Mathematical Methods for 2D-3D Cardiac Image Registration

by

Lok Wan Lorraine Ma

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Science

in

The Faculty of Science

Modelling and Computational Science

University of Ontario Institute of Technology

December 2016

Abstract

We propose a mathematical formulation aimed at intensity-based slice-to-volume registration, aligning a cross-sectional slice of a 3D volume to a 2D image. The approach is flexible and can accommodate various regularization schemes, similarity measures, and optimizers. We evaluate the framework by registering 2D and 3D cardiac magnetic resonance (MR) images obtained *in vivo*, aimed at image-guided surgery applications that utilise real-time MR imaging as a visualization tool. Rigid-body and affine transformations are used to validate the parametric model. Target registration error (TRE), Jaccard, and Dice indices are used to evaluate the algorithm and demonstrate the accuracy of the registration scheme on both simulated and clinical data. Registration with the affine model appeared to be more robust than the rigid model in controlled registration experiments. By simply extending the rigid model to an affine model, alignment of the cardiac region generally improved, without the need for complex dissimilarity measures or regularizers.

Keywords: image registration, inverse problems, slice-to-volume registration model, 2D to 3D alignment, cardiac MRI, multi-level, multi-resolution, optimization

Acknowledgements

First, I would like to thank my supervisor Dr. Mehran Ebrahimi for his guidance throughout this project. He introduced me to medical image processing and I am very grateful for his tremendous patience and support.

I would like to thank Drs. Graham Wright and Robert Xu of Sunnybrook Research Institute, Toronto, Canada for providing us with the MR data, especially Dr. Xu for his time spent over valuable discussions with us. I would also like to thank Dr. Jan Modersitzki for sharing his image registration codes with us.

I will always be grateful for my family, for all their support and wisdom. I would also like to express my gratitude to my friends, especially those in the Modelling and Computational Science program, who shared this journey with me, and to the members of the Science Faculty here at UOIT for fostering such a friendly learning environment.

Contents

List of Tables

List of Figures

List of Symbols

$\mathcal{T}$, $\mathcal{R}$, $\mathcal{J}$, $\mathcal{S}$, $\mathcal{D}$, $\mathcal{L}_z$	(Uppercase, stylized) Functions or functionals over a continuous domain
T, R, J, S, D, L_z	(Uppercase, italicized) Discrete operators, matrices
w, z, x, y	(Lowercase, italicized) Variables or coordinates in a continuous domain
`w`, `x`, `y`	(Lowercase, fixed-width) Variables or coordinates in a discretized domain
x^1, x^2, x^3	(Superscripted numerals) Denote correspondence to each of the 3 dimensions in the Cartesian coordinate system
Ω, $\mathcal{Z}$	Continuous domains
I	The identity matrix
$\mathbb{R}$	The set of all real numbers
$\|\cdot\|_\infty$	Infinity norm of a matrix
$\|\cdot\|_F$	Frobenius norm of a matrix
$\|\cdot\|_2$	Vector 2-norm
$^{-1}$	Inverse
^{T}	Transpose
$\leq$	Less than or equal to
$\approx$	Approximately equal to
$\lesssim$	Approximately equal to but never greater than

δ_{ij}	Kronecker delta
$\cap$	Intersection
$\cup$	Union

Chapter 1

Introduction

1.1 Problem Statement

Image registration is the task of aligning one image to another image. In a modern age where cameras are ubiquitous items, registration techniques find many applications, such as in image stitching. Image registration also finds important applications in clinical settings, where imaging equipment are indispensable diagnostic tools. It is useful at times for physicians to obtain images of the same region with different imaging methods, such as X-rays and magnetic resonance imaging (MRI), since different imaging methods contrast various body tissues differently. To combine information from two images taken at different times and see certain features in relation to other anatomical features, it may be necessary to align or *register* the images together. Other clinical applications include image-guided surgery, where registration can be used to combine pre-operative information with intra-operative information during surgery.

Cardiovascular disease is the leading cause of death globally, claiming more lives than cancer and chronic lower respiratory disease combined [M$^+$15]. In Canada, cardiovascular disease is responsible for approximately 1 in every 3 deaths, with a quarter of those deaths resulting from myocardial infarction (MI) [Pub09].

Reentrant arrhythmias are a common complication in patients with a history of MI [Ros13, KS09, KCKK06]. The arrhythmia is triggered when scarring disrupts electrical activity in the heart. Arrhythmias in the ventricles are potentially life threatening because they can render the heart unable to effectively circulate blood through the body, and are associated with increased risk of sudden cardiac death [KS09, KCKK06].

A range of treatment options exist, from non-invasive antiarrhythmic drugs to invasive procedures such as catheter ablation of problem areas of the heart [Ros13, SS07b, S$^+$08]. Catheter ablation is traditionally guided by X-ray fluoroscopy [BDTM12], but because of the ionizing radiation involved, alternative visualization methods have been proposed for use during surgical procedures [Lar00, TPL$^+$14].

Studies on carcinogenesis related to partial-body radiation exposure at levels used in medical imaging have produced conflicting results, so patient and staff exposure to X-rays should be kept minimal as a precaution [Jar16, G$^+$09]. For example, [CCB$^+$12] and [L$^+$04] report an association between dental X-rays and the risk of meningioma, while [XLH$^+$15] found no association between X-rays and meningioma. MRI is a well-known imaging technique that exploits the way different atoms behave in a magnetic field when perturbed by pulses of radio-frequency waves. Unlike X-rays, MRI does not measure tissue density along a projection axis and does not require the use of ionizing radiation. Different tissues contain different proportions of atoms and thus behave differently depending on the frequency of the radio waves, allowing MRI to distinguish between different types of soft tissues without the injection of contrast agents.

Additionally, because of its superior soft tissue contrast compared to X-rays, MRI better captures anatomical features in and around the heart and has been proposed as an alternative to guide catheter ablation and to aid pre-procedural planning [TPL$^+$14]. Recently, hybrid X-ray and MR (XMR) systems have been developed and tested for use in cardiac procedures, including catheterization, as a step towards reducing patient exposure to X-rays and toward fully MRI-guided intervention [BDTM12, R$^+$03a, R$^+$05,

A$^+$97, vV98, D$^+$05, Moo05].

MRI-guided procedures necessitate fast imaging techniques to capture 2D images in real-time. Fortunately, MR sequences for real-time visualization exist and have been used to guide catheterization procedures in studies on swine [B$^+$02, SSH$^+$03, R$^+$03b, H$^+$10, S$^+$09], dogs [L$^+$00, S$^+$06, SYF$^+$03], and humans [PTR14] and have been shown to be feasible. Other advantages of MRI over X-ray fluoroscopy commonly cited are the ability to easily adjust the positions of the imaging planes to access areas of interest [KHB$^+$08, PTR14, R$^+$03b, SS07a] and capture depth information without the need for multiple projections (as is the case for X-rays) as well as the ability to provide 3D anatomical information in the form of multi-planar volumes [SSH$^+$03, PTR14, KHB$^+$08, R$^+$05, SS07a].

One obstacle to MRI-guided catheterization is the limited availability of MRI-safe tools and electronic equipment to be used during intervention [Moo05]. Most catheters were designed to be used with X-rays and potentially contain materials rendering them hazardous in a strong magnetic field or induce artefacts, or are almost invisible in MR scans [MR05, R$^+$05, KHB$^+$08]. Implantable cardioverter defibrillators (ICDs) in patients must also be MRI-safe.

The comparison between pre-operative images and intra-operative images is essential to image-guided procedures. Pre-operative images are used in treatment planning while 2D intra-operative images provide live positional updates to aid surgeons in carrying out treatment. The tradeoff between image quality and acquisition time means that high-quality 3D MR image volumes can be acquired prior to intervention without the constraints of producing images at real-time frame rates. Each volume consists of a stack of 2D image slices, with each of those slices imaged at a resolution higher than that of the 2D real-time images, allowing anatomical features to be seen in detail. During MRI-guided procedures, surgeons can only depend on a mental picture of how the 2D real-time images are positioned relative to the 3D pre-procedural images, so ideally, one

would register the 3D pre-procedural images to 2D real-time images. Registration would combine the advantages of both image types – the image quality of the pre-procedural images and the live positional updates of the intra-operative images – to obtain high-quality images that account for small amounts of motion, such as motion due to respiration, in real time.

In this thesis we present a mathematical framework for slice-to-volume 2D-3D registration, implement registration on cardiac MR images, and investigate the ill-posedness that arises from an affine registration model.

1.2 Literature Survey

Many papers have been published presenting methods to register 3D pre-operative images to 2D intra-operative images in clinical applications. [TLSP03] presents a solely intensity-based approach to registering a pre-operative 3D CT or MR image to an intra-operative 2D X-ray image, with no fiducial markers or need for intra-operative X-ray segmentation or any sort of digital reconstruction. [PBH$^+$01] validates an algorithm that registers 3D pre-operative CT volumes to intra-operative 2D (X-ray) fluoroscopy images. [TLP06] presents a novel 2D-3D registration method that registers a 3D CT or MR image to a 3D image reconstructed from a set of two or more intra-operative X-ray images. [XW08] presents an efficient 2D-3D registration method to register a 3D volume to a simulated X-ray image and implements it in parallel. [MTPL08] presents a gradient-based method to register pre-operative 3D CT or MRI volumes to intra-operative 2D X-rays. [LYJ03] presents a gradient-based rigid-body registration method for registering pre-operative CT or MRI to intra-operative fluoroscopic X-ray or ultrasound images with applications to image-guided surgery and robotic positioning and validates the method on various simulated and *in vitro* conditions.

Numerous studies in the literature focus on alignment of pre-operative images to X-

ray intra-operative images (fluoroscopy and CT use X-rays). Due to the drawbacks of X-rays, however, more recent studies have begun to examine the use of MRI instead of X-rays in some applications, especially when imaging areas containing large amounts of soft tissue.

During MRI-guided procedures, acquisition time for real-time, intra-operative images is limited and therefore noisier compared to pre-operative images. Real-time images are also only available in 2D, whereas the high-quality pre-operative images are available in 3D. One way of combining information from both pre- and intra-operative images is to register them.

In registering 3D to 2D images, there is a distinction between the 2D image being an X-ray or an MR image. X-rays is a projection imaging modality, whereas a 2D MR image is obtained by imaging a cross-sectional slice of the 3D object. In registering a 3D image to a cross-sectional slice of a volume, information outside of the plane is lost, whereas in registering a 3D volume to a projection, elements of the 3D volume along projection rays are retained.

Methods to perform slice-to-volume registration have been proposed. [DA08] presents a slice-to-volume rigid registration method based on phase congruency and validates the method on simulated MR images, but registers multiple slices to a volume instead of a single slice and is feature-based. [OK10] presents a non-rigid slice-to-volume registration method that combines a rigid registration step with a deformable registration step. [Z$^+$10] proposed a registration using Markov random fields (MRF) and optimization methods and performed registration with rigid, similarity, and affine transformations. [FP13] and [FFP15] also present methods based on MRF and optimization, but for deformable (not rigid) registration. [B$^+$07] implements rigid slice-to-volume registration between CT and FluoroCT images, which are cross-sectional slices, not projective images, and notes the need for strict breath-hold techniques to reduce artefacts due to respiratory motion when rigid registration is employed.

Of interest to us is the application of slice-to-volume registration to MRI-guided procedures. [XLF+15] developed a multi-slice-to-volume registration algorithm to align images in the context of MRI-guided biopsy, but on a more relevant note, work has also been done *in vivo* on pigs to register 2D intra-operative cardiac MR image slices to pre-operative MR image volumes [SWD04, SWWFD05].

In humans, 2D-3D slice-to-volume registration between high-quality pre-operative MR image volumes and live-time intra-operative MR images have been studied for applications to MRI-guided ablation of prostate cancer [FDB+03]. [XW15] also present registration of high-quality pre-operative MR image volumes to live cardiac MR images on human volunteers *in vivo* with applications to MRI-guided radiofrequency ablation of substrate in the heart, but also focuses on registration incorporating rigid-body transformations.

In most of the studies mentioned above, rigid-body registration was employed. Alignments are only made by translating and rotating the 3D volume to match a 2D projection or slice. While rigid registration is generally employed to reduce computational cost and to speed up the registration process, it risks oversimplifying the displacement of body tissues, which are generally not rigid. The highly deformable nature of the heart and displacement at various stages of the breathing cycle make registration of the cardiac region more challenging. Registration accuracy at the millimetre scale is important during image-guided cardiac intervention, but is influenced by various imaging parameters. Deformable registration may be more accurate, but is computationally much more expensive.

1.3 Objective

2D-2D and 3D-3D rigid and affine registration is widely employed and readily available in software such as Insight Segmentation and Registration Toolkit (ITK), and methods

for projective 2D-3D registration have been and are being researched, particularly for applications involving X-rays or X-ray fluoroscopy (such as some image-guided electrophysiology procedures).

Work involving 2D-3D slice-to-volume registration has been fairly recent and not nearly as numerous as projective 2D-3D registration, especially with respect to applications in MRI-guided procedures. In addition, there seems to be a lack of precise model in the literature, in contrast to 2D-2D or 3D-3D registration [Mod09].

We propose a general mathematical framework for slice-to-volume registration which can accommodate parametric and non-parametric transformation models. A rigid transformation model can be used in this framework, but the user can easily adapt a different parametric transformation model.

We will demonstrate this framework on parametric models, specifically, using this framework to extend existing 2D-3D rigid registration to affine registration. Although the number of parameters in an affine parametric model (12 parameters) is twice the number of parameters in a rigid model (6 parameters), the figure dwarfs in comparison to the number of parameters dealt with in deformable registration, and thus is still a computationally inexpensive method that accounts for some non-rigid deformations.

The intensity-based registration framework is flexible and can accommodate various models and parameters. We demonstrate by registering high-resolution 3D MR images to noisier 2D real-time MR images, using rigid and affine parametric models, and investigate the ill-posedness of 2D-3D registration as an inverse problem.

1.4 Outline

In Chapter 2, I will briefly discuss some background to the thesis – notes on software used, a review of parametric transformations, and a brief introduction to mathematical optimization, the core of image registration. In Chapter 3, 2D-2D image registration will

be introduced as an optimization problem to familiarize the reader with the basic concepts, then present a 2D-3D slice-to-volume registration model, which is investigated as an inverse problem. Chapter 4 discusses various registration experiments performed on cardiac MR images to test our model. It also describes the data used as well as methods used in validating the results and some discussion of the results. Finally, Chapter 5 summarizes our observations of experimental results and includes general remarks regarding the slice-to-volume registration model, and concludes with topics of further interest to us should we be presented with the opportunity to expand on the work presented here.

Chapter 2

Preliminary Notes

2.1 Software

MATLAB[1] was chosen for implementation due to its simplicity, its arsenal of built-in functions and image viewing tools, and handy debugging tools. Though not the fastest or most efficient language (compared to C, for example), MATLAB is geared towards mathematical applications so one can write technical code in a straightforward manner without having to deal with the intricacies of the programming task. Because this project focusses on the mathematical modelling aspects of image registration, speed and efficient use of computational resources were not top priority.

Jan Modersitzki's Flexible Algorithms for Image Registration (FAIR) [Mod09] is a code package written mostly in MATLAB to perform 2D-2D registration and 3D-3D registration. Modersitzki also has an accompanying book of the same name documenting the code.

To avoid re-inventing the wheel, our 2D-3D registration codes are written as an add-on to FAIR, keeping a similar workflow and using FAIR code whenever possible.

[1] ©2016 The MathWorks, Inc. MATLAB and Simulink are registered trademarks of The MathWorks, Inc. See mathworks.com/trademarks for a list of additional trademarks. Other product or brand names may be trademarks or registered trademarks of their respective holders.

2.2 Parametric Transformations and Inverses

Various 2D parametric transformations on a point $x = (x^1, x^2)^T \in \mathbb{R}^2$ or 3D parametric transformations on a point $x = (x^1, x^2, x^3)^T \in \mathbb{R}^3$ are defined in this section. Each transformation is parameterized by w. Superscripts 1, 2, and 3 each relate to the x-, y-, and z-coordinates in the Cartesian coordinate system, respectively.

2D Rigid In 2D, w contains 3 parameters: 1 rotation (w_1) and 2 translations (w_2, w_3). (Some conventions refer to translations as *shifts*.) A rigid transformation on x is

$$y(w, x) = \begin{pmatrix} \cos w_1 & -\sin w_1 \\ \sin w_1 & \cos w_1 \end{pmatrix} x + \begin{pmatrix} w_2 \\ w_3 \end{pmatrix}.$$

2D Rigid – Inverse The inverse transform can be obtained by inverting the translation and then the rotation.

The inverse transformation would simply involve negating the translation terms and rotating in the opposite direction:

$$\begin{aligned} x &= \begin{pmatrix} \cos(-w_1) & -\sin(-w_1) \\ \sin(-w_1) & \cos(-w_1) \end{pmatrix} \left[y - \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right] \\ &= \begin{pmatrix} \cos(w_1) & -\sin(w_1) \\ \sin(w_1) & \cos(w_1) \end{pmatrix}^T \left[y - \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right]. \end{aligned}$$

Note that rotation matrices are orthogonal (i.e. if a matrix Q is orthogonal, $Q^{-1} = Q^T$), so the inverse of a rotation matrix is its transpose.

3D Rigid In 3D, the transformation consists of 3 rotations followed by 3 translations:

$$y(w, x) = R^1(w_1)R^2(w_2)R^3(w_3)x + \begin{pmatrix} w_4 \\ w_5 \\ w_6 \end{pmatrix}. \tag{2.1}$$

w_4, w_5, and w_6 are translations in the x-, y-, and z-directions, respectively. If w_1, w_2, and w_3 are angles of rotation about the x-, y-, and z-axes, rotation operations about the x-, y-, and z-axes are defined to be

$$R^1(w_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos w_1 & -\sin w_1 \\ 0 & \sin w_1 & \cos w_1 \end{pmatrix}, \quad R^2(w_2) = \begin{pmatrix} \cos w_1 & 0 & \sin w_1 \\ 0 & 1 & 0 \\ -\sin w_1 & 0 & \cos w_1 \end{pmatrix}, \tag{2.2}$$

$$\text{and} \quad R^3(w_3) = \begin{pmatrix} \cos w_3 & -\sin w_3 & 0 \\ \sin w_3 & \cos w_3 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.3}$$

3D Rigid – Inverse For convenience, let us define the composition of the three rotations to be $R = R^1(w_1)\, R^2(w_2)\, R^3(w_3)$. As in the 2D case, if the forward transform is

$$y(w, x) = Rx + \begin{pmatrix} w_4 \\ w_5 \\ w_6 \end{pmatrix}, \tag{2.4}$$

then

$$x = R^{-1} \left[y - \begin{pmatrix} w_4 \\ w_5 \\ w_6 \end{pmatrix} \right] = R^T \left[y - \begin{pmatrix} w_4 \\ w_5 \\ w_6 \end{pmatrix} \right], \tag{2.5}$$

since the matrix R_w is orthogonal.

2D Affine In 2D, the affine transformation is parametrized by 6 parameters $w = (w_1, w_2, \ldots, w_6)$. If we define

$$A = \begin{pmatrix} w_1 & w_2 \\ w_4 & w_5 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} w_3 \\ w_6 \end{pmatrix}, \tag{2.6}$$

the transformation on a point $x = (x^1, x^2)^T$ is

$$y(w, x) = Ax + b.$$

2D Affine – Inverse Assuming that A is invertible, the inverse process is similar to the rigid case. We obtain

$$x = A^{-1}(y - b). \tag{2.7}$$

In registering images of real objects, transformations are required to be realistic and physically feasible. It is safe to assume that all parametric transformations in this thesis, which are applied to cardiac MR images, are physically feasible and therefore non-degenerate and invertible.

3D Affine If we define

$$A = \begin{pmatrix} w_1 & w_2 & w_3 \\ w_5 & w_6 & w_7 \\ w_9 & w_{10} & w_{11} \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} w_4 \\ w_8 \\ w_{12} \end{pmatrix}, \tag{2.8}$$

then an 3D affine transformation on a point with coordinates $x = (x^1, x^2, x^3)^T$ is defined as

$$y(w, x) = Ax + b.$$

3D Affine – Inverse Again, assuming that the transformation A is physically feasible and therefore invertible,

$$x = A^{-1}(y - b). \tag{2.9}$$

2.3 Optimisation

Mathematical optimization involves determining parameters that will maximize or minimize the value of a function. We will use the terms optimization and minimization interchangeably in this thesis, since the maximization of a function is equivalent to the minimization of the negative of the function. The function is usually called the objective

function or the cost because it is the value in question we want to reduce or a penalty to be avoided.

Some objective functions are fairly simple to optimize and can be done analytically. One such example is finding the quickest path on which an ideal point object would roll down under the influence of gravity, for which the solution is called a brachistone curve. The solution is a general one, so parameters can be substituted into the solution to tailor it to a specific situation.

Other objective functions are more complex and may not have a solution in the form of an analytical expression. Or, the number of variables and parameters involved is too large to be solved using pen and paper. These problems may be difficult on paper, but such problems can be discretized and solved iteratively to approximate a solution. Many numerical methods have been developed and used to solve various problems computationally [CdB80, NW06, QSS07, BGLS06].

Image processing is often a repetitive task, where evaluations and calculations are repeatedly performed over individual pixels or voxels in an image. The problem in question may also be a complex one, necessitating brute-force numerics in determining an approximate solution. This calls for numerical methods.

In the next section, image registration is introduced as a minimization problem and discussed.

Chapter 3

Image Registration

Image registration is the process of finding a transformation on the *coordinates* of the pixels or voxels of one image to align it to another. Registration techniques can generally be divided into two categories: intensity-based and landmark-based methods. Intensity-based methods perform registration based on pixel or voxel intensities of the images, whereas landmark-based or geometry-based methods use markers or features of the image and track the movement of these markers or features. Because landmark-based methods rely on features in the image, such methods require fiducial markers or segmentation to be performed beforehand, whereas intensity-based methods do not single out features and rely only on the intensity values of an image. This thesis will solely relate to intensity-based methods.

3.1 Introduction to 2D-2D Registration

It is helpful to first familiarize ourselves with the language and conventions used in 2D-2D registration before 2D-3D registration is introduced.

Consider the problem of aligning a 2D 'template' image $\mathcal{T}$ to another 2D 'reference' image $\mathcal{R}$, where $\mathcal{R}$ is a realization of $\mathcal{T}$ deformed via a transformation. The reference and template images are defined over the 2D domain Ω and are represented by mappings

$\mathcal{R} : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ and $\mathcal{T} : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ of compact support.

