FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Fully Automated Modelling, Localisation and Segmentation of 3D CT Hand Data

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Medizinische Informatik

eingereicht von

Helmut Franz Geza Steiner, BSc.
Matrikelnummer 0325571

an der
Fakultät für Informatik der Technischen Universität Wien

Betreuung
Betreuer: a.o.Univ.-Prof. Dipl.-Ing. Dr.techn. Robert Sablatnig
Mitwirkung: René Donner

Wien, 28.04.2011

(Unterschrift Verfasser) (Unterschrift Betreuer)

Abstract

Rheumatoid arthritis is a problem in our aging society. To slow down the progression of the disease early diagnosis and treatment is needed. A continuing quantification of joint space widths and the tracking of loci on bone contours are the main subjects of medical interest. Image segmentation techniques have been successfully used on X-ray images in the past, to speed up medical practice and the evaluation of longitudinal studies on the background of rheumatoid arthritis. As imaging technology advances the use of computed tomography and magnetic resonance imaging in medical diagnoses and studies rises. Therefore existing segmentation techniques like Sparse MRF Appearance Models and Active Shape Models (ASMs) have to be adapted to work on three-dimensional data as well.

This thesis proposes a fully automated segmentation pipeline to handle 3D imaging data. It uses Minimum Description Length and Coherent Point Drift to find point correspondences to build Point Distribution Models. The localisation of anatomical structures is achieved with Random Forests and Discrete Optimisation giving a starting point for the final ASM segmentation.

The proposed methods are evaluated on two different types of data. CT volumes of hands from patients suffering from rheumatoid arthritis are used for the localisation of joint positions and the segmentation of os metacarpale I, phalanx proximalis pollicis and phalanx distalis pollicis. Additionally, synthetic volumes containing bone like objects are created for the evaluation of segmentation accuracy. All methods are analysed individually to yield unbiased results.

The outcome of the experiments on the localisation of anatomical structures are adequate to segment the bones with ASMs. If the object searched for in the unseen image is within the ASM's capture range, precise segmentation results are achieved. Cases where the ASM search is unsuccessful are analysed to find the failures' causes. These and other problems that occurred during the testing of the methods are pointed out and suggestions to possible solutions are given.

Kurzfassung

Rheumatoide Arthritis ist ein Problem in unserer alternden Gesellschaft. Eine frühe Diagnose und Behandlung kann das Voranschreiten der Krankheit verlangsamen. Das kontinuierliche Messen der Gelenkspaltsabstände und die Verfolgung von Loci auf den Knochenkonturen sind die Hauptinteressen im Zusammenhang mit medizinischen Studien. Bildsegmentierungstechniken wurden in der Vergangenheit mit Erfolg auf Röntgenbildern eingesetzt, um die medizinische Praxis und die Evaluierung von Longitudinalstudien zum Thema rheumatoide Arthritis zu beschleunigen. Da bildgebende Verfahren immer ausgereifter werden, ist ein stetiger Anstieg der Benutzung von Computertomographie und Magnetresonanz bei der medizinischen Befundung und in Studien zu bemerken. Demzufolge müssen auch existierende Segmentierungs- methoden, wie Sparse MRF Appearance Models und Active Shape Models (ASMs), an die nun dreidimensionalen Daten angepasst werden, um zu funktionieren.

Diese Diplomarbeit schlägt folgende Segmentierungspipeline für die Handhabung von dreidimensionalen Bilddaten vor: Für das Finden von Punktkorrespondenzen wird Minimum Description Length und Coherent Point Drift eingesetzt, um Point Distribution Models zu bilden. Die Lokalisierung von anatomischen Strukturen wird mit Random Forests und Discrete Optimisation bewerkstelligt, welche einen Startpunkt für die finale ASM Segmentierung liefern.

Die vorgeschlagenen Methoden werden auf zwei unterschiedlichen Datentypen evaluiert. Einerseits werden CT Volumen von Händen von Patienten, die unter rheumatoider Arthritis leiden, für die Lokalisierung von Gelenksspalten und die Segmentierung der Daumenknochen Os Metacarpale I, Phalanx Proximalis Pollicis und Phalanx Distalis Pollicis benutzt. Andererseits werden zusätzlich synthetische Daten für Experimente zur Segmentierungsgenauigkeit generiert. Eine individuelle Analyse aller Methoden wird vollzogen, um von einander unabhängige Resultate zu erhalten.

Die durchgeführten Experimente zur Lokalisierung anatomischer Strukturen liefern brauchbare Resultate für die Knochensegmentierung mit ASMs. Falls das Objekt im ungesehenen Bild sich innerhalb des Suchbereichs des ASM befindet, werden präzise Segmentierungsresultate erzielt. Fälle in denen die ASM Suche nicht erfolgreich ist, werden analysiert, um die Ursachen des Scheiterns zu finden. Diese und andere Probleme, die sich beim Testen der Methoden ergaben, werden aufgezeigt und Lösungsvorschläge zu deren Bewältigung gegeben.

Acknowledgements

First of all I want to thank my parents for their support and financing my studies. Thank you for your patience and believing in me even though completing this thesis took a lot longer than expected. Furthermore I want to thank all of my colleagues and friends at the CIR Lab, especially Lukas Fischer for revising this thesis and always being there for me when I needed help, Michael Sprinzl for the excellent collaboration, Ernst Schwartz for his MATLAB support, Eva Dittrich for her patience, Andreas Burner for his words of advice on all topics of life, Erich Birngruber for his effort and help visualising my results, Georg Langs for the time he spent answering my questions and of course René Donner for his expert advice, revising and co-supervising this thesis. A big thank you also goes to my friend Alexander Valentinitsch who helped me through hard times and always cheered me up when I was frustrated. Finally, I also want to thank Professor Robert Sablatnig for the supervision of this thesis.

Contents

CHAPTER 1

Introduction

The aim of medical image segmentation is to segment known anatomical structures in medical images. Such structures are e.g. bones, vessels, organs, such as lungs or kidneys, as well as brain structures. Segmenting an object of interest allows to draw conclusions about pathological or abnormal conditions of the respective object. The segmentation pipeline proposed in this thesis provides the possibility to locate and segment objects without user interaction. Section 1.1 describes the motivation and the aim of this thesis. An introduction to the medical background is given in Section 1.2. Fundamental terms are described in Section 1.3. The state of the art in medical image segmentation is depicted in Section 1.4. Section 1.5 gives an overview of the main contributions. The chapter is concluded with Section 1.6 which gives a brief outline of this thesis.

1.1 Motivation and Goal of this Thesis

As stated in [Müller10] the use of imaging modalities in hospitals was ever increasing in the last decade yielding following effects: The amount of produced data rises and conventional X-ray images are superseded by three dimensional (3D) imaging modalities such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI).

The manual annotation and segmentation of these datasets is not only a time consuming task but it is also reader dependent [Langs03]. To be able to deliver repeatable results deterministic segmentation approaches like Active Shape Models (ASMs) are used in the medical segmentation process. With the increasing application of MRI and CT the methods to analyse the resulting data have to be broadened to be capable of higher dimensional data. [Heimann09]

The goal of this thesis is to build an automated segmentation pipeline that can handle 3D medical imaging data to aid and accelerate clinical practice during longitudinal studies. A foundation for an automated quantification of Rheumatoid Arthritis (RA) in CT data should be established. So far the degree of joint destruction is either manually determined or automatically quantified by means of image analysis in two dimensional (2D) X-ray images.

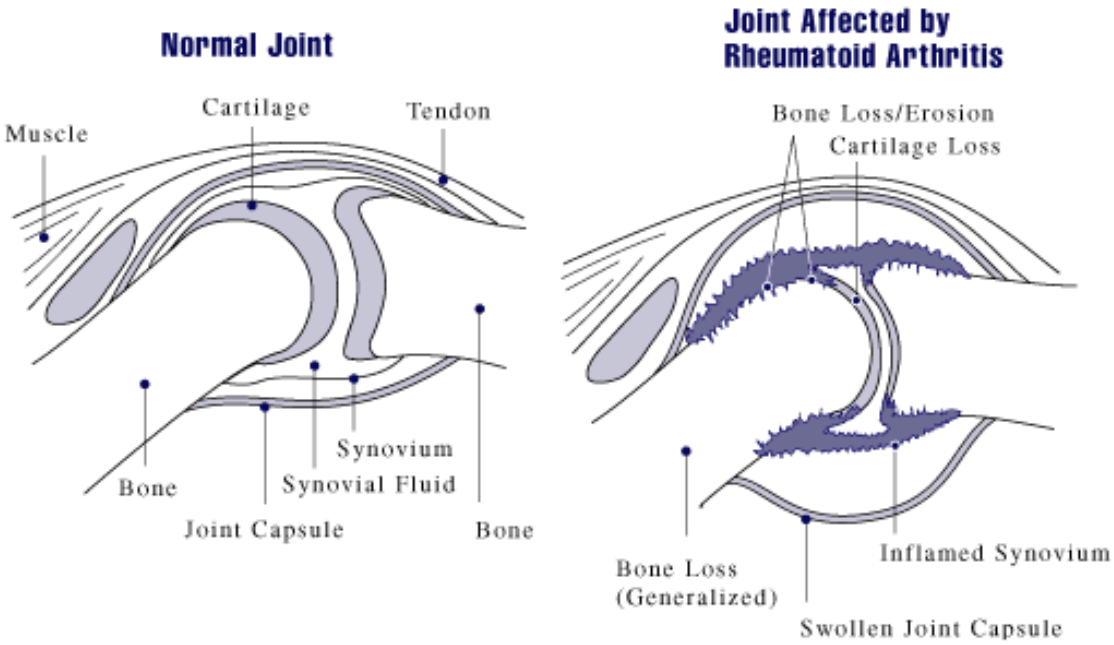


Figure 1.1: Effects of rheumatoid arthritis on a joint (Source: [NIAMS09])

1.2 Medical Background - Rheumatoid Arthritis

RA is a systemic inflammatory disease that primarily affects synovial joints and surrounding tissues (Figure 1.1) but can also have impact on other organs, e.g. the lungs [NIAMS09]. It progresses in three stages [Valentinitsch06]:

1. Stage: Pain, warmth, redness, stiffness and swelling around the joint due to swelling of the synovial lining.
2. Stage: Rapid division and growth of pannus, or cells, causing the synovium to thicken.
3. Stage: Inflamed cells release enzymes that may digest cartilage and bone, which often causes the involved joint to lose its alignment and shape, therefore a loss of movement and more pain.

1.2.1 Causes and Prevalence

[Hamprecht05, Smith10] state the following: The cause of RA is unknown but it is considered a chronic, autoimmune disease that can occur at any age. About one percent of the world's population is affected with an annual incidence of approximately 3 cases per 10,000 population. RA can appear at any age but usually occurs between 40 and 60 years. Women develop RA two to three times more often than men.

RA usually affects fingers, wrists, ankles and feet on both sides of the body equally [Borigini10]. The severity and course of the illness is varying individually and depend on other infections, genes and hormones [Smith10].

1.2.2 Symptoms, Tests and Treatment

RA starts with symptoms that can also be seen in other diseases like loss of appetite, fatigue, swollen glands, low fever and weakness. Eventually, joint pain appears and the disease manifests with following symptoms [Borigini10]:

- Morning stiffness for more than 1 hour / Warm, tender and stiff joints when not used
- Joints are swollen and feel boggy/spongy and warm to the touch
- Loss of range of motion and deformation of joints

Within 1-2 years after manifestation of RA joint destruction may occur [Borigini10]. On further clinical features of RA please see [Harris Jr.08].

To distinguish RA from other types of arthritis like osteoarthritis and psoriatic arthritis a specific blood test (the anti-CCP antibody test) is available [Borigini10]. Following examinations are done additionally to determine the progression of the disease:

- Joint ultrasound
- Joint X-rays
- MRI or CT of the joints

RA requires a lifelong treatment with medications, physical therapy and possibly surgery. An early and aggressive treatment can delay joint destruction [Borigini10]. This has shown the need for accurate early detection, which due to the larger inter-reader differences can only be achieved in a computerised system.

1.2.3 Long-term Prognosis

The progression of RA differs from person to person. "*Spontaneous clinical remission is uncommon (approximately 5-10%). After 5 years of disease, approximately 33% of patients are unable to work; after 10 years, approximately half have substantial functional disability.*" [Smith10]

RA is one of the leading reasons for disability among people aged 15+ and in about 10% of the cases RA leads to death directly or indirectly [Langs03]. This not only means a health issue but also an economic impact as the mean annual direct and indirect costs per person with RA were found to be €4019 and €4090 as stated in [Cooper00].

1.3 Fundamental Terms

This section explains the terms landmarks, shapes and Hounsfield Units (HUs) - a measuring unit used in CT.

1.3.1 Landmarks and Shapes

Landmarks are points of interest in an image. As stated in [Stegmann00] there are three different kinds of landmarks:

- **Anatomical Landmarks** are points that can be clearly identified by an expert (e.g. a doctor) and have some biological meaning.
- **Mathematical Landmarks** are points that are derived from a mathematical or geometrical calculation and are lying at the outline of or within an object.
- **Pseudo Landmarks** are artificially created points that are lying on a border or between other landmarks. The most common method to obtain pseudo landmarks is to sample the outline of an object between landmarks equidistantly.

A shape can be seen as sets of landmarks hence it is a collection of corresponding border points or the characteristic surface configuration of an object: a contour or an outline.

In medical applications landmarks are obtained through manual annotation using tools such as Livewire [Mortensen95]. This method is time consuming and the results are partly subjective and sub-optimal as the annotation depends on the training level of the reader [Langs07]. Due to the progress in medical imaging the number of 3D (e.g. MRI, CT) and even four-dimensional (4D) (e.g. imaging over time) data has increased from year to year yielding in about 100 gigabyte of produced data per day in one hospital [Müller10]. A manual annotation in those volumes is impracticable because of the circumstances to navigate within a 3D space in order to annotate volumes [Dittrich09].

User-guided, semi automatic tools like ITK-SNAP [Yushkevich06] are also used for the annotation of medical images. Unfortunately the level set algorithm used for image segmentation in ITK-SNAP is not capable of dealing with touching bones and erosions/holes inside the bone structure which can be seen in medical images of RA patients. This means that the annotation process has to be done slice by slice.

1.3.2 Hounsfield Units

In CT imaging HUs, named after Sir Godfrey Hounsfield, are computed for each voxel in a CT volume. The HU scale is a linear transformation of the original linear attenuation coefficient measurement. The radiodensity of distilled water at standard pressure and temperature is defined as 0 HU whereas air at the same conditions is defined as -1000 HU. For a material mat with linear attenuation coefficient μ_{mat}, the corresponding HU value is given by

$$HU = \frac{\mu_{mat} - \mu_{water}}{\mu_{water} - \mu_{air}} * 1000$$

where μ_{water} and μ_{air} are the linear attenuation coefficients of water and air. [Brooks77]

Typical values for objects in CT images can be seen in Table 1.1.

Substance	HU
Bone	+400 or more
Soft tissue	+40 to +80
Water	0
Fat	-100 to -60
Lung	-600 to -400
Air	-1000

Table 1.1: Typical Hounsfield values for different parts of the human body

1.4 State of the Art

In this Section methods used for medical image segmentation and RA quantification are described. [Heimann09] states that "*In the last two decades, model-based segmentation approaches have been established as one of the most successful methods for image analysis.*" As described in Section 1.3 level set algorithms are not applicable for the segmentation of RA patients' data. Due to the similar appearance of different hand bones global search algorithms were not taken into account for the segmentation of individual hand bones. With regard to the use of the segmentation pipeline within longitudinal studies only landmark based segmentation techniques were examined to ensure the comparison of loci on the bones. With these constraints in mind the focus of examination laid on Statistical Shape Models (SSMs), which are "*more stable against local image artifacts and perturbations than conventional low-level algorithms*" [Heimann09], due to the a-priori information.

1.4.1 Image Segmentation with SSMs

Segmenting an unseen object in new image data with SSMs consists of the following parts [Heimann09]:

1. Building a model with a set of training examples,
2. search initialisation and
3. the actual object search.

Landmark Correspondences

Before modelling the statistics of a class of shapes a set of training shapes with correspondences is needed. To achieve well-defined shape correspondences several approaches exist which can be categorised in the following way [Heimann09]:

1. **Mesh-to-mesh registration:** A landmarked template mesh is fitted to all surfaces that have been extracted from the training images to propagate its landmarks to all samples.

2. **Mesh-to-volume registration:** A deformable surface with landmarks is adapted to the training images. The correspondences are then determined by the final position of the landmarks.

3. **Volume-to-volume registration:** Instead of adapting a template mesh to all training images, a volumetric atlas is matched. The resulting deformation field is used to propagate landmarks on the atlas to the training data.

4. **Parameterization-to-parameterization registration:** Extracted surfaces from the training images are mapped one-to-one to a common base domain. Landmarks which are defined in the base domain can then be directly transferred to the training surfaces by using the inverse of the mapping function.

5. **Population-based optimisation:** Similar to the parameterization approach, mappings to a common base domain with landmarks are created for all training meshes. The mapping functions are modified in an iterative optimisation resulting in correspondence changes until a cost function which describes the quality of the model has converged.

As Davies et al. state, the Minimum Description Length (MDL) approach (a population-based optimisation) outperformed uniform and SPHARM (SPherical HARMonics, a parameterization based registration) methods in the model building process and is "*both practicable, superior in performance, and suitable for real-world applications*" [Davies10].

SSM Search Initialisation

"*Due to the large size of the search space in 3D, most methods applied to locate an SSM in new image data use local search algorithms that require an initial estimate of the model pose.*" [Heimann09]

One solution to this problem is to initialise a SSM by user interaction. A rough alignment of the position and rotation of the mean shape in the new image is in general sufficient and can be accomplished in negligible time [Heimann09]. In [Hug00] another approach for a manual initialisation is given. A principal control polygon is defined by setting a small number of points (usually three to four). These points can also be detected automatically by using learned confidence regions.

Alternatively the following image processing techniques are used for automated SSM initialisation [Heimann09]:

- Extracted histogram information
- Affine registration using an atlas
- Global search on the entire image using evolutionary algorithms

Heimann et al. used an algorithm based on evolutionary programming and evolution strategies similar to particle filtering for the initialisation of a 3D SSM [Heimann07]. Similarly, particle filtering based approaches inherently combine initialization and segmentation [Fischer10].

Donner et al. have presented a method to automatically locate complex anatomical structures without initialisation by using Markov Random Fields (MRFs) and Sparse MRF Appearance Models (SAMs) [Donner07] [Donner10b]. This approach provides a starting position for further refinement with SSMs and has been extended to 3D with Random Forests (RFs) and discrete optimisation [Donner10a].

SSM Search

No matter how the initialisation is done, the final step in the segmentation process is to find the model of the object in the new image data. Therefore several possibilities to fit a SSM to new images exist, but the most frequently employed constrained search algorithms are Active Shape Models (ASMs) and Active Appearance Models (AAMs) [Heimann09]. This is not only because of their accuracy for finding anatomical structures but also due to their advantageous runtime performance.

Despite being better at segmenting objects, e.g. faces, in 2D images, AAMs have following disadvantage compared to ASM in 3D medical imaging: "*the excessive memory usage of the 3D texture model: To be able to run the method on a normal desktop computer, texture resolution for medium to large structures needs to be scaled down radically. In addition, it appears that the mostly homogeneous texture encountered in typical biological shapes does not deliver the same quality of features as in, e.g. face recognition.*" [Heimann09]

3D SSMs

ASMs have been extended to 3D and used for the segmentation of synthetic generated kidney volumes [Dickens02]. Unfortunately this paper does not deliver any quantitative information on their performance. They were, as well as 3D AAMs, also applied to cardiac MRI data [vanAssen06, Zambal05] but not on bones of the hand providing no direct comparison to the results in this thesis. Nonetheless, a juxtaposition of the resulting error rates is given in Chapter 3.

1.4.2 Measurement of RA in X-Ray Images

Current manual scoring methods for RA have one issue: High inter-reader variability even after intensive training. This results in imprecise and unrepeatable outcomes within studies. [Langs03]

Langs et al. introduced a method to quantify joint space narrowing, using Local Linear Mappings (LLMs) based on texture features to detect the joint positions and ASMs for the bone contours which where refined by snakes to increase the accuracy and allow a fitting in pathological images. The described method also measures the amount of erosion in the bone structure [Langs09].

1.5 Contribution

As part of this thesis methods for the proposed segmentation pipeline were analysed with varying parameter settings. 3D ASMs were implemented for an ASM framework that is capable of 2D and 3D data. The methods proposed for the pipeline were evaluated on medical and synthetic datasets in a Leave-One-Out Cross Validation (LOOCV) framework.

1.6 Synopsis

In Chapter 2 the proposed segmentation pipeline as seen in Figure 2.1 is presented.

Section 2.1 describes the two methods evaluated for finding point correspondences across the training sets: Coherent Point Drift (CPD), a method to perform point set registration was first intended to be used only as an initialisation for the landmark selection with the MDL principle. As the experimental outcomes in Chapter 3 show the MDL implementation did not work properly and therefore was not usable for the landmark creation. This is why CPD replaced MDL in the final segmentation framework.

The localisation of anatomical structure using RF and MRF is described in Section 2.2. It delivers the starting point for the final bone segmentation with ASMs. Principal Component Analysis (PCA) is used for the creation of Point Distribution Models (PDMs) as part of the ASM generation as described in Section 2.3.

The topic of Chapter 3 are the conducted experiments on all methods and their outcome. The testing environment is described in Section 3.1. For the experiments CT volumes as well as synthetic data sets were used. They are described in Section 3.2. Section 3.3 covers the model building experiments with CPD and MDL. The anatomical structure localisation which was tested on manual annotated CT volumes is described in Section 3.4. For the experiments on ASMs (Section 3.5) synthetic as well as CT volumes were used. In specific the thumb bones os metacarpale I, phalanx proximalis pollicis and phalanx distalis pollicis (see Figure 1.2) were used for the segmentation experiments.

The best segmentation results were achieved on phalanx proximalis pollicis whereas the segmentation of os metacarpale I only worked in one out of six cases. An interpretation of the results as well as the preliminaries are given in each section.

The thesis is concluded with Chapter 4, which dissects the findings and gives an outlook on further optimisation. Additional material and results can be found in Appendix A and the acronyms used in this thesis are found in Appendix B.

