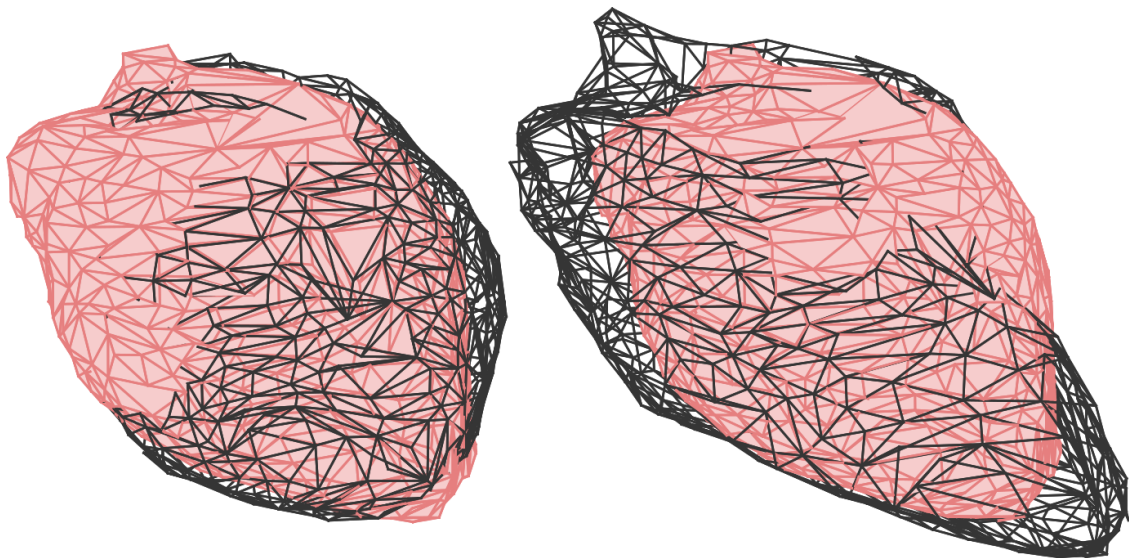




CHALMERS
UNIVERSITY OF TECHNOLOGY



Automated construction of statistical deformation models for non-rigid registrations of the pericardium and hippocampus

Master's thesis in Complex Adaptive Systems

MÅNS LARSSON

MASTER'S THESIS 2015:EX014

**Automated construction of statistical deformation models
for non-rigid registrations of the pericardium and
hippocampus**

Måns Larsson



Department of Signals and Systems
Image Analysis and Computer Vision
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2015

Automated construction of statistical deformation models for non-rigid registrations of the pericardium and hippocampus

Måns Larsson

© Måns Larsson, 2015.

Examiner: Fredrik Kahl, Department of Signals and Systems

Master's Thesis 2015:EX014
Department of Signals and Systems
Image Analysis and Computer Vision
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Visualization of a statistical deformation model of pericardium. The red surface structures are unaltered pericardiums whereas the black structures are pericardiums that have been transformed, downwards to the left and upwards to the right, in the main mode of variation described by the model.

Typeset in \LaTeX
Gothenburg, Sweden 2015

Automated construction of statistical deformation models for non-rigid registrations of the pericardium and hippocampus

MÅNS LARSSON

Department of Signals and Systems

Chalmers University of Technology

Abstract

Image segmentation plays a crucial role in many medical applications such as treatment planning, diagnosis and quantification of tissue volume. A key part of several segmentation methods is image registration where spatial correspondences between two images described by a transformation is calculated. The fact that local warpings often are needed to fully describe these correspondences makes this a difficult task that many standard methods fail to solve in a time efficient and reliable way.

In this thesis, a feature-based registration method, Shapereg, is presented. The method uses statistical deformation models to describe the non-rigid part of the transformation. The models are built using deformation fields from pair-wise registrations of the atlas images. The Shapereg method was applied to two segmentation tasks, pericardium segmentation in CT images of the heart and hippocampus segmentation in MRI images of the brain.

The results from the segmentations were compared to two other methods, one feature-based and one intensity-based. The Shapereg method outperformed the other two methods for the pair-wise segmentation of pericardium and produced comparable results for the pair-wise segmentations of the hippocampus as well as the multi-atlas segmentations.

Keywords: medical image analysis, computer vision, statistical deformation models, statistical shape models, feature-based registration, pericardium segmentation, hippocampus segmentation, multi-atlas segmentation, robust linear regression

Acknowledgements

I would like to start of by thanking my supervisor Fredrik Kahl for his guidance and helpful input during this project. Also thanks to everyone in the Computer Vision Group at Chalmers, especially Jennifer Alvéen for her support and unwavering patience for all my stupid questions during the beginning of the project. A special thanks to my office mate Fredrik Elofsson for his help with the report and good company.

Måns Larsson
Gothenburg, March 2015

Contents

List of Figures	x
List of Tables	xiii
1 Introduction	1
1.1 Problem Formulation	2
1.2 Data Sets	2
1.3 Proposed Solution	3
1.4 Related Work and Contributions	4
1.5 Structure of the Report	5
2 Theory	7
2.1 Image Registration	7
2.1.1 Transformation models	7
2.1.1.1 Affine Transformation	7
2.1.1.2 Non-rigid Transformation	8
2.2 Statistical Shape Models	8
2.2.1 Statistical Deformation Models	10
2.3 Robust Linear Regression	10
2.3.1 L_1 -norm	11
2.3.2 Huber Loss Function	11
2.3.3 Iteratively Reweighted Least Squares	11
3 Method	13
3.1 Feature-Based Registration	14
3.1.1 Detection	14
3.1.2 Description	15
3.1.3 Matching	15
3.2 Affine Registration	15
3.3 Offline Building of the Models	15
3.3.1 Training Set	16
3.3.1.1 Inter-Atlas Registration	16
3.3.1.2 Extra-Atlas Registration	17
3.3.2 Alignment of the Training Set	17
3.3.3 Downsampling	17
3.3.4 Statistical Deformation Models	18
3.4 Online Segmentation Using an SDM	18
3.4.1 Feature Detection	18
3.4.2 Affine Transformation Initialization	19

3.4.3	Transformation Estimation	19
3.4.3.1	Feature Selection	19
3.4.3.2	Interpolation	19
3.4.3.3	Optimization	19
3.4.3.4	Creation of the Warp Field	21
3.4.4	Transformation of the Source Segmentation	21
3.4.5	Multi-Atlas Segmentation	21
4	Implementation and Baseline Methods	23
4.1	Implementation	23
4.1.1	Feature Extraction	23
4.1.2	Affine Registration	23
4.1.3	Offline Registration of the Atlas Set	23
4.1.4	Online Segmentation Using an SDM	23
4.2	Baseline Methods	24
4.2.1	Baseline 1: Feature-Based	24
4.2.2	Baseline 2: Intensity-Based	24
5	Results	25
5.1	Segmentation	25
5.1.1	Pair-Wise Atlas Segmentation	25
5.1.2	Multi-Atlas Segmentation	25
5.2	Model Evaluation	27
5.2.1	Effect of Training Set Size	32
5.3	Feature Matching	35
5.4	Runtimes	35
6	Discussion	37
6.1	Creation of Training Set	37
6.2	Segmentation	38
6.3	Feature Matching	39
6.4	Runtimes	40
6.5	Model	40
6.6	Future Work	41
7	Conclusion	43
	Bibliography	47

List of Figures

1.1	A slice of a CT image of the heart with corresponding manual segmentation. To the left is a slice of the CT image with the pericardium marked with a red line and to the right is the corresponding segmentation where pixels inside pericardium are white and pixels outside pericardium are black. . . .	2
1.2	A slice of an MRI image of the brain with corresponding segmentation. To the left is a slice of the MRI image with the hippocampus marked with a red line and to the right is the corresponding segmentation where pixels inside the hippocampus are white and pixels outside the hippocampus are black.	3
2.1	Comparison of different loss functions for fitting a straight line to a set of 500 data points. The points are simulated on a line described by the equation $y = .5x - .7$ with small uniformly distributed randomized deviations on the interval $]0 \ 0.1]$. 10 % of the points are outliers with a randomized y-value uniformly distributed on the interval $1 \leq y < 3$	11
3.1	Example of matched features from two CT images of the heart. The features with the same color have been matched to each other. Note that this is only one slice of a 3D image, the features included in the figure are located at a maximum distance of 5 voxels in the direction perpendicular to the slices. .	14
5.1	Cumulative variance, V_T , described by the statistical shape model as a function of number of modes included, k , see Equation (2.10). These results are from all the 380 models created during the leave-one-out experiments done on the heart data set . The blue lines are individual results from each model and the red line is the mean of all models. The green line is located at $V_T = .98$	28
5.2	Cumulative variance, V_T , described by the statistical shape model as a function of number of modes included, k , see Equation (2.10). These results are from all the 870 models created during the leave-one-out experiments done on the brain data set . The blue lines are individual results from each model and the red line is the mean of all models. The green line is located at $V_T = .98$	28
5.3	Histogram over the number of modes, k , needed to be included to describe 98 % of the variance of the statistical shape model. Note that this number equal the numbers of parameters needed to be optimized during the online process. These results are from all the 380 models created during the leave-one-out experiments done on the heart data set	29

5.4	Histogram over the number of modes, k , needed to be included to describe 98 % of the variance of the statistical shape model. Note that this number equal the numbers of parameters needed to be optimized during the online process. These results are from all the 970 models created during the leave-one-out experiments done on the brain data set	29
5.5	Visualization of the statistical deformation model for atlas image 1 in the heart data set . The topmost image is an unaltered slice of the atlas image, the rest of the images has been generated by varying the two main modes of variation between $-3\sqrt{\lambda_i}$ (to the left) and $3\sqrt{\lambda_i}$ (to the right) and applying the warping to the slice of image 1. In row two mode 1 is varied and in row three mode 2 is varied.	30
5.6	Visualization of the statistical deformation model for atlas image 1 in the brain data set . The topmost image is an unaltered slice of the atlas image, the rest of the images has been generated by varying the two main modes of variation between $-3\sqrt{\lambda_i}$ (to the left) and $3\sqrt{\lambda_i}$ (to the right) and applying the warping to the slice of image 1. In row two mode 1 is varied and in row three mode 2 is varied.	31
5.7	Mean euclidean distance error \mathbf{D} in mm as defined in Equation (5.1). \mathbf{D} is shown as a function of the training set size n for the heart data set . For n larger than 19 deformation fields from extra-atlas registrations were included. Atlas image 1 was used as target and atlas images 2, . . . , 9 were used as sources, the \mathbf{D} shown in the plot is the mean of the 8 results.	32
5.8	Mean euclidean distance error \mathbf{D} in mm as defined in Equation (5.1). \mathbf{D} is shown as a function of the training set size n for the brain data set . Atlas image 1 were used as target and atlas images 2,3 and 4 were used as sources, the \mathbf{D} shown in the plot is the mean of the 3 results.	33
5.9	Segmentation results when transforming the segmentation using the recreated fields as described in Section 5.2. The Jaccard index \mathbf{J} is shown as a function of the training set size n for the heart data set . For n larger than 19 deformation fields from extra-atlas registrations were included. Atlas image 1 were used as target and atlas images 2, . . . , 9 were used as sources, the \mathbf{J} shown in the plot is the mean of the 8 segmentations.	33
5.10	Segmentation results when transforming the segmentation using the recreated fields as described in Section 5.2. The Jaccard index \mathbf{J} is shown as a function of the training set size n for the brain data set . Atlas image 1 was used as target and atlas images 2,3 and 4 were used as sources, the \mathbf{J} shown in the plot is the mean of the 3 segmentations.	34
6.1	Error-bar plot of the Jaccard index for the transformations in the heart training set as a function of target image number. The markers denote the means and the bars denote the intervals that are within one standard deviation.	38
6.2	Visualization of the multi-atlas voting on the heart data set for all three methods. The images show all the pair-wise segmentations added together as well as the edge of the true target segmentation. From left to right the result of Shapereg, Featreg and Niftyreg is shown.	39
6.3	Visualization of the multi-atlas voting on the brain data set for all three methods. The images show all the pair-wise segmentations added together as well as the edge of the true target segmentation. From left to right the result of Shapereg, Featreg and Niftyreg is shown.	39

List of Tables

3.1	Jaccard indices of the inter-atlas registrations for the two data sets, mean ± 1 standard deviation.	17
5.1	Jaccard indices of the pair-wise atlas segmentation of the pericardium for the three methods, <i>mean ± 1 standard deviation</i>	25
5.2	Jaccard indices of the pair-wise atlas segmentation of the hippocampus for the three methods, <i>mean ± 1 standard deviation</i>	26
5.3	Jaccard indices of the multi-atlas segmentation of the pericardium for the three methods, <i>mean ± 1 standard deviation</i>	26
5.4	Jaccard indices of the multi-atlas segmentation of the hippocampus for the three methods, <i>mean ± 1 standard deviation</i>	26
5.5	Mean euclidean distance error D in mm as defined in Equation (5.1) for the two data sets. The result is shown as <i>mean ± 1 standard deviation</i> for all recreated fields.	27
5.6	Jaccard index when transforming the segmentation using the recreated fields as described in Section 5.2 for the two data sets. The result is shown as <i>mean ± 1 standard deviation</i> for all recreated fields.	27
5.7	Mean euclidian distance error in mm between target feature points and transformed source feature points. The result is shown as <i>mean ± 1 standard deviation</i> for all recreated fields.	35
5.8	Mean runtime for the non-rigid part of a pair-wise registration in seconds. .	35

1

Introduction

Diagnostic imaging is an important and valuable tool in medicine. The ability to map the anatomy of a subject is often a critical component in diagnosis and treatment planning. Image segmentation is the process of partitioning an image into multiple segments delineating anatomical structures and other regions of interest. It plays a crucial role in applications such as quantification of tissue volumes, diagnosis, localization of pathology, study of anatomical structure, treatment planning, and computer-integrated surgery, for more information see [1].

Segmentation of the pericardium can be used to determine if fat cells are located inside the pericardial sac, so called epicardial fat. Several studies have during recent years shown a relationship between the amount of epicardial fat and several severe cardiac diseases, see [2]. Pericardium is a barely visible line in a CT scan, hence manual delineation requires a lot of knowledge and is time consuming. Therefore there is a great need of reliable automatic methods for pericardium segmentation. The hippocampus structure is among the first parts of the brain affected by Alzheimer's disease. Segmentation of the hippocampus hence constitutes a useful diagnostic tool for detection of Alzheimer's disease, for details see [3].