The following is written for 2D-2D registration, but can easily be generalized to 3D-3D registration.

3.2 2D-2D Objective Function

The goal is to find a transformation $y : \mathbb{R}^2 \to \mathbb{R}^2$ on the set of coordinates $x \in \Omega \subset \mathbb{R}^2$ that makes the transformed template image $\mathcal{T}$ most similar to the reference image $\mathcal{R}$. The objective function to be minimized is

$$\mathcal{J}[y] := \mathcal{D}[\mathcal{T}(y(x)), \mathcal{R}(x)] + \mathcal{S}[y(x)], \tag{3.1}$$

$\mathcal{D}$ is a measure of dissimilarity between images $\mathcal{T}$ and $\mathcal{R}$. $\mathcal{S}$, called the regularizer, imposes restrictions on the transformation y.

Registration is ill-posed in general, so there may be more than a single solution for a given problem. We can impose restrictions on the transformation in the form of a regularizer, which imposes penalties on certain transformations by increasing the value of the objective function. An obvious example is registering the two images of a square in Figure 3.1. If θ represents the rotation angle about the origin, there are already an infinite number of solutions $\theta = \frac{\pi}{4} + k\frac{\pi}{2}, k \in \mathbb{Z}$. A regularizer could be added to restrict the number of solutions, for example, one which keeps the amount of rotation small by penalising the value of $|k|$ or k^2.

Both $\mathcal{D}$ and $\mathcal{S}$ will be discussed in further detail in Sections 3.4, 3.8, and 3.9. Knowledge about the imaging methods used to obtain $\mathcal{T}$ and $\mathcal{R}$ and about the objects imaged can aid in choosing a dissimilarity measure and a regularizer.

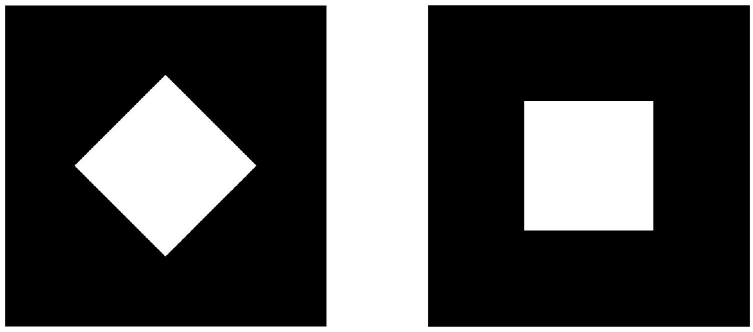


Figure 3.1: A registration problem with infinitely many solutions. There is not one unique solution but an infinite number of transformations that would register one to the other.

3.3 2D-2D Parametrized Objective Function

Suppose that the transformation y can be parametrized via parameters w. The number of parameters contained in w will depend on the transformation chosen. Common transformations are described in 2.2. The parametrized objective function is

$$\mathcal{J}[w] := \mathcal{D}[\mathcal{T}(y(w,x)), \mathcal{R}(x)] + \mathcal{S}(w). \tag{3.2}$$

Instead of minimizing $\mathcal{J}$ over y, it is minimized over w, and $\mathcal{S}$ now acts on the parameters w.

The motivation for parameterization is explained in the next section.

3.3.1 2D-2D Discretized Objective Function

We employ a discretize-then-optimize paradigm, discretizing before numerically minimizing the objective. Although each component of the objective can be chosen such that the problem appears to be quadratic, the problem is actually highly non-linear – more details can be found in Section 3.12. This contrasts to an optimize-then-discretize paradigm

where the problem is optimized analytically before the rest of the problem is computed numerically. Examples employing an optimize-then-discretize approach include certain registration problems solved using a partial differential equations (PDEs).

Discretizing Ω into $m \times n$ pixels, we can define a discretized grid $[x_k^1, x_k^2]_{k=1,\ldots,mn}$ and call this $\mathtt{x}$. Images $\mathcal{R}$ and $\mathcal{T}$ are approximated on the discrete grid to obtain cell-centered discretizations $R \approx \mathcal{R}(\mathtt{x})$ and $T \approx \mathcal{T}(\mathtt{x})$, each with $m \times n$ pixels. The discretized transformed coordinates are $\mathtt{y} \approx y(\mathtt{w}, \mathtt{x})$ where $\mathtt{w} = w$, for consistency of notation, and the transformed template image $T[\mathtt{y}] \approx \mathcal{T}(\mathtt{y})$. Furthermore, D and S represent the discretized versions of $\mathcal{D}$ and $\mathcal{S}$.

The discretized, parametrized objective function is

$$J[w] := D[T(\mathtt{y}(\mathtt{w}, \mathtt{x})), R(\mathtt{x})] + S(w), \tag{3.3}$$

For ease of computation, the elements of $\mathtt{w}$, $\mathtt{y}$, R, and T are laid out into ordered column vectors.

The motivation for parameterization is to reduce computational cost when performing registration. Parameterization reduces the number of variables over which to minimize the objective function. Without parameterization, the objective function is minimized over $\mathtt{y}$, which contains the same number of elements as the number of pixels in the discretized image. With parametrization, the objective function is is minimized over $\mathtt{w}$, which is typically chosen to contain far less elements compared to $\mathtt{y}$. For example, the 2D rigid-body and affine transformations can be parameterized by 3 and 6 variables, respectively. The advantage is faster registration, but because parameterization restricts y to a certain class of transformations, it is important to choose a model wisely.

3.4 Dissimilarity measures and derivatives (2D-2D)

Since the goal of registration here is to align the images by minimization of an objective function, the dissimilarity measure plays a central role in how aligned two images are.

The dissimilarity measure should be chosen to be small when the template and reference images appear to be aligned or similar and large when they are not. The choice of dissimilarity measure therefore depends on the type of images to be registered.

In the numerical minimization of J, most algorithms require the derivative with respect to the variable over which J is minimized, and some require the second derivative as well. The following are several common dissimilarity measures and their derivatives. The derivative of each $D[T, R]$ is computed with respect to the template image T, which is needed during numerical optimization of the objective function.

3.4.1 Sum of Squared Distances

The sum of squared distances (SSD) measures differences in pixel values between two images. Since this measure depends on crude image values, without extracting and using any additional information, it works best on images of the same modality, i.e. images taken with the same type of imaging equipment.

Continuous domain

In the continuous domain, the SSD is defined as

$$\mathcal{D}^{\text{SSD}}[\mathcal{T}, \mathcal{R}] = \frac{1}{2} \int_{\Omega} (\mathcal{T}(x) - \mathcal{R}(x))^2 \, \mathrm{d}x. \tag{3.4}$$

Discretized domain

In the discrete case, the SSD sums up the squared differences in individual pixel values between the reference and template images. Because the sum depends on the level of discretization, it is scaled by `hd`, the area of each pixel.

$$D^{\text{SSD}}[T, R] = \frac{1}{2} \, \texttt{hd} \, \|T - R\|^2 = \frac{1}{2} \, \texttt{hd} \sum_{i=1}^{mn} (T_i - R_i)^2 \tag{3.5}$$

An obvious example of a case where the SSD is not expected to measure alignment well is one where the template image is a negative of the reference image, so the measured

dissimilarity is large when the two images are perceived to be most aligned.

Derivative over the discretized domain

The dissimilarity measure $D[T, R]$ can be written as $\psi(r)$ (see p. 76 of [Mod09]). For the SSD, define $r_i = T_i - R_i$. Then the dissimilarity measure is

$$\psi(r) = \frac{1}{2}\,\mathtt{hd}\sum_{i=1}^{mn} r_i^2. \tag{3.6}$$

To compute $\frac{\partial\psi}{\partial T}$, the derivative of the dissimilarity measure with respect to T, use the chain rule:

$$\frac{\partial\psi}{\partial T} = \frac{\partial\psi}{\partial r}\frac{\partial r}{\partial T}. \tag{3.7}$$

Because $\frac{\partial\psi}{\partial r_i} = \mathtt{hd}\, r_i$,

$$\frac{\partial\psi}{\partial r} = \mathtt{hd}\begin{pmatrix}\frac{\partial\psi}{\partial r_1} & \frac{\partial\psi}{\partial r_2} & \dots & \frac{\partial\psi}{\partial r_n}\end{pmatrix} = \mathtt{hd}\begin{pmatrix} r_1 & r_2 & \dots & r_n\end{pmatrix} = \mathtt{hd}\, r^T. \tag{3.8}$$

Note that $\frac{\partial r}{\partial T} = I$, the identity matrix. Putting the two terms together yields

$$\frac{\partial\psi}{\partial T} = \mathtt{hd}\, r^T\, I = \mathtt{hd}\, r^T. \tag{3.9}$$

Since $\frac{\partial^2\psi}{\partial r_i^2} = \mathtt{hd}$, second derivative of ψ with respect to r can be computed as

$$\frac{\partial^2\psi}{\partial r^2} = \begin{pmatrix} \frac{\partial^2\psi}{\partial r_1\,\partial r_1} & \frac{\partial^2\psi}{\partial r_1\,\partial r_2} & \cdots & \frac{\partial^2\psi}{\partial r_1\,\partial r_n} \\ \frac{\partial^2\psi}{\partial r_2\,\partial r_1} & \ddots & & \\ \vdots & & \ddots & \\ \frac{\partial^2\psi}{\partial r_n\,\partial r_1} & & & \frac{\partial^2\psi}{\partial r_n\,\partial r_n}\end{pmatrix} = \mathtt{hd}\, I. \tag{3.10}$$

3.4.2 Normalized Cross-Correlation

Continuous domain In the FAIR book, it is not documented how the approximation used in the code for the derivative of the NCC distance measure is obtained. Although the NCC was not used as a distance measure in our registration experiments, it is worthwhile to present here an explanation for the approximation, as our codes are written as an add-on to FAIR and allows the user flexibility in choosing a dissimilarity measure.

A measure of *similarity* between $\mathcal{T}$ and $\mathcal{R}$ is $\langle \mathcal{T}, \mathcal{R} \rangle = \int_\Omega \mathcal{T}(x)\,\mathcal{R}(x)\,\mathrm{d}x$, known as the cross-correlation. One would maximize the similarity when registering two images.

From that, we can define a measure of *dissimilarity* by normalizing $\langle \mathcal{T}, \mathcal{R} \rangle$, squaring, and subtracting from 1. (The negation of the normalized cross-correlation term converts it into a minimization problem.)

Assuming that $\mathcal{T}$ and $\mathcal{R}$ are non-zero, content-less images so $\|T\| \neq 0$ and $\|R\| \neq 0$, the normalized cross correlation distance measure (NCC) is defined to be

$$\mathcal{D}^{\mathrm{NCC}}[\mathcal{T}, \mathcal{R}] = 1 - \left(\frac{\langle \mathcal{T}, \mathcal{R} \rangle}{\|\mathcal{T}\|\,\|\mathcal{R}\|} \right)^2 = 1 - \frac{\langle \mathcal{T}, \mathcal{R} \rangle^2}{\|\mathcal{T}\|^2\,\|\mathcal{R}\|^2} \tag{3.11}$$

where

$$\|\mathcal{T}\| = \sqrt{\langle \mathcal{T}, \mathcal{T} \rangle}. \tag{3.12}$$

Discretized domain Discretization of the normalized cross-correlation dissimilarity measure (NCC) is straightforward. The images $\mathcal{T}$ and $\mathcal{R}$ are each discretized into $m \times n$ pixels, and each integral discretized into a sum, as was done in the discretization of the SSD. Again, assuming that both T and R do not contain pixels all with values of 0 (so the expression is always defined) and re-writing the distance measure as ψ,

$$\psi(T, R) = D^{\mathrm{NCC}}[T, R] = 1 - \frac{\left(\sum_{i=1}^{mn} T_i\,R_i\right)^2}{\left(\sum_{i=1}^{mn} T_i^2\right)\left(\sum_{i=1}^{mn} R_i^2\right)}. \tag{3.13}$$

Derivative over the discretized domain Using the quotient rule and combining terms, the derivative of the distance measure with respect to a pixel in T is

$$\frac{\partial \psi}{\partial T_j} = -\frac{2R_j \left(\sum_{i=1}^{mn} T_i R_i\right)\left(\sum_{i=1}^{mn} T_i^2\right)\left(\sum_{i=1}^{mn} R_i^2\right) - 2T_j \left(\sum_{i=1}^{mn} R_i^2\right)\left(\sum_{i=1}^{mn} T_i R_i\right)^2}{\left(\sum_{i=1}^{mn} T_i^2\right)^2 \left(\sum_{i=1}^{mn} R_i^2\right)^2} \tag{3.14}$$

$$= -\frac{2R_j \langle T, R \rangle \|T\|_F^2 \|R\|_F^2 - 2T_j \|R\|_F^2 \langle T, R \rangle^2}{\|T\|_F^4 \|R\|_F^4} \tag{3.15}$$

$$= -\frac{2 \langle T, R \rangle \|R\|_F^2}{\|T\|_F^4 \|R\|_F^4} \left[R_j \|T\|_F^2 - T_j \langle T, R \rangle \right] \tag{3.16}$$

$$= -\frac{2 \langle T, R \rangle}{\|T\|_F^2 \|R\|_F^2} \left[R_j - \frac{\langle T, R \rangle}{\|T\|_F^2} T_j \right]. \tag{3.17}$$

The Jacobian with respect to the entire image T is

$$\frac{\partial \psi}{\partial T} = -\frac{2\langle T, R\rangle}{\|T\|_F^2 \|R\|_F^2}\left[R^T - \frac{\langle T, R\rangle}{\|T\|_F^2} T^T\right] \tag{3.18}$$

$$= \frac{2}{\|R\|_F^2}\left[\frac{\langle T, R\rangle^2}{\|T\|_F^4} T^T - \frac{\langle T, R\rangle}{\|T\|_F^2} R^T\right]. \tag{3.19}$$

The second derivatives can be found by taking derivatives in a similar manner. The Hessian of ψ with respect to the elements of T is:

$$\frac{\partial^2 \psi}{\partial T^2} = \frac{\partial}{\partial T}\left(\frac{\partial \psi}{\partial T}\right) \tag{3.20}$$

$$= \begin{pmatrix} \frac{\partial^2 \psi}{\partial T_1 \partial T_1} & \frac{\partial^2 \psi}{\partial T_1 \partial T_2} & \cdots & \frac{\partial^2 \psi}{\partial T_1 \partial T_{mn}} \\ \frac{\partial^2 \psi}{\partial T_2 \partial T_1} & \frac{\partial^2 \psi}{\partial T_2 \partial T_2} & & \\ \vdots & & \ddots & \\ \frac{\partial^2 \psi}{\partial T_{mn} \partial T_1} & & & \frac{\partial^2 \psi}{\partial T_{mn} \partial T_{mn}} \end{pmatrix}. \tag{3.21}$$

Each entry of the Hessian matrix above can be found individually, i.e.

$$\frac{\partial^2 \psi}{\partial T_i \partial T_j} = \frac{\partial}{\partial T_i}\left(\frac{\partial \psi}{\partial T_j}\right) \tag{3.22}$$

$$= \frac{2}{\|R\|_F^2}\frac{\partial}{\partial T_i}\left[\frac{\langle T, R\rangle^2}{\|T\|_F^4} T_j - \frac{\langle T, R\rangle}{\|T\|_F^2} R_j\right] \tag{3.23}$$

$$= \frac{2}{\|R\|_F^2}\left[\frac{\partial}{\partial T_i}\frac{\langle T, R\rangle^2}{\|T\|_F^4} T_j - \frac{\partial}{\partial T_i}\frac{\langle T, R\rangle}{\|T\|_F^2} R_j\right] \tag{3.24}$$

$$= \frac{2}{\|R\|_F^2}\left[\frac{\partial}{\partial T_i}\frac{\left(\sum_{k=1}^{mn} T_k R_k\right)^2 T_j}{\left(\sum_{k=1}^{mn} T_k^2\right)^2} - \frac{\partial}{\partial T_i}\frac{\left(\sum_{k=1}^{mn} T_k R_k\right) R_j}{\left(\sum_{k=1}^{mn} T_k^2\right)}\right] \tag{3.25}$$

$$\stackrel{\text{def}}{=} \frac{2}{\|R\|_F^2}\left[\frac{\partial}{\partial T_i} A - \frac{\partial}{\partial T_i} B\right]. \tag{3.26}$$

Each term in the brackets expands into a long expression, so we treat them separately

and simplify as follows:

$$\frac{\partial}{\partial T_i}A = \frac{\left[R_i\, 2\,\left(\sum_{k=1}^{mn} T_k R_k\right) T_j + \left(\sum_{k=1}^{mn} T_k R_k\right)^2 \left(\frac{\partial T_j}{\partial T_i}\right)\right]\left(\sum_{k=1}^{mn} T_k^2\right)^2}{\left(\sum_{k=1}^{mn} T_k^2\right)^4} \tag{3.27}$$

$$- \frac{\left(\sum_{k=1}^{mn} T_k R_k\right)^2 T_j \left[2\, T_i\, 2\,\left(\sum_{k=1}^{mn} T_k^2\right)\right]}{\left(\sum_{k=1}^{mn} T_k^2\right)^4} \tag{3.28}$$

$$= \frac{\left[2\,\langle T, R\rangle\, T_j R_i + \langle T, R\rangle^2\, \delta_{ij}\right] \|T\|_F^4 - 4\,\|T\|_F^2\,\langle T, R\rangle^2\, T_i T_j}{\|T\|_F^8} \tag{3.29}$$

$$= \frac{2\,\langle T, R\rangle}{\|T\|_F^4} T_j R_i + \frac{\langle T, R\rangle^2}{\|T\|_F^4}\delta_{ij} - \frac{4\,\langle T, R\rangle^2}{\|T\|_F^6} T_i T_j \tag{3.30}$$

$$\frac{\partial}{\partial T_i}B = \frac{R_i R_j \left(\sum_{k=1}^{mn} T_k^2\right) - \left(\sum_{k=1}^{mn} T_k R_k\right) R_j\, 2\, T_i}{\left(\sum_{k=1}^{mn} T_k^2\right)^2} \tag{3.31}$$

$$= \frac{1}{\|T\|_F^2} R_i R_j - \frac{2\,\langle T, R\rangle}{\|T\|_F^4} T_i R_j. \tag{3.32}$$

Substituting the simplified expressions back into Equation 3.26 and combining terms, we obtain

$$\frac{\partial^2 \psi}{\partial T_i\, \partial T_j} = \frac{2}{\|R\|_F^2}\left[\frac{2\,\langle T, R\rangle}{\|T\|_F^4} T_j R_i + \frac{\langle T, R\rangle^2}{\|T\|_F^4}\delta_{ij} - \frac{4\,\langle T, R\rangle^2}{\|T\|_F^6} T_i T_j - \frac{1}{\|T\|_F^2} R_i R_j + \frac{2\,\langle T, R\rangle}{\|T\|_F^4} T_i R_j\right] \tag{3.33}$$

$$= \frac{2}{\|R\|_F^2 \|T\|_F^2}\left[\frac{\langle T, R\rangle^2}{\|T\|_F^2}\delta_{ij} - \frac{4\,\langle T, R\rangle^2}{\|T\|_F^4} T_i T_j + \frac{2\,\langle T, R\rangle}{\|T\|_F^2} T_i R_j + \frac{2\,\langle T, R\rangle}{\|T\|_F^2} T_j R_i - R_i R_j\right] \tag{3.34}$$

$$= \frac{2}{\|R\|_F^2 \|T\|_F^2}\left[\frac{\langle T, R\rangle^2}{\|T\|_F^2}\delta_{ij} + \left(-\frac{2\,\langle T, R\rangle}{\|T\|_F^2} T_i + R_i\right)\left(\frac{2\,\langle T, R\rangle}{\|T\|_F^2} T_j - R_j\right)\right] \tag{3.35}$$

$$= \frac{2}{\|R\|_F^2 \|T\|_F^2}\left[\frac{\langle T, R\rangle^2}{\|T\|_F^2}\delta_{ij} - \left(\frac{2\,\langle T, R\rangle}{\|T\|_F^2} T_i - R_i\right)\left(\frac{2\,\langle T, R\rangle}{\|T\|_F^2} T_j - R_j\right)\right] \tag{3.36}$$

$$= \frac{2}{\|R\|_F^2 \|T\|_F^2}\frac{\langle T, R\rangle^2}{\|T\|_F^2}\left[\delta_{ij} - \left(\frac{2}{\|T\|_F} T_i - \frac{\|T\|_F}{\langle T, R\rangle} R_i\right)\left(\frac{2}{\|T\|_F} T_j - \frac{\|T\|_F}{\langle T, R\rangle} R_j\right)\right] \tag{3.37}$$

$$\stackrel{\text{def}}{=} \frac{2\,\langle T, R\rangle^2}{\|R\|_F^2 \|T\|_F^4}\left[\delta_{ij} + CD\right], \tag{3.38}$$

where

$$C = -\left(\frac{2}{\|T\|_F} T_i - \frac{\|T\|_F}{\langle T, R\rangle} R_i\right), \quad D = \left(\frac{2}{\|T\|_F} T_j - \frac{\|T\|_F}{\langle T, R\rangle} R_j\right). \tag{3.39}$$

The expression for $\frac{\partial^2 \psi}{\partial T_i \, \partial T_j}$ is long and going to be expensive to compute. Let us simplify the expression by showing that $|C| \ll 1$ and $|D| \ll 1$ and then making approximations. For all i,

$$\frac{T_i}{\|T\|_F} \leq \frac{\|T\|_\infty}{\|T\|_F} \quad \Big(\|T\|_\infty = \max_i |T_i| \Big) \quad \text{and} \quad R_i \leq \|R\|_\infty. \tag{3.40}$$

Applying the triangle inequality to C,

$$|C| \leq \left| \frac{2\, T_i}{\|T\|_F} \right| + \left| -\frac{\|T\|_F \, R_i}{\langle T, R \rangle} \right| \tag{3.41}$$

$$\leq \frac{2\, \|T\|_\infty}{\|T\|_F} + \frac{\|T\|_F \, \|R\|_\infty}{\langle T, R \rangle} \tag{3.42}$$

$$\leq \frac{2\, \|T\|_\infty}{\|T\|_F} + \frac{\|T\|_F \, \|R\|_\infty}{\|T\|_F \, \|R\|_F}. \tag{3.43}$$

Cancelling terms,

$$|C| \leq \frac{2\, \|T\|_\infty}{\|T\|_F} + \frac{\|R\|_\infty}{\|R\|_F}. \tag{3.44}$$

A similar argument applies to D.