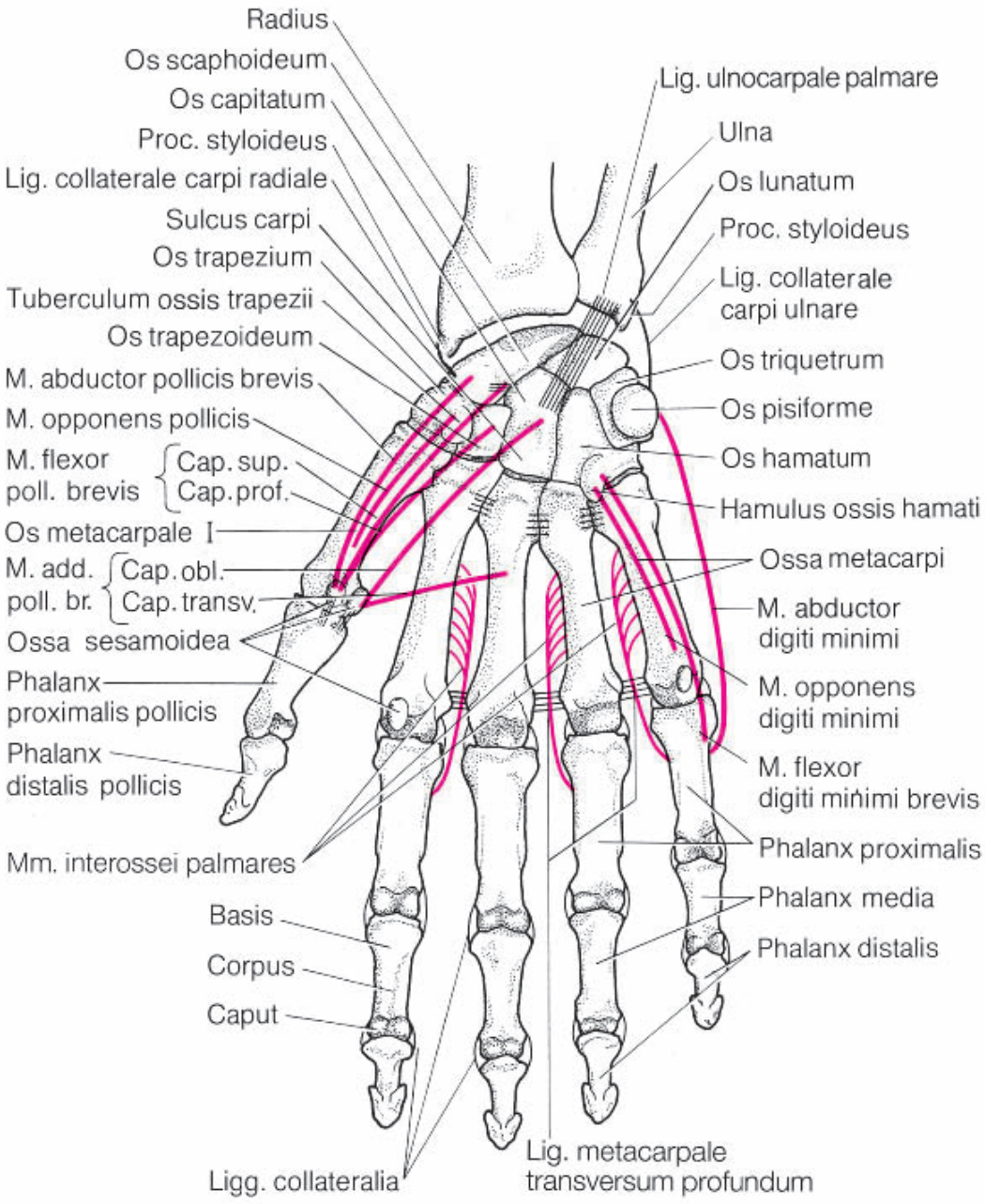


Figure 1.2: Bones of the human hand (Source: [Schiebler07])

CHAPTER 2

Methods

The segmentation pipeline proposed in this thesis consists of three main parts. The first part is to find point correspondences (Section 2.1) over the training sets using CPD and MDL to build SSMs of the bones. The next step is the initial anatomical structure localisation which is performed by means of RFs and discrete optimisation (Section 2.2). It delivers the starting point for the third and final step: the segmentation of bones in the new volume with ASMs (Section 2.3). A flowchart of the segmentation pipeline is shown in Fig. 2.1

2.1 Finding Point Correspondences with CPD and MDL

In this section CPD and MDL are described which where evaluated for the landmark creation. CPD was first intended to be used as an initialisation for the MDL algorithm. As the used implementation of MDL was not capable of finding point correspondences across the data sets (see Chapter 3 for details) only CPD is used in the final segmentation pipeline.

2.1.1 Coherent Point Drift

CPD is a method for point set registration and was introduced by Myronenko et al. in 2007 [Myronenko07]. Its goal is to recover the transformation that maps one point set to the other and/or find the correspondences between them. To be practicable, registration algorithms should have three properties [Myronenko10]:

1. The ability to build an accurate model of the transformation required to align the point sets with each other but maintaining a low computational complexity.

2. The ability to handle also higher dimensioned point sets (>3D) and

3. the robustness to noise, missing points, outliers and other degradations.

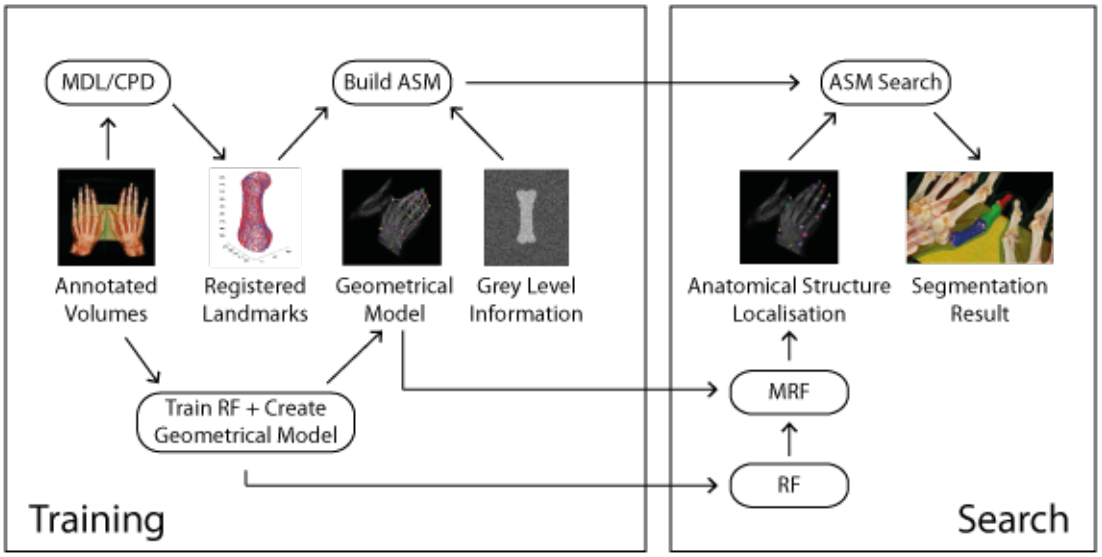


Figure 2.1: Flowchart of the proposed segmentation pipeline.

Transformations are usually distinguished by means of two categories: rigid or non-rigid. Rigid transformations allow rotation, translation and scaling only whereas non-rigid ones, e.g. the simplest form - affine transformation, allow skewing and anisotropic scaling of the point set too.

Due to its low computational complexity and simplicity one of the most popular rigid point set registration methods is the Iterative Closest Point (ICP) algorithm introduced by Besl and McKay [Besl92]. It iteratively assigns correspondences based on a closest distance criterion and finds the least squares transformation relating the point sets. The closest point set is redetermined and the algorithm continues until a local minima is reached. That is why ICP requires a close initial position of the two point sets [Myronenko07].

Among methods that use a probabilistic approach to overcome the ICP limitations the Robust Point Matching (RPM) algorithm introduced by Gold et. al [Gold98] has to be mentioned. It is a non-rigid approach which allows global to local search and uses soft assignment of correspondences between two point sets. Chui and Rangarajan proposed to use Thin Plate Spline (TPS) parametrisation of the transformation which resulted in the TPS-RPM method [Chui00b]. The TPS-RPM algorithm is similar to Expectation Maximisation (EM) for Gaussian Mixture Models (GMMs), where one point set represents centroids of the GMM with equal isotropic covariances and the other set is treated as data points [Chui00a].

CPD was introduced as a non-rigid transformation method but has been extended to be also capable of rigid-only transformations. It is a robust probabilistic multidimensional registration algorithm that considers the alignment of two point sets as a probability estimation problem [Myronenko10]. CPD uses a GMM to fit one point set on the other by maximising the likelihood. The point sets are aligned at the optimum and the correspondences are obtained using the GMM components' posterior probabilities. CPD is similar to [Chui00a] but does not make any assumption of the transformation model.

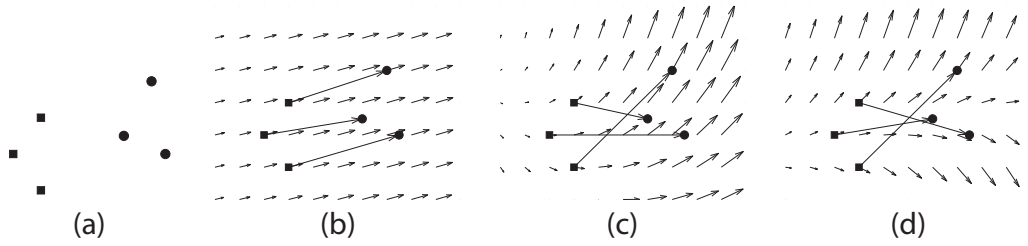


Figure 2.2: (a) Two point sets. (b) Coherent movement. (c & d) Less coherent velocity fields for the given correspondences. (Source: [Myronenko07])

The process of adapting the Gaussian centroids from initial to their final positions is seen as a temporal motion process instead. Therefore a motion coherence constraint is imposed over the velocity field which gives smoothness on the underlying transformation [Myronenko07]. The Motion Coherence Theory [Yuille88] states that points close to each other tend to move coherently. This constraint penalises the underlying velocity field's derivatives. In Figure 2.2 examples of velocity fields, two point sets and their more or less coherent movements are given. In the CPD algorithm the coherence constrained is imposed by regularisation of the displacement field (for smooth non-rigid transformations) or by explicit re-parametrisation of GMM centroid locations (for affine and rigid transformations), [Myronenko10].

In real world problems, as medical image registration, non-rigid transformations are more prominent than rigid transformations [Myronenko10]. In this thesis only non-rigid CPD was used and therefore only this method is further described.

Expectation Maximisation

In CPD the Expectation Maximisation (EM) algorithm is used [Dempster77]. The EM algorithm consists of two steps:

1. **Expectation or E-step:** The values of parameters are first estimated to further compute a posteriori probability distributions of mixture components via the Bayes' theorem.

2. **Maximisation or M-step:** The new parameter values are then computed by minimising the expectation of the negative log-likelihood function (containing the estimated values) with respect to the new parameters.

For a detailed description see [Myronenko10].

The Coherent Point Drift Algorithm

To align the point set $\mathbf{Y}_{M\times D} = (\mathbf{y}_1, \ldots, \mathbf{y}_i, \ldots, \mathbf{y}_M)^T$ with the reference point set $\mathbf{X}_{N\times D} = (\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N)^T$, the points in $\mathbf{Y}$ are considered as GMM centroids ($\mathbf{y}_i = (y_{i1}, \ldots, y_{iD})$, $\mathbf{x}_i$ likewise). The first step is to initialise the free parameters w, λ and β, the weight matrix $\mathbf{W}$,

1. Initialisation:
   - Initialise $\mathbf{W}_{M\times D} = (\mathbf{w}_1, \ldots, \mathbf{w}_M)^T$, a matrix of Gaussian kernel weights, with zero.
   - $\sigma^2 = \frac{1}{DNM}\sum_{m,n=1}^{M,N} \|\mathbf{x}_n - \mathbf{y}_m\|^2$
   - $0 \leq w \leq 1, \beta > 0, \lambda > 0$
2. Construct the Gaussian kernel matrix $\mathbf{G}_{M\times M}$ with $g_{ij} = G(\mathbf{y}_i, \mathbf{y}_j) = e^{-\frac{1}{2\beta^2}\|\mathbf{y}_i - \mathbf{y}_j\|^2}$
3. EM optimisation (repeated until convergence)
   - E-step: Compute the matrix of posterior probabilities $\mathbf{P}_{M\times N}$
   - M-step:
     - Solve $(\mathbf{G} + \lambda\sigma^2 \mathrm{d}(\mathbf{P1})^{-1})\mathbf{W} = d(\mathbf{P1})^{-1}\mathbf{PX} - \mathbf{Y}$
     - $N_\mathbf{P} = \mathbf{1}^T\mathbf{P1}$
     - Compute the aligned point set $\mathbf{T} = \mathbf{Y} + \mathbf{GW}$
     - Anneal $\sigma^2 = \frac{1}{N_\mathbf{P} D}(\mathrm{tr}(\mathbf{X}^T \mathrm{d}(\mathbf{P}^T\mathbf{1})\mathbf{X}) - 2\mathrm{tr}((\mathbf{PX})^T\mathbf{T}) + \mathrm{tr}(\mathbf{T}^T\mathrm{d}(\mathbf{P1})\mathbf{T}))$

Figure 2.3: Meta-code for the CPD algorithm , where $\mathrm{d}(\mathbf{a})$ is a diagonal matrix formed from a vector $\mathbf{a}$, $\mathrm{tr}(\mathbf{A})$ is the trace of a square matrix $\mathbf{A}$, i.e. the sum of the elements on the main diagonal, $\mathbf{I}$ is the identity matrix, and $\mathbf{1}$ is a column vector of ones [Myronenko10].

and the variance σ^2 of the GMM. w reflects the assumption on the amount of noise in the point sets, β defines the model of the smoothness regulariser and λ represents the trade-off between the regularisation and the goodness of maximum likelihood fit [Myronenko10].

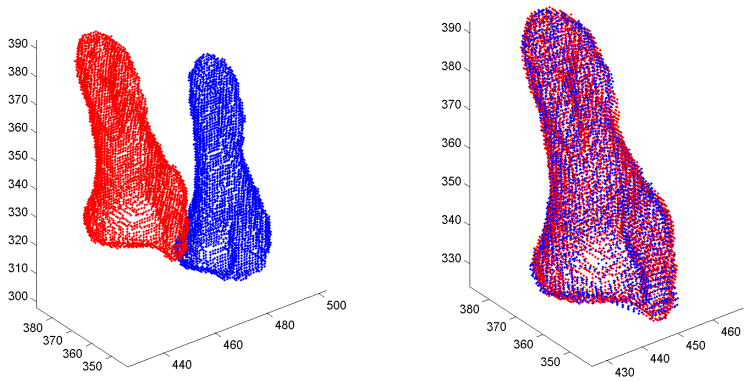
In the next step a Gaussian kernel matrix $\mathbf{G}$ is constructed which is used as a low-pass filter for a regularisation term to ensure smooth movement in the EM according to the Motion Coherence Theory. $\mathbf{G}$ is not related to the Gaussian form of the mixture model.

In the last step the EM is optimised. Therefore the matrix $\mathbf{P}$ of posterior probabilities p_{mn} is computed in the expectation step:

$$p_{mn} = \frac{e^{-\frac{1}{2\sigma^2}\|\mathbf{x}_n - (\mathbf{y}_m + \mathbf{G}(m,\cdot)\mathbf{W})\|^2}}{\sum_{k=1}^{M} e^{-\frac{1}{2\sigma^2}\|\mathbf{x}_n - (\mathbf{y}_k + \mathbf{G}(k,\cdot)\mathbf{W})\|^2} + \frac{w}{1-w}\frac{(2\pi\sigma^2)^{D/2}M}{N}} \tag{2.1}$$

In the maximisation step the resulting system is solved for $\mathbf{W}$ which is used to compute the aligned point set $\mathbf{T}$. The EM optimisation is repeated until convergence by annealing the value σ^2 [Myronenko10].

The meta-code for the CPD algorithm can be seen in Fig. 2.3. CPD shows accurate alignments even with outliers or missing points and it has been shown that it outperforms ICP and TPS-RPM. For a detailed description of the results, the algorithm and the computation see [Myronenko07] and [Myronenko10]. In Figure 2.4 an example of a non-rigid registration of phalanx distalis pollicis is shown.



(a) (b)

Figure 2.4: (a) Shows two phalanx distalis pollicis point sets before registration and (b) after.

2.1.2 Minimum Description Length

The MDL Principle has been mainly developed by J. Rissanen in 1978 [Rissanen78]. Its roots lie in the theory of Kolmogorov or algorithmic complexity. MDL was also inspired by the first model selection method based on information-theoretic ideas, the AIC (Akaike Information Criterion) method, developed by Akaike in 1973, but the underlying philosophy and the actual method are both substantially different from MDL [Grünwald05].

The MDL Principle is closer related to the Minimum Message Length Principle developed by Wallace et al. in 1968 [Wallace68]. It is a method for inductive and statistical inference that provides a generic solution to the model selection problem.

MDL is based on the idea that any regularity in the data can be used to compress it, which means fewer symbols to describe the data are needed than the number of symbols which would be used to describe it literally. More regularities in the data therefore mean more possible compression. This equates finding regularity with learning. Hence the more the data can be compressed the more has been learned about it.

If formalised, this leads to a general theory of inductive inference with following five properties [Grünwald05]:

1. **Occam's Razor** MDL embodies a form of Occam's Razor by choosing a model that trades off goodness of fit on the observed data with complexity of the model.

2. **No over-fitting** MDL protects automatically and inherently against over-fitting and can be used to estimate the parameters and the structure of a model. Traditional methods for estimating the structure have to be modified and/or extended to ensure no over-fitting.

3. **Bayesian interpretation** The MDL Principle is closely related to Bayesian inference but avoids some difficulties of the Bayesian approach especially in the (realistic) case when it's a priori known that none of the considered models is true. Note that in contrast to

Bayesian inference which is formulated in a probabilistic framework, MDL is purely data driven [Vitanyi00].

4. **No need for underlying truth** Whether or not there exists an underlying true model, MDL procedures have a clear, independent interpretation.

5. **Predictive interpretation** MDL methods can be interpreted as searching for a model with good predictive performance on unseen data because of probabilistic prediction been seen as formally equivalent to data compression.

The difference between a hypothesis and a model is declared by [Grünwald05]:

- **Point hypotheses** are single probability distributions or functions, also known as simple hypotheses in statistical literature (e.g. the polynomial $3x^2 + 5x + 4$).
- **Models** are sets or families of probability distributions or functions with the same functional form, also known as composite hypotheses (e.g. all third-degree polynomials).
- **Hypothesis** as a generic term that refers to both point hypotheses and models.

Crude MDL

The first implementation of the MDL idea is the so called two-part code version [Grünwald05]: To explain the data D the best point hypothesis $H \in \mathcal{H}^{(1)} \cup \mathcal{H}^{(2)} \cup \ldots \cup \mathcal{H}^{(n)}$ is the one that minimises the sum

$$L(H) + L(D|H), \text{ where} \tag{2.2}$$

$\mathcal{H}^{(1)}, \mathcal{H}^{(2)}, \ldots$ is a list of candidate models, each containing a set of point hypotheses,

$L(H)$ is the length of the hypothesis' description (in bits) and

$L(D|H)$ is the length of the data's description (in bits) when encoded by means of the hypothesis.

Hence, the best model to explain the data D is the smallest model containing the selected hypothesis H. Neither a very complex point hypothesis with a perfect fit (large $L(H)$ - small $L(D|H)$) nor a very simple point hypothesis with a bad fit (small $L(H)$ - large $L(D|H)$) will minimise the sum of the two description lengths. This means that the best results are achieved with a trade-off between $L(H)$ and $L(D|H)$ [Grünwald05].

Refined MDL

While it can be shown that finding a code for $L(D|H)$ is not a problem and such a (Shannon-Fano) code always exists [Grünwald05] the definition of $L(H)$ is not trivial and gives arbitrary results. To circumvent this Rissanen introduced a one-part rather than a two-part code which is also called refined MDL [Rissanen96]:

The data D is not encoded with a single $H \in \mathcal{H}$ but with the full model $\mathcal{H}$ in a one-part code with length $\bar{L}(D|\mathcal{H})$. The code-length $\bar{L}(D|\mathcal{H})$ will be small if $L(D|H)$ is small

(a $H \in \mathcal{H}$ fits the data well). The ability of a model $\mathcal{H}$ to fit random data is indicated by its parametric complexity, $\mathbf{COMP}(\mathcal{H})$, which is the second fundamental concept of refined MDL [Grünwald05].

In the one-part code version of MDL the model with the smallest stochastic complexity of D is chosen which is given by

$$L(D|\hat{H}) + \mathbf{COMP}(\mathcal{H}), \text{ where} \tag{2.3}$$

$\hat{H}$ is the distribution in $\mathcal{H}$ that minimises the code-length $L(D|\hat{H})$.

This still means a trade-off between a complexity and a goodness-of-fit term but avoids the explicit encoding of a hypothesis H.

Usage of MDL for Landmark Generation

In [Thodberg03] and [Davies10] the MDL principle is used to find point correspondences in 2D and 3D training examples for SSMs. The two approaches reduce the number of shape pixels/voxels (represented as the shape's contour) of the used data sets resulting from manual annotation and are used to obtain landmarks for the creation of PDMs. In contrast to Davies and Thodberg who work on surfaces, Langs uses MDL directly on volumes [Langs07]. His approach does not require any annotation and the topology can change throughout the data set.

2.2 Anatomical Structure Localisation with RFs and MRFs

As described in Chapter 1 a segmentation via local SSMs needs an initial model position. For this purpose a generic method that identifies anatomical structures in a global search framework is used. The local appearance of landmarks and an elastic shape constraint from annotated training volumes are learned. A classifier is used for the generation of candidate points during search and the final location is computed by discrete optimisation [Donner10a]. An overview of the method is shown in Figure 2.5.

2.2.1 Sparse MRF Appearance Models

The most closely related approach to the method described are SAMs. By solving a MRF, they match shape and appearance models to query images. SAMs are based on interest points and an elastic geometric model of their spatial configuration. Local descriptors around the interest points and along the connecting edges of a Delaunay triangulation encode the appearance of an anatomical structure. The model thus encompasses

- information about the mesh topology,
- mean and standard deviation of edge lengths,
- circular statistics of edge directions relative to interest point orientation and
- the local point and edge descriptors.

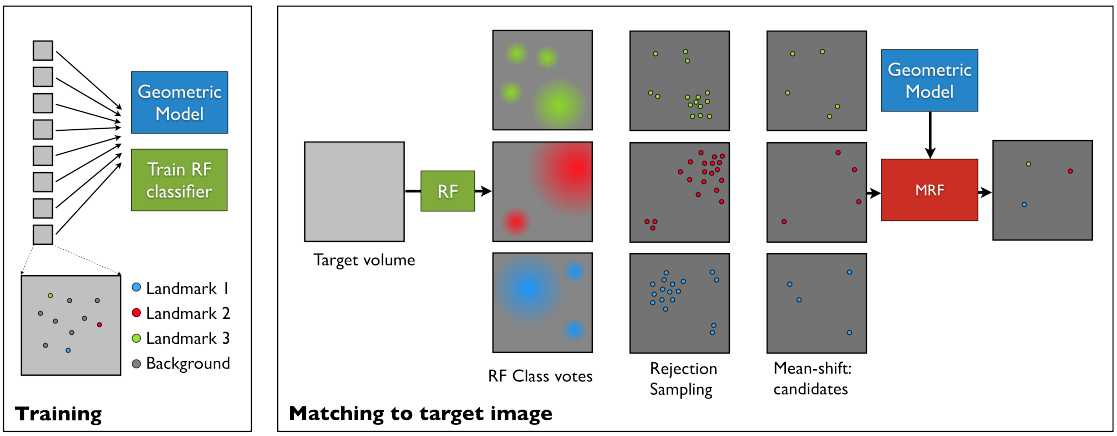


Figure 2.5: Outline of the interest point detection and model matching method used. (Source: [Donner10a])

Interest points and descriptors are computed to localise the structure in a target image. By solving a MRF the most probable match of the model onto the target image is found [Donner07].

Despite good results following issues remain, that prevent SAMs from having been applied to 3D data sets yet [Donner10a]: Different interest point detectors may be needed for the different sub-parts, depending on the structure [Donner10b] yielding in a greatly increased number of labels for the MRF. This inhibits fast inference and increases memory requirements. Due to the typically overwhelming number of detected interest points the approach is infeasible in 3D. To localise anatomical structures in 3D several changes to the SAM approach were made [Donner10a].

2.2.2 Domain Specific Interest Points

To detect interest points [Donner10a] trains a RF classifier on Haar wavelet-like descriptors around the model landmarks. The resulting classification volumes are sampled into point sets and clustered to obtain target point candidates specific for each landmark. From small, local regions which are used as unary node costs during discrete optimisation, interest points are derived.

Haar-like Features

A set of 3D features computed using a basis of filters similar to Haar wavelets is employed for describing the local appearance around the model landmarks. Using integral volumes [Ke05]These features [Viola01] can be computed in a highly efficient manner [Donner10a].

Random Forest based Appearance Learning and Search

RFs [Breiman01] are ensemble classifiers which learn a set of decision trees by randomly sampling from training feature vectors and corresponding labels. The decision trees vote for a class

label for each query feature vector during search.