Many segmentation methods involve image registration which is the process of establishing a spatial correspondence between two images, i.e. finding a transform that maps pixels in one image to corresponding points in the other image. This can be a difficult task since most of the images in the field of medical imaging are in 3D and hence contains a lot of data making the registration process computation heavy. In addition a lot of medical images are noisy and there can be large inter-subject variations. Registration methods are usually divided into two groups: feature-based methods where the transform is estimated based on the location and properties of specific feature points (contours, surfaces or curves) and intensity-based methods where the transform estimation is based on the pixel intensities of the two images [4, 5].

In this thesis, a feature-based registration method using prior information of the subject is presented. The method uses available atlas images to build statistical deformation models in order to perform image registration. The method is compared to two other methods, a standard feature based method and an intensity based method. The three methods are applied to two segmentation tasks, pericardium segmentation in CT images of the heart and hippocampus segmentation in MRI images of the brain, and are compared with respect to segmentation accuracy and speed.

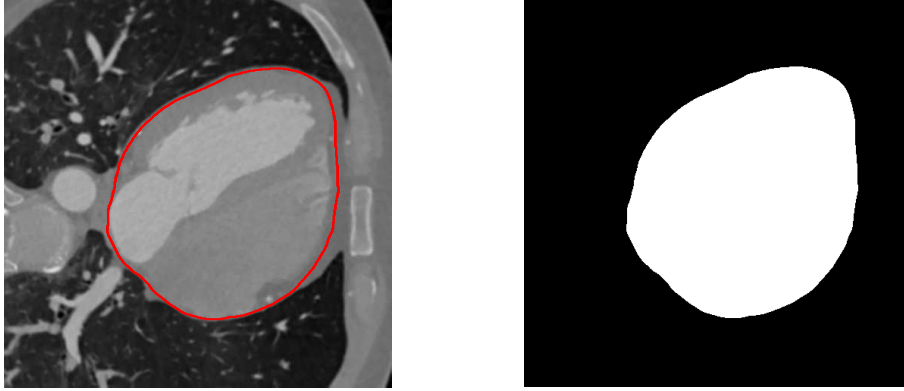


Figure 1.1: A slice of a CT image of the heart with corresponding manual segmentation. To the left is a slice of the CT image with the pericardium marked with a red line and to the right is the corresponding segmentation where pixels inside pericardium are white and pixels outside pericardium are black.

1.1 Problem Formulation

The main goal with this thesis is to be able to as precisely as possible determine a labeling of the pericardium in CT-images of the heart as well as the hippocampus in MR-images of the brain.

An image can be described as a set of voxels (3D pixels) P with corresponding intensities i_p , $p \in P$. It can hence be regarded as a function that maps the coordinates of all voxels to an intensity, $I : \Omega \rightarrow \mathbb{R}$ where Ω is the set of all voxel coordinates, $\Omega \subset \mathbb{R}^3$.

A segmentation of an image means to mark what voxels belong to the same anatomical structure. In this thesis a voxel has the label 1 if it is located inside pericardium and 0 if it is located outside pericardium. The segmentation can hence also be seen as a function mapping voxel coordinates to the set of allowed labels $L : \Omega \rightarrow \mathbb{L}$, in this case $\mathbb{L} = \{0, 1\}$.

Manual segmentations have been performed by experts and these are referred to as the gold standard. Examples of manual segmentations are shown in Figures 1.1 and 1.2. The estimated segmentations are evaluated by how similar they are to the corresponding gold standard. The similarity measure used in this thesis is the Jaccard index defined as

$$J(A, B) := \frac{|A \cap B|}{|A \cup B|}, \quad (1.1)$$

where A and B are two sets. Let the estimated segmentation be described as the set of voxels labeled 1 and denoted Γ_e . If the gold standard segmentation are described in the same way and denoted Γ_m the similarity between the labels can be quantized as $J(\Gamma_e, \Gamma_m)$. This similarity measure will have a value between 0 and 1, where 1 means that the segmentations are identical and 0 means that the segmentations have no overlapping elements.

1.2 Data Sets

In this thesis two different data sets are used. The first one consists of 50 CT, or Computed Tomography, images of the heart. CT is a method that produces virtual slices of



Figure 1.2: A slice of an MRI image of the brain with corresponding segmentation. To the left is a slice of the MRI image with the hippocampus marked with a red line and to the right is the corresponding segmentation where pixels inside the hippocampus are white and pixels outside the hippocampus are black.

a scanned object using X-rays. Three-dimensional images of specific areas of the object are then generated using digital geometry processing on a large series of two-dimensional radiographic images taken around a single axis of rotation. The images used were collected as a part of the Swedish Cardiopulmonary Bioimage Study (SCAPIS) which is a research study that aims to build an extensive medical knowledge bank consisting of CT, MR and ultrasound images from 30 000 healthy subjects between the age of 50 and 64 years old. The goal is to improve the ability to predict and prevent the risk of cardio-pulmonary diseases [6].

Before the acquisition of the CT images the patients were given contrast material which enhances the visibility of the blood and the tissue. This makes the areas around the left atrium and ventricle appear more detailed as well as making the pericardium easier to distinguish. The CT images in the SCAPIS data set have resolutions ranging from $512 \times 512 \times 342$ to $512 \times 512 \times 458$ voxels with voxel sizes between $0.3320 \times 0.3320 \times 0.3000$ and $0.4297 \times 0.4297 \times 0.3000$. For 20 of these 50 CT images a manual labeling of the pericardium has been done by an expert, an example of this can be seen in Figure 1.1. The delineation was done for every 10th slice of the image in all three viewing directions. These were interpolated into a final segmentation that was approved by the same expert. These manual segmentations will be referred to as the gold standard.

The second data set consists of 30 images of the brain. The images were taken using Magnetic Resonance Imaging (MRI) and the test subject were young adults. A manual labeling of 83 regions has been done for these 30 images according to the protocol described in [7]. In this thesis only the labels corresponding to the left and the right hippocampus are used and the labeling is again referred to as the gold standard. An example of an MRI image and the corresponding labeling of the hippocampus can be seen in Figure 1.2. The resolution of the images is between $165 \times 198 \times 155$ and $195 \times 199 \times 175$ voxels with voxel sizes ranging from $0.9350 \times 0.9350 \times 0.9350$ to $0.9375 \times 0.9375 \times 0.9375$ mm³. The data can be found online at <http://www.brain-development.org>.

1.3 Proposed Solution

In this thesis a feature-based registration method using prior information of the subject is proposed as solution to the segmentation problem formulated in Section 1.1. The method is

divided into an offline (performed once for each atlas set) part and an online (performed for each segmentation) part. During the offline process accurate registration are done between all the atlas images using a reliable but time-consuming method. The transformations acquired during these registrations are then used to build statistical deformation models that represent the warping for each atlas image. During the online process the atlas images are registered to the target image by firstly estimating an affine transformation using a robust feature-based method. Secondly a non-rigid transformation is estimated using the statistical deformation models created during the offline process. Here the statistical deformation models are used to represent the non-rigid transformations using few parameters. The parameters are then optimized using feature locations.

1.4 Related Work and Contributions

Feature-based multi-atlas methods have previously been used for segmentation of the pericardium by Alvén [8] and Norlén [9]. Norlén used a feature based registration method optimizing a free form deformation model based on cubic B-splines. This method is similar to the feature-based baseline used in this work. Alvén presented a method called Überatlas where the atlases are first co-registered and clusters of feature points present in the atlases are created. This makes it possible to match features in the target image with the Überatlas feature clusters in an attempt to speed up the process.

In the field of cardiac and pericardium segmentation Kirsli et al. [10] and Shahzad et al. [11] have both developed multi-atlas based methods for fully automatic cardiac segmentation. Both methods achieved good segmentation results and are considered state-of-the-art, although the registrations are intensity-based which usually are very time consuming. An effort to create a less time-consuming method was made by Dey et al. [12] where the atlases are co-registered beforehand and the target image is only registered to a single atlas. The labels are then propagated from all the atlases and combined using a decision weighting function.

Statistical shape models have previously been used as a tool for segmentation and structure analysis in medical imaging. Rueckert et al. [13] used statistical deformation models to construct a model of the average shape and variability of the brain. This was done by first performing non-rigid co-registration of the atlases using a B-spline representation of the non-rigid transformation. Statistical analysis was then done on the control points describing the transformation. Loeckx et al. presented in [14] a voxel-based non-rigid registration algorithm for subtraction of two-dimensional X-ray images of the thorax where the spline parameters of the non-rigid transformation were constrained by a statistical deformation model.

The method developed in this thesis is partly based on the work by Norlén and Alvén although the non-rigid registration method used is based on statistical deformation models. It also differs from previous work in several aspects. One of these are that the parameters of the statistical deformation models are optimized using feature point locations. The fact that feature points are used instead of image intensity matching results in a reduction of computation time. Estimating the non-rigid transformation using a low-dimensional statistical deformation model also introduces robustness in the method counteracting the effect of possible outliers in the feature points. It is also shown how including unlabeled data in the training set when building the statistical deformation models can improve segmentation results.

1.5 Structure of the Report

The remainder of this thesis is structured as follows. In Chapter 2 the theoretic concepts that this thesis is built on are presented. The main parts of this chapter are image registration, statistical deformation models and robust linear regression. If the reader is already familiar with these concept this chapter can easily be skipped. Chapter 3 describes the method and is divided into three parts, firstly the feature extraction and matching as well as the affine registration method is described, secondly the offline part of the method is described and lastly the online part. In Chapter 4 implementation details such as programming languages and parameter settings are presented. In addition this chapter also describes the two baseline methods compared to the method developed in the thesis. The results are presented in Chapter 5 and these results are discussed in Chapter 6. The thesis is concluded in Chapter 7.

2

Theory

2.1 Image Registration

Image registration is the process of establishing a spatial correspondence between two images, i.e. finding a transform \mathbf{T} that maps points in one image to corresponding points in the other image. Which points that are corresponding in the two images differs depending on the application. Given a target image I_t , also referred to as reference or fixed image and a source image I_s , also referred to as float or moving image, the image registration problem can be formulated as minimizing the cost function according to

$$\arg \min_{\mathbf{T}} \rho(I_t, \mathbf{T} \circ I_s) \quad (2.1)$$

where ρ is a cost function that quantifies the dissimilarity between two images, in this case the target image and $\mathbf{T} \circ I_s$ which should be interpreted as the transformed source image.

There are a lot of different options for the cost function and the choice is generally application dependent. The different types of cost functions are usually divided into two groups: feature-based functions where the cost is based on the location and properties of specific feature points (contours, surfaces or curves) and intensity-based functions where the cost is based on the pixel intensities of the two images [4, 5].

2.1.1 Transformation models

The transformation $\mathbf{T} : \mathbb{R}^d \mapsto \mathbb{R}^d$, where d is the dimension of the image, can be described in various ways. Usually one distinguishes between affine or linear transformation models and non-rigid or elastic transformation models. It is also common to combine the two, i.e. start with estimating the affine transformation that minimizes (2.1) and then add a non-rigid part to the transformation model. In this way one reduces the risk of getting stuck in a local minima of the cost function.

2.1.1.1 Affine Transformation

An affine transformation is a global transformation that preserves straight lines, points and planes. It is a composition of a linear map and translation and can be represented by a multiplication by a matrix \mathbf{A} and an addition of a translation vector \mathbf{t} in the following way

$$\mathbf{y} = \mathbf{Ax} + \mathbf{t}. \quad (2.2)$$

A more compact way to represent an affine transform is to use a so called augmented matrix. This enables the transform to be performed with one matrix multiplication in the

following way

$$\begin{pmatrix} \mathbf{y} \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} = \mathbf{T}^{\mathbf{A}} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \quad (2.3)$$

where the augmented matrix is denoted $\mathbf{T}^{\mathbf{A}}$, see [4, 5].

2.1.1.2 Non-rigid Transformation

Non-rigid transformations allow local warping of the source image to align with the target image which is often necessary in medical imaging. When choosing a non-rigid transformation model there is often a trade-off between computational efficiency and richness of description. This is decided by the number of parameters or degrees of freedom of the model.

A straightforward way to describe a non-rigid transformation is by using a deformation field where for each voxel in the target image the coordinates of the corresponding voxel in the source image is saved. This can be represented as

$$\mathbf{y} = \mathbf{u}(\mathbf{x}) \quad (2.4)$$

where the deformation field $\mathbf{u} : \Omega \mapsto \mathbb{R}^d$ is a function on the domain Ω that is the voxel coordinates of the target image, $\Omega \subset \mathbb{N}^d$. Using this model, neighboring voxels have no information about each other and there is a risk that the topology of the object to register is broken. Hence, usually some kind of regularization is needed when optimizing the deformation field [15, 16].

A common parameterized model used in medical imaging is a free-form deformation model based on cubic B-splines. This model was proposed by Rueckert et al. in 1999 [17] and the basic idea is that a grid of control points for the B-splines are superimposed on the image. The transformation is then optimized by adjusting the position of the control points. The composition of the B-splines ensures that the transformation will be twice continuously differentiable.

2.2 Statistical Shape Models

The basic idea behind statistical shape models is to use prior knowledge of a certain class of objects to get a better segmentation. A statistical shape model describes the variability of a class of objects and is created using a set of examples, a training set. The model is able to describe new shapes if they are of the same class of objects and within the variability of the training set. In this way a model with rather few parameters will be created to describe a class of objects. Statistical shape models are an effective and widely used tool in image segmentation and image analysis [18].

The shapes used to create the models are easily distinguishable features e.g. landmarks, curves or surfaces. The simplest case is when feature points are used to create the model. In this case step one is to find a set of correctly annotated landmarks that are present in each of the objects in the training set. This step can be problematic since it is hard to perform automatically and time consuming to do manually [13, 18].

The next step is to align the landmarks of all the objects in the training set. This can be done by using for instance so called Procrustes analysis or by estimating an affine

transformation between each training image that minimizes the distance between the corresponding landmarks by some measure [18, 19].

