Matthias Fenchel

Statistical Models for Segmentation from MR Localizer Images

Cuvillier Verlag Göttingen
Internationaler wissenschaftlicher Fachverlag

Matthias Fenchel

Statistical Models for Segmentation from MR Localizer Images

Cuvillier Verlag Göttingen
Internationaler wissenschaftlicher Fachverlag

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.
1. Aufl. - Göttingen: Cuvillier, 2010
Zugl.: Tübingen, Diss., Univ., 2009

978-3-86955-439-6

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2010
Gedruckt auf säurefreiem Papier

978-3-86955-439-6

STATISTICAL MODELS FOR SEGMENTATION FROM MR LOCALIZER IMAGES

DISSERTATION

der Fakultät für Informations- und
Kognitionswissenschaften
der Eberhard-Karls-Universität Tübingen

zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von
MATTHIAS FENCHEL
aus Tübingen

Tübingen
2009

Tag der mündlichen Qualifikation:	*16.06.2010*
Dekan:	Prof. Dr. Oliver Kohlbacher
1. Berichterstatter:	Prof. Dr. A. Schilling
2. Berichterstatter:	Prof. Dr. W. Straßer

Zusammenfassung

In dieser Dissertation werden Methoden zur Segmentierung anatomischer Strukturen in Planungsbildern der Magnetresonanztomographie (MRT), sogenannten Localizer-Bildern, vorgestellt. Localizer sind schnelle MR-Scanprotokolle zur Untersuchungsplanung. Segmentierungen anatomischer Strukturen aus diesen Bildern können für Anwendungen zur vollautomatischen Untersuchungsplanung, z.B. Organlokalisierungen, Schichtpositionierungen, Sequenzanpassungen, etc. verwendet werden. Da Localizer-Bilder nicht hinsichtlich Bildqualität sondern hinsichtlich Messzeit und Abdeckung optimiert sind, sind modellbasierte statistische Verfahren für die Segmentierung vorteilhaft.

Zwei Methoden werden vorgestellt: Die erste ist eine Methode zur Rekonstruktion von Leberform, -position und -orientierung aus einer Serie von wenigen 2D-Planungsschichtbildern mit großem Schichtabstand. Dazu wird ein Active Shape Model aus manuellen Lebersegmentierungen von 3D Trainingsbildern erstellt, das die durchschnittliche Leberform und die Hauptkomponenten seiner Varianz beschreibt. Korrespondierende Landmarkenpunkte auf der Oberfläche werden durch Remeshing mit Hilfe konformer Abbildungen in der sphärischen Domäne initialisiert und verfeinert durch Optimierung eines Korrespondenzmaßes, welches auf Minimum Description Length (MDL) basiert und die Kompaktheit des generierten statistischen Modells beschreibt. Die Segmentierung der Leber aus den gestapelten 2D-Schichtbildern erfolgt durch durch die Berechnung derjenigen Modellinstanz des Active Shape Models, welche bestmöglich die Bilddaten beschreibt. Man erreicht dies durch iterative Berechnung optimaler Verschiebungen der Landmarken. Die optimalen Verschiebungen beruhen auf Grauwertprofilen in den Bildern und einer normalisierten lokalen Statistik der Grauwertverteilungen in den Trainingsbildern. Die Instanz des Active Shape Models, die die gefundenen Verschiebungen der Landmarken am besten repräsentiert, wird durch eine Projektion auf den Linearraum des Active Shape Models gefunden. Daraus erhält man eine gültige Modellinstanz, die die Verschiebungen der Landmarken bestmöglich beschreibt.

Die Ergebnisse der Segmentierung aus generierten Localizer-Bildern werden mit den manuellen Segmentierungen mittels 4 Fehlermetriken verglichen. Die Ergebnisse zeigen, dass die Methode gegenüber Lebersegmentierungen mittels Active Shape Models aus 3D Daten konkurrenzfähig ist, wenn auch mit geringerer Präzision aufgrund der geringeren Bildqualität.

Die zweite Methode, die in dieser Dissertation vorgestellt wird, ist ein automatisches, anatomisches Labeling oder eine Multiorgansegmentierung anatomischer Strukturen in FastView-Bildern. FastView ist ein modernes MR-Protokoll, welches 3D Localizer-Bilder produziert, indem 2D-Schichten während kontinuierlichem Vorschub des Patiententisches gemessen werden. Die Segmentierung basiert auf einem statistischen Atlas des menschlichen Körpers, der aus einer Gruppe repräsentativer FastView Datensätze gewonnen wird. Der Atlas enthält einerseits ein statistisches Deformationsmodell, das verwendet werden kann, um unbekannte Datensätze auf die durchschnittliche Körperform des Atlas zu verformen. Zusätzlich enthält der Atlas ein statistisches Modell der Grauwertverteilungen, das verwendet werden kann, um gültige Atlasbilder zu erzeugen. Beide statistischen Modelle können verwendet werden, um einen unbekannten Datensatz auf den Atlas zu registrieren, indem die Modellinstanzen des Deformationsfeldes und des Atlas berechnet werden, welche die Ähnlichkeit zwischen dem Atlas und dem, durch das Deformationsfeld verformten, unbekannten Bilddatensatz maximiert. Das Labeling des unbekannten Datensatzes erhält man dann aus der Propagierung der anatomischen Labels

des Atlas auf den registrierten Datensatz. Da der Registrierungsprozeß die Optimierung einer Multiparameter-Zielfunktion mit vielen Freiheitsgraden und die Berechnung ihrer Ableitungen erfordert, und da jede Auswertung der Kostenfunktion eine 3D-Verformung der Eingabe erfordert, müssen Strategien zur Verbesserung der Laufzeit in Betracht gezogen werden. Dazu wurden das statistische Deformationsmodell und das statistische Grauwertmodell komplett auf die GPU (Graphics Processing Unit) portiert, was die Berechnung der kompletten deformierbaren Registrierungen mit Grafikhardwarebeschleunigung erlaubt. Die Hardwarebeschleunigung umfaßt somit die gesamte Berechnungskette der Registrierungen inklusive der Erzeugung der Instanzen des Deformationsfeldes und der Grauwertverteilungen, welche die Auswertung der linearen Modelle durch Matrixmultiplikationen und Vektoradditionen erfordert. Die Laufzeiten des Registrierungsprozesses sind etwa 10-30s, was der Größenordnung der Dauer der Bildakquisition selbst entspricht. Diese Laufzeiten erlauben auch einen praktischen Einsatz in der klinischen Routine.

Die Validierung der atlas-basierten Segmentierungen erfolgt durch manuelles Setzen korrespondierender Landmarken an definierbaren Punkten in den Eingabebildern und im Atlas. Die Abweichungen der Landmarken im Atlas von den registrierten Landmarken im Bilddatensatz werden als Fehlerkriterium für die Güte der Methode verwendet und mit der Reproduzierbarkeit der manuellen Markierung von Landmarken verglichen.

Abstract

This thesis presents methods for the segmentation of anatomical structures from magnetic resonance (MR) localizer images. Localizer images are obtained from fast pre-scan protocols and are usually used for scan planning. Segmentations of anatomical structures from these images can leverage applications in automated scan planning like organ localization, slice positioning tasks, sequence adaptations, etc. Since localizer images are not optimized for image quality but for scan time and scan range, statistical and model based approaches are preferred for the segmentations.

Two methods are proposed: The first is a method for reconstructing liver shape, position and orientation from a set of stacked sparse 2D localizer images. For this purpose an active shape model, which represents the average liver shape and its major modes of statistical variation, is created from a set of liver segmentations from 3D training data images. Corresponding landmark vertices on the surfaces are found by remeshing the surfaces using conformal mappings in the spherical domain and tuning of the correspondences by means of a measure based on minimum description length (MDL) which describes the compactness of the model. Segmentation of the liver from the stacked localizer image data is done by computing the model instance of the active shape model which best fits the given image data. This is achieved by iteratively computing optimal displacements of the landmark vertices based on image profiles according to the statistics of local grey value appearance in the training image data. The instance of the active shape model which best represents the displaced landmark vertices is found by a projection to the linear space of the model. This yields the instance of the model which is closest to the landmark displacements and ensures that the shape is valid.

The results of the segmentations from generated localizer images are compared to the ground truths of the manual segmentations by comparing them according to four error metrics. The results show that the method can compete with standard 3D active shape model based segmentation approaches, though with less precision due to lower input image quality.