On a new target volume all voxels are classified by the RF and descriptors for every voxel are computed. The RFs votes are normalised for each class yielding in L volumes $\mathbf{C}^l$ that contain the classification probabilities for class l in each voxel. In the next step landmark candidates are generated by means of these volumes. For details see [Donner10a].

Mean-shift based Interest Point Generation

Mean-shift [Cheng95, Comaniciu02] is a method for density estimation and cluster analysis of a sparse set of points in a feature space which is potentially high-dimensional. Each data point $\mathbf{d}_i$ in a d-dimensional dataset $\mathbf{D}$ is moved iteratively towards the mean of the data points within a chosen bandwidth around $\mathbf{d}_i$. The process is repeated until equilibrium is reached, i.e. when all data points have reached their final position and are not shifted any more [Donner10a].

In the test volume regions within the classification probability volume $\mathbf{C}^l$ are searched with high local support to find the most probable candidates for each model landmark l, i.e. regions where their Gaussian weighted integral yields high values. This process results in sets of cluster centres which represent interest points, or candidates, $\mathbf{p}_i^l$ for each model landmark l together with estimates of their local support s_i^l, i.e. the number of points converged to the cluster centre [Donner10a].

2.2.3 3D Geometric Model Matching using Discrete Optimisation

To constrain the set of landmark candidates, the spatial configuration of the landmarks is used, yielding in a highly probable landmark assignment in the search volume. The assignment is a trade-off between the plausibility of the spatial configuration and the local appearance at the landmark position [Donner10a].

An elastic geometric model, together with confidences about point and edge similarity is obtained from the training data. The local similarities of the interest points found in the query volume to the model learnt are then encoded in a MRF. The localisation of the anatomical structure in the target volume is equal to solving the graphical model which matches the model to the query image [Donner10a].

Formulating the MRF

To match the learnt model to a new volume the objective function is

$$Conf(\mathcal{S}) = \sum_{l=1\ldots L} \mathcal{L}(l, \mathcal{S}(l)) + \sum_{e=1\ldots E} \mathcal{E}(e, \mathcal{S}(e)), \tag{2.4}$$

where $\mathcal{L}$ are unary terms at the nodes of the graphical model describing the point candidate to L model landmarks similarities. $\mathcal{E}$ are binary terms that capture the similarities of the target edges to the E model edges. For the MRF's solution

$$\mathcal{S}^* = \text{argmax}_{\mathcal{S}} Conf(\mathcal{S}) \tag{2.5}$$

each model node l is assigned to one point candidate $\mathbf{p}_i^l$ in the target image. Thus matching the model to the target volume [Donner10a].

2.3 Bone Segmentation via ASMs

Active Shape Models (ASMs) were first introduced by Tim Cootes and Chris Taylor in 1992 (see [Cootes92a]). ASMs use PDMs to represent objects as sets of labelled points and to find new examples of modelled objects in images.

Since a search using ASMs is based on local adjustments the starting position must be close to the object that is searched for. Multi-resolution ASMs where introduced by Cootes in 1994 (see [Cootes94b]) to loosen the constraint given by the starting position.

In this section a short introduction to PCA is given. Then PDMs are described followed by a description of ASMs and their extension to 3D.

2.3.1 Principal Component Analysis

PCA was first introduced by Pearson in [Pearson01]. It is a mathematical technique to convert a set of observations of possibly correlated variables into principal components (a set of values of uncorrelated variables) by using an orthogonal transformation. Pearson defined PCA as the linear projection that minimises the average projection cost (the mean squared distance between the data points and their projection) [Pearson01]. In other words, the PCA is the orthogonal projection of data onto a lower dimensional linear space, which maximises the variance of the projected data.

PCA by Eigenvalue Decomposition

Supposing a $m \times nd$ - matrix $\mathbf{A}$, where m denotes the number of samples, e.g. the number of shapes, and nd is the number of landmarks n times their dimensionality d. The first step in PCA is to normalise all samples $\mathbf{A}_i = (\mathbf{A}_{1i}, \ldots, \mathbf{A}_{mi})^T$, where $i \in \{1, \ldots, nd\}$, by subtracting their mean. This is done to assure that the data is centred around the origin. Otherwise a wrong direction for the eigenvectors is calculated. The next step is to compute the covariance matrix $\mathbf{C}$ of the nd-dimensional data [Smith02]:

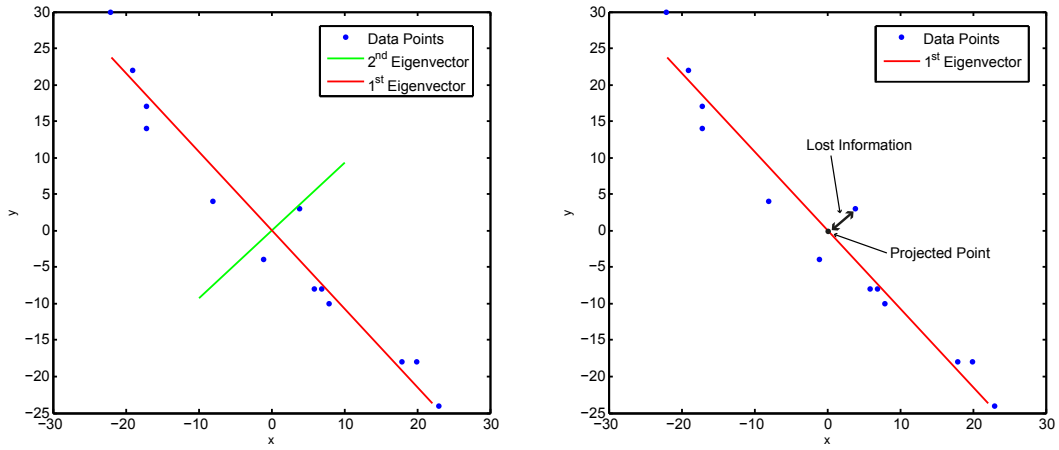
$$\mathbf{C} = \begin{pmatrix} cov(\mathbf{A}_1, \mathbf{A}_1) & \ldots & cov(\mathbf{A}_1, \mathbf{A}_i) \\ \vdots & \ddots & \vdots \\ cov(\mathbf{A}_i, \mathbf{A}_1) & \ldots & cov(\mathbf{A}_i, \mathbf{A}_i) \end{pmatrix}, \tag{2.6}$$

where $i \in \{1, \ldots, nd\}$. Having calculated $\mathbf{C}$, the matrix $\mathbf{V}$ of eigenvectors ϕ which diagonalises the covariance matrix $\mathbf{C}$ is computed with

$$\mathbf{V^{-1}CV} = \mathbf{D}, \tag{2.7}$$

where $\mathbf{D}$ is the diagonal matrix of eigenvalues λ of $\mathbf{C}$. This results in $\min(m, nd)$ eigenvectors. The eigenvector with the largest eigenvalue is the data set's principle component.

The eigenvectors ϕ_i $(nd \times 1)$ with a norm of 1 are sorted in decreasing order of their corresponding eigenvalues λ_i. This results in the rotated basis $\mathbf{F} = (\phi_1, \ldots, \phi_k)$, where ϕ_i $(i \in (1, \ldots, k))$ are the sorted eigenvectors. To decrease the dimensionality of the data only the first t eigenvectors $(\phi_1, \ldots, \phi_t, \ldots, \phi_k)$ are used which cover a certain percentage (e.g. 95%) of



(a) (b)

Figure 2.6: (a) shows the result of a PCA of m 2D data points (blue) with their two principal components/eigenvectors (red, green). In (b) the data's dimensionality is reduced to one dimension. The orthogonal information (black) is lost.

data variation, leaving out the last $k-t$ ones with smallest eigenvalues. The lost information can be neglected if the discarded eigenvectors are small (e.g. they only cover 5% of data variation) [Smith02].

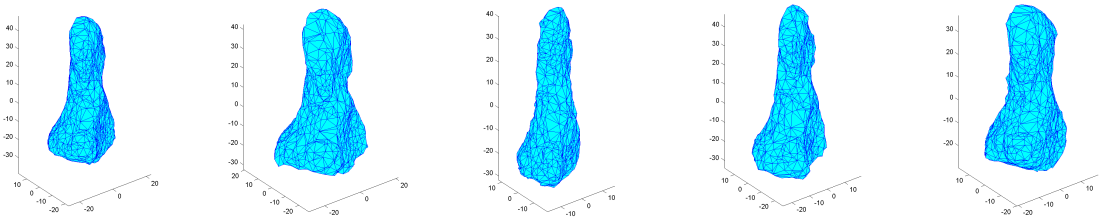
An example for the reduction of the dimensionality from m 2D data points ($nd = 2$) to one dimension is given in Figure 2.6. In (a) the red and green line represent the eigenvectors. The data's dimensionality is reduced to one dimension (b) by projecting onto the red line. The orthogonal information (black) is lost.

PCA is a statistical method that can be found in applications like structural tensors, AAMs, PCA-SIFT (Scale-Invariant Feature Transform), etc. In this thesis the PCA is used for the construction of PDMs and Grey Level Models (GLMs) for the ASM generation.

2.3.2 Point Distribution Model

The PDM is a way to represent a class of shapes using labelled points (landmarks) and their position as described in [Cootes92b].

The first step in creating a shape model is to align all shapes gathered in the training process with the first one. This is achieved by Procrustes analysis [Goodall91] which finds the scale, rotation and translation that minimises the sum of squared distances between their corresponding landmarks. Using PCA the shape vector dimensionality is reduced to the t most significant modes of variation. The modes of variations are the ways in which the landmarks of a shape tend to move together. The first t eigenvectors ϕ_i that correspond to the largest eigenvalues λ_i describe the t most significant modes of variation [Cootes92b]. Any shape seen in the training



(a) (b) (c) (d) (e)

Figure 2.7: A phalanx distalis pollicis model with (a) $\mathbf{b} = 0$ (the mean shape). (b & c) show the variations of the first mode with (b) $b_1 = 2\sqrt{\lambda_1}$ and (c) $b_1 = -2\sqrt{\lambda_1}$. (d & e) show the variations of the second mode likewise.

set can be approximated using the model

$$\mathbf{x} \approx \bar{\mathbf{x}} + \mathbf{\Phi}\mathbf{b}, \tag{2.8}$$

where the vector $\mathbf{x}$ of length nd represents the n landmarks of the d-dimensional shape

$$\mathbf{x} = (x_{11}, \ldots, x_{1d}, x_{21}, \ldots, x_{2d}, \ldots, x_{n1}, \ldots, x_{nd})^T.$$

$\bar{\mathbf{x}}$ is the mean shape. $\mathbf{\Phi} = (\phi_1\ \phi_2\ \ldots\ \phi_t)$ is the matrix of the first t eigenvectors, ϕ_i, derived from a PCA of the training shapes and $\mathbf{b} = (b_1\ b_2\ \ldots\ b_t)^T$ is a vector of weights for each eigenvector.

The model can only vary in ways seen in the training set and is constrained by the weight vector $\mathbf{b}$. As the columns of $\mathbf{\Phi}$ are orthogonal $\mathbf{\Phi}^T\mathbf{\Phi} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, and

$$\mathbf{b} = \mathbf{\Phi}^T(\mathbf{x} - \bar{\mathbf{x}}) \tag{2.9}$$

can be derived from Equation 2.8. $\mathbf{b}$ is used to constrain the model to vary only in ways seen in the training set. The eigenvectors $\mathbf{\Phi}$ and the mean shape $\bar{\mathbf{x}}$ are estimated from a set of training examples and the above equations allow to generate new examples from the class of shapes. This is done by varying the parameters b_i. As stated in [Cootes95] the limits of variation for b_i are typically within three standard deviations $\sigma_i = \sqrt{\lambda_i}$ of the corresponding eigenvalue λ_i of the training set,

$$-3\sqrt{\lambda_i} \leq b_i \leq 3\sqrt{\lambda_i}. \tag{2.10}$$

Figure 2.7 shows a model of phalanx distalis pollicis with variations of the first two modes within those limits. For figures of all thumb bone models please see Appendix A.

The shape of a model instance can be defined (in an object centred co-ordinate frame) by choosing values for $\mathbf{b}$. An instance, $\mathbf{X}$, of the model in the image frame is created by defining the position, orientation and scale:

$$\mathbf{X} = M(s, \theta)[\mathbf{x}] + \mathbf{X}_c \tag{2.11}$$

where $\mathbf{X}_c = (X_{c1}, \ldots, X_{cd}, \ldots, X_{c1}, \ldots, X_{cd})^T$ is n times the d-dimensional model centre position $(X_{c1}, \ldots, X_{cd})$ in the image frame. $M(s, \theta)[\mathbf{x}]$ is a scaling by s and rotation by angle θ of the model $\mathbf{x}$. See [Cootes92b] and [Cootes95] for detailed explanation.

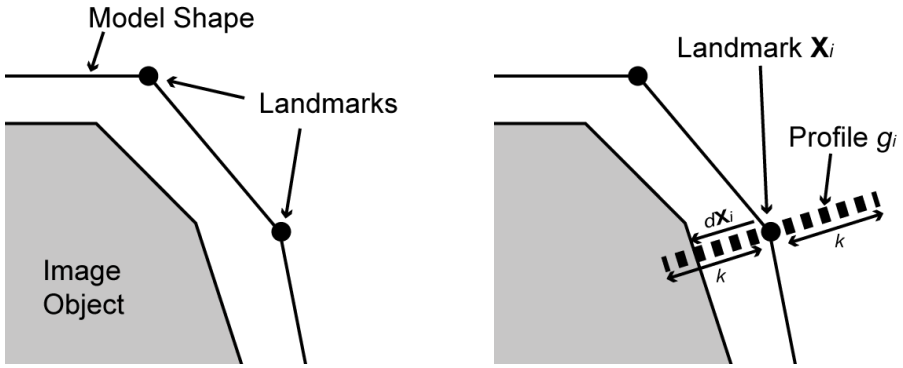


Figure 2.8: (a) Model shape near the edge of an image object. (b) Suggested movement $d\mathbf{X}_i$ of landmark i according to a sampled grey level profile g_i.

2.3.3 Image Search Using ASMs

Given a rough starting approximation ASM search will fit the model $\mathbf{x}$ to the unseen object in the new image by using an iterative scheme [Cootes95]:

After generating a flexible model, a new instance $\mathbf{X}$ is created using Equation 2.11. By choosing values for $\mathbf{b}$ (see Equation 2.9), which define the shape of the model object, the initial object can be varied.

The current estimate of $\mathbf{X}$ is placed into the image. A region of the image around each landmark is examined to calculate a displacement which moves the landmark to a better position. These local adjustments are then transformed to update the pose, scale and shape parameters of the PDM. Limiting the shape parameters, which means applying global shape constraints from Equation 2.9, ensures that the shape of the model example remains similar to the shapes seen in the training set.

These steps are repeated until there are no significant changes, i.e. the mode changes from one iteration to the next are below a user defined threshold.

Calculating Suggested Movements of Landmarks

In Figure 2.8(a) landmarks which represent a shape are shown. To find new objects in an image those landmarks have to be moved closer to the actual image object.

One way of finding displacements for new landmark positions during the ASM search is to find the strongest edge along a normal of the model's shape and move the point towards this edge proportional to the strength of the edge [Cootes95].

Another approach which is used in this thesis to improve the segmentation precision [Cootes94a] is to compute the displacement $d\mathbf{X}_i$ according to local grey level environments of each landmark i ($i \in \{1, \ldots, n\}$). Using these environments a GLM for each point is computed.

Therefore k pixels on both sides of the landmark i are sampled in a training image l ($l \in \{1, \ldots, m\}$) along a normal to the shapes boundary (Figure 2.8(b)) which results in a profile

vector $\mathbf{g}_i^l = (\mathbf{g}_{i1}^l, \ldots, \mathbf{g}_{ij}^l)$ with j pixel samples ($j \in \{1, \ldots, 2k+1\}$). The best results are achieved ([Cootes93]) when using either the original grey-values or a normalised profile of the first derivative $\mathbf{h}_i^l$ of the grey-values,

$$\mathbf{g}_i^l \rightarrow \frac{1}{\sum_j |h_{ij}^l|} \mathbf{h}_i^l \tag{2.12}$$

This is done for each landmark in every training image to obtain sets $g_i = \{\mathbf{g}_i^1, \ldots, \mathbf{g}_i^m\}$. A statistical model $\bar{\mathbf{g}}, \mathbf{\Phi}_g$ for the grey level profile at landmark i is retrieved (similar to the PDM creation) by computing the mean $\bar{\mathbf{g}}$ and the covariance $\mathbf{S}_g$ of the set g_i and using PCA [Cootes94a]:

$$\mathbf{g}_{new} \approx \bar{\mathbf{g}} + \mathbf{\Phi}_g \mathbf{b}_{gnew}, \tag{2.13}$$

where $\mathbf{\Phi}_g$ contains the eigenvectors of $\mathbf{S}_g$ corresponding to the t_g largest eigenvalues λ_j ($j \in \{1, \ldots t_g\}$) and $\mathbf{b}_{gnew}$ is a set of t_g parameters describing the profile model.

During search a profile $\mathbf{g}$ for each point is sampled with length $2(q-k)+1$ where q is the amount of pixels on each side of the current point ($q > k$). The parameters to best fit $\mathbf{g}$ to the model at a point are given by

$$\mathbf{b}_g = \mathbf{\Phi}_g^T (\mathbf{g} - \bar{\mathbf{g}}) \tag{2.14}$$

and therefore the best fit of the model to $\mathbf{g}$ is according to Equation 2.13

$$\mathbf{g}_{best\_fit} = \bar{\mathbf{g}} + \mathbf{\Phi}_g \mathbf{b}_g \tag{2.15}$$

The sum of squares of differences R^2 between $\mathbf{g}_{best\_fit}$ and $\mathbf{g}$ is

$$R^2 = (\mathbf{g} - \mathbf{g}_{best\_fit})^T (\mathbf{g} - \mathbf{g}_{best\_fit}) \tag{2.16}$$

which is equal to [Cootes94a]

$$R^2 = (\mathbf{g} - \bar{\mathbf{g}})^T (\mathbf{g} - \bar{\mathbf{g}}) - \mathbf{b}_g^T \mathbf{b}_g. \tag{2.17}$$

A measure of how well the model fits the profile is given by an approximation of the Mahalanobis distance [Cootes94a]

$$F = \sum_{j=1}^{t_g} \frac{\mathbf{b}_{gj}}{\lambda_j} + \frac{2R^2}{\lambda_{t_g}}, \tag{2.18}$$

where all $\lambda_j = 0.5\lambda_{t_g}$ for $j > t_g$. As depicted in Figure 2.9, F approaches zero as the quality of fit improves. For a detailed description and the derivation of the formulas see [Cootes94a].

This method improves the performance of ASMs as can be seen in [Cootes93].For further details also see [Cootes04].

The set of adjustments calculated for each point (Figure 2.10) with either of the methods described above are denoted as a vector

$$d\mathbf{X} = (dX_{11}, \ldots, dX_{1d}, \ldots, dX_{n1}, \ldots, dX_{nd})^T, \tag{2.19}$$

where n is the number of landmarks in the d dimensional shape $\mathbf{X}$.

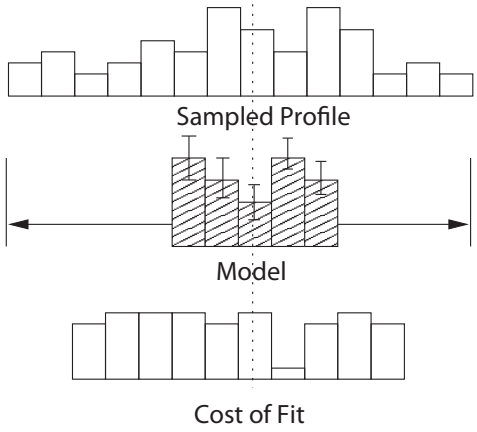


Figure 2.9: ASM search using a local grey level model (Source: [Cootes04])

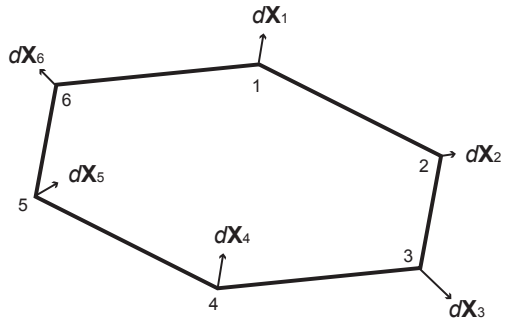


Figure 2.10: Calculated adjustments $d\mathbf{X}$ of a shape $\mathbf{X}$ with six landmarks.

Given $d\mathbf{X}$, pose and shape parameters have to be found which move the landmarks as close as possible to their suggested new positions $(\mathbf{X} + d\mathbf{X})$ while still satisfying the model's shape constraints.

Therefore the rotation, scaling and translation differentials between the current shape $\mathbf{X}$ and the new landmark positions $(\mathbf{X} + d\mathbf{X})$ are found with Procrustes analysis. Having obtained the scaling differential $(1 + ds)$, the rotation differential $d\theta$ and the translation differential $d\mathbf{X}_c$, the residual adjustments, $d\mathbf{x}$, in the local model coordinate frame are calculated [Cootes95]:

$$d\mathbf{x} = M((s(1 + ds))^{-1}, -(\theta + d\theta))[M(s, \theta)[\mathbf{x}] + d\mathbf{X} - d\mathbf{X}_c] - \mathbf{x} \tag{2.20}$$

Due to the fact that these suggested movements are in general not consistent with the shape model [Cootes95], $d\mathbf{x}$ has to be transformed into model parameter space to apply the shape

constraints. The transformation of $d\mathbf{x}$ gives

$$d\mathbf{b} = \mathbf{\Phi}^T d\mathbf{x} \tag{2.21}$$

Updating Shape and Pose Parameters

Having calculated the changes to the pose variables and adjustments to the shape parameter, those are applied to update the model parameters in an iterative scheme [Cootes95]:

$$\mathbf{X}_c \rightarrow \mathbf{X}_c + w_t d\mathbf{X}_c \tag{2.22}$$
$$\theta \rightarrow \theta + w_\theta d\theta \tag{2.23}$$
$$s \rightarrow s(1 + w_s ds) \tag{2.24}$$
$$\mathbf{b} \rightarrow \mathbf{b} + \mathbf{W}_b d\mathbf{b} \tag{2.25}$$

where w_t, w_s and w_θ are scalar weights and $\mathbf{W}_b$ is a diagonal weight matrix (one weight for each mode). $\mathbf{W}_b$ can be either the identity matrix or each weight is proportional to the standard deviation of the corresponding shape parameter over the training set. The latter allows rapid movement in modes with larger shape variations [Cootes95].

Restricting the values of $\mathbf{b}$ as seen in Equation 2.10 ensures that the model only varies in ways seen in the training set.

2.3.4 Extending ASMs to 3D

Similar to [Dickens02] the ASM concept has been extended to 3D:

The PDM (Equation 2.8) $\mathbf{x}$ and the model instance $\mathbf{X}$ (Equation 2.11) are extended so that

$$\mathbf{X} = M(\mathbf{s}, \mathbf{\Theta})[\mathbf{x}] + \mathbf{X}_c \tag{2.26}$$

and

$$\mathbf{x} = (x_0, y_0, z_0, x_1, y_1, z_1, \ldots, x_k, y_k, z_k, \ldots, x_{n-1}, y_{n-1}, z_{n-1})^T, \tag{2.27}$$

where (x_k, y_k, z_k) is the position of landmark k, $\mathbf{X}_c = (X_c, Y_c, Z_c, X_c, Y_c, Z_c, \ldots, X_c, Y_c, Z_c)^T$ is n times the centre of the model (X_c, Y_c, Z_c) in the image frame, $\mathbf{s}$ is a scaling vector (s_x, s_y, s_z) and $\mathbf{\Theta}$ is a rotation vector $(\theta_x, \theta_y, \theta_z)$.

