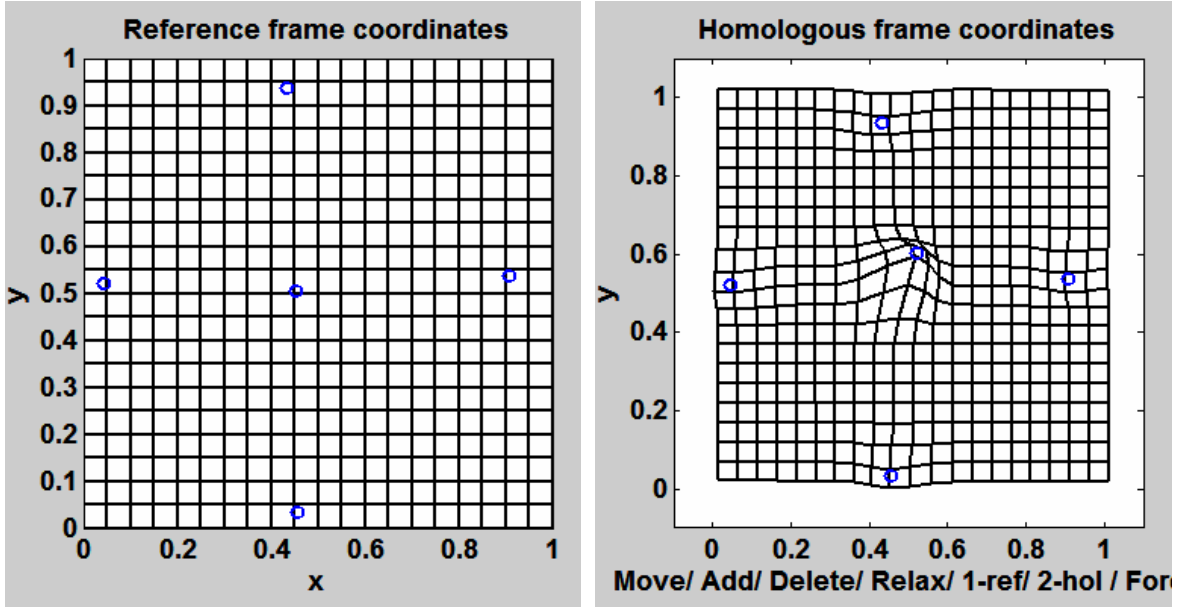


Fig. 2.4. Sample Wu's radial basis function based non-linear transform. Five control point pairs are placed in “+” shape across the [0 1]x[0 1] square. Control point in the middle in the homologous side is moved towards northeast direction.

### 2.3.5. Viscous Fluid Model

Christensen et al. [30, 58-60] have modeled the non-linear transforms as non-linear displacements at every voxel in the data set. The deformation field is set to satisfy certain partial differential equations borrowed from continuum mechanics for deformable bodies. The mechanical model is the Navier-Poisson fluid. In this approach, all voxels act as control points, thus the overall transform can be very flexible.

$$\begin{aligned}
 T(x) &= x + u(x, t) & (2.15) \\
 v &= \frac{\partial u}{\partial t} + \sum_{i=1}^3 v_i \frac{\partial u}{\partial x_i} \\
 u \nabla^2 v + (\lambda + u)(\nabla \cdot v) + b(u) &= 0 \\
 b(u) &= -\alpha [B(x + u(x, t)) - A(x)] \nabla B \Big|_{x+u(x, t)}
 \end{aligned}$$

where  $u(x,t)$  is the displacement,  $v(x,t)$  is the velocity,  $\lambda$  is the rate of change in the partial differential equation (PDE) and  $b(u)$  is the body force that drives the PDE. Note that the PDE is dependent on the body force  $b(u)$ , which in turn is dependent on differences in gray scale values and gradient of the homologous image. Thus, this approach is well suited for intra-modality registration. There have been some attempts to change  $b(u)$  such that it will work for inter-modality registrations [86]. Assuming that  $u(x,t=0)=0$ , set of PDEs are solved iteratively until  $b(u)$  falls below some threshold.

#### **2.4. Local Effects of Control points in TPS**

With B-splines and Wu's radial basis function cases, the non-linear transforms are constructed to have the local support property. But with TPS, the local support property is not strictly true because the basis function is globally defined everywhere. Still TPS is fairly local if the DOF are high enough. This is a commonly misunderstood point regarding TPS. While the basis functions are proportional to a positive integer power of  $|r|$ , these basis functions cause the effect, i.e., the weighting of control points to be inversely related to the distances from them. Fig. 2.5 clearly shows that effects of control points are local.

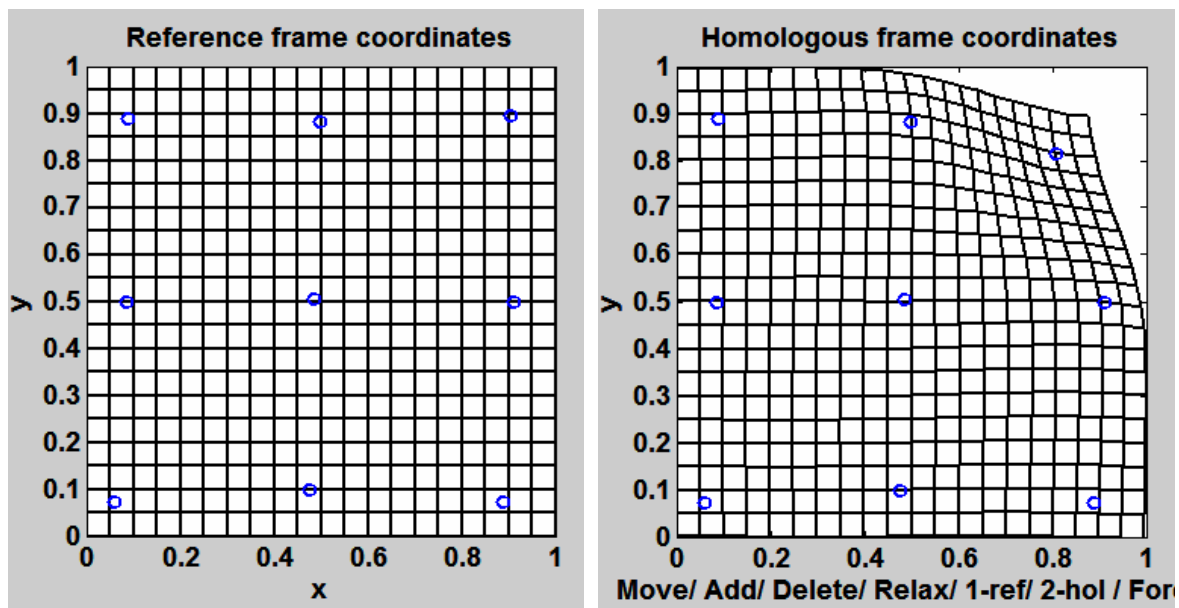


Fig. 2.5. Nine control point pairs are placed uniformly across the  $[0\ 1] \times [0\ 1]$  square. Control point of the northeast side in the homologous side is moved towards southwest direction. Effects of control points of moving the control point in the northeast region are clearly localized to the northeast region.

## 2.5. MIAMI Fuse

Our research group has compiled a set of registration software packages based on Advanced Visual System 5 (AVS5) toolkit. Our implementation is MI-based TPS non-linear registration and we call it Mutual Information-based Automatic Multimodality Image fusion (MIAMI Fuse) [9, 11, 15].

## CHAPTER 3

### ADAPTIVE REGISTRATION

In this chapter, we will explore the shortcomings of traditional non-linear high DOF warping registrations, and suggest an alternative which overcomes those shortcomings by using an adaptive registration method where DOF are selectively increased. This chapter is based on Park and Meyer [87].

#### **3.1. Shortcoming of high DOF non-linear registrations**

The degrees of Freedom (DOF) related to a non-linear transform can vary greatly. The more DOF we have, the more flexible the transform is. For full affine transforms, the needed DOF are 6 and 12 for 2D and 3D respectively. In the case of non-linear transforms, the needed DOF may increase tremendously depending on the geometric complexity between the data sets. For example, abdominal data sets with many soft organs, like liver and kidneys, will require 100 or more DOF. In case of the TPS, the DOF are specified by the location and number of control points. Thus the operator is required to manually identify many control points in the data sets. In the case of abdominal data sets in 3D, users are required to identify 30 or more control points (>90 DOF). This manual specification is laborious, biased and frequently inconsistent. Another way to specify high DOF is to lay a dense grid of control points, which is common in B-spline based, non-linear registrations [26]. This avoids the manual specification of control points, but the DOF associated are very high, typically in the thousands or more, thus

rendering this approach very computationally expensive. In Reuckert's case [26], final registration involves a dense grid of regularly spaced B-splines, whose DOF are determined by the number of B-splines used and is on the order of hundreds to thousands. While many algorithms have demonstrated their ability to register using large DOF quite successfully [15, 26], this additional flexibility requires increased computational power, and leads to potential convergence problems.

### **3.2. General Approach for Adaptive Registration**

Our work tries to automatically place control points without increasing the density of control points globally everywhere. In our mutual information-based automatic multimodality image fusion software (i.e., MIAMI Fuse) [9, 11, 15], the manual selection of *initial* control points is the only user intervention, and in many cases, the registration process can be automated. The central issue here is to find where to increase the density of control points. That issue will be addressed using local information measures. Compared to laying a dense grid of control points, our method will result in irregularly spaced control points with far fewer number of control points; thus it is computationally more efficient.

### **3.3. Local Mismatch Measure M**

Our goal is to increase DOF selectively in regions where they are needed rather than increase DOF globally in the data set. In context of TPS, we are trying to increase the density of control points where additional control points are needed. Noting that effects of control points are local, a good candidate region for increased control point density is the region with the largest local mismatch under the current geometric transform. Regional mismatches can be detected by assessing regional/local MI.



Mismatched local regions are poorly “correlated” by definition, and thus have low regional MI. Additionally the regional MI between two regions can also be low if the entropy of either of the corresponding regions is low because MI is bounded by the entropies of the individual data sets (i.e.,  $MI(a,b) \leq \min\{H(a),H(b)\}$ ). Because regions of low entropy are basically featureless, they are not good candidate loci for additional placement of control points. Therefore, local regions of interest (ROI) for the placement of additional control point pairs are confined to those with high local entropy and low local MI. We propose a normalized measure of mismatch  $M$ ,  $M = 1 - MI(a,b)/\min\{H(a),H(b)\}$ , formed from local measures. The mismatch measure is largest (i.e., close to 1) when MI is zero or MI is relatively smaller than  $\min\{H(a),H(b)\}$ , which implies locally mis-registered area with some texture; also note that the mismatch measure goes to zero as MI increases, which implies a locally registered area. Use of  $\min\{H(a), H(b)\}$  in the denominator is to provide a tighter bound than  $H(a)$  or  $H(b)$  at the cost of mathematical tractability, i.e., continuity. Thus, regions with large local mismatch measure  $M$  are good candidates for an additional placement of control points. The mismatch measure always lies between 0 and 1. We are not claiming that above approach is the optimal way to quantify a local mismatch. In summary, we have just proposed the following method to quantify a local mismatch:

$$M = 1 - \frac{MI(a,b)}{\min(H(a),H(b))}. \quad (3.1)$$

The way regional/local MI or entropy is calculated is a direct extension from the calculation of global MI/entropy by histogram. To calculate a local MI at some given location, a subblock of voxels centered at the given location is defined both on the

reference data set and homologous data set under the existing transform between two data sets, then marginal and joint PDFs are calculated from the joint histogram formed from both subblocks. For chapters 3 and 4, the subblock size is fixed to  $41 \times 41$  (i.e., 1681 voxels) and the number of bins for the histogram is set by assessing optimal bin width from Izenman [20] from equation 2.6, which depends on the data distribution in each subblock. Note that the subblock size cannot be too small because with only a few samples available, PDF estimation by histogram becomes unreliable. Since we need to assess local MIs for all subblocks across the image, we adaptively change bin widths for each subblock.

### **3.4. Triad of Control Points**

After identifying a candidate mismatched region to increase DOF, we must determine how many additional control points are needed. In fact, we may need more than just a single control point depending on the deformation present in the mismatched region. Since control points have local effects, to compensate for a local translation we primarily need one control point pair, but for local 2D rotate/scale/shear/translate (i.e., affine) compensation we need 3 local control point pairs. Thus, in the 2D registration problem, we put a triad of control point pairs in the mismatched region. Assuming that distances between 3 control points in the triad are significantly less than the average distance between previously existing control points, adding a triad hardly affects much of the existing long range geometric transform since distances from the 3 control points of the triad to the existing control points are almost identical and movements of the 3 control points with respect to the center of the triad tend to be differential with no net centroid shift. Maintaining the existing long range transform is vital since we are interested in correcting local deformations without affecting the long range transform.

The locus of greatest mismatch serves as the center of the triad and the 3 control points are distributed in an equilateral triangle fashion. The larger the spacing between control points in the triad, the larger the region the triad will affect. This is the fundamental issue of scale that will be addressed further in the separate scale issue section, but for the remainder of this chapter we will address a single level of scale, which translates into some small fixed size of the triad. Note also that an affine transform can be represented by triads of control point pairs of any orientation. Thus the orientation of the additional triad that goes into the mismatched region does not matter since from a piecewise linear viewpoint, we are trying to correct a local affine deformation in that region.

### **3.5. Iterative algorithm**

Given a current geometric mapping between two data sets, we add one triad of control point pairs at a time to the existing control point pairs to increase DOF locally. For each iteration, a triad of control point pairs is added. To determine where to put a new triad of control point pairs given the current transformation, the region with highest mismatch measure as described previously is located in the reference coordinate frame. Note that when adding a triad, we are confining it to be apart from the existing control points since it is meaningless to put a triad to places where control points already exist. This is done by imposing a constraint to the minimum distance between the center of the triad and existing control points. In our implementation, the minimum distance is set to be twice the triad size. A formal statement of the center location of the triad is the following.

$$x_{new} = \arg \max_{x, d_{min} > 2 * \text{triad size}} \left\{ 1 - \frac{MI(a, b)}{\min(H(a), H(b))} \right\} \quad (3.2)$$

$x_{new}$ ; center position of the triad

$x$ ; coordinates in the reference frame

$d_{min}$ ; minimum distance between control points

Once the reference triad center locus is found, a local search for the placement of the corresponding center locus in the homologous data is performed. If we do not search on the homologous side, adding control point pairs under the current mapping will always lead to a local mismatch since we intentionally selected the region with the most mismatch. A window of size 30x30 at the center of the added triad on the reference side is defined in the search for the corresponding center and a 30x30 pixel window is moved in the homologous image within a 40x40 pixel search window centered around the current mapping of the new reference triad center onto the homologous image. The local MI is then calculated between the reference and homologous windows for each position of the homologous window using the same adaptive bin width joint histogram method previously described; the center of the homologous window that yields the highest local MI is chosen as the center location of the corresponding new triad in the homologous image. The window size needs to be similar to the subblock size from the mismatch measure considering that local MIs of the same position may change when number of voxels from which local MIs are calculated change. In addition, it is beneficial to confine the search size according to the triad size since smaller triads create more local deformations which leads to smaller search size.

The size of triad (i.e., distance between vertex and its center position of the equilateral triangle) is fixed to 3 mm to operate at a single chosen level of scale in the 240x240 mm<sup>2</sup> image. After this new triad of control points is added, registration is

repeated by using the MIAMI Fuse and the final global MI is recorded. If the resulting global MI has not increased more than some user defined threshold value, the iterative process of adding control points stops by discarding the last added triad pairs and falls back to the control point configuration of the previous iteration. Below is the flow chart for the overall algorithm at a single level of scale.

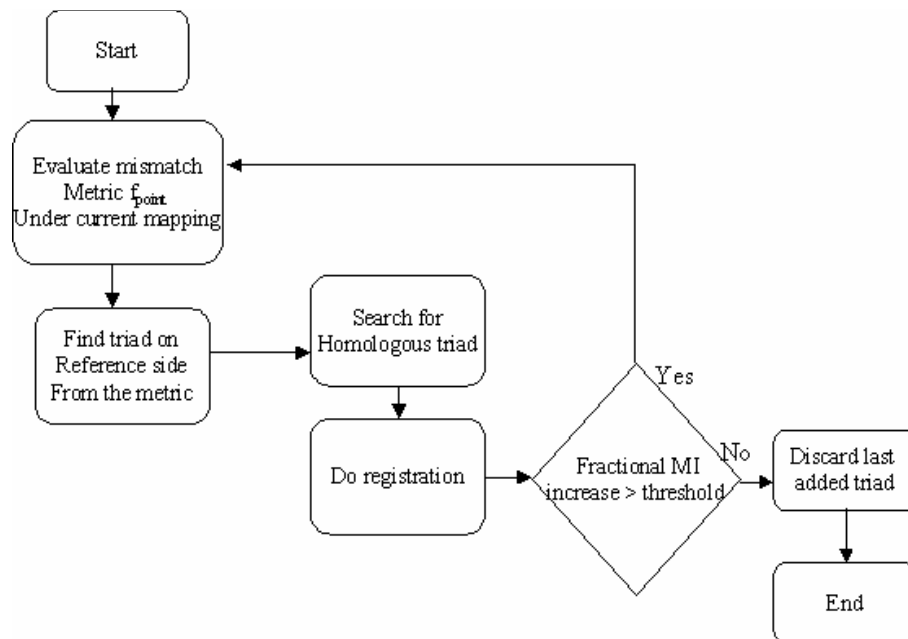


Fig. 3.1. Flow chart of the iterative algorithm (Fixed scale)

For the first iteration, an affine registration is computed to give the algorithm a starting transform. Although the user's initially defined control point pairs are used to compute the affine transformation, because the affine transform is computed globally, the same transformation would be computed given any reasonable set of initial control points. When transitioning from the affine transform as defined by 3 control points in the 2D case to the non-linear TPS transform (i.e., number of control points > 3 for 2D), not only should the 4th control point pair be defined as previously described, but also should

the user defined initial 1-3rd control point pairs be replaced. Since there is no existing TPS transform at this stage, adding a triad will not work. Note that effects of control points are most global when control points are far apart due to the fact that the distal control points control the warping in each of their respective locales. Since global deformations need to be addressed before smaller and local deformations, the second iteration's control points are centered in each of 4 quadrants. For third and later iterations, triads of control points are added by the iterative algorithm.

## CHAPTER 4

### 2D FEASIBILITY DEMONSTRATION OF ADAPTIVE REGISTRATION

In this chapter we will demonstrate effectiveness of the proposed adaptive registration algorithm of chapter 3 in the case of MR data sets with known deformations. We will try to recover the known deformation with our adaptive registration approach. This chapter is based on Park and Meyer [87].

#### 4.1. Accuracy of Triad Position

We first illustrate the actual measure of mismatch between two MR images. First, two data sets previously well registered are obtained (i.e., T1 and T2 weighted MR images of visible male from National Library of Medicine). Then one of the pair is deformed in a known way, and the algorithm previously described in Chapter 3 is used to recover the known deformation. To reduce noise, a median filter of 5x5 is applied before assessing local MIs and entropies.