The second method presented in this thesis is an automatic anatomical labeling or multi-organ segmentation of anatomical structures from FastView images, which is a modern MR imaging protocol producing 3D localizer images by acquiring 2D slices during continuous movement of the patient table. The segmentation is based on a statistical atlas of the human body created from an image data set of representative FastView images from a group of volunteers. The statistical atlas on the one hand consists of a statistical model of deformation, which can be used to create valid instances of deformation fields to warp unseen image data sets to the mean shape of the atlas. On the other hand, the atlas contains a statistical model of grey value appearance, which can be used to create instances of atlas images with the mean shape and valid grey value appearance. Both statistical models can be used to register an unseen data set by finding the instances of the deformation and local model of appearance, which maximize the similarity between the atlas with the grey values from the local model of appearance and the unseen image, which is warped according to the deformation model. Propagating anatomical labels from the atlas to the registered image then yields the anatomical labeling of the image data. Since the registration process requires optimizing a multi-parameter cost function with various degrees of freedom and computing its derivatives and since each evaluation of the cost function requires a 3D image warp, performance speed-up strategies have to be considered. It was decided to shift the statistical models of deformation and appearance completely to the

GPU (Graphics Processing Unit), which permits computing the complete non-rigid registrations with GPU hardware acceleration, including the calculations of the instances of both the deformation field and grey value appearance which require evaluating the linear models by matrix multiplications and vector additions. The running times of the registration process are about 10 to 30 seconds which is the same magnitude as the image acquisition itself. Running times of this magnitude allow for practical usage of the method in clinical MR routine.

Validation of the atlas based segmentations is done by manually placing corresponding landmarks at descernible points in the input FastView image data sets and the atlas. The displacements between the landmarks in the registered data sets and the atlas are taken as an error criterion for the exactness of the method and can be compared to the reproducibility of the manual landmark placements.

Acknowledgements

Writing this thesis would not have been possible without the contributions of whole lot of people, who gave advice, ideas or just mental support. First of all I would like to thank my supervisor, Prof. Andreas Schilling, for ideas, thoughts, proofreads and critics and above all his willingness to do the supervision of an external thesis. He was always eager to help, listen to problems and suggest solutions when it was necessary. Besides I would like to thank all the people at GRIS, the chair for Graphical Interactive Systems for fruitful discussions, hints and ideas, help on search for literature, exchange of source code, reviews, proofreading and more.

Furthermore, I am grateful to the people at the Department for Applications and Workflow of the MR division at Siemens Healthcare (formerly Siemens Medical Solutions) in Erlangen for providing me the scholarship. Many people at Siemens MR helped me during this thesis, e.g. by acquiring image data, by getting in touch with cutting-edge MR applications and techniques, and by sharing their tremendous knowledge and experience in MR imaging. The most important person to mention is Dr Stefan Thesen, who supported this thesis in every imaginable way. Without his contributions this project would not have been possible. I am also grateful to the colleagues, the members of the former team for visualization & processing for the kind admission to the group, for giving me advice, help and fun and motivation at work and also for letting me participate in the company's day-to-day business. I would also like to mention Dr Christoph Zindel for giving me the opportunity to join his department as a PhD student.

I would like to thank the collaboration partners and colleagues at the Imaging Science Institute and UKT in Tübingen for their discussions and feedback.

Last but not least, special thanks goes to my parents for making all my academic efforts possible and for supporting me in so many ways. I would also like to thank my brother Michael for his contributions and discussions from a clinical point of view.

A final dedication goes to all my friends and Anette who supported me by providing distraction in good times and strength and motivation to carry on in bad times.

Ever tried. Ever failed. No matter. Try again. Fail again. Fail better.
Samuel Beckett, Worstward Ho, 1983

Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

Automation and workflow improvements of magnetic resonance (MR) tomography examinations become more and more important. With patient numbers increasing and reimbursement rates being reduced, there is growing need for automation, simplicity and reproducibility of MR examinations. As a matter of fact, the complexity, the lack of reproducibility and the long examination times are among the main disadvantages of MR as an imaging modality in clinical practice in comparison to other modalities like computed tomography despite all its other advantages.

Automation and simplicity helps accelerating and standardizing examination processes and reduces the need for highly trained staff to operate the scanners. Both of these factors lead to a reduction of costs. The effects of shorter slot times have been studied in a variety of studies like [JDBH07] and publications by [RSC02], [Ess00] or [WL01] . It is often also the only solution to cope with increasing patient and case numbers [Yea07].

Increasing the reproducibility enhances the comparability of examinations of different individuals and follow-up examinations of identical individuals. Standardization neutralizes the variability induced by different human operators. This can also lead to a significant gain in quality because simplicity and standardization helps minimizing the likelihood of mistakes and acquisition of useless images and data. Avoiding re-examinations and rescans also alleviates the patients' stress and discomfort.

Scan automation may include auto-align or positioning tasks, like automatic positioning of slices, automatic adaptation of sequences and protocols, automatic hardware selection, etc. to name but a few. Since MR is a non-invasive imaging modality and no long-term harmful effects to patients undergoing MR exams are known, it is possible to acquire additional data apart from the actual diagnostic images without ethical or regulatory objections. It is even possible to scan clinically irrelevant image data, which are later discarded. This opens up a variety of possible applications by performing dedicated pre-scans and implementing post-processing algorithms, e.g. for the automation of certain workflow steps in the MR examination workflow which are currently done manually. Scan planning and positioning of slices and ROIs are examples of manual workflow steps that would benefit from algorithmic automation based on processing of image data which are acquired using certain pre-scan protocols. The pre-scan protocols which are used for scan planning and positioning purposes are also called scouts or localizers. Protocols for the acquisition of localizer images are optimized with respect to

different objectives than diagnostic images. Localizer protocols usually have short acquisition times, large fields of view and moderate image quality. Often, the localizers only sparsely cover the volume of interest, i.e. the volume is scanned in 2D slices with large gaps between the slices. The general purpose of a localizer image series is to provide a large range of topographic overview of an individual's anatomy within the shortest possible time. For automatic positioning tasks, structural information has to be inferred from the localizer images, resp. the structures have to be segmented automatically from the images. The goal is to establish spatial correspondence in anatomical structures from one individual to another and equally important for the follow-up case of one individual being scanned multiple times. Once the segmentations have been calculated and the anatomical structures extracted, it is straightforward to find spatially corresponding points, planes and slice positions, which are necessary for reproducible positioning. As image quality can be poor in MR localizer images, fully automatic segmentation can be challenging. For this reason, model and learning based approaches along with statistical methods, which take prior knowledge of the structures to be segmented into account, are most promising.

1.2 Contributions

This thesis presents two methods for the segmentation of anatomical structures from MR localizer images based on statistical models.

The first method uses a 3D active shape model to reconstruct liver position, orientation and shape from a set of sparsely covering image slices. This can be used for scan automation tasks like slice and navigator positioning and other applications that require 3D anatomical information from localizer slice image data. The active shape model is created beforehand from a set of 3D training data images by statistical analysis of the 3D liver surfaces obtained from manual segmentations of the liver. The active shape model search algorithm then computes the most probable position, orientation and shape for each individual being scanned given the sparsely covering 2D image data of the individual. The contributions consist of the creation of the active shape model from manual segmentations, including a solution to the correspondence problem of the landmark vertices in 3D. This solution is based on remeshing and tuning the positions of the corresponding landmarks by optimizing a compactness measure from information theory based on minimum description length (MDL). For the segmentation resp. the active shape model search, two strategies are proposed. The two strategies differ by the data structures and comparison algorithms which are used as a local model of grey value appearance. The local model of grey value appearance is necessary to quantify the quality of the segmentations by establishing a correlation between the statistical shape model and the image data, i.e. how well the model fits to the image data. The segmentation algorithm iteratively first searches for the most probable landmark displacements of the surface of the active shape model. A subsequent subspace shape model projection finds the most probable instance from the model space by considering just a few hundred reliable vertex landmark displacements. This algorithm and the results have been published in [FTS08b].

The purpose of the second method is a fully automatic labeling of anatomical structures from a fast 3D whole-body localizer, the so called FastView imaging protocol. To this end, a

statistical atlas is created consisting of a statistical model of deformation and a statistical model of grey value appearance. The former is obtained from statistical analysis of the deformation fields from non-rigid registrations of a set of representative training images to a reference image. The latter is calculated by pixelwise statistical analysis of the normalized grey values from the registered images. The organs and structures of interest are segmented in the atlas by a human expert and the anatomical labels are assigned according to these segmentations. The actual segmentation of an image from an individual is done by registering unseen image instances to the atlas and propagating the anatomical labels from the atlas. A major contribution is the complete implementation of the atlas and its statistical models on the graphics processing unit (GPU), which provides a a huge gain in performance. All steps of the atlas based registration, i.e. the task of finding the instance of the atlas which is closest to the unseen image, are computed using GPU based fragment shaders which benefit from the massive computation power provided by modern GPUs. This algorithm and the results were published in [FTS08a].

1.3 Outline

The thesis is structured as follows: First, a short introduction into MR imaging is given with an emphasis on the MR imaging techniques which are the basis for the scan protocols used in this thesis. The purpose of this chapter is to make the reader familiar with the basic concepts and terms of MR imaging in order to understand the possibilities and limitations, the strengths and weaknesses also in comparison to other medical imaging modalities like computed tomography from the point of view of a reader with a background in image post-processing.

The next chapter describes the algorithm and results for reconstructing liver shape and position from sparsely covering 2D MR slice image data. The first part of the chapter is dedicated to the creation of a statistical surface model of the liver. The second part deals with the model based segmentation guided by the active shape model.

The third chapter deals with the automatic labeling or multi-class segmentation of FastView images using a statistical atlas. The first part of this chapter describes the creation of the atlas, while the second part explains the segmentation method based on registering unseen data sets to the atlas and propagating the labels from the atlas.