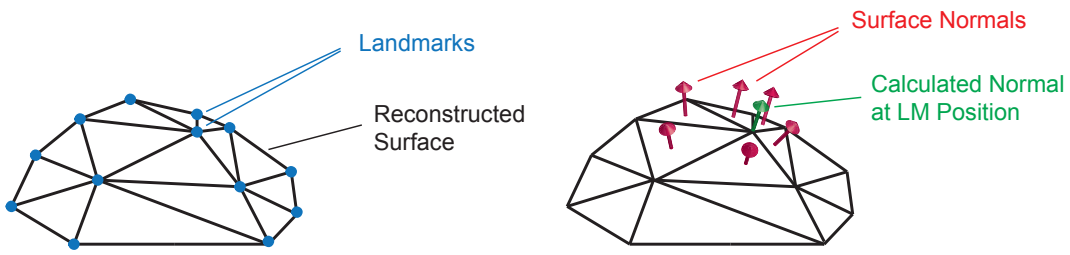
The vector $d\mathbf{X}$ (Equation 2.19) is extended so that

$$d\mathbf{X} = (dX_0, dY_0, dZ_0 \ldots, dX_{n-1}, dY_{n-1}, dZ_{n-1})^T \tag{2.28}$$

yielding in residual adjustments

$$d\mathbf{x} = M((\mathbf{s}(1 + d\mathbf{s}))^{-1}, -(\mathbf{\Theta} + d\mathbf{\Theta}))[M(\mathbf{s}, \mathbf{\Theta})[\mathbf{x}] + d\mathbf{X} - d\mathbf{X}_c] - \mathbf{x}, \tag{2.29}$$

where $(1 + d\mathbf{s})$ is a scaling differential, $d\mathbf{\Theta}$ is a vector of rotation differentials $(d\theta_x, d\theta_y, d\theta_z)$ and $d\mathbf{X}_c$ is the translation differential.



(a) (b)

Figure 2.11: (a) Reconstructed surface from triangulation of landmark point cloud. (b) Triangulated object with normal at the landmark's position

The model parameters are extended so that

$$X_c \rightarrow X_c + w_t dX_c \tag{2.30}$$
$$Y_c \rightarrow Y_c + w_t dY_c \tag{2.31}$$
$$Z_c \rightarrow Z_c + w_t dZ_c \tag{2.32}$$
$$\mathbf{\Theta} \rightarrow \mathbf{\Theta} + w_\Theta d\mathbf{\Theta} \tag{2.33}$$
$$\mathbf{s} \rightarrow \mathbf{s}(1 + w_s d\mathbf{s}) \tag{2.34}$$
$$\mathbf{b} \rightarrow \mathbf{b} + \mathbf{W}_b d\mathbf{b} \tag{2.35}$$

where w_s and w_Θ are scalar weights.

Grey Level Sampling in 3D

3D ASMs also need a new grey level sampling method [Dickens02]. As described in Subsection 2.3.3 the profile is sampled along a normal to the boundary in 2D. The same idea is used in 3D.

The landmarks (representing an object's shape) are used to create a tight, manifold triangulation [Amenta98]. Thus giving a surface reconstruction as seen in Figure 2.11(a). It is used to calculate normals at the landmarks position:

In the first step a normal is computed for each triangular plane. The next step is to find all adjacent triangles for each landmark. Given all those planes with their corresponding normals, the landmark's normal can be calculated by computing the normalised mean of all surrounding normals (Figure 2.11(b)). After calculating all normals the grey level profiles (Figure 2.12) can be sampled along those and the search performed like in the 2D case. It is essential that all normals look either to or away from the model centre to have a consistent direction over all training and search profiles. Otherwise a wrong GLM is computed and the ASM segmentation fails.

(a)

(b)

Figure 2.12: Two example profiles: (a) is a normalised mean profile of the grey levels sampled along the surface normal and (b) are the original grey levels.

2.4 Synopsis

In this chapter the methods used in the segmentation pipeline were presented. For the automatic landmarking process CPD and MDL are used to find point correspondences across the training sets. The anatomical structure is located by means of RFs and discrete optimisation. The landmarks obtained with MDL and CPD are used to build PDMs. Together with a GLM describing the local grey level environment at each landmark they build ASMs which are used for the final bone segmentation.

CHAPTER 3

Experiments

As the segmentation pipeline consists of three major parts, this chapter is organised similarly. At first the test environments and the data used for the experiments are described. For every part of the automatic segmentation process experiments have been executed. Not only the results but also the problems that occurred during the process are pointed out.

3.1 Test Environments

In the experiments made, two environments have been used differing both in hardware and software configuration. Table 3.1 shows a juxtaposition of the differences.

The code was mostly written in Matlab. For better runtime performance computational expensive parts were written in C (mex-files). The datasets were manually annotated with ITK-SNAP [Yushkevich06] and the experimental outcome was visualised with Matlab and matVTK [Birngruber09].

3.2 Data

In this section the data that was used for the experiments is described. Two different kinds of data were used:

1. **Medical data:** 8 CT scans of patients' hands, suffering from RA.
2. **Synthetic data:** User generated sets of 30 bone-like structure volumes

3.2.1 Medical Data

The medical data used in this thesis was captured by a Philips CT machine at the Vienna General Hospital. All data has been rendered anonymously for further use.

The minimum and maximum HU in the medical data used as well as it's size and spacing can be seen in Table 3.2.

Component	Sony Vaio	MacPro
CPU Type	Intel(R) Core(TM) i7	Intel(R) Xeon(R)
CPU Cores (phys/virt)	4/8	4/8
CPU Frequency	1.73 GHz	2.26 GHz
RAM	8 GB	16 GB
OS	Windows 7 64 Bit	OS X 10.6.2
Matlab Version	R2010a	R2009b

Table 3.1: Description of the test environments used containing hardware and software configuration.

No.	Slice Size	Slices	HU Max	HU Min	Pixel Spacing	Slice Spacing
1	768×768	592	2976	-1024	0.2604	0.3301
2	768×768	666	2976	-1024	0.2943	0.3301
3	768×768	612	2976	-1024	0.3242	0.3300
4	768×768	680	2976	-1024	0.2982	0.3301
5	768×768	627	2976	-1024	0.3255	0.3300
6	768×768	701	2976	-1024	0.2982	0.3300
7	768×768	655	2976	-1024	0.2721	0.3301
8	768×768	653	2976	-1024	0.3242	0.3301

Table 3.2: Details for all CT volumes used (numbered from 1 to 8), their slice size (in pixels), amount of slices, spacing (in mm) and HUs.

Image Artefacts

The medical data worked with had image artefacts and noise the algorithms had to challenge with:

- **Noise:** The CT volumes have been captured using a kernel that enhances contrast. Unfortunately also the noise value was risen leading to HUs of the surrounding tissue that is almost or even as high as the typical HU of bone structure (400 or above).

- **Image Artefacts:** In addition to noise, two out of three slices in each volume were affected by line-artefacts of about 20 pixels thickness. These artefacts have a vast effect on the voxels' HUs. The intensity values captured by the machine are added to the wrong CT slice meaning a lack of them in the slice before or after. In the data used the affected areas varied on average about 900 HUs from one slice to the next. Figure 3.1 shows the artefact in two succeeding slices.

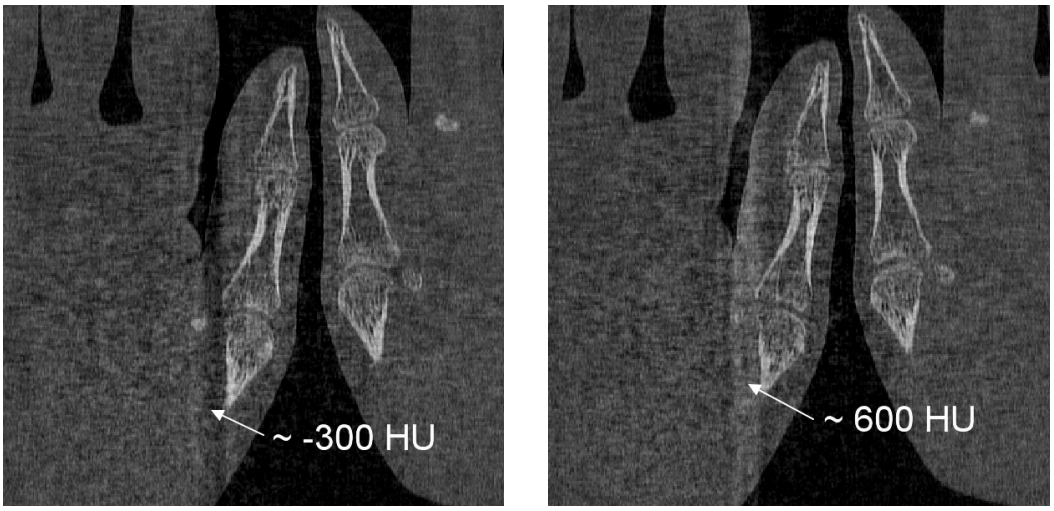


Figure 3.1: Example of the line artefact that appears in every second to third slice of the CT volumes. The HU difference of the affected voxels is about 900 from one slice to the next.

Effects of RA Present in the CT Volumes

As the hand data used was captured from people suffering from RA following effects can be seen on the CT slices:

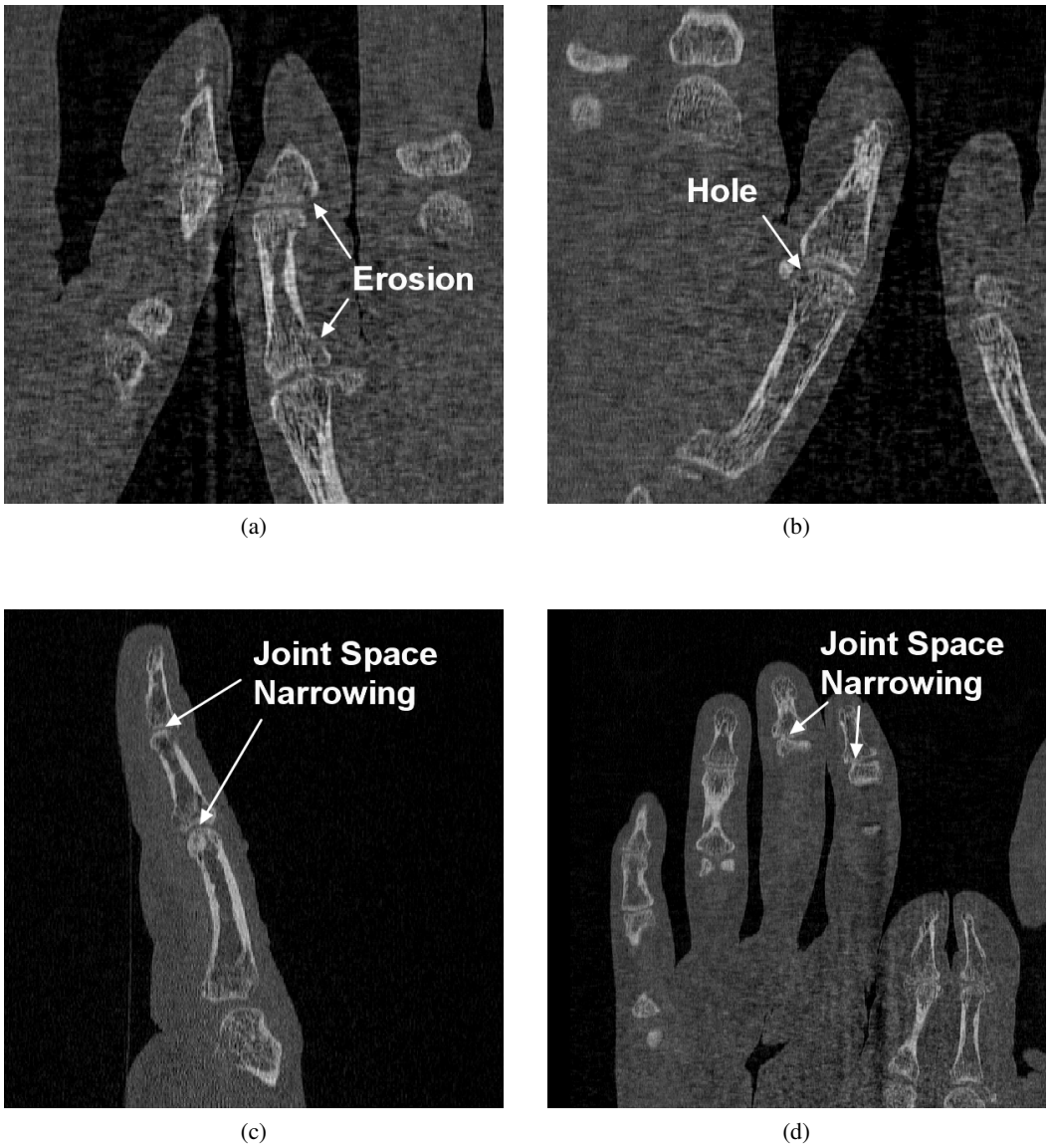
- **Bone Structure:** The structure of the bones was already damaged and even holes within the bone corticalis can be seen and measured analysing the HU intensities at those points. Examples images can be seen in Figure 3.2(a & b).

- **Joint Space:** The joint space between bones narrowed, leading to in some cases no joint space at all, i.e. bones touching each other, and also abnormal position of the fingers which can be seen in Figure 3.2(c & d)

3.2.2 Synthetic Data

For the segmentation process and the automatic modelling experiments, synthetic data was generated. The volumes' basis was a hand drawn bone-like model in a volume of size $100 \times 100 \times 100$ which can be seen in Figure 3.3(a).

For generating the synthetic bone models a slice in the middle of the volume was replicated to stretch the bone. The number of replications was random (using a uniform distribution) between 5 and 40 times giving a volume height of the synthetic data between 105 and 140. To create further differences between the volumes each one of them was dilated with a disk of radius between 0 and 10 (again randomly chosen). As can be seen in Figure 3.3(b & c) the generated bone models differ both in height and general look, especially at the top.

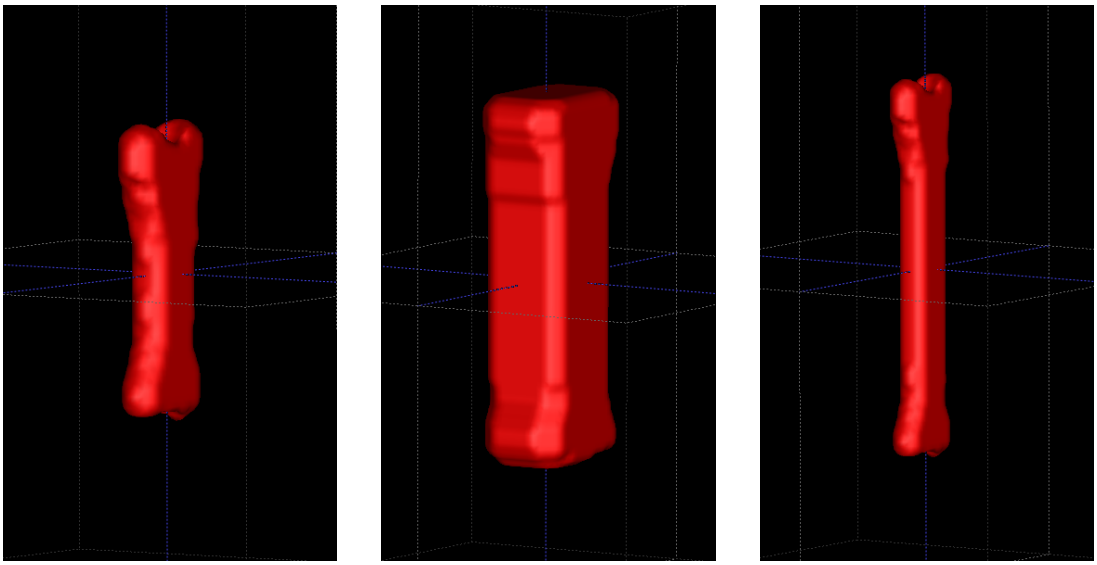


(a) (b)

(c) (d)

Figure 3.2: Effects of RA present in the CT data: In (a & b) erosions and holes in the bone structure can be seen. Joint space narrowing and its effects can be seen in (c & d).

To give the volumes a more natural appearance different grey values have been assigned similar to medical data captured by a CT machine. Two different approaches were used to create the voxels' grey values. The first one was to assign grey values of uniform distributions to the data set. One distribution for each object and one for each background in the set was created giving a dataset with vast variation between corresponding voxels. The second approach was to use normal distributions with less variation. Thus giving two sets of data for further experiments



(a) Model Basis (b) Synthetic Model 3 (c) Synthetic Model 12

Figure 3.3: Basis for all synthetic generated models (a) and two models created from it for the experiments (b & c).

	Backgrounds' Grey Values	Objects' Grey Values	Distribution Type
Set 1	$[-1024, 200]$	$[500, 1500]$	Uniform
Set 2	$\mu = 0, \sigma = 200$	$\mu = 1000, \sigma = 200$	Normal

Table 3.3: Grey values assigned to the synthetic data set

with ASMs. Table 3.3 shows details on the used grey values.

In Figure 3.4 a slice of the first volume of data set 2 is shown (a) and the corresponding model is added as an overlay in (b).

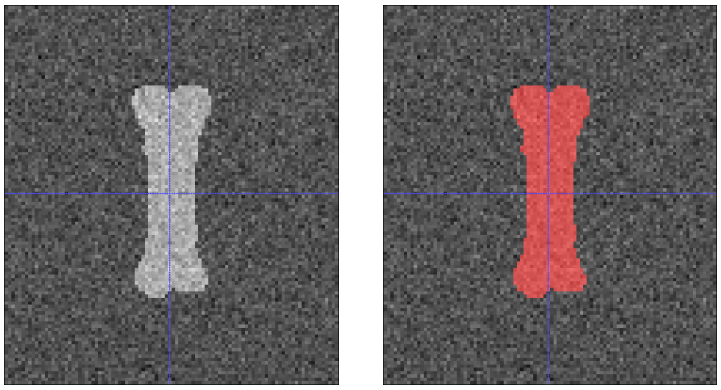
3.3 Finding Point Correspondences

In this section the experiments using MDL and CPD for finding point correspondences across the training sets are described. All synthetic data sets and the CT volumes no. 1-6 have been used for the tests.

3.3.1 Preliminaries

Annotation of the CT Volumes

Two methods were tried to annotate the medical data described in Section 3.2 semi-automatically:



(a) (b)

Figure 3.4: In (a) a slice of one final volume of set 2 with normal distributed grey values and in (b) the same slice with the model's overlay is shown.

- **Graph Cuts** are a technique to solve computer vision problems that can be formulated in terms of energy minimisation. This can be reduced to a maximum flow problem in a graph and is equivalent to a minimal cut of the graph. Graph Cuts were introduced by Greig et al. in 1989 [Greig89].

- **ITK-SNAP** is a software application using smart snakes for semi-automatic medical image segmentation. [Yushkevich06]

Due to the high rate of noise and the holes inside the bone structure neither of them was sufficient in segmenting the bones. Taking the segmentation of ITK-SNAP as a starting point the further annotation process was done manually for each slice in the volumes. Following bones have been annotated:

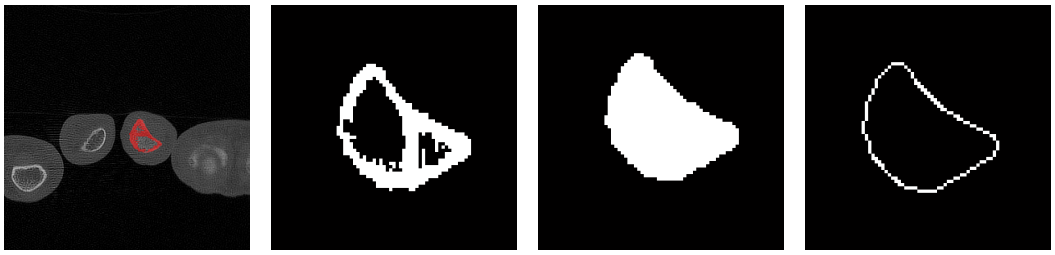
- Os metacarpale I

- Phalanx proximalis pollicis

- Phalanx distalis pollicis

Creating Bone Contours from Annotations

To get the coordinates of the voxels defining the bones hull, the annotations had to be refined by morphological operations. To close small holes inside the annotations and fill it a morphological closing was performed with a disc structure element of size 10 followed by a filling operation. The last step was to extract the edge thus giving the final bone contour (Figure 3.5). For further use the coordinates of every point on the contour were extracted.



(a) CT slice with annotation (b) ITK-SNAP annotation refined manually (c) Annotation after closing and filling (d) Final edge contour

Figure 3.5: This Figure shows the extraction of the bone contour from the annotation through morphological operations.

Option	Value(s)
Nneighbours	5, 15, 25
MutationRange	2, 5, 10
CoefficientRestriction	0
MaxIts	500
EntropyPreservation	5
TextureWeight	0

Table 3.4: Options set for the MDL algorithm

3.3.2 Experiments and Results

MDL

For the initialisation of the MDL algorithm used ([Langs07]) the hull coordinate points of the objects in each model set (three real bones and a synthetic one) had to be aligned close to each other. In the first attempt they were shifted to the origin having their mean coordinates at (0, 0, 0). Different settings used for the MDL algorithm are shown in Table 3.4. Both global as local shape weights were tried. The genome size for the experiments was chosen to be 1000 which results in the same amount of landmarks computed for each model.

Unfortunately all runs with different settings had the same result. The coordinate points across the model sets were not registered correctly as shown in Figure 3.6.

The second attempt for initialising the MDL algorithm was to align the point sets using CPD. The outcomes were similar to the results after only shifting them to the coordinate system's origin. The error rate after the MDL was even higher than after the initialisation with CPD. Therefore, the use of MDL was abandoned from the segmentation pipeline using CPD for registration instead.

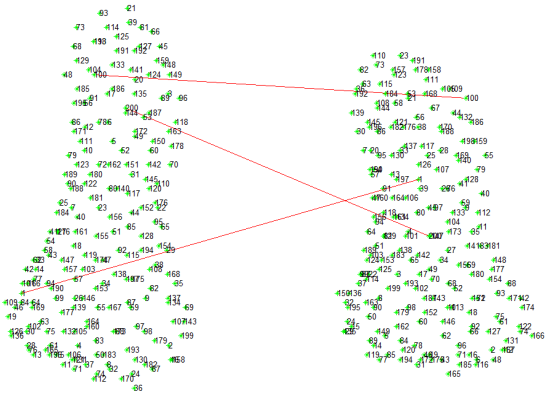


Figure 3.6: False registration of two phalanx distalis pollicis as a result of the MDL algorithm. The red lines connect landmarks that should be at the same position on each bone.

Option	Value	Effect/Meaning
method	'nonrigid'	Registration method
corresp	1	Estimates the correspondence vector at the end of the registration
normalize	1	Normalises both point sets to zero mean and unit variance before registration and denormalises them afterwards
max_it	150	Maximum number of iterations
tol	1e-5	Tolerance stopping criterion
outliers	0.1	Weight of noise and outliers
beta	2	Gaussian smoothing filter size
lambda	3	Regularisation weight

Table 3.5: Options set for the CPD algorithm

CPD

The CPD implementation from [Myronenko10] was used to find point correspondences across the data sets to circumvent the need for MDL during the landmarking process. As the landmarks have to be at the same position on every training set and the constraints of a rigid transformation were too strict to register the data sets precisely, non-rigid point registration was used.The other options set are shown in Table 3.5. For every type of object the one with smallest amount of hull coordinate points in the set was aligned on the others, yielding point correspondences across the data sets.

As Figure 3.7 shows the registration using CPD found the correct correspondences between the landmarks in the sets. Note that only the first 200 landmarks are displayed. In Figure 3.8 three examples of the registration process are shown with the initial starting position and the final result.