For the images we will deal with, $\|T\| \neq 0$ and $\|R\| \neq 0$. Also, for natural images, it is a reasonable assumption that $\|T\|_\infty \ll \|T\|_F$ and $\|R\|_\infty \ll \|R\|_F$, so $|C| \ll 1$ and $|D| \ll 1$.

Approximating Equation 3.38 with $|C| \ll 1$ and $|D| \ll 1$,

$$\frac{\partial^2 \psi}{\partial T_i \, \partial T_j} \approx \frac{2 \, \langle T, R \rangle^2}{\|R\|^2 \|T\|^4} \, \delta_{ij} \tag{3.45}$$

$$\frac{\partial^2 \psi}{\partial T^2} \approx \frac{2 \, \langle T, R \rangle^2}{\|R\|^2 \|T\|^4} \, I. \tag{3.46}$$

Apply the Holder inequality

$$\|fg\|_1 \leq \|f\|_p \, \|g\|_q, \quad \text{where } \frac{1}{p} + \frac{1}{q} = 1 \quad (\text{and } p, q \in [1, \infty]) \tag{3.47}$$

to T and R, and choose $p, q = 2$. Then

$$\|TR\|_1^2 \leq \|R\|_F^2 \, \|T\|_F^2. \tag{3.48}$$

For $T, R \geq 0$, i.e. T and R contain no negative pixels, $\langle T, R \rangle = \|TR\|_1$. The inequality becomes

$$\langle T, R \rangle^2 \leq \|R\|_F^2 \, \|T\|_F^2. \tag{3.49}$$

Rearranging the expression,

$$\frac{\langle T, R \rangle^2}{\|R\|_F^2 \, \|T\|_F^2} \leq 1. \tag{3.50}$$

For $T \approx R$ (i.e. the pixel values in T are similar to the pixel values in R), the value for the normalized cross correlation between T and R will be close to (but never greater than) 1:

$$\frac{\langle T, R \rangle^2}{\|T\|_F^2 \|R\|_F^2} \lesssim 1. \tag{3.51}$$

Multiplying both sides by $\frac{2}{\|T\|_F^2}$, we obtain an upper bound on part of the right side of the expression in 3.46

$$\frac{2 \langle T, R \rangle^2}{\|R\|_F^2 \|T\|_F^4} \leq \frac{2}{\|T\|_F^2}, \tag{3.52}$$

and

$$\frac{\partial^2 \psi}{\partial T^2} \lesssim \frac{2}{\|T\|_F^2} I. \tag{3.53}$$

To simplify and speed up computation, instead of computing the entries of $\frac{\partial^2 \psi}{\partial T^2}$ individually as given in the expression in 3.37, we showed that the second derivative was approximately equal to the expression in 3.46. Additionally, under the assumption that the the images to be registered will appear similar ($T \approx R$), 3.46 was further simplified to 3.53. The right hand side of 3.53 is much less expensive to compute compared to the original expression in 3.46 and can be used to approximate $\frac{\partial^2 \psi}{\partial T^2}$. This becomes significant during the optimization process, where derivatives are computed at every iteration.

3.5 Derivatives of the 2D-2D Objective Function

Let us set aside discussion of regularizers for now to explore the rest of the steps in the image registration process.

3.5.1 Non-parametric Model

For a general non-parametrized objective function to be minimized over $\mathtt{y}$, the derivative is with respect to $\mathtt{y}$:

$$\frac{\partial J}{\partial \mathtt{y}} = \frac{\partial \psi}{\partial \mathtt{y}} + \frac{\partial S}{\partial \mathtt{y}} \tag{3.54}$$

$$= \frac{\partial \psi}{\partial T}\frac{\partial T}{\partial \mathtt{y}} + \frac{\partial S}{\partial \mathtt{y}} \tag{3.55}$$

by the chain rule.

$\frac{\partial \psi}{\partial T}$ depends on the choice of dissimilarity measure and has been derived for various dissimilarity measures in Section 3.4. $\frac{\partial T}{\partial y}$ depends on the input image T and is calculated via finite differences. The derivative of the regularizer S, $\frac{\partial S}{\partial y}$, depends on how S is defined.

3.5.2 Parametric Model

If $\mathtt{y}$ is parametrized by w, then J will be minimized over w, and $\frac{\partial J}{\partial w}$ is required instead of $\frac{\partial J}{\partial \mathtt{y}}$. By the chain rule again,

$$\frac{\partial J}{\partial w} = \frac{\partial \psi}{\partial T}\frac{\partial T}{\partial \mathtt{y}}\frac{\partial \mathtt{y}}{\partial w} + \frac{\partial S}{\partial w}. \tag{3.56}$$

Except for the addition of $\frac{\partial \mathtt{y}}{\partial w}$, the rest of the terms are identical to the non-parametrized case. $\frac{\partial \mathtt{y}}{\partial w}$ can easily be obtained as a function by analytically computing the derivative of the transformation $\mathtt{y}$, which is known.

3.6 Introduction to 2D-3D Registration

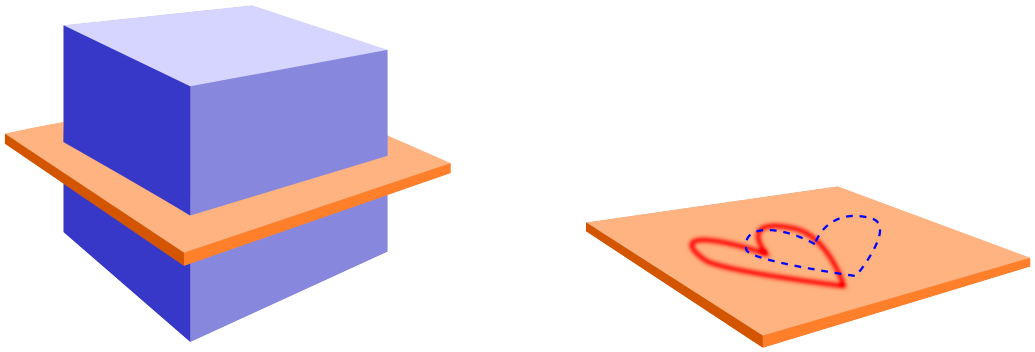
In registering two 2D images, an assumption is made about the objects represented by the images that they both physically reside within the same plane. When $\mathcal{T}$ and $\mathcal{R}$ are 2D slice realizations of physical, 3D objects, such as MR images of the heart, 2D-2D registration does not account for movements perpendicular to the plane of the image. In clinical applications, the assumption does not hold in general because 2D-2D registration

does not reflect the physical situation well, due to respiratory motion and as the patient will be in different positions between pre-operative and intra-operative scans, neither of which are limited to in-plane motions.

Although 2D-2D registration may return satisfactory transformation parameters where motion perpendicular to the plane is estimated to be small, results may be non-physical and difficult to evaluate without information in the third dimension.

In clinical applications of real-time MRI, a 2D image is taken as a slice of a 3D object. In imaging short-axis slices of the heart, for example, a 2D cross section is taken of the chest cavity. We want to simulate this in forming our 2D-3D registration model.

In this 2D-3D registration model, the template image remains 3D and the reference image is 2D. At each iteration during registration, transformations are made on a 3D image before a slice of it is compared to the 2D reference image. If registration is successful, a transformation on the 3D template image is found such that the slice of it is optimally aligned with the 2D reference image. Figures 3.2 and 3.3 illustrate.

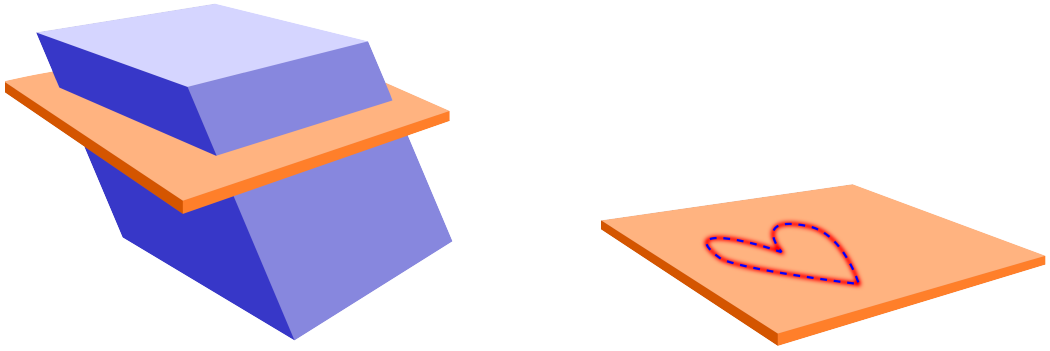


(a) Reference image (orange) and initial template image (blue).

(b) The 2D projections of the heart are initially misaligned.

Figure 3.2: Before registration

In 2D-2D registration, the template and reference images were both 2D, but in 2D-3D registration, the template image $\mathcal{T} : \Omega \times \mathcal{Z} \subset \mathbb{R}^3 \to \mathbb{R}$ is defined over the 3D domain $\Omega \times \mathcal{Z}$ while the reference image $\mathcal{R} : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ is defined over the 2D domain Ω, both



(a) Reference image (orange) and transformed template image (blue).

(b) The 2D projections of the heart are aligned.

Figure 3.3: After 2D-3D registration

compactly supported.

3.7 2D-3D Objective Function

The goal is to find a transformation $y : \mathbb{R}^3 \to \mathbb{R}^3$ on the set of 3D coordinates $x \in \Omega \times \mathcal{Z} \subset \mathbb{R}^3$ such that a slice of the template image $\mathcal{T}$ taken at a location z is similar to the reference image $\mathcal{R}$. The objective function to be minimized is

$$\mathcal{J}[y] := \mathcal{D}[\mathcal{L}_z(\mathcal{T}(y(x))), \mathcal{R}(x)] + \mathcal{S}[y(x)]. \tag{3.57}$$

Here, $\mathcal{D}$ is a measure of dissimilarity between two 2D images $\mathcal{L}_z(\mathcal{T}(y(x)))$ and $\mathcal{R}$ in the objective function above. $\mathcal{S}$ is a regularizer that imposes restrictions on the transformation y.

One important piece that is not needed in 2D-2D or 3D-3D registration is the slicing operator $\mathcal{L}_z : \mathbb{L}^2(\Omega \times \mathcal{Z}) \to \mathbb{L}^2(\Omega)$ that extracts a 2D slice from a 3D volume. The operator $\mathcal{D}$ can only take arguments of the same dimension. $\mathcal{T}$ and $\mathcal{R}$ are of different dimensionalities, so the slicer reduces $\mathcal{T}$ to the same dimension as $\mathcal{R}$. The slicer $\mathcal{L}_z$ extracts a slice at location $z \in \mathcal{Z} \subset \mathbb{R}$ such that $\mathcal{L}_z(\mathcal{T}(x^1, x^2, x^3)) := \mathcal{T}(x^1, x^2, z)$ for $(x^1, x^2, x^3) \in \mathbb{R}^3$.

3.7.1 2D-3D Parametrized Objective Function

As explained earlier, the transformation y can be parametrized via parameters w. Note that y is a transformation on the template image, and therefore acts on 3D coordinates, not 2D coordinates. The task is now to search for parameters w that minimize the objective function

$$\mathcal{J}[w] := \mathcal{D}[\mathcal{L}_z(\mathcal{T}(y(w,x))), \mathcal{R}(x)] + \mathcal{S}(w). \tag{3.58}$$

3.7.2 Multiple Solutions - 3D Affine Transformations

The objective function above applies in general for any parametric transformation on $\mathcal{T}$. For certain parameterizations, however, some parameters can be redundant due to the reduction from 3 dimensions to 2 in taking a slice of the 3D template.

Here, we investigate the question: If $y(w,x)$ is a 3D affine transformation (see Section 2.2), can different sets of parameters w yield the same $\mathcal{L}_z(\mathcal{T}[y(w,x)])$?

Theorem 1. *Consider a given z. Any two affine transformations w^A and w^B that satisfy the following conditions yield $\mathcal{L}_z(\mathcal{T}[y(w^A;x)])$=$\mathcal{L}_z(\mathcal{T}[y(w^B;x)])$.*

$$\begin{pmatrix} w_1^A \\ w_5^A \\ w_9^A \end{pmatrix} = \begin{pmatrix} w_1^B \\ w_5^B \\ w_9^B \end{pmatrix}, \begin{pmatrix} w_2^A \\ w_6^A \\ w_{10}^A \end{pmatrix} = \begin{pmatrix} w_2^B \\ w_6^B \\ w_{10}^B \end{pmatrix} \quad \textit{and} \quad \begin{pmatrix} w_3^A - w_3^B \\ w_7^A - w_7^B \\ w_{11}^A - w_{11}^B \end{pmatrix} z + \begin{pmatrix} w_4^A - w_4^B \\ w_8^A - w_8^B \\ w_{12}^A - w_{12}^B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Proof. Note that for any given z and w

$$\begin{aligned} \mathcal{L}_z[\mathcal{T}[y(w;x)]] &= \mathcal{L}_z \left[T \left[\begin{pmatrix} w_1 & w_2 & w_3 \\ w_5 & w_6 & w_7 \\ w_9 & w_{10} & w_{11} \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} + \begin{pmatrix} w_4 \\ w_8 \\ w_{12} \end{pmatrix} \right] \right] \\ &= T \left[\begin{pmatrix} w_1 & w_2 & w_3 \\ w_5 & w_6 & w_7 \\ w_9 & w_{10} & w_{11} \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ z \end{pmatrix} + \begin{pmatrix} w_4 \\ w_8 \\ w_{12} \end{pmatrix} \right]. \end{aligned}$$

Now consider w^A and w^B that for any x^1, x^2

$$\begin{pmatrix} w_1^A & w_2^A & w_3^A \\ w_5^A & w_6^A & w_7^A \\ w_9^A & w_{10}^A & w_{11}^A \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ z \end{pmatrix} + \begin{pmatrix} w_4^A \\ w_8^A \\ w_{12}^A \end{pmatrix} = \begin{pmatrix} w_1^B & w_2^B & w_3^B \\ w_5^B & w_6^B & w_7^B \\ w_9^B & w_{10}^B & w_{11}^B \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ z \end{pmatrix} + \begin{pmatrix} w_4^B \\ w_8^B \\ w_{12}^B \end{pmatrix}.$$

Therefore for any x^1, x^2

$$\begin{pmatrix} w_1^A - w_1^B \\ w_5^A - w_1^B \\ w_9^A - w_1^B \end{pmatrix} x^1 + \begin{pmatrix} w_2^A - w_2^B \\ w_6^A - w_6^B \\ w_{10}^A - w_{10}^B \end{pmatrix} x^2 = - \left[\begin{pmatrix} w_3^A - w_3^B \\ w_7^A - w_7^B \\ w_{11}^A - w_{11}^B \end{pmatrix} z + \begin{pmatrix} w_4^A - w_4^B \\ w_8^A - w_8^B \\ w_{12}^A - w_{12}^B \end{pmatrix} \right].$$

Equating the right-hand-side and left-hand-side to zero completes the proof. □

This suggests that if no restrictions are imposed on w, the first two columns of w^A and w^B have to match. In addition, for any given third columns of w^A and w^B, a given z, and a given fourth column of w^A, we can always compute the fourth column of w^B that yields the same sliced result. This suggests that the number of parameters of w can be reduced to 9 instead of 12 to obtain a unique solution.

Intuitively, this makes sense because w_3 and w_7 correspond to shearing along an xy-plane, parallel to the slice taken by $\mathcal{L}_z$, and w_{11} corresponds to scaling in the z-direction, normal to the slice.

3.7.3 Discretized Objective Function

As stated earlier, to optimize the problem numerically, the problem first needs to be discretized. This contrasts with the optimize-then-discretize paradigm, where the functional is first optimized analytically before the rest of the calculations are performed numerically for specific input images.

Discretizing Ω into $m \times n$ pixels and $\mathcal{Z}$ into l pixels, we can define $\mathtt{x}_R$ and $\mathtt{x}_T$ relating to $\mathcal{R}$ and $\mathcal{T}$, respectively, to be the discretizations of Ω and $\Omega \times \mathcal{Z}$. The discretized, transformed coordinates are $\mathtt{y} \approx y(\mathtt{w}, \mathtt{x}_T)$ (where $\mathtt{w} = w$), the discretized images are

$T \approx \mathcal{T}(\mathtt{x}_T)$ and $R \approx \mathcal{R}(\mathtt{x}_R)$, and discretization of the operators $\mathcal{D}$ and $\mathcal{S}$ are represented by D and S.

Assuming the elements of T are laid out as a column vector in column-major ordering, for a given z, the discretization of the operator $\mathcal{L}_z$, denoted by L_z, can be computed as

$$L_z = I_{mn\times mnl} := I_{mn\times mn} \otimes \overbrace{[0,\ldots,0,\underbrace{1}_{\lceil mnl(z+\omega)/\omega\rceil\text{-th component}},0,\ldots,0]}^{1\times mnl \text{ size}}, \tag{3.59}$$

where $\mathcal{Z} \in (-\frac{\omega}{2}, -\frac{\omega}{2})$. Practically, to perform the operator on the 3D image in MATLAB, the 3D image is reshaped into a 3D array, then sliced using an indexing operation.

The discretized problem is now to minimize

$$J[w] := D[L_z(T(\mathtt{y}(\mathtt{w},\mathtt{x}))), R(\mathtt{x})] + S(w). \tag{3.60}$$

3.8 Dissimilarity Measures and Derivatives (2D-3D)

In the 2D-2D case or the 3D-3D case, dissimilarity measures were applied to two images that were the same dimension, i.e. T and R were both 2D or both 3D. In slice-to-volume 2D-3D registration, we are no longer interested in comparing T to R, but rather L to R, where L denotes the image $L_z(T)$, which is a slice of T taken at location z. When measuring the dissimilarity, $D[L, R]$ is computed instead of $D[T, R]$, but the definitions for each of the dissimilarity measures listed in subsection 3.4 remain the same.

3.9 Regularizers

The registration problem is not guaranteed to have a unique solution. For implementation in image-guided surgery, where a 2D image obtained in live time is registered to a 3D image obtained prior to surgery, a question that arises is - if an infinite number of transformations produce the same registered, end-result image, which transformation should be chosen?

Given information about the problem, however, we can choose a regularizer to penalize unlikely or improbable transformations. Earlier, we found that in affine 2D-3D registration, the parameters w needed to obtain identical slices $\mathcal{L}$ were not unique. Without regularizing w, the affine 2D-3D registration problem is ill-posed. To yield a unique w, we require a regularization functional $\mathcal{S}$ independent of the input images, for example,

$$\mathcal{S}[w - w^{\text{ref}}] := \frac{1}{2} \times (w - w^{\text{ref}})^T \mathbf{M} (w - w^{\text{ref}}), \tag{3.61}$$

where $\mathbf{M}$ is a symmetric positive definite weight matrix that acts as a regularizer.

The regularizer above penalizes the value of $w - w^{\text{ref}}$, so we can choose a set of transformation parameters w^{ref} we do not want w to stray too far from, chosen using what is known about the system.

In practice, since we typically have information about w^{ref}, we can simply impose regularization to restrict w instead of reducing the number of parameters in the affine transformation to yield a unique w.

3.10 Derivatives of the 2D-3D Objective Function

The derivative of the objective function with respect to the parameters over which it is minimized is needed for most optimization algorithms, and some require the second derivative as well.

For a non-parametrized objective function minimized over the coordinates $\mathtt{y}$, by the chain rule,

$$\frac{\partial J}{\partial \mathtt{y}} = \frac{\partial \psi}{\partial \mathtt{y}} + \frac{\partial S}{\partial \mathtt{y}} \tag{3.62}$$

$$= \frac{\partial \psi}{\partial L} \frac{\partial L}{\partial T} \frac{\partial T}{\partial \mathtt{y}} + \frac{\partial S}{\partial \mathtt{y}}. \tag{3.63}$$

$\frac{\partial \psi}{\partial L}$, the derivative of the dissimilarity measure with respect to a 2D image, is known from the definition of the dissimilarity measure. Since $L = L_z(T) = I_{mn \times mnl}\, T$, $\frac{\partial L}{\partial T} = I_{mn \times mnl}$.

$\frac{\partial T}{\partial \mathtt{y}}$ depends on the image T and is calculated via finite differences. $\frac{\partial S}{\partial \mathtt{y}}$ is also known from the definition of S.

If $\mathtt{y}$ is parametrized by $\mathtt{w}$, J is minimized over $\mathtt{w}$, and

$$\frac{\partial J}{\partial \mathtt{w}} = \frac{\partial \psi}{\partial L}\frac{\partial L}{\partial T}\frac{\partial T}{\partial \mathtt{y}}\frac{\partial \mathtt{y}}{\partial \mathtt{w}} + \frac{\partial S}{\partial \mathtt{y}}\frac{\partial \mathtt{y}}{\partial \mathtt{w}}. \tag{3.64}$$

$\frac{\partial \mathtt{y}}{\partial \mathtt{w}}$ can be obtained analytically by computing the derivative of the transformation $\mathtt{y}$, which is known [Mod09].

Recall that ψ can be written as $\psi(r(\mathtt{w}))$. A common approximation for the Hessian of a function $\psi(r(w))$ obtained using chain rules and Taylor series expansions is

$$H_\psi \approx \left(\frac{\partial r}{\partial \mathtt{w}}\right)^T \left(\frac{\partial^2 \psi}{\partial r^2}\right)\left(\frac{\partial r}{\partial \mathtt{w}}\right). \tag{3.65}$$

The Hessian of J, denoted by H_J is

$$H_J = H_\psi + H_S \approx \left(\frac{\partial r}{\partial \mathtt{w}}\right)^T \left(\frac{\partial^2 \psi}{\partial r^2}\right)\left(\frac{\partial r}{\partial \mathtt{w}}\right) + H_S, \tag{3.66}$$

if we assume the regularizer described previously.

Using 3D affine registration with the SSD dissimilarity measure as an example, after putting it all together, the derivatives are

$$\frac{\partial J}{\partial \mathtt{w}} = \frac{\partial \psi}{\partial r}\frac{\partial r}{\partial L}\frac{\partial L}{\partial T}\frac{\partial T}{\partial \mathtt{y}}\frac{\partial \mathtt{y}}{\partial \mathtt{w}} + \frac{\partial S}{\partial \mathtt{y}}\frac{\partial \mathtt{y}}{\partial \mathtt{w}} \tag{3.67}$$

$$= \mathtt{hd}\, r^T \times I_{mn} \times I_{mn\times mnl} \times dT \times dy + (\mathtt{w} - \mathtt{w}^{\text{ref}})^T \mathbf{M}, \tag{3.68}$$

where we have assumed the regularizer described earlier and used $dT := \frac{\partial T}{\partial \mathtt{y}}$, $dy := \frac{\partial \mathtt{y}}{\partial \mathtt{w}}$, and $\mathtt{w}$ is a 12×1 vector of parameters $w_1, w_2, \ldots, w_{12}$. The Hessian is

$$H_J \approx dr^T dr + \mathbf{M}, \tag{3.69}$$

using $dr = \frac{\partial r}{\partial w}$.