The thesis concludes with a summary and discussion of the methods described and provides an outlook on possible future developments, applications and limitations.

2 MR Imaging

2.1 Introduction and History

This chapter provides a brief introduction into magnetic resonance imaging as a medical imaging modality and as the basic technique used for the acquisition of the images for which the concepts developed throughout this thesis are developed. The purpose of this chapter is twofold. One the one hand the reader should be given a brief introduction into the basic concepts and modern developments of magnetic resonance imaging (MRI or NMR[1]). On the other hand, from a point of view of a reader with a background in image processing, it is important to be aware of the possibilities and limitations of the MR image acquisition process. Often people with profound knowledge in image processing make an enormous effort to come up with intricate solutions for processing and analysis of MR images. However, tuning the acquisition side, e.g. by making changes in the pulse sequence, the scanning parameters resp. protocol or the image reconstruction from the raw data, the problems could be solved easily without sophisticated post-processing algorithms. For example, the segmentation of a particular anatomical structure may be impossible given a standard image contrast. By making use of another MR imaging contrast, the segmentation can be done by a simple thresholding or windowing of the grey values. It is therefore important in MR image processing to keep an eye on both the acquisition and the processing of the images, the more so in the case of this thesis where steps of human interaction the examination workflow should be automatized by means of image processing algorithms.

Magnetic resonance imaging is a modality for in-vivo imaging of structures or functions of the body. The term magnetic arises from its basic principle of the interaction of nuclear magnetic spins with external magnetic fields. The external magnetic fields may be static or varying over space and time. In the latter case they are also referred to as magnetic gradients resp. as a sequence of gradients. The term resonance insinuates the interaction with external radio frequency fields. In simple words, MR data represent the induction signal of macroscopic magnetization of aligned nuclear spins in receiver coils.

The huge success of MRI is based on two facts: First, the 'non-invasive' nature of magnetic fields allows for examinations without harming effects and without the use of ionizing radiation. Second, MRI is extremely rich in information, as its data may not only deliver information about the anatomical structure but also about physical properties like flow, temperature,

[1] nuclear magnetic resonance

diffusion, etc. Physiological effects, like stimulation of brain activity as in fMRI [HBTV99] can be observed, or even metabolic processes of certain parts of tissue, as in MR spectroscopy applications.

Historically, the roots of NMR go back to the 1940s. Bloch [BHP46] and Purcell[PTP46] described the quantum mechanical effects of nuclear spins in magnetic fields in 1946. They managed to measure the signals of water and paraffin and were awarded the Nobel Prize for their contributions in 1952. In 1973, Lauterbur [Lau73] and Mansfield [MK73] laid the foundation for MR as an imaging technique by making spatial encoding possible, which 30 years later resulted in being awarded the Nobel Prize for their breakthrough concepts.

The value of MRI as a diagnostic medical imaging modality was first discovered by Damadian [Dam71]. He pointed out the value of MRI for medicine by examining that tumors can have significant effects on the behavior of the magnetic signal, which means that lesions and tumors may be detected from irregularities in the MR images.

2.2 Comparison to Computed Tomography

In common discussions, there is often a lot of confusion about the differences and similarities of magnetic resonance imaging (MRI) and computed tomography (CT). Often the terms are mixed up, and even if there is knowledge about the physical principles of the image acquisition that the two modalities are based on, people are not familiar with the particularities of the two. Often, it is assumed that the two modalities could produce equal images or that they could be used interchangeably. As a matter of fact, this is not the case. This section will explain in short terms the major particularities and differences between the two modalities.

The basic differences between MR and CT are best explained, if one considers the physical principles that the image acquisition is based on. In MR, the interactions of the nuclear spins with external magnetic fields and RF pulses are measured by detecting the resonance signal of nuclei after excitation. CT images are reconstructed from multiple projection images, acquired by a rotating X-Ray tube. Contrast in CT images is therefore mainly determined by the different densities of the tissues with respect to the absorption of the X-rays. As for MR images, contrast is determined by a whole set of parameters, like the density of protons in the tissue, their relaxation times, chemical properties, etc. Depending on the scanning technique, it is also possible to measure other physical values which may be related to physiological properties, like flow or temperature, diffusion in diffusion weighted imaging, brain activity in functional MRI, concentration of metabolites in chemical shift imaging, etc. [HBTV99].

In general, a CT exam is faster, cheaper and easier to perform than an MR exam. It is usually the first imaging modality of choice for emergency patients with multiple traumata. Because of its speed and high reproducibility it is also often and commonly used in many clinical standard examination procedures. There is no general answer to the question which modality is superior or produces better images. Both modalities have weaknesses and strengths. Basically, MR is better for contrasts in soft tissue [DHL+89], whereas CT produces better morphological contrasts. Therefore, MR is often superior for detection and identification of tumors while CT is unrivaled in detecting bones and calcification.

Both modalities have certain hazards for patients being examined. Patients with metallic

implants, cardiac pace makers or tattoos containing iron should not undergo MR examinations, because movement of these objects caused by the magnetic fields or induction heating may occur. Other hazards include claustrophobia from which patients being examined may suffer, eventual peripheral nerve stimulation in the case of magnetic gradients being switched too quickly and hyperthermia from absorption of radio frequency energy. Hyperthermia is avoided by consideration of the specific absorption rate (SAR), which defines the ratio of temporal increase in temperature caused by RF energy [DWRB02]. These calculations often require complex calculations and modeling. Many of the other effects are avoided by hardware and software limitations for the scanning parameters. Apart from that, no major negative long-term effects have been observed for MR examinations in contrast to CT where ionizing radiation is used, although there is a possibility of minimizing this effect by low-dose CT scans. As a consequence of the fact that no long-term harming effects have been observed for MR, there are no regulatory objections against acquiring additional MR images, which are of no diagnostic value. This allows for setting up image processing applications which require additional special images. Since these images are not further used, their contrast can be designed and optimized for the need of the application or processing algorithm. For instance, this is the case for the localizer images used in this thesis.

2.3 Basic Principles

2.3.1 Spins, Excitation and Relaxation

Spin Angular Moment

Depending on their composition of neutrons and protons, atom nuclei can have an intrinsic spin angular moment or shorter a spin. In accordance with quantum theory, this spin can only have discrete quantized values.

It turns out that nuclei with an impair number of protons and an impair number of neutrons have integral spin values (e.g. deuterium ^2H) and nuclei with with an even pair of protons and an impair number of neutrons have half-integral spin values (like the most abundant isotope of hydrogen ^1H [2]). Nuclei of atoms with even numbers of both protons and neutrons do not have any value of spin at all. The latter type of atoms is not suitable for MRI.

The quantized angular spin moment can be expressed by the following equation:

$$L = \hbar\sqrt{I(I+1)} \tag{2.1}$$

where L is the angular spin moment, $\hbar$ Planck's constant[3] divided by 2π and I the respective integral or half-integral quantum number.

The nuclear spins cause a magnetic field of a strength B that is directly proportional to the angular spin moment. The proportionality constant γ is called the gyromagnetic ratio, which is a specific constant of every isotope.

[2] since ^1H hydrogen nucleus consists of only one proton, it is common to use the expression proton equivalently

[3] $\hbar = 1.0545Js$

$$\vec{B} = \gamma \vec{L} \tag{2.2}$$

where $\vec{B}$ is the magnetic field strength, γ the gyromagnetic ratio and $\vec{L}$ the angular spin moment.

Because of its abundance in the human body and its high gyromagnetic ratio, ${}^1H$ is the most interesting isotope in terms of magnetic resonance imaging with a half-integral angular spin moment of $\frac{1}{2}$.

Another phenomenon of quantum mechanics is that the expectation values of the angular spin moments are not only quantized in their values but also in their orientations with respect to a specific axis. It turns out that, if the quantum number of the angular spin moment is I, there exist $2I + 1$ possibilities for the angular spin to orientate along a specific axis. These orientations are characterized by the $2I + 1$ values between $-I$ and $+I$. These values are also called the orientation quantum numbers.

For protons, for example, with an angular spin moment of $\frac{1}{2}$ there exist two orientations along a certain axis with the orientation quantum numbers $-\frac{1}{2}$ and $\frac{1}{2}$. For deuterium with an angular spin moment of 1 there exist three orientations with the orientation quantum numbers -1, 0 and $+1$, etc.

Interaction with an External Magnetic Field

Under normal conditions, the different states of orientation are energetically equivalent. Given the presence of an external magnetic field, however, an effect called Zeeman interaction can be observed: the energetic equivalence is split up into $2I + 1$ different states, one for each orientation quantum number along the axis defined by the direction of the given external magnetic field, as in figure 2.1 a): on the left side there is energetic equivalence, on the right side, under the presence of the external magnetic field B_0 the equivalence is split up in a state of anti-parallel orientation and high energy and a state of parallel orientation and low energy. The direction of the external magnetic field B_0 defines the z-axis of the coordinate system.