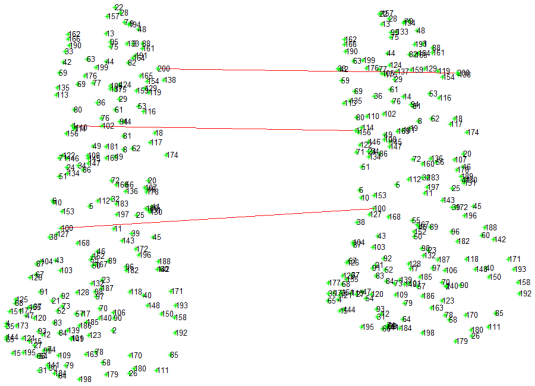


Figure 3.7: Correct registration of two phalanx distalis pollicis using the CPD algorithm. The red lines show example correspondences of landmarks between the two point sets.

Having registered all data sets correctly, the next step was to create PDMs. For this purpose the first p modes of the PCA covering 95% of the eigenvalues were used. Further details on the PDMs are described in Section 3.5 due to their relation to the experiments with ASMs.

3.4 Anatomical Structure Localisation

Experiments for anatomical structure localisation were only run on the medical data due to the fact that the synthetic data did not contain repetitive structures. The task to achieve was to locate joints of a hand. Having located these joints they can then be used to get a initial position for the ASM search, i.e. initial size, orientation and centre of the model to be placed on the image can be computed. For the localisation experiments 28 landmarks were manually set on all CT volumes with ITK-SNAP according to Figure 3.9 on the right hand on each volume.

[Donner10a] downsized the CT volumes to a resolution of $256 \times 384 \times 330$ voxels.

The experiments were made with a LOOCV framework using the manual annotations of 7 training volumes to construct the geometric model which can be seen in Figure 3.10(a).

[Donner10a] extracted 33250 descriptors from the background and around the landmarks of the training images. The RF classifier was trained on them using 200 trees. This led to the detection of around 50 candidate interest points for each landmark in the test images. [Donner10a] used a simple random walk approach for approximately inferring the MRF. The residual voxel distances between the selected interest points and the corresponding ground truth landmarks was recorded after matching, giving the following results (the corresponding size in mm was computed using the spacing of the CT data described in table 3.2) [Donner10a]:

- **Mean residual:** 10.13 voxels, respectively 12.87 mm
- **Median residual:** 5.59 voxels, respectively 7.01 mm
- **Standard deviation residual:** 16.99 voxels, respectively 21.57 mm

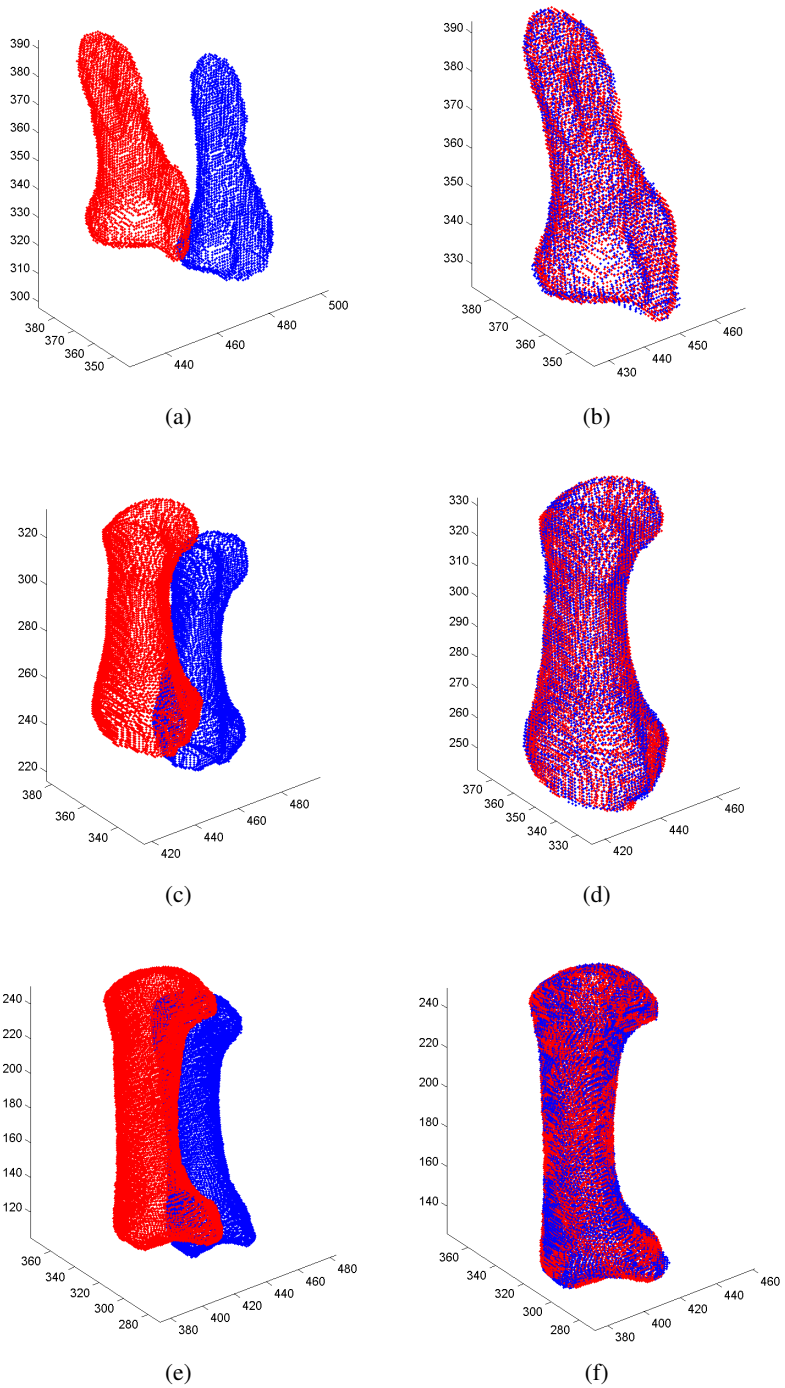


Figure 3.8: (a) Two phalanx distalis pollicis point sets before registration and (b) after. (c & d) Registration of two phalanx proximalis pollicis. (e & f) The same for two os metacarpale I.

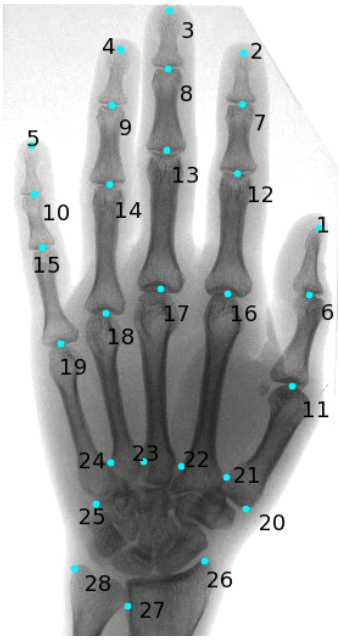


Figure 3.9: 28 landmarks set on each training volume for the anatomical structure localisation experiments.

One hand with spheres of radii corresponding to the median landmark residuals from all LOOCV runs is shown in Figure 3.10(b) to display the quality of the results.

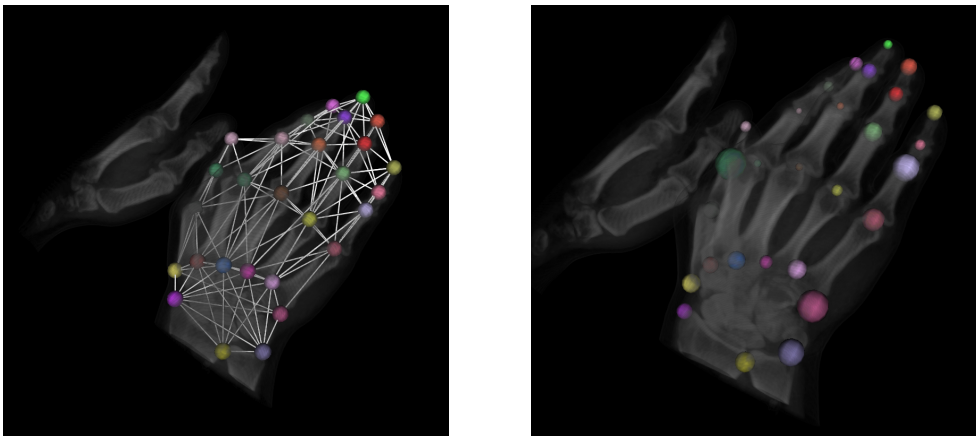
The residual distances for all leave-one-out runs can be seen in Figure 3.11. The average finger width being around 32 voxels, the median shows that the vast majority of landmarks are localised at the proper position. Mean and standard deviation values are deteriorated by 15 outliers. Five of them can be explained by following facts [Donner10a]:

- One CT volume is cut off too close to the carpus, thus leading the solver to choose another landmark not related to the anatomical structure.
- In two other volumes fingertips crossed over.

Due to the small set of training volumes it is believed that the geometrical model is too restrictive, yielding the MRF solution not representing the global optimum in certain cases. The proposed approach (implemented in Matlab except the C-based RF) had a runtime of about 5 minutes per single localisation (on the MacPro) [Donner10a].

3.5 Segmentation using ASMs

In this section the experiments with the ASM implementation are described. Both synthetic and medical data were used. As for the structure localisation, a LOOCV framework was chosen for the ASM experiments to examine how the shape model fits to unseen objects.



(a) (b)

Figure 3.10: (a) Training data including the manually annotated landmarks and the connectivities building the geometric model. (b) The result of the anatomical structure localisation method with median residual distance between ground truth and localisation result from all leave-on-out runs used as radius for each sphere. (Source: [Donner10a])

3.5.1 Preliminaries

This subsection gives an overview of the preliminaries needed for the ASM experiments.

PDMs

The results of Section 3.3, i.e. point correspondences for each data set, were used to create PDMs. From the bone contours 1000 landmarks were chosen randomly for each medical data set and 200 landmarks for the synthetic data. The PDMs created with PCA cover 95% of the variation seen in the data sets. In Figure 3.12 an example of a PDM for phalanx distalis pollicis is shown. (a) is the PDM with mean value for the t first modes. (b) shows the mean model plus and (c) minus three times the standard deviation of the first mode. In Appendix A.1 examples for all models with their first t modes are given.

Noise Filtering

For the ASM search on the medical data to work properly, a Gaussian filter was used to reduce the noise in the images. Three different types of Median filter implementations were tried first but all of them resulted in Matlab to hang-up during execution due to memory shortage and deadlocks. Figure 3.13 shows CT volume no. 1 before (a) and after noise filtering (b).

3.5.2 Experiments and Results

On every volume the mean PDM was posed without altering any modes. To be independent of the structure localisation, the initial search position was set randomly around the mean landmark

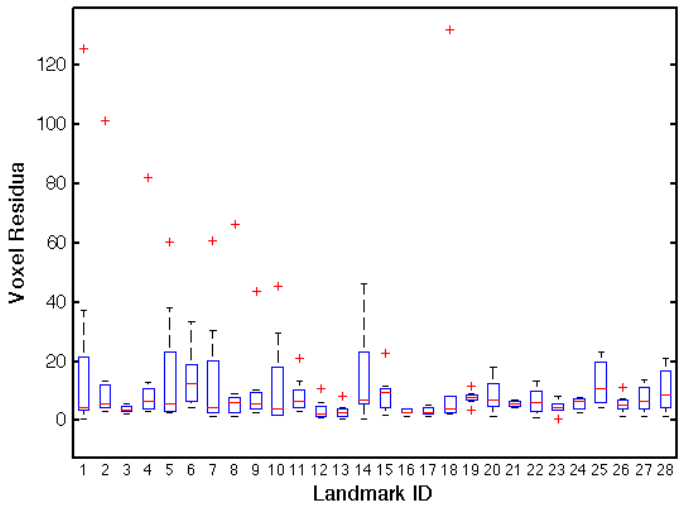
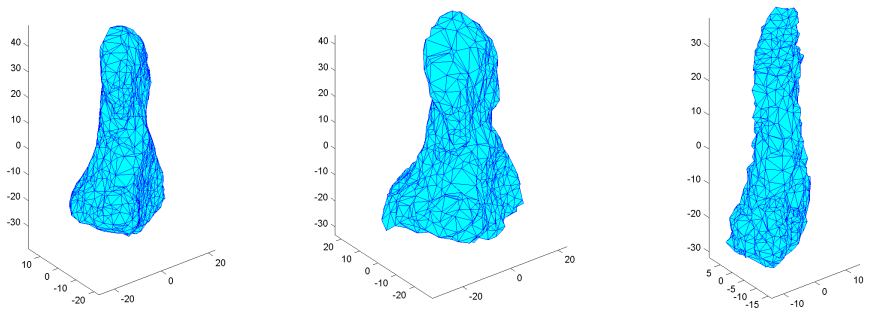


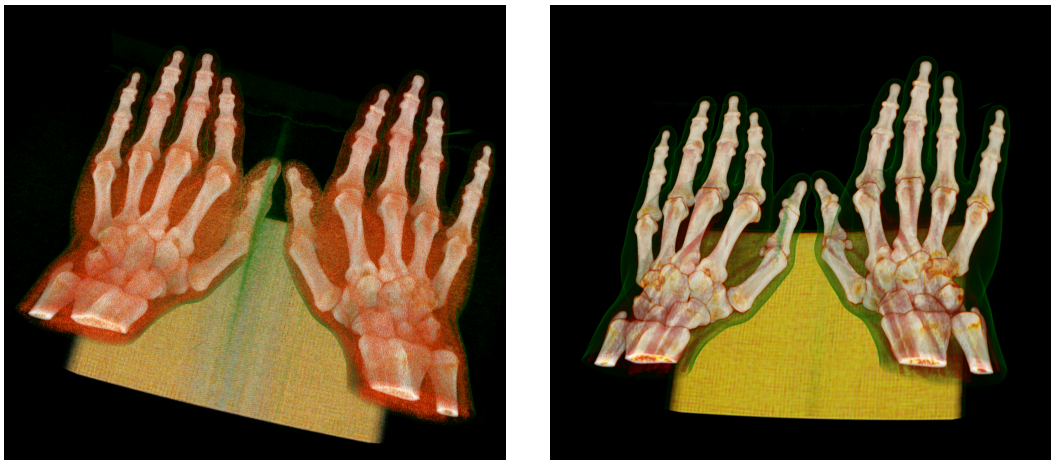
Figure 3.11: Boxplot of the residual distances from all leave-one-out runs for each landmark. (Source: [Donner10a])



(a) (b) (c)

Figure 3.12: (a) Shows the phalanx distalis pollicis model of the first LOOCV run. (b) shows the mean model plus and (c) minus three times the standard deviation of the first mode.

position of the ground truth. Between -8 and +8 voxels were added on every axis to the mean position leading to an additional Euclidean distance of up to 13.86 voxels. Thus also giving the possibility to analyse the ASM's search performance in dependence of the initialisation position. As stopping criterion for the search loop, the difference between the eigenvalues computed in the current iteration and those of the last iteration, was chosen. The eigenvalues were restricted to not exceed two times the standard deviation seen in the PDM's creation thus explaining around 95.45% of the data's variation. For evaluation purposes the Landmark to Landmark Error (LLE) was computed, i.e. the Euclidean distance between the manual annotated ground truth landmark



(a) (b)

Figure 3.13: (a) CT volume no. 1 before and (b) after filtering with a Gaussian kernel to reduce image noise.

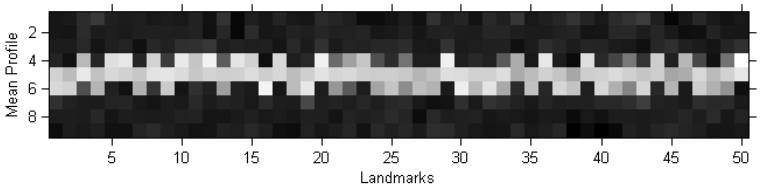


Figure 3.14: The mean grey level profile of the first 50 landmarks of one LOOCV run on synthetic data set 1.

and the model landmark position after search.

Synthetic Data - Set 1

The normalised first derivative of the grey levels along the surface normals were captured during the training process. The length of these grey level profiles was set to 9 voxels. A grey level model explaining 68.72% of all grey level profiles was used to reduce the data's dimensionality and the mean profiles were computed over all corresponding landmarks on the training images. An example of the mean grey level profiles for synthetic data set 1 can be seen in Figure 3.14.

For the ASM search a profile length of 25 voxels was chosen. The maximum number of search iterations was set to 200 and the stopping criterion set to a maximum difference of 0.001. As can be seen in Figure 3.15 the outcome of the experiment is twofold. Either the LLE is close to zero, meaning an almost perfect segmentation, or the model drifted away significantly. The average mean LLE over all LOOCV runs is 6.99 voxels which means, assuming a spacing of

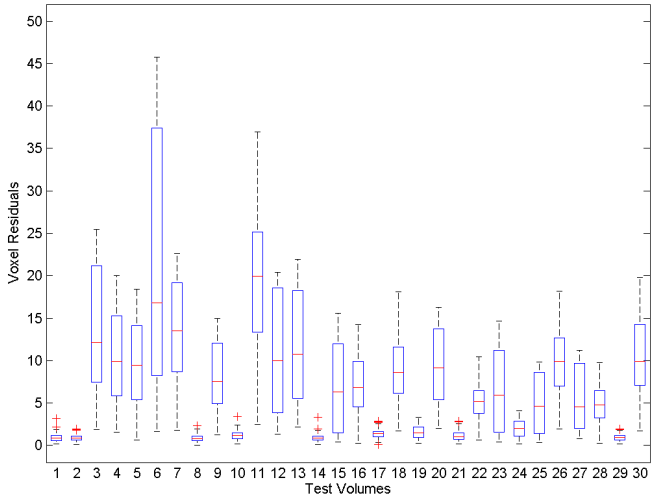


Figure 3.15: Boxplot showing the ASM search results on the synthetic data (set 1). Notice the twofold outcome.

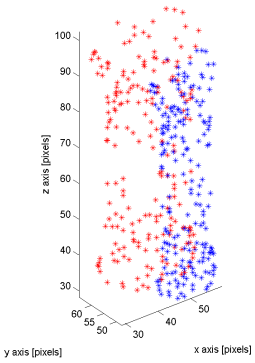
0.25mm, a distance of 1.75mm. Accordingly the average median LLE is 6.58 voxels respectively 1.64mm.

To demonstrate the search process three examples are given:

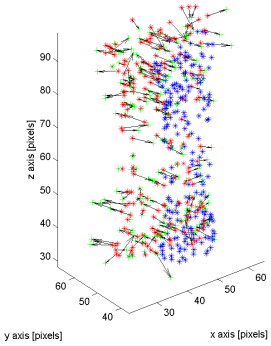
- In Figure 3.16 the ASM search performed on test volume 8 is shown. The starting mean LLE was 12.24 voxels. Although the model's inital position is that far away, the ASM search algorithm manages to fit the model to the unseen object in the volume, leading to an average LLE of 0.38 voxels, respectively 0.09mm when assuming a spacing of 0.25mm.
- Figure 3.17 shows test volume 12. The mean LLE at initialisation was 12.53 voxels. The model manages to fit to the top of the unseen object but gets stuck in the local shape constraints yielding the search to stop after 168 iterations. The final mean LLE was 10.71 voxels or 2.68mm.
- The third example shown in Figure 3.18 shows one of the main problems the ASMs have to scope with in 3D space: rotation. The further the model is rotated away from the image object, the less possible it is that it rotates back to its initial position because of the GLM not finding a match along the surface normals. The example shows test volume 6 which had the biggest LLEs (as seen in Figure 3.15).

Synthetic Data - Set 2

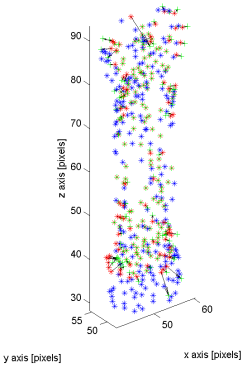
For the next experiment on synthetic data set 2, the number of iterations was set to 500 to see if the segmentation result changes significantly during later iterations. In contrast to Section 3.5.2



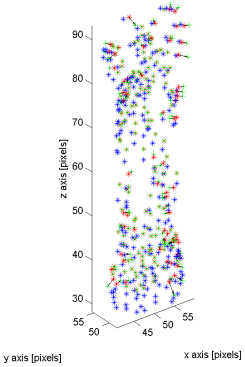
(a) Search Initialisation



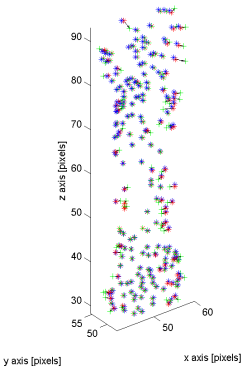
(b) Iteration 20



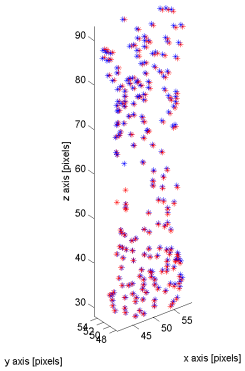
(c) Iteration 40



(d) Iteration 60

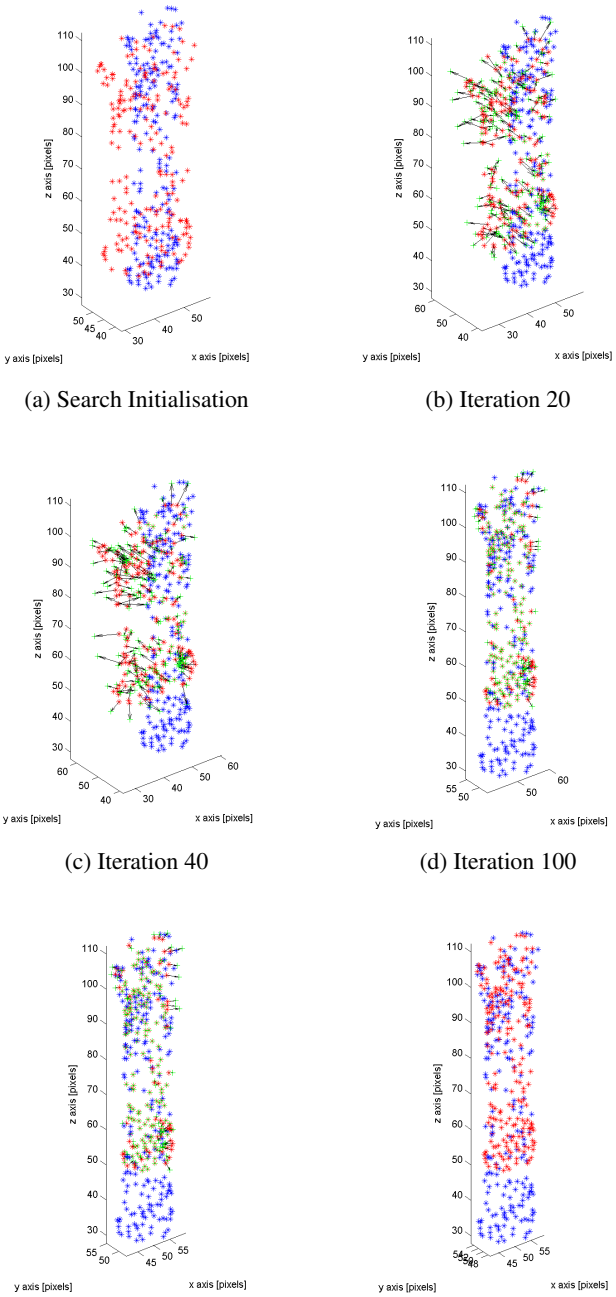


(e) Iteration 140



(f) Final Position

Figure 3.16: ASM search example: LOOCV run on volume 8 of the synthetic data set 1. The red points represent the model's landmarks and the blue points the ground truth. The green points and black arrows show the suggested movement of the model's landmarks computed from the best fit of the GLM to the search profile.