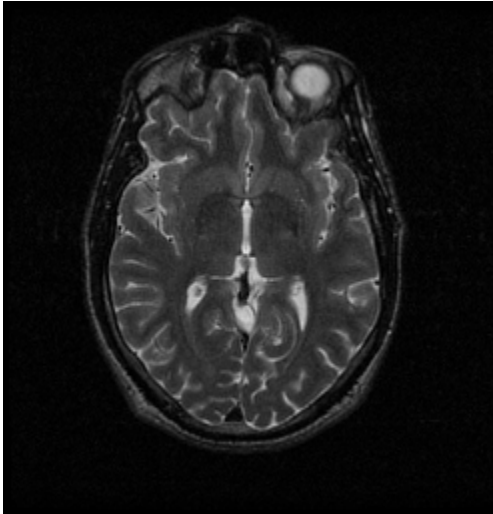


Fig. 4.1.a: Reference image

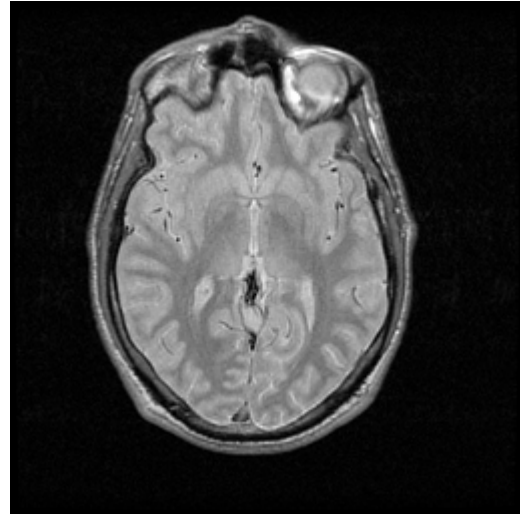


Fig. 4.1.b: Initial homologous image

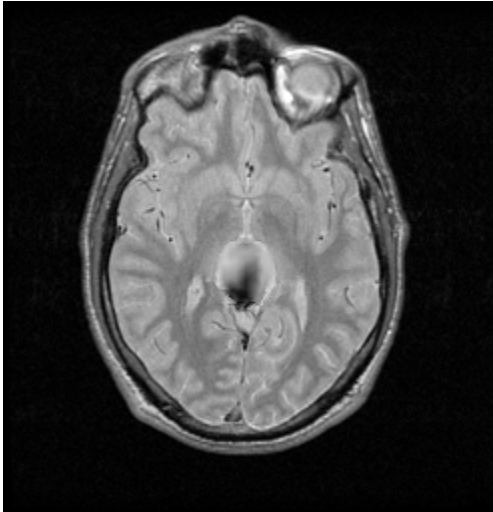


Fig. 4.1.c: Distorted homologous image; radial dilation is applied to the center of image

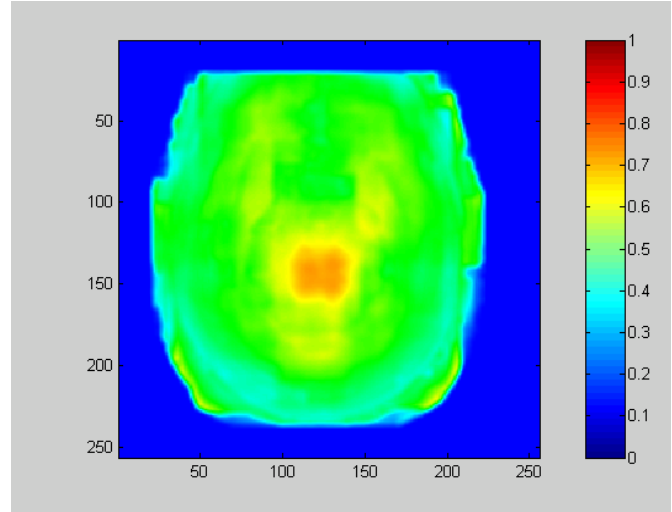


Fig. 4.1.d: Mismatch measure computed between Fig. 4.1.a and 4.1.c.

Fig. 4.1. Illustration of mismatch measure.

Blue represents low values and red represents high values. Note that mismatch is pronounced around the center.

To demonstrate that the choice for the center of the added triad by the algorithm of Fig. 3.1 is accurate, a known dilation is introduced in various areas of the homologous image and the resulting center locations of the triad placements from the algorithm are recorded. The radial dilation used here is defined by the following formula.

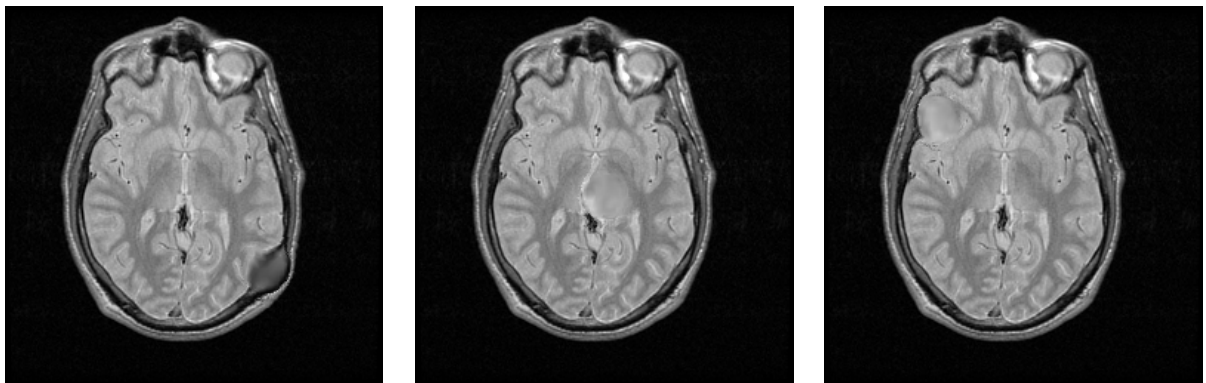


$$d(r) = r(5e^{-r/5} + 1) \tag{4.1}$$

$r$ ; radius measured from center of subblock to be deformed

$d(r)$ ; newly assigned radius from center of subblock

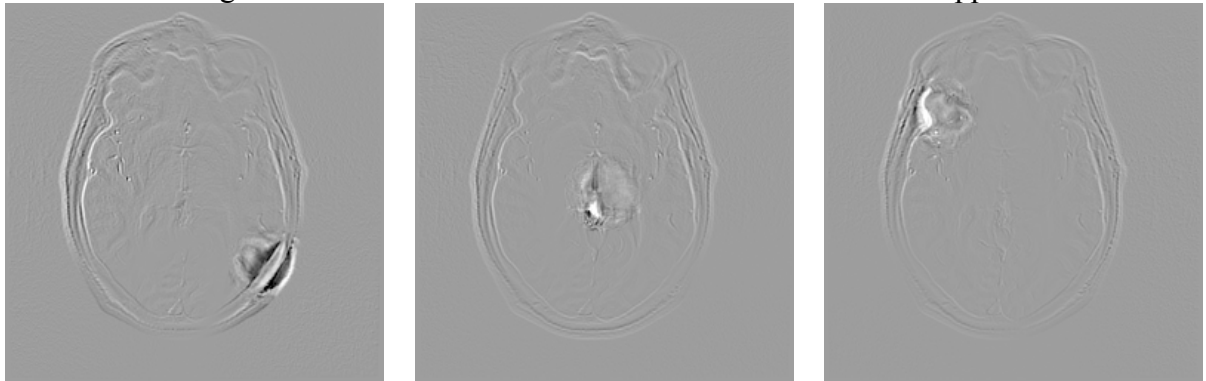
The deformation is applied to a 51x51 subblock. Differences between the loci of the center of the subblock that is deformed and the center of the added triad in the reference frame are tabulated. Two errors,  $(\text{err}_{\text{horizontal}}, \text{err}_{\text{vertical}}) = x_{\text{triad}} - x_{\text{true}}$ , where  $x_{\text{true}}$  is the true center of dilation and  $x_{\text{triad}}$  is the center of the triad, are recorded. The following are the 3 examples taken from the 30 dilations deformations used here.



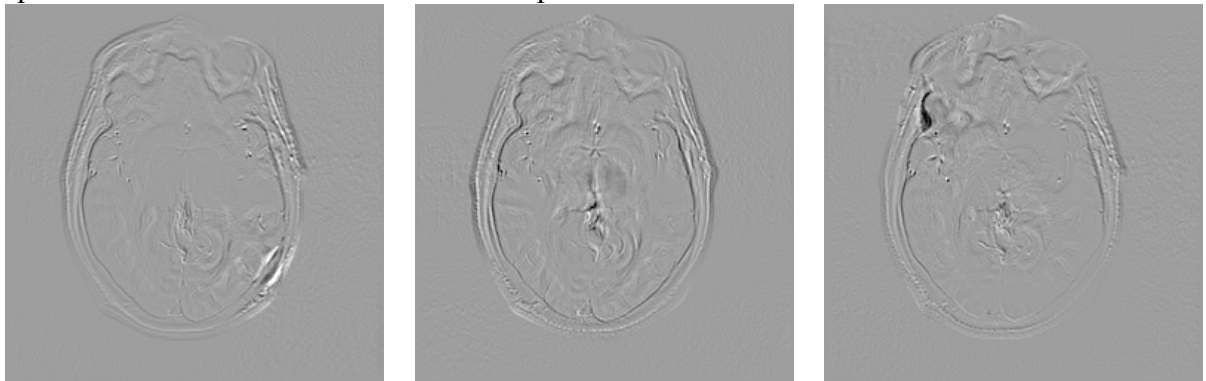
dilation: lower right corner

dilation: center

dilation: upper left corner



Difference between original image and unwarped image using 9 control points uniformly spread out. Note that there are no control points at locus of local dilation.



Difference between original image and unwarped image using triad at locus of maximum mismatch. Note that there are existing control points other than the added triad. Also note that dilations are significantly less than the previous row of images.

Fig. 4.2. Various dilation deformations in homologous image.

For horizontal errors, the mean and standard deviation are 0.2667 and 5.2780 mm respectively. For vertical errors, the mean and standard deviation are 0.2333 and 4.6586 mm respectively. Note that the errors are not significantly different from zero. Thus the algorithm is able to detect the center of dilation reliably with small errors and no demonstrated bias. Images in Fig. 4.2 demonstrate that by placing a triad where maximum mismatch occurs, the dilation can be significantly reduced. Though radial dilation is just one form of countless possible deformations, the ability to correct dilation strongly suggests the capability to correct other types of deformations.

#### **4.2. Issue of Scale**

Recall that when calculating the mismatch measure, both local MI and entropies are calculated on a subblock of size 41x41 using a joint histogram with adaptive bin width. We have tested the sensitivity of the triad center error with respect to the subblock size on which the mismatch measure is calculated. In short, we have measured errors between triad center and dilation center by varying the subblock size of the mismatch measure. The first nine dilation deformations out of 30 are used to plot the average of error magnitude below.

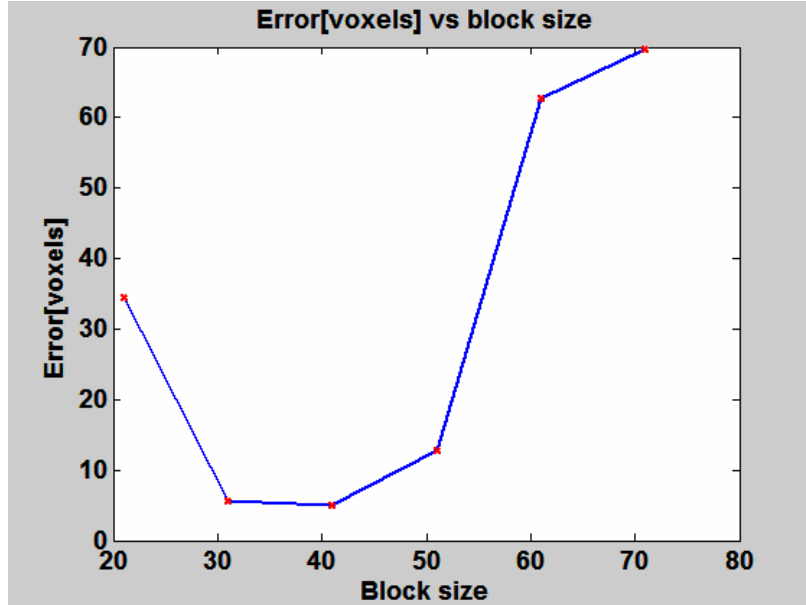


Fig. 4.3. Error vs. subblock size of mismatch metric

From Fig 4.3, it is evident that there is some optimal subblock size that leads to minimum error. Note that from the dilation equations, the maximum radial displacement occurs at the diameter of 10 pixels. At the diameter of 27 pixels the radial displacement is 50% of maximum and at the diameter of 40 pixels the radial dilation is still 20% of maximum. By 40 pixels 80% of dilation has occurred, thus obtaining the smallest error from subblock size 41 is very reasonable. If the subblock size is larger than the size of deformation, then the mismatch measure will find the correct center of deformation less reliably since the mismatch measure is calculated over larger subblock than the deformation, and thus the peak location where the maximum mismatch occurs is blurred (i.e., obscured). For similar reasons, the mismatch measure based on smaller subblock size than the deformation may lead to multiple local maxima loci within the same deformation. In result, there are scale issues that can be evaluated with “matched filter like” approaches. Thus, the mismatch measure can reliably detect deformations whose sizes are similar to sizes of the subblock used to compute the mismatch measure. As

noted earlier, triad size is directly related to the size of the deformation the algorithm is intended to correct. In summary, two things are determined by the scale of the deformation; triad size and subblock size of the mismatch measure.

### 4.3. Results using 2D MR Scans with Known Deformations

Simulations are performed to see how the algorithm actually works where the deformations are not dilations. First, two data sets previously well registered are obtained (i.e., again T1 and T2 weighted MR images of visible male from National Library of Medicine). Then one of the pair is deformed in a known way, and then the algorithm previously described is used to recover the known deformation. The known deformation used here is a 9-point TPS. The T2-weighted MR image is chosen as the reference data set and the T1-weighted MR image as homologous data set.

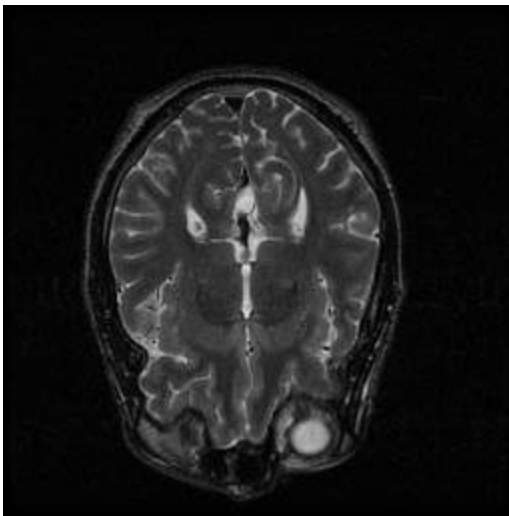


Fig. 4.4.a: Reference image – T2 MR

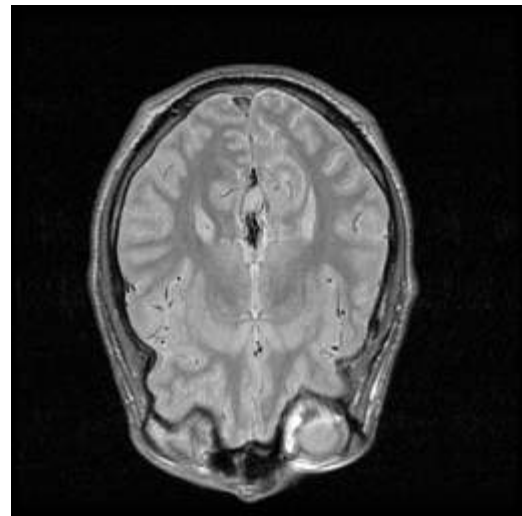


Fig. 4.4.b: Homologous image – T1 MR

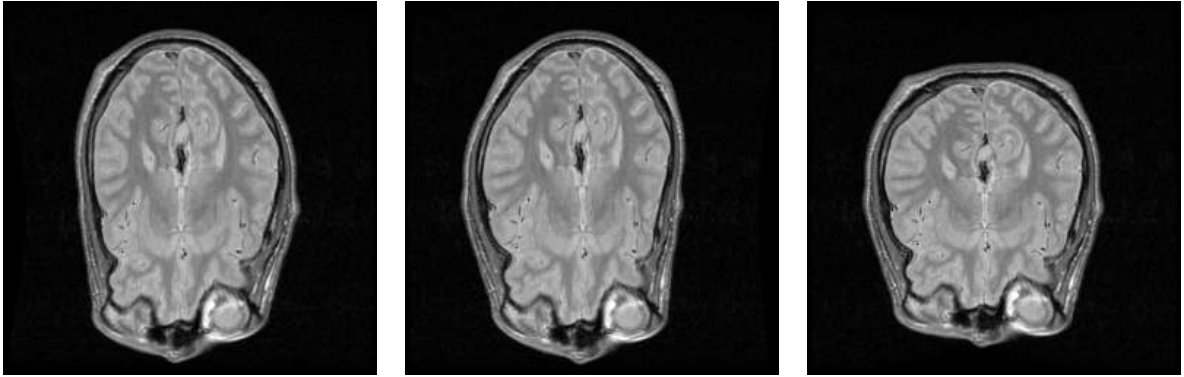


Fig 4.4.c: Three initially distorted homologous starting images

Fig. 4.4. Images for simulation

The upper bound on achievable global MI for the algorithm can be determined from the global MI of the undistorted co-registered pair of reference and homologous images. Note that our registration software uses TPS, but with a finite number of control points, it is impossible to invert or recover the TPS induced deformation exactly. More on the topic of invertibility of prior TPS deformations can be found in the appendix. The algorithm tries to find a set of control points that approximately recovers the deformation. Three experiments have been performed. For each experiment triad size is fixed to 3[mm] and fractional MI threshold is set to 1%. A median filter of size 5x5 is applied to reduce noise when calculating regional entropies, MIs and the global MI. To compare different global MIs from different experiments, the adaptive bin width technique is not applied, instead the global MIs are measured with fixed bin width across all three experiments. For the reference image 88 bins are used and for homologous images 40 bins are used. The former value comes from using adaptive bin width approach and the latter value comes from an average of adaptive bin width approach across different homologous data sets to eliminate this possible confounding effect. For the first iteration, an affine registration is performed. For the second iteration a 4-point-warp is done. For the third and later iterations a triad of control points is added according to the algorithm previously

described. For each experiment global MI values are recorded as a function of iteration for the goodness measure of registration. Since we have the “ground truth” displacement/deformation field for the experiments, average root mean square displacement error is also calculated over the uniformly sampled grid spaced at 0.9375[mm] (i.e.,1 voxel) excluding regions with little information (i.e., excluding regions outside of the head). This type of displacement error is relevant considering that we are only interested in recovering deformation from regions where there is support to observe the deformation. The experiments comparing between globally increasing DOF and our algorithm and the comparison between adding a triad and adding a single point as shown in Fig. 4.5 are done in the same manner.

In Experiment 1, deformations are introduced in the left and upper part of the image. With 4 iterations, algorithm stops with 10 control points and recovers 94.4% of the MI's upper bound. Effective displacement error after 4 iterations is 0.39mm.