In a classical description, under the presence of an external magnetic field, the angular magnetic moment rotates or precesses around the axis of the magnetic field. The axis of precession is given by the direction of the z-axis as shown in figure 2.1. This phenomenon is caused by the physical law of the constancy of an angular moment which is the same that makes a spinner precess around the direction of gravity. The frequency of the precession is called the Larmor frequency:

$$\omega_L = \frac{\gamma B_0}{2\pi} \tag{2.3}$$

The difference in energy between the two states of spins of the protons is directly proportional to the Larmor frequency and therefore proportional to the external magnetic field as well:

$$\Delta E = h\omega_L = \hbar\gamma B_0 \tag{2.4}$$

where h is Planck's constant, $\hbar = \frac{h}{2\pi}$ as before, B_0 the magnetic field strength and ω_L the respective Larmor frequency.

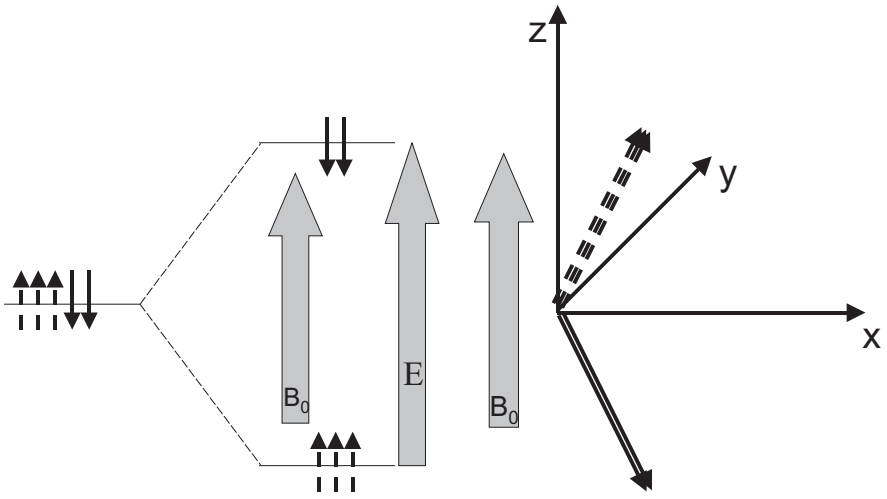


Figure 2.1: a) shows a Zeeman diagram. Whereas on the left side there is energetic equivalence of the up and down spin states of a proton. On the right, situation under the presence of an external magnetic field B_0. The state with spin parallel to the external magnetic field is energetically lower than the state anti-parallel. b) shows the orientation of the spins parallel (dashed) or anti-parallel (solid) to the external magnetic field B_0. The direction of B_0 determines the z-axis of the given coordinate frame.

In MR experiments the induction signal caused by precessing protons is measured. The vast majority of the protons in biological tissues are bound either in water or in fat. Due to their different chemical surroundings their Larmor frequencies differ by a small value. This difference is called the chemical shift and can be a source for artifacts as section 2.5.6 explains.

Macroscopic Description and MR Signal

Because of the energy difference between the two states, they are not occupied equally by the protons. The occupation difference can be quantized by a result from statistical thermodynamics in the so called Boltzmann equation:

$$\frac{N_{upper}}{N_{lower}} = e^{-\frac{\Delta E}{kT}} \tag{2.5}$$

In this equation N_{upper} and N_{lower} describe the number of nuclei in the upper respectively lower level of energy. ΔE describes the energy difference between the states. T is the temperature and k the Boltzmann constant[4].

At room temperature of 293 K and a magnetic field strength B of 1 Tesla there is an excess of $1 : 10^6$ nuclei in the lower state in comparison to the upper state. To give an idea, in a probe of 9 g water containing roughly $6 \cdot 10^{23}$ protons, this means $6 \cdot 10^{17}$ protons. This is enough to

[4] $k = 1.3806 \frac{J}{K}$

observe a macroscopic net magnetization of the tissue under influence of the external magnetic field: since more nuclei have their spins oriented parallel to external magnetic field than in the opposite direction, the tissue has experienced a net magnetization in the direction of the external magnetic field. That is the macroscopic description which is the overall vectorial sum of all nuclear magnetic moments. The stronger the basic magnetic field, the larger the excess of magnetization. Most modern clinical MR systems have magnetic fields of 1.5 or 3 Tesla. However, there are also research systems with higher basic magnetic fields.

As the spins of the nuclei precess incoherently, their phases sum up to 0, so there is no macroscopic net magnetization in a plane perpendicular to the basic magnetic field. As the B_0 direction is the z axis, this perpendicular plane is called the transversal plane or the xy-plane. If all nuclei were in phase, a net magnetization in the xy-plane precessing around the z-axis could be observed. The next section shows how phase coherence and net magnetization in the xy-plane can be achieved by means of RF pulses.

The measured signal in an MR experiment stems from the detection of the electromotive force induced by the precessing magnetization. As the receiver coils are aligned perpendicular to the xy-plane, the transversal net magnetization M_{xy} is the one which is measured as a complex MR signal with the x component being the real and the y component being the imaginary part.

RF Pulses

An RF pulse or simply pulse can be described as an electro-magnetic wave with periodic magnetic content B_1, interfering with the net magnetization M and causing a rotation of the net magnetization. The angle of the rotation of the magnetization is determined by the length, the strength and the bandwidth of the pulse. The rotation axis is determined by the directions of B_0 and B_1.

For example, a 90° pulse causes the net magnetization in z-direction to rotate into the xy-plane, as delineated in figure 2.2. In the same way, other pulses, for example 180° pulses can be applied, which cause the z magnetization to be inverted into -z direction.

After the application of the pulse the protons will gradually return to the equilibrium state, which is a net magnetization in z direction, in the same way that a ball positioned at a steep hill rolls downhill towards the lowest state of energy. This phenomenon is called relaxation and has various reasons, among them the tendency for energy stability. The process of applying a pulse to achieve a higher energetic state by rotating the net magnetization is called excitation.

Spin Lattice Relaxation

Spin lattice relaxation or longitudinal relaxation is the process in which macroscopic net magnetization along the z axis is restored. It is an enthalpy driven process, and it can be described by an exponential decay:

$$M_z(t) = M_{z_0} \cdot e^{-\frac{t}{T_1}} \tag{2.6}$$

$M_z(t)$ is the net magnetization in z direction after time t. M_{z_0} is the net magnetization in z direction before the experiment starts at $t = 0$. T_1 is defined as the time needed by the system

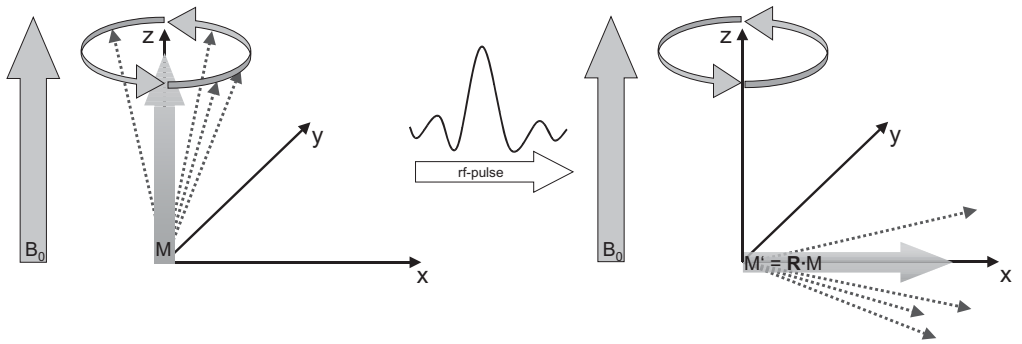


Figure 2.2: The effect of a 90° RF-pulse on net magnetization M. The direction of M is rotated according to the rotation matrix R perpendicular to B_0 and the rotation axis. R is defined by the length and magnitude of the pulse. In this case it is a 90° rotation. The rotated net magnetization M' keeps precessing around the z-axis.

to recover all but $1/e$ of the original net magnetization in z direction.

Considering T_1 relaxation is very important to understand the effects of subsequent pulse application experiments. If the delay time between two pulses is smaller than $3T_1$, which normally is the case, it is insufficient for the tissue to recover all its initial z magnetization. Therefore following pulses will result in considerably less net absorption and less measurable signal. T_1 depends on the type of tissue and its composition of fat and water but also on certain extrinsic factors. T_1 can be measured using the inversion recovery experiment [HBTV99]. Typical values of T_1 for biological tissues are typically in the range of seconds.

Spin Spin Relaxation

Spin spin relaxation or transversal relaxation is an entropy driven process in contrast to longitudinal relaxation. Entropy is a measure of disorder of a system and according to the second fundamental law of thermodynamics, entropy in a closed system increases steadily towards a maximum value. After the application of the pulse and the rotation of net magnetization into the xy-plane all precessing protons are in phase, which corresponds to a high state of order. As time elapses and the protons precess according to their Larmor frequencies this coherence of phases gradually gets lost as the state of high order decreases. Eventually, all protons will return to be out of phase. This state causes zero net magnetization because the vectors in the transversal xy-plane sum up to 0.

$$M_{xy}(t) = M_{xy0} \cdot e^{-\frac{t}{T_2}} \tag{2.7}$$

where M_{xy} is the net magnetization in the xy-plane immediately after the application of the pulse and T_2 the time after which the net magnetization has decreased to $1/e$ of its initial value.