When the training set has been aligned the variability of the model is captured using principal component analysis (PCA). Suppose that each shape i in the training set is described by m landmark points whose coordinates are assembled in a shape vector \mathbf{X}_i of length $3m$ (for the 3-dimensional case). The distribution of \mathbf{X} for the different shapes will now be approximated by a linear model according to

$$\mathbf{X} = \hat{\mathbf{X}} + \Phi \mathbf{b} \quad (2.5)$$

where $\hat{\mathbf{X}}$ is the average of the n shape vectors in the training set,

$$\hat{\mathbf{X}} = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \quad (2.6)$$

and \mathbf{b} is a vector containing the model parameters. The matrix Φ is formed by the eigenvectors (as columns) of the covariance matrix \mathbf{S}

$$\mathbf{S} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{X}_i - \hat{\mathbf{X}})(\mathbf{X}_i - \hat{\mathbf{X}})^T. \quad (2.7)$$

The principal modes of variation of the shape vectors can now be calculated as the eigenvectors Φ_i and corresponding eigenvalues λ_i (sorted in descending order of magnitude) of \mathbf{S} . Any new shape described by the shape vector \mathbf{Y} within the population group can now be approximated using (2.5) where the parameters \mathbf{b} are given by

$$\arg \min_{\mathbf{b}} \|\mathbf{Y} - \hat{\mathbf{X}} - \Phi \mathbf{b}\| \quad (2.8)$$

where $\|\cdot\|$ denotes a suitable norm. Here \mathbf{Y} and $\hat{\mathbf{X}}$ are column vectors of length $3m$ whereas Φ is of size $3m \times n$ and \mathbf{b} is of size $n \times 1$. The variance described by each eigenvector is equal to the corresponding eigenvalue, hence there is no need to include the eigenvectors that have corresponding eigenvalues equal to zero. If $t < \min(3m, n-1)$ is the number of non-zero eigenvalues the total variance λ_T of the training set can thus be calculated as

$$\lambda_T = \sum_{i=1}^t \lambda_i. \quad (2.9)$$

To reduce the dimensionality of the model, a variance threshold V_t can be set as a fraction of how much of the total variance of the training set that should be included in the model. This is done by only including the first k eigenvectors Φ_i in Φ such that the following statement holds

$$V_t > \frac{1}{\lambda_T} \sum_{i=1}^k \lambda_i. \quad (2.10)$$

Limits to the parameters \mathbf{b} can also be applied by assuming that the points in the shape vectors follow a multidimensional Gaussian distribution. Then the majority of the population will be within three standard deviations of the mean and the parameters b_i can be restricted by

$$|b_i| \leq 3\sqrt{\lambda_i}. \quad (2.11)$$

For further details and a thorough derivation, see [19, 13]

2.2.1 Statistical Deformation Models

The difference between statistical deformation models (SDMs) and statistical shape models (SSMs) is that the statistical analysis is carried out on the deformation fields that describe a correspondence between two anatomies instead of landmark points of an object. This means that the models are not limited to only one object and can describe the variability of a structure on various levels of detail. There are also a lot of ways to calculate the deformation fields for the training set using various non-rigid registration methods [13].

2.3 Robust Linear Regression

The presence of outliers, i.e. observations that do not follow the pattern of other observations, in data related to medical image analysis is a common phenomenon. The least square estimate for linear regression is, when this is the case, inefficient and can produce unwanted results as the parameters calculated using least squares will be heavily affected by the outlier data points.

Given a data set consisting of the dependent variable $\mathbf{y} = (y_1, \dots, y_n)^\top$ and the set of independent variables \mathbf{x}_{ij} , $i = 1 \dots n$ and $j = 1 \dots p$. The goal is to find a set of parameters $\boldsymbol{\beta}$ that approximates

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}, \text{ where } \mathbf{X} = \begin{pmatrix} \mathbf{x}_1^\top \\ \vdots \\ \mathbf{x}_n^\top \end{pmatrix}. \quad (2.12)$$

If $p > n$ there are more independent variables than parameters. The system is then overdetermined and can not be solved exactly in most cases. This means that $\boldsymbol{\beta}$ needs to be approximated in a way that "best" fits the data. Using ordinary least squares, the approximation of $\hat{\boldsymbol{\beta}}$ is given by solving

$$(\mathbf{X}^\top \mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}^\top \mathbf{y}. \quad (2.13)$$

This approximation will minimize the quadratic residual, i.e.

$$S(\boldsymbol{\beta}) = \sum_{i=1}^n |y_i - \sum_{j=1}^p \mathbf{x}_{ij}\boldsymbol{\beta}_j|^2 = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 \quad (2.14)$$

which would make the effect of data outliers great. There are various methods that deal with this problem and provide an efficient estimate for data containing outliers. In this thesis methods insensitive to outliers will be referred to as robust.

A straightforward approach is to use an alternative norm to L_2 , minimizing

$$S(\boldsymbol{\beta}) = \sum_{i=1}^n \phi \left(\left| y_i - \sum_{j=1}^p \mathbf{x}_{ij}\boldsymbol{\beta}_j \right| \right), \quad (2.15)$$

where ϕ is a loss function, instead of (2.14) [20, 21]. Below two different choices of norms or loss functions are described and in Figure 2.1 they have been applied to a simulated data set. The figure shows that the presence of outliers affects the L_2 -solution greatly while the alternative methods are less influenced.

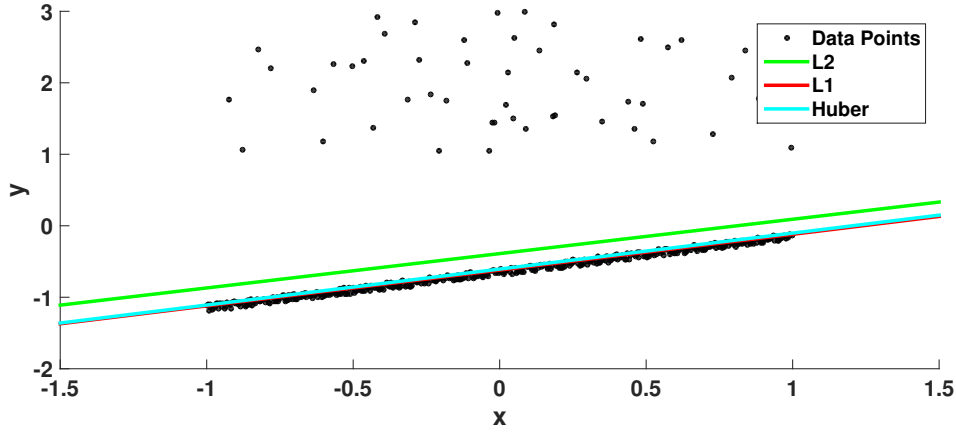


Figure 2.1: Comparison of different loss functions for fitting a straight line to a set of 500 data points. The points are simulated on a line described by the equation $y = .5x - .7$ with small uniformly distributed randomized deviations on the interval $]0 \ 0.1]$. 10 % of the points are outliers with a randomized y -value uniformly distributed on the interval $1 \leq y < 3$.

2.3.1 L_1 -norm

The L_1 -norm problem can be formulated as minimizing

$$S(\beta) = \sum_{i=1}^n |y_i - \sum_{j=1}^p x_{ij}\beta_j|. \quad (2.16)$$

This problem is also referred to as *least absolute deviations* or *least absolute errors* and can be solved by using for example simplex-based methods (or other linear programming techniques) or iteratively reweighted least squares [22, 23].

2.3.2 Huber Loss Function

The Huber loss functions combines the properties of a quadratic loss function (L_2 -norm) and an absolute loss function (L_1 -norm). It is defined as

$$\phi(r) = \begin{cases} \frac{1}{2}r^2 & |r| \leq \delta \\ \delta(|r| - \frac{1}{2}\delta) & \text{otherwise,} \end{cases} \quad (2.17)$$

where δ is a parameter and

$$r = y - \sum_{j=1}^p x_j\beta_j. \quad (2.18)$$

For further details and motivation see [24].

2.3.3 Iteratively Reweighted Least Squares

Iteratively Reweighted Least Squares (IRLS) is a method used to minimize other loss functions of the residuals than the quadratic one (L_2 -norm). It is used to solve the following problem

$$\arg \min_{\beta} \sum_{i=1}^n \phi(y_i - \mathbf{X}_i\beta), \quad (2.19)$$

where ϕ is a robust loss function. This is done by iteratively solving a weighted least squares problem minimizing

$$\boldsymbol{\beta} = \arg \min_{\boldsymbol{\beta}} \sum_{i=1}^n w_i^{(t)} |y_i - \mathbf{X}_i \boldsymbol{\beta}|^2 \quad (2.20)$$

where w_i^t are weights updated each iteration. The solution to this problem is efficiently calculated by

$$\boldsymbol{\beta} = (\mathbf{X}^\top \mathbf{W}^{(t)} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W}^{(t)} \mathbf{y} \quad (2.21)$$

where $\mathbf{W}^{(t)}$ is a diagonal matrix containing the weights. Denoting the residuals as $r_i = |y_i - \mathbf{X}_i \boldsymbol{\beta}^{(t)}|$ the weights are chosen according to

$$w_i = \frac{1}{2r_i} \frac{\partial \phi(r_i)}{\partial r_i}. \quad (2.22)$$

This expression is deduced from the fact that equation (2.21) minimizes $w_i r_i^2$, setting the derivatives of this and $\phi(r_i)$ with respect to r_i equal according to

$$\frac{\partial \phi(r_i)}{\partial r_i} = \frac{\partial (w_i r_i^2)}{\partial r_i} \quad (2.23)$$

will hence lead to a minimization of $\phi(r_i)$. The algorithm is summarized in Algorithm 1, for more details see [25, 26].

```

Initialize  $\mathbf{W} = \mathbf{I}$ 
while not converged do
     $\boldsymbol{\beta} = (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W} \mathbf{y}$ 
     $r_i = |y_i - \mathbf{X}_i \boldsymbol{\beta}|$ 
     $\mathbf{W}_{ii} = \phi'(r_i)/(2r_i)$ 
end

```

Algorithm 1: Iteratively Reweighted Least Squares

It has been shown that the algorithm converges as long as the constructed function $\psi(x) := \phi(\sqrt{x})$ is concave and differentiable except at a finite number of points [27]. This is the case for both the L_1 -norm and the Huber loss function.

3

Method

The method used in this thesis can be divided into an offline and an online part. The offline part is only performed once for each atlas data set. During this part the statistical deformation models (SDMs) are created for each atlas image. The online part is where the SDMs are used to register target images to the source images of the atlas set. This part is performed once for each new segmentation and is therefore called online. The work-flow of both parts will be shown in bullet form below.

Offline building of the models:

- Perform pair-wise registrations between all atlas images, resulting in transformation fields $\mathbf{u}_{ij}(\mathbf{x})$, using a slow but accurate registration method.
- For each atlas image i align all the transformations from that image by estimating an affine transformation and removing the effect of it from the transformation fields $\mathbf{u}_{ij}(\mathbf{x})$, resulting in a set of warp fields $\mathbf{w}_{ij}(\mathbf{x})$.
- Downsample the warp fields.
- Build the SDMs for each atlas image using the downsampled warp fields.

Online segmentation using an SDM:

- Extract features in the target and source images.
- Estimate an affine transformation used as initialization.
- Select the features located closest to the source segmentation surface.
- Interpolate in the SDM to get the statistical model of the warping at said feature locations.
- Calculate the model parameters and affine transformation using an alternating optimization method.
- Recreate the warp field using the parameters and the SDM
- Use the affine transformation as well as the warp field to transform the source segmentation to the target.

The rest of the method chapter is divided into four main parts. Firstly, a description of the basics of the feature-based method used in this work is given. Secondly, a short description of the method for estimating an affine transformation for the registration is given. Then, the offline procedure will be described in more detail. Lastly, a description of the online procedure is given. Note that implementation details will be represented in Chapter 4 and not here.

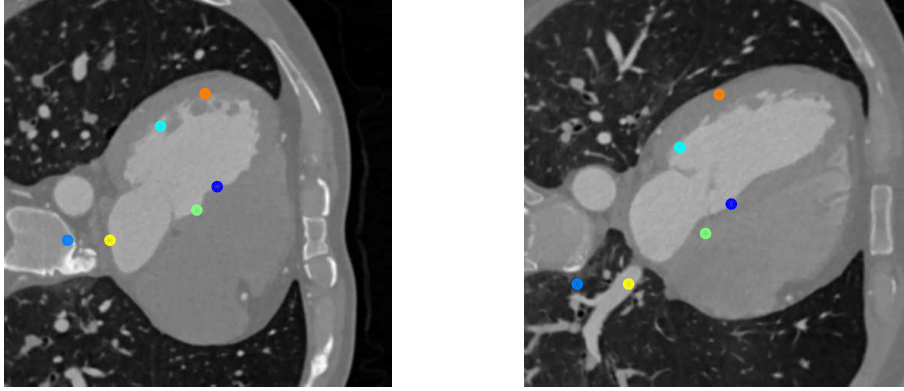


Figure 3.1: Example of matched features from two CT images of the heart. The features with the same color have been matched to each other. Note that this is only one slice of a 3D image, the features included in the figure are located at a maximum distance of 5 voxels in the direction perpendicular to the slices.

3.1 Feature-Based Registration

In feature-based registration correspondences between image features such as corners, lines and contours are used to perform the registration. This differs from intensity-based methods where the image intensities are used to evaluate the similarities between the images. A feature-based registration algorithm is usually carried out in three steps

- **Detection:** Feature point locations are extracted from the target and the source image.
- **Description:** For each feature point a descriptor is calculated based on the properties of the neighborhood of the feature point.
- **Matching:** Using the descriptors the feature points in the target image are matched to the feature points in the source image.

The goal is to have feature points describing the corresponding anatomical structure in the target and the source images matched with each other. The matched feature points can now be used to align the target and the source images using equation (2.1) with a cost function based on the distance between matched feature point locations after transformation. In Figure 3.1 an example of matched features from two CT images of the heart is shown.

3.1.1 Detection

The method for feature detection used in this work is called Scale Invariant Feature Transform (SIFT) and was proposed by Lowe [28]. The SIFT feature point detector uses a so called staged filtering approach and hence searches for feature points at different scales of the image. The method is also invariant to image scaling, rotation and translation.

3.1.2 Description

The method for calculating the descriptors used is the one described by Bay et al. in SURF: Speeded Up Robust Features [29] where the descriptor for each feature point is calculated based on gradient calculations of image patches in the neighbourhood of the feature point. In this thesis, a rotation variant version of SURF is used sometimes referred to as U-SURF, for details see [30].

3.1.3 Matching

Given a set of N feature point locations from the source image $\mathbf{x}_1^s, \mathbf{x}_2^s, \dots, \mathbf{x}_N^s$ and M from the target image $\mathbf{x}_1^t, \mathbf{x}_2^t, \dots, \mathbf{x}_M^t$ where the descriptor for a feature point \mathbf{x} is denoted $\mathbf{d}(\mathbf{x})$ and is vector-valued. A pair of feature points are then matched if

$$\mathbf{x}^s = \arg \min_{\mathbf{x}_i^s} \mathbf{d}(\mathbf{x}^t) \cdot \mathbf{d}(\mathbf{x}_i^s), \quad i = 1, \dots, N \quad (3.1)$$

$$\mathbf{x}^t = \arg \min_{\mathbf{x}_j^t} \mathbf{d}(\mathbf{x}_j^t) \cdot \mathbf{d}(\mathbf{x}^s), \quad j = 1, \dots, M \quad (3.2)$$

which means that for all matched feature points, there cannot be another feature point in the other picture with a descriptor more similar than the one of the matched feature.