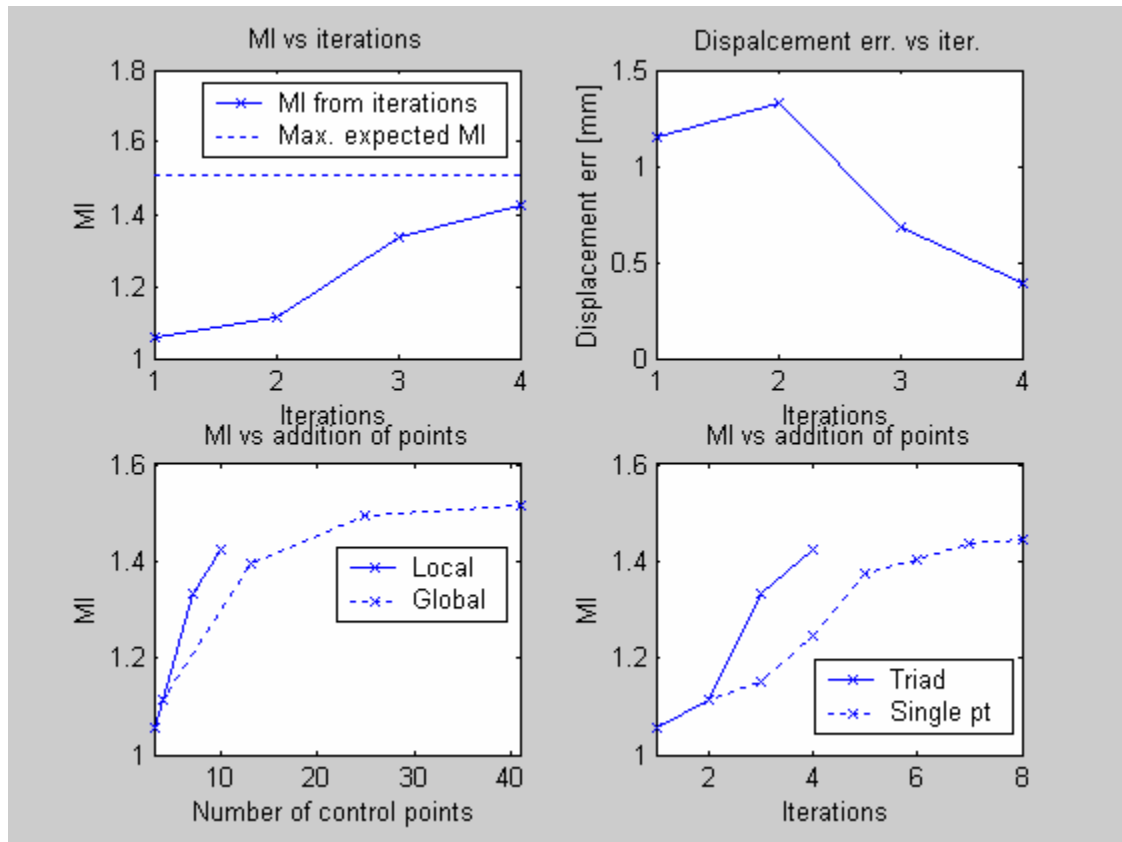


Fig. 4.5. Experiment 1 result

From the Lower left plot, advantages of our algorithm are evident that we achieved almost the same global MI without adding so many control points. From Lower right corner plots, with adding a triad we are able to converge faster than adding a single point.



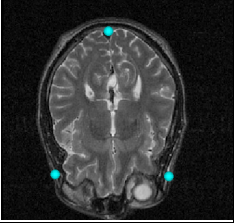
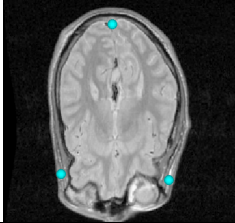
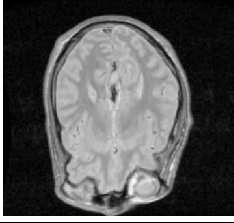
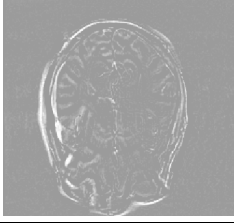
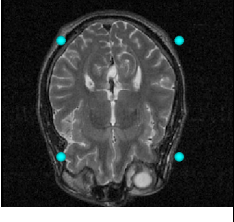
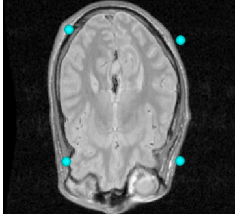
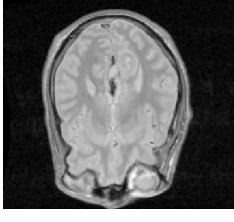
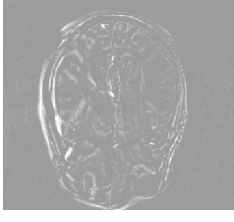
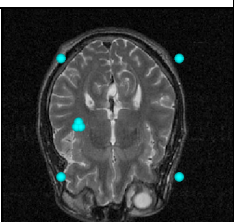
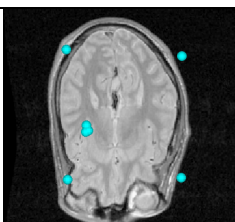

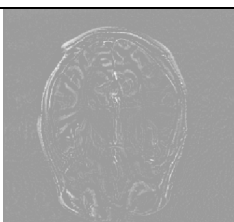
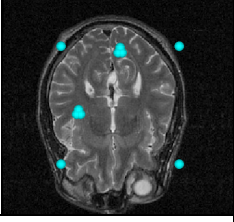
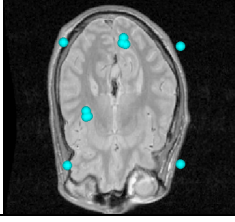
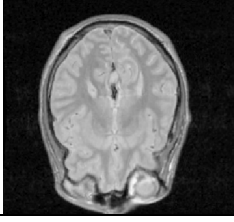
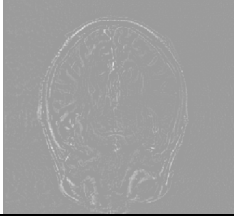
Iteration 1				
Iteration 2				
Iteration 3				
Iteration 4				
	Reference image with control points	Initial homologous image with control points placed incrementally as described in text	Recovered (unwarped) homologous image after optimizing global MI	Difference image between recovered homologous and original homologous image

Fig. 4.6. Intermediate stage results in Experiment 1

For Experiment 2, deformations are introduced to homologous image in the upper region and right region. After 4 iterations with 10 control points, algorithm recovers 93.0% of the MI's upper bound. Effective displacement error after 4 iterations is 0.56mm.

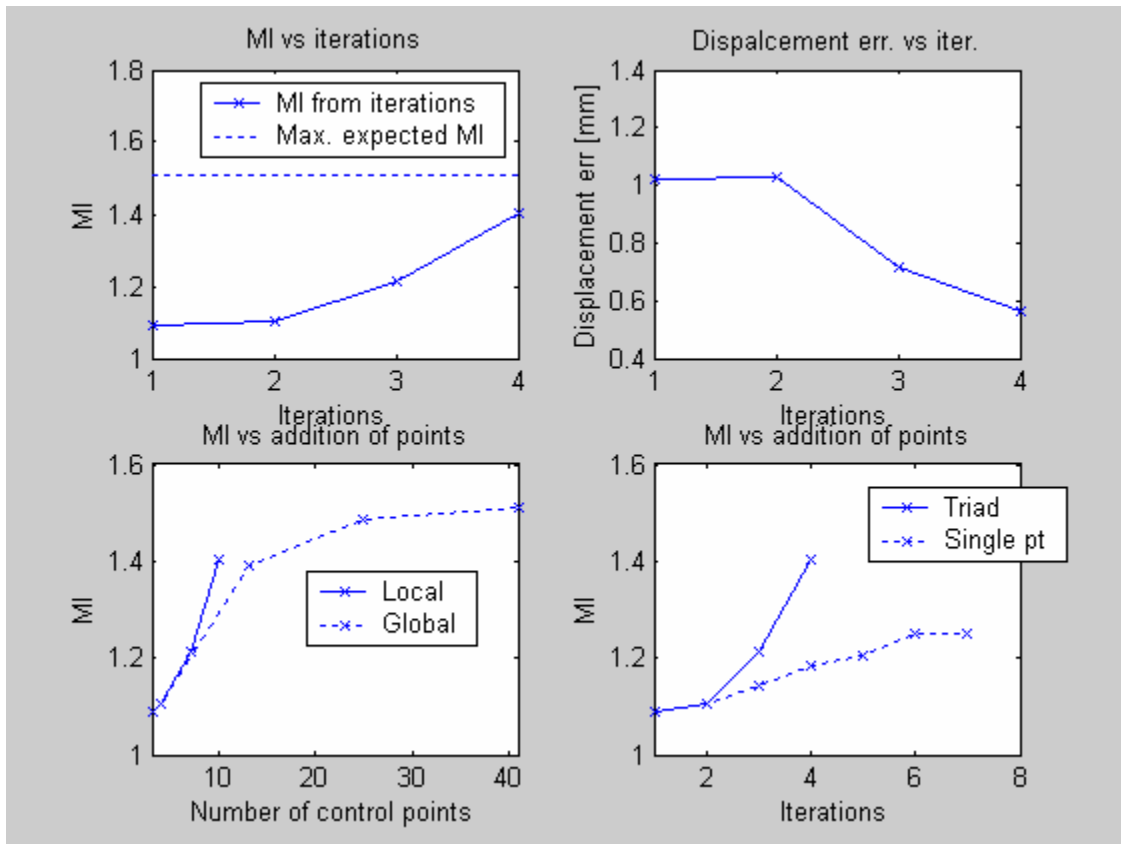


Fig. 4.7. Experiment 2 results

In this case adding a single point did not achieve as high a global MI as adding a triad. This shows that in some instances, we need more than a single point to correct a local deformation.

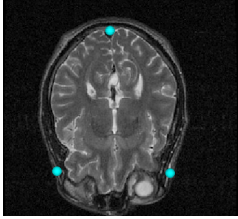
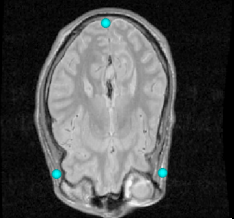
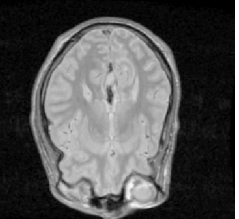
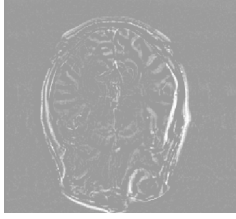
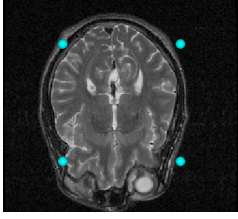
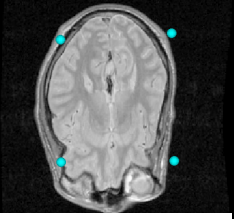
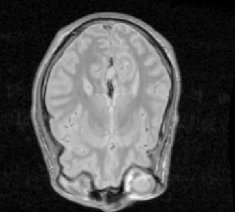
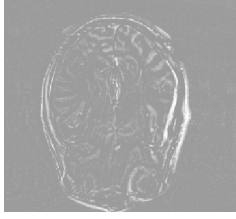
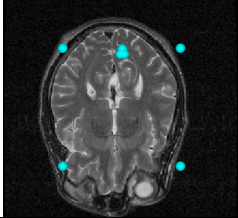
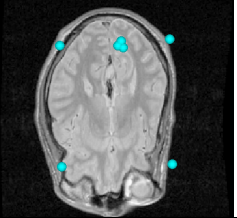
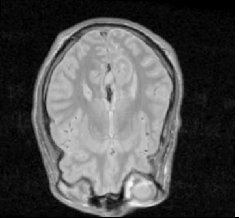
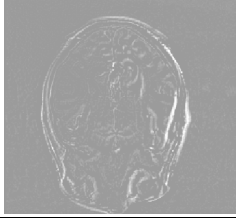
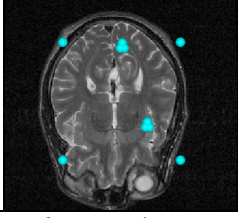
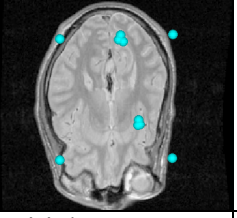
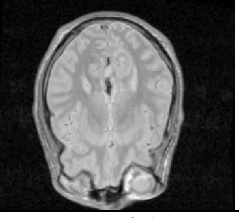
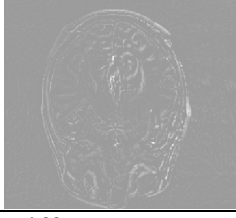
Iteration 1				
Iteration 2				
Iteration 3				
Iteration 4				
	Reference image with control points	Initial homologous image with control points placed incrementally as described in text	Recovered (unwarped) homologous image after optimizing global MI	Difference image between recovered homologous and original homologous image

Fig. 4.8. Intermediate stage results for Experiment 2.

For Experiment 3, deformations are introduced to the homologous image in the lower and upper region. After 5 iterations with 13 control points, algorithm recovers 93.3% of MI's upper bound. Effective displacement error after 5 iterations is 0.51mm.

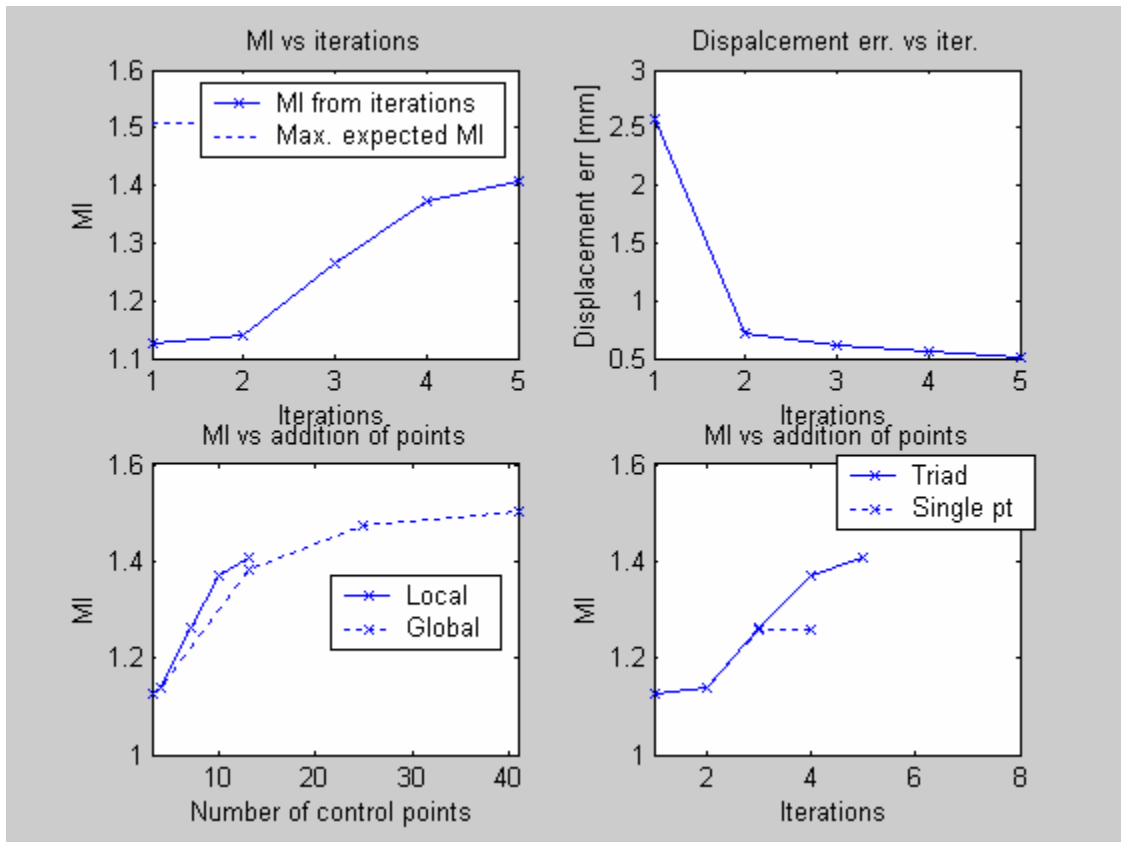


Fig. 4.9. Experiment 3 results

In this case adding a single point did not achieve as high a global MI as adding a triad for the same reason stated for experiment 2.

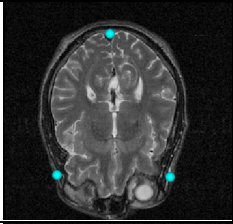
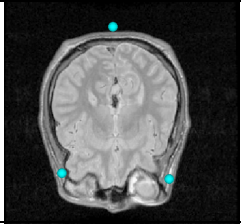

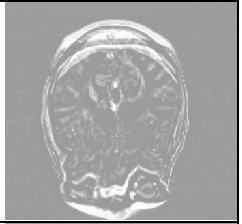
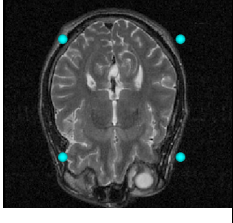
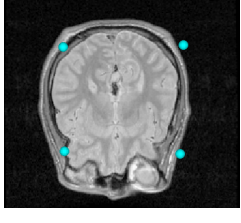
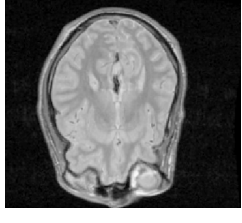
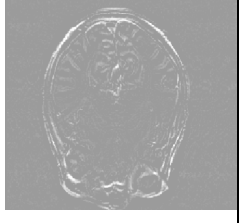
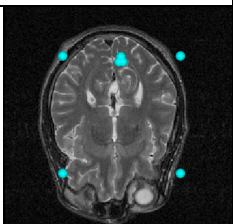
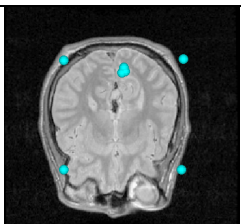
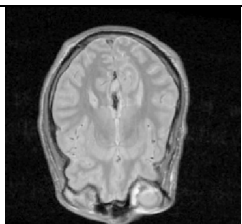
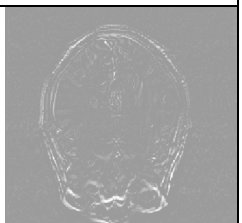
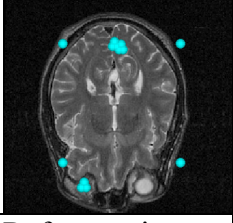
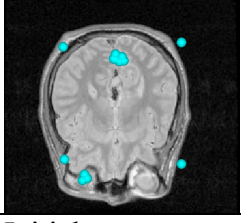
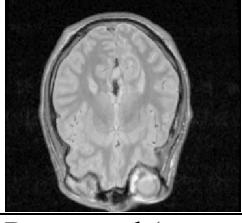
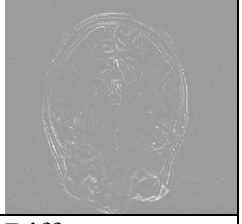
Iteration 1				
Iteration 2				
Iteration 3				
Iteration 5				
	Reference image with control points	Initial homologous image with control points placed incrementally as described in text	Recovered (unwarped) homologous image after optimizing global MI	Difference image between recovered homologous and original homologous image

Fig. 4.10. Intermediate stage results in Experiment 3

For iteration 5, triads seem to be overlapping but the markers for the controls points are big so that it is in fact not overlapping.