The value of T_2 also depends on the type of tissue and additional extrinsic factors. Typically, liquids have small values of T_2, while solids exhibit long T_2 times. This makes liquid structures appear bright in T_2 weighted images.

In reality the decay of the signal is stronger than only the thermodynamic effect T_2 could explain. There is also a loss of transversal magnetization due to external effects like the inhomogeneity of the basic magnetic field caused by the probe. Both effects are combined to a constant T_2^*.

T_2 and T_2^* can be measured using the spin echo experiment [HBTV99]. T_2 values are typically in the range of tens of milliseconds.

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'} \tag{2.8}$$

Contrast Agents

A boost in signal intensity can be achieved if contrast agents are applied to a patient being examined. Usually, organic gadolinium (Gd) complexes and compounds are injected. These agents have T_1 and often also T_2 shortening effects, resulting in a massive increase in signal due to the shortened relaxation times. This makes organs or vessels containing the contrast agent "stand out" against other structures that do no contain these agents [HBTV99].

2.3.2 Spatial Encoding

Magnetic Gradients

The former sections described a lot about nuclei, spins and their interaction with magnetic fields. In order to obtain images, information about the spatial distribution of the nuclei has to be inferred. Spatial information can be deduced from the Larmor precessing frequencies by using a magnetic gradient G as proposed by Lauterbur [Lau73] and Mansfield [MK73]. Formally, the gradient is a magnetic field varying over space. It is preferable to use linear gradients.

$$G = \frac{\partial B}{\partial x} \tag{2.9}$$

where G is the gradient, B the magnetic field strength and x a measure in a spatial dimension. The application of a gradient results in a spatial dependence of the external magnetic field and of the Larmor frequency of the precessing protons. Figure 2.3 a) shows the principle of spatial encoding.

$$\omega_L(x_i) = \gamma(B_0 + G \cdot x_i) \tag{2.10}$$

$\omega_L(x_i)$ is the Larmor frequency of a proton at position x_i, B_0 the basic magnetic field and G the gradient applied.

Equation 2.10 states that each proton resonates at a unique frequency which is defined by its spatial position within the gradient field.

2.3.3 Effective Spin Density and k-Space Formalism

As described before, the MR signal is measured as the electromotive force induced by precessing net magnetization. The image is obtained by converting the MR signal s to a spatial spin density $\rho(\vec{r})$, where $\vec{r}$ is a spatial position within the probe. It can be shown that the following relation holds [HBTV99]:

$$s(\vec{k}) \propto \int d^3\vec{r}\rho(\vec{r})e^{-i2\pi\vec{k}\vec{r}} \tag{2.11}$$

where $\vec{k}(t)$ is a function of time and often called k-space. $\vec{k}$ itself is defined in the following way as the value to which the integral of all gradients over time sums up:

$$\vec{k}(t) = \gamma \int_0^t d\tau G(\tau) \tag{2.12}$$

The effect of the temporal variation of G accounts for a phase difference in the precessing protons depending on their positions. This becomes clearer if a look at figure 2.3 b) is taken, which presents a 2D MR imaging example.

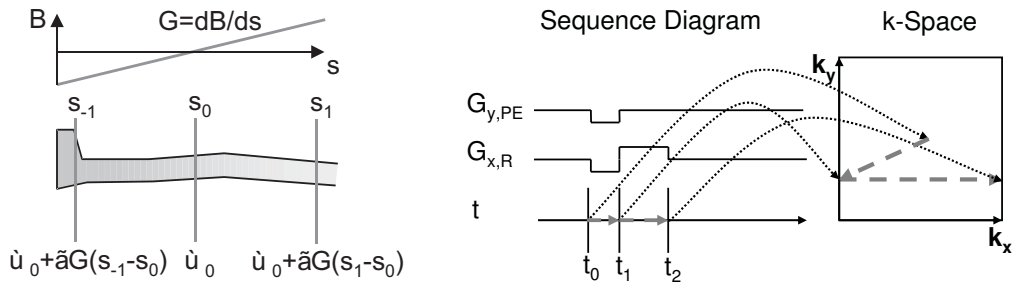


Figure 2.3: a) shows how magnetic gradients contribute to spatial resolution. The magnetic field with a constant gradient G varies over the spatial domain s. The Larmor frequencies at positions of higher magnetic field strength are slightly higher than at positions of lower magnetic field strengths.
b) shows the relation of gradients over time and the respective positions in k-space. The corresponding points are connected by dotted arrows. The position at time t_i is defined by the integral over all gradients applied for $t < t_i$. This yields the trajectory indicated by the dashed grey arrows. Between t_0 and t_1 negative x and negative y gradients are applied, followed by a positive x gradient.

Two things have to be kept in mind from the latter two equations: First, the signal $s(\vec{k})$ is the Fourier transform of the spin density of the sample. This means that after collecting all sample data from k-space, one can reconstruct $\rho(\vec{r})$ from $s(\vec{k})$ by an inverse Fourier transform as follows:

$$\rho(\vec{r}) \propto \int d\vec{k}s(\vec{k})e^{+i2\pi\vec{k}\vec{r}} \tag{2.13}$$

The second thing to memorize is that the gradients are used to navigate through k-space according to equation 2.12 where G denotes speed and direction. For being at a certain spot $\vec{k}'(t')$ at time t' it has to be made sure that the gradients $G(t)$ for $t < t'$ are selected in a suitable way such that one ends up in $\vec{k}'$. Usually, the gradients are selected in a way that ensures a desired trajectory in k-space. Mostly, the trajectories are defined to cover the portion of k-space which is necessary to reconstruct an image.

In a standard 2D MR imaging sequence, the gradients can be separated into 3 components: a slice selection gradient, that ensures that during excitation only nuclei with a certain range of Larmor frequencies are excited, a readout gradient, which is applied while data are sampled, and the phase-encoding gradient which is applied between each two readout steps. The directions which correspond to the latter two gradients are also referred to as the readout direction and the phase-encoding direction.

2.4 Sequence Families

A temporal succession of gradients, pulses and readouts, as described before, is called a "pulse sequence" or shorter "sequence" of an MR imaging experiment. Based on the primary order of RF pulses and gradient switches, the sequences can be classified into sequence families that share a common name and paradigm. The parameters of a sequence, like the times between pulses and gradients, the gradient strength, etc, are usually not specified or only limited to certain bounds or ranges. For an actual scan with a given sequence, a specific parameter setting must be defined. The combination of the sequence and a parameter setting are called an MR protocol.

The simplest sequence is the one for measuring the free induction decay (FID). Other common sequence families are the spin echo sequences and the gradient echo sequences.

2.4.1 FID sequence

The FID sequence experiment is the most simple of all sequences. It consists of just one RF pulse, followed by a sampling phase. This is repeated after a time T_R. The signal corresponds to the overlay of damped complex exponentials. The inverse Fourier transform of the signal generated by this sequence, yields the spectrum of Larmor frequencies contained in the probe.

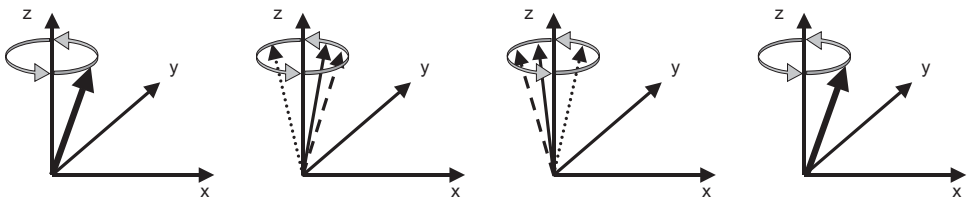
2.4.2 Spin Echo Sequence

The spin-echo sequence is one of the most important ones in MRI routine. Depending on the choice of the parameters T_E and T_R, contrast variations from T_1 to T_2 can be achieved. Most importantly, spin echo sequences are capable of producing pure T_2 weighted contrast without T_2^* effects. The spin-echo sequence consists of the following phases:

1. application of a 90° pulse
2. short delay of time $\tau = T_E/2$
3. application of the readout gradient G_x and the phase-encoding gradient G_y for a time t_g

4. application of a 180° pulse
5. second delay of time $\tau = T_E/2$
6. application of the readout gradient G_x and sampling for a time t_g

The effect of this pulse sequence experiment is the following (see figure 2.4): by application of the 90° pulse net magnetization is rotated from the z-axis into the xy-plane. At this moment all precessing protons are in phase as shown in figure 2.4 a). During the following delay time, protons lose their phase coherence because of spin spin interactions of local field inhomogeneities. Fast precessing protons rush ahead, slow precessing ones lag behind. After the delay $T_E/2$ each proton i has a phase of $\omega_{Li}T_E/2$ (figure 2.4 b)). The application of the 180° pulse inverts this phase to $\phi - \omega_{Li}T_E/2$, which means that now the slowest protons are in front and the fastest behind, as in figure 2.4 c). Allowing $T_E/2$ to elapse again, causes all protons to focus again at the phase of ϕ, as figure 2.4 d) shows. This is called the echo of the original signal.