3.2 Affine Registration

The method to estimate an affine transformation between two images described in this chapter is used throughout this thesis. It is a feature-based method designed for robustness considering that the set of matched features usually contains outliers. Using previous notation, the goal is to find an affine transformation that minimizes

$$\sum_{i=1}^N \min(|\mathbf{A}\mathbf{x}_i^t + \mathbf{t} - \mathbf{x}_i^s|, \epsilon) \quad (3.3)$$

where ϵ is a parameter that represents the threshold between inliers and outliers and N is the number of matched features. Truncated L_1 -norm is used as a cost function to achieve robustness. This means that for inliers the L_1 -distance is used and for outliers a constant cost ϵ is used. The optimization was done using the Random Sample Consensus method (RANSAC) which was introduced by Fischler and Boyles [31]. RANSAC is an iterative method that uses a minimal subset of the data to estimate the model parameters. The steps of RANSAC are shown in Algorithm 2. An affine transformation has 12 degrees of freedom, hence a set of four randomly selected matched 3D features are used in each iteration.

3.3 Offline Building of the Models

The offline procedure is done once for each set of atlases in order to build the statistical model for each atlas image. The models are built from non-rigid deformation fields describing how points in the atlas image maps to points in other images.

For the case of the heart data set an SDM is created for each image paired with a manual

```

while not converged do
    Randomly select a subset of the original data large enough to estimate the model parameters.
    Use the selected data to estimate the model parameters.
    Evaluate the model using the original data set and the specified cost function.
    If the cost of this model is the lowest so far, save the parameters.
    Set to converged if a certain amount of iterations has been done or cost is below a threshold.
end
Return parameters with lowest cost

```

Algorithm 2: General description of RANSAC.

segmentation but not for the images lacking a manual segmentation. The statistical deformation models are supposed to describe the average deformation and the variability of the deformation for each atlas image. Intuitively the model should describe how each voxel of an atlas image usually warps when registered to other images. Note that the SDMs only model the local warping for the atlas voxels, a global affine transformation is estimated independently of the SDM during registration.

3.3.1 Training Set

To be able to create statistical deformation models for each atlas image, the atlas image needs to be registered to the other images available. For each registration a deformation field as described in Section 2.1.1.2 is calculated. These fields are denoted $\mathbf{u}_{ij}(\mathbf{x})$ and describe how points \mathbf{x} in image i are transformed to corresponding points $\mathbf{y} = \mathbf{u}_{ij}(\mathbf{x})$ in image j . Since this method is tested on a data set containing both images paired with manual segmentations (atlas images) as well as images with no paired manual segmentation (from now on referred to as non-atlas images) different methods must be used when registering an atlas image to another atlas image (referred to as an inter-atlas registration) and registering an atlas image to a non-atlas image (referred to as an extra-atlas registration).

3.3.1.1 Inter-Atlas Registration

The registrations performed between atlas images are called inter-atlas registrations. For these registrations both the target and the source images have a segmentation which enables a more accurate registration. The registrations are done in two steps. To start off an affine registration is done between the target and the source using a feature-based method where feature points in the target and the source images are detected using the method described in Section 3.1. The affine transformation is then estimated as described in Section 3.2 and used as an initialization to the non-rigid registration. The non-rigid registration algorithm used is the Fast Free-Form Deformation algorithm developed by Rueckert et al. in [17] and implemented by Marc Modat as part of NiftyReg [15] which is an image registration toolbox publicly available. The non-rigid registration is done on multi-channelled images where the first channel is the original 3D intensity image and the second channel is a mask describing the segmentations. The accuracy of the registrations are evaluated by calculating the Jaccard index of the target segmentation and the transformed source segmentation. The results for the different data sets are shown in Table 3.1.

Data Set	Jaccard Index
Heart	0.9715 ± 0.0562
Brain	0.6928 ± 0.0327

Table 3.1: Jaccard indices of the inter-atlas registrations for the two data sets, mean ± 1 standard deviation.

3.3.1.2 Extra-Atlas Registration

The extra-atlas registrations are done in a similar way as the inter-atlas ones. The only difference is that multi-channel images can not be used since the source image does not have a manual segmentation. The non-rigid part is in other words done on the intensity images only. The result of the registrations can for the same reason not be evaluated using the Jaccard index. Instead a visual inspection of a subset of the registrations was performed.

3.3.2 Alignment of the Training Set

The SDMs are only supposed to model the local warping of the atlas voxels since a global affine transformation will be estimated independently of the SDM during registration. Given the total transformation field $\mathbf{u}_{ij}(\mathbf{x})$ modeled as

$$\mathbf{u}_{ij}(\mathbf{x}) = \mathbf{A}_{ij}(\mathbf{x} + \mathbf{w}_{ij}(\mathbf{x})) + \mathbf{t}_{ij} \quad (3.4)$$

where $\mathbf{w}_{ij}(x)$ is the local warping function in the coordinate frame of the atlas image and \mathbf{A}_{ij} together with \mathbf{t}_{ij} describes the affine transformation converting points in image i to image j . Note that one statistical deformation model will be built for each atlas image in its original coordinate frame using the local warping functions $\mathbf{w}_{ij}(x)$, for j, \dots, N where N is the size of the training set. $\mathbf{w}_{ij}(x)$ is given by

$$\mathbf{w}_{ij}(\mathbf{x}) = \mathbf{A}_{ij}^{-1}(\mathbf{u}_{ij}(\mathbf{x}) - (\mathbf{A}_{ij}\mathbf{x} + \mathbf{t}_{ij})) = \mathbf{A}_{ij}^{-1}\mathbf{u}_{ij}(\mathbf{x}) - \mathbf{x} - \mathbf{A}_{ij}^{-1}\mathbf{t}, \quad (3.5)$$

which follows directly from Equation (3.4).

The affine transformation used in the alignment is chosen such that the transformation of four points denoted $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ in image i is completely described by the affine transformation, i.e.

$$\mathbf{u}_{ij}(\mathbf{x}_k) = \mathbf{A}_{ij}\mathbf{x}_k + \mathbf{t}_{ij}, \quad k \in \{1, 2, 3, 4\}, \forall j. \quad (3.6)$$

This also means that $\mathbf{w}_{ij}(\mathbf{x}_k) = 0$ for $k \in \{1, 2, 3, 4\}$ and all j . The four points are choosen at positions close to the surface of the segmentations far from each other. It is important that the same four points are used for all j when aligning the training set for image i .

3.3.3 Downsampling

In its matrix representation, each warp field $\mathbf{w}_{ij}(\mathbf{x})$ has the dimension $N_x \times N_y \times N_z \times 3$ where $N_x \times N_y \times N_z$ is the size of image i . This is in many cases too much data to efficiently be able to build an useful SDM from. Hence the warp field needs to be downsampled. This is done using independently in the three warping directions (\hat{x}, \hat{y} and \hat{z}) using cubic interpolation. If a warp field is downsampled with a factor k the downsampled field will have the dimension $N_x/k \times N_y/k \times N_z/k \times 3$ which means that the amount of data has been reduced by a factor of k^3 .

3.3.4 Statistical Deformation Models

The Statistical Deformation Models (SDM) can now be built from the downsampled warp fields. For atlas image i the warp fields $\mathbf{w}_{ij}(\mathbf{x})$, where $j = 1, \dots, N$ and N is the number of available deformation fields in the training set, are used to build the model. These fields $\mathbf{w}_{ij}(\mathbf{x})$ are reshaped into column vectors, denoted $\mathbf{W}_{ij}(\mathbf{x})$, and collected into the shape vector matrix \mathbf{X} , see Section 2.2. A weighted version of PCA is then used on \mathbf{X} where each column of \mathbf{X} is multiplied with a weight corresponding to the estimated reliability of the training sample. The weights are manually chosen as 1 for columns based on inter-atlas registration fields and 0.5 for columns based on extra-atlas registration fields. When the weighting has been applied, PCA is carried out normally according to the steps in Chapter 2.2. In addition to this, a column vector $\hat{\mathbf{X}}$ containing the means of all row of \mathbf{X} is calculated. For each atlas image an SDM has now been build consisting of $\hat{\mathbf{X}}$, Φ which is a matrix containing the eigenvectors of the covariance matrix formed using \mathbf{X} and λ which is a vector of the eigenvalues corresponding to the eigenvectors in Φ . Note that only the eigenvectors corresponding to non-zero eigenvalues are saved, the maximum size of Φ will hence be $3n \times (N - 1)$ where n is the size of the downsample image i ($n = N_x/k \cdot N_y/k \cdot N_z/k$, where k is the downsampling factor).

To recreate a warp field using the SDM a set of parameters \mathbf{b} are estimated. The column representation of the warp field can then be calculated as

$$\mathbf{W} = \hat{\mathbf{X}} + \Phi \mathbf{b} \quad (3.7)$$

which can be reshaped into the matrix representation of a warp field \mathbf{w} .

3.4 Online Segmentation Using an SDM

During the online process the calculated statistical deformation models are used for image registration. This thesis is focused on the pair-wise registrations meaning the registration of one atlas image (source) to a target image. The results of the pair-wise registrations can afterwards be combined giving a multi-atlas segmentation result. This subject will be described briefly in the last paragraph of this chapter. For the pair-wise registration a transformation $\mathbf{T}_{ij}(\mathbf{x})$ that transforms points in the source image I_i to the corresponding points in the target image I_j will be estimated. The transformation will be estimated in two parts, one affine describing the global part of the deformation $\mathbf{T}_{ij}^A(\mathbf{x})$ and a non-rigid one describing local warpings $\mathbf{w}_{ij}(\mathbf{x})$, hence

$$\mathbf{T}_{ij}(\mathbf{x}) = \mathbf{T}_{ij}^A(\mathbf{x} + \mathbf{w}_{ij}(\mathbf{x})). \quad (3.8)$$

The estimated transformation is used to transform the segmentation of the source image L_i to give a segmentation guess for the target image $L_j^g = \mathbf{T}_{ij} \circ L_i$. To evaluate the segmentation guess of the target image the Jaccard index between the guess and the gold standard segmentation is calculated.

3.4.1 Feature Detection

For each registration a set of matched features is calculated using the method described in Section 3.1. These features are used for estimation of both the affine transformation and the non-rigid transformation. The set of matched features will be denoted $(\mathbf{x}_i^t, \mathbf{x}_i^s)$ where \mathbf{x}_i^t is the position of feature i in the target image and \mathbf{x}_i^s is the position of feature j in the source image, $i = 1, \dots, N$ where N is the number of matched features.

3.4.2 Affine Transformation Initialization

The affine transformation is initialized using the method described in Section 3.2. The transformation is denoted $\mathbf{T}^A(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{t}$.

3.4.3 Transformation Estimation

3.4.3.1 Feature Selection

Since this method is feature based the first step is to decide which of the detected and matched features to use. Generally not all features are used since some of them are located in sections of the image that are not interesting for the segmentation. Instead the features located closest to the surface of the segmentation are used. The number of features included is a tunable parameter. From this point $(\mathbf{x}_i^t, \mathbf{x}_i^s)$ will refer to the set of chosen features to use for the transformation estimation and not all the detected features.

3.4.3.2 Interpolation

The SDM used for the non-rigid transformation only describes the statistics of how the source image warps in a discrete and evenly spaced grid points in the image. This means that interpolation needs to be done to make the SDM describe the warping at the feature locations. The goal with the interpolation is to create a vector $\hat{\mathbf{X}}_k^o$ that describes the mean warping at each feature location and Φ_k^o that contains the variation information from the SDM in such a way that

$$\mathbf{x}_k^t \approx \mathbf{A}(\mathbf{x}_k^s + \hat{\mathbf{X}}_k^o + \Phi_k^o \mathbf{b}) + \mathbf{t}. \quad (3.9)$$

Here \mathbf{x}_k^t are the feature locations in the target image and \mathbf{x}_k^s are the feature locations in the source image. $\hat{\mathbf{X}}_k^o$ and Φ_k^o are calculated according to

$$\hat{\mathbf{X}}_k^o = \sum_i \mu_i \hat{\mathbf{X}}_i, \quad (3.10)$$

$$\Phi_k^o = \sum_i \mu_i \Phi_i, \quad (3.11)$$

where $\hat{\mathbf{X}}_i$ denotes the rows of $\hat{\mathbf{X}}$ corresponding to grid point i and the sum is over all grid points. The weights μ_i equal zero for all rows not corresponding to the grid points that together form the smallest cube containing the feature point. For the non-zero elements the value of μ_i is inversely proportional to the distance from the feature point to the corresponding grid point.

3.4.3.3 Optimization

At this point we can formulate the estimation of the transformations as the following optimization problem

$$\min_{\mathbf{b}, \mathbf{A}, \mathbf{t}} \sum_{k=1}^n \phi(|\mathbf{x}_k^t - \mathbf{A}(\mathbf{x}_k^s + \hat{\mathbf{X}}_k^o + \Phi_k^o \mathbf{b}) - \mathbf{t}|) \quad (3.12)$$

where ϕ is the Huber loss function described in Section 2.3.2. Both the affine transformation and the parameters \mathbf{b} are optimized using an alternating optimization described in Algorithm 3.

Initialize $\mathbf{r}_{old} = \infty$ and $i = 0$

while *not converged* **do**

- Calculate $\mathbf{a} = \mathbf{A}^{-1}(\mathbf{x}^t - \mathbf{t}) - \hat{\mathbf{X}}^o - \mathbf{x}^s$.
- Reshape \mathbf{a} to a column vector.
- Calculate $\mathbf{b} = \arg \min_{\mathbf{b}} \phi(\mathbf{a} - \Phi^o \mathbf{b})$ using IRLS.
- Calculate $\mathbf{x}_a^s = \mathbf{x}^s + \hat{\mathbf{X}}^o + \Phi^o \mathbf{b}$.
- Reshape \mathbf{x}_a^s and \mathbf{x}^t to a $4 \times n$ matrix where the first three rows are the feature coordinates and the fourth row is set to ones.
- Calculate $\mathbf{A}_{aug} = \arg \min_{\mathbf{A}_{aug}} \phi(\mathbf{x}^t - \mathbf{A}_{aug} \mathbf{x}_a^s)$ using IRLS.
- Calculate the residual $\mathbf{r} = \phi(\mathbf{x}^t - \mathbf{T}^A(\mathbf{x}^s + \hat{\mathbf{X}}^o + \Phi^o \mathbf{b}))$.
- If $(\mathbf{r}_{old} - \mathbf{r}) < \epsilon$ or $i > i_{max}$, set to converged.
- Set $\mathbf{r} = \mathbf{r}_{old}$ and $i = i + 1$.

end

Return \mathbf{b} and \mathbf{T}^A .