During computation, 0th-order variables (i.e. not derivatives) are ordered lexicographically into vectors. Their derivatives (Jacobians and Hessians) are matrices for obvious reasons. Table 3.1 summarizes the sizes of each variable used during computation.

Variable	Size
R, L, r	$mn \times 1$
T	$mnl \times 1$
$\mathtt{x}_R$	$2mn \times 1$
$\mathtt{x}_T, \mathtt{y}$	$3mnl \times 1$
$\mathtt{w}, \mathtt{w}^{\text{ref}}$	$p \times 1$
S, J, ψ	1×1
dT	$mnl \times 3mnl$
$d\mathtt{y}$	$3mnl \times p$
dr	$mn \times p$
dJ, dS	$1 \times p$
$H_J, H_S, \mathbf{M}$	$p \times p$

Table 3.1: Sizes of discrete variables during computation. 2D or 3D arrays are laid out lexicographically. p is the number of parameters in the parametric tranformation.

3.11 Optimizers

The last step is to numerically optimize the objective function.

The gradient descent and Gauss-Newton methods (Algorithms 1 and 2) are two common algorithms for minimizing a function. They are straightforward and simple to use, but the gradient descent method can be quite inefficient and the Gauss-Newton method requires second derivatives. For problems where the Hessian is unavailable or too expensive to compute, other optimization schemes, such as Quasi-Newton methods, can be chosen that do not require the Hessian [NW06].

Algorithm 1 Minimizing $J[\mathtt{w}]$ using the Gradient (Steepest) Descent Approach

Initialize $\begin{bmatrix} \mathtt{w} \end{bmatrix} \leftarrow \begin{bmatrix} \mathtt{w}_0 \end{bmatrix}$.

while not converged **do**

Evaluate dJ at $[\mathtt{w}]$.

Solve the descent direction from the linear equation $\delta\mathtt{w} = -dJ^T$.

Find a positive scalar step-size $\mathtt{s}$ using a line-search scheme.

Update $\begin{bmatrix} \mathtt{w} \end{bmatrix} \leftarrow \begin{bmatrix} \mathtt{w} \end{bmatrix} + \mathtt{s} \begin{bmatrix} \delta\mathtt{w} \end{bmatrix}$.

end while

Algorithm 2 Minimizing $J[\mathtt{w}]$ using the Gauss-Newton Approach

Initialize $\begin{bmatrix} \mathtt{w} \end{bmatrix} \leftarrow \begin{bmatrix} \mathtt{w}_0 \end{bmatrix}$.

while not converged **do**

Evaluate H_J and dJ at $[\mathtt{w}]$.

Solve the descent direction from the linear equation $H_J \begin{bmatrix} \delta\mathtt{w} \end{bmatrix} = -dJ^T$.

Find a positive scalar step-size $\mathtt{s}$ using a line-search scheme.

Update $\begin{bmatrix} \mathtt{w} \end{bmatrix} \leftarrow \begin{bmatrix} \mathtt{w} \end{bmatrix} + \mathtt{s} \begin{bmatrix} \delta\mathtt{w} \end{bmatrix}$.

end while

3.12 Interpolation

Since the registration problem is a discretized one, interpolation is needed to compute the discrete transformed coordinates and the transformed template image.

$T(\mathtt{x})$ is well-defined, but for a given $\mathtt{w}$, $\mathtt{y}(\mathtt{w},\mathtt{x})$ likely will not coincide with gridpoints of $\mathtt{x}$. Thus, $T(\mathtt{y}(\mathtt{w},\mathtt{x}))$ is technically not defined. An interpolation step is necessary to compute the transformed template image over the discrete domain. $T(\mathtt{y}(\mathtt{w},\mathtt{x}))$ was used previously with the implication of an interpolation step, so technically, we were referring $T(\mathtt{P}\cdot\mathtt{y}(\mathtt{w},\mathtt{x}))$ where P is an interpolant.

Since P depends on y and changes with every iteration of the optimizer, the registration problem is a highly non-linear one and a closed-form solution can not be found. For this reason we must numerically optimize the problem to find a solution.

For simplicity, a linear interpolator will be used for registration experiments presented in this thesis.

3.13 Multi-Level Framework

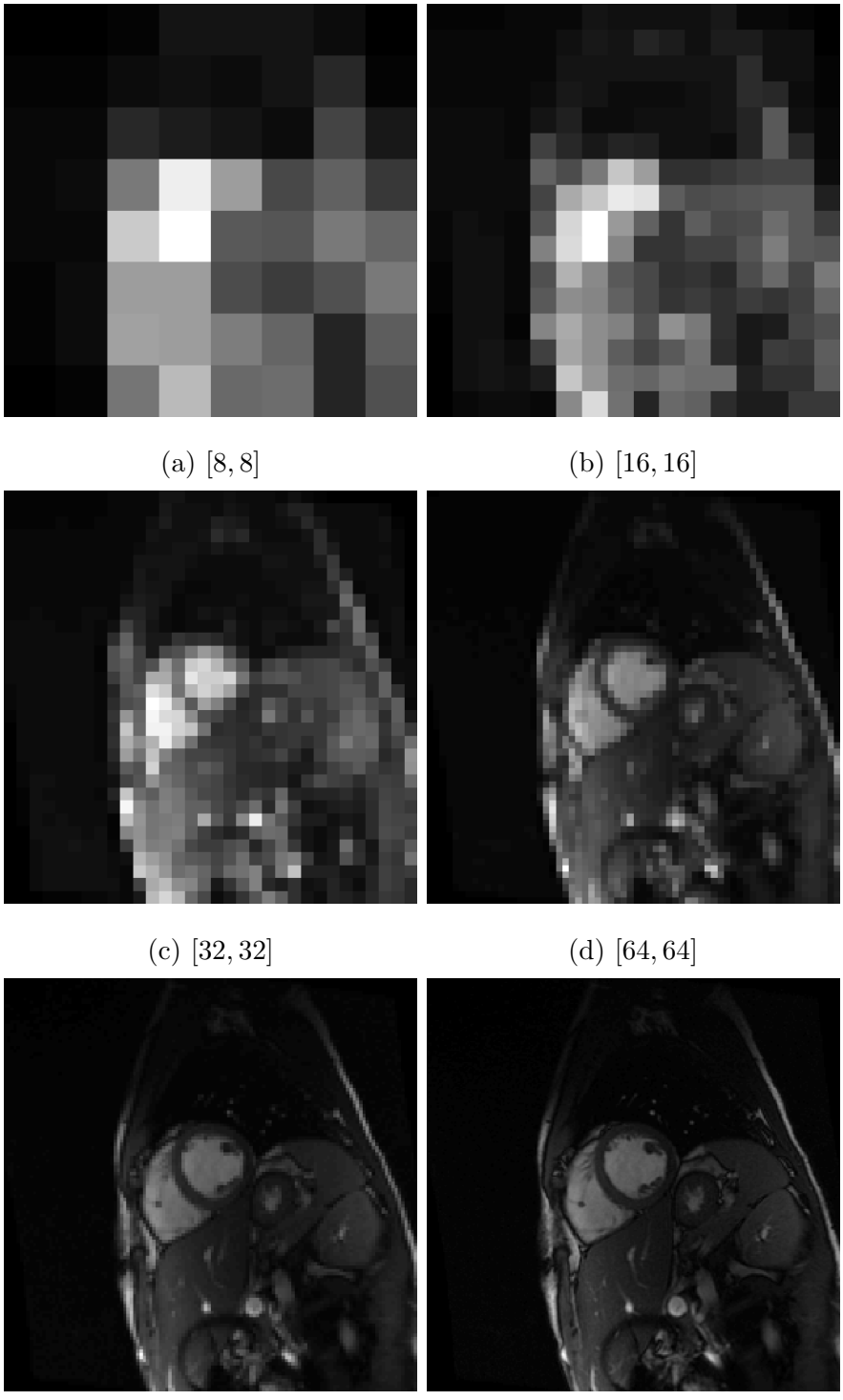
It is apparent that the minimization problem depends on how the problem is discretized. Solving the coarsely discretized problem is computationally inexpensive, but at the expense of large errors as information is lost. On the other hand, a fine-grained discretization of the problem may produce more precise results, but may also more easily be trapped in local minima.

Let us consider different discrete representations of the image registration problem and address the discrete problems sequentially in the so-called multi-level approach. Starting with the coarsest level, a solution is computed, which then serves as a starting point for the next finer discretization.

There are several advantages to this. It is efficient since most of the work is done at the computationally inexpensive coarse representations, and only small adjustments

are required at the more costly fine representations. Small variations in intensity present in the full-resolution image do not appear in coarse discretizations, so the optimizer is less likely to be trapped in local minima. It also adds additional regularization by giving more weight to more important, large-scale structure.

An objective function is optimized at each discretization level, preserving the optimization character of the problem and allowing the use of established schemes for line searches and stopping.



(a) $[8, 8]$ (b) $[16, 16]$

(c) $[32, 32]$ (d) $[64, 64]$

(e) $[128, 128]$ (f) $[256, 256]$

Figure 3.4: An image at coarse to fine levels of discretization.

Chapter 4

Experiments and Results

4.1 Data

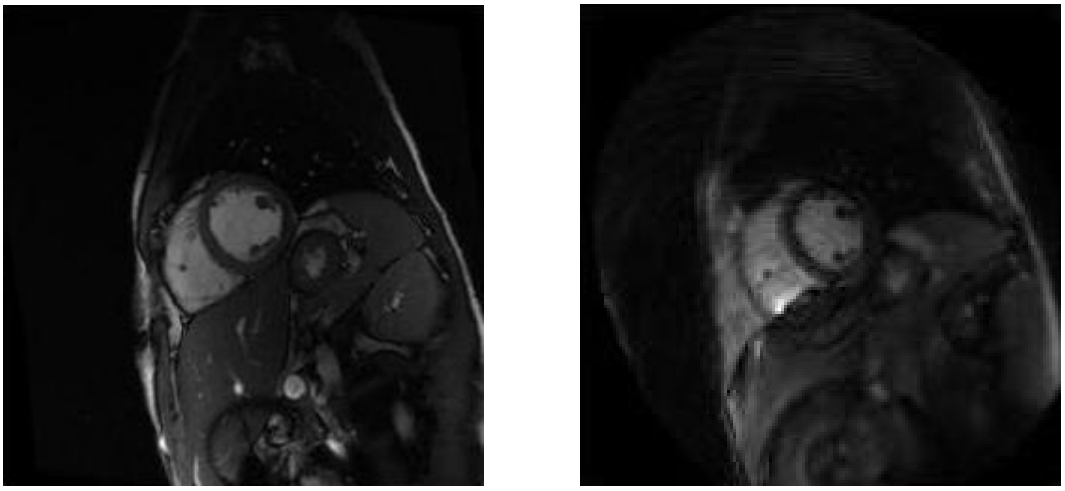
3D pre-procedural and 2D real-time cardiac MRI were acquired from 6 volunteers using a 1.5T MRI scanner (GE Signa Excite, GE Healthcare, Waukesha, WI).

4.1.1 3D Prior (Cine) Images:

Each pre-procedural 3D volume consists of a stack of 12 to 14 short-axis (SAX) slices of the heart with a resolution of $1.37 \times 1.37 \times 8$ mm^3 and a field of view (FOV) of 350×350 mm^2 (i.e. 256×256 pixels) for each slice. The images were acquired at end-expiration breath-hold with an electrocardiogram (ECG) triggered GE FIESTA pulse sequence.

ECG triggering works by measuring the amount of time that has passed since the last R-peak of the cardiac cycle, called the trigger time, and capturing images at specified trigger times. Images in this dataset were acquired for 20 phases – 20 trigger times equally spaced in time.

It should be noted that these 3D images are not true 3D volumes but, rather, conventional, clinically-used multi-slice acquisitions. Instead of acquiring an entire 3D volume at



(a) A high-resolution 'cine' MR image. (b) A noisy, 'real-time' MR image.

Figure 4.1: Examples of pre-operative and intra-operative short axis MR images of the heart. If the long axis runs from the apex of the heart to the base, a short axis slice images a cross-section of the heart perpendicular to the long axis. The light-coloured circular region near the centre of each image is the left ventricle (LV) and the light-coloured region immediately to its left is the right ventricle (RV).

a single temporal point, a 3D volume is formed by acquiring multiple 2D slices through multiple cardiac cycles, and then piecing together slices acquired at the same cardiac phase. More details can be found in [PGR15].

An example of a cine image is shown in Figure 4.1a.

4.1.2 2D Real-Time Images:

2D real-time images were aquired at the same slice locations as in the pre-procedural scans, but under free-breathing conditions. The images were obtained with a fast spiral balanced steady state free precession sequence with an in-plane resolution of 2.2×2.2 mm^2, slice thickness 8 mm, and a FOV of 350×350 mm^2 (i.e. 158×158 pixels).

Image capture was not triggered as in the cine images. Instead, images at each slice

location were continuously acquired at a frame rate of 8 frames per second (fps) and ECG-gated to record the time passed since the last R-peak. There is no synchronization between different slices, either, so image stacks do not produce meaningful volumes.

An example of a real-time image is shown in Figure 4.1b.

4.2 Validation of Results

If registration is successful, a slice obtained from transforming the 3D template image with the transformation parameters obtained from registration and then slicing at a predetermined slice location would yield a 2D image similar to the 2D reference image. In general, registration of two images taken at different times is an ill-posed probem. While 'eyeballing' the end-result images can give us a subjective impression of whether registration was successful, no 'ground truth' is available and we therefore have no way of quantitatively measuring the errors on the transformation parameters.

We can, however, validate the end-result images for its purpose in application. The images in question are cardiac MRI, where the region of interest is the left ventricle (LV). The LV is fairly circular, so landmarks within the LV, such as the papillary muscles, can act as reference points.

One way of measuring how well images have been aligned by registration is to measure how much the LVs in the template and reference images overlap before and after registration. The Dice coefficient and Jaccard index (also known as the Jaccard similarity coefficient) are two ways of quantifying the overlap between two regions and are defined as:

$$\text{Jaccard}(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad \text{and} \quad \text{Dice}(A, B) = \frac{2|A \cap B|}{|A| + |B|}. \tag{4.1}$$

To compute these values for our data, the LV is segmented to produce a binary map. Each pair of vertical bars in the definitions above indicate a count of the number of pixels satisfying the criteria between the 2 bars. Both functions return values of 0 when regions

A and B do not overlap, and return values of 1 when A and B overlap perfectly.

We can also measure how well landmarks in the LV align by measuring how far landmarks in the template image are from the landmarks in the reference image, before and after registration, assuming the spatial location of the image slices are known. This quantity, called the target registration error (TRE), is the l^2-normed distance between landmarks in the template image and the corresponding landmarks in the reference image.

Landmark locations are known for un-transformed template images and real-time reference images, so the TRE before registration is computed by $|x_T - x_R|_2$, where x_T and x_R refer to the coordinates of the landmark locations in the template and reference images, respectively. To compute the location of the same points in the transformed template images, the inverse transformation is applied to the coordinates in the un-transformed images, so the location of the transformed coordinates after registration is $y^{-1}(x_T)$ and the TRE after registration is computed by $|y^{-1}(x_T) - x_R|_2$. In controlled test experiments where the reference image is a sliced of a transformed version of the template image, the transformation y_{init} is known, so $x_R = y_{\text{init}}^{-1}(x_T)$, and the TRE before registration is $|x_T - y_{\text{init}}^{-1}(x_T)|_2$.

The LV and landmarks in the LV were manually segmented and pinpointed in the cine image volume. The endocardium of the LV was outlined for each slice, and the in-plane segmentations stacked to form a 3D segmentation mask. To obtain a 2D segmentation mask of L after registration, the 3D segmentation mask is transformed using the parameters obtained from registration, and then sliced at location z.

For the real-time images, the LV and landmarks were also segmented, by an expert. The coordinates for the landmarks in the image are 2D, but knowing the location where the slice was taken from allows us to append an approximate third coordinate to the landmarks.

4.3 Cine/cine Controlled Experiments – Rigid Initial Transformation

We set out to develop a mathematical model for 2D-3D registration and to test the model on cardiac MR images, with applications to image-guided procedures.

We were particularly interested in using an affine parametric model in registering images of the heart. Because the heart is a highly mobile organ, we hypothesized that an affine model better accounts for the deformable nature of the heart than a rigid model. The rigid model is often employed for its simplicity and speed, as opposed to an entirely deformable model, but an affine model only doubles the number of parameters from 6 to 12 and should not consume much more resources than the rigid model.

Before testing the affine model on registration between a cine image and a real-time image, we first perform controlled experiments to gauge how well the model works. In these controlled experiments, the 2D reference image was taken to be a slice of the 3D cine volume that was transformed with known transformation parameters. To maintain a fair comparison between rigid and affine registration, the initial transformation applied to the 3D cine volume to obtain the reference image was a rigid one, with 6 parameters. Registration was then performed starting with the original 3D cine volume as the template image. If registration is successful, the parameters returned should account for the initial transformation and the resultant images should match up perfectly. Rigid registration should return transformation parameters identical or close to the known initial transformation parameters. Due to Theorem 1, it is possible for affine registration to return parameters that do not correspond to the initial transformation.

For affine registration, let us choose a positive definite 12×12 diagonal regularizer $\mathbf{M}$ with unit entries on the main diagonal except for locations 3, 7, and 11 where entries are 10^6, i.e. large. By restricting the three terms to values close to the corresponding entries of $\mathtt{w}^{\text{ref}}$, the regularizer practically limits the number of parameters in the affine transfor-

mation to 9 terms, returning a unique set of transformation parameters. For example, if $\mathtt{w}^{\text{ref}}$ is chosen to be the identity transformation, the regularizer ensures parameters $[\mathtt{w}_3, \mathtt{w}_7, \mathtt{w}_{11}]$ returned from registration are close to $[0, 0, 1]$.

We perform these experiments because they have some form of a 'ground truth' available. Because the initial transformation made to obtain R is known, we know what the parameters returned by rigid registration should be. Also, because the reference image was taken from the same source as the template image, the difference between L and R after registration should be perfectly aligned and difference between corresponding pixels close to zero. Visually, a perfectly aligned image after registration corresponds to a flat difference image – the image obtained by subtracting the R from the transformed L. One way to objectively measure this is to measure the variance of the pixel values in the difference image. Another way is to evaluate the Frobenius norm of the difference image, summing up the squared pixel value differences in the entire image. However, since the Frobenius norm depends on image content (image pixel values), there is no absolute threshold under which the norm of the difference image is considered small or over which the norm is considered large. One can consider the relative change in the norm of the difference image before and after registration in deciding whether the norm after registration is considered 'close' to zero, but the problem of choosing a threshold remains – what relative change is considered 'good enough'? And how reliably and how well does a reduction of the norm indicate that images are being aligned?

In clinical situations, where the relation between the reference and template image will not be precisely known (for example, the subject has moved) or where the image acquisition differs between the reference and the template image, none of the methods outlined in the previous paragraph reliably evaluate how well the images have aligned. What we will be looking for in the registration of two images obtained in a clinical setting (Section 3.3b) is how well the features represented in the images have aligned, by measuring LV overlap (quantified by the Jaccard and Dice indices) and by measuring

the distance between landmarks in the template and reference image (quantified by the TRE). For consistency, the same methods used in evaluating the results of registration on clinical data will be used in control experiments as well.

For all experiments following, it will be assumed that $\Omega = (-175, 175) \times (-175, 175)$ mm^2, $\mathcal{Z} = (-48, 48)$ mm, $m = n = 128$, and $l = 12$. The Gauss-Newton algorithm seemed sufficient in registering the images in our data set and more robust than the Gradient Descent algorithm, so we will use the Gauss-Newton approach with an Armijo line search scheme [NW06].

Figures 4.2 and 4.3 show an example of a control experiment where reference image was a slice of the rigid-transformed 3D cine volume with arbitrarily chosen parameters (except the last entry) $\mathtt{w} = [-0.1, 0.05, \pi/16, 50, -5, 32]$. Slicing operations were applied at $z = -36$ mm. The last entry of $\mathtt{w}$, the z-direction component of translation, was chosen so that landmarks would be reasonably close to the location of the slicing operator after successful registration and the papillary muscles would be visible in the end-result images. It would also reduce z-direction (normal to the plane of an image slice) uncertainty so the TRE would be a reasonable indicator of whether registration was successful. Both rigid and affine registration performed well in this case. Values for the Jaccard, Dice, and TRE measures before and after rigid and affine registration are shown in Table 4.1.

Multiple controlled cine-cine registration were performed on pairs of images where the reference image was a slice of the transformed template image, as described above. For a range of 1715 different initial transformations performed to obtain the reference image, rigid and affine registrations were performed on 6 data sets to compare their differences. The 1715 initial transformations consisted of combinations of 7 translations ranging from -24mm to 24mm in all three Cartesian directions and 5 rotations along the z-axis between $-\frac{\pi}{16}$ and $\frac{\pi}{16}$.

For each of the 6 data sets available, affine and rigid registration was performed for the 1715 reference images obtained from various initial parameters described above.