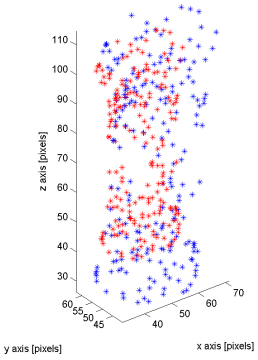


(a) Search Initialisation (b) Iteration 20

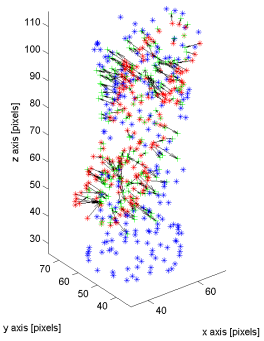
(c) Iteration 40 (d) Iteration 100

(e) Iteration 140 (f) Final Position

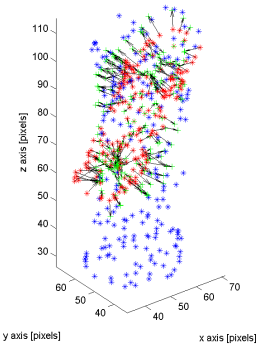
Figure 3.17: ASM search example: LOOCV run on volume 12 of the synthetic data set 1. The red points represent the model's landmarks and the blue points the ground truth. The green points and black arrows show the suggested movement of the model's landmarks computed from the best fit of the GLM to the search profile.



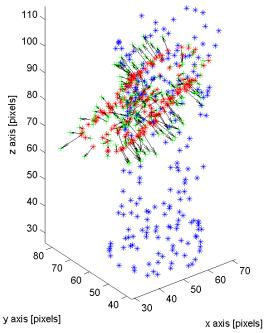
(a) Search Initialisation



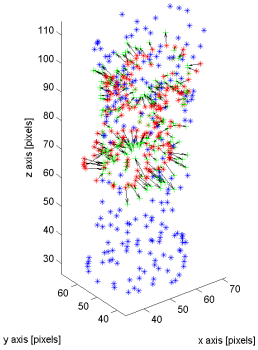
(b) Iteration 20



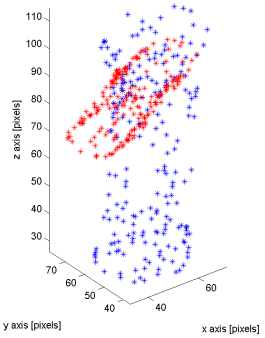
(c) Iteration 40



(d) Iteration 60



(e) Iteration 140



(f) Final Position

Figure 3.18: ASM search example: LOOCV run on volume 6 of the synthetic data set 1. The red points represent the model's landmarks and the blue points the ground truth. The green points and black arrows show the suggested movement of the model's landmarks computed from the best fit of the GLM to the search profile.

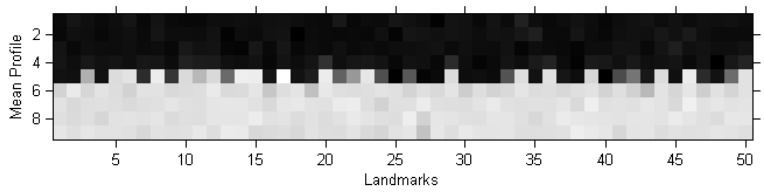
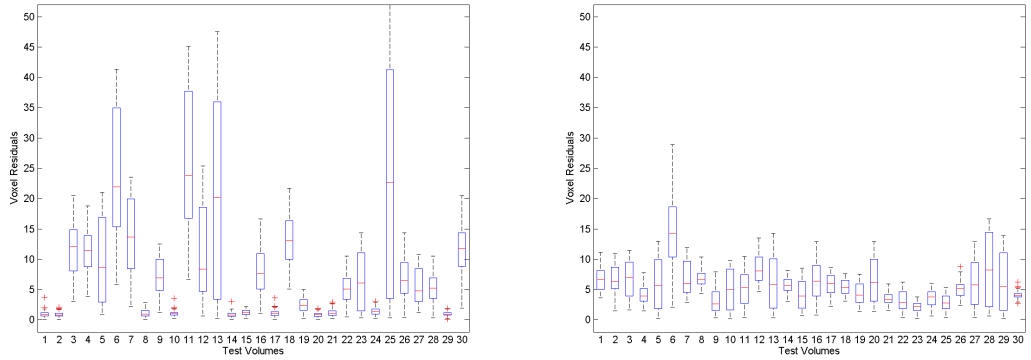


Figure 3.19: The mean grey level profile of the first 50 landmarks of one LOOCV run on synthetic data set 2.



(a) (b)

Figure 3.20: Boxplot showing the ASM search results on the synthetic data (set 2) with the stopping criterion changed from (a) 0.001 to (b) 0.1.

the raw grey levels (see Figure 3.19 for an example) were used to build the GLM $\mathbf{g}_{new}$. This way the difference between the ASM segmentation results using raw grey level GLMs and normalised derivative GLMs is analysed. As stated in [Cootes94a] the segmentation results should be similar. All other settings remained the same.

As can be seen from the boxplot in Figure 3.20(a) the results do not differ widely from the ones seen in Sec. 3.5.2. The average mean LLE over all LOOCV runs is 7.84 voxels or 1.96mm and the mean median LLE is 7.44 voxels or 1.86 mm respectively.

When changing the stopping criterion of the search loop from 0.001 to 0.1 the search is finished in all cases before reaching the 32$^{\text{nd}}$ iteration giving the results seen in Figure 3.20(b). Thus leading to on the one hand segmentations that only fit roughly but on the other hand avoiding the model to rotate away from the object's position. This is due to the fact that the biggest movement can be seen in the first iterations with the refinement process starting afterwards. When points on one side of the object are segmented correctly but others (neighbouring points) further away tend to move in one direction, the model starts to rotate and drift as seen in Figure 3.21. The average mean LLE is 5.67 voxels/1.42mm. The mean median LLE is 5.49 voxels

No.	Mean LLE min	Mean LLE max	Mean LLE average	Median LLE min	Median LLE max	Median LLE average
1	6.18	6.58	6.39	5.64	5.81	5.74
2	4.67	23.70	6.60	4.14	25.67	6.33
3	6.17	6.42	6.25	5.98	6.27	6.07
4	4.01	9.20	4.55	3.66	8.93	4.21
5	3.68	5.07	4.73	3.53	4.94	4.63
6	11.75	12.04	11.89	11.85	12.17	12.00

Table 3.6: Phalanx distalis pollicis segmentation results: The minimum, maximum and average mean and median LLE in voxels of 10 search runs on CT volumes no. 1-6 is given.

/ 1.37mm. Using a search length of 25 voxels and 200 landmarks one search iteration took about 0.645s on the Sony Vaio. This is equivalent to 129s for 200 iterations.

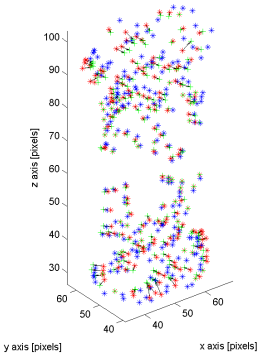
Medical Data

For the ASM search experiments with the CT volumes following settings were used:

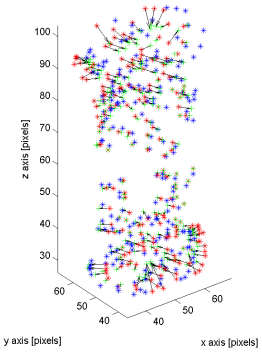
- The grey level profile search length was set to 25 voxels. This length allows to capture the bone surface even if the model is initialised roughly and is small enough to avoid overlap with neighbouring bones.
- For the GLM creation a profile length of 9 voxels was used. The profile captures the bone surface and its surrounding area without neighbouring bones.
- As profile type original grey values were chosen to be able to ensure (with visual inspection) that the cut out profiles are correct.
- Two different stopping criterions (0.1/0.01) were tested to examine the segmentation difference.

For each of the three annotated bones per volume 10 ASM search runs were performed, every run having a different random model centre initialisation around the mean landmark position of the ground truth. For the first experiments a maximum of 150 iterations and 0.1 as stopping criterion was used. Following results were achieved:

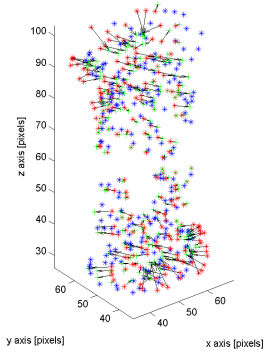
Phalanx distalis pollicis As can be seen in Table 3.6 the minimum mean LLE was varying from 3.68 to 11.75 voxels on the CT volumes 1 to 6. ASM search lacked overall performance on the sixth volume. An explanation for this is that the model did not cover the variation needed to segment the bone in the volume. This leads to almost twice the error rate of the second biggest minimum mean LLE found in volume 1.



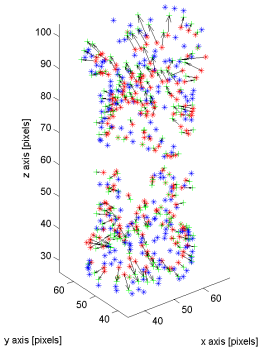
(a) Iteration 15



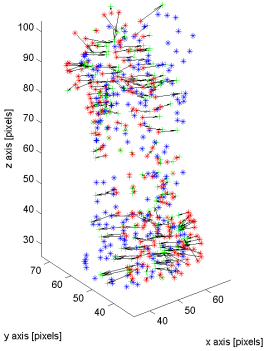
(b) Iteration 19



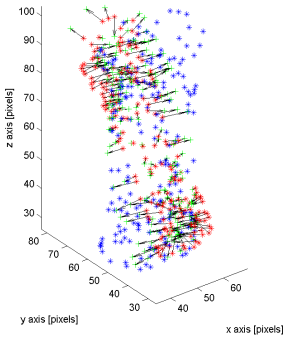
(c) Iteration 21



(d) Iteration 24



(e) Iteration 27



(f) Iteration 29

Figure 3.21: Example of an ASM segmentation failure: The ASM rotates away from the ground truth after having a rough fit in iteration 15. The red points represent the model's landmarks and the blue points the ground truth. The green points and black arrows show the suggested movement of the model's landmarks computed from the best fit of the GLM to the search profile.

	Mean LLE			Median LLE		
No.	min	max	average	min	max	average
1	2.84	3.89	3.19	2.65	3.71	3.04
2	3.58	3.99	3.69	3.36	3.93	3.51
3	2.33	2.69	2.54	2.10	2.35	2.25
4	3.78	3.92	3.85	3.25	3.39	3.34
5	2.24	2.39	2.33	2.04	2.26	2.18
6	4.75	4.89	4.83	4.04	4.16	4.12

Table 3.7: Phalanx proximalis pollicis segmentation results: The minimum, maximum and average mean and median LLE in voxels of 10 search runs on CT volumes no. 1-6 is given.

	Mean LLE			Median LLE		
No.	min	max	average	min	max	average
1	22.12	33.35	25.50	20.43	31.36	23.57
2	15.44	26.14	22.87	14.51	26.17	22.45
3	12.39	22.59	16.86	11.66	20.55	15.74
4	13.49	24.08	16.70	13.36	23.58	16.45
5	5.50	5.63	5.57	5.20	5.32	5.26
6	22.20	24.66	23.65	22.57	24.53	23.72

Table 3.8: Os metacarpale I segmentation results: The minimum, maximum and average mean and median LLE in voxels of 10 search runs on CT volumes no. 1-6 is given.

Phalanx proximalis pollicis The second bone segmented with ASMs yields the best results as can be seen in Table 3.7. The minimum mean LLE varies from 2.24 to only 4.75 voxels (again on volume 6). An average median LLE of 2.18 to 4.12 voxels is achieved over all segmentation runs. These results can be interpreted as the model satisfying the variance seen in the volumes even though only 5 were used as training data. The average mean LLE varied from 2.33 to 4.83 voxels which is equivalent to 0.76mm and 1.49mm with respect to the voxel spacing on the different volumes. 3D ASMs used on cardiac MRI data in [vanAssen06] resulted in mean landmark to surface distances between 1.27mm and 2.57mm. As the landmark to landmark distance is greater or equal to the landmark to surface distance, the experimental outcome in this experiment is almost two times better.

Os metacarpale I The search for os metacarpale I yielded unexpected results (Table 3.8). The bone was segmented correctly only on volume 5.

After investigating the problem two reasons for the experimental outcome are believed to be found:

1. As can be seen in Figure 3.22 one problem is believed to be the strong trabecular structure

Figure 3.22: Os metacarpale I: The trabecular structure inside the bone is more present than the bone contour resulting in wrong landmark positions during ASM search.

inside os metacarpale I. The trabeculas' position vary from volume to volume and therefore are not represented in the GLM. Due to this the GLM fits to the trabecular structure inside the bone rather than to the contour when searching for new landmark positions.

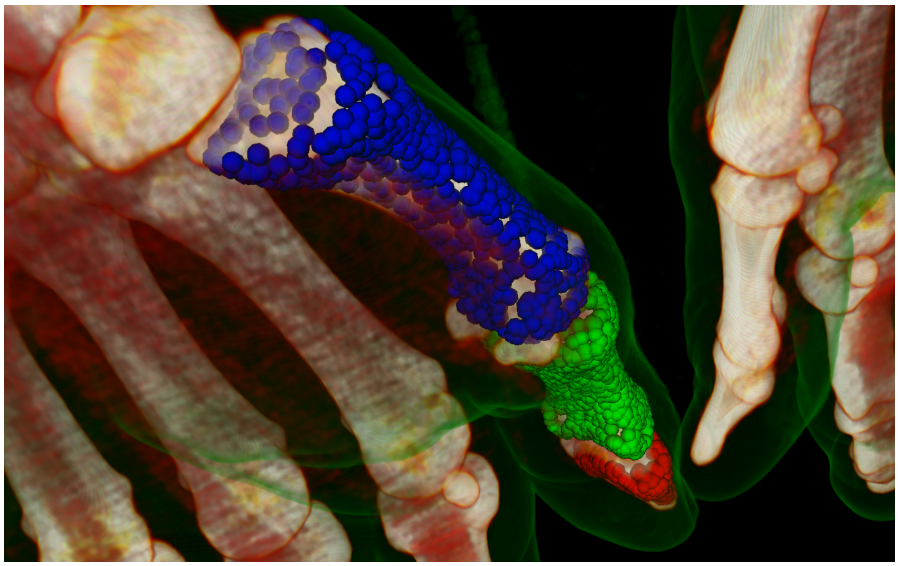
2. Depending on the ASM's initial position the GLM tends to fit to the ossa sesamoidea (see Figure 1.2) which are located close to os metacarpale I, rather than to the desired bone contour. This only happens when the ossa sesamoidea are not present in the GLM.

These reasons lead to new landmark positions that are not on the bones surface and the model to drift resulting in segmentation failure.

A complete segmentation of the thumb bones on CT volume 5 is shown in Figure 3.23. The landmarks of phalanx distalis pollicis are coloured red, landmarks for phalanx proximalis pollicis green and landmarks of os metacarpale blue. Mind the ossa sesamoidea close to the segmented bones.

Boxplots of the results of all test runs can be seen in Figure 3.24 (phalanx distalis pollicis), Figure 3.25 (phalanx proximalis pollicis) and Figure 3.26 (os metacarpale I). For the experimental results using a maximum of 200 iterations and 0.01 as stopping criterion see Appendix A.2.

(a)

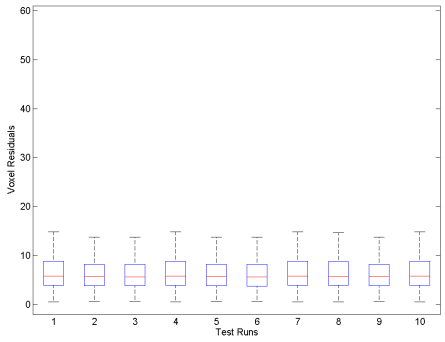


(b)

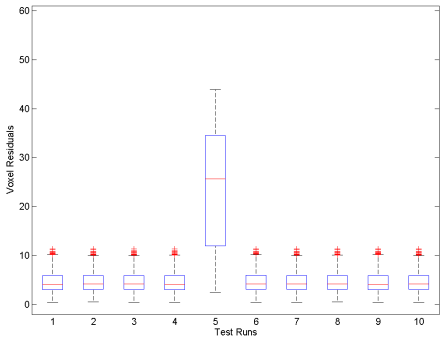
Figure 3.23: Final ASM segmentation of phalanx distalis pollicis (red), phalanx proximalis pollicis (green) and os metacarpale (blue) on CT volume no. 5.

3.6 Synopsis

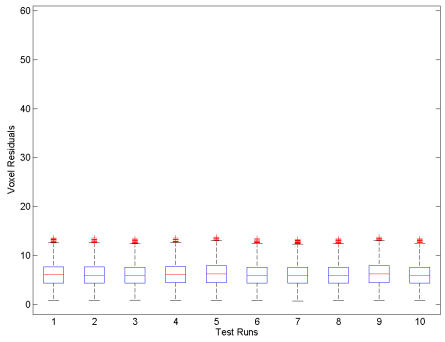
As shown in the experiments the MDL implementation did not find the correct point correspondences. As the bug in the MDL code could not be found and the CPD implementation was computational less expensive it was used as a replacement to find the point correspondences



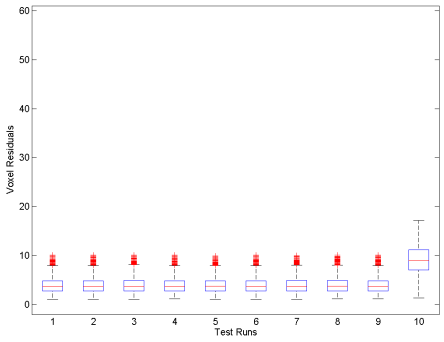
(a) CT Volume 1



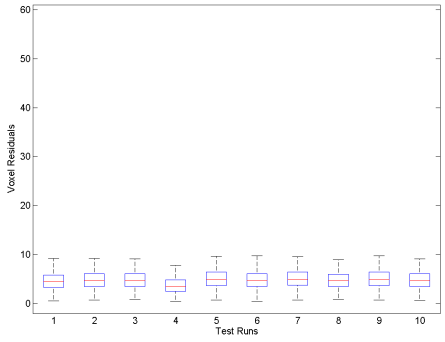
(b) CT Volume 2



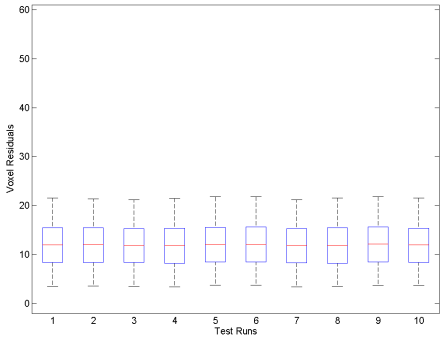
(c) CT Volume 3



(d) CT Volume 4

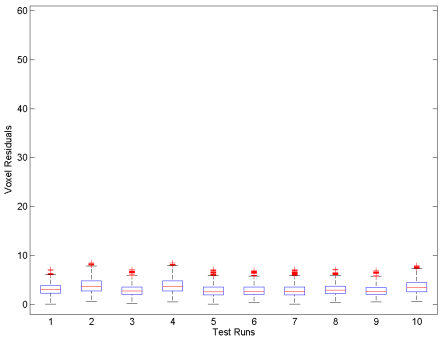


(e) CT Volume 5

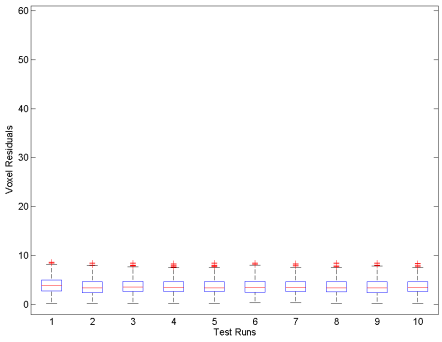


(f) CT Volume 6

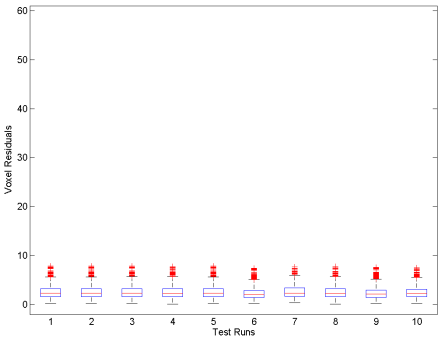
Figure 3.24: ASM search results of phalanx distalis pollicis on CT volumes no. 1-6



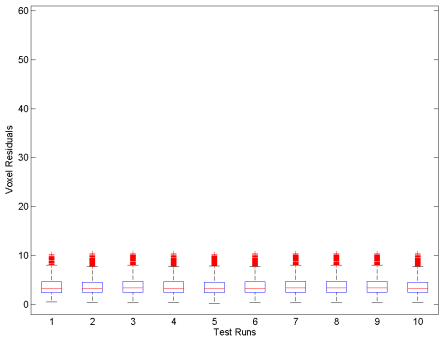
(a) CT Volume 1



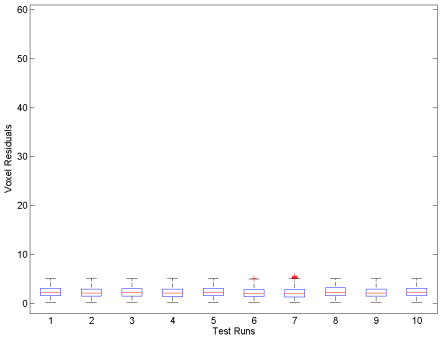
(b) CT Volume 2



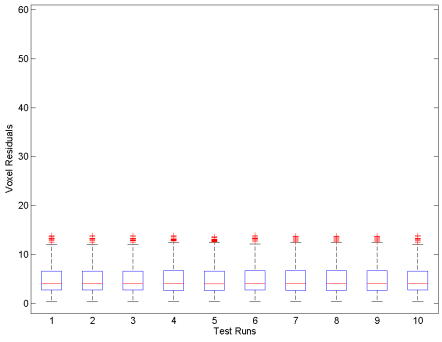
(c) CT Volume 3



(d) CT Volume 4

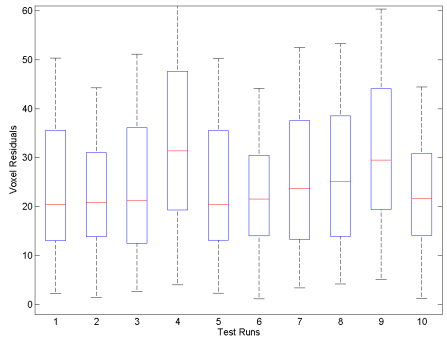


(e) CT Volume 5

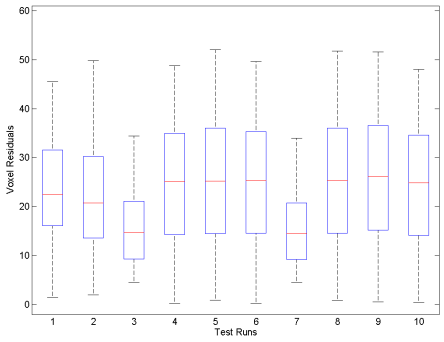


(f) CT Volume 6

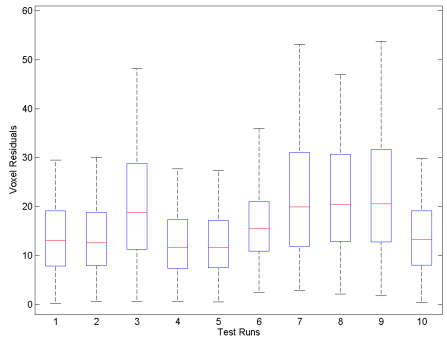
Figure 3.25: ASM search results of phalanx proximalis pollicis on CT volumes no. 1-6



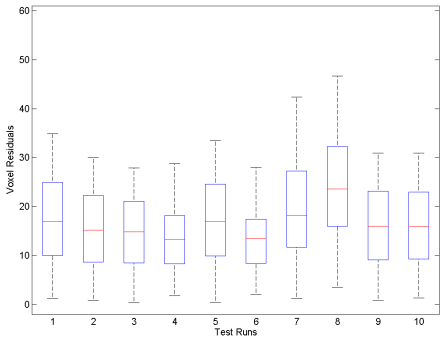
(a) CT Volume 1



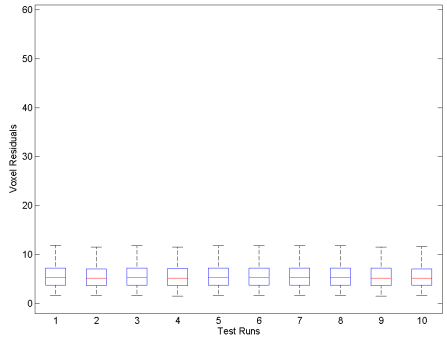
(b) CT Volume 2



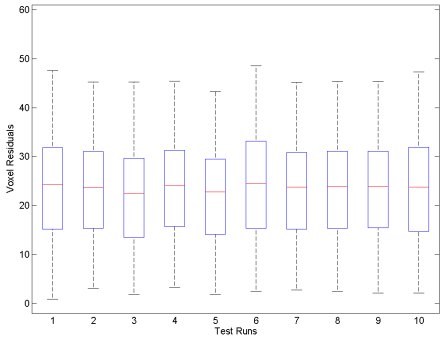
(c) CT Volume 3



(d) CT Volume 4



(e) CT Volume 5



(f) CT Volume 6

Figure 3.26: ASM search results of os metacarpale I on CT volumes no. 1-6

across the training sets.