Algorithm 3: Algorithm for optimization of the affine and non-rigid transformations. IRLS refers to Iteratively Reweighted Least Squares, see Section 2.3.3. $\hat{\mathbf{X}}^o$ and Φ^o are the mean warping and modes of variation at the feature locations. \mathbf{A}_{aug} is the augmented matrix describing the affine transformation \mathbf{T}^A , see chapter 2.1.1.1, ϵ is a set threshold and i_{max} is the maximum number of iterations allowed. This algorithm is used to solve the problem presented in Equation (3.12).

3.4.3.4 Creation of the Warp Field

When the parameters \mathbf{b} have been calculated the entire warping field can be recreated by

$$\mathbf{W} = \hat{\mathbf{X}} + \Phi \mathbf{b} \quad (3.13)$$

and reshaping \mathbf{W} to matrix field representation $\mathbf{w}(\mathbf{x})$ that describes the non-rigid transformation. The affine transformation $\mathbf{T}^{\mathbf{A}}(\mathbf{x})$ is already calculated in its final form during optimization. The entire transformation on the form of equation (3.8) has now been calculated and can be used to transform the source segmentation to the target image.

3.4.4 Transformation of the Source Segmentation

The transformation $\mathbf{T}(\mathbf{x})$ calculated during registration describes how points in the source image transforms to the corresponding points in the target image. To transform the labeling of the source image a polyhedron representation of the labeling is created. The polyhedron P consists of a set of vertex coordinates \mathbf{v} and a matrix \mathbf{f} describing which vertices are connected with a surface, hence $P = \{\mathbf{v}, \mathbf{f}\}$. The surface of the polygon describes the limit where the labeling changes from zero to one. Given the polygon describing the source labeling $P_s = \{\mathbf{v}_s, \mathbf{f}_s\}$ the polygon describing the target segmentation is given as

$$P_t = \{\mathbf{T}(\mathbf{v}_s), \mathbf{f}_s\}. \quad (3.14)$$

The segmentation can then be transformed to a voxel representation L_t^g by determining which voxels are located inside the polyhedron (labeled as ones) and which are outside of the polyhedron (labeled as zeros).

3.4.5 Multi-Atlas Segmentation

Given a set of N segmentations from the pair-wise registration L_i^g , where $i = 1, \dots, N$ a multi-atlas segmentation can be estimated according to

$$L^g = \left\lceil \frac{1}{N} \sum_{i=1}^N L_i^g \right\rceil \quad (3.15)$$

where $\lceil \cdot \rceil$ is the nearest integer function. This is a simple majority voting method where a voxel is labeled as one if more than half of the pair-wise registrations have labeled it as one.

4

Implementation and Baseline Methods

4.1 Implementation

The method developed in this thesis, from here on referred to as ShapeReg, was implemented in MATLAB with a few exceptions where Mex-files written in C++ was used. The implementations of the feature extraction and affine registration are described separately since they are included in both the offline and the online process.

4.1.1 Feature Extraction

The feature points were extracted using the program `ff3d` implemented by Svärm et al. [30]. The program returns feature locations and descriptors and was run with default setting for rotation variant features. The brain images were upsampled with a factor 2 before feature extraction.

4.1.2 Affine Registration

The affine registration was done using the method described in Section 3.2. The threshold between inliers and outliers, ϵ , was set to 10 mm while the maximum number of iteration was set to 10^6 .

4.1.3 Offline Registration of the Atlas Set

The non-rigid registrations for the inter-atlas and extra-atlas registration were done using the program `reg_f3d` from NiftyReg. The program was run with default settings as described in Section 3.3.1. For the heart data set a downsampling factor of 4 was used when downsampling the deformation fields, see Section 3.3.3, for the brain data set the downsampling factor was 2.

4.1.4 Online Segmentation Using an SDM

The number of features used to estimate the non-rigid transformation during the online process is set to 400 for the heart data set and 200 for the brain data set. The optimization was done using the Huber loss function with $\delta = 1$, see Section 2.3.2.

4.2 Baseline Methods

The method developed in this thesis is compared to two baseline methods. The first one is a feature-based method which estimates the transformations using the same features as the method developed in this thesis. The other one is an intensity-based method widely used in medical imaging.

4.2.1 Baseline 1: Feature-Based

The feature-based method used as baseline in this thesis will be referred to as Featreg and uses the methods described in Section 3.1 and 3.2 to detect and match features as well as estimate the affine transformation. The implementations of these methods are described in Section 4.1.1 and 4.1.1. The non-rigid transformation uses a free-form deformation model based on cubic B-splines proposed by Rueckert et al. [17]. The transformation field is optimized using a method described in the article "Scattered Data Interpolation with Multilevel B-splines" by Lee et al. [32]. The method was implemented in MATLAB by Dirk-Jan Kroon and is used to find a transformation between the affinely transformed feature points in the source image $\mathbf{T}^A(\mathbf{x}^s)$ to the feature points in the target image \mathbf{x}^t . The transformation is then used to transform the source segmentation as described in Section 3.4.4 giving the segmentation of the target.

4.2.2 Baseline 2: Intensity-Based

The intensity-based method used as baseline in this thesis will be referred to as Niftyreg and uses methods from the program package NiftyReg developed by Marc Modat [15].

The affine transformation is estimated with the program `reg_aladin` which uses an intensity-based block matching algorithm proposed by Ourselin et al. [33]. The program was run with default settings.

The non-rigid transform is estimated with the program `reg_f3d` which uses a fast free-form deformation algorithm based on the work done by Rueckert et al. [17]. The program was run with default settings.

When the transform had been estimated the program `reg_transform` was used to transform the source segmentation to the target. Nearest neighbour interpolation was used.

5

Results

5.1 Segmentation

The segmentation results are presented for the pair-wise segmentations in Section 5.1.1 and for the multi-atlas segmentation in Section 5.1.2.

5.1.1 Pair-Wise Atlas Segmentation

A leave-one-out cross-validation approach was used to evaluate the capacity of Shapereg for pair-wise segmentation. For each image i in an atlas a statistical shape models were built using all images in the atlas except image i . pair-wise registration were then done with image i as target and the rest of the atlas as sources. The registration was evaluated by calculating the Jaccard index of the target segmentation and the transformed source segmentation. This process were repeated for each image i in the atlas giving $N(N - 1)$ segmentations, where N is the number of atlases in the data set, registrations (20 for the heart data set and 30 for the brain data set). The same registrations were also done for the baseline methods for comparison. The result of the pair-wise segmentations can be seen in table 5.1 and 5.2.

5.1.2 Multi-Atlas Segmentation

For the multi-atlas segmentations the propagated labels created during the pair-wise registration were merged to a multi-atlas segmentation. Hence the training and evaluation were performed using the same leave-one-out approach. This gives N segmentations where N is the number of atlases in the data set (20 for the heart data set and 30 for the brain data set). These registrations were done for the ShapeReg method as well as the two baseline methods and the segmentations were evaluated using the Jaccard index. The results are show in Table 5.3 for the heart and Table 5.4 for the brain.

Method	Jaccard Index
Shapereg	0.8307 ± 0.0707
Featreg	0.8245 ± 0.0846
Niftyreg	0.7737 ± 0.1388

Table 5.1: Jaccard indices of the pair-wise atlas segmentation of the pericardium for the three methods, *mean* \pm 1 *standard deviation*.

Method	Jaccard Index
Shapereg	0.5296 ± 0.0552
Featreg	0.4702 ± 0.0667
Niftyreg	0.5700 ± 0.0587

Table 5.2: Jaccard indices of the pair-wise atlas segmentation of the hippocampus for the three methods, *mean* \pm 1 *standard deviation*.

Method	Jaccard Index
Shapereg	0.8906 ± 0.0341
Featreg	0.8979 ± 0.0357
Niftyreg	0.8779 ± 0.0800

Table 5.3: Jaccard indices of the multi-atlas segmentation of the pericardium for the three methods, *mean* \pm 1 *standard deviation*.

Method	Jaccard Index
Shapereg	0.5957 ± 0.0504
Featreg	0.6049 ± 0.0443
Niftyreg	0.6803 ± 0.0427

Table 5.4: Jaccard indices of the multi-atlas segmentation of the hippocampus for the three methods, *mean* \pm 1 *standard deviation*.

Data Set	D
Heart	4.0569 ± 2.6319
Brain	2.7601 ± 0.3718

Table 5.5: Mean euclidean distance error \mathbf{D} in mm as defined in Equation (5.1) for the two data sets. The result is shown as *mean* \pm 1 *standard deviation* for all recreated fields.

Data Set	Jaccard Index
Heart	0.8335 ± 0.0729
Brain	0.5336 ± 0.0536

Table 5.6: Jaccard index when transforming the segmentation using the recreated fields as described in Section 5.2 for the two data sets. The result is shown as *mean* \pm 1 *standard deviation* for all recreated fields.

5.2 Model Evaluation

The statistical deformation models created during the offline process are evaluated by looking at their generality and compactness. With generality means the models ability to describe warping fields not in the training set. This is evaluated using a leave-one-out approach where the warping field $\mathbf{w}_{ij}(x)$ is recreated using the model for atlas i from the training done when atlas j was left out, the recreated field is denoted $\tilde{\mathbf{w}}_{ij}(x)$. The parameters for the recreated field are calculated using $\mathbf{w}_{ij}(x)$ instead of feature location as in the online process. The number of parameters used for the model is decided in the same way as for the online process, for these results enough parameters to describe 98 % of the total variance of the model is included. The recreated and original warping fields are then compared by calculating the mean euclidean distance error for all voxels

$$D_{ij} = \frac{1}{N_x} \sum_{\mathbf{x}} \|\tilde{\mathbf{w}}_{ij}(x) - \mathbf{w}_{ij}(x)\| \quad (5.1)$$

where $\|\cdot\|$ is the L_2 -norm, the sum is over all x for which $\mathbf{w}_{ij}(x)$ is defined and N_x is the size of this domain. This is done for all i and j in the atlases. The result is shown in Table 5.5. The recreated fields are also used to transform the source segmentation and the Jaccard index between the target segmentation and the transformed source segmentation is calculated. The result is shown in Table 5.6.

The compactness of the models are also evaluated by looking at the cumulative variance as a function of the number of parameters, or modes, included. This is done for all models created during the offline procedure and the result is shown in Figure 5.1 for the heart and Figure 5.2 for the brain. In addition a histogram over the number of modes needed to be included to describe 98 % of the variance of the statistical shape model is shown in Figure 5.3 for the heart and Figure 5.4 for the brain.

A simple visualization of the statistical shape models are shown in Figure 5.5 for the heart and in Figure 5.6 for the brain. The figures show examples of how the main modes of variation are warping the atlas image.

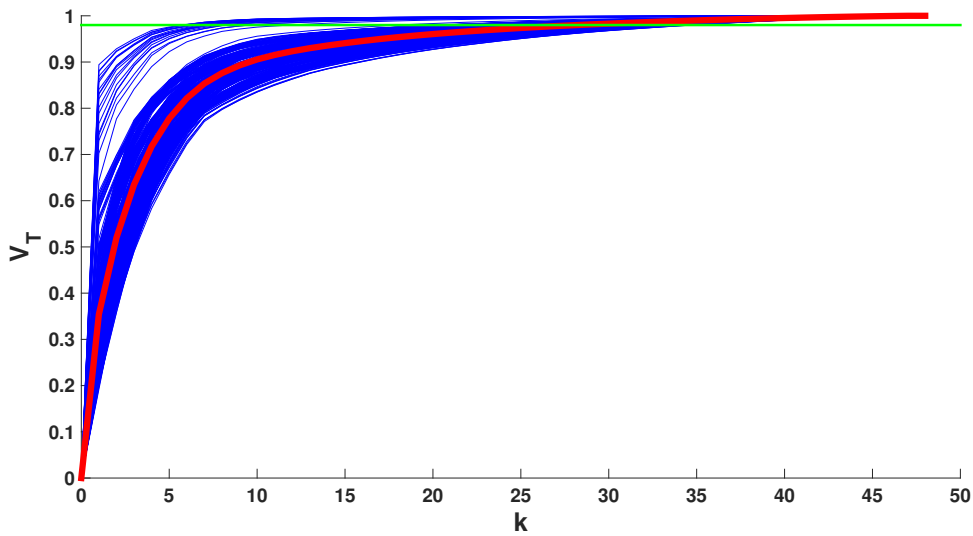


Figure 5.1: Cumulative variance, V_T , described by the statistical shape model as a function of number of modes included, k , see Equation (2.10). These results are from all the 380 models created during the leave-one-out experiments done on the **heart data set**. The blue lines are individual results from each model and the red line is the mean of all models. The green line is located at $V_T = .98$.

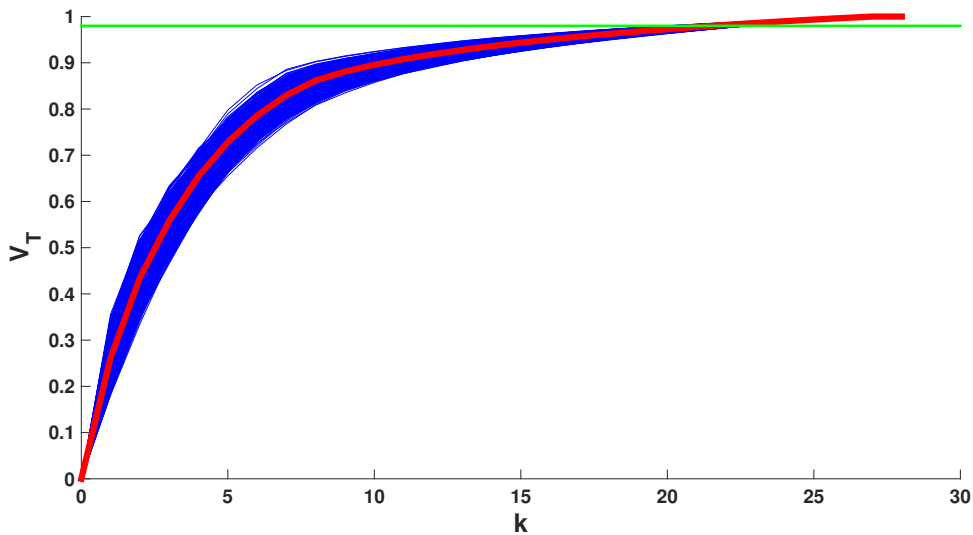


Figure 5.2: Cumulative variance, V_T , described by the statistical shape model as a function of number of modes included, k , see Equation (2.10). These results are from all the 870 models created during the leave-one-out experiments done on the **brain data set**. The blue lines are individual results from each model and the red line is the mean of all models. The green line is located at $V_T = .98$.