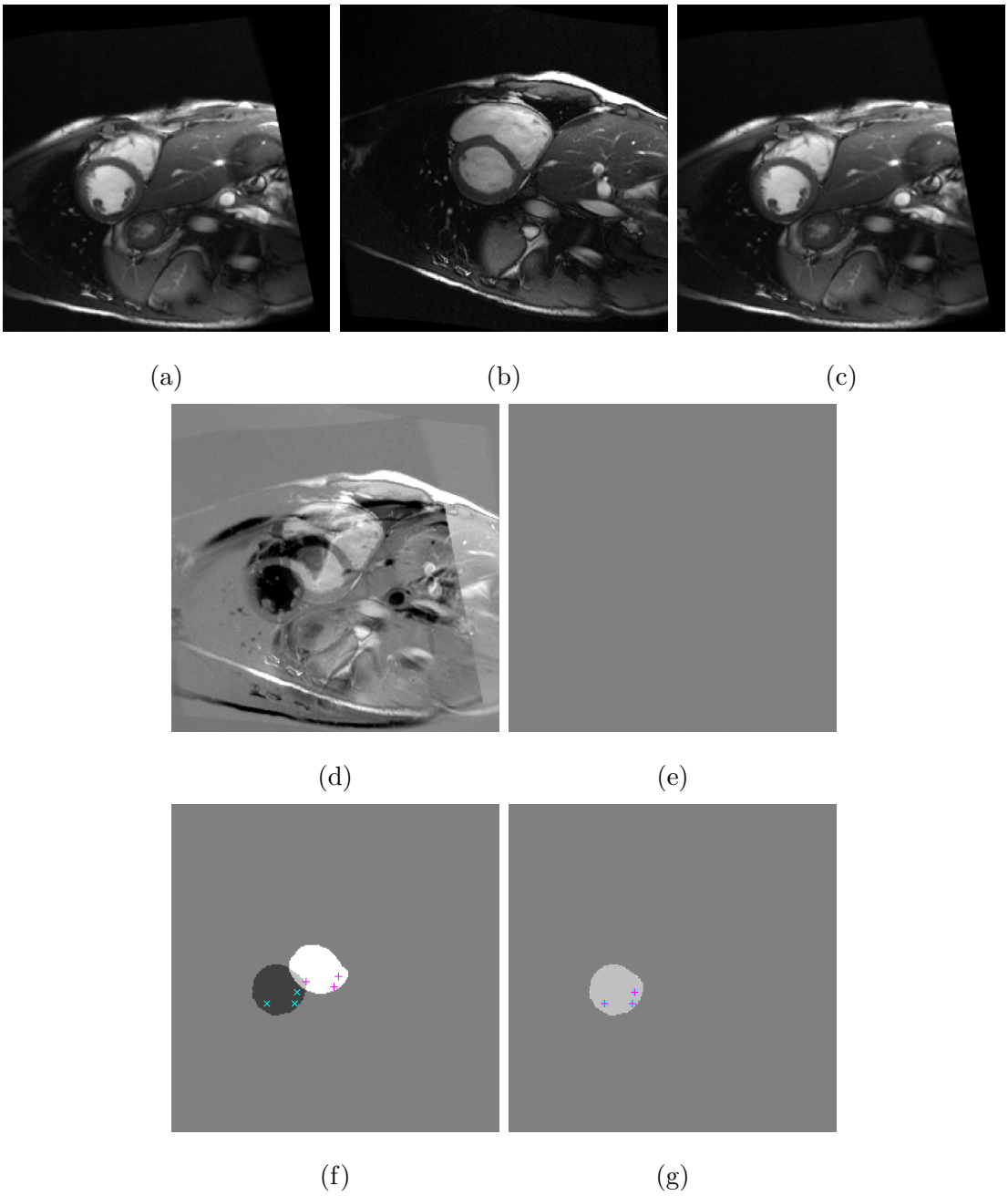


Figure 4.2: Results of rigid registration between a 3D image and a 2D cine image in a controlled experiment. (a) Reference image R. (b),(c) Template slice L before and after registration. (d),(e) Difference between the reference image and template slice $(L - R)$ before and after registration. (f),(g) Segmentation masks showing left ventricle overlap before and after registration, with out-of-plane reference image landmarks projected onto image ($\times$) and in-plane template image landmarks ($+$).

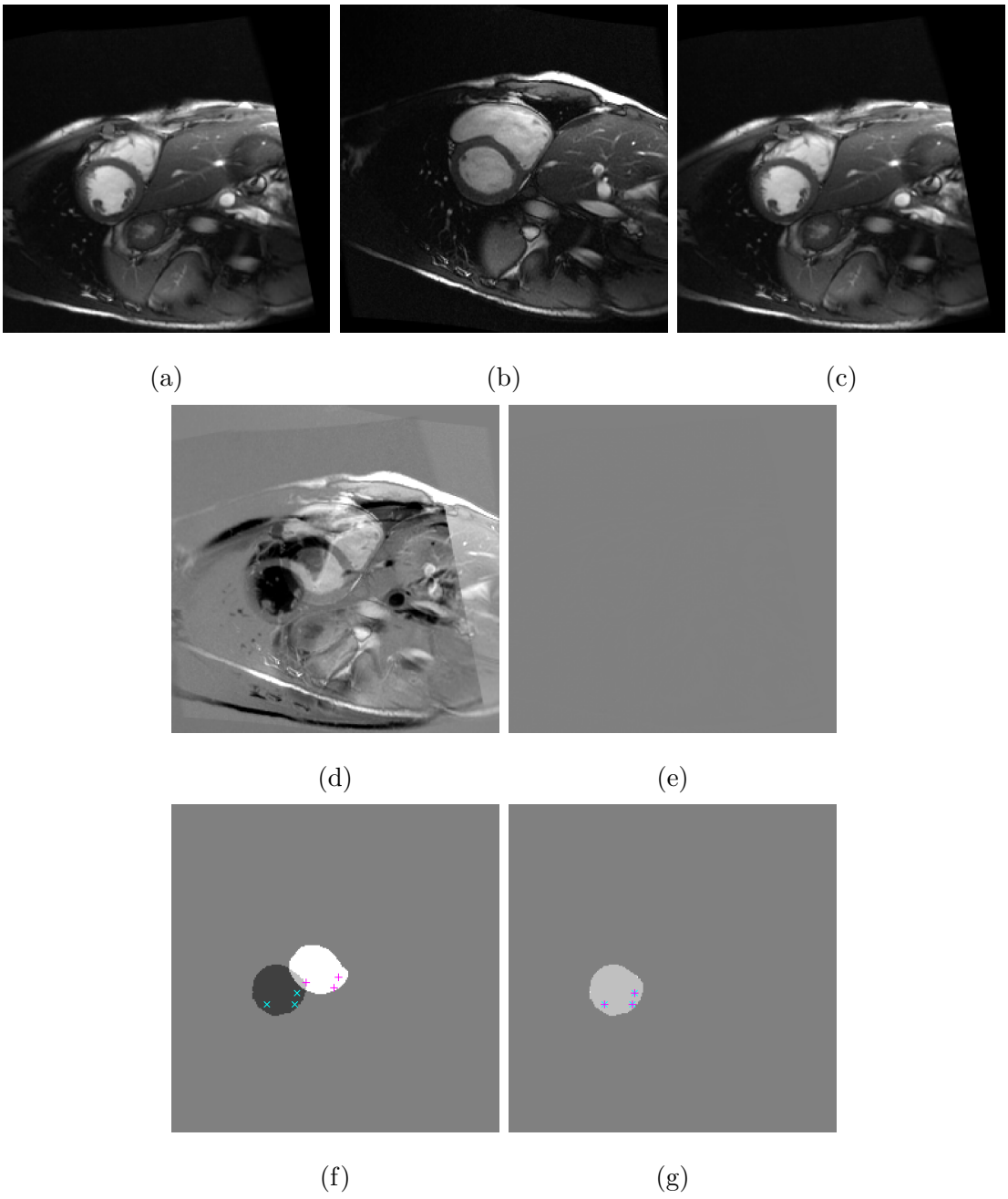


Figure 4.3: Results of affine registration between a 3D image and a 2D cine image in a controlled experiment.

(a) Reference image R. (b),(c) Template slice L before and after registration. (d),(e) Difference between the reference image and template slice $(L - R)$ before and after registration. (f),(g) Segmentation masks showing left ventricle overlap before and after registration, with out-of-plane reference image landmarks projected onto image ($\times$) and in-plane template image landmarks ($+$).

	Jaccard	Dice	TRE (mm)
Before registration	0.03	0.06	57.1 ± 1.8
After rigid registration	1.00	1.00	0.02 ± 0.01
After affine registration	1.00	1.00	0.10 ± 0.08

Table 4.1: Jaccard indices and Dice coefficients of left ventricle overlap before and after registration in a controlled experiment where the template image is a 3D cine volume and the reference image is a slice of the same cine volume transformed (rigid-body) with known parameters.

Registration was performed on a 64-bit Lenovo S30 ThinkStation with a 3.7 GHz Intel Xeon E5-1620 v2 processor. The average (of 1715) parametric registration time for a single rigid registration ranged from 36 seconds to 55 seconds between 6 data sets. For affine registration, it ranged between 49 seconds to 71 seconds, .

We found that although the affine model registered images correctly in more registration experiments than the rigid model, there were also cases where the rigid model registered the images successfully while the affine model did not. However, those tended to be isolated cases, where if the the initial transformation parameters were adjusted slightly, changing the initial misalignment between the template and the reference images by a little, rigid registration would fail. It was observed that, when each of the initial transformation parameters in w_{init} were incremented across a range of values, the affine model tended to perform more consistently compared to the rigid model and register images successfully for a broader range of initial misalignment.

Figures 4.4 to 4.9 show results of all 1715 registrations performed on Data Set 1.

Figure 4.4 shows the amount of LV overlap before registration, plotted against the initial translation distance – the total Euclidian (l^2-normed) distance the reference image R is in relation to the initial template image T. These include instances in which the

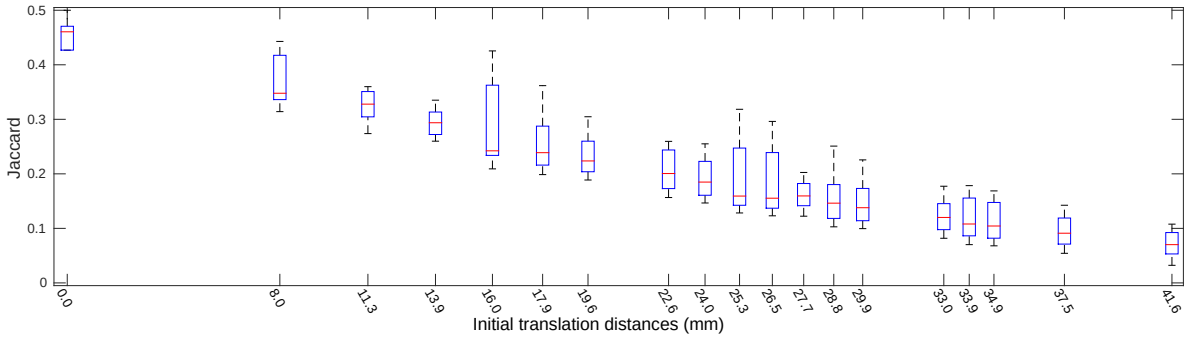


Figure 4.4: LV overlap as a function of initial translation distance before registration to a reference image obtained by an initial rigid transformation, Data Set 1.

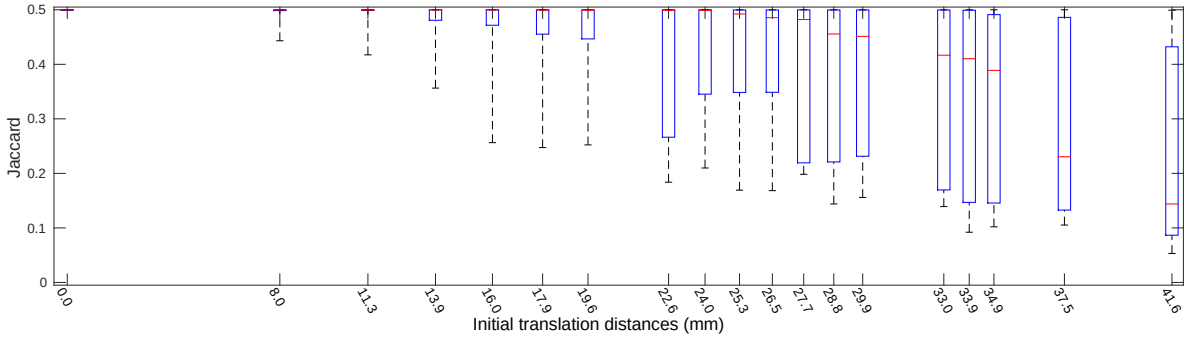


Figure 4.5: LV overlap as a function of initial translation distance after rigid registration to a reference image obtained by an initial rigid transformation, Data Set 1.

initial transformation involved a rotation as well, so an initial translation of 0 can correspond to a Jaccard index of less than 1 before registration. Whiskers mark the locations of the 10th and 90th percentiles (of 1715 registrations), the box represents the 25th-75th percentile, and the line in the box marks the median. As expected, as initial translation distances increases, alignment and therefore the amount of LV overlap decreases. Figure 4.5 and 4.6 show the amount of LV overlap after rigid and affine registration, respectively. Both rigid and affine registrations improve LV overlap, but affine registration appears to show more improvement over rigid registration for initial misalignments due to translation for a broader range of distances.

As mentioned earlier, the Jaccard index here quantifies the overlap between projec-

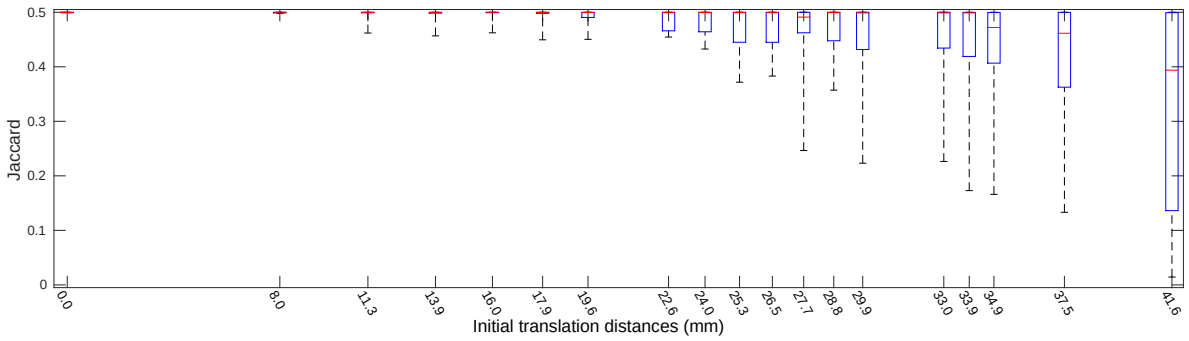


Figure 4.6: LV overlap as a function of initial translation distance after affine registration to a reference image obtained by an initial rigid transformation, Data Set 1.

tions of the LV. The Jaccard index is only a reliable indicator of good alignment if the physical slice location of the registered template slice is close to that of the reference. The TRE measures landmark distances between the template and reference images, so an indicator of good alignment would be simultaneously small TRE values and large Jaccard indices. Figures 4.7, 4.8, and 4.9 show the TRE before registration, after rigid registration, and after affine registration, respectively. As expected, before registration, the TRE varies directly with the initial translation distance. Both rigid registration and affine registration appear to improve results, but affine registration appears to reduce the TRE more compared to rigid registration.

As stated previously, each box (and its whiskers) summarizes a collection of points, each point indicating 1 of 1715 registration experiments with a different initial misalignment. For a given initial translation distance, the translations are in various directions and with varying degrees of rotation, so it is possible for a set of parameters with a translation of 0mm to produce a non-zero TRE before registration, due to rotation.

The same was done for other data sets as well; collective results are shown in Figures 4.10 to 4.15. Only the medians are shown. When initial translation distance is large, results are more varied – some TREs remain large, meaning affine and rigid registration did not improve results, which is expected because the algorithm may not find a solution

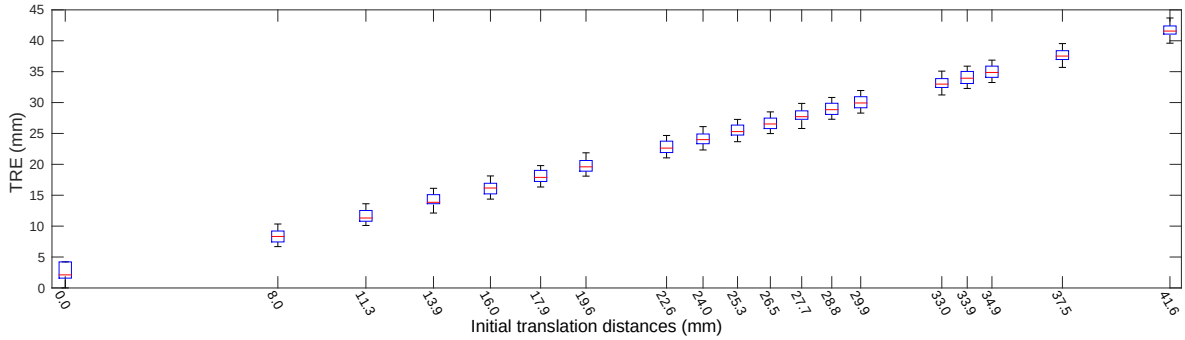


Figure 4.7: Target registration error (TRE) as a function of initial translation distance before registration to a reference image obtained by an initial rigid transformation, Data Set 1.

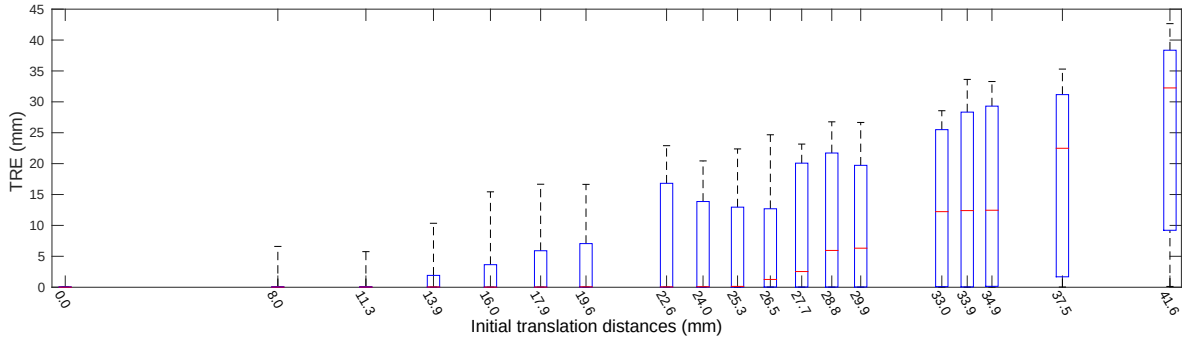


Figure 4.8: Target registration error (TRE) as a function of initial translation distance after rigid registration to a reference image obtained by an initial rigid transformation, Data Set 1.

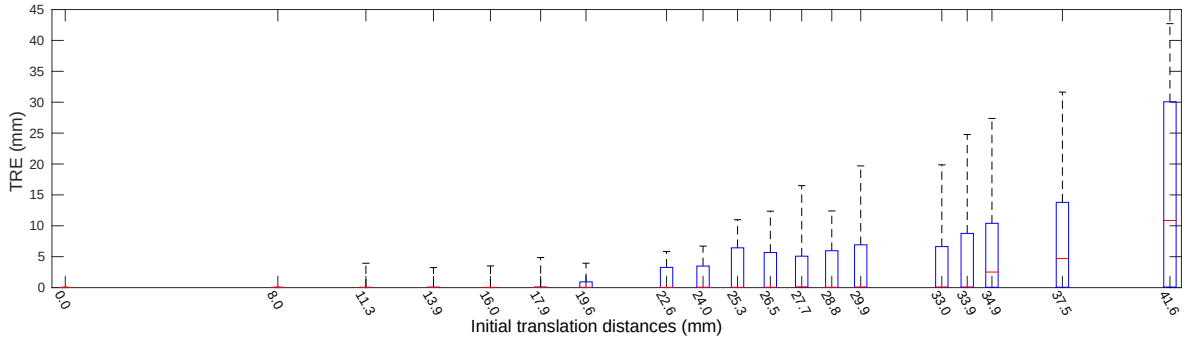


Figure 4.9: Target registration error (TRE) as a function of initial translation distance after affine registration to a reference image obtained by an initial rigid transformation, Data Set 1.

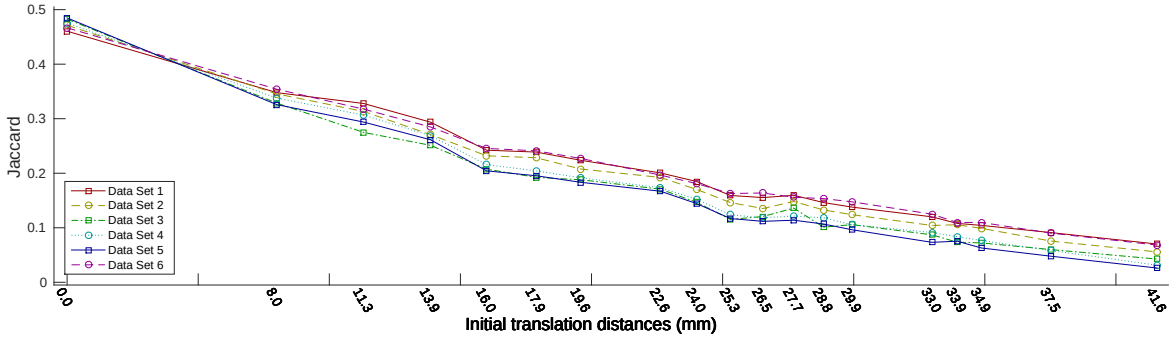


Figure 4.10: LV overlap as a function of initial translation distance before registration to a reference image obtained by an initial rigid transformation, all data sets.

more optimal than a solution near the initial state when the images are far apart to begin with. Similarly, the LV overlap as quantified by the Jaccard index is improved by both rigid and affine registration, and it appears that affine registration improves LV overlap more than rigid registration, but since the limited number of data sets is limited, we can say that the affine model is comparable to the rigid model.

The same data was analysed, but in relation to the amount of rotation made in the initial transformation to obtain the reference image R. The results are shown in Figures 4.16 to 4.21.

Since affine registration worked more consistently compared to rigid registration and

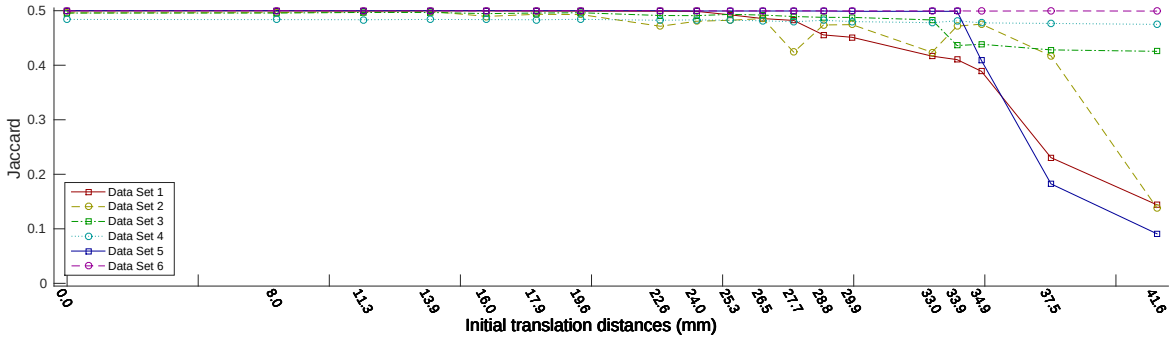


Figure 4.11: LV overlap as a function of initial translation distance after rigid registration to a reference image obtained by an initial rigid transformation, all data sets.