#### 4.4. Summary and Discussion

We have demonstrated in 2D an iterative algorithm that increases DOF, i.e., adds control points, to recover manually induced deformations. There is no claim nor proof that our algorithm produces the optimal (i.e., minimal) set of control points that recovers the deformation, but we have shown that it does recover the deformation quite successfully with far fewer control points than the method where DOF are increased globally. The algorithm first identifies the region of greatest mismatch where there is sufficient information to support the decision; the center of that region is the locus for an additional triad of control points in the reference image. Then the appropriate locus for the additional homologous triad is identified based on a local search for largest local MI under the existing mapping. Instead of placing an additional single control point pair at the identified locus, an equilateral triad of control points is added. The use of the triad, instead of a single control point is based on two factors: 1) at this level of scale in 2D a triad supplies most of the DOF necessary to support any local deformation required, and 2) by moving the control points of the triad differentially with respect to the center of the triad, the algorithm can avoid long range warping effects. This latter property is especially important since it is vital not to destroy the long range solution while trying to ameliorate local mismatches.

While the focus of this chapter is the automatic increase in DOF (i.e., the addition of control point pairs) to compensate for residual local mis-registration, an important component not addressed is the ability to examine the resulting solution at each step increase in DOF to determine if the DOF may be reduced without significant loss in accuracy. Future efforts in addressing the possible reduction in DOF while maintaining

nearly the same warping accuracy, i.e., the ability to fall back from the additional triad of points to only one or two additional points, will be addressed in future work using the notions of principal warps and kriging theory [19].

Another component that needs to be addressed is the deformation scale. As noted earlier two things are determined by the scale of the deformation; size of the triad and size of the subblock used to compute the mismatch measure. Obviously, if a triad is large, then it can be used to correct large deformations. Additionally, we have shown that the mismatch measure could only reliably detect deformations of size scale related to the region of support used to compute the mismatch measure. Thus, the triad size has to be “matched” to the subblock size used to compute the mismatch measure. In this chapter we fix the triad and subblock sizes thus confining ourselves to a fixed scale of deformation. Future work must involve repeating the process described here at several different levels of scale, varying from large to small. The smallest scale may be determined by the fundamental inability to estimate entropy/MI from a very small number of voxels. Entropy estimation based on minimal spanning trees [21] shows promise for estimating entropy/MI with sparse histograms; further efforts are needed to incorporate that estimation method.

Note that the recovered global MI is always 6 or 7 % less than the ideal global MI. There may be several reasons that the recovered global MI does not reach 100% of the global MI of the initial undistorted image pair. First, by introducing the initial known deformation some of the information is lost due to the low pass filtering effect of bilinear interpolation in creating the initial deformed homologous image. Second, it is important to observe that deformation recovery is limited to the amount of information present in

the image. There is little possibility of recovering deformation in low entropy regions; without texture it is impossible to observe the deformation and correct it. However the deformation in a richly textured region (i.e., high regional entropy) is likely to be recovered. Lastly, though our algorithm has been tested in 2D, the extension to 3D is obvious by calculating the mismatch measure in 3D and adding tetrahedrons instead of triads.



## CHAPTER 5

### COMPARISON WITH OTHER ADAPTIVE REGISTRATION ALGORITHMS

#### 5.1. Overview of Adaptive Registration Algorithms

This chapter is based on Park and Meyer [88]. While adaptive registration methods may have different geometric interpolants and similarity measures, they all share a common approach, i.e., they first detect areas where registration can be improved, and then they refine the grid in those areas. While there may be many ways to detect where to refine the grid, most current techniques observe differences in a *global* similarity metric; we suggest that a logical alternative is the use of a *local* similarity metric. Rhode *et al* use the gradient of global MI to refine grids [61, 62]. Others use methods based on entropies [63, 64]. In this chapter, we compare two methods to refine the grid, one global and one local, and show that for the two common methods chosen, the local measure method is better suited for improving the registration. We also compare our local mismatch measure from chapter 3 with Rhode’s method. Note that for global measures, cost functions are calculated across the whole data set, whereas for local measures, cost functions are calculated only over a finite local region. For this chapter, MI and B-splines are used as the similarity measure and the geometric transform respectively.

## 5.2. Common Framework in Adaptive registration

Before attempting any high order registration, a simple, low order affine registration removes any large, linear global effects. This ensures that the subsequent high order transform deals with only relatively small local transforms. After the affine registration, a high DOF registration is performed in an iterative multi-scale fashion to save computation time. For lower resolution data, i.e., sub-sampled data, a sparse grid of B-splines is used. As resolution increases, the grid becomes denser. Also within a specific resolution, different sized grids from large to small scale are used to speed convergence of the optimization. Any time grid density is increased, the new denser grid is initialized using the previously optimized sparse grid. This multi-scale grid optimization continues for a fixed number of resolutions and a fixed number of scales within a resolution, as long as the global MI from the previous optimization has increased more than a user defined threshold. Typically, at every grid refinement in 3D, the existing grid is halved in all dimensions resulting in an 8-fold DOF increase.

Reuckert's B-spline non-rigid registration works well at lower DOF, but suffers major computational complexity at high DOF [26]. For a typical CT data set of  $512 \times 512 \times 60$  with voxel dimension  $1 \times 1 \times 5 \text{ mm}^3$ , B-splines placed every 5 mm results in 1,880,000 DOF, which leads to significant computational and convergence problems. Additionally, with B-splines placed regularly everywhere, there are many B-splines placed in background areas where there is little information to judge the registration and little interest in the local result. To remedy these computational and convergence issues, adaptive non-rigid registration algorithms have been developed [61-64]. They place or use B-splines, or control points only in areas where they are needed to improve the overall registration. Thus, the registration is computed using an irregularly spaced grid

having fewer DOF. Moreover, for a given DOF adaptive algorithms can allocate a dense grid of B-splines or control points to areas of interest without wasting them in backgrounds, and thus can achieve better registration accuracies. The following is the summary of the general adaptive non-rigid registration framework.

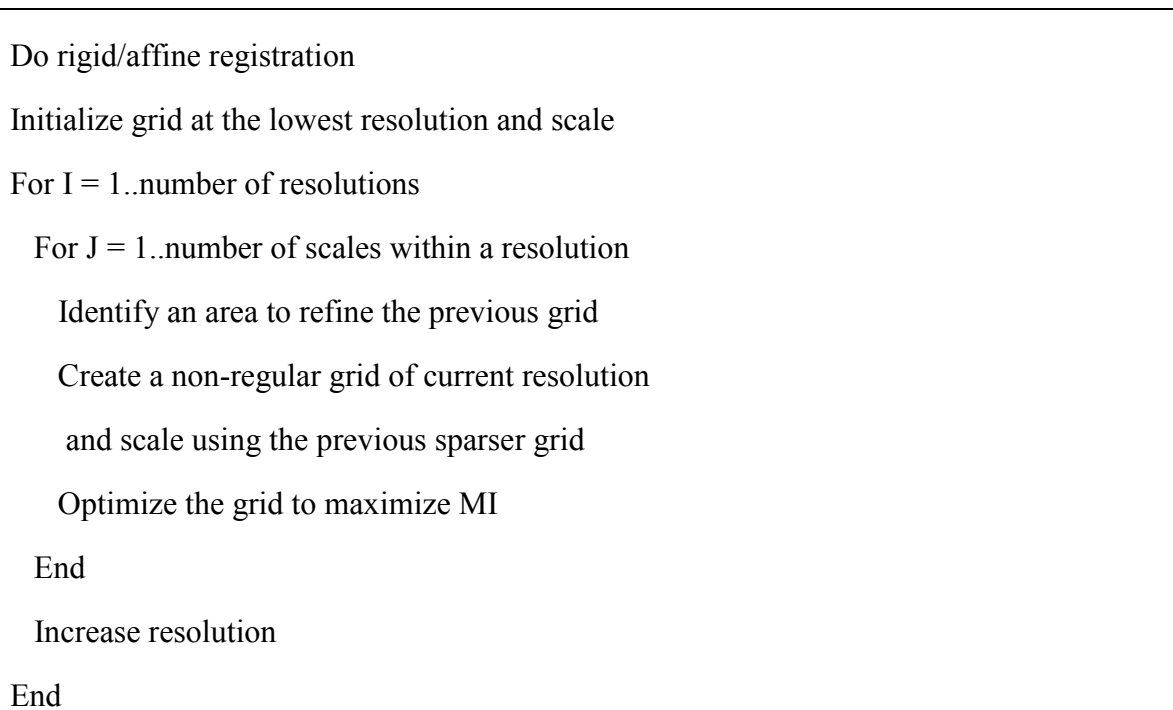


Fig. 5.1. General adaptive non-rigid registration algorithm

### 5.3. Other Adaptive Registration Algorithms

#### 5.3.1. Rhode paper

Rhode *et al* use the gradient of global MI to adaptively refine a grid [61, 62]. They use Wu's radial basis function for the geometric transform, which has a finite local support property as with B-splines, and MI as the similarity measure. The authors use the following algorithm to accomplish grid refinement. At a given scale of a grid, they put

basis functions at regular intervals and then move one basis function at a time and observe the change in global MI. They argue that if the gradient of global MI is large, then the global MI is not maximum with respect to that specific location of basis function. Thus, that location is where two data sets are locally mis-registered and registration can be improved further. Once the new location of basis function is determined, 8 smaller (i.e., octree) scale basis functions occupy the area where there previously was one larger basis function.

### 5.3.2. Other Papers

Rohlfing *et al* use B-splines and a modified similarity function where one term is the global normalized MI and the other term is the smoothness of the deformation [63]. They use local entropies to determine locally mis-registered areas and simultaneously use active B-splines and inactive (i.e., not allowed to move) B-splines to reduce DOF and effectively implement an irregular B-spline grid. Schanbel *et al* also use the same B-spline and two-term similarity measure, but refine their grid using local entropy, local standard deviation or gradient of the global cost function [64].

## 5.4. Experiments

Data sets used here are T1 and T2 weighted MR images from BRAINWEB [66] where the noise level is 3% in both images. The T1 slice is chosen as the reference data set and the T2 slice is the homologous data set. The homologous set is deformed locally in a known way using the following formula,

$$d(r) = r(5e^{-r/rate} + 1), \quad (5.1)$$

where  $r$  is the radius measured from the center of subblock to be deformed,  $d(r)$  is the newly assigned radius from the center of the subblock and rate is the scale at which deformation occurs. With this radial dilation deformation, the scale of deformation can be manipulated to be large or small. If the rate is large, then  $d(r)$  decays slowly and thus, generates a large scale deformation.

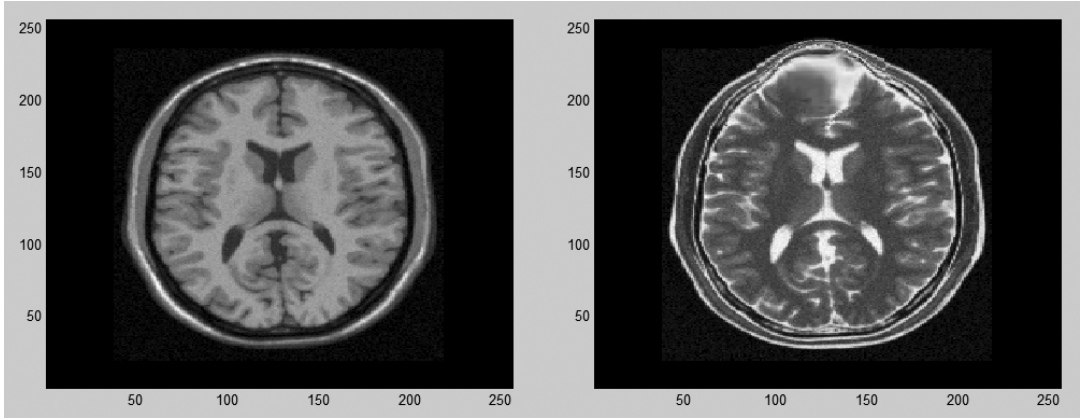


Fig. 5.2. Data sets used. Left: reference data set. Right: homologous data set with dilation centered at (125, 215).

#### 5.4.1. Sensitivity of Local vs. Global Measurements

We now compare the sensitivity of two registration cost functions, local MI and global MI, with respect to various scales of deformation. Both measures can be thought of as observations of a matched filter in performing a detection task. Borrowing from standard methods to compare outputs of matched filters [65], the ratio  $C$  is defined as the following as a metric to compare different objective functions,

$$C = \frac{\frac{\partial^2}{\partial \theta^2} s(\theta) \Big|_{\theta=\theta_0}}{\sqrt{N_0}} = \frac{\frac{\partial^2}{\partial \theta^2} E[MI(\theta)] \Big|_{\theta=\theta_0}}{Std(MI(\theta))}, \quad (5.2)$$

$$\theta_0 = \arg \max_{\theta} E[MI(\theta)]$$

where  $s(\Theta)$  is the noiseless output,  $N_0$  is the noise variance, and the peak of the objective function occurs at  $\theta = \theta_0$ . This ratio  $C$  favors objective functions with less noise and high curvature at the expected extremum point because such objective functions have a smaller error distribution of the loci of the extremum. We computed  $C$  by fitting the observations of global and local MI with 2<sup>nd</sup> degree polynomials and evaluating their second partials at the ideal maximum. The polynomial is assumed to be the noiseless objective function and noise is calculated from the mean of variances of MI values at different deformation scales. Here  $N_0$  includes all noise sources, and specifically the effects of local minima and maxima. The plots have been scaled and shifted to have the same noise and peak so that the signal to noise ratio differences between the two curves can be easily observed.

In this first experiment, we assume that the center of dilation in the homologous data set is known and the scale of dilation varies from small to large. Both global MI and local MI are recorded as we increase the scale of deformation. Note that local MI is confined to a subblock centered on the dilation center. Optimal bin width is applied to both measures and independent zero mean Gaussian noise is added to both reference and homologous data sets. From Fig. 5.3, it is evident that using local MI is more sensitive to local deformations than global MI. The ratio  $C$  for this case is 0.0020 (global MI) and 0.0038 (local MI). We observed similar trend (i.e., local MI observations have more sensitivity than global MI observations) over a range of image noise variances as in Table 5.1. This result suggests that local measures (i.e., local MI) have better sensitivity than global measures (i.e., global MI), and are better suited for detecting local mis-registrations.

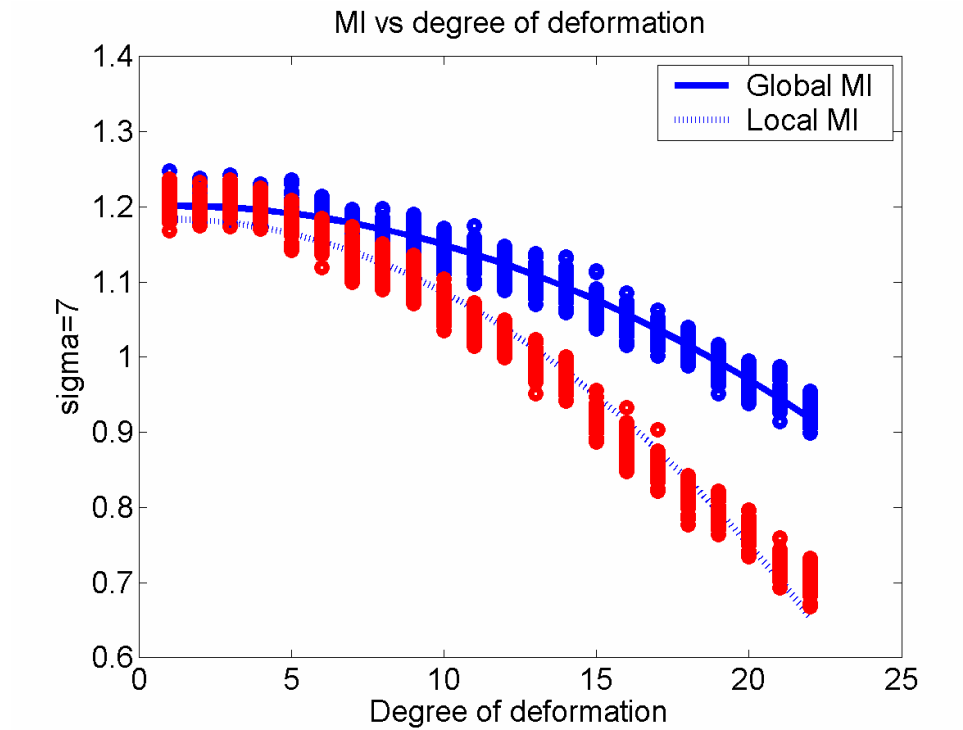


Fig. 5.3. Global/Local MI vs. degree of deformation. Solid line is the global MI fit and dotted line is the local MI fit. Circles are the MI values of different realization. Added noise variance of the image is 49. Note that there are 50 realizations for each scale of deformation.

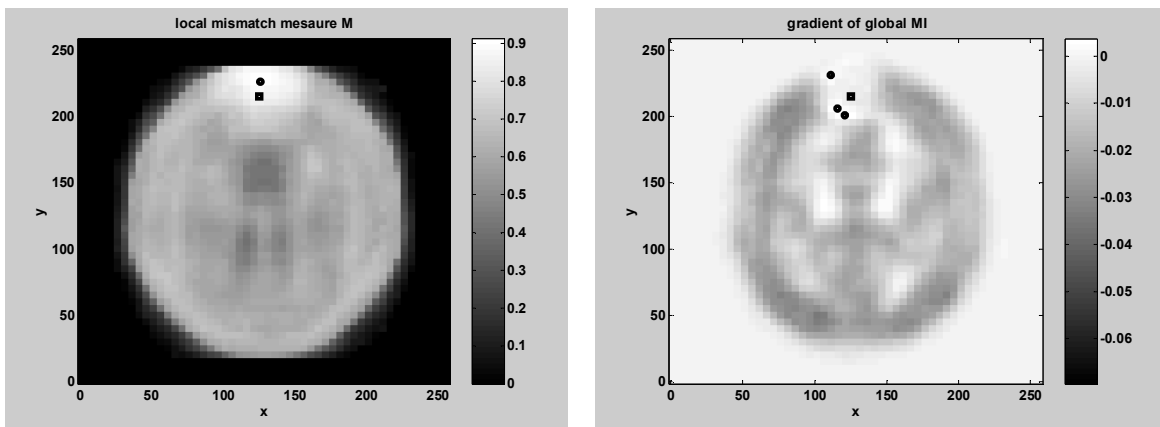
Noise variance	Global MI			Local MI		
	Ratio C	Curvature	Noise	Ratio C	Curvature	Noise
9	0.0035	0.0014	0.3995	0.0043	0.0017	0.3995
25	0.0026	0.0008	0.3169	0.0045	0.0014	0.3169
49	0.0021	0.0009	0.3240	0.0037	0.0012	0.3240
81	0.0016	0.0005	0.3310	0.0031	0.0010	0.3310

Table 5.1. Comparison of global MI and local MI

#### 5.4.2. Local Mismatch Measure M vs. Gradient of Global MI

In the second experiment, we use the more realistic scenario where the dilation center is unknown and compare the performance of our local mismatch measure and

gradient of global MI. B-splines are used instead of Wu's radial basis function for evaluating the gradient of MI as described in Section 5.3.1. We apply a known dilation of a given scale and try to detect the center of dilation by both our local mismatch measure,  $M$ , and gradient of global MI by Rhode *et al.* Both mismatch measures are evaluated every 5 pixels. Subblock size and B-spline spacing are dependent on deformation scale for local mismatch measure,  $M$ , and gradient of global MI respectively. Both scale dependent parameters are chosen so that both methods have maximum ability to detect the given deformation. We have tried different scales of deformation ( $rate = 2, 3, 5$  and  $8$ ) have found that for large scale deformations (e.g.  $rate = 5, 8$ ) both mismatch measures perform well, while for small scale deformations (e.g.  $rate = 2, 3$ ) our local mismatch measure,  $M$ , performs better. The results in Section 5.4.1 suggest that global measures may have less sensitivity than local measures. Although for large scale deformations the gradient of global MI has enough sensitivity to detect the dilation, as the deformation scale decreases, global measures may not have the sensitivity needed to detect small dilations.