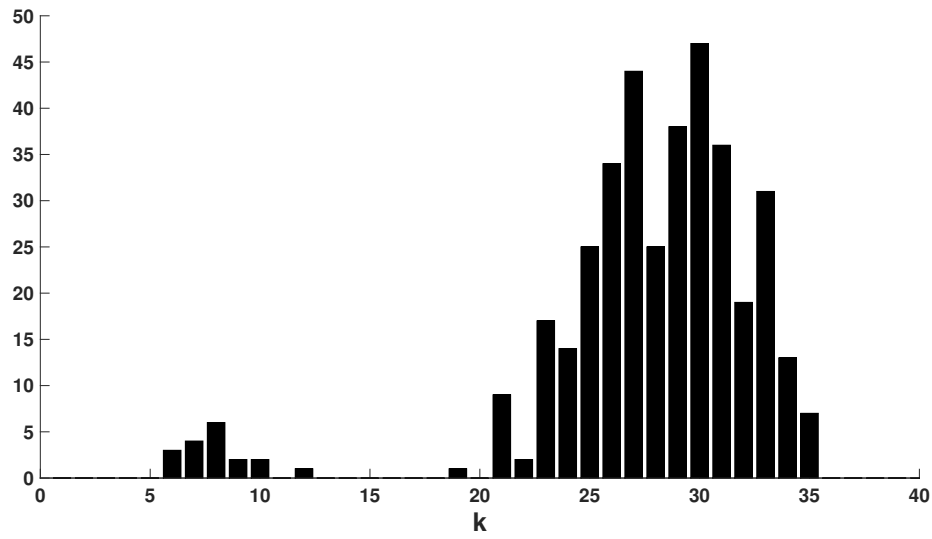


Figure 5.3: Histogram over the number of modes, k , needed to be included to describe 98 % of the variance of the statistical shape model. Note that this number equal the numbers of parameters needed to be optimized during the online process. These results are from all the 380 models created during the leave-one-out experiments done on the **heart data set**.

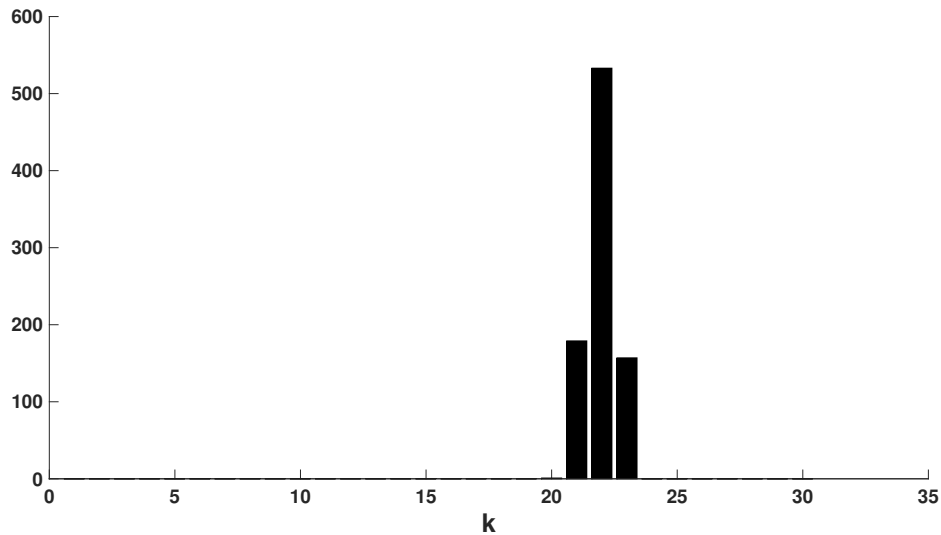


Figure 5.4: Histogram over the number of modes, k , needed to be included to describe 98 % of the variance of the statistical shape model. Note that this number equal the numbers of parameters needed to be optimized during the online process. These results are from all the 970 models created during the leave-one-out experiments done on the **brain data set**.

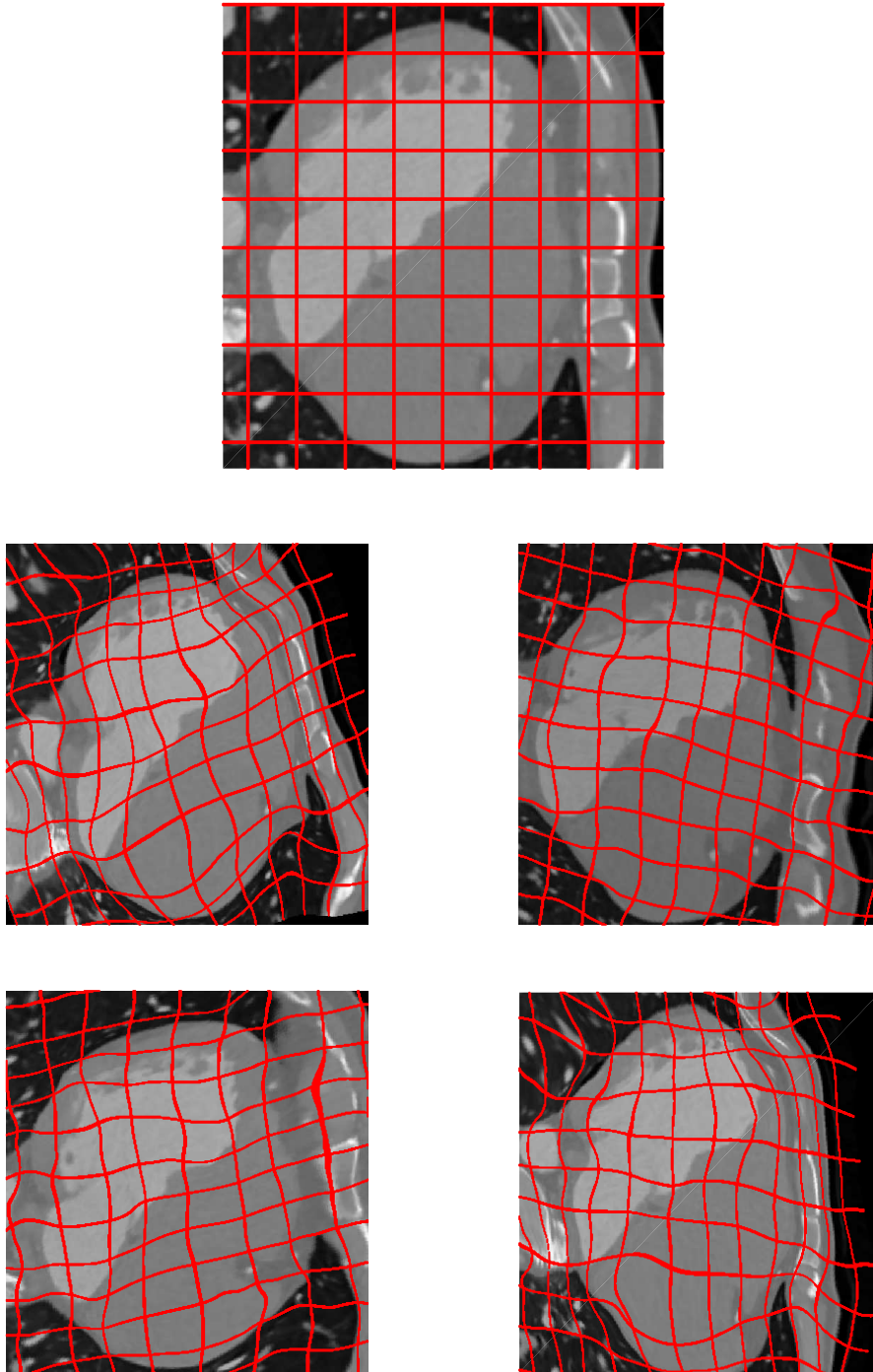


Figure 5.5: Visualization of the statistical deformation model for atlas image 1 in the **heart data set**. The topmost image is an unaltered slice of the atlas image, the rest of the images has been generated by varying the two main modes of variation between $-3\sqrt{\lambda_i}$ (to the left) and $3\sqrt{\lambda_i}$ (to the right) and applying the warping to the slice of image 1. In row two mode 1 is varied and in row three mode 2 is varied.

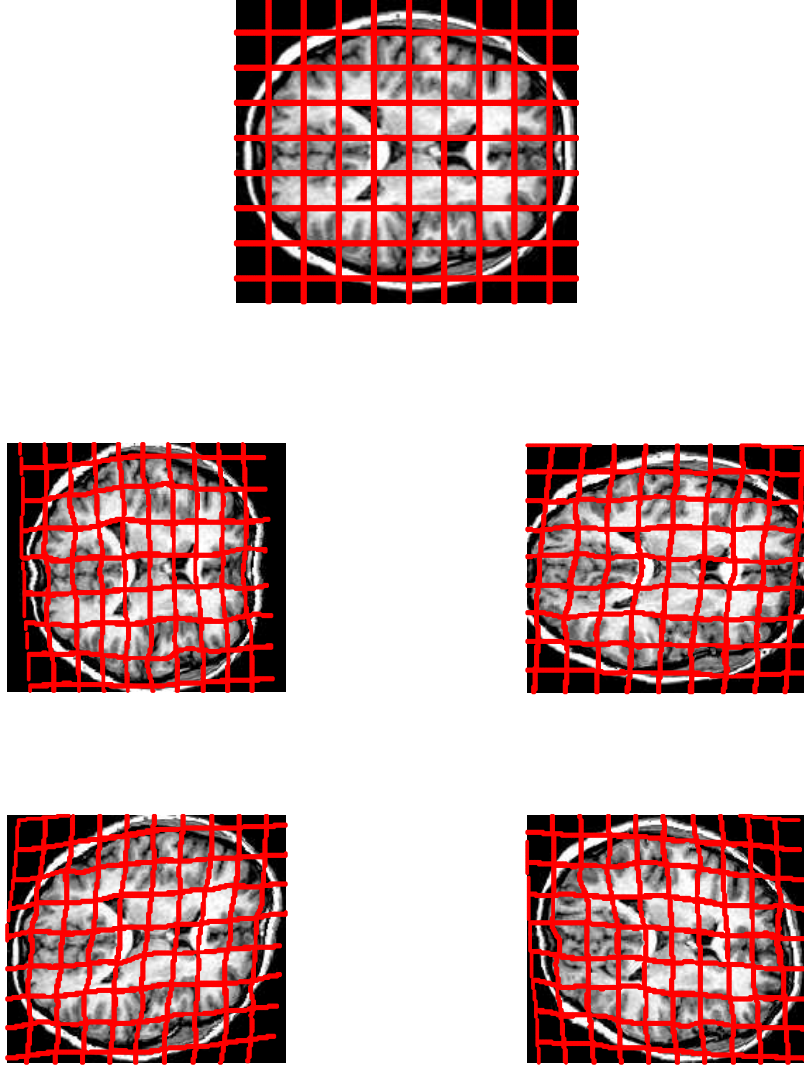


Figure 5.6: Visualization of the statistical deformation model for atlas image 1 in the **brain data set**. The topmost image is an unaltered slice of the atlas image, the rest of the images has been generated by varying the two main modes of variation between $-3\sqrt{\lambda_i}$ (to the left) and $3\sqrt{\lambda_i}$ (to the right) and applying the warping to the slice of image 1. In row two mode 1 is varied and in row three mode 2 is varied.

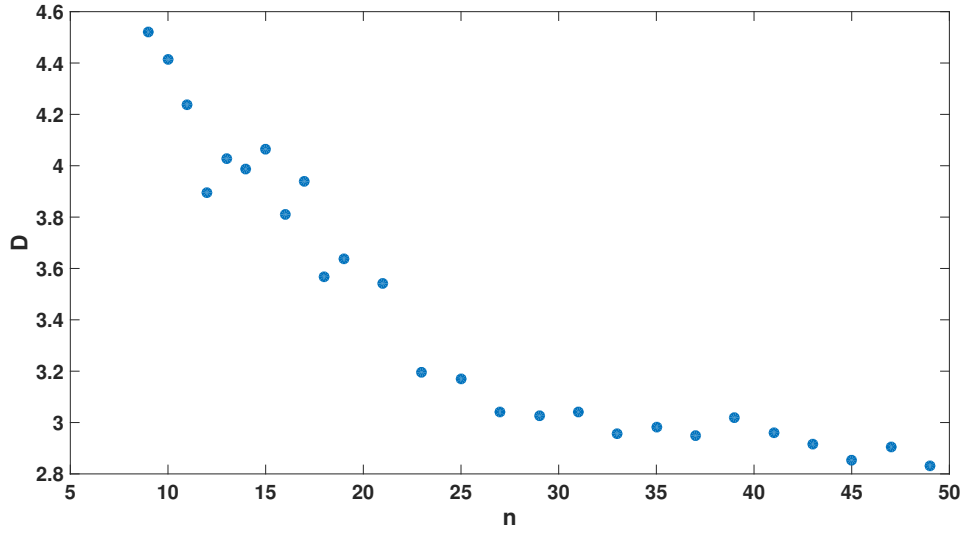


Figure 5.7: Mean euclidean distance error D in mm as defined in Equation (5.1). D is shown as a function of the training set size n for the **heart data set**. For n larger than 19 deformation fields from extra-atlas registrations were included. Atlas image 1 was used as target and atlas images $2, \dots, 9$ were used as sources, the D shown in the plot is the mean of the 8 results.

5.2.1 Effect of Training Set Size

To evaluate the effect the training set size has on the segmentation result the size of the training set was varied when building the statistical shape models. This was done using atlas image 1 as a target and the rest of the atlas as sources, several models were built using different sizes of the training set (atlas image 1 was never included). The models were then evaluated using the methods described in Section 5.2. The mean euclidean distance error result is shown in Figure 5.7 for the heart and Figure 5.8 for the brain. The segmentation result is shown in Figure 5.9 for the heart and Figure 5.10 for the brain. Note that the absolute numbers in these results should not be compared with other segmentation results since different targets were used.