The anatomical structure localisation of [Donner10a] shows promising results for the automatic localisation of the joints. The joint localisation is adequate to be used as an initialisation for the ASM segmentation.

For the ASM experiments three thumb bones and one synthetic generated bone-like structure were used. The settings used for the ASM segmentation were varied to examine the experimental outcome. The best results were achieved on phalanx proximalis pollicis with a mean LLE of 2.24 voxels. As there are no similar works on hand bones no direct comparison to the experimental outcome can be drawn. Nonetheless if compared to another 3D ASM implementation that uses cardiac MRI data [vanAssen06] the results are almost two times better.

CHAPTER 4

Conclusion and Outlook

In this chapter a conclusion about the presented results achieved in the scope of this thesis is drawn (Section 4.1). In Section 4.2 an outlook containing possible future improvements is given.

4.1 Conclusion

The following conclusion can be drawn by summing up the results presented in the scope of this thesis. The methods proposed for the segmentation pipeline are applicable. The pipeline was tested on 3D CT hand data and is working. The 3D ASM implementation was tested both on synthetic and medical data. Even on rheumatoid arthritis affected bones promising segmentation results were achieved.

The main contribution of this thesis was the analysis of different methods for the proposed fully automatic segmentation pipeline. As a part of an ASM framework that is capable of 2D and 3D data 3D ASMs have been implemented. All methods used in the pipeline have been analysed individually in a LOOCV framework to deliver unbiased results.

The final pipeline contains CPD instead of MDL for the automatic landmarking process. The MDL implementation used did not work properly and found wrong point correspondences across the data sets. The bug in the MDL code could not be tracked down and MDL was therefore abandoned from the pipeline. With the CPD algorithm a computational cheap and accurate way for replacing the MDL implementation was found.

The anatomical structure localisation [Donner10a] delivers the position of the joints which are used for the initialisation of the ASM segmentation. The results are adequate for the ASM initialisation. With the automatic localisation of anatomical structures a manual initialisation of the segmentation process like in ITK-SNAP is not necessary any more.

Four different types of objects were used for the ASM segmentation experiments (a synthetic bone-like object and all three thumb bones). The synthetic data sets were used to analyse the ASM behaviour in 3D space and point out problems that can occur during segmentation. Possible solutions to these problems are proposed in Section 4.2.

The ASM segmentation using medical data was twofold. On the one hand the experiments on the medical data sets showed promising results. The best results were achieved on phalanx proximalis pollicis with segmentation accuracy up to 2.24 voxels (mean LLE). Nonetheless the small amount of training data and therefore a too restrictive model of phalanx distalis pollicis the segmentation was accurate. On the other hand the GLM fitted to the wrong contour during the experiments with os metacarpale I due to the fact that ossa sesamoidea was not taken into account when annotating the CT volumes and the strong trabecular structure. This led to only one successful segmentation in all six volumes. However, in those cases where the ASM search worked correctly it delivered equal or better results than seen in similar works. Possible solutions to the problems pointed out and improvements to the 3D ASM implementation are proposed in Section 4.2.

4.2 Outlook

Following changes to the segmentation pipeline and especially to the 3D ASM implementation used in this thesis are expected to improve performance and accuracy:

- Using MDL not for finding correspondences between point sets but for generating models that capture the biggest data variance with as few as possible LMs as described in [Davies10]. As more LMs also mean more computational costs for the single ASM steps (e.g. LGE extraction, GLM computation, surface reconstruction) this measure is believed to speed up the search process.
- Grey level volumes (e.g. cubes) at the LMs position are suggested to be used instead of grey level profiles along the surface's normals for the ASM search. Finding the highest correlation between volumes cut out at the LMs' positions during search and a model generated during the training phase provides the new LM position. It's believed that this change would counteract the influence of rotation. Thus delivering better results.
- Continuously decreasing the size of the search LGE when the ASM modes' change rate is below a certain rate would increase segmentation accuracy.
- Adding ossa sesamoidea to the model of os metacarpale I is believed to help segmenting this bone.
- At least 50 or more data sets should be used for the PDM creation to be able to project a bigger model variance during search. Also the changing trabecular structure inside os metacarpale I is believed to be modelled with the GLM this way.
- If possible data from people not suffering from RA should be used to build GLMs that depict healthy bone structure.

APPENDIX A

Additional Material and Results

In this appendix additional material, figures and results of the experiments with PDMs and ASMs are given.

A.1 PDMs

This section contains additional figures, showing examples of the variation of all PDMs created in this thesis.

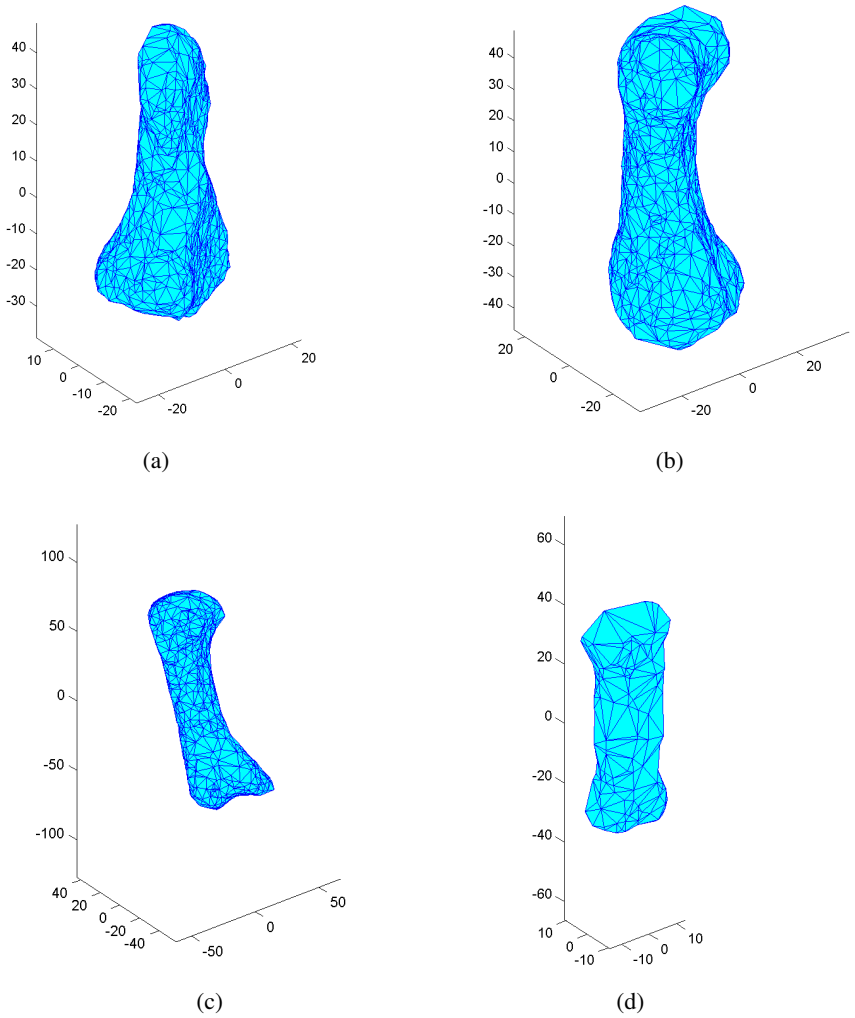
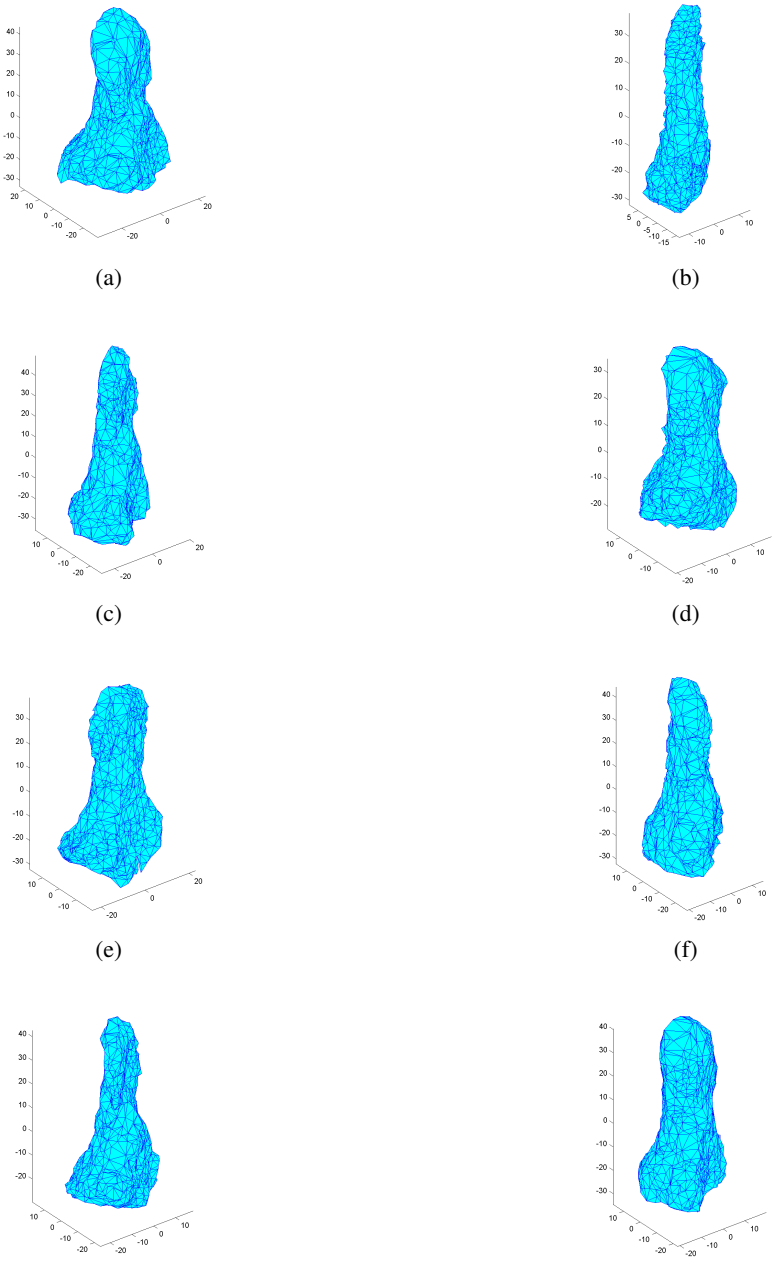


Figure A.1: Model of (a) phalanx distalis pollicis, (b) phalanx proximalis pollicis, (c) os metacarpale I, and (d) the synthetic data of the first leave-one-out runs with $\mathbf{b} = 0$.



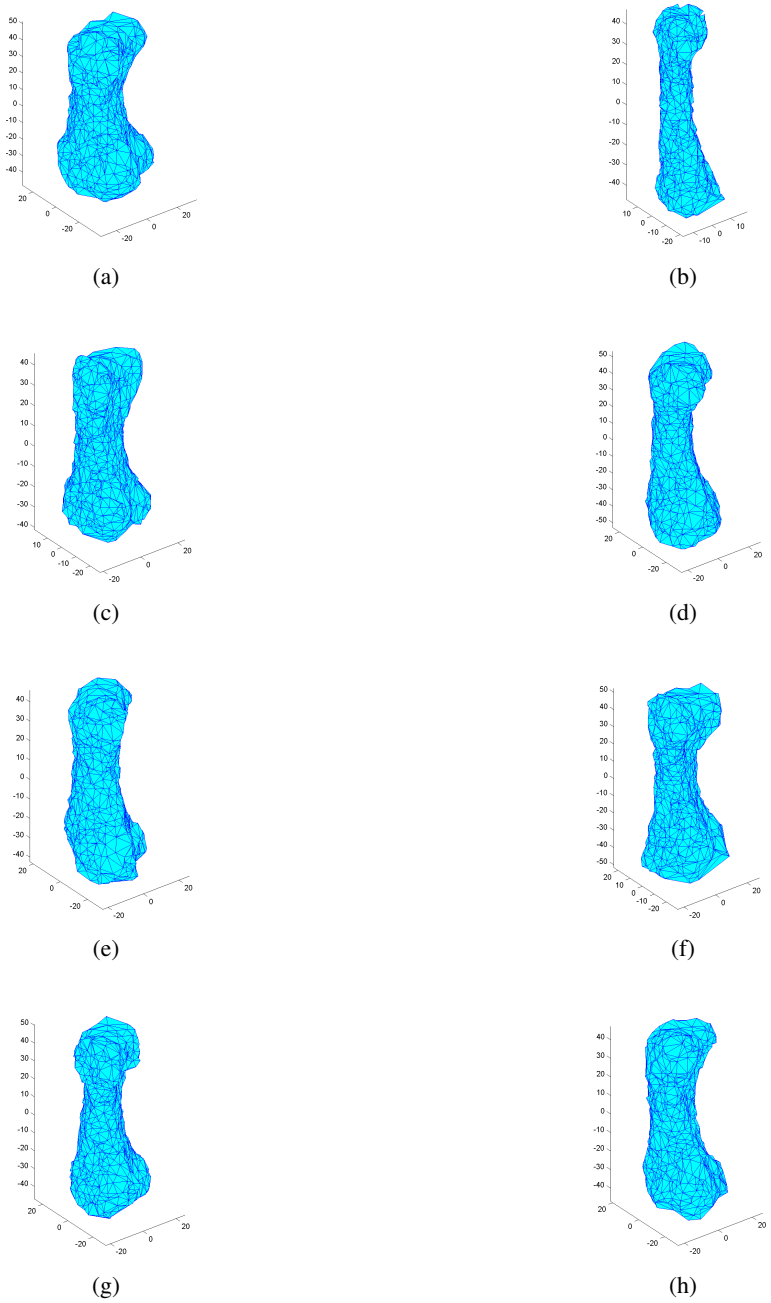
(a) (b)

(c) (d)

(e) (f)

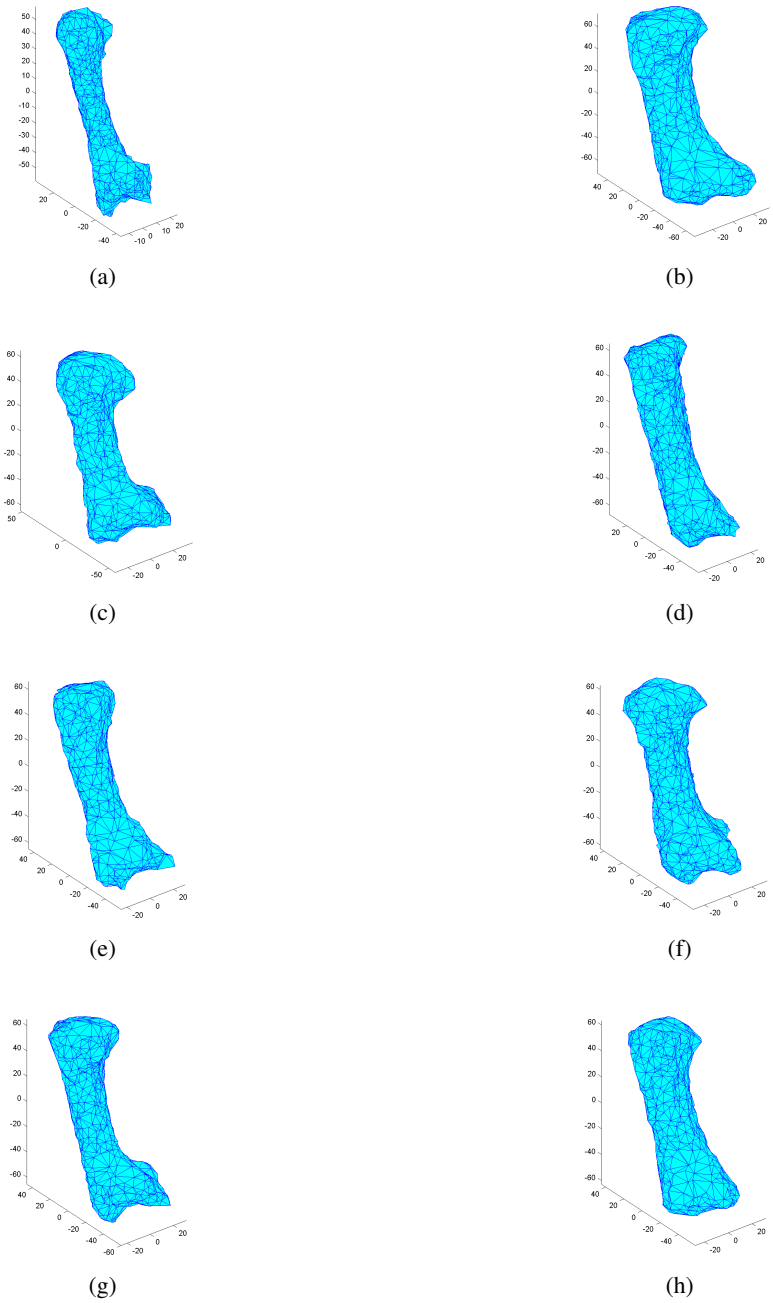
(g) (h)

Figure A.2: Phalanx distalis pollicis model of the first leave-one-out run with mode ranges from plus (first column) and minus (second column) three times the standard deviation of the first 4 modes. (a & b) show the variation in the first mode, (c & d) in the second mode, (e & f) in the third mode and (g & h) in the fourth mode.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure A.3: Phalanx proximalis pollicis model of the first leave-one-out run with mode ranges from plus (first column) and minus (second column) three times the standard deviation of the first 4 modes. (a & b) show the variation in the first mode, (c & d) in the second mode, (e & f) in the third mode and (g & h) in the fourth mode.



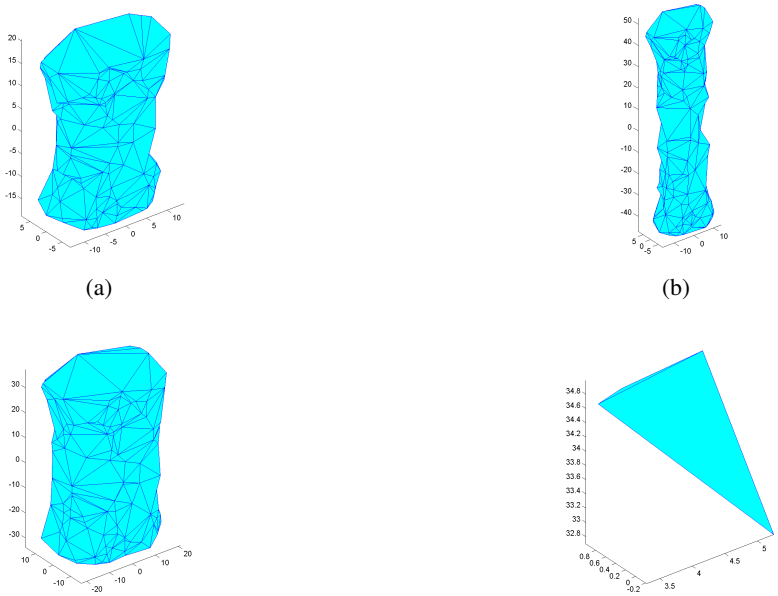
(a) (b)

(c) (d)

(e) (f)

(g) (h)

Figure A.4: Os metacarpale I model of the first leave-one-out run with mode ranges from plus (first column) and minus (second column) three times the standard deviation of the first 4 modes. (a & b) show the variation in the first mode, (c & d) in the second mode, (e & f) in the third mode and (g & h) in the fourth mode.



(a) (b)

(c) (d)

Figure A.5: Synthetic data model of the first leave-one-out run with mode ranges from plus (first column) and minus (second column) three times the standard deviation of the first 2 modes. (a & b) show the variation in the first mode and (c & d) in the second mode.

A.2 ASM Search

This section covers the outcome of the experiments with a maximum of 200 iterations and 0.01 as stopping criterion.

	Mean LLE			Median LLE		
No.	min	max	average	min	max	average
1	8.51	21.88	13.81	7.22	22.95	13.45
2	5.03	12.69	7.62	4.48	12.95	7.31
3	6.71	6.72	6.71	6.55	6.56	6.56
4	4.57	32.83	7.87	4.24	34.78	7.77
5	5.96	9.96	6.76	5.84	9.61	6.61
6	11.51	17.85	13.19	11.74	16.87	13.26

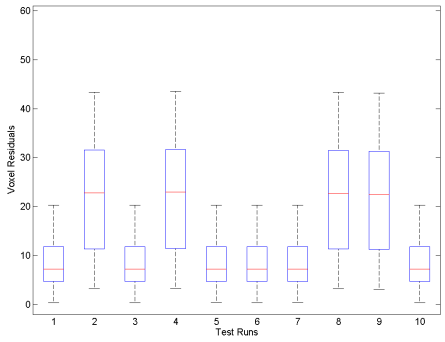
Table A.1: Phalanx distalis pollicis segmentation results: The minimum, maximum and average mean and median LLE in voxels of 10 search runs on CT volumes no. 1-6 is given.

	Mean LLE			Median LLE		
No.	min	max	average	min	max	average
1	4.48	5.29	4.88	4.30	4.97	4.63
2	8.66	8.80	8.72	8.83	8.95	8.89
3	2.41	2.92	2.60	2.17	2.70	2.38
4	3.56	15.69	4.81	3.09	14.72	4.28
5	1.79	1.80	1.80	1.74	1.74	1.74
6	5.82	6.54	5.98	5.04	5.60	5.17

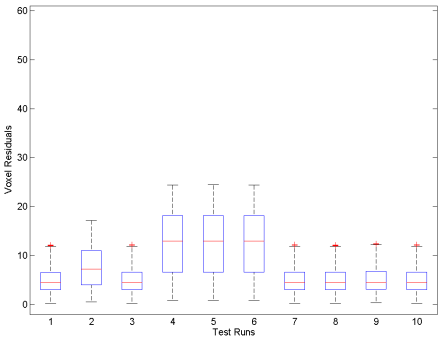
Table A.2: Phalanx proximalis pollicis segmentation results: The minimum, maximum and average mean and median LLE in voxels of 10 search runs on CT volumes no. 1-6 is given.