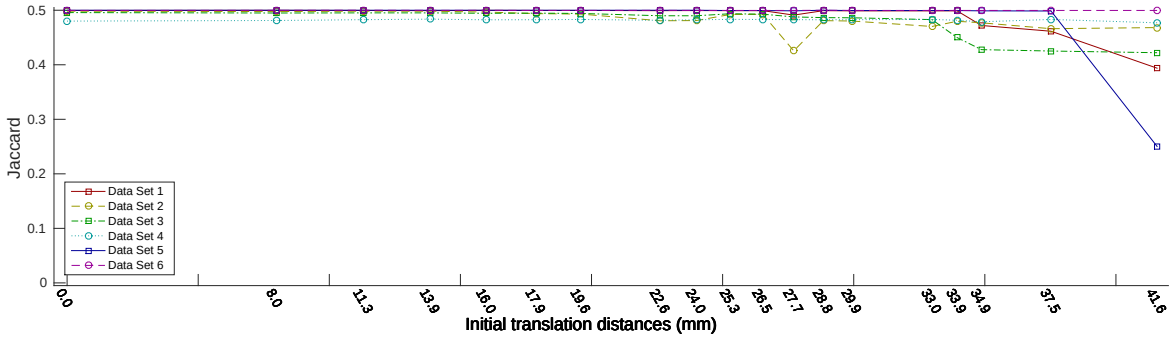


Figure 4.12: LV overlap as a function of initial translation distance after affine registration to a reference image obtained by an initial rigid transformation, all data sets.

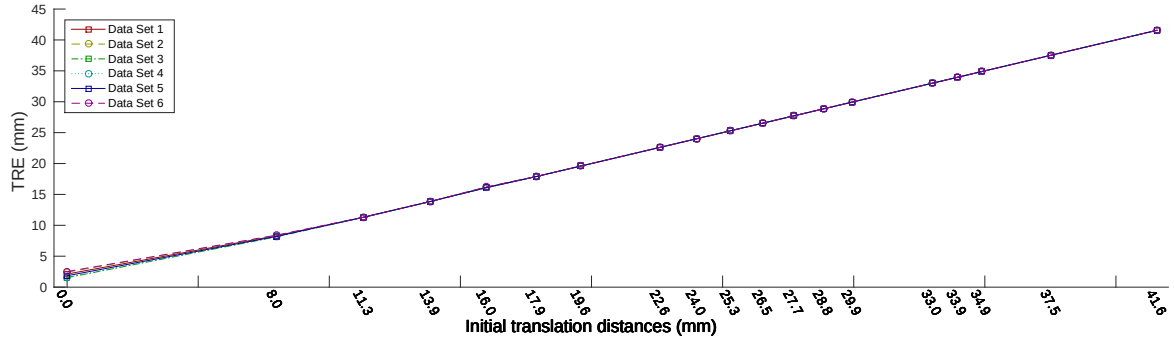


Figure 4.13: Target registration error (TRE) as a function of initial translation distance before registration to a reference image obtained by an initial rigid transformation, all data sets.

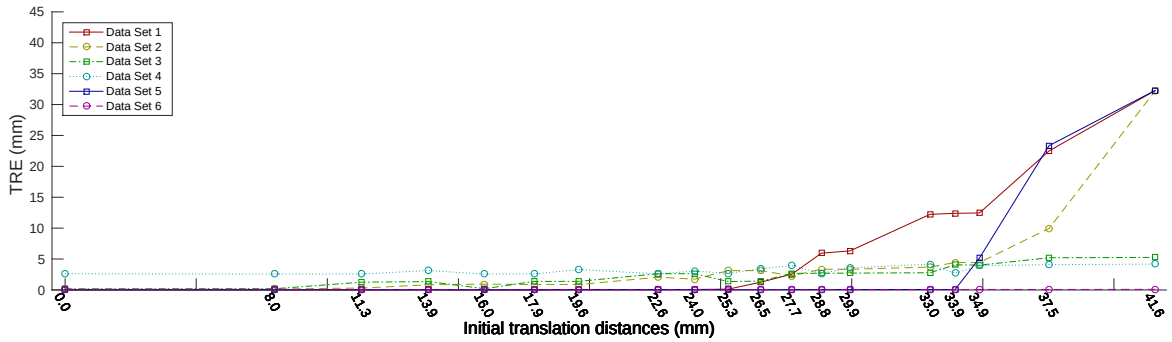


Figure 4.14: Target registration error (TRE) as a function of initial translation distance after rigid registration to a reference image obtained by an initial rigid transformation, all data sets.

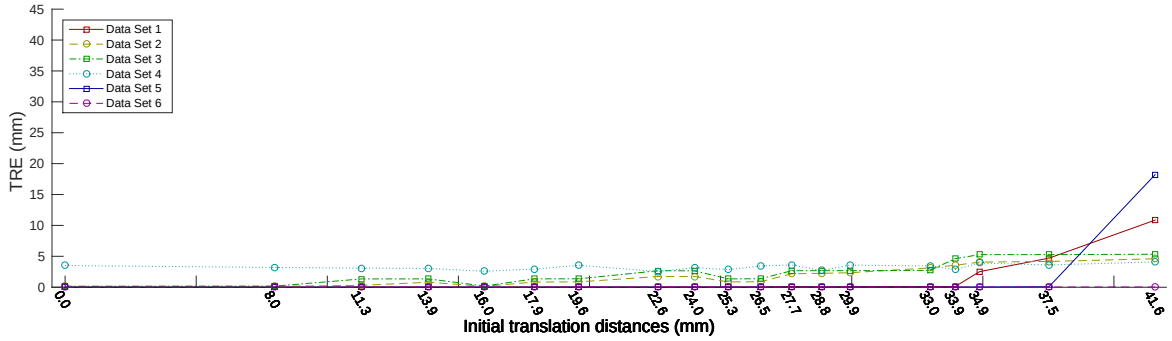


Figure 4.15: Target registration error (TRE) as a function of initial translation distance after affine registration to a reference image obtained by an initial rigid transformation, all data sets.

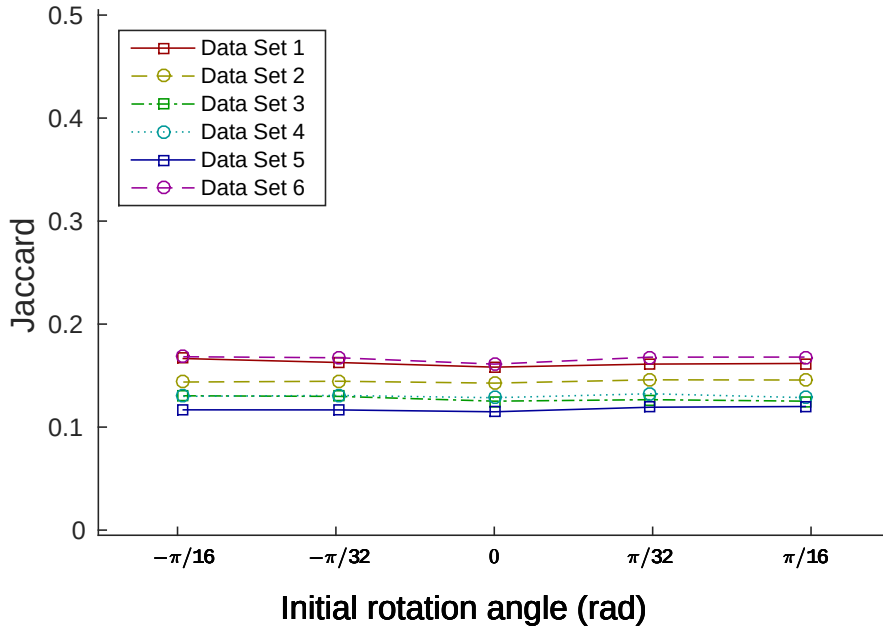


Figure 4.16: LV overlap as a function of initial rotation angle before registration to a reference image obtained by an initial rigid transformation, all data sets.

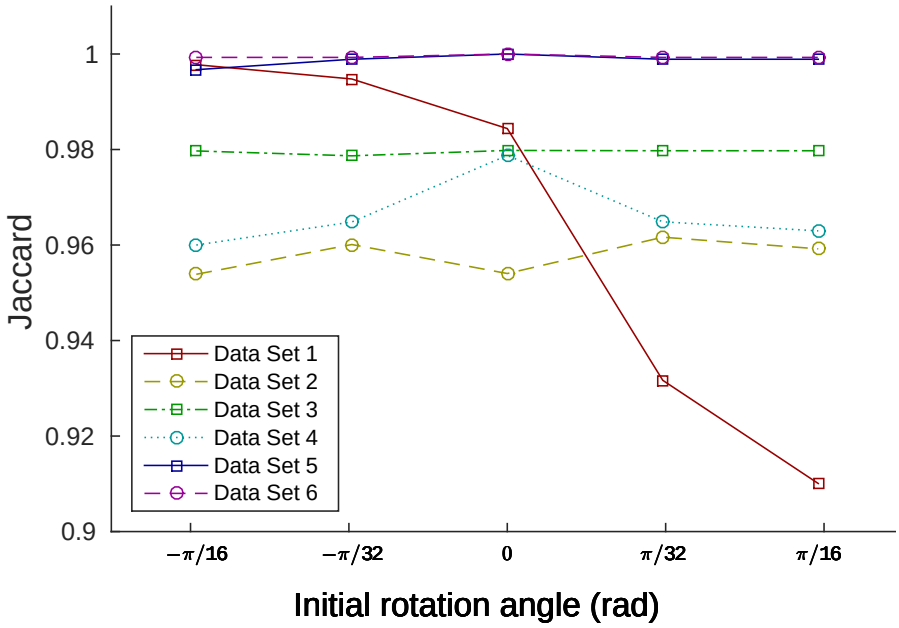


Figure 4.17: LV overlap as a function of initial rotation angle after rigid registration to a reference image obtained by an initial rigid transformation, all data sets. Note the change in scale of the vertical axis.

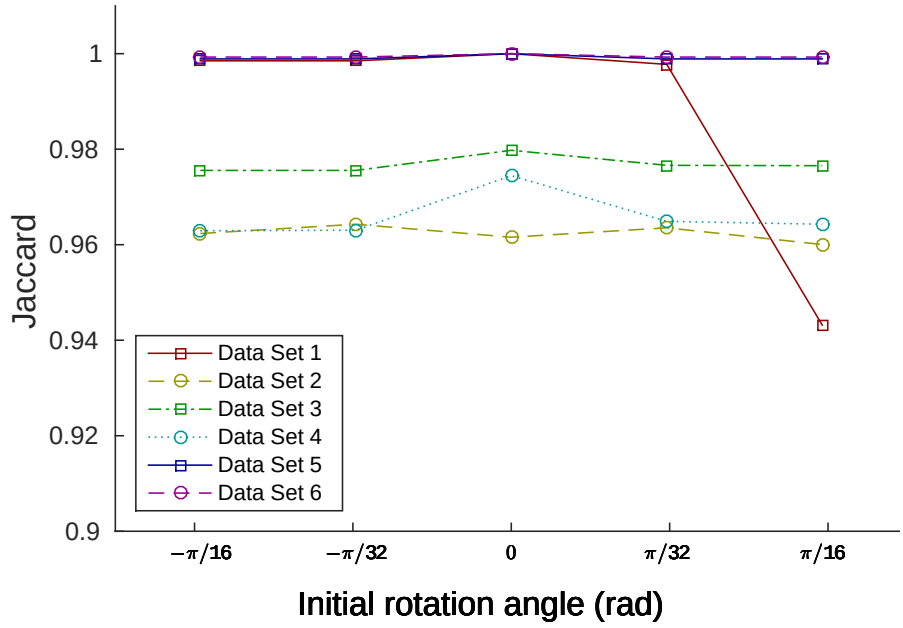


Figure 4.18: LV overlap as a function of initial rotation angle after affine registration to a reference image obtained by an initial rigid transformation, all data sets. Note the change in scale of the vertical axis.

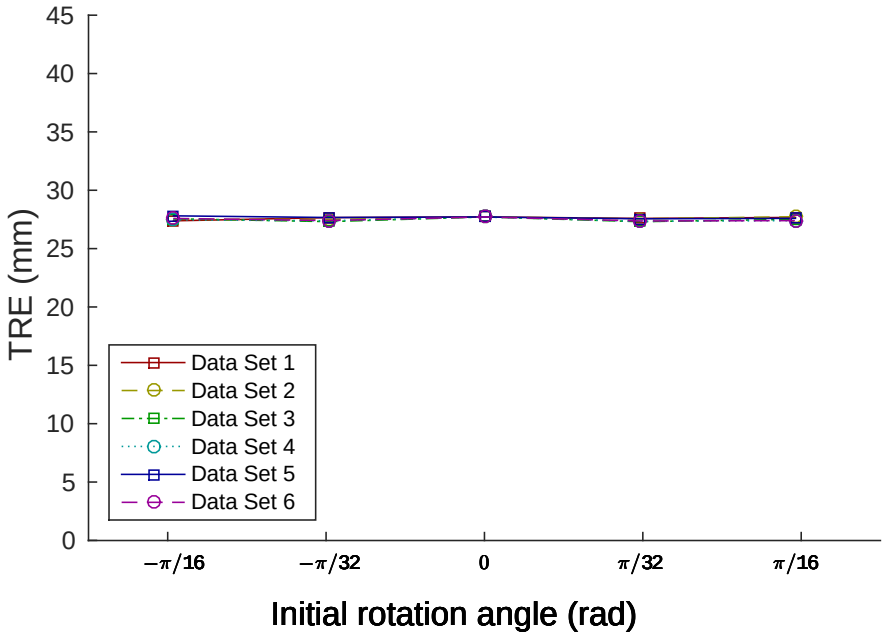


Figure 4.19: Target registration error (TRE) as a function of initial rotation angle before registration to a reference image obtained by an initial rigid transformation, all data sets.

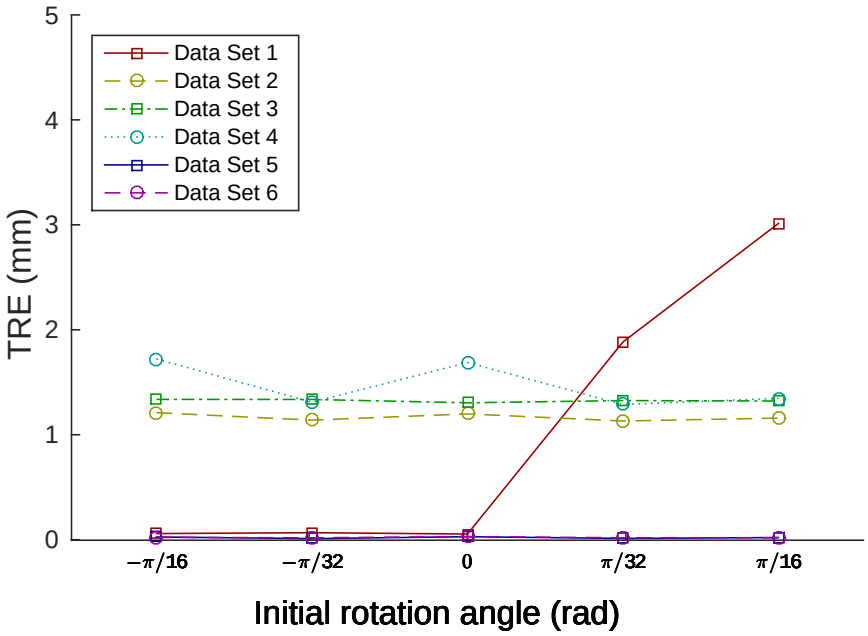


Figure 4.20: Target registration error (TRE) as a function of initial rotation angle after rigid registration to a reference image obtained by an initial rigid transformation, all data sets. Note the change in scale of the vertical axis.

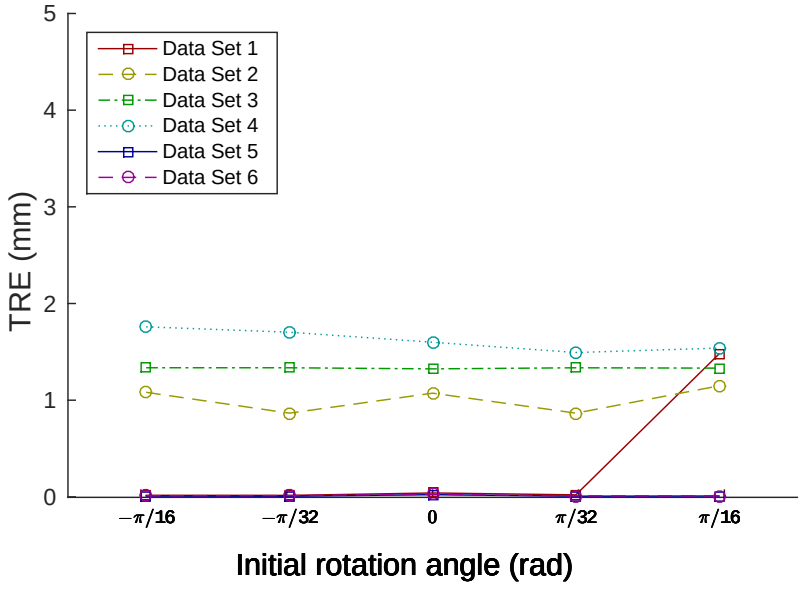


Figure 4.21: Target registration error (TRE) as a function of initial rotation angle after affine registration to a reference image obtained by an initial rigid transformation, all data sets. Note the change in scale of the vertical axis.

performed better for a wider range of initial transformation parameters, it suggests that it is more robust compared to the rigid model, albeit slower.

4.4 Cine/cine Experiments – Affine Initial Transformation

In the previous section, control experiments were done where the initial transformation applied to obtain the reference image was a rigid one. Recall that the motivation behind using an affine model as opposed to a rigid was to more accurately represent the deformable nature of organs in the body. To demonstrate that the rigid model does indeed fail when the nature of the deformation applied to the reference image R is not rigid, we individually perturbed each entry of the identity transform $\mathtt{w} = [1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0]$, and applied the perturbed set of parameters in the initial transformation to obtain R. For each R that was obtained, rigid and affine registration was performed and results evaluated. Due to Theorem 1, perturbing entries w_3, w_7, and w_{11} is equivalent (in terms of producing the same template image slice) to perturbing w_4, w_8, and w_{12} but scaled by a factor of z, the location of the slicing operator, so only 9 entries of $\mathtt{w}$ need to be perturbed; w_4, w_8, and w_{12} were not perturbed. For each of the 9 entries, an ϵ between -0.5 to 0.5 was added to the entry to produce a set of initial parameters used to obtain R.

Figures 4.22, 4.23, and 4.24 show the effects of perturbing the entries of w_2 on LV overlap before registration, after rigid registration, and after affine registration. As expected, affine registration improves results over rigid registration. Similar results found for entries w_1, w_5, w_6, w_9, and w_{10}, but for the sake of brevity, no figures will be shown for those entries.

Figures 4.25, 4.26, and 4.27 show the effects of perturbing the entries of w_3 on LV overlap before registration, after rigid registration, and after affine registration. Due to

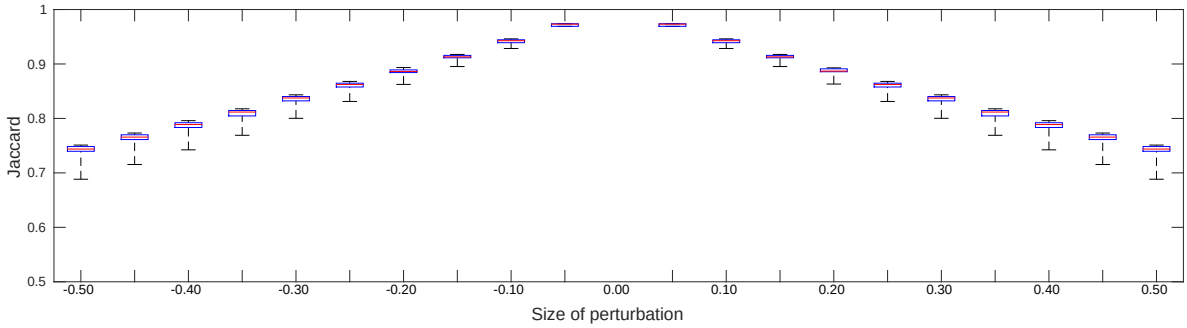


Figure 4.22: LV overlap as a function of the perturbation on w_2 before registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_2.

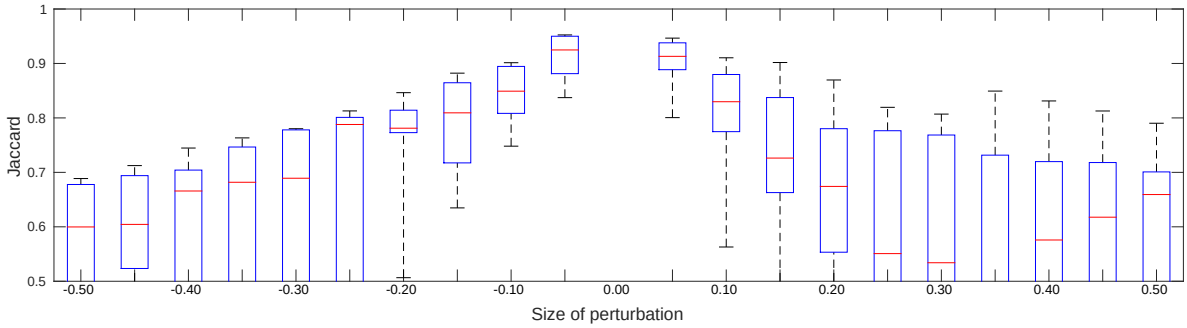


Figure 4.23: LV overlap as a function of the perturbation on w_2 after rigid registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_2.

Theorem 1, perturbations in w_3, w_7, and w_{11} can be compensated for during registration by changing the values of w_4, w_8, and w_{12}, which are translation parameters and therefore a rigid transformation. Rigid registration was therefore comparable to affine transformation for perturbations on w_3, as seen in Figures 4.26 and 4.27. The same was found for entries w_7 and w_{11}, but for the sake of brevity, no figures will be shown for those entries.