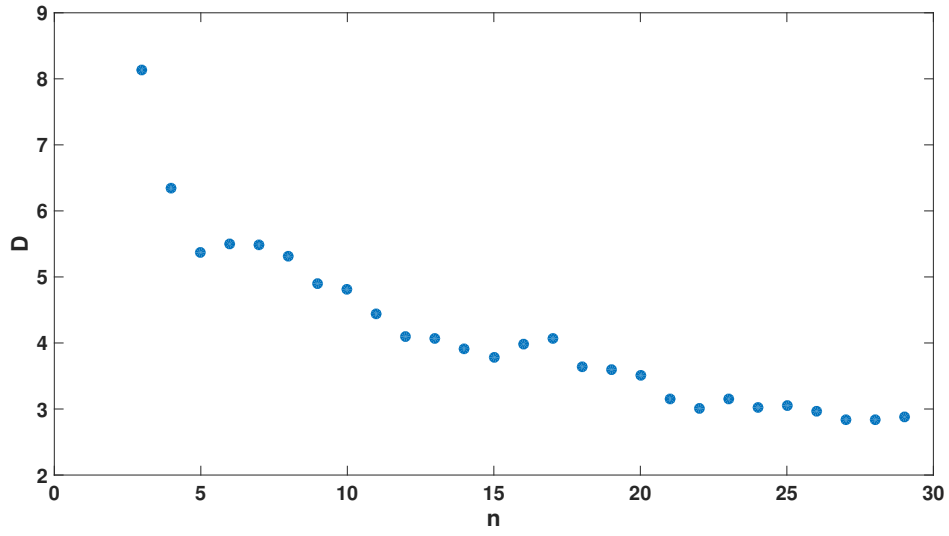


Figure 5.8: Mean euclidean distance error \mathbf{D} in mm as defined in Equation (5.1). \mathbf{D} is shown as a function of the training set size n for the **brain data set**. Atlas image 1 were used as target and atlas images 2,3 and 4 were used as sources, the \mathbf{D} shown in the plot is the mean of the 3 results.

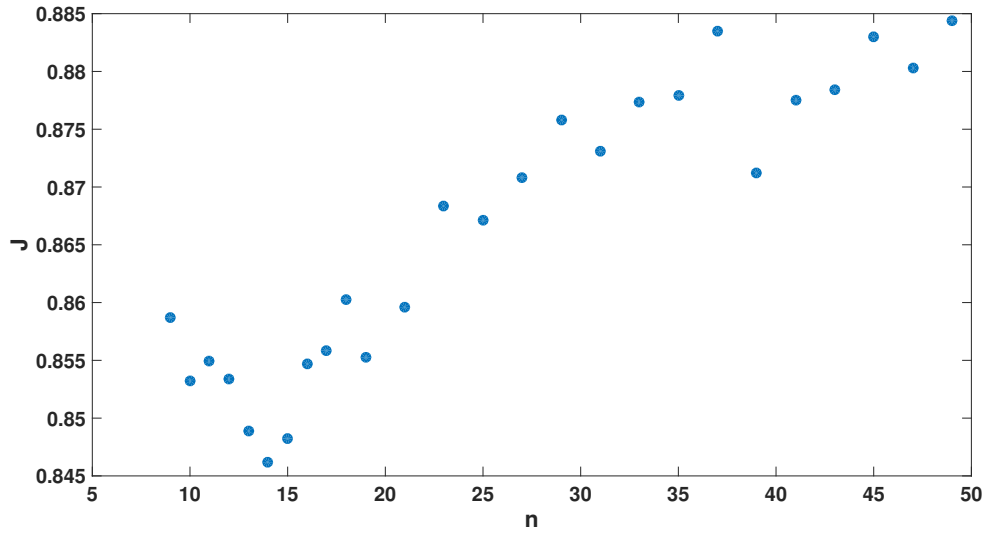


Figure 5.9: Segmentation results when transforming the segmentation using the recreated fields as described in Section 5.2. The Jaccard index \mathbf{J} is shown as a function of the training set size n for the **heart data set**. For n larger than 19 deformation fields from extra-atlas registrations were included. Atlas image 1 were used as target and atlas images 2, ..., 9 were used as sources, the \mathbf{J} shown in the plot is the mean of the 8 segmentations.

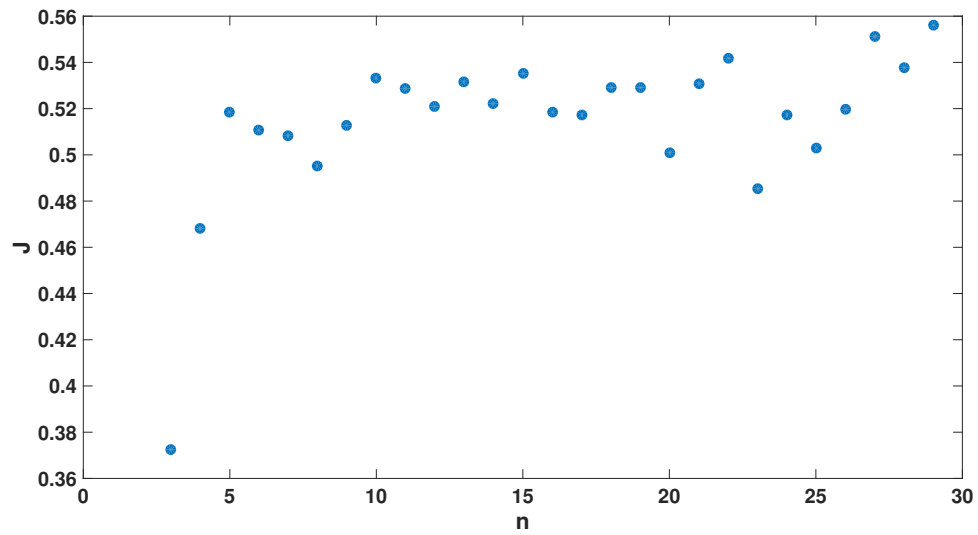


Figure 5.10: Segmentation results when transforming the segmentation using the recreated fields as described in Section 5.2. The Jaccard index J is shown as a function of the training set size n for the **brain data set**. Atlas image 1 was used as target and atlas images 2,3 and 4 were used as sources, the J shown in the plot is the mean of the 3 segmentations.

Data Set	Method	Shapereg	Featreg
Heart		5.3392 ± 0.2914	0.3474 ± 0.0692
Brain		3.6247 ± 0.2367	0.5772 ± 0.0530

Table 5.7: Mean euclidian distance error in mm between target feature points and transformed source feature points. The result is shown as *mean \pm 1 standard deviation* for all recreated fields.

Data Set	Method	Shapereg	Featreg	Niftyreg
Heart		46.47	4.06	1945.00
Brain		14.20	1.21	100.00

Table 5.8: Mean runtime for the non-rigid part of a pair-wise registration in seconds.

5.3 Feature Matching

The feature matching capacity of each method are evaluated by calculating the mean euclidean distance error between the target feature points and the transformed source feature points. The source feature points have been transformed using the transformation calculated during registration. This is done for the Shapereg method and the Featreg method and the result is shown in Table 5.7.

5.4 Runtimes

The runtimes of the three methods were also evaluated by timing the non-rigid part of the registration. This was done for all pair-wise registrations and the result can be seen in Table 5.8. Note that to achieve a multi-atlas segmentation this part needs to be performed once for each atlas image.

6

Discussion

In this chapter the method and results of the thesis will be discussed. First, the method and accuracy of the method used to create the training set will be reviewed. Second follows a discussion of the segmentation results as well as the model evaluation. Last, some ideas for future work and improvements are presented.

6.1 Creation of Training Set

The creation of the the two training sets is described in Section 3.3.1. The accuracy of the pair-wise registrations used to build the statistical deformation models are evaluated using the Jaccard index and the result is shown in Table 3.1. The Jaccard index for the heart training set is 0.9715 ± 0.0562 which may seem like almost perfect registrations but this is not entirely true. Looking at each atlas images separately one can see that atlas image 14 has an average Jaccard index of 0.9056 for registrations when it is used as the target image, this is shown in Figure 6.1. This is slightly lower than for the rest of the atlas which might result in a less accurate and descriptive statistical deformation model for this image. For the brain training set the Jaccard index is 0.6928 ± 0.0327 . This is quite low and limits the maximum achievable segmentation result for the models based on this training data.

Worth to note is the difference between the shapes of the objects that are segmented in the two data sets. The inside of the pericardium forms a large sphere-like volume, a segmentation that is skewed by one pixel in any direction would then still give quite a high Jaccard index. The hippocampi on the other hand are two relatively small volumes in relation to the entire brain image which means that a segmentation off by one pixel in any direction would give a significant decrease in Jaccard index. This effect is tested by loading the segmentation of atlas image 1 for both training set, creating a secondary segmentation that is skewed by one pixel relative to the original in the positive x -direction and calculating the Jaccard index of the two segmentations. For the heart segmentation this gives a Jaccard index of 0.9817 and for the brain segmentation 0.5116. Hence a small segmentation error will give a much larger effect on the Jaccard index for the brain data set.

There are several aspects of the pair-wise atlas registrations that could be investigated further. As of now, the segmentation is treated as an extra channel of the image during inter-atlas registrations. This means that the intensity and the segmentation is treated equally when aligning the images which might not be optimal. Further, the similarity measure used for the optimization could be changed, for this thesis a similarity measure based on bi-variate normalized mutual information was used. Alternatives to this was not

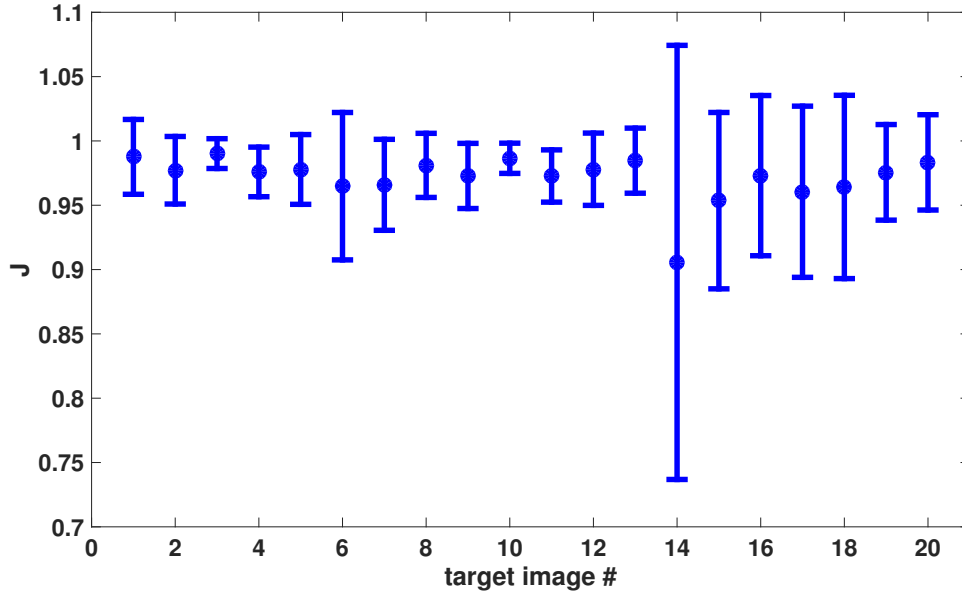


Figure 6.1: Error-bar plot of the Jaccard index for the transformations in the **heart training set** as a function of target image number. The markers denote the means and the bars denote the intervals that are within one standard deviation.

tested thoroughly and there might be an alternative similarity measure that gives better results.

6.2 Segmentation

The results for the pair-wise segmentations from the data set of the heart and the brain are shown in Table 5.1 and in Table 5.2 respectively. As can be seen in the tables, the Shapereg method performs better than the Featreg method on both data sets but is beaten by the Niftyreg method on the brain data set. One should point out that the Niftyreg method is intensity-based and do not use the feature locations during registration. Featreg on the other hand uses the same features as Shapereg and might for that reason be considered a more fair comparison.

The segmentation problems in the heart data set and the brain data set have some major differences. Segmentation of pericardium is done on large CT-images with high resolution while the brain images are smaller MRI-images with lower resolution, also the inside of the pericardium is a significantly larger volume than the hippocampus. From the segmentation results in this thesis one can draw the conclusion that the feature-based methods work better when segmenting larger objects in images with high resolution while the intensity-based method is more suited for segmentation of smaller objects in images with low resolution. In addition the runtime for the registration scales better with image size for feature-based methods. This can be seen in Table 5.8.

The results for the multi-atlas segmentations are shown in Table 5.3 for the heart data set and in Table 5.4 for the brain data set. As can be seen from the results, the Featreg method achieves slightly better multi-atlas segmentation results than Shapereg even though the pair-wise registrations for the Shapereg method are more accurate. One can also see

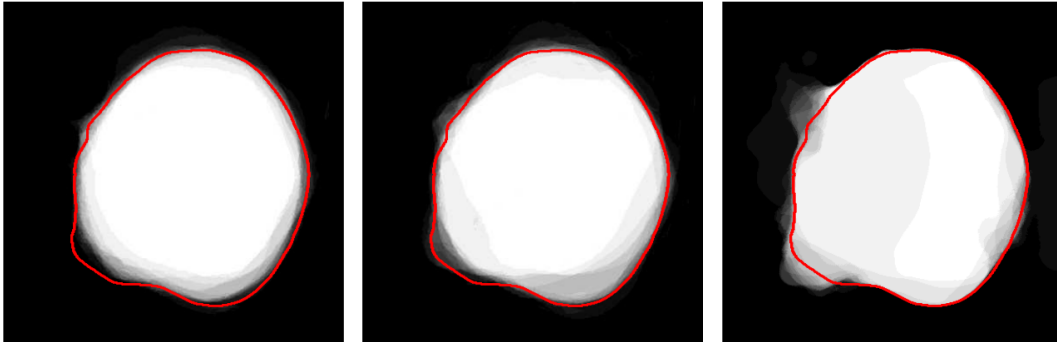


Figure 6.2: Visualization of the multi-atlas voting on the **heart data set** for all three methods. The images show all the pair-wise segmentations added together as well as the edge of the true target segmentation. From left to right the result of Shapereg, Featreg and Niftyreg is shown.

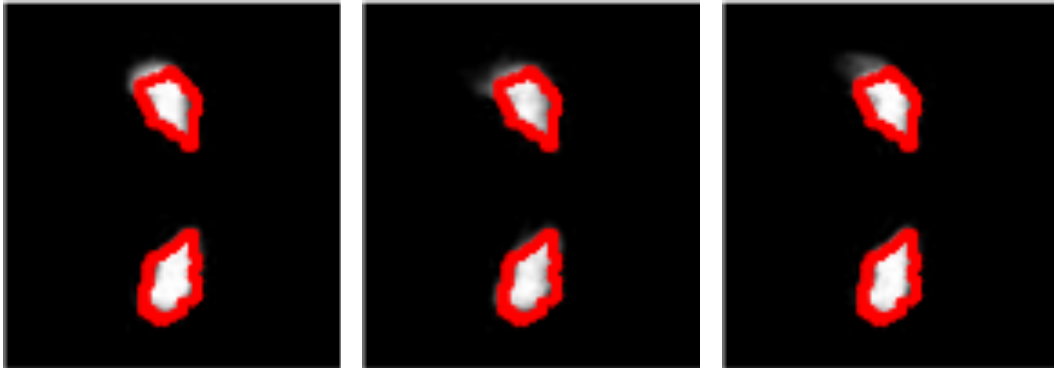


Figure 6.3: Visualization of the multi-atlas voting on the **brain data set** for all three methods. The images show all the pair-wise segmentations added together as well as the edge of the true target segmentation. From left to right the result of Shapereg, Featreg and Niftyreg is shown.

that Niftyreg achieves higher multi-atlas results in relation to the pair-wise segmentation results. Figure 6.2 and Figure 6.3 shows a visualization of an example of the voting process for the heart data set and the brain data set. Looking at the figures, one can see that the white areas for the Shapereg method are more coherent while for the other two methods the white areas are more spread out. This signifies that the different votes from the pair-wise registrations are more alike for the Shapereg method than for the other two methods. A reason for this may be that the pair-wise registrations for the Shapereg method are not independent of each other since all the atlas images are used during the training phase. Consequently, the errors for each pair-wise segmentations will not cancel out each other as efficiently as for the other two methods and the multi-atlas result will be worse in relation to the pair-wise segmentation result.