a) b) c) d)

Figure 2.4:
a) after phase 1
b) after phase 2, fast precessing protons dotted, slow ones dashed
c) after phase 3
d) after another elapse of $T_E/2$ the echo has formed

2.4.3 Gradient Echo Sequence

The gradient echo is another important sequence in clinical practice. First, there is an initial pulse, that can be a 90° pulse but the flip angle can also be smaller. This is followed by a phase-encoding gradient and a defocusing gradient, making nuclei at higher gradient fields rush ahead, while those at the smaller gradient field lag behind. Instead of refocusing the spins by a 180° pulse as in the spin echo sequence, in a gradient echo sequence this is done by first applying an inverse re-phasing gradient. Spins that have rushed ahead during the de-phasing phase are then refocused. So the entire sequence can be characterized in the following way:

1. application of a 90° pulse
2. application of the de-phasing readout gradient $-G_x$ and the phase-encoding gradient G_y for a time $\frac{\tau}{2}$

3. application of the re-phasing readout gradient G_x and sampling for a time τ

2.4.4 Contrast, Resolution, SNR and the Relations

In order to obtain an optimal result from an MR examination a multitude of parameters of the MR imaging sequences must be considered. Since standardization in contrast to CT is difficult, there no set of parameters that produces the "best" images. Hence, choosing the optimal parameters for the acquisition of an MR image is a difficult task that strongly depends on the requirements. It turns out that there are three major objectives in an MR exam. The three objectives cannot be optimized simultaneously and they are not independent but correlated. So for each examination a tradeoff has to be made that yields the best result for the given setting. The three major objectives are: Maximizing the signal-to-noise-ratio (SNR), maximizing the resolution and minimizing the scan time. Figure 2.6 shows the triangle with the three objectives.

For an emergency patient, for example, a fast imaging sequence would be preferred, while more noise, respectively less resolution would be acceptable. For tumor screening however, the highest resolution is desirable while scan time is of minor importance.

As a rule of thumb, for a fixed magnetic field strength B_0, the following relations hold:

$$SNR \propto V_{Voxel} \propto Resolution^{-3} \tag{2.14}$$

where V_{Voxel} is the volume of one voxel. The power of three applies to isotropic 3D acquisitions. In other cases, the dimensions can be treated accordingly.

$$SNR \propto \sqrt{T_{Scan}} \tag{2.15}$$

where the T_{Scan} is the total time of scanning consisting of the complete time, during which data are sampled. The scan time is directly proportional to the number of sequence cycle repetitions and to the repetition time of each cycle during which data are sampled.

Scan Time

Scan time is an important objective in MR imaging for two reasons. On the one hand, the longer the scan time, the more discomfort is imposed on the patient, which may in turn affect the patient's willingness and ability for cooperation. On the other hand, longer scan times mean longer examination slots and less scanner utilization. Increasing the utilization of a scanner is an important issue, if the huge costs for investment and running expenses of an MR imaging device are considered.

As mentioned before, the scan time is mainly determined by the number of sequence repetition cycles, which is given by the number of phase encoding steps resp. resolution in phase encoding direction, and the repetition time T_R of the cycles themselves. The number of sequence repetitions is the product of the number of slices N_Z, the number of phase encoding steps N_{PE} and the number of averages N_{avg} per cycle.

$$T_{Scan} = N_Z N_{PE} N_{avg} T_R \tag{2.16}$$

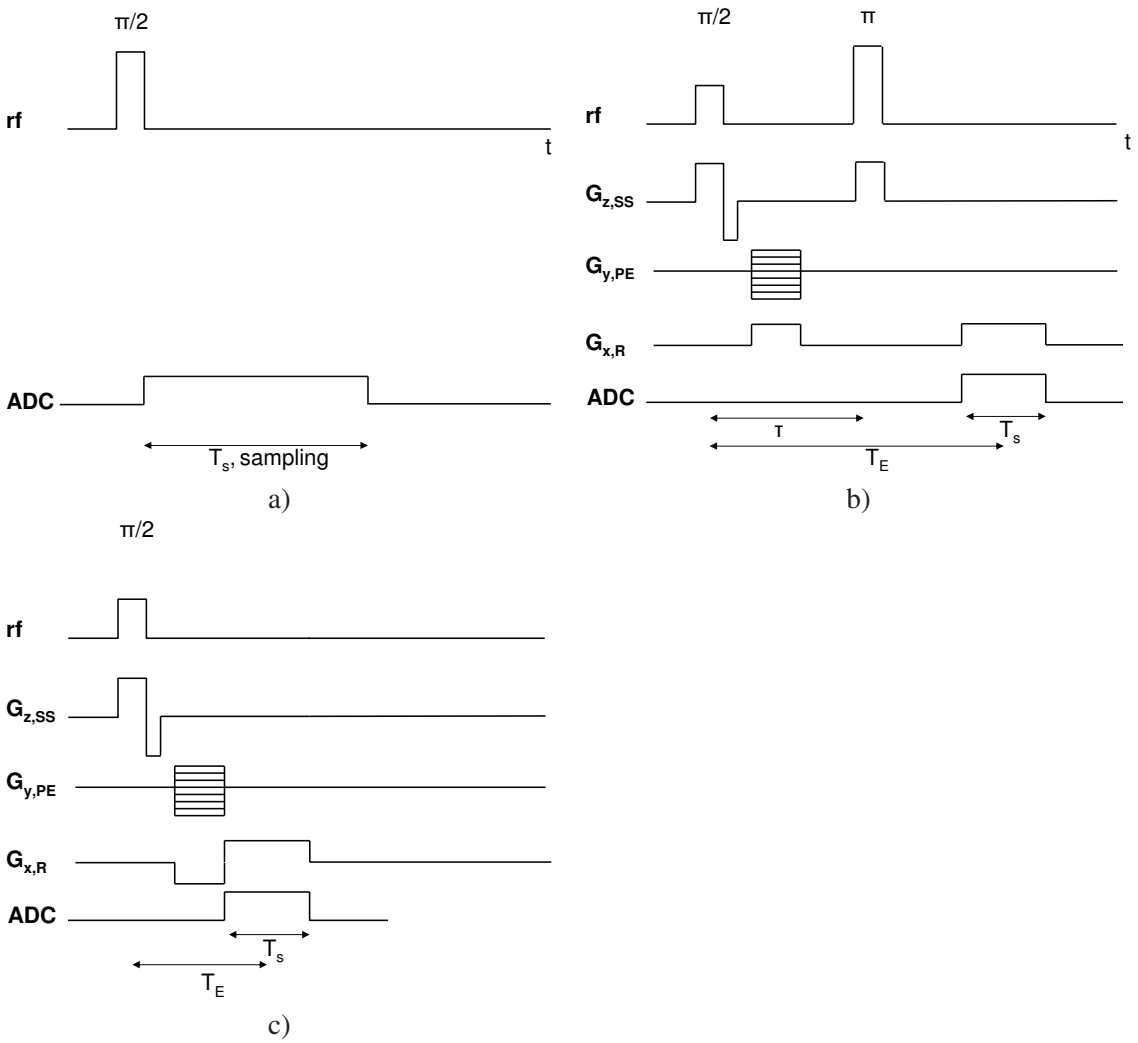


Figure 2.5: Sequence diagrams of a) the FID sequence experiment, b) a 2D spin echo experiment and c) a 2D gradient echo experiment.
The horizontal axis of the sequence diagrams shows the time domain. Sequence diagrams indicate how and at which time point gradients are switched on and off, RF pulses are applied and the sampling of values takes place. The sampling of data is also called ADC time.

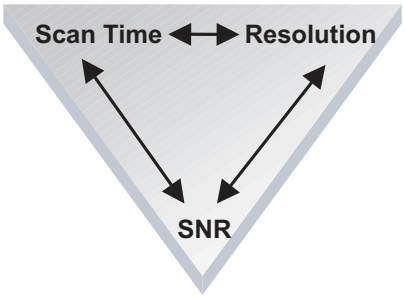


Figure 2.6: The triangle of the three major objectives in an MR exam.

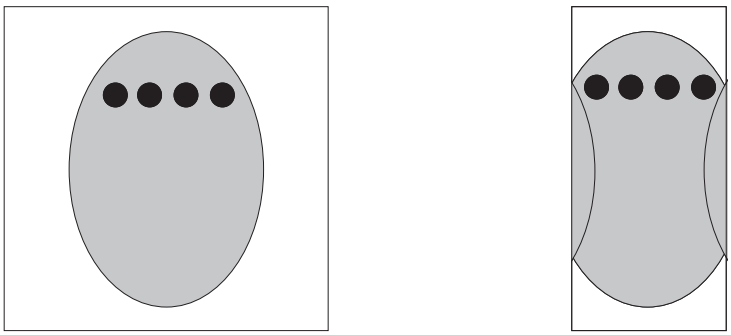
The easiest way of reducing T_{Scan} is to reduce the number of averages to 1, which causes a loss in SNR though. Another way is to decrease T_R. This is not always possible though, because T_R is determined by factors that may be beyond the operator's influence, like relaxation times, or a fixed time for T_E needed for a certain contrast, etc.