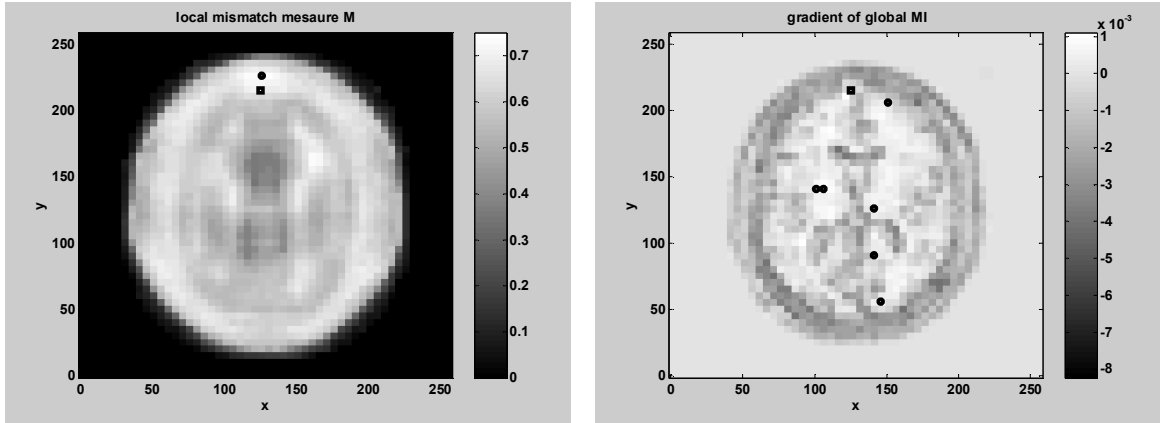


Fig. 5.4. Local mismatch measure  $M$  vs. gradient of global MI for large and small deformations (rate = 5, 2). Black box is the true dilation center and black circle is the dilation center found by mismatch measures over 5 trials. First row is the large deformation case and the second row is the small deformation case. For the large deformation, both measures perform well while for the small deformation only local mismatch measure  $M$  is successful.

## 5.5. Summary and Discussion

We show a general framework in adaptive non-rigid registration algorithms and note that there are many ways to refine a grid within the adaptive non-rigid registration framework. In Section 5.4.1, we show that local MI is a better detector of dilation than global MI. We suggest that this is probably due to the increased sensitivity of local measurements. Though we only show this for local vs. global MI, we suggest the same trend will be observed for most pairs of local vs. global similarity functions. In Section 5.4.2, we show that our local mismatch measure,  $M$ , works well over a large range of deformation scale, while the gradient of global MI only works for large scale deformations. While the mismatch measure we present,  $M$ , is only one of many possible mismatch measures, we suggest that grid refinement, or the detection of the locally mis-registered areas, should be based on local measures rather than global ones.

## **CHAPTER 6**

### **ABDOMEN ATLAS**

#### **6.1. Introduction**

This chapter is based on Park et al. [89]. There have been significant efforts to build a probabilistic atlas of the brain and to use it for many common applications, like segmentation and registration. Though the work related to brain atlases can be applied to non-brain organs, less attention has been paid to actually building an atlas for organs other than the brain. Motivated by the automatic identification of normal organs for applications in radiation therapy treatment planning, we present a method to construct a probabilistic atlas of an abdomen consisting of 4 organs (i.e., liver, kidneys and spinal cord). Using 32 non-contrast abdominal CT scans, 31 are mapped onto one individual scan using TPS (Thin Plate Spline) as the non-linear transform and MI (Mutual Information) as the similarity measure. Except for an initial coarse placement of 4 control points by the operators, the MI based registration is automatic. Additionally, the four organs in each of the 32 CT data sets were manually segmented. The manual segmentations are warped onto the “standard” patient space using the same transform computed from their gray scale CT data set and a probabilistic atlas was calculated.

## 6.2. Probabilistic Atlas on Abdomen

Recently, significant effort has been directed towards the development of deformable templates typically for segmentation of the human brain [31-41]. Only one recent publication addresses segmentation of the liver and kidneys using a deformable surface model [42]. Rich literature has been compiled on probabilistic atlases of the brain [43-53]; while most of the literature is extendable to abdominal organs, these extensions exist only in theory and have not been actually demonstrated on abdominal organs. Thus, we have pursued the simultaneous construction of a probabilistic atlas that includes a set of abdominal organs similar to that describe by Evans [57]. The term “probabilistic atlas” does not simply mean the average boundary of an organ, but rather the complete spatial distribution of probabilities that a voxel belongs to one or more organs, i.e., each voxel has an  $n$ -vector, where  $n$  depends on the number of organs in the modeling system. We continue to focus on probabilistic atlases in spite of the remarkable, even spectacular results of deformable templates, because of the belief that atlases, if well formed, bring more prior information to the process of defining complex organs in low signal-to-noise settings. After mapping a new patient onto the atlas using the same methods used to construct the atlas, the computed inverse transform maps the probability distributions of the atlas back onto the patient. Then the atlas can be used to find the most probable loci of edges of the patient’s organ. Furthermore, the derivative of the atlas defines a Bayesian weighting (i.e., probability density function), which provides very rich information for identifying the organ’s boundaries.

Here we present the results of a full probabilistic atlas construction for 32 patients, where 31 of the patients have been mapped onto one individual using thin plate splines (TPS) as the non-linear transform and mutual information (MI) as the similarity

measure. Registering 31 patients onto one arbitrarily chosen reference patient can introduce bias towards that specific reference patient, especially if the reference patient is far from the average population for which the atlas is intended. Thus, the reference patient selected best represents the population to our knowledge aided by an expert inspection. By iterating this construction process; i.e., using the resulting atlas as the reference target of another atlas construction phase, other groups have shown that the variance of the atlas and the dependence of the reference patient can be reduced further [44, 51]. While in the past, many groups focused only on the mean surface, the Bayesian value of the probabilistic atlas in the segmentation task lies not only in the mean but also in the variance of the atlas.

### **6.3. Methods**

#### **6.3.1. MIAMI Fuse**

The registration of the individual data sets onto the target reference is implemented using mutual information for automatic multimodality image fusion, i.e., MIAMI Fuse, which implements TPS as the non-linear transform and MI as the similarity measure [9, 11, 15]. Since the abdomen consists of organs that are compressible and have different sizes and shapes across individuals, the use of a non-linear transform is required instead of a rigid transform. TPS is expressed in terms of control points, which are placed in pairs in the corresponding loci of the related data sets. TPS deformation arises from surface interpolation literature where control points are constraints of the interpolating function. From the interpolating perspective each constraint (control point) primarily affects interpolating function in its vicinity. Detailed visualization of local effects of TPS in 2D can be found in chapter 2. Thus control points primarily affect deformations near

the control points. Note also that the density of control points is related to the DOF of the non-linear transform. Thus, when the region has high density of control points the DOF of the non-linear transform in that region is high. TPS is typically reserved for non-linear transforms, but TPS reduces to affine or rigid transform if the number of control points is less than 5 in 3D. Details of TPS and control points are found in these texts [22, 23]. MI is used as the similarity measure of choice. It basically exploits the fact that two co-registered data sets yield a joint probability density function (PDF) with tight clusters, whereas un-registered data sets yield a joint PDF with disperse clusters. Tighter clusters (i.e., more correlation) translate into higher MI values than more disperse ones (i.e., less correlation). In our implementation, all PDFs (both marginal and joint) are estimated by histograms with fixed bin width.

### **6.3.2. Control Points Used**

The liver, kidneys and spinal cord of the target reference data set were manually segmented by an expert using a window between  $-123$  HU (Hounsfield Unit) and  $131$  HU. Boundaries were traced using hand controlled optical mouse and 36 control points were placed in the target reference data set. The distribution of control points in the reference data set were chosen as follows: 17 control points were chosen in the reference data set's liver, 7 control points were chosen in and around the spinal cord, and 6 were chosen in each kidney as seen Fig. 6.1. The only criterion for the placement of these points was to achieve an approximately uniform density of control points for their distribution in each organ. The voxel dimensions of the target reference CT data set is  $(1.29 \text{ mm})^2 \times 5 \text{ mm}$ . Again, note that control points in each organ control deformations/transforms in their locality.

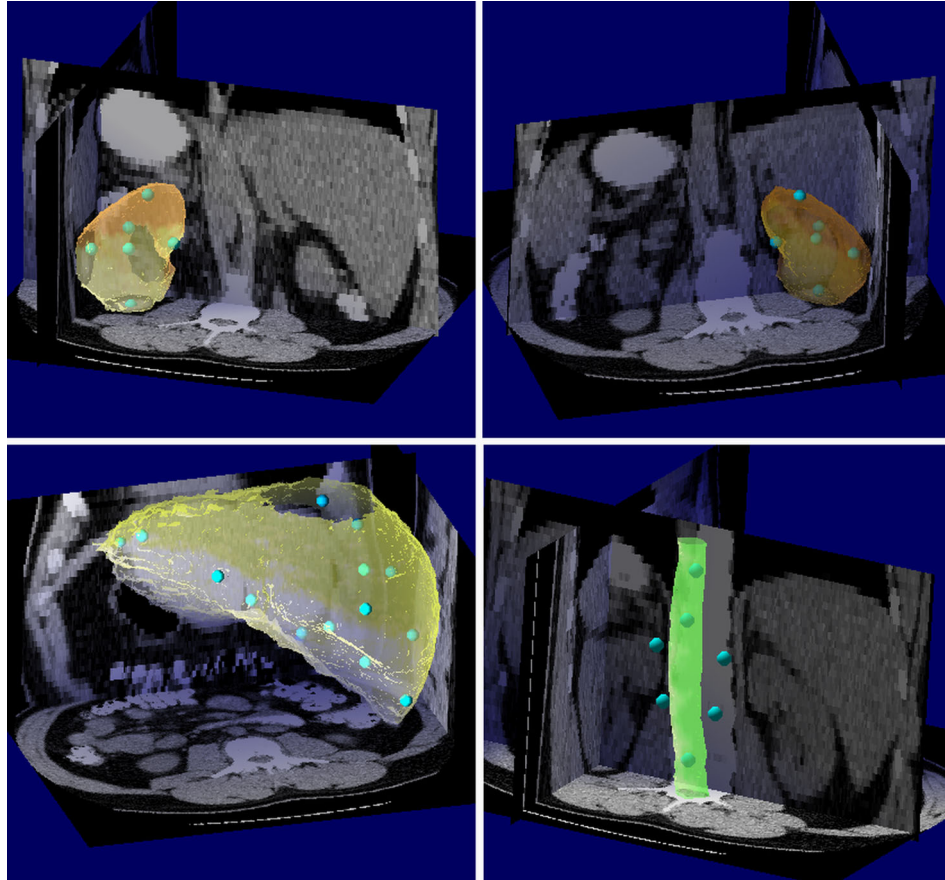


Fig. 6.1. Distribution of control points for organs of interests. From top left, left kidney model and associated 6 control points , top right, right kidney model and associated 6 control points, bottom left, liver model and associated 17 control points and bottom right, spinal cord model and associated 7 control points in the reference volume.

### 6.3.3. Construction of Atlas

We observed that if we registered the whole abdomen between reference and patient volume, liver primarily drives the registration since it has the largest voxel count of any organ, and therefore the liver affects the joint histogram and the resulting MI more than other organs. Thus, smaller organs like kidneys are not accurately aligned if jointly registered with other organs. By registering each organ separately, we obtained better overall registration accuracy. Each organ was registered onto its homologous pair in the reference volume separately, and then one final optimization using all 36 control points

was used to correct for any control point interactions that may have occurred between adjacent control points from the separate organ registrations.

Separate registrations for each organ are performed by masking out, i.e., zeroing, the remainder of the gray scale reference data so that it only contains the organ of interest. Since the usual implementation of our registration algorithm ignores zeros if they occur in either data set, the only data used to construct the joint histogram derives from the chosen reference organ and the mapping of the patient's data onto that organ only.

The registration of a single patient onto the target reference geometry begins by the operator's selection of just 4 homologous control points in the patient's data set corresponding the loci of the first 4 control points in the target reference, i.e., two in the liver, and one at the center of each kidney. This process essentially identifies to the algorithm the pose of the patient in the scanner, i.e., head/feet first and lying prone/supine. The loci of these 4 points need not be exact since the optimizing process that follows will find the correct corresponding loci under the MI similarity measure. Then the algorithm repeats the process of registering each organ separately. The optimizer moves the first 4 control points in the patient's data space to obtain the best isotropically scaled, rotate-translate mapping of the patient's organ onto the target reference organ. Then based on the optimized geometry mapping just obtained, the remaining points (out of the first 4 plus the  $N$  for that organ) are mapped onto the patient's data set to initialize the  $N+4$  control point TPS transform of the patient's organ onto the target reference organ and the  $N+4$  point TPS optimization follows. After the fit of one organ has been optimized, the process of focusing the attention of the registration

on each remaining organ separately continues until all of the specified organs have been fit. Basically for each organ, first affine geometry is optimized and then based on the obtained affine geometry N+4 point TPS is initialized, which removes the burden of specifying N+4 control points in the patient's volume. The resulting loci of all the organ-specific control points are then simultaneously optimized using a 36 control point TPS to correct for the small interplay that may occur between relatively close control points in the global solution, which is initially ignored in the individual focused organ refinement process.

Except for the placement of the 4 initial control points in each patient, the process of mapping the patient onto the target reference geometry is completed by the algorithm totally automatically. Indeed we may be able to achieve the first full affine registration fully automatically by placing the first 4 control points based only on the extents of the dataset and then implementing a limited DOF registration somewhere between full affine (12 DOF) and that of the rigid registration (6 DOF), e.g. isotropically scaled (7 DOF) or even anisotropically scaled (9 DOF). We have not as yet tried such fully automatic approaches but they clearly are important topics for further research.

## **6.4. Results**

### **6.4.1. Qualitative Analysis**

After registering each patient onto the reference, the probabilistic atlas is computed by applying the same registration transform to each patient's manually segmented data set. Then for each voxel in the reference volume, a vector float value is computed where the value of each vector component for each voxel represented the



fractional percentage of registered patient data sets that have a label at that reference voxel location corresponding the particular organ represented by that vector component. Volume rendering of the resulting probabilistic atlas is shown in Fig. 6.2.a. For the sake of comparison the full affine atlas construction is shown in Fig. 6.2.b. Volume rendering for both is via the ray tracing algorithm where the final color and brightness of the rendered pixel depend on the integrated local attenuation of the ray as it passes through the data volume. There are no gradient lighting effects used here. In Fig. 6.2.a the appearance of darker "shadows" around each of the individual organ models is due to the fact that at the edge of each model its probability, represented here by a particular intensity and hue corresponding to the organ, drops sharply to zero. The sharper the gradient, i.e., the better the registration process in constructing the probabilistic atlas, the more limited and thus visible are the dark fringes. The difference between two construction methods is most evident between the right kidney and the liver. Note in Fig. 6.2.a the sharp dark demarcation between the right kidney and liver where the probability that voxels in that region are either kidney or liver are low. In contrast, note in Fig. 6.2.b how the kidney and liver blend into one another in the fuzzy, affine atlas construction.

In general, the variance of the atlas will decrease as we increase the DOF of the transform, but it may not necessarily increase the effectiveness of the atlas to other applications (e.g., segmentation), which is commonly known as the "curse of dimensionality". We believe that we are operating at DOF where moderate increase in DOF will increase the effectiveness of the atlas, but it still requires validation. The validation part is left as future works. The variability of our atlas is largely affected by the presence/absence of features across patients, e.g. some patients do not have a left lobe

of the liver, and shape differences that exceed the ability of our non-linear transform to compensate.

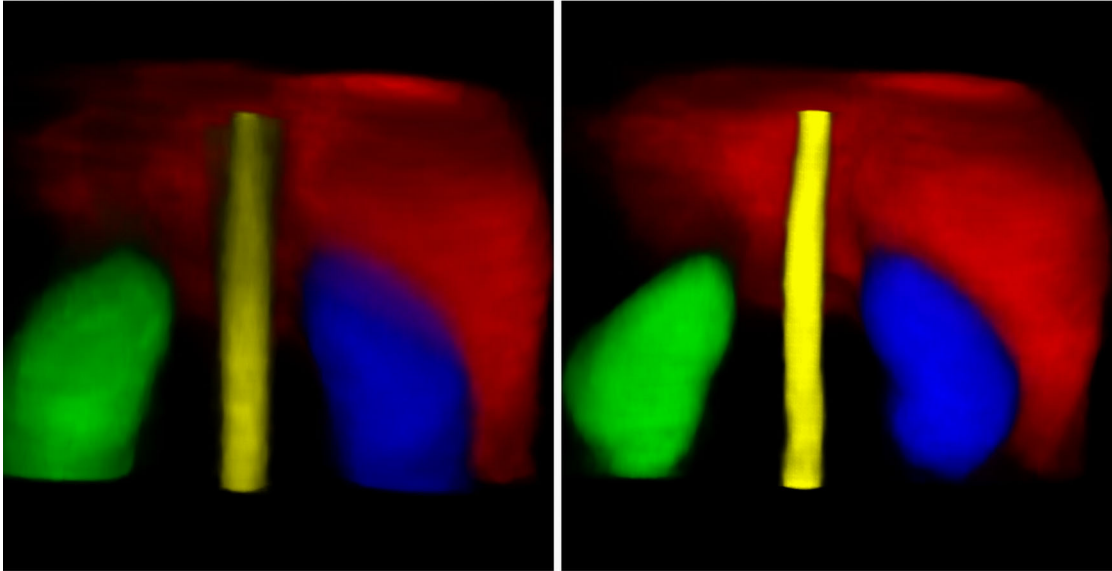


Fig. 6.2. Comparison of probabilistic atlases. The left figure (Fig. 6.2.a) is the probabilistic atlas constructed by a 36 control point TPS transform while the right figure (Fig 6.2.b) is the probabilistic atlas constructed by 4 control point affine transform. Red corresponds to liver and green, blue and yellow correspond to right kidney, left kidney and spinal cord respectively.

#### 6.4.2. Quantitative Analysis

To quantify results we have computed average standard deviation over different slices of the resulting atlas. Occurrence of the organ edges are assumed at 50 % threshold of the vector value and Gaussian curve with different variances are convolved with a step function occurring at the organ edges to best match in the least mean squared error (LMSE) sense the actual probabilistic atlas. The results are summarized in the table below:

Organ of Volumetric Probabilistic Atlas	TPS Construction, Average $\sigma$ , mm	Affine Construction, Average $\sigma$ , mm
Midhepatic slice, entire	6.8	11.2
Fossa of R. kidney	4.7	13.7
L. lobe	12.1	14.9
R. Kidney, midslice	6.1	13.0

Table 6.1. Variances of the probabilistic map in some slices

In general the variances of TPS based atlas construction are on the order of 4 to 8 times smaller than the affine atlas construction. The structure that is a notable exception to this generalization is the left lobe of the liver. Here the similarity in variance between the two methods may reflect more on the anatomical variation, i.e., presence or absence of this structure across patients.