	Mean LLE			Median LLE		
No.	min	max	average	min	max	average
1	17.73	40.04	22.27	17.21	38.96	21.91
2	13.22	23.10	18.12	13.18	23.69	17.71
3	12.75	23.90	17.74	11.71	23.48	16.35
4	17.77	23.28	19.50	17.18	21.59	19.02
5	4.04	4.06	4.05	4.01	4.03	4.01
6	17.15	30.28	23.09	16.24	32.07	22.96

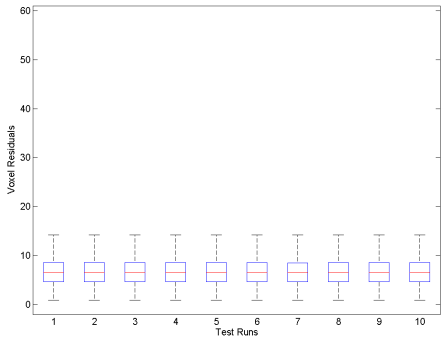
Table A.3: Os metacarpale I segmentation results: The minimum, maximum and average mean and median LLE in voxels of 10 search runs on CT volumes no. 1-6 is given.



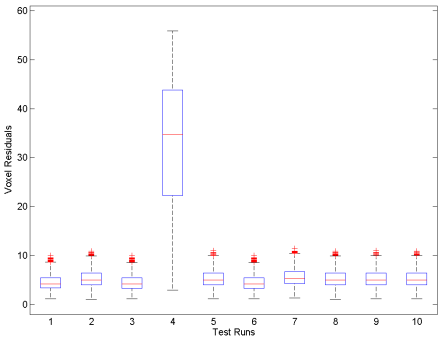
(a) CT Volume 1



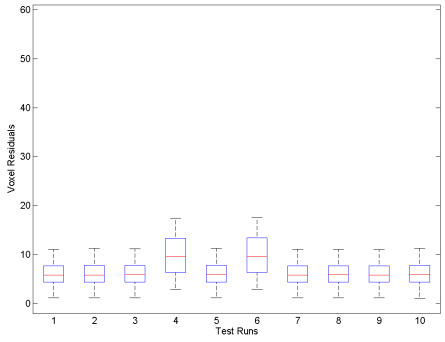
(b) CT Volume 2



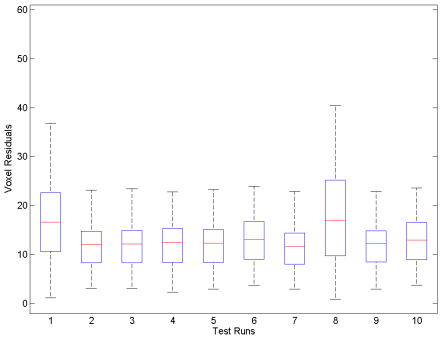
(c) CT Volume 3



(d) CT Volume 4

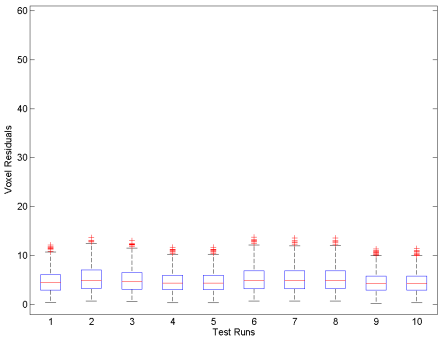


(e) CT Volume 5

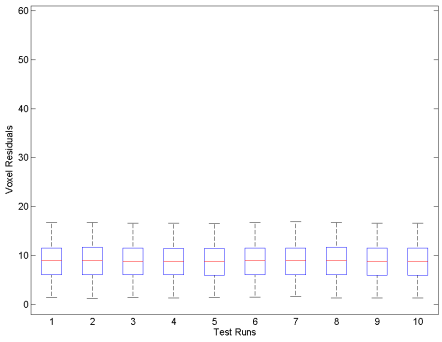


(f) CT Volume 6

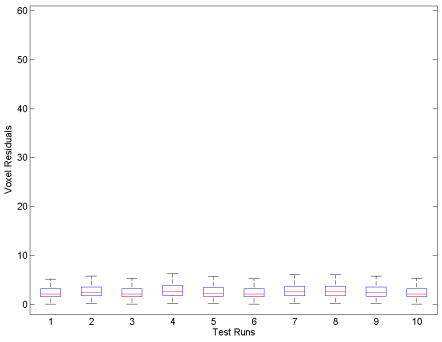
Figure A.6: ASM search results of phalanx distalis pollicis on CT volumes no. 1-6



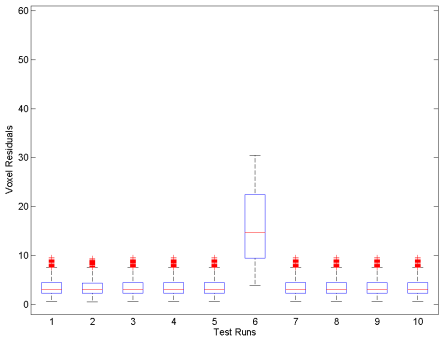
(a) CT Volume 1



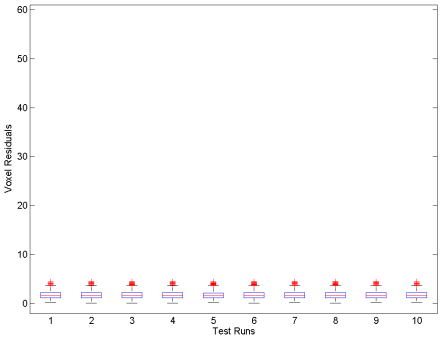
(b) CT Volume 2



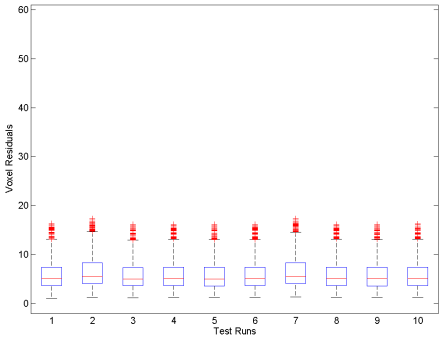
(c) CT Volume 3



(d) CT Volume 4

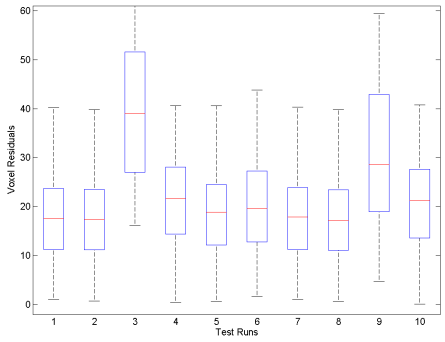


(e) CT Volume 5

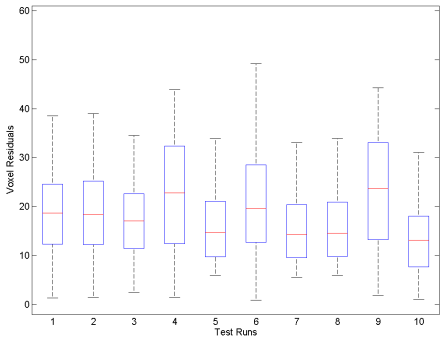


(f) CT Volume 6

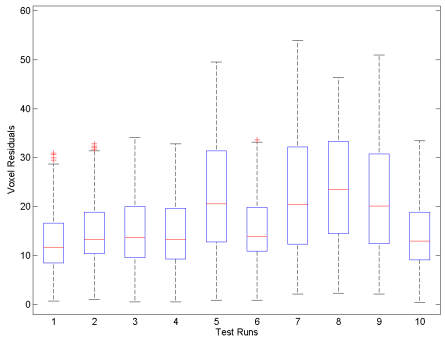
Figure A.7: ASM search results of phalanx proximalis pollicis on CT volumes no. 1-6



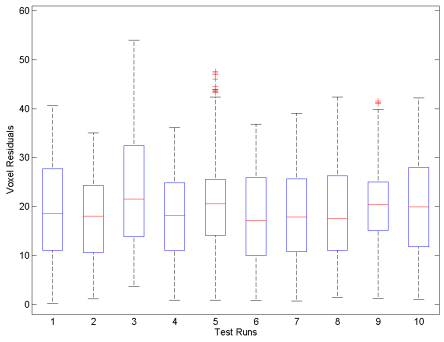
(a) CT Volume 1



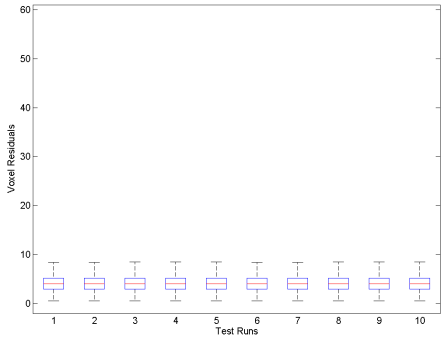
(b) CT Volume 2



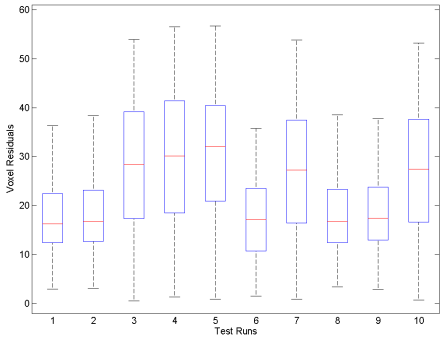
(c) CT Volume 3



(d) CT Volume 4



(e) CT Volume 5



(f) CT Volume 6

Figure A.8: ASM search results of os metacarpale I on CT volumes no. 1-6

APPENDIX B

Acronyms

2D	Two Dimensional
3D	Three Dimensional
4D	Four Dimensional
AAM	Active Appearance Model
AIC	Akaike Information Criterion
ASM	Active Shape Model
CPD	Coherent Point Drift
CT	Computed Tomography
EM	Expectation Maximisation
GLM	Grey Level Model
GMM	Gaussian Mixture Model
HU	Hounsfield Unit
ICP	Iterative Closest Point
LLE	Landmark to Landmark Error
LLM	Local Linear Mapping
LOOCV	Leave-One-Out Cross Validation
MDL	Minimum Description Length

MRF	Markov Random Field
MRI	Magnetic Resonance Imaging
PCA	Principal Component Analysis
PDM	Point Distribution Model
RA	Rheumatoid Arthritis
RF	Random Forest
RPM	Robust Point Matching
SAM	Sparse MRF Appearance Model
SIFT	Scale-Invariant Feature Transform
SPHARM	SPherical HARMonics
SSM	Statistical Shape Model
TPS	Thin Plate Spline

APPENDIX C

Erklärung zur Verfassung der Arbeit

Helmut Steiner
Nordwestbahnstraße 101/2/2
1200 Wien

"Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe."

Wien, 28.04.2011

Bibliography

[Amenta98] N. Amenta, M. Bern, M. Kamvysselis. A New Voronoi-based Surface Reconstruction Algorithm. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '98, pp. 415–421. ACM, New York, NY, USA, 1998.

[Besl92] P. J. Besl, H. D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14(2):pp. 239–256, 1992.

[Birngruber09] E. Birngruber, R. Donner, G. Langs. MatVTK - 3D Visualization for Matlab. The MIDAS Journal - Systems and Architectures for Computer Assisted Interventions, 2009. `http://hdl.handle.net/10380/3076` accessed 24-January-2011.

[Borigini10] M. J. Borigini. Rheumatoid arthritis. MedlinePlus Medical Encyclopedia, 2010. `http://www.nlm.nih.gov/medlineplus/ency/article/000431.htm` accessed 24-January-2011.

[Breiman01] L. Breiman. Random Forests. *Machine Learning*, vol. 45:pp. 5–32, 2001.

[Brooks77] R. A. Brooks. A Quantitative Theory of the Hounsfield Unit and Its Application to Dual Energy Scanning. *Journal of Computer Assisted Tomography*, vol. 1(4):pp. 487–493, 1977.

[Cheng95] Y. Cheng. Mean Shift, Mode Seeking, and Clustering. *IEEE TPAMI*, vol. 17(8):pp. 790–799, 1995.

[Chui00a] H. Chui, A. Rangarajan. A feature registration framework using mixture models. In *Proc. IEEE Workshop Mathematical Methods in Biomedical Image Analysis*, pp. 190–197. 2000.

[Chui00b] H. Chui, A. Rangarajan. A new algorithm for non-rigid point matching. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 44–51. 2000.

[Comaniciu02] D. Comaniciu, P. Meer, S. Member. Mean Shift: a Robust Approach Toward Feature Space Analysis. *IEEE TPAMI*, vol. 24:pp. 603–619, 2002.

[Cooper00] N. J. Cooper. Economic burden of rheumatoid arthritis: a systematic review. *Rheumatology*, vol. 39(1):pp. 28–33, 2000.

[Cootes92a] T. F. Cootes, C. J. Taylor. Active Shape Models - 'Smart Snakes'. In *Proceedings of the British Machine Vision Conference*, pp. 266–275. 1992.

[Cootes92b] T. F. Cootes, C. J. Taylor, D. H. Cooper, J. Graham. Training Models of Shape from Sets of Examples. In *Proceedings of the British Machine Vision Conference*, pp. 9–18. 1992.

[Cootes93] T. Cootes, C. Taylor. Active Shape Model Search using Local Grey-Level Models: A Quantitative Evaluation. In *4th British Machine Vision Conference*, pp. 639–648. BMVA Press, 1993.

[Cootes94a] T. F. Cootes, C. J. Taylor. Using grey-level models to improve active shape model search. In *Proc. 12th IAPR Int Pattern Recognition Vol. 1 - Conf. A: Computer Vision & Image Processing. Conf*, vol. 1, pp. 63–67. 1994.

[Cootes94b] T. F. Cootes, C. J. Taylor, A. Lanitis. Multi-resolution search with active shape models. In *Proc. 12th IAPR Int Pattern Recognition*, vol. 1, pp. 610–612. 1994.

[Cootes95] T. F. Cootes, C. J. Taylor, D. H. Cooper, J. Graham. Active Shape Models-Their Training and Application. *Computer Vision and Image Understanding*, vol. 61(1):pp. 38–59, 1995.

[Cootes04] T. Cootes, C. Taylor. *Statistical Models of Appearance for Computer Vision*. Tech. rep., Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, University of Manchester, 2004.

[Davies10] R. H. Davies, C. J. Twining, T. F. Cootes, C. J. Taylor. Building 3-D Statistical Shape Models by Direct Optimization. *IEEE Transactions on Medical Imaging*, vol. 29(4):pp. 961–981, 2010.

[Dempster77] A. P. Dempster, N. M. Laird, D. B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39(1):pp. 1–38, 1977.

[Dickens02] M. M. Dickens, S. S. Gleason, H. Sari-Sarraf. Volumetric Segmentation via 3D Active Shape Models. In *Proc. Fifth IEEE Southwest Symp. Image Analysis and Interpretation*, pp. 248–252. 2002.

[Dittrich09] E. Dittrich. *Automatic Model Generation for Sparse MRF Appearance Models using Minimum Description Length*. Tech. Rep. 120, Pattern Recognition and Image Processing Group, Institute of Computer Aided Automation, Vienna University of Technology, 2009.

[Donner07] R. Donner, B. Micusik, G. Langs, H. Bischof. Sparse MRF Appearance Models for Fast Anatomical Structure Localisation. In *Proc. BMVC'07*. 2007. `http://www.bmva.org/bmvc/2007/papers/paper-274.pdf` accessed 24-January-2011.

[Donner10a] R. Donner, E. Birngruber, H. Steiner, H. Bischof, G. Langs. Localization of 3D Anatomical Structures Using Random Forests and Discrete Optimization. In *Proceedings of Medical Computer Vision Workshop MICCAI 2010*, pp. 86–95. 2010.

[Donner10b] R. Donner, G. Langs, B. Mičušík, H. Bischof. Generalized Sparse MRF Appearance Models. *Image and Vision Computing*, vol. 28(6):pp. 1031–1038, 2010.

[Fischer10] L. Fischer. *Optimizing Shape Particle Filters for the Detection and Segmentation of Medical Images*. Master's thesis, Institute of Computer Aided Automation, Vienna University of Technology, 2010.

[Gold98] S. Gold, A. Rangarajan, C.-P. Lu, S. Pappu, E. Mjolsness. New algorithms for 2D and 3D point matching: pose estimation and correspondence. *Pattern Recognition*, vol. 31(8):pp. 1019–1031, 1998.

[Goodall91] C. Goodall. Procrustes Methods in the Statistical Analysis of Shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 53(2):pp. 285–339, 1991.

[Greig89] D. Greig, B. Porteous, A. Seheult. Exact Maximum A Posteriori Estimation for Binary Images. *Journal of the Royal Statistical Society Series B*, vol. 51(2):pp. 271–279, 1989.

[Grünwald05] P. Grünwald. A Tutorial Introduction to the Minimum Description Length Principle. In P. Grünwald, I. Myung, M. Pitt (Editors) *Advances in Minimum Description Length: Theory and Applications*, chap. 1–2. MIT Press, 2005.

[Hamprecht05] A. Hamprecht. *Genetische Polymorphismen in Toll-like-Rezeptoren, rheumatoide Arthritis und Höhe von Rheumafaktor im Serum*. Dissertation, Humboldt-Universität zu Berlin, Medizinische Fakultät - Universitätsklinikum Charité, 2005.

[Harris Jr.08] E. D. Harris Jr., G. S. Firestein. Clinical features of rheumatoid arthritis. In G. S. Firestein, R. C. Budd, E. D. Harris Jr., I. B. McInnes, S. Ruddy, J. S. Sergent (Editors) *Kelley's Textbook of Rheumatology*, chap. 66. Saunders Elsevier, 8 ed., 2008.

[Heimann07] T. Heimann, S. Münzing, H.-P. Meinzer, I. Wolf. A Shape-Guided Deformable Model with Evolutionary Algorithm Initialization for 3D Soft Tissue Segmentation. In N. Karssemeijer, B. Lelieveldt (Editors) *Information Processing in Medical Imaging*, vol. 4584 of *Lecture Notes in Computer Science*, pp. 1–12. Springer Berlin / Heidelberg, 2007.

[Heimann09] T. Heimann, H.-P. Meinzer. Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis*, vol. 13(4):pp. 543–563, 2009.

[Hug00] J. Hug, C. Brechbühler, G. Székely. Model-Based Initialisation for Segmentation. In D. Vernon (Editor) *Computer Vision — ECCV 2000*, vol. 1843 of *Lecture Notes in Computer Science*, pp. 290–306. Springer Berlin / Heidelberg, 2000.

[Ke05] Y. Ke, R. Sukthankar, M. Hebert. Efficient visual event detection using volumetric features. In *Proc. Tenth IEEE Int. Conf. Computer Vision ICCV 2005*, vol. 1, pp. 166–173. 2005.

[Langs03] G. Langs. *Automatic Quantification of Destructive Changes caused by Rheumatoid Arthritis*. Master's thesis, Institute of Computer Aided Automation, Vienna University of Technology, 2003.

[Langs07] G. Langs, R. Donner, P. Peloschek, H. Bischof. Robust autonomous model learning from 2D and 3D data sets. In *Proceedings of the 10th international conference on Medical image computing and computer-assisted intervention - Volume Part I*, MICCAI'07, pp. 968–976. Springer-Verlag, Berlin, Heidelberg, 2007.

[Langs09] G. Langs, P. Peloschek, H. Bischof, F. Kainberger. Automatic Quantification of Joint Space Narrowing and Erosions in Rheumatoid Arthritis. *IEEE Transactions on Medical Imaging*, vol. 28(1):pp. 151–164, 2009.

[Mortensen95] E. N. Mortensen, W. A. Barrett. Intelligent Scissors for Image Composition. In *SIGGRAPH '95: Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pp. 191–198. 1995.

[Müller10] H. Müller, J. Kalpathy-Cramer, B. Caputo, T. Syeda Mahmood, F. Wang. Overview of the first workshop on Medical Content-Based Retrieval for Clinical Decision Support at MICCAI 2009. In *MCBR-CDS 2009: Medical Content-based Retrieval for Clinical Decision Support*, no. 5853 in Lecture Notes in Computer Science (LNCS), pp. 1–17. 2010.

[Myronenko07] A. Myronenko, X. Song, M. Carreira-Perpinan. Non-rigid point set registration: Coherent Point Drift. In B. Schölkopf, J. Platt, T. Hoffman (Editors) *Advances in Neural Information Processing Systems (NIPS) 19*, pp. 1009–1016. MIT Press, Cambridge, MA, 2007.

[Myronenko10] A. Myronenko, X. Song. Point Set Registration: Coherent Point Drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32(12):pp. 2262–2275, 2010.

[NIAMS09] NIAMS. Handout on Health: Rheumatoid Arthritis. NIH Publication No. 09-4179, 2009. NIAMS (National Institute of Arthritis and Musculoskeletal and Skin Diseases).

[Pearson01] K. Pearson. On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine*, vol. 2:pp. 559–572, 1901.

[Rissanen78] J. Rissanen. Modeling by shortest data description. *Automatica*, vol. 14(5):pp. 465–471, 1978.

[Rissanen96] J. Rissanen. Fisher information and stochastic complexity. *Information Theory, IEEE Transactions on*, vol. 42(1):pp. 40–47, 1996.

[Schiebler07] T. H. Schiebler, H.-W. Korf. *Anatomie Histologie, Entwicklungsgeschichte, makroskopische und mikroskopische Anatomie, Topographie*. Steinkopff Verlag, 10 ed., 2007.

[Smith02] L. I. Smith. *A tutorial on Principal Components Analysis*. Tech. rep., Cornell University, 2002.

[Smith10] H. R. Smith. Rheumatoid Arthritis. Website, 2010. `http://emedicine.medscape.com/article/331715-overview` accessed 24-January-2011.

[Stegmann00] M. B. Stegmann. *Active Appearance Models*. Master's thesis, Technical University of Denmark, 2000.

[Thodberg03] H. Thodberg. Minimum Description Length Shape and Appearance Models. In *Information Processing in Medical Imaging*, vol. 2732 of *Lecture Notes in Computer Science*, pp. 51–62. Springer Berlin / Heidelberg, 2003.

[Valentinitsch06] A. Valentinitsch. *Documentation of radiographic follow-up studies on the musculoskeletal system focusing on rheumatic diseases: a systematic review*. Diplomarbeit, Carinthia University of Applied Sciences, Klagenfurt, Austria, 2006.

[vanAssen06] H. C. van Assen, M. G. Danilouchkine, A. F. Frangi, S. Ordás, J. J. Westenberg, J. H. Reiber, B. P. Lelieveldt. SPASM: A 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data. *Medical Image Analysis*, vol. 10(2):pp. 286–303, 2006.

[Viola01] P. Viola, M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition CVPR 2001*, vol. 1, pp. 511–518. 2001.

[Vitanyi00] P. M. B. Vitanyi, M. Li. Minimum description length induction, Bayesianism, and Kolmogorov complexity. *IEEE Transactions on Information Theory*, vol. 46(2):pp. 446–464, 2000.

[Wallace68] C. S. Wallace, D. M. Boulton. An Information Measure for Classification. *The Computer Journal*, vol. 11(2):pp. 185–194, 1968.

[Yuille88] A. L. Yuille, N. M. Grzywacz. The Motion Coherence Theory. In *Proc. Second Int Computer Vision. Conf*, pp. 344–353. 1988.

[Yushkevich06] P. A. Yushkevich, J. Piven, H. Cody Hazlett, R. Gimpel Smith, S. Ho, J. C. Gee, G. Gerig. User-Guided 3D Active Contour Segmentation of Anatomical Structures: Significantly Improved Efficiency and Reliability. *Neuroimage*, vol. 31(3):pp. 1116–1128, 2006.

[Zambal05] S. Zambal. *3D Active Appearance Models for Segmentation of Cardiac MRI Data*. Master's thesis, Institute of Computer Graphics and Algorithms, Vienna University of Technology, Favoritenstrasse 9-11/186, A-1040 Vienna, Austria, 2005.