A huge number of approaches for decreasing scan time focuses on reducing the number of phase encoding steps N_{PE}. This can either be achieved by redesigning the sequences in a way that less phase encoding lines are needed. Multi echo sequences like turbo spin echo for example differ from the original sequences by the fact that multiple phase encoding steps are sampled after one excitation. In EPI sequences or HASTE sequences, all phase encoding steps are done after a single excitation [HBTV99]. As a general rule, however, the more fast imaging techniques are used, the higher the probability of image artifacts.

Other approaches like steady-state sequences or FLASH techniques have also reduced scanning time significantly but are based on different premises. A powerful concept for fast MR imaging that emerged in recent years is parallel imaging. The key idea in parallel imaging is to reduce the number of phase encoding steps by omitting a fraction of them. Missing information can for instance be recovered by collecting the data using multiple receiver coils and knowledge of their sensitivity profiles. A more detailed description of parallel imaging can be found in section 2.5.4.

Resolution

From equation 2.14 one can conclude that resolution has a strong influence on the SNR ratio. However, if resolution is smaller than a threshold, small objects may not be distinguishable from the surrounding any more. To be able to detect tiny lesions, the resolution of the image should exceed a certain threshold. This can be best explained by illustrating the effects of the resolution during image reconstruction. If the resolution is chosen to be $\Delta x = \frac{L}{N}$, where L is the field of view (FoV) of the image and N the number of samples in a given dimension, the following effect on the image quality can be deduced: As the samples are taken in the inverse k-space domain, this transforms to $\Delta x = \frac{1}{N\Delta k} = \frac{1}{W}$. W is called the sampling window. It corresponds to the range of frequencies in the k-space domain that are taken into account for the image reconstruction. This is equivalent to a multiplication with a boxcar or rectangular function in the frequency domain. According to the convolution theorem [HBTV99], in the



a) b)

Figure 2.7: This figure shows a schematic example of aliasing. In a) the FoV is larger than the object scanned. In b) the FoV is smaller. Aliasing occurs by a wrap around of the outer parts of the image folding in on the other side.

spatial image domain, this results in a convolution of the image data with the point spread function of the boxcar. A point spread function is defined as the response of an imaging system to a point source or point object [HBTV99].

The point spread function of the boxcar is its corresponding Fourier pair, the sinc function $\frac{sin(x)}{x}$. The period of the point spread function decreases with increasing resolution, so the image appears to be less blurred, if the sampling window W resp. N are large.

A different reason for maintaining a minimum resolution arises from the fact, that the samples measured in k-space with a resolution of Δk between each two of them are the Fourier transform of the actual spin densities $\rho(x)$. Due to the periodicity of the discrete inverse Fourier transform, it turns out that the following spin densities are equal [HBTV99]:

$$\rho(x) = \rho(x) + \frac{1}{\Delta k} \tag{2.17}$$

$\frac{1}{\Delta k} = L$ is the FoV. This measure delineates the aperture of the object being scanned in which the spin densities can be assigned to different locations. If now for some reason L is chosen to be smaller than the size of the object A, then consequently Δk would be larger than $\frac{1}{A}$, thus causing the period in 2.17 to become smaller than the object. This means that parts of the object outside the FoV would be wrapped around appearing on the opposite side of the image. This phenomenon is called aliasing and is caused by undersampling. Figure 2.7 shows an aliasing example caused by a FoV smaller than the object.

Another source of artifacts in MR images is, that if the sampling window W is chosen too small, big steps in intensity in the image may not be reconstructed correctly but instead an overshoot and undershoot in intensity can be observed. The effect is called Gibbs ringing [HBTV99]. It is a direct consequence of the Fourier basis which is the canonical representation of the MR data. Although the effect may be weakened if a Hanning filter with an apodizing effect is applied to the raw data, it can still be a problem.

Another important problem arising from too small resolutions and which can be observed

above all in slice selection direction is the partial volume effect. This is caused if slices are chosen so thick that they contain various tissues. The reconstructed spin density is then a linear combination of the spin densities of the separate tissues. Partial volume effects lead to blurred images with weak edges although the in-plane resolution may seem adequate. A common technique in MR imaging is to work with high in-plane resolutions and compensate for the loss in SNR by extending the slice thickness. Thus, particularly anisotropic voxels are obtained, which are prone to partial volume effects.

So obviously, there are many reasons for increasing resolution, however, as it has an influence on the SNR, a tradeoff has to be made. Increasing resolution, at least in phase encoding direction, also has a direct influence on scan time.

SNR

The signal-to-noise ratio or SNR is the most important measure concerning MR image quality. It makes a quantitative statement on how much of the scanned signal which is assigned to a voxel is caused by noise. If the SNR is not high enough, an object may become indistinguishable from the background, so that maybe a lesion cannot be distinguished from the surrounding tissue. It is evident that a minimum SNR is required for a reliable diagnosis from an MR image. A shortcoming in MRI is that the SNR cannot be determined in absolute values, since the values of the voxels have no absolute but only relative meaning. Therefore only relative changes of the SNR can be quantified.

As stated before, the SNR depends on the resolution, the imaging time and the magnetic field strength B_0. At higher field strengths, the SNR is almost proportional to B_0. This gain in SNR can be used to increase resolution, however, it has to be kept in mind that other MR imaging factors also depend on B_0 so that not all sequences show better results if they are run at higher fields. The receiver bandwidth, i.e. the strength of the readout gradient, also has an effect on the SNR ratio. For example, more noise is collected if a larger bandwidth is chosen.

Contrast

The term contrast refers to a difference in visual properties that makes an object distinguishable from others. According to Michelson [Mic27], contrast between two objects A and B is the relative or absolute difference in signal intensity between the two objects.

$$C_{AB} = S_A - S_B \tag{2.18}$$

In the case of a gradient echo experiment with a flip angle of 90° as described before, this definition converts to the following equation [HBTV99]

$$C_{AB} = S_A(T_E) - S_B(T_E) = \rho_{0,A}(1 - e^{-\frac{T_R}{T_{1,A}}})e^{-\frac{T_E}{T^*_{2,A}}} - \rho_{0,B}(1 - e^{-\frac{T_R}{T_{1,B}}})e^{-\frac{T_E}{T^*_{2,B}}} \tag{2.19}$$

where $\rho_{0,XY}$ is the spin density of tissue X, $T_{1,X}$ the T_1 relaxation time of tissue X and so on. The difference in signal between the two tissues is then a function of the parameters T_R and T_E. The difference in signal is directly proportional to the spin densities of the respective

tissues. The effects of the loss of signal due to spin-lattice relaxation, as described before, account for the first exponential term. The entropy effects caused by the spin-spin relaxation are responsible for the second exponential term.

Several special cases can be derived from this equation:

1. Proton Density Weighting:
   If $T_E << T^*_{2,A}$ and $T_R >> T_1$ equation 2.19 can be written like this:

$$C_{AB} = S_A(T_E) - S_B(T_E) = \rho_{0,A} - \rho_{0,B} \tag{2.20}$$

   This special case is called proton density weighting because the contrast in the image is mainly determined by differences in spin density of the different tissues. Figure 2.8 c) shows an example of a proton density weighted scan.

2. T1 Weighting:
   If only $T_E << T^*_{2,A}$ equation 2.19 can be simplified like:

$$\begin{aligned} C_{AB} &= S_A(T_E) - S_B(T_E) = \rho_{0,A}(1 - e^{-\frac{T_R}{T_{1,A}}}) - \rho_{0,B}(1 - e^{-\frac{T_R}{T_{1,B}}}) \\ &= (\rho_{0,A} - \rho_{0,B}) - (\rho_{0,A}(e^{-\frac{T_R}{T_{1,A}}}) - \rho_{0,B}(e^{-\frac{T_R}{T_{1,B}}})) \end{aligned} \tag{2.21}$$

   This contrast is called the T_1 weighting because the differences in signal are mainly determined by the different values of T_1 of the tissues A and B. See figure 2.8 a) for an example of a T_1 weighted scan.