Recall that an indicator of good alignment is a simultaneously large Jaccard index and small TRE. Figures 4.28 to 4.33 show the effects of perturbation on w_2 and w_3 on the TRE before registration, after rigid registration, and after affine registration. Consistent

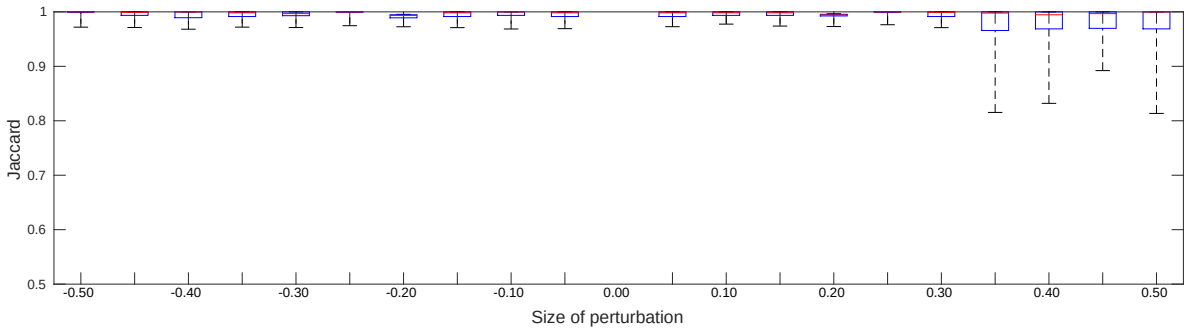


Figure 4.24: LV overlap as a function of the perturbation on w_2 after affine registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_2.

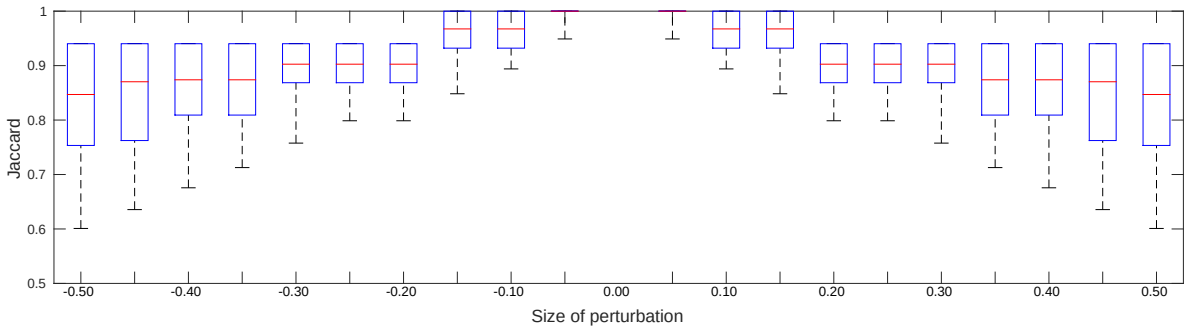


Figure 4.25: LV overlap as a function of the perturbation on w_3 before registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_3.

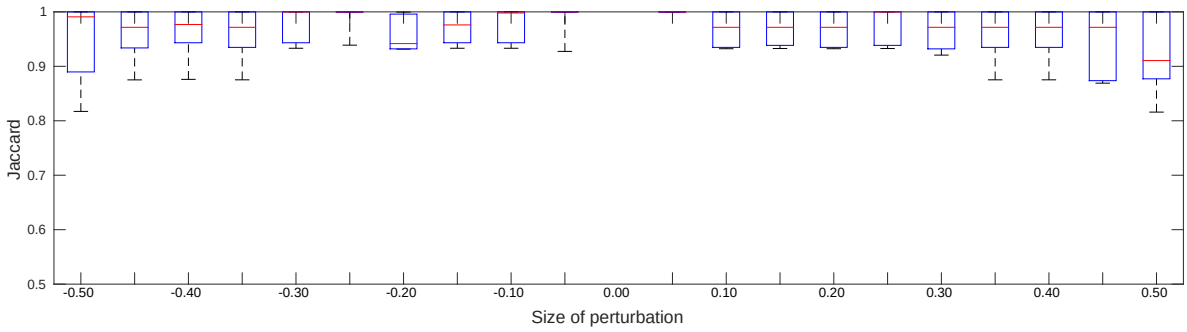


Figure 4.26: LV overlap as a function of the perturbation on w_3 after rigid registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_3.

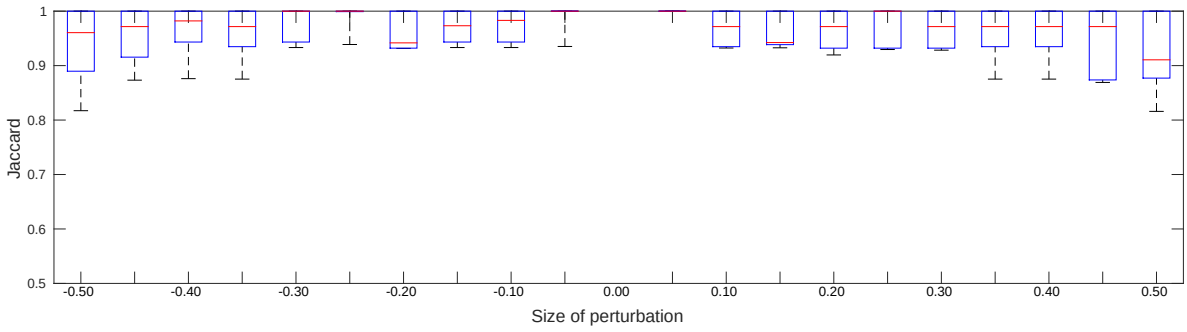


Figure 4.27: LV overlap as a function of the perturbation on w_3 after affine registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_3.

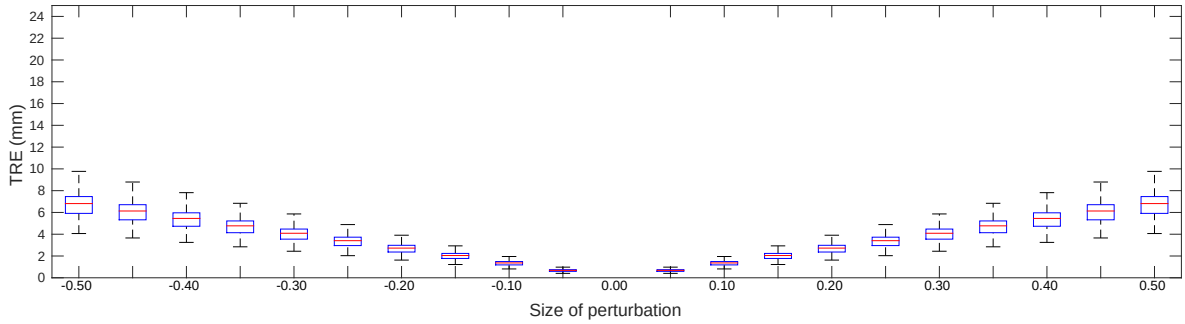


Figure 4.28: The TRE as a function of the perturbation on w_2 before registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_2.

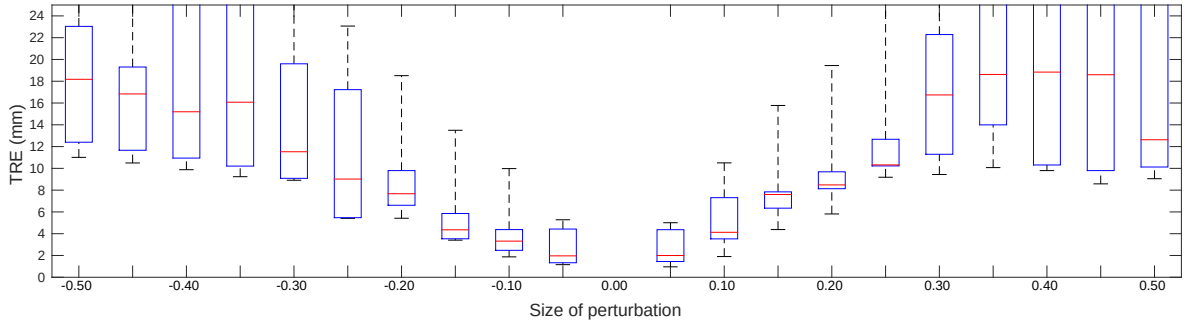


Figure 4.29: The TRE as a function of the perturbation on w_2 after rigid registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_2.

with the previous paragraph, it was also found that perturbing the entries of w_2 produced results similar to those obtained by perturbing the entries of w_1, w_5, w_6, w_9, and w_{10}, and perturbing the entries of w_3 produced results similar to those obtained by perturbing the entries of w_7 and w_{11}.

For perturbations on w_2 (and $w_1, w_5, w_6, w_9,, w_{10}$), affine registration generally did well, increasing Jaccard indices and reducing TREs. Rigid registration did not improve results; Jaccard indices after rigid registration became more varied and generally appear to worsen.

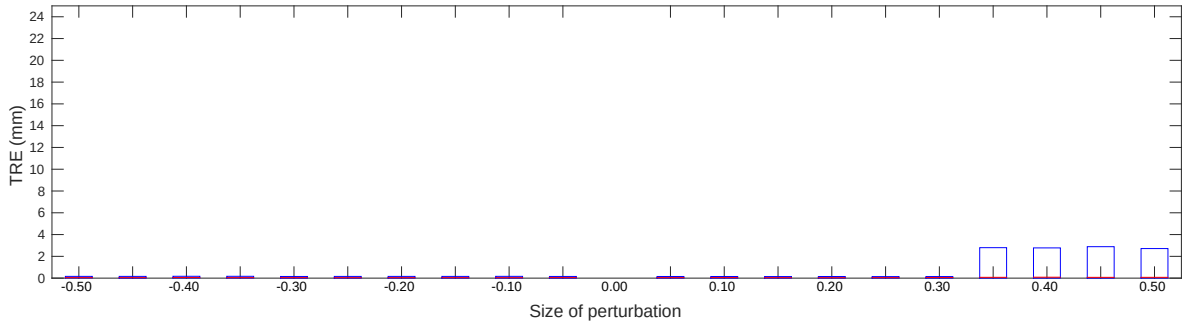


Figure 4.30: The TRE as a function of the perturbation on w_2 after affine registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_2.

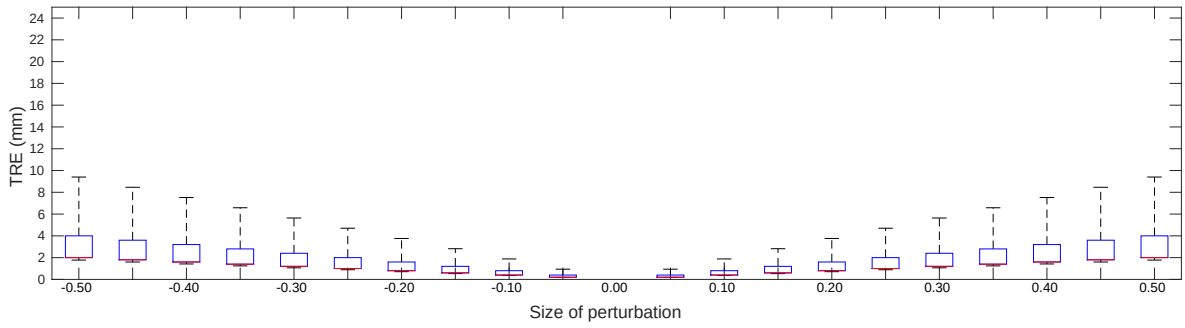


Figure 4.31: The TRE as a function of the perturbation on w_3 before registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_3.

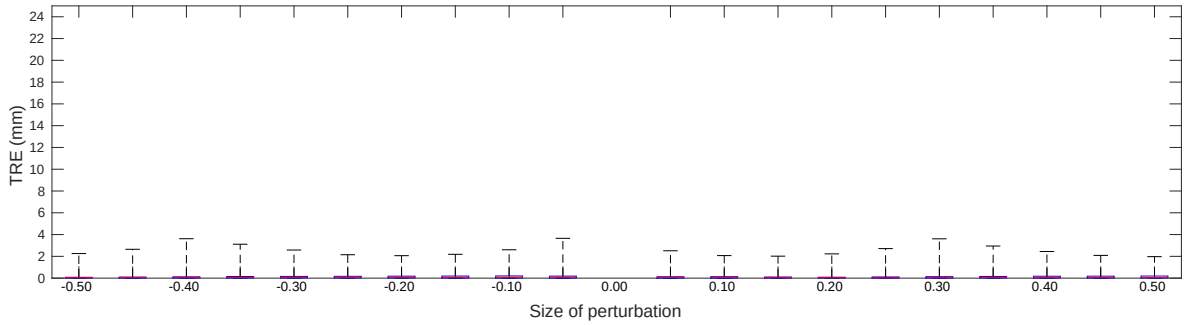


Figure 4.32: The TRE as a function of the perturbation on w_3 after rigid registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_3.

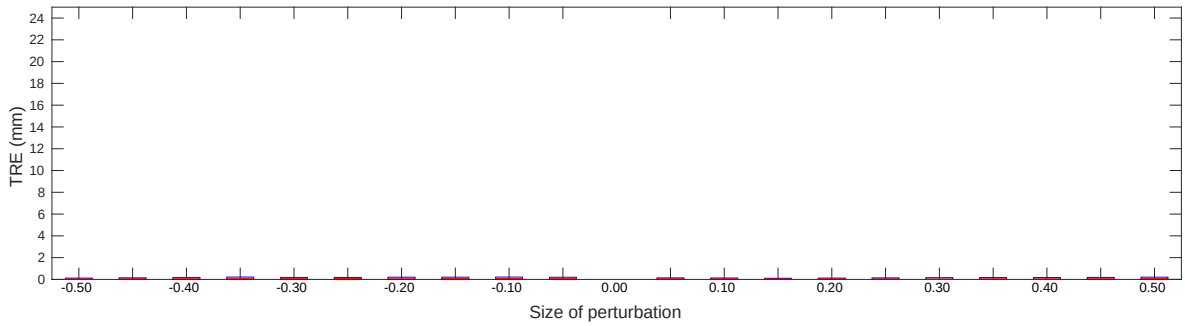


Figure 4.33: The TRE as a function of the perturbation on w_3 after affine registration for all data sets. Reference image obtained by an affine transformation that is the identity transformation except for the addition of the perturbation to w_3.

For perturbations on w_3 (and w_7, w_{11}), the results of rigid registration were comparable to affine registration. This can be explained by Theorem 1, which states that variations in w_3, w_7, and w_{11} can be compensated for by changing the values of w_4, w_8, and w_{12} to obtain the same 2D slice of a 3D volume. Since w_4, w_8, and w_{12} are translation parametres in the x-, y-, and z-directions, respectively, and translation is allowed in rigid transformations, rigid registration was able to compensate for perturbations on w_3 (and w_7, w_{11}) and produce results similar to that of affine registration.

4.5 Real-time/Cine Experiments

After we are satisfied that the model works for controlled test cases, we would like to test the model on clinical data.

There is no initial 3D transformation applied to obtain the reference image as was done on the test cases, since the reference images here are 2D real-time images. It is also not meaningful to perform a 2D transformation on a real-time image to obtain the reference image for registration, since a modified image no longer represents an actual clinical situation. Because the slice locations in the real-time and cine cardiac MRI are already rather aligned initially in the z-direction, registration between images from same the slice prescription would align things mostly within the xy-plane, and give little indication of how well the algorithm works when the images are taken at different slice locations. Performing registration between different slices would be a better indicator of how well the algorithm improves alignment in the z-direction. For the following example (Figures 4.34 and 4.35), the real-time slice was taken at spatial location $z = -4$ mm while the slicing operation was applied on the template image at $z = -36$ mm, so the initial slice of the 3D template is at $z = -36$ mm. To register the images successfully, the registration algorithm must return transformation parameters that translate the template image by approximately 32 mm (the physical distance between the spatial locations of the reference

	Jaccard	Dice	TRE (mm)
Before registration	0.67	0.80	32.8 ± 0.2
After rigid registration	0.71	0.83	6.6 ± 0.8
After affine registration	0.87	0.93	4.5 ± 0.1

Table 4.2: Jaccard indices and Dice coefficients of left ventricle overlap before and after registration in an experiment where the template image is a 3D cine volume and the reference image is a real-time image at $z = -4$ mm. Slicing operations performed at $z = -36$ mm.

image and slicing operator) in the z-direction, along with appropriate alignments in the x- and y-directions.

Figures 4.34 and 4.35 and Table 4.2 show the results of an experiment. The affine model appears to produce slightly better results for this experiment, due to its ability to deform, apparent in the LV overlap after rigid and affine registration (Figures 4.34g and 4.35g).

In most clinical applications, initial misalignment will not be as large and the two images registered will be slices in close proximity to one another. Affine and rigid registration was performed on real-time images from 6 data sets, each contributing 1 cine image and between 17 to 29 real-time images, to a total of 143 real-time images between 6 data sets. Each real-time image was registered to a cine image of the same volunteer at the same slice location and cardiac phase. Although the slice prescriptions are identical, there may be small motion normal to the image plane. The results are listed in Table 4.3.

With the exception of Data Set 3 and Data Set 5, rigid registration improves or leaves results unchanged. Affine registration improves results for all data sets except Data Set 3. For Data Set 3, rigid registration returned values worse than what was initially given

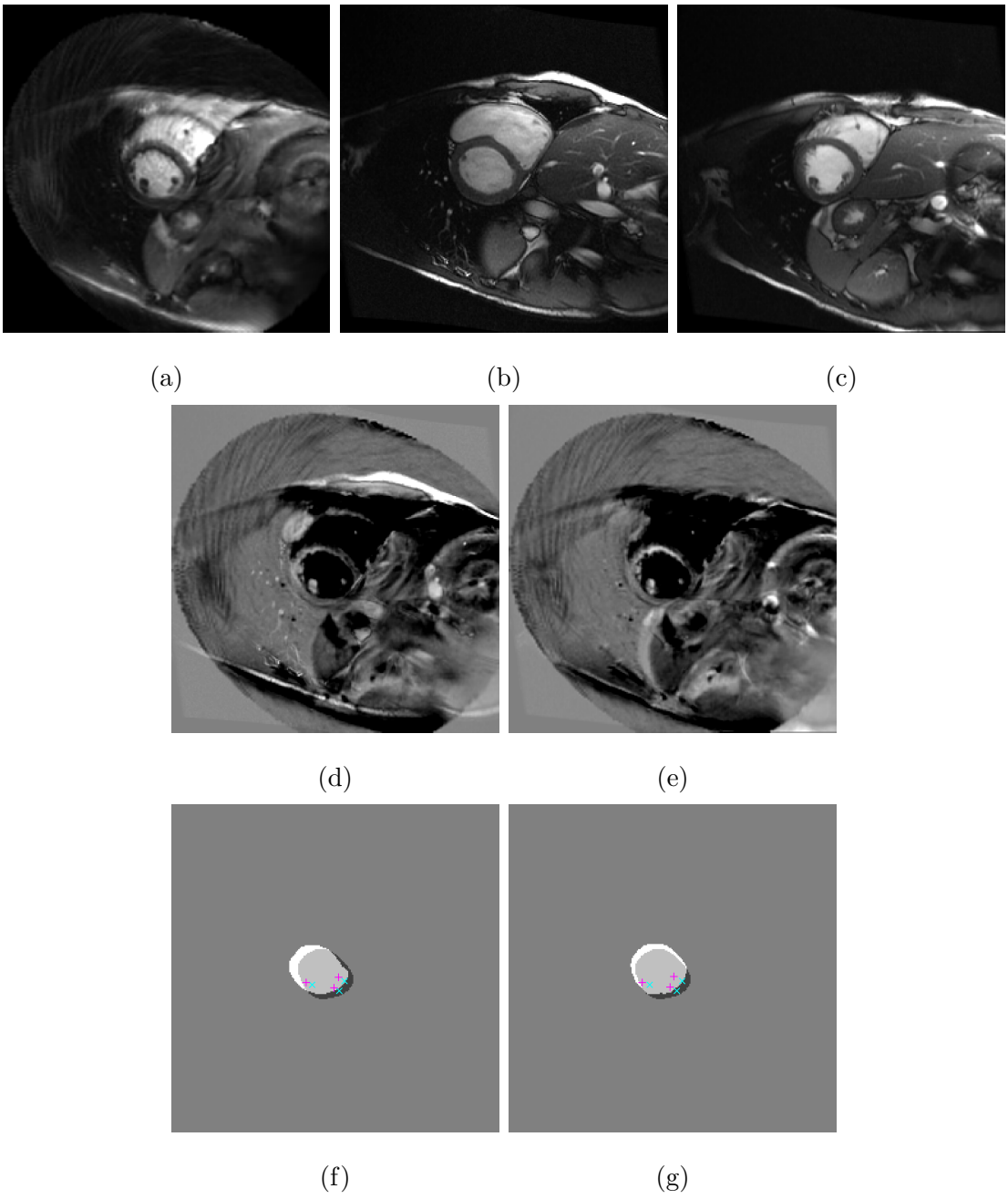


Figure 4.34: Results of rigid registration between a 3D cine image and a 2D real-time image on the same subject as in the controlled experiment, with an initial misalignment of approximately 32mm in the z-direction (through the image plane). (a) Reference image R. (b),(c) Template slice L before and after registration. (d),(e) Difference between the reference image and template slice $(L - R)$ before and after registration. (f),(g) Segmentation masks showing left ventricle overlap before and registration, with in-plane reference image landmarks ($\times$) and out-of-plane template image landmarks projected onto image ($+$).