## **CHAPTER 7**

### **ATLAS BASED SEGMENTATION ON ABDOMEN**

#### **7.1. Segmentation Overview**

This chapter is based on Park et al. [89]. Segmentation can be thought as the process of assigning labels to individual voxels in the data set. Extensive review of the segmentation methods and deformable surface models can be found in these papers [67, 84]. Most segmentation methods can be divided into two groups. The first one is supervised segmentation where certain Volume Of Interest (VOI) has to be manually specified for training different tissue types. The other is unsupervised segmentation where the discovery of different tissue types is automated. Supervised segmentation typically uses k-nearest neighbors (k-NN) algorithm followed by some regularization [68, 69]. Unsupervised segmentation typically uses Expectation Maximization (EM) algorithm to estimate parameters for different tissue types and feeds those parameters to Bayes rule followed by some regularization [70-72]. Both segmentation algorithms include regularization in their steps. A common way to regularize or smooth is to incorporate a Markov Random Field (MRF) priors. Basically, the Markov Random Field model encourages adjacent voxels having similar labels. Extensive reviews and implementation on MRF regularization are available [73-77]. In the execution of Bayes rule of the unsupervised segmentation, some algorithms use Maximum a posteriori

(MAP) approach while others use “maximization of the posterior marginals” (MPM) approach [78]. Our work takes the route of the unsupervised segmentation with MAP.

Our contribution is to extend the unsupervised segmentation with the additional information provided by the probabilistic atlas. Improvements in segmentation quality compared to cases with no atlas information are readily noticeable, especially where there is little contrast between organs of interest. Specifically, we can eliminate most of the soft tissue false positives around liver/kidneys/spinal cord and distinguish liver, spinal cord, left kidney and right kidney. Note that to apply the atlas information, a new CT data set should be mapped onto the atlas space by the same method that is originally used to construct the atlas.

We have applied our segmentation algorithm retrospectively to 20 abdominal CT scans of anonymous patients who have established lymphoma and received these scans as part of a standard clinical protocol to follow possible lymph node enlargement. These are all 7-10 mm thick slices acquired during a breath-hold on a helical scanner. Only oral contrast material was administered; no intravenous (IV) contrast was used. Segmentation of such non-IV-contrast CT scans is very difficult because CT values for non-lipid soft tissues are primarily related to density, and all non-lipid soft tissues have the density of water. False positive rates and false negative rates are reported as a measure of segmentation performance. Also, comparisons with cases where no atlas information is used are made.

## 7.2. Methods

In this chapter, the true label field (i.e., segmented field) will be denoted  $\mathbf{X}$ , the given observed data set will be denoted  $\mathbf{Y}$  and the probabilistic atlas is denoted by  $\mathbf{A}$ . Elements of  $\mathbf{X}$ ,  $\mathbf{Y}$  and  $\mathbf{A}$  are arranged by spatial position denoted by  $i \in I$ , where  $I$  is the simple index in a rectangular grid. Throughout the chapter,  $\mathbf{x} = (x_1, x_2, \dots, x_N)$ ,  $\mathbf{y} = (y_1, y_2, \dots, y_N)$  and  $\mathbf{a} = (a_1, a_2, \dots, a_N)$ , where  $N$  is the total number of voxels, will represent sample realizations of  $\mathbf{X}$ ,  $\mathbf{Y}$  and  $\mathbf{A}$  respectively. Sample space of  $X$  will be denoted  $\Omega_x, \Omega_x = \{\mathbf{x} : x_i \in \{1, 2, 3, 4, 5\}, \forall i \in I\}$ . Label 1 is the liver, 2 is the right kidney, 3 is the left kidney, 4 is the spinal cord and 5 is the ‘‘None of above’’ label. Note that the probability atlas is 5-vector,  $a_i = (a_{i1}, a_{i2}, a_{i3}, a_{i4}, a_{i5})$ , where each component corresponds to probability of a specific organ being there. The atlas information is realized as  $P(X_i = k) = a_{ik}$ ,  $k = 1..5$ ,  $a_{i5} = 1 - \sum_{k=1}^4 a_{ik}$ ,  $i \in I$ . As noted before data set  $\mathbf{Y}$  is already mapped onto the spatial frame of the probabilistic atlas using the same techniques that are used to construct the atlas.

### 7.2.1. Bayesian Framework

The problem is to estimate the label  $\mathbf{X}$  that best explains the given observation  $\mathbf{Y}$  according to some cost function. Different cost functions can be applied, such as MAP which aims to maximize the global a posteriori probability  $P(\mathbf{X}|\mathbf{Y})$  and MPM which aims to maximize the posterior marginal distribution  $P(X_i|\mathbf{Y})$  for each voxel  $i$ . We chose MAP and the following is the formula for the realization of estimate of  $\mathbf{X}$ .

$$\hat{x} = \arg \max_x P(X = x | Y = y) \quad (7.1)$$

### 7.2.2. Tissue Type Modeling (Gaussian)

We will assume that random variable  $Y_1, Y_2, \dots, Y_N$  are conditionally independent given the true labels  $\mathbf{X}$ . We will also assume that probability density function of  $Y_i$  given  $\mathbf{X}$  depends only on  $X_i$ . Furthermore, the probabilities of  $Y_i$ 's are modeled as conditional Gaussians given mean and variance of the true label  $X_i$ . There are other advanced models for different tissue types, especially for brain tissues [81], but still many papers use this simple Gaussian model and have been successful [71, 72]. Formulation for the conditional probability follows.

$$p(\mathbf{Y} | \mathbf{X}) = \prod_{i=1}^N p(Y_i | X_i = k) \quad (7.2)$$
$$p(y_i | X_i = k) = \left\{ \begin{array}{l} \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(\frac{-(y_i - \mu_k)^2}{2\sigma_k^2}\right); \text{ for } k = 1, 2, 3, 4 \\ \text{histogram}; \text{ for } k = 5, \text{ i.e. 'None of above'} \end{array} \right\}$$

For the tissue type ‘None of above’, there may be multiple clusters in the gray scale distribution since it includes background, all intestine, spleen, stomach, body wall and bones. Thus, using a uni-modal Gaussian will not be sufficient. Note that our typical abdomen CT scans of slice thickness 5 mm have  $512 \times 512 \times 40 \approx 10^8$  voxels. If we can use only 0.1% of all voxels (i.e.,  $10^5$  counts) to train the tissue type ‘None of above’ using a histogram, we have enough samples to make the 256 bin histogram closely approximate the true distribution. Using a very conservative estimate, which assumes ‘None of above’ tissue type occupies 5% of the abdomen volume (i.e.,  $5 \times 10^6$  counts), we are guaranteed to have a reasonable estimate of the ‘None of above’ tissue type.

### 7.2.3. Automatic Training

Mean  $\mu_k$  and variance  $\sigma_k^2$  of the Gaussian model in eq. 7.2. have to be estimated for the first four tissue types. A simple sample mean and variance over VOIs are used. The selection of VOI is automated by the aid of the probabilistic atlas. With the atlas information, we have a very strong prior knowledge where the organs are. Thus the VOI for the training of a specific organ can be identified where the atlas predicts nearly 100 % organ presence. We have used above 96 percent of being the desired organ for specifying VOIs. Lowering the threshold for the VOIs may induce better robustness of parameters (i.e., mean and variance), but it may also bias the estimates, e.g. lowered threshold VOIs may included multiple organs. Actual segmentation results with different threshold for VOIs will be discussed later. For the tissue type ‘None of above’, a histogram is calculated over the automatically selected VOI.

The standard method used in unsupervised segmentation is the combination of EM/MAP. In those cases iteration between EM and MAP to ensure convergence of the hyper parameters (i.e., mean and variance) is required [71, 72]. Our method is estimating hyper parameters over the areas where the likelihood of a specific organ presence is very high, so the estimates of hyper parameters are very stable. In result, our method need not iterate to achieve stable hyper parameter estimates and segmentation. We have confirmed with our data sets that our segmentation results did not improve with iteration.

#### **7.2.4. Markov Random Field (MRF) regularization**

We have included the usual penalty for dissimilar adjacent labels in the form of Markov Random Field regularization where the underlying label is modeled as the Gibbs distribution. The MRF penalty favors adjacent label fields to have similar labels, which is the same general spatial correlation that exists in medical data sets. Vast amount of



literature exists on MRF models [73, 74]. Here we use a very simple implementation of the MRF model. A 6-voxel neighborhood system that is called clique is used. Our Clique consists of 4 nearest neighbor voxels in the slice plane direction and 2 nearest neighbor voxels in the out of plane direction. In the case of corner voxels and outer edge voxels, Cliques are accordingly reduced to a smaller neighborhood. There can be other complicated Cliques that span many more voxels than 6, which may be tailored for other applications. According to Hammersley-Clifford theorem, the probability density function of the MRF model takes the form of the Gibbs distribution [79].

$$p(x) = Z^{-1} \exp(-U(x)) = Z^{-1} \exp(-\beta \sum_{\{r,s\} \in C} t(x_r, x_s)) \quad (7.3)$$

$$t(x_r, x_s) = \begin{cases} 0; x_r = x_s \\ 1; x_r \neq x_s \end{cases}$$

$Z$  is the normalizing factor,  $\beta$  is the positive constant which controls regularization,  $C$  is the Clique and  $U(x)$  is the energy function that increases with inhomogeneous adjacent labels. We assume that the value  $\beta$  is known a priori, which is assumed in many MRF based segmentations [82, 83]. It is also found that the optimal value of  $\beta$  is not highly dependent of the image content and the performance of the segmentation algorithm remains relatively unchanged with respect to  $\beta$  [71].

### 7.2.5. Posterior Probability Formula

Optimizing the posterior probability is not an easy task, especially because there are so many possible realizations of the MRF model and the optimization is prone to be caught in local maximas. There are many methods for optimizing the posterior probability in the MRF setup. Typical methods are Simulated Annealing [75] and genetic algorithms [76] that are theoretically proven to converge to the global maximum, but all

these methods are extremely computationally expensive. In this chapter, we have used a method similar to Iterated Conditional Mode (ICM) by Besag [77], which is a greedy approach to those slower methods. Additional information available from the atlas is reflected on the Markovian assumption of the label field.

$$p(X_i = k | x_{\partial i}) = P(X_i = k) \frac{e^{\beta s_i(k)}}{Z}; \text{ Markovian assumption} \quad (7.4)$$

$x_{\partial i}$ ; clique of voxel  $i$

$s_i(k)$ ; number of voxels with the same label as voxel  $i$  in the clique

$Z$ ; normalizing factor

Thus, we are favoring similar adjacent labels in addition to the most probable label from the atlas. In ICM, maximizing the posterior probability is maximizing the product of two probabilities;  $p(y_i | x_i)$  and  $p(x_i | x_{\partial i})$  [77]. The  $p(x_i | x_{\partial i})$  term is further divided into atlas term and the Markov regularization term. Basically, we have included one additional term into the posterior probability. The first term,  $p(y_i | x_i)$  is the difference in the observation from the model mean in Gaussian sense, the second term is the contribution from the probabilistic atlas and the third term is the MRF regularization term. Typical methods have only two terms, the first and the third term.

$$\begin{aligned} & \arg \max p(y_i | X_i = k) p(X_i = k | x_{\partial i}) \\ & = \arg \max p(y_i | X_i = k) P(X_i = k) e^{\beta s_i(k)}, i \in I \end{aligned} \quad (7.5)$$

The ICM method needs an initial segmentation to begin with. The initial segmentation is provided by maximizing the posterior probability without the MRF term. In detail, it is optimizing the only two terms excluding the MRF term in the posterior probability formula provided above. In the implementation of ICM, regular ICM requires that all current labels (i.e. labels upto index  $i-1$  and labels from the previous iteration) are used to update the new label. Our implementation only uses labels from the previous iteration.

Our modified ICM iteration stops when label field changes between iterations are less than a given threshold.

### **7.3. Segmentation Results**

Our segmentation method has been tested on 20 non-IV-contrast abdominal CT scans. Patient CT voxels that map onto probabilistic atlas organ regions of 96% or greater are used for training, the threshold for the ICM iteration is set to 5% label change and  $\beta$ , the smoothness parameter, is set to 1.5. Performance of the segmentation is measured using false positive rate and false negative rates. We assume that the ground truth segmentation comes from an expert labeling, the null hypothesis is the absence of a specific organ and the alternative hypothesis is the presence of a specific organ. Thus, a false positive error occurs when a voxel is not an organ of interest by the manual segmentation, but is classified as an organ of interest by the segmentation algorithm. Below is one example of segmentation out of 20 cases.

#### **7.3.1. Qualitative Results**

Note in Fig. 7.1.g that there is very little contrast among different Gaussian modelings of organs. The four Gaussian curves for different tissue types have very large overlapping regions. Thus, without the atlas information the segmentation algorithm cannot distinguish well among organs of interest. Without the atlas information as in Fig. 7.1.c, right kidney, left kidney and spinal cord are all assigned the same label and there are lots of other soft tissues that are still classified as organs of interest. With the added atlas information as in Fig. 7.1.d, we now gain contrast over kidneys and spinal cord, and soft tissues that are not of interest (e.g. intestines, spleen etc.) are eliminated. Fig. 7.1.f

shows the difference between Fig. 7.1.c and Fig. 7.1.d. In Fig. 7.1.f, the background gray is the zero difference and all other values reflect differences between the two segmentation methods. From Fig. 7.1.f, mis-classification of spinal cord and left kidney is evident and differences in the two segmentation methods are found at body walls near the liver, organs of non-interest and centers of kidneys.

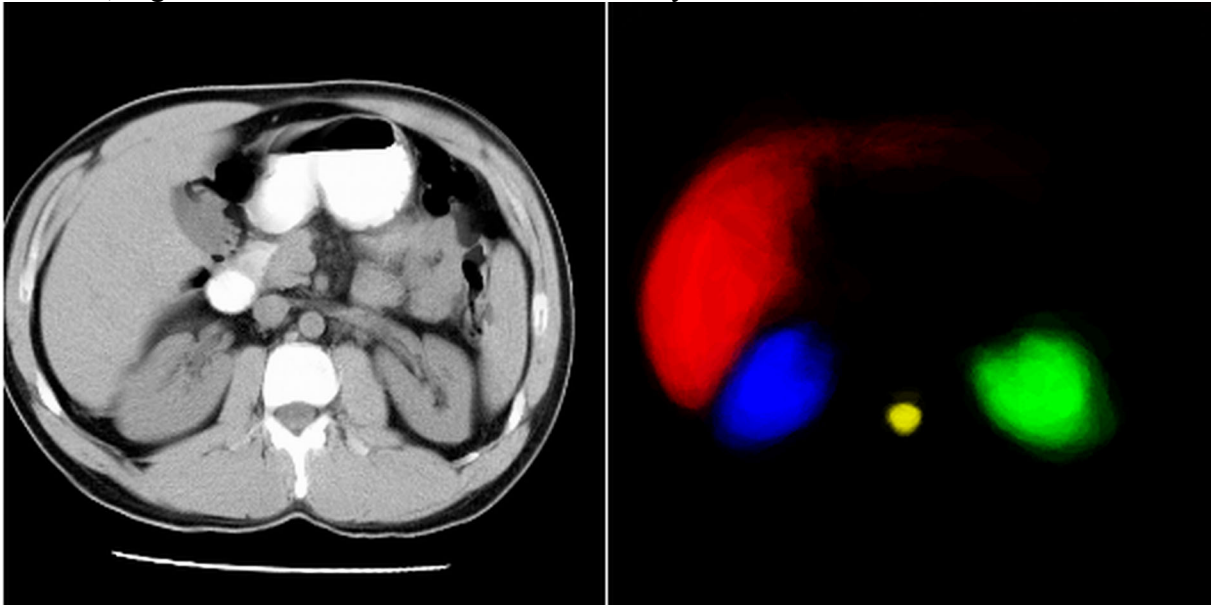


Fig. 7.1.a, 7.1.b. Left : one slice of the data set. Right : Atlas information.

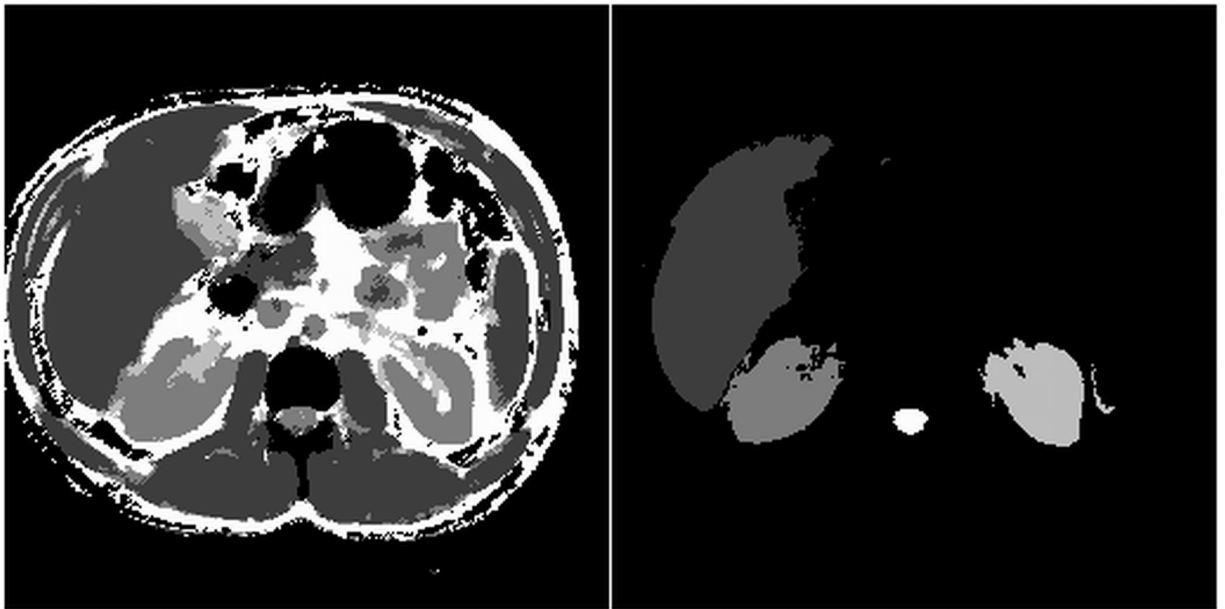


Fig. 7.1.c, 7.1.d. Left : segmentation results without atlas information. Right : segmentation results with atlas information.