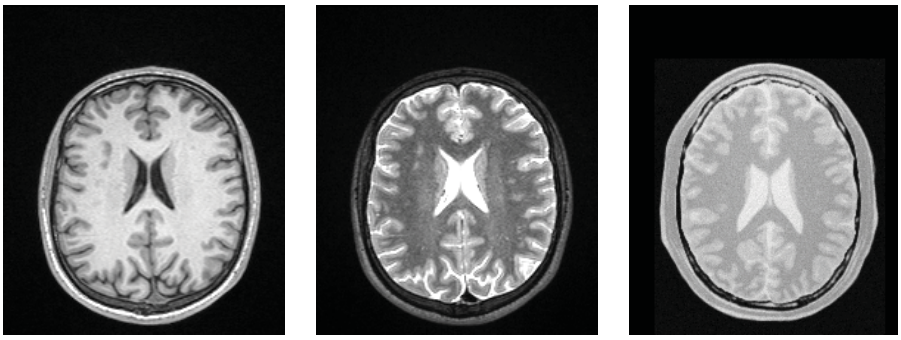
3. T2 Weighting:
   If $T_R >> T_1$ and T_E in the magnitude of T^*_2 we get T^*_2 weighting, with an example shown in figure 2.8 b):

$$C_{AB} = S_A(T_E) - S_B(T_E) = \rho_{0,A} e^{-\frac{T_E}{T^*_{2,A}}} - \rho_{0,B} e^{-\frac{T_E}{T^*_{2,B}}} \tag{2.22}$$

So, for the gradient echo sequence with flip angles smaller than $90°$ contrast is influenced by the choice of the parameters T_E and T_R. Depending on the purpose of an examination, the desired contrast can be determined. From the upper equations, also optimal parameter settings for differentiating tissue A from tissue B can be derived by computing the derivatives with respect to T_R or T_E, if the relaxation times T_1 resp. T_2 are known.

Similar observations, though not that straightforward, hold for other types of sequences. However, by choosing T_E and T_R to generate a desired contrast, a lower bound for the scan time is already defined since the latter two parameters have a direct influence on the time needed for one scan cycle.

From figure 2.8 it can be concluded that small changes in the parameter settings of the sequences, cause images to appear with completely different contrasts. Liquids like the cerebrospinal fluid for example, which fills the ventricular system of the brain, show low signal (dark spots) in T_1 weighted scans as in a) but high signal (bright spots) in T_2 weighted images like b).



a) b) c)

Figure 2.8: Similar head slices scanned at different contrast. a) shows an image of T_1 weighted MP-RAGE scan, b) is a T_2 weighted turbo spin echo and in c) a proton density weighted slice is shown.

2.5 Fast MR Imaging

After giving a general introduction into the concepts of MR image acquisition, this section is dedicated to the methods and techniques for making an acquisition fast. Since this thesis mainly deals with methods for extracting information from fast pre-scans, the focus on the MR image acquisition side is to make the pre-scans really fast, and be more generous in terms of resolution and SNR. Pre-scan images like localizers are images that are used for orientation and positioning purposes but not for diagnosis, so the described shortcomings can be accepted.

2.5.1 Multi Echo Sequences

In multi echo sequences, n phase encoding steps are sampled in k-space after the excitation of a single pulse. According to equation 2.16 this causes a reduction of scan time by the factor n. In the extreme cases of EPI or HASTE sequences, all phase encoding steps are done after a single excitation. These sequences are very fast but have shortcomings in image quality. They often show artifacts that result from a variety of reasons, like motion, etc. As a rule of thumb the more of the acceleration techniques are used, the higher the probability of imaging artifacts.

2.5.2 Half-Fourier-Techniques

The idea for Half-Fourier techniques stems from the observation of symmetry that is inherent in k-space data. It is a well known property of Fourier transformations that the transformation of a real object (which means that it has 0 content for all imaginary parts in complex representation), is conjugate symmetric with respect to the origin. This means that $s(\vec{k}) = s(-\vec{k})^*$ where the * denotes the complex conjugate value. As all scanned objects are real, in theory, it would be sufficient to measure only one half of the k-space and extrapolate the rest to the data by complex conjugation. However, as information content in the center of k-space is higher, since the values belonging to low frequencies are found there, usually a certain fraction more than half of the data (for example $\frac{6}{8}$ or $\frac{7}{8}$) is collected.

2.5.3 Steady-State Sequences, FLASH

Steady-state imaging is a technique that evolutionarized MR imaging, when conventional scanning methods reached their limits with T_R times becoming shorter and shorter and reaching the ranges of T_1 and T_2 of various tissues. The key idea in steady-state imaging is to apply a new RF pulse long before the transversal magnetization has recovered its original state. This way, only the longitudinal magnetization that has already recovered is flipped into to transversal plane by the new pulse. If this experiment is repeated various times, the transversal magnetization gradually approaches a steady state where relaxation and newly flipped magnetization are at an equilibrium. This allows for very short T_R times and thus greatly reduces acquisition times, which leverages fast imaging of organs where performance is crucial like breath hold methods of the abdomen or cardiac imaging. The approach to the steady-state depends on the relaxation times of the tissues, on the repetition time T_R and also on the flip angle θ caused by the pulse. The signal and contrast depends on all of these parameters. A perfect flip angle θ_{Ernst}, which maximizes transversal magnetization and therefore also the MR signal, can be calculated if the upper constants and parameters are known. The optimal angle is called the Ernst angle [HBTV99].

Steady-state imaging methods are classified into coherent techniques, with the abbreviation SSC, or incoherent techniques also referred to as SSI [HBTV99]. The main difference is the way in which the transversal magnetization is treated between two consecutive RF pulses. Both methods entail different magnetization responses and therefore different contrasts. While the incoherent steady-state approaches produce spin density or T_1-weighted images, the coherent state produces a T_1/T_2 weighting. Whereas in the case of coherent steady state imaging, the transversal magnetization is left untouched, in the incoherent state the transversal magnetization is destroyed or spoilt. Spoiling can either be done by applying gradients that effectively defocus the spins in the xy-plane or by applying spoiling RF pulses.

One of the first steady-state techniques and one of the most successful sequences overall was FLASH [HFM$^+$86]. This is an incoherent steady-state gradient echo sequence with gradient spoiling of the transverse magnetization. The FLASH sequence was the basis of many sequences that were used for image acquisition in this thesis.

2.5.4 Parallel Imaging

Parallel imaging has emerged during the past couple of years. The gain in scan time by a factor f is achieved by scanning only every fth line in phase encoding direction. Parallel imaging methods generally need multiple receiver coils collecting data simultaneously resp. in parallel. Since each receiver coil has a different sensitivity profile, which means that the signal from one spatial sample causes different signal intensities in each coil, missing data can be reconstructed by making use of the additional information content that is given by multiple coils and multiple sensitivity profiles

In SMASH [SM97] or GRAPPA methods, the holes caused by the missing lines in k-space are filled by using basis sets of spatial harmonics that are generated from the sensitivity profiles. This procedure allows for smooth interpolation of the missing data. In short terms, in SMASH the complete k-space data is restored prior to the Fourier transform.

In the SENSE method [PWSB99], aliased images are reconstructed and the original image is obtained by inverting the system of linear combinations given by the weightings from the sensitivity profiles. As the field of view of the acquisition remains unchanged, this procedure results in an undersampling of the data and thus overfolding respectively aliasing. This means that each voxel contains information which is a linear combination of the intensities from various spatial locations. However, by using multiple receiver coils, with each coil having a particular sensitivity profile for a given spatial location, which can be determined by a pre-scan, this overfolding can be undone. Using the sensitivity profiles, the linear coefficients of the aliased image can be determined and the original non-aliased image can be reconstructed. Figure 2.9 shows the single steps of an acquisition sped up by SENSE.

As a matter of fact, the number of coils in phase encoding direction must exceed the factor f of reduction of the number of phase encoding steps.

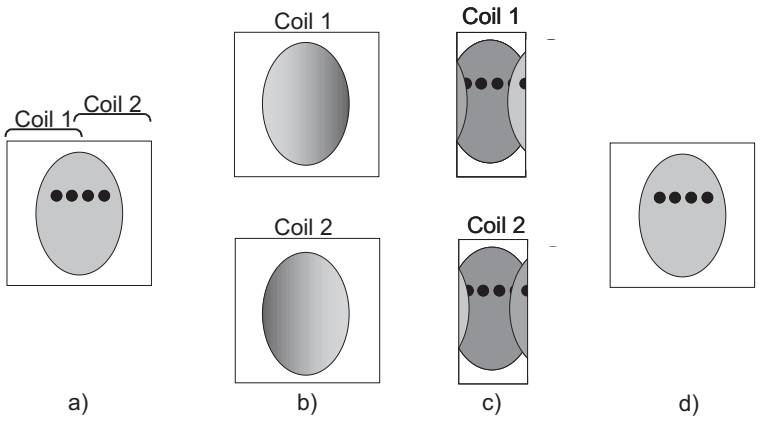


Figure 2.9: This figure shows a schematic example of parallel imaging using SENSE. a) shows the schematic phantom that is scanned and the placement of the multiple coils in readout direction. b) shows the sensitivity profiles of the two coils depending on their positions relative to the object. In c) the images are shown which have been independently reconstructed from each coil, Aliasing due to the undersampling can be observed. d) shows the resulting image which is reconstructed by solving the linear system of the aliased images and the coil sensitivity profiles.

Unfortunately, it turns out that the SNR ratio also drops by a factor of $\sqrt{f}$.

Still both methods produce good image quality. The crucial point in both methods is the determination of the coil sensitivity profiles for arbitrary coil geometries. Since the sensitivity varies with coil load, the profiles must be determined prior to each scan in an autocalibration step. A shortcoming of parallel imaging with SENSE is that one has to make sure that all aliasing effects are caused by the reduction of phase encoding steps and not by a field of view that is chosen too small because otherwise the combined effects would cause strange image artifacts. Figure 2.10 shows the differences between SMASH and SENSE.