6.3 Feature Matching

Looking at Table 5.7, it can be seen that the Featreg method matches the features significantly better than the Shapereg method does. This is expected since Featreg uses a more general and adaptable representation of the non-rigid transformation. Since the matched feature contains outliers and does not always perfectly describe the correspondences be-

tween the target and source image one can argue that the features should not always be matched as good as possible. A fact that points to this is that the segmentation results for Shapereg are better than the Featreg ones even though it matched the features worse.

6.4 Runtimes

As can be seen in Table 5.8, the Shapereg method is faster than the Niftyreg method which is to be expected since it is feature-based instead of intensity-based. It is however slower than the Featreg method. This can be explained by the fact that Shapereg needs to load the data describing the SDMs which can be time demanding considering these files are relatively large. Furthermore, the deformation fields are calculated several times for each registration since the alternating optimization method described in Section 3.4.3.3 is used. One estimation of the non-rigid part of the deformation field using the Shapereg method is considerably faster than Featreg but since it is done up to 100 times in addition to the data loading the total process of performing the non-rigid transformation is more time consuming. Finding a more data efficient way to represent the models as well as developing another way of estimating the affine transformation, so that the non-rigid transformations would only have to be estimated once, are examples of improvements that could be done to decrease the runtime of Shapereg.

6.5 Model

The evaluation of the compactness of the SDMs are shown in Figure 5.1 and 5.3 for the heart and in Figures 5.2 and 5.4 for the brain. For the heart models there are a few models that are a bit more compact than the others, this is most easily seen in Figure 5.3 where there is a small bump in the histogram around $n = 9$. The models corresponding to this bump all come from atlas image 17 which was discovered to have a somewhat inaccurate manual labeling. This unfortunately affects the atlas registrations and hence also the creation of the SDMs which might be the reason for this result. This miss-labeling could also influence the generality as well as the segmentation results of the models. Unfortunately, the miss-labeling was noticed too late for all calculations do be remade.

The generality of the models is evaluated in two ways where the parameters of the SDM are calculated using the true warp field that was left out during the training. This means that all potential error sources induced during registration using the online method is removed except for the ones caused by the actual SDM. The generality is hence a good measure of what the maximum capacity for segmentation using the SDM is. In Table 5.5 the mean euclidean distance error is shown and in Table 5.6 the segmentation result is shown. The mean euclidean distance error is small for the brain data set where 2.76 mm corresponds to about 3 voxels. For the heart data set the error is a bit larger, in addition the variance is quite large. Upon further investigation, it can be seen that the SDMs describing the variability of atlas image 17 generally had larger mean distance error than the rest (about twice as large). This is quite certainly because of the miss-labeling mentioned above. Looking at the segmentation results in Table 5.6 one can see that these values are not substantially higher than the ones achieved during the online process, shown in Table 5.1 and Table 5.2. This indicates that the model generality is the bottleneck that limits the segmentation results for this method. Finding a way to increase the model

generality e.g. by increasing the training size or using a different way to represent the models would consequently increase the segmentation results.

The effect of the training size is evaluated in Section 5.2.1. Looking at the figures the general trend is that an increase in training size gives a decrease in mean euclidean distance error and an increase in segmentation performance i.e. an increase in generality. This effect somewhat decreases as the training set increases. Worth noting is that including unlabeled data, as done in the heart data set for $n > 19$, also increases the generality of the model. Because of the limited amount of training data these experiments were carried out on quite few targets which may be the reason that some of the results are a bit noisy.

6.6 Future Work

The method used for the alignment of the training set is not quite satisfactory and has a few drawbacks. The thought behind the method used now is that all transformation fields for each target should be aligned in the same way in order for the SDMs to more efficiently capture the variability of the training set. The drawback however, is that the same method can not be used during the online registration process. Hence the affine transformation is recalculated using an alternating optimization method. Changing to an alignment method that can also be used in the online registration method would remove the need to recalculate both the affine transformation and the non-rigid transformation decreasing the running time of the method greatly.

At this moment, the non-rigid transformations used as the training data are represented by deformation fields. Using a more compact representation of the non-rigid transformation such as e.g. a B-spline grid might be a way to build more compact models requiring less data space. This is something that should be investigated in order to speed up the registration process as well as being able to build more compact SDMs.

As previously mentioned, the multi-atlas segmentation results in relation to the pair-wise segmentation results are worse for the Shapereg method compared to the two baseline methods. This is when using a simple majority voting algorithm for label merging and is most likely because of the dependencies between the pair-wise votes induced by the training method. Using a more sophisticated method of merging the labels might suppress this effect and produce better multi-atlas segmentation results. One such method is using Markov Random Fields to produce the multi-atlas segmentation, an example of that is described in [9]. Additionally, the labels produced during the pair-wise segmentations could be weighted differently based on the generality of the SDM used during the registration.

7

Conclusion

In this thesis, the feature-based registration method Shapereg was presented. The method uses statistical deformation models (SDMs) to estimate the non-rigid part of the transformation. The SDMs are built using deformation fields from pair-wise registrations of the atlas images using a reliable but time-consuming method. The Shapereg method was applied to two segmentation tasks, pericardium segmentation in CT images of the heart and hippocampus segmentation in MRI images of the brain.

For the pericardium segmentation in CT images the presented method outperformed the two state-of-the-art methods, referred to as baselines. it was compared against. For the leave-one-out experiments Shapereg achieved a Jaccard index of 0.8307 ± 0.0707 . For hippocampus segmentation in MRI images a Jaccard index of 0.5296 ± 0.0552 was achieved. For this task the presented method was beaten by the intensity-based and more computationally heavy Niftyreg method.

For multi-atlas segmentation the Shapereg method achieved a Jaccard index of 0.8906 ± 0.0341 for pericardium segmentation and 0.5957 ± 0.0504 using a simple majority vote method for merging the labels from the pair-wise registrations. This is slightly lower than for the baseline methods where the feature-based one, Featreg, performed marginally better for the pericardium segmentation and both Featreg and Niftyreg performed slightly better for the hippocampus segmentation.

In addition to showing that SDMs can be a viable option for representing the non-rigid part of the transformation during 3D segmentation tasks, it was also shown that including unlabeled data in the training set may improve the generality of the SDMs as well as increase the segmentation results achieved when they are used.

Bibliography

- [1] D. L. Pham, C. Xu, J. L. Prince, A Survey of Current Methods in Medical Image Segmentation, in: Annual Review of Biomedical Engineering, Vol. 2, 2000, pp. 315–338.
- [2] D. Dey, R. Nakazato, D. Li, D. S. Berman, Epicardial and thoracic fat-Noninvasive measurement and clinical implications, Cardiovascular Diagnosis and Therapy 2 (2) (2012) 85–93.
URL <http://www.thecdt.org/article/view/666/720>
- [3] M. Chupin, E. Gérardin, R. Cuingnet, C. Boutet, L. Lemieux, S. Lehericy, H. Benali, L. Garnero, O. Colliot, Fully automatic hippocampus segmentation and classification in alzheimer’s disease and mild cognitive impairment applied on data from adni, PMC 2.
URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2837195/>
- [4] B. Zitová, J. Flusser, Image registration methods: a survey, Image and Vision Computing 21 (11) (2003) 977 – 1000.
URL <http://www.sciencedirect.com/science/article/pii/S0262885603001379>
- [5] F. Khalifa, G. Beache, G. Gimel’farb, J. Suri, A. El-Baz, State-of-the-art medical image registration methodologies: A survey, in: A. S. El-Baz, R. Acharya U, M. Mirmehdi, J. S. Suri (Eds.), Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies, Springer US, 2011, pp. 235–280.
URL http://dx.doi.org/10.1007/978-1-4419-8195-0_9
- [6] Scapis @ONLINE (Dec. 2014).
URL <http://http://www.hjart-lungfonden.se/Forskning/Vara-forskningssatsningar/scapis/Darfor-ar-SCAPIS-ett-unikt-forskningsprogram/>
- [7] A. Hammers, R. Allom, M. J. Koeppe, S. L. Free, R. Myers, L. Lemieux, T. N. Mitchell, D. J. Brooks, J. S. Duncan, Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe, Human Brain Mapping 19 (4) (2003) 224–247.
URL <http://dx.doi.org/10.1002/hbm.10123>
- [8] J. Alvéén, A. Norlén, O. Enqvist, F. Kahl, Überatlas: Robust speed-up of feature-based registration and multi-atlas segmentation, SCIA, 2015.

- [9] A. Norlén, Automated segmentation of the pericardium using a feature based multi-atlas approach, Master's thesis, Lund University, Centre for Mathematical Sciences (11 2014).
- [10] H. A. Kirisli, M. Schaap, S. Klein, L. A. Neefjes, A. C. Weustink, T. Van Walsum, W. J. Niessen, Fully automatic cardiac segmentation from 3d cta data: a multi-atlas based approach (2010).
URL <http://dx.doi.org/10.1117/12.838370>
- [11] R. Shahzad, D. Bos, C. Metz, A. Rossi, H. Kirisli, A. van der Lugt, S. Klein, J. Witteman, P. de Feyter, W. Niessen, L. van Vliet, T. van Walsum, Automatic quantification of epicardial fat volume on non-enhanced cardiac ct scans using a multi-atlas segmentation approach, *Medical Physics* 40 (9) (2013) –.
URL <http://scitation.aip.org/content/aapm/journal/medphys/40/9/10.1118/1.4817577>
- [12] D. Day, A. Ramesh, P. J. Slomka, R. Nakazato, V. Y. Cheng, G. Germano, D. S. Berman, Automated algorithm for atlas-based segmentation of the heart and pericardium from non-contrast ct, *Proceedings of SPIE* 7623 (2010) 762337–.
URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2953476/>
- [13] J. A. S. Daniel Rueckert, Alejandro F. Frangi, Automatic construction of 3-d statistical deformation models of the brain using nonrigid registration, *IEEE Transaction on medical imaging* 22 (8).
- [14] D. Loeckx, F. Maes, D. Vandermeulen, P. Suetens, Temporal subtraction of thorax cr images using a statistical deformation model, *Medical Imaging, IEEE Transactions on* 22 (11) (2003) 1490–1504.
- [15] M. Modat, Efficient dense non-rigid registration using the free-form deformation framework, Ph.D. thesis, University College London, Gower Street (2012).
- [16] A. Sotiras, C. Davatzikos, N. Paragios, Deformable medical image registration: A survey, *Medical Imaging, IEEE Transactions on* 32 (7) (2013) 1153–1190.
- [17] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, D. Hawkes, Nonrigid registration using free-form deformations: application to breast mr images, *Medical Imaging, IEEE Transactions on* 18 (8) (1999) 712–721.
- [18] J. Karlsson, Towards fully automatic optimal shape modeling, Ph.D. thesis, Lund University (2008).
- [19] T. Cootes, C. Taylor, D. Cooper, J. Graham, Active shape models-their training and application, *Computer Vision and Image Understanding* 61 (1) (1995) 38 – 59.
URL <http://www.sciencedirect.com/science/article/pii/S1077314285710041>
- [20] R. Andersen, *Modern Methods for Robust Regression*, Vol. 152, SAGE Publications, 2455 Teller Road, 2008, Ch. 1, pp. 1–8.
- [21] R. L. Branham, Jr., Alternatives to least squares, *Astronomical Journal* 87 (1982) 928–937.

-
- [22] I. Barrodale, F. D. K. Roberts, An improved algorithm for discrete l_1 linear approximation, *SIAM Journal on Numerical Analysis* 10 (5) (1973) pp. 839–848.
URL <http://www.jstor.org/stable/2156318>
- [23] E. J. Schlossmacher, An iterative technique for absolute deviations curve fitting, *Journal of the American Statistical Association* 68 (344) (1973) 857–859.
URL <http://www.tandfonline.com/doi/abs/10.1080/01621459.1973.10481436>
- [24] P. J. Huber, Robust estimation of a location parameter, *Ann. Math. Statist.* 35 (1) (1964) 73–101.
URL <http://dx.doi.org/10.1214/aoms/1177703732>
- [25] J. E. Gentle, *Matrix Algebra - Theory, Computations and Applications in Statistics*, Springer Science, 233 Spring Street, New York, 2007, Ch. 6.8.1, pp. 229–233.
- [26] R. Chartrand, W. Yin, Iteratively reweighted algorithms for compressive sensing, in: *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 2008, pp. 3869–3872.
- [27] V. Larsson, C. Olsson, F. Kahl, A simple and fast method for robust subspace estimation, to be published.
- [28] D. Lowe, Object recognition from local scale-invariant features, in: *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, Vol. 2, 1999, pp. 1150–1157 vol.2.
- [29] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool, Speeded-up robust features (surf), *Computer Vision and Image Understanding* 110 (3) (2008) 346 – 359, similarity Matching in Computer Vision and Multimedia.
URL <http://www.sciencedirect.com/science/article/pii/S1077314207001555>
- [30] L. Svärm, O. Enqvist, F. Kahl, M. Oskarsson, Improving robustness for inter-subject medical image registration using a feature-based approach, in: *International Symposium on Biomedical Imaging*, 2015.
- [31] M. A. Fischler, R. C. Bolles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM* 24 (6) (1981) 381–395.
URL <http://doi.acm.org/10.1145/358669.358692>
- [32] S. Lee, G. Wolberg, S. Y. Shin, Scattered data interpolation with multilevel b-splines, *IEEE Transactions on Visualization and Computer Graphics* 3 (3) (1997) 228–244.
URL <http://dx.doi.org/10.1109/2945.620490>
- [33] S. Ourselin, A. Roche, G. Subsol, X. Pennec, N. Ayache, Reconstructing a 3d structure from serial histological sections, *Image and Vision Computing* 19 (1–2) (2001) 25 – 31.
URL <http://www.sciencedirect.com/science/article/pii/S0262885600000524>