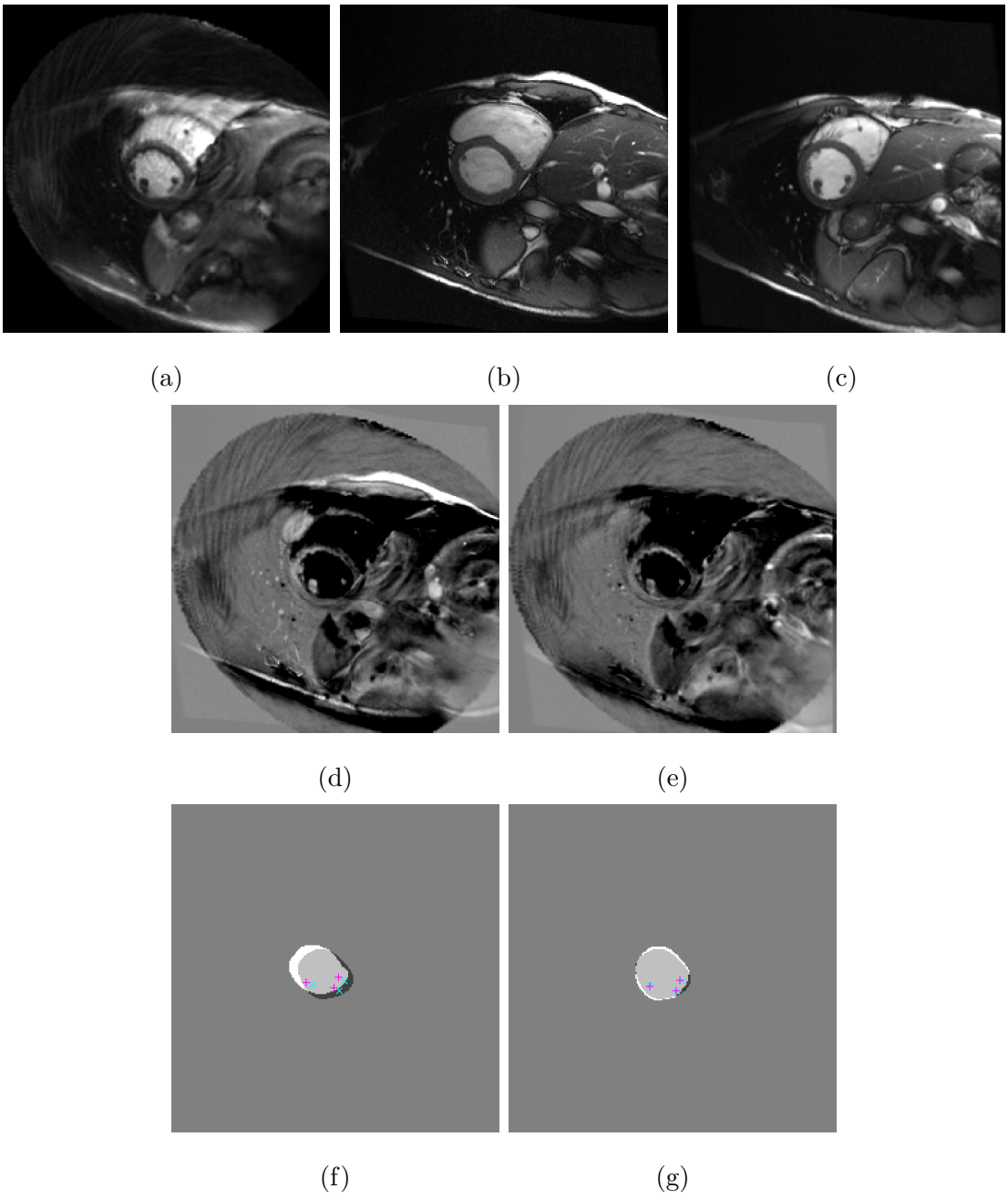


Figure 4.35: Results of affine registration between a 3D cine image and a 2D real-time image on the same subject as in the controlled experiment, with an initial misalignment of approximately 32mm in the z-direction (through the image plane). (a) Reference image R. (b),(c) Template slice L before and after registration. (d),(e) Difference between the reference image and template slice $(L - R)$ before and after registration. (f),(g) Segmentation masks showing left ventricle overlap before and after registration, with in-plane reference image landmarks ($\times$) and out-of-plane template image landmarks projected onto image ($+$).

	Data Set	Before registration	After rigid registration	After affine registration
Jaccard	1	0.86 ± 0.06	0.87 ± 0.07	0.92 ± 0.02
Dice	1	0.92 ± 0.04	0.93 ± 0.04	0.96 ± 0.01
Jaccard	2	0.75 ± 0.02	0.86 ± 0.02	0.87 ± 0.02
Dice	2	0.86 ± 0.02	0.93 ± 0.01	0.93 ± 0.01
Jaccard	3	0.77 ± 0.06	0.47 ± 0.13	0.21 ± 0.08
Dice	3	0.87 ± 0.04	0.63 ± 0.13	0.34 ± 0.12
Jaccard	4	0.49 ± 0.08	0.66 ± 0.05	0.73 ± 0.08
Dice	4	0.65 ± 0.07	0.80 ± 0.04	0.84 ± 0.06
Jaccard	5	0.80 ± 0.04	0.73 ± 0.03	0.83 ± 0.04
Dice	5	0.89 ± 0.03	0.84 ± 0.02	0.91 ± 0.02
Jaccard	6	0.76 ± 0.09	0.76 ± 0.09	0.80 ± 0.09
Dice	6	0.86 ± 0.06	0.86 ± 0.06	0.88 ± 0.06

Table 4.3: LV overlap before registration, after rigid registration, and after affine registration between a pre-operative 3D cine volume and a noisier, lower-resolution intra-operative 2D real-time image, as in a clinical setting.

and affine registration performed even worse. This was due to local deformation in the cardiac region, consistent throughout the data available for Data Set 3. Since the body cavity is considerably larger than the cardiac region and comprises most of the content in each image, the algorithm accounted for the body cavity, not the heart. Thus, the LV becomes more misaligned after registration. For Data Set 5, rigid registration returned slightly worse values than what the algorithm initially started with, but affine registration produced values that were a slight improvement over the initial data. From the values for the rest of the data sets, however, affine registration returns better results in general compared to rigid registration. The TRE was not calculated in this set of experiments because the images are from the same slice locations, image resolution in the z-direction (the direction normal to a short-axis slice) is much coarser than in the x- and y-directions (within the plane of a short-axis image slice), and there is no ground truth available for us to obtain more precise landmark locations. z-direction uncertainty would dominate and render the results meaningless.

Chapter 5

Conclusions and Future Work

5.1 Discussions and Conclusions

In the previous section, it was demonstrated that the algorithm performs well in control experiments where the reference image is a transformed and sliced version of the template image. The left ventricle (LV) overlap between the 3D cine volume and a 2D cine image in the previous section aligned well after registration, as quantified by the increased Jaccard and Dice indices and by the lowered TRE. The affine model aligned images well for a broader range of initial transformation parameters used to obtain R, and returned better results compared to the rigid model for the 6 data sets available here, so we can conclude that the affine model performs better than, or is at least comparable to, the rigid model for control experiments, but at a small expense of computational time.

It was also demonstrated that, in control experiments, rigid registration does not sufficiently account for deformations that are affine in nature. Rigid registration can, however, return the same end-result image slice (slice of transformed template) if the only affine deformations are scaling in the z-direction (w_{11}) or shearing in directions normal to the z-axis (w_3 and w_7).

The algorithm was also tested on other clinical data, registering high-resolution cine

images to noisier, lower-resolution real-time images. An example of registration performed between a real-time image slice and a cine image sliced at a location approximately 32 mm away in the normal direction from the real-time image was shown in the previous section. The registration algorithm corrected the large z-direction translation and produced a resulting image with structural features in the heart very similar to those in the reference image, shown in Figures 4.34 and 4.35. It is worth noting that, the regions outside the heart may not look as similar due to large motion of surrounding organs such as the lungs and diaphragm. The Jaccard and Dice measures are two-dimensional measures, therefore it is also important to note that their values are projections onto a 2D plane. Visually, the cine slice aligns well after registration and appears to be in the same plane as the real-time image, but before registration, the cine slice is out of the plane so measures of overlap before registration may not be meaningful. In computing the TRE before and after registration, the z-component is also an estimate based on the geometrical locations of the image slices, assuming that landmark motion normal to the image plane is small compared to the slice thickness (8mm). The only way of verifying the true location of a landmark, pinpointing its precise location, is to attach fiducial markers to the landmarks – impossible for *in vivo* samples for obvious ethical reasons. For the reasons described above, for such experiments where there is no 'ground truth' and where misalignment is attributed mainly to a component normal to the image plane, i.e. the reference and template images are of objects in different geometrical spaces, it is important to look at the end-result images to judge how well the registration algorithm has aligned the images.

In registration between real-time images and cine images of the same slice location and cardiac phase, affine registration generally performed better than rigid registration, presumably due to its greater flexibility over the rigid model, allowing it to deform the cine image to more closely match the real-time image. In Data Set 3, neither rigid nor affine registration improved LV overlap after registration, and for Data Set 5, rigid

registration did not improve LV overlap and affine registration only improved results slightly.

We can conclude that between images of the same modality, the proposed multi-level parametric 2D-3D registration scheme can align images well for misalignments within reasonable limits encountered in clinical applications, such as motion due to respiration. Despite different acquisition methods in the real-time and prior cine MR images, the registration algorithm improved alignment with the SSD dissimilarity measure.

Affine registration was found to be a generally more robust model than rigid registration in this framework. This suggests that in attempting to improve results for applications employing 2D-3D rigid registration with the SSD, one can first consider simply expanding the transformation model to an affine one before considering more complex dissimilarity measures and regularizers. The advantage of the affine model is its simplicity, allowing more accurate registration at a small cost.

For multi-modality registration where intensities of the template and reference images differ more drastically, one can consider using other dissimilarity measures and/or optimizers [Gos05, Mod09] that can fit well within the context of the general proposed model.

5.2 Future Work

Although the template and reference MR images used in our experiments were acquired using different pulse sequences, they were similar enough that the SSD was sufficient as a distance measure. The SSD was initially not thought to be ideal as a distance measure, since it depends on absolute values of intensity values, which can vary between different image acquisitions. For registration between differing image types and to see if current results can be improved, we could compare results obtained using other dissimilarity measures. In registering MR images, we are most interested in aligning the boundaries

of organs and the outlines of their features. One candidate distance measure is the Normalized Gradient Field (NGF), which interprets intensity gradients as surface normals and measures dissimilarity based on how well aligned the normals in one image are to the normals in the other. It is however, susceptible to noise. One possible solution is to smooth both images prior to measuring dissimilarity.

Other areas we could investigate are different choices of regularizers and optimizers. Of interest are optimizers that may be more efficient and less taxing on computational resources.

With data for 6 volunteers, our observations may generalize on cine and real-time cardiac MR images, but with large uncertainty. If more data becomes available, we would test our model again to see if similar results are observed and how consistent those observations are.

Bibliography

[A+97] Gerhard Adam et al. Interventional magnetic resonance: Initial clinical experience with a 1.5-tesla magnetic resonance system combined with C-arm fluoroscopy. *Investigative Radiology*, 32:191–197, April 1997.

[B+02] Arno Buecker et al. Magnetic resonance–guided placement of atrial septal closure device in animal model of patent foramen ovale. *Circulation*, 106(4):511–515, 2002.

[B+07] Wolfgang Birkfellner et al. Rigid 2D/3D slice-to-volume registration and its application to fluoroscopic CT. *Medical Physics*, 34:246–255, 2007.

[BDTM12] Jan Bogaert, Steven Dymarkowski, Andrew M. Taylor, and Vivek Muthurangu. *Clinical Cardiac MRI*. Springer, 2nd edition, 2012.

[BGLS06] J. Frédéric Bonnans, J. Charles Gilbert, Claude Lemaréchal, and Claudia A. Sagastizábal. *Numerical Optimization: Theoretical and Practical Aspects*. Springer, 2nd edition, 2006.

[CCB+12] Elizabeth B. Claus, Lisa Calvocoressi, Melissa L. Bondy, Joellen M. Schildkraut, Joseph L. Wiemels, and Margaret Wrensch. Dental X-rays and risk of meningioma. *Cancer*, 118(18):4530–4537, 2012.

[CdB80] Samuel D. Conte and Carl de Boor. *Elementary Numerical Analysis: An Algorithmic Approach*. McGraw-Hill, 3rd edition, 1980.

[D+05] Alexander J. Dick et al. Invasive human magnetic resonance imaging: Feasibility during revascularization in a combined XMR suite. *Catheterization and Cardiovascular Interventions*, 64(3):265–274, 2005.

[DA08] R. Dalvi and R. Abugharbieh. Fast feature based multi slice to volume registration using phase congruency. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5390–5393, Aug 2008.

[FDB+03] Baowei Fei, Jeffrey L. Duerk, Daniel T. Boll, Jonathan S. Lewin, and David L. Wilson. Slice-to-volume registration and its potential application to interventional MRI-guided radio-frequency thermal ablation of prostate cancer. *IEEE Trans. Med. Imag.*, 22(4):515–525, 2003.

[FFP15] Enzo Ferrante, Vivien Fecamp, and Nikos Paragios. Slice-to-volume deformable registration: efficient one-shot consensus between plane selection and in-plane deformation. *Int. J CARS*, 10:791–800, 2015.

[FP13] Enzo Ferrante and Nikos Paragios. Non-rigid 2D-3D medical image registration using Markov random fields. In *16th International Conference on Medical Image Computing and Computer Assisted Intervention – MICCAI 2013*, volume 8151 of *Lecture Notes in Computer Science*, pages 163–170, September 2013.

[G+09] Thomas C. Gerber et al. Ionizing radiation in cardiac imaging – a science advisory from the American Heart Association Committee on Cardiac Imaging of the Council on Clinical Cardiology and Committee on Cardiovascular Imaging and Intervention of the Council on Cardiovascular Radiology and Intervention. *Circulation*, 119:1056–1065, 2009.

[Gos05] A. Ardeshir Goshtasby. *2-D and 3-D Image Registration*. Wiley Press, New York, 2005.

[H+10] Boris A. Hoffmann et al. Interactive real-time mapping and catheter ablation of the cavotricuspid isthmus guided by magnetic resonance imaging in a porcine model. *European Heart Journal*, 31:450–456, 2010.

[Jar16] Sergei V. Jargin. On the tumor risk from dental diagnostic X-ray exposure. *Annals of Oncology*, 27(6):1177, 2016.

[KCKK06] Spyros Kokolis, Luther T. Clark, Rodamanthos Kokolis, and John Kassotis. Ventricular arrhythmias and sudden cardiac death. *Progress in Cardiovascular Diseases*, 48(6):426–444, 2006.

[KHB+08] Sebastian Kos, Rolf Huegli, Georg M. Bongartz, Augustinus L. Jacob, and Deniz Bilecen. MR-guided endovascular interventions: a comprehensive review on techniques and applications. *European Radiology*, 18(4):645–657, 2008.

[KS09] Bruce A. Koplan and William G. Stevenson. Ventricular tachycardia and sudden cardiac death. *Mayo Clin. Proc.*, 84(3):289–297, March 2009.

[L+00] Albert C. Lardo et al. Visualization and temporal/spatial characterization of cardiac radiofrequency ablation lesions using magnetic resonance imaging. *Circulation*, 102:698–705, 2000.

[L+04] W. T. Longstreth et al. Dental X-rays and the risk of intracranial meningioma. *Cancer*, 100(5):1026–1034, 2004.

[Lar00] Albert C. Lardo. Real-time magnetic resonance imaging: Diagnostic and interventional applications. *Pediatr. Cardiol.*, 21(1):80–98, 2000.

[LYJ03] Harel Livyatan, Ziv Yaniv, and Leo Joskowicz. Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. *IEEE Trans. Med. Imag.*, 22(11):1395–1406, 2003.

[M+15] D. Mozaffarian et al. Heart disease and stroke statistics – 2015 Update: A report from the American Heart Association. *Circulation*, 131:e29–e322, 2015.

[Mod09] Jan Modersitzki. *FAIR: Flexible Algorithms for Image Registration*. SIAM, 2009.

[Moo05] Phillip Moore. MRI-guided congenital cardiac catheterization and intervention: The future? *Catheterization and Cardiovascular Interventions*, 66(1):1–8, 2005.

[MR05] Vivek Muthurangu and Reza S. Razavi. The value of magnetic resonance guided cardiac catheterisation. *Heart*, 91(8):995–996, August 2005.

[MTPL08] Primož Merkelj, Dejan Tomaževič, Franjo Pernuš, and Boštjan Likar. Robust gradient-based 3-D/2-D registration of CT and MR to X-ray images. *IEEE Trans. Med. Imag.*, 27(12):1701–1714, 2008.

[NW06] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, New York, 2nd edition, 2006.

[OK10] Sergey Osechinskiy and Frithjof Kruggel. Slice-to-volume nonrigid registration of histological sections to MR images of the human brain. *Anatomy Research International*, 2010, 2010.

[PBH+01] Graeme P. Penney, Philipp G. Batchelor, Derek L. G. Hill, David J. Hawkes, and Juergen Weese. Validation of a two- to three-dimensional

registration algorithm for aligning preoperative CT images and intraoperative fluoroscopy images. *Med. Phys.*, 28(1024):1024–1032, 2001.

[PGR15] Sven Plein, John Greenwood, and John P. Ridgway. *Cardiovascular MR Manual*. Springer, 2nd edition, 2015.

[PTR14] Kuberan Pushparajah, Aphrodite Tzifa, and Reza Razavi. Cardiac MRI catheterization: a 10-year single institution experience and review. *Interventional Cardiology*, 6, June 2014.

[Pub09] The Public Health Agency of Canada. *Tracking Heart Disease and Stroke in Canada*, 2009.

[QSS07] Alfio Quarteroni, Riccardo Sacco, and Fausto Saleri. *Numerical Mathematics*. Springer, 2nd edition, 2007.

[R+03a] Kawal S. Rhode et al. Registration and tracking to integrate X-ray and MR images in an XMR facility. *IEEE Trans. Med. Imag.*, 22(11):1369–1378, Nov 2003.

[R+03b] Carsten Rickers et al. Magnetic resonance image-guided transcatheter closure of atrial septal defects. *Circulation*, 107(1):132–138, 2003.

[R+05] K. S. Rhode et al. A system for real-time XMR guided cardiovascular intervention. *IEEE Transactions on Medical Imaging*, 24(11):1428–1440, Nov 2005.

[Ros13] Clive Rosendorff. *Essential Cardiology Principles and Practice*. Springer, New York, 3rd edition, 2013.

[S+06] Jan-Henry Seppenwoolde et al. Fully MR-guided hepatic artery catheterization for selective drug delivery: A feasibility study in pigs. *Journal of Magnetic Resonance Imaging*, 23(2):123–129, 2006.

[S+08] William G. Stevenson et al. Irrigated radiofrequency catheter ablation guided by electroanatomic mapping for recurrent ventricular tachycardia after myocardial infarction. *Circulation*, 118:2773–2782, 2008.

[S+09] Ehud J. Schmidt et al. Electroanatomic mapping and radiofrequency ablation of porcine left atria and atrioventricular nodes using magnetic resonance catheter tracking. *Circulation*, 2:695–704, 2009.

[SS07a] Olaf Saborowski and Maythem Saeed. An overview on the advances in cardiovascular interventional MR imaging. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 20(3):117–127, 2007.

[SS07b] William G. Stevenson and Kyoko Soejima. Catheter ablation for ventricular tachycardia. *Circulation*, 115:2750–2760, 2007.

[SSH+03] Simon Schalla, Maythem Saeed, Charles B. Higgins, Alastair Martin, Oliver Weber, and Phillip Moore. Magnetic resonance–guided cardiac catheterization in a swine model of atrial septal defect. *Circulation*, 108(15):1865–1870, 2003.

[SWD04] Renata. Smolíková, Mark P. Wachowiak, and Maria Drangova. Registration of fast cine cardiac MR slices to 3D preprocedural images: toward real time registration for MRI-guided procedures. In *Proceedings of SPIE*, volume 5370, pages 1195–1205, May 2004.

[SWWFD05] Renata. Smolíková-Wachowiak, Mark P. Wachowiak, Aaron Fenster, and Maria Drangova. Registration of two-dimensional cardiac images to preprocedural three-dimensional images for interventional applications. *J. Magn. Reson. Im.*, 22:219–228, 2005.

[SYF+03] Jean-Michel Serfaty, Xiaoming Yang, Thomas K. Foo, Ananda Kumar, Andrew Derbyshire, and Ergin Atalar. MRI-guided coronary catheterization

and PTCA: A feasibility study on a dog model. *Magnetic Resonance in Medicine*, 49(2):258–263, 2003.

[TLP06] Dejan Tomaževič, Boštjan Likar, and Franjo Pernuš. 3-D/2-D registration by integrating 2-D information in 3-D. *IEEE Transactions on Medical Imaging*, 25(1):17–27, Jan 2006.

[TLSP03] Dejan Tomaževič, Boštjan Likar, Tomaž Slivnik, and Franjo Pernuš. 3-D/2-D registration of CT and MR to X-ray images. *IEEE Transactions on Medical Imaging*, 22(11):1407–1416, Nov 2003.

[TPL+14] Qian Tao, Sebastiaan R. D. Piers, Hildo J. Lamb, Katja Zeppenfeld, and Rob J. van der Geest. Preprocedural magnetic resonance imaging for image-guided catheter ablation of scar-related ventricular tachycardia. *The International Journal of Cardiovascular Imaging*, 31(2):369–377, 2014.

[vV98] Joop J. van Vaals. Interventional MR with a hybrid high-field system. In Jörg F. Debatin and Gerhard Adam, editors, *Interventional Magnetic Resonance Imaging*, chapter 3, pages 19–32. Springer, 1998.

[XLF+15] Helen Xu, Andras Lasso, Andriy Fedorov, Kemal Tuncali, Clare Tempany, and Gabor Fichtinger. Multi-slice-to-volume registration for MRI-guided transperineal prostate biopsy. *Int. J. CARS*, 10(1):563–572, 2015.

[XLH+15] Ping Xu, Hong Luo, Guang-Lei Huang, Xin-Hai Yin, Si-Yang Luo, and Ju-Kun Song. Exposure to ionizing radiation during dental X-rays is not associated with risk of developing meningioma: A meta-analysis based on seven case-control studies. *PLoS One*, 10:1–12, 2015.

[XW08] Lin Xu and Justin W. L. Wan. Real-time intensity-based rigid 2d-3d medical image registration using RapidMind Multi-core Development Platform.

In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5382–5385, Aug 2008.

[XW15] Robert Xu and Graham A. Wright. *Statistical Atlases and Computational Models of the Heart - Imaging and Modelling Challenges: 5th International Workshop, STACOM 2014, Held in Conjunction with MICCAI 2014, Boston, MA, USA, September 18, 2014, Revised Selected Papers*, chapter Registration of Real-Time and Prior Imaging Data with Applications to MR Guided Cardiac Interventions, pages 265–274. Springer International Publishing, Cham, 2015.

[Z+10] Darko Zikic et al. Linear intensity-based image registration by Markov random fields and discrete optimization. *Medical Image Analysis*, 14(4):550–562, 2010.