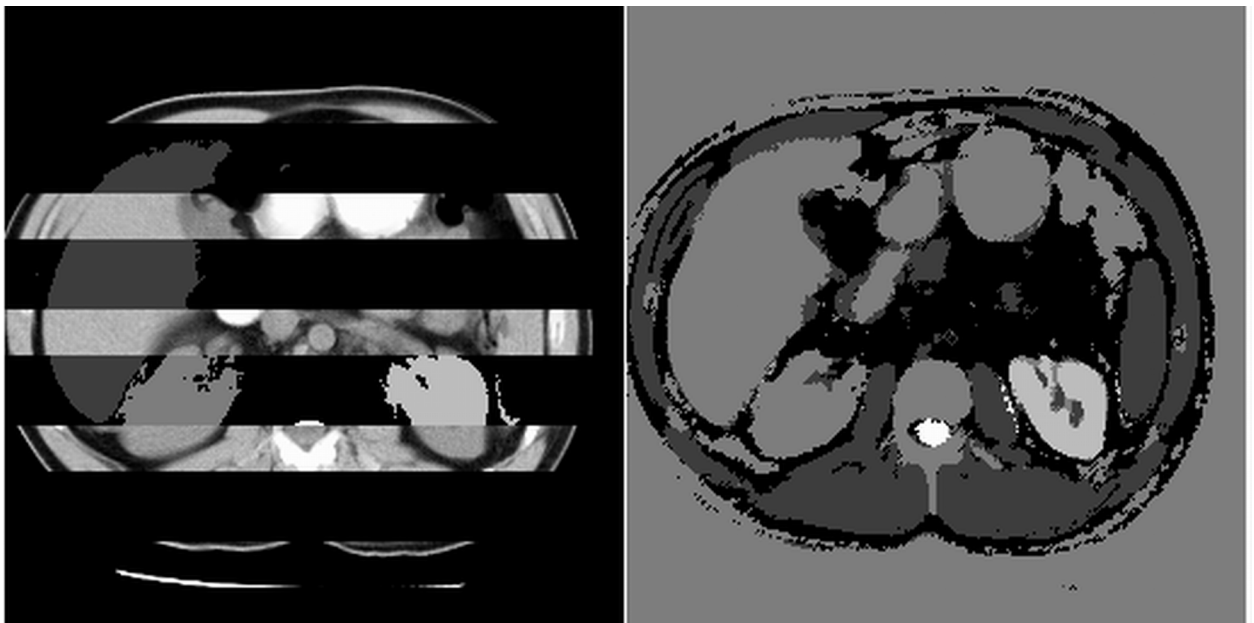


Fig. 7.1.e, 7.1.f. Left : comparison between grayscale data set and segmentation with atlas information. Right: difference between segmentation with and without atlas information.

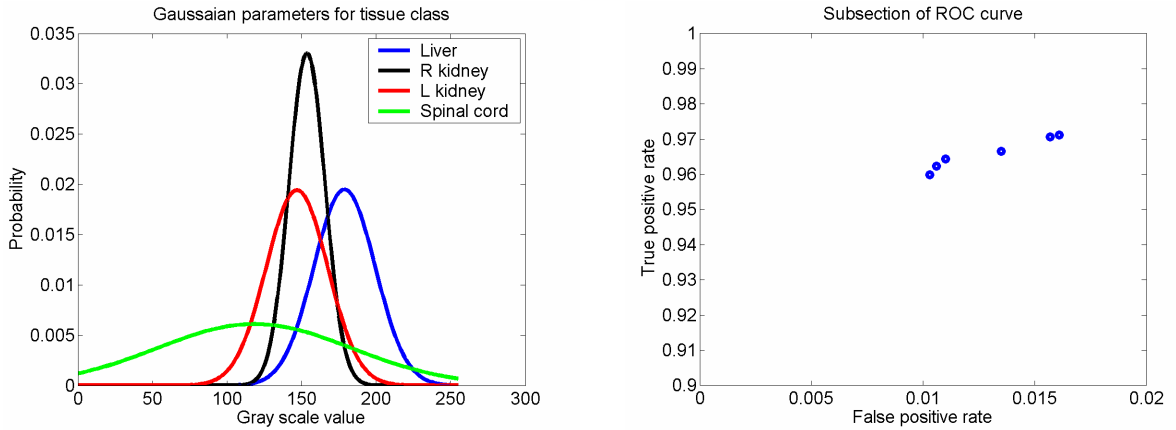


Fig. 7.1.g, 7.1.h. Left : Gaussian modeling for different tissue types. Right : subsection of the ROC curve.

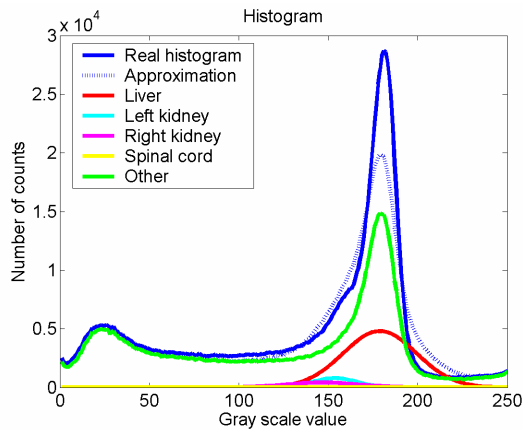


Fig. 7.1.i. Actual histogram and modeling of the histogram.  
Fig 7.1. Sample segmentation results.

### 7.3.2. Quantitative Results

The false negative rates for left kidney and spinal cord without the atlas are over 90%; it comes as no surprise that the segmentation algorithm without the atlas cannot distinguish left kidney and spinal cord from other organs. From the first two rows of Table 7.1, false negative rate values are generally larger than false positive values, which is the result of a conservative segmentation. By overlaying Fig. 7.1.a and Fig. 7.1.d or by

inspecting Fig. 7.1.e, you can observe that the segmentation results are not really extended towards kidney and liver boundaries. In other words, our results are conservatively biased, i.e., we are more likely to call a voxel not an organ when there is an organ present especially around the edges of the organs. Overall performance measures with the atlas information are very close to the performance measures of the liver, which is not surprising since liver is the biggest organ in our 4-organ abdomen model and thus has the most voxel counts. The last two rows of Table 7.1 show the effect of changing the training threshold from the 96% to 90%. The resultant false positive rates and false negative rates for the 90% case compared to 96% are slightly worse, but possibly within the realm of no statistical significance. If the threshold is further lowered for instance 60%, it is likely that we would observe biased estimates for the Gaussian parameters leading to worse segmentation results. For a moderate change in threshold (i.e., 96% to 90%) changes in segmentation results are minimal.

There remain fragments of organs around the liver and the kidneys in the segmented results (Fig. 7.1.d), these fragments can be further eliminated if we had an atlas with lower variance. Atlases with lower variances will fall off very rapidly near the organ boundaries, thus the algorithm will be unlikely to classify tissue fragments that are not only similar to the organ itself in grayscale values but also spatially close, as organs of interest. Still, the atlas constructed from 32 subjects is sufficient to remove most of the fragments. A subsection of the Receiver Operating Characteristic (ROC) curve is presented in Fig. 7.1.h. In this case we are not able to adjust the false positive rate at our will since the change of the only maneuverable parameter  $\beta$  yielded very limited range of false positive rates. Thus, in this implementation we are operating at the low false positive rate and high true positive rate region. Performance measures for all 20 data sets

are given in table 7.2. For table 7.2, ICM threshold is set to 0.5% and other parameters are the same as before.

	Liver	Right Kidney	Left Kidney	Spinal Cord	Overall
False positive rate (with atlas)	0.0064	0.0004	0.0004	0.0002	0.0054
False negative rate (with atlas)	0.0639	0.1447	0.1081	0.2116	0.0759
False positive rate (w/o atlas)	0.1452	0.0476	0.0130	0.1200	0.1271
False negative rate (w/o atlas)	0.0579	0.1278	0.9195	0.9843	0.1405
False positive rate (with atlas and lowered threshold for training)	0.0072	0.0004	0.0004	0.0002	0.0060
False negative rate (with atlas and lowered threshold for training)	0.0641	0.1365	0.1166	0.1982	0.0758

Table 7.1. Performance measures for the sample case

	Liver	Right Kidney	Left Kidney	Spinal Cord	Overall
False positive rate	0.0077	0.0005	0.0007	0.0002	0.0064
False negative rate	0.0782	0.1053	0.1066	0.1355	0.0839

Table 7.2. Performance measures for 20 data sets

#### 7.4. Summary and Discussions

Segmentation is one of the obvious applications where atlas information can help. In case of the non-contrast CT scans, atlas information aided the segmentation algorithm to support differentiation between similar gray valued organs where previously no other differentiations have been possible. Atlas information also led us to remove organs of lesser interest. All these will be helpful in clinical setups like radiation therapy treatment planning. Our algorithm showed underestimation of organs largely due to inability to assign correct labels near organ boundaries. This effect can be remedied by having an atlas with lower variance or a better modeling of ‘None of above’ label other than a



simple histogram. Other methods of finding specific organ edges in the patient's data set using the model information may further improve the results. In summary, we have demonstrated successful integration of the atlas information into the standard unsupervised segmentation and the results for the non-contrast CT cases are excellent. Further research should include assimilation of the atlas information into other well established segmentation algorithms including level sets [80].

## **CHAPTER 8**

### **SUMMARY AND DISCUSSIONS**

#### **8.1. Summary and Discussion for Adaptive Registration**

Recently we have seen non-linear registration applied to many common soft organs in the body and witnessed their successes. Through chapters 3 and 4, we have pointed out the shortcomings of traditional high DOF non-linear registration. Those shortcomings include computational complexity and convergence issues related to the high DOF nature of the particular geometric transform considered. To remedy those problems, we have proposed an adaptive registration algorithm, where the DOF are increased selectively only in local areas where they are needed rather than increasing the DOF globally. The central issue here is where to selectively increase DOF, namely where the locally mismatched areas occur. We have proposed to use our local mismatch measure  $M$  based on local MI and entropy. Others have proposed to use global MI or a particular form of global measures to find locally mismatched areas. In Chapter 5, we have shown that local measure based methods are better suited for detecting small scale deformations (i.e., small scale mismatches). We also have proposed an iterative adaptive registration algorithm that uses our local mismatch measure  $M$  and triad of control points. A triad of control points is used to accommodate local affine transformation in 2D. Our adaptive registration algorithm has been demonstrated to work in 2D MRI scans where the ground truth deformation is known.

In our implementation of the adaptive registration algorithm, the deformation scale is fixed. Note that the deformation scale is related to the size of the triad and the subblock size of the local mismatch measure  $M$ . For future work, adaptive registration should be repeated for several scales from large to small. Note also that if we decrease the scale, the subblock size of the local mismatch measure will eventually be so small that we cannot estimate the PDF reliably from a small number of samples. For the simulations in section 4.3, since the true deformation is known, the performance of the adaptive registration can be better visualized by overlaying the true deformation and the recovered deformation from the algorithm for figures 4.6, 4.8 and 4.10.

## **8.2. Summary and Discussion for Atlas Based Segmentation**

Atlas construction has been thought to bring good prior information to common image analysis techniques, like segmentation and registration. In Chapter 6, we have shown how to build an atlas of four abdominal organs using MI as the similarity measure and TPS as the geometric transform from 32 CT scans. Registration focuses on a separate organ at a time so that the registration process is more sensitive to that particular organ than registering the whole abdomen. After the atlas is built, we create a very strong spatial prior because of the reduced spatial variance. This is evident when compared with the affine constructed atlas.

We also have demonstrated how the atlas information can be incorporated in the standard unsupervised segmentation scheme. Atlas information is incorporated as an additional prior to the posterior probability. Note that a data set has to be registered onto the atlas space by the same algorithm used to construct the atlas to use any atlas

information. The training of different tissue types is automated by identifying highly probable regions of a particular tissue type from the atlas. The result of the proposed segmentation scheme on 20 non-contrast CT scans shows that there is a marked improvement on segmentation results. For example, compared to the standard unsupervised segmentation, we have eliminated organs of non-interest (e.g., spleen and intestines etc.) and been able to distinguish between two kidneys, where grayscale values are very similar.

In the construction of the atlas, we have registered all 32 patients on to an arbitrarily chosen reference patient, which may introduce bias towards that particular patient's geometry. Though the patient is chosen to approximately represent the intended population, for future work, we need to iterate the whole atlas building procedure on the previously built atlas until the atlas converges to the unbiased atlas. In addition, atlas priors can be incorporated to other well established segmentation algorithms, like level set theory. Also, performances of the segmentation algorithm with and without the MRF priors can be compared to assess the contribution of the MRF prior to the segmentation results.

### **8.3. Future work**

For adaptive registration, future work should include multiple deformation scales from large to small. Also, future work should include calculating the local mismatch measure  $M$  over a "tapered kernel" subblock to reduce the bias caused by a rectangular shaped subblock. This approach will be especially effective if the local deformation is known apriori so that the kernel can be "matched" to that deformation. For atlas based segmentation, future work should include segmentation with the aid of the atlas

constructed from various DOF to explore the relationship between DOF used to construct the atlas and effectiveness of the atlas. Also, assimilation of atlas information to other well established segmentation algorithms should be explored.

## APPENDIX

### Invertibility of TPS Deformation

Theoretically the TPS can generate warpings that are not invertible because the determinant of the local Jacobian is negative; in practice, this rarely happens for two reasons: a) the solution to the TPS warping minimizes bending energy which penalizes folding, and b) the DOF typically implemented by applications in our experiments are relatively low, i.e.,  $< 50$  for 2D and  $< 120$  for 3D. Although, currently there is no closed form for the analytic inverse, the inverse may be calculated using iterative numerical methods. Others have shown that the inverse of TPS cannot be represented by another TPS with the same number of control points [30]. In addition we will also demonstrate that the inverse of TPS needs to be represented by far more control points than the original deformation.

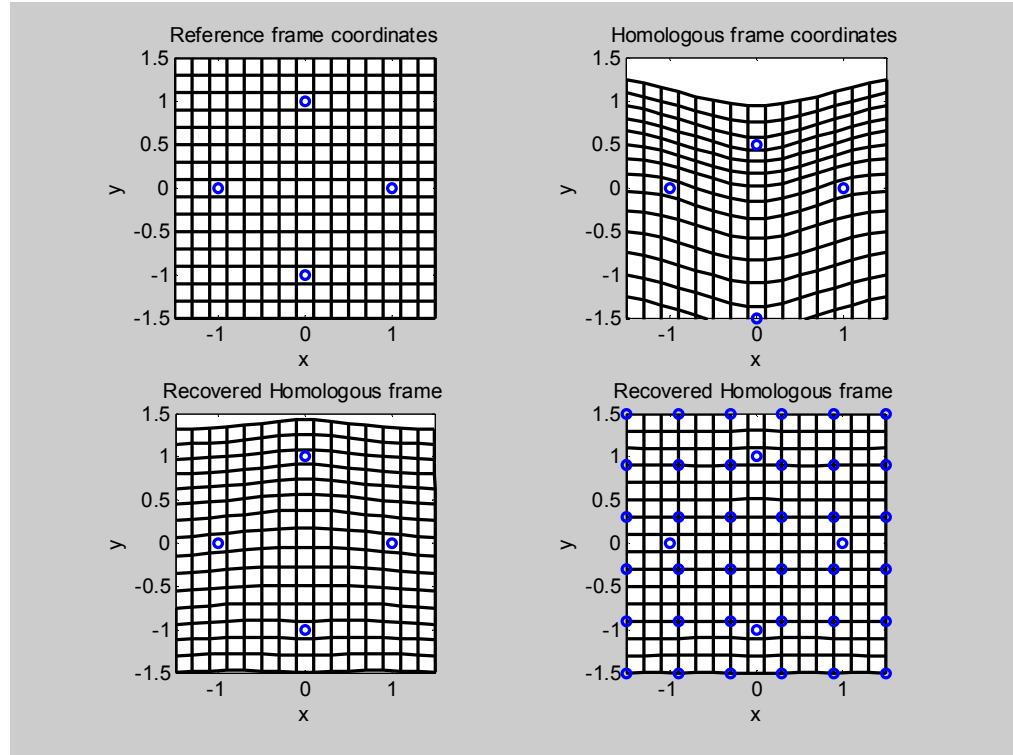


Fig. A.1. 4-point example for invertibility of TPS

Two points at  $(0,1)$  and  $(0,-1)$  are moved to  $(0, 1/2)$  and  $(0, -3/2)$  respectively. Two other points on the  $x$  axis are not moved.

The top-left plot is the reference frame and the top-right plot is the deformed homologous frame using the TPS that maps reference control points onto homologous control points (i.e., forward TPS). The bottom-left plot is the recovered homologous frame using the TPS that maps homologous control points onto reference control points (i.e., reverse TPS). Notice the non-square grid structures on the bottom-left plot. It clearly shows that the inverse of the TPS deformation cannot be represented by the same number of control points that caused it. The bottom-right plot is the inversion using 40 control points. Notice the far more regular shape of grids than the bottom-left plot.

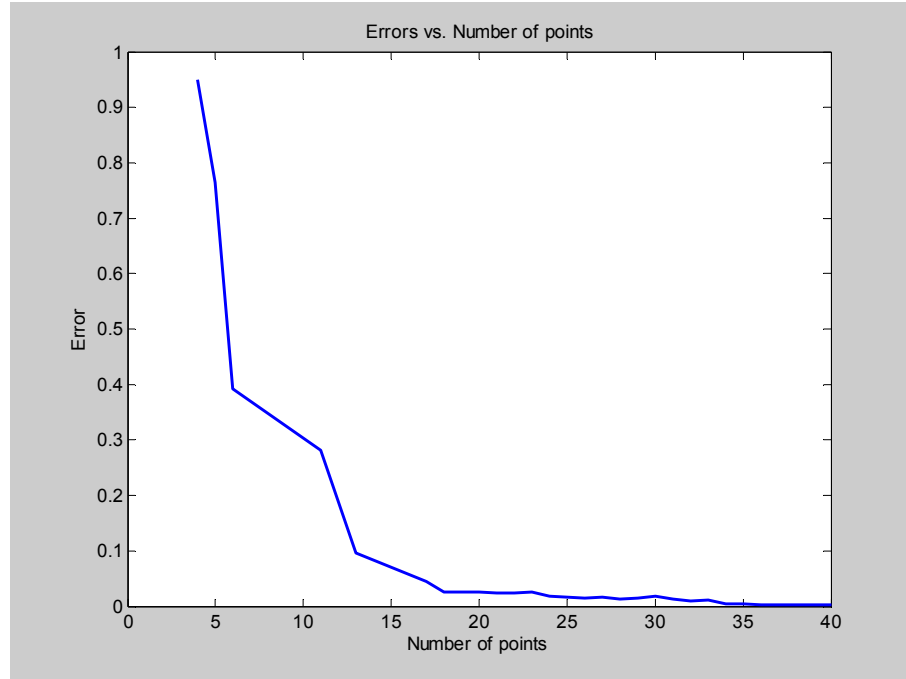


Fig. A.2. Errors with respect to number of control points used for inversion

For the plot above, error is measured as squared error from the ideal square grid for the same 4-point case from Fig. A.1. The plot shows that errors are decreasing with more control points used in the inversion. The way we added control points for inversion need not be optimal since we are only interested in observing the trend between the number of points used for inversion and the error. It is clear that many more control points than 4 are needed to recover a TPS induced deformation in a TPS fashion. It is very likely that we need infinite number of control points to exactly invert a TPS induced deformation.



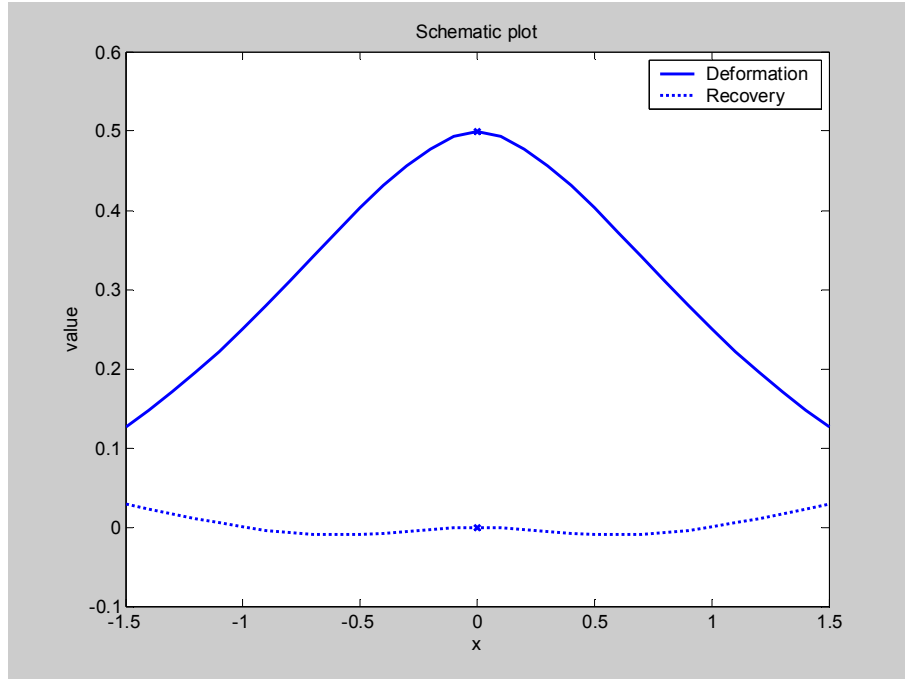


Fig. A.3. Schematic for invertibility.  
Solid line is the deformation and dotted line is the recovered deformation.

Above schematic also implies that inverse of TPS cannot be represented by finite number of control points. Above plots show a deformation of TPS in 1D, we have moved a control point (i.e. a knot) at  $x=0$  from base a value 0 to a value 0.5, resulting in the solid line. Even if we tack down the cross to the axis (i.e., from 0.5 to 0), the smooth basis function (i.e., dotted line) that connects the tack is still not a flat line.

## **BIBLIOGRAPHY**

## BIBLIOGRAPHY

- [1] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, et al., "Automated multimodality image registration using information theory," *Computational Imaging and Vision, Ile de Berder*, FR, 1995.
- [2] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal, "3D multimodality medical image registration using feature space clustering," *Computer Vision and Robotics in Medicine*, Nice, FR, 1995.
- [3] B. Kim, K. A. Frey, S. Mukhopadhyay, B. Ross, et al., "Co-registration of MRI and autoradiography of rat brain in 3D following automatic reconstruction of 2D data set," *Computer Vision and Robotics in Medicine (CVRMed)*, Nice, FR, 1995.
- [4] B. Kim, J. L. Boes, and C. R. Meyer, "Mutual information for automated multimodal image warping," *Second International Conference on Functional Mapping of the Human Brain*, Boston, MA, 1996.
- [5] B. Kim, J. L. Boes, K. A. Frey, and C. R. Meyer, "Mutual information for automated unwarping of rat brain autoradiographs," *NeuroImage*, vol. 5, pp. 31-40, 1997.
- [6] B. Kim, J. L. Boes, P. H. Bland, T. L. Chenevert, et al., "Motion Correction in fMRI via Registration of Individual Slices into an Anatomical Volume," *Magn. Reson. Med.*, vol. 41, pp. 964-972, 1999.
- [7] B. Kim, J. L. Boes, P. H. Bland, and C. R. Meyer, "Out-of-plane nonlinear warping of a slice into volume," *Medical Image Computing and Computer Assisted Intervention*, Cambridge, U.K., 1999.
- [8] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, et al., "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imaging*, vol. 16, pp. 187-98, 1997.
- [9] C. R. Meyer, J. L. Boes, B. Kim, P. Bland, et al., "Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin plate spline warped geometric deformations," *Medical Image Analysis*, vol. 3, pp. 195-206, 1997.

- [10] C. R. Meyer, J. L. Boes, B. Kim, P. H. Bland, et al., "Semiautomatic registration of volumetric ultrasound scans," *Ultrasound in Medicine and Biology*, vol. 25, pp. 339-347, 1999.
- [11] C. R. Meyer, J. L. Boes, B. Kim, and P. H. Bland, "Probabilistic Brain Atlas Construction: Thin-Plate Spline Warping via Maximization of Mutual Information," *Medical Image Computing and Computer Assisted Intervention*, Cambridge, U.K., 1999.
- [12] P. Viola and W. M. Wells, "Alignment by maximization of mutual information," MIT, 1995.
- [13] W. Wells, P. Viola, H. Atsumi, S. Hakajima, et al., "Multimodal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, pp. 35-51, 1996.
- [14] W. Wells, P. Viola, and R. Kikinis, "Multimodal volume registration by maximization of mutual information," *Second Annual Symposium on Medical Robotics and Computer Assisted Surgery*, Philadelphia, 1995.
- [15] C. R. Meyer, J. L. Boes, B. Kim, P. H. Bland, et al., "Warping normal patients onto the ICBM atlas by maximizing MI," *Proceedings of 4th Int'l Conference of Functional Mapping of the Human Brain*, vol. 7, T. Paus, A. Gjedde, and A. Evans, Eds. Montreal, Quebec, CA: Academic Press, 1998.
- [16] C. R. Meyer, J. L. Boes, B. Kim, and P. H. Bland, "Evaluation of control point selection in automatic mutual information driven, 3D warping," *Medical Image Computing and Computer-Assisted Interventions*, Boston, MA, 1998.
- [17] M. Holden, D. L. G. Hill, E. R. E. Denton, J. M. Jarosz, et al., "Voxel similarity measures for 3D serial MR brain image registration," *IEEE Trans Med Imaging*, vol. 19, pp. 94-102, 2000.
- [18] F. L. Bookstein, "Quadratic variation of deformations," *Information Processing in Medical Imaging*, vol. 1230, Lecture Notes in Computer Science, 1997, pp. 15-28.
- [19] J. T. Kent and K. V. Mardia, "Link between kriging and thin-plate splines," *Probability, Statistics and Optimization*, F. P. Kelley, Ed.: Wiley, 1994, pp. 325-339.
- [20] A. Izenman, "Recent development in nonparametric estimation" *Journal of American Statistical Association*, vol. 86, Issue 413 (Mar., 1991), pp 205-224

- [21] A. O. Hero, B. Ma, O. Michel, and J. Gorman, “ Applications of entropic spanning graphs,” *IEEE Signal Processing Magazine*, vol. 19, no. 5, pp. 85-95, 2002.
- [22] F. L. Bookstein, “*Morphometric Tools for Landmark Data*,” Cambridge Univ. Press, Cambridge, UK, 1991.
- [23] I. L. Dryden and K. V. Mardia, “*Statistical Shape Analysis*,” Wiley, Chichester, 1998.
- [24] S. Lee, G. Wolberg, K.-Y. Chwa, and S.Y. Shin, “Image Metamorphosis with Scattered Feature Constraints,” *IEEE Trans. Visualization and Computer Graphics*, vol. 2, pp. 337-354, 1996
- [25] N. Arad and D. Reisfeld, “Image warping using few anchor points,” *Comput. Graph. Forum*, vol. 14, pp. 35–46, 1995.
- [26] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, “Non-rigid registration using free-form deformations : application to breast MR images,” *IEEE Trans. on Medical Imaging*, vol. 18, pp. 712-721, 1999.
- [27] M. Satter, and A. Goshtasby, “Registration of deformed images,” *Proc. Image Registration Workshop*, Washington, DC, Nov., 1997
- [28] D. L. G. Hill, P. G .Batchelor, M. Holden, and D. J. Hawkes, “Medical image registration,” *Physics in medicine and biology*, vol. 46, pp. r1-r45, 2001
- [29] C. Studhome, D. L. G. Hill, and D. J. Hawkes, “An overlap invariant entropy measure of 3D medical image alignment,” *Pattern Recognition*, vol. 32, pp. 71-86, 1999.
- [30] H. J. Johnson and G. E. Christensen, “Landmark and Intensity-based, consistent thin plate spline image registration,” *IPMI 2001, LNCS 2082*, pp. 329-343, 2001
- [31] A. Kelemen, G. Szekely, and G. Gerig, “Elastic model-based segmentation of 3-D neuroradiological data sets,” *IEEE Transactions On Medical Imaging*, vol. 18, pp. 828-839, 1999.
- [32] D. MacDonald, D. Avis, and A. C. Evans, “Proximity constraints in deformable models for cortical surface identification,” in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 98, Lecture Notes in Computer Science*, vol. 1496, pp. 650-659, 1998.

- [33] D. MacDonald, N. Kabani, D. Avis, and A. C. Evans, "Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI," *Neuroimage*, vol. 12, pp. 340-356, 2000.
- [34] T. McInerney and D. Terzopoulos, "Topology adaptive deformable surfaces for medical image volume segmentation," *IEEE Transactions On Medical Imaging*, vol. 18, pp. 840-850, 1999.
- [35] J. Montagnat and H. Delingette, "Volumetric medical images segmentation using shape constrained deformable models," in *CVRMED-MRCAS'97, Lecture Notes in Computer Science*, vol. 1205, pp. 13-22, 1997.
- [36] W. J. Niessen, B. M. T. Romeny, and M. A. Viergever, "Geodesic deformable models for medical image analysis," *IEEE Transactions On Medical Imaging*, vol. 17, pp. 634-641, 1998.
- [37] D. G. Shen and C. Davatzikos, "An adaptive-focus deformable model using statistical and geometric information," *IEEE Transactions On Pattern Analysis and Machine Intelligence*, vol. 22, pp. 906-913, 2000.
- [38] A. Garrido and N. P. de la Blanca, "Applying deformable templates for cell image segmentation," *Pattern Recognition*, vol. 33, pp. 821-832, 2000.
- [39] H. Y. Wang and B. Ghosh, "Geometric active deformable models in shape modeling," *IEEE Transactions On Image Processing*, vol. 9, pp. 302-308, 2000.
- [40] X. L. Zeng, L. H. Staib, R. T. Schultz, and J. S. Duncan, "Segmentation and measurement of the cortex from 3-D MR images using coupled-surfaces propagation," *IEEE Transactions On Medical Imaging*, vol. 18, pp. 927-937, 1999.
- [41] X. C. Zhao and F. H. Qi, "Automatic image segmentation based on deformable models and pixel affinity statistic," *Journal of Infrared and Millimeter Waves*, vol. 19, pp. 33-37, 2000.
- [42] L. M. Gao, D. G. Heath, and E. K. Fishman, "Abdominal image segmentation using three-dimensional deformable models," *Investigative Radiology*, vol. 33, pp. 348-355, 1998.
- [43] I. D. Dinov, M. S. Mega, P. M. Thompson, L. Lee, et al., "Analyzing functional brain images in a probabilistic atlas: A validation of subvolume thresholding," *Journal of Computer Assisted Tomography*, vol. 24, pp. 128-138, 2000.

- [44] A. Guimond, J. Meunier, and J. P. Thirion, "Average brain models: A convergence study," *Computer Vision and Image Understanding*, vol. 77, pp. 192-210, 2000.
- [45] G. Le Goualher, E. Procyk, D. L. Collins, R. Venugopal, et al., "Automated extraction and variability analysis of sulcal neuroanatomy," *IEEE Transactions On Medical Imaging*, vol. 18, pp. 206-217, 1999.
- [46] G. Le Goualher, D. L. Collins, C. Barillot, and A. C. Evans, "Automatic identificaiton of cortical sulci using a 3D probabilistic atlas," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98, Lecture Notes in Computer Science*, vol. 1496, pp. 509-518, 1998.
- [47] P. Thompson and A. W. Toga, "Visualization and mapping of anatomic abnormalities using a probabilistic brain atlas based on random fluid transformations," in *Visualization in Biomedical Computing, Lecture Notes in Computer Science*, vol. 1131, pp. 383-392, 1996.
- [48] P. M. Thompson, D. MacDonald, M. S. Mega, C. J. Holmes, et al., "Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces," *Journal of Computer Assisted Tomography*, vol. 21, pp. 567-581, 1997.
- [49] P. M. Thompson, J. Moussai, S. Zohoori, A. Goldkorn, et al., "Cortical variability and asymmetry in normal aging and Alzheimer's disease," *Cerebral Cortex*, vol. 8, pp. 492-509, 1998.
- [50] P. M. Thompson, M. S. Mega, R. P. Woods, C. I. Zoumalan, et al., "Cortical change in Alzheimer's disease detected with a disease- specific population-based brain atlas," *Cerebral Cortex*, vol. 11, pp. 1-16, 2001.
- [51] A. W. Toga and P. M. Thompson, "The role of image registration in brain mapping," *Image and Vision Computing*, vol. 19, pp. 3-24, 2001.
- [52] K. L. Narr, P. M. Thompson, T. Sharma, J. Moussai, et al., "Three-dimensional mapping of gyral shape and cortical surface asymmetries in schizophrenia: Gender effects," *American Journal of Psychiatry*, vol. 158, pp. 244-255, 2001.
- [53] K. L. Narr, P. M. Thompson, T. Sharma, J. Moussai, et al., "Mapping morphology of the corpus callosum in schizophrenia," *Cerebral Cortex*, vol. 10, pp. 40-49, 2000.
- [54] J. L. Boes, P. H. Bland, T. E. Weymouth , L. E. Quint, et al., "Generating a normalized geometric liver model using warping," *Investigative Radiology*, vol. 29, pp. 281-286, 1994.

- [55] J. L. Boes, C. R. Meyer, and T. E. Weymouth, "Liver definition in CT using a population-based shape model," in *Computer Vision, Virtual Reality and Robotics in Medicine - CVRMed'95, Lecture Notes in Computer Science*, N. Ayache, Ed. Nice, FR: Springer-Verlag, Berlin, vol. 905, pp. 506-512, 1995.
- [56] J. L. Boes, T. E. Weymouth, and C. R. Meyer, "Multiple organ definition in CT using a Bayesian approach for 3D model fitting.," in *SPIE Vision Geometry IV*, vol. 2753. San Diego: SPIE, pp. 244-251, 1995.
- [57] A. C. Evans, <http://www.bic.mni.mcgill.ca/demos/animal/>.
- [58] M. I. Miller, G. E. Christensen, Y. Amit, and U. Grenander, "Mathematical textbook of deformable neuroanatomies.," *Proc. Natl. Acad. Sci. USA*, vol. 90, pp. 11944-8, 1993.
- [59] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinetics," *IEEE Trans. Image Proc.*, vol. 5, pp. 1435-1447, 1996.
- [60] G. E. Christensen, S. C. Joshi, and M. I. Miller, "Volumetric transformation of brain anatomy," *IEEE Trans. Medical Imaging*, vol. 16, pp. 864-877, 1997.
- [61] G. K. Rhode, A. Aldroubi, and B. M. Dawant, "Adaptive free form deformation for inter-patient medical image registration," *Information Processing in Medical Imaging (IPMI 2001)*, Proceedings of SPIE Vol. 4322, p. 1578-1787, 2001.???
- [62] G. K. Rhode, A. Aldroubi, and B. M. Dawant, "The adaptive bases algorithm for nonrigid image registration," *Information Processing in Medical Imaging (IPMI 2002)*, Proceeding of SPIE Vol. 4684, p. 933-944, 2002.
- [63] T. Rohlfing and C. R. Maurer, "Intensity based non-rigid registration using adaptive multilevel free form deformation with an incompressibility constraint," *MICCAI 2001, LNCS 2208*, p.111-119, 2001
- [64] A. Schnabel, D. Reuckert, M. Quist, et al., "A generic framework for non-rigid registration based on non-uniform multi-level free form deformations," *MICCAI 2001, LNCS 2208*, p.573-581, 2001.
- [65] M. I. Skolnik, "Introduction to radar systems," 3<sup>rd</sup> ed., McGraw-Hill, 2000.
- [66] C. A. Cocosco, V. Kollokian, R. K. - S. Kwan, and A. C. Evans, "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database," *NeuroImage*, vol.5, no.4, part 2/4, S425, 1997



- [67] J. C. Bezdek, L. O. Hall, and L. P. Clark, "Review of MR segmentation images using pattern recognition," *Medical Physics*, vol. 20, pp.1033–1048, 1993.
- [68] S. Warfield, "Fast k-NN classification for multichannel image data," *Pattern Recognition Letters*, vol. 17, pp. 713-721, 1996.
- [69] S. M. Belkasim, M. Shridhar, and M. Ahmadi, "Pattern classification using an efficient KNNR," *Pattern Recognition*, vol. 25, pp. 1269-1274, 1992.
- [70] T. K. Moon, "The Expectation-Maximization algorithm," *IEEE Signal Processing Magazine*, pp. 47-60, Nov. 1996.
- [71] M. L. Comer and E. J. Delp, "The EM/MPM algorithm for segmentation of textured images: Analysis and further experiment results," *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1731-1744, 2000.
- [72] M. L. Comer and E. J. Delp, "Multiresolution image segmentation," *Proceedings of the 1995 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Detroit, Michigan, pp. 2415-2418, May 1995.
- [73] P. Barone, A. Frigessi, and M. Piccioni, "*Stochastic models, statistical methods, and algorithms in image analysis*," Springer-Verlag, Berlin, 1992.
- [74] G. R. Cross and A. K. Jain, "Markov random field texture models," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 5, no. 1, pp.25-39, 1983.
- [75] T. Simchony, R. Chellappa, and Z. Lichtenstein, "Relaxation algorithms for MAP estimation of gray-level images with multiplicative noise," *IEEE Trans. on Information Theory*, vol. 36, no. 3, pp. 608-613, 1990.
- [76] D. E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning," Addison-Wesley, 1989.
- [77] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society B*, vol. 48, no. 3, pp. 259-302, 1986.
- [78] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic solution of ill-posed problems in computational vision," *Journal of the American Statistical Association*, vol. 82, pp. 76-89, 1987.
- [79] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society B*, vol. 36, no. 2, pp. 192-236, 1974.

- [80] X. Zeng, L.H.Staib, R.T.Schultz, and J.S.Duncan, "Segmentation and measurement of the cortex from 3-D MR images using coupled-surfaces propagation," *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 927 – 937, 1999.
- [81] D. H. Lawrence, K. W. Fleischer, and A. H. Barr, "Partial volume Bayesian classification of material mixtures in MR volume data using voxel histograms," *IEEE Transactions on Medical Imaging*, vol. 17, no. 1, pp. 74-86, 1998.
- [82] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images," *IEEE Transactions on Pattern Anal. Machine Intelli.*, vol. PAMI-6, no. 1, pp. 721-741, 1984.
- [83] J. Zhang, J. W. Modestino, and D. A. Langan, "Maximum-likelihood parameter estimation for unsupervised model based image segmentation," *IEEE Transactions on Image Processing*, vol. 3, pp. 404-420, 1994.
- [84] J. Montagnat, H. Delingette, and N. Ayache, "A review of deformable surfaces: topology, geometry and deformation," *Image and vision computing*, vol. 19, no. 14, pp. 1023-1040, 2001.
- [85] L. Limieux, U. C. Wiesmann, N. F. Moran, D. R. Fish, and S. D. Shorovon, "The detection and significance of subtle changes in mixed-signal brain lesion by serial MRI scan matching and spatial normalization," *Medical Image Analysis*, vol. 2, pp. 227-242, 1998.
- [86] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "A Viscous Fluid Model for Multimodal Non-rigid Image Registration Using Mutual Information," *MICCAI 2002*, LNCS vol. 2489, pp. 541-548, 2002.
- [87] H. Park and C. R. Meyer, "Locally adaptive registration using local information measures," submitted to *Medical Image Analysis*, Dec., 2002;
- [88] H. Park and C. R. Meyer, "Grid refinement in adaptive non-rigid registration," submitted *MICCAI 2003*, Apr., 2003.
- [89] H. Park, P. H. Bland, and C. R. Meyer, "Construction of an abdominal probabilistic atlas and its application in segmentation," to appear in *IEEE Transactions on medical imaging*, 2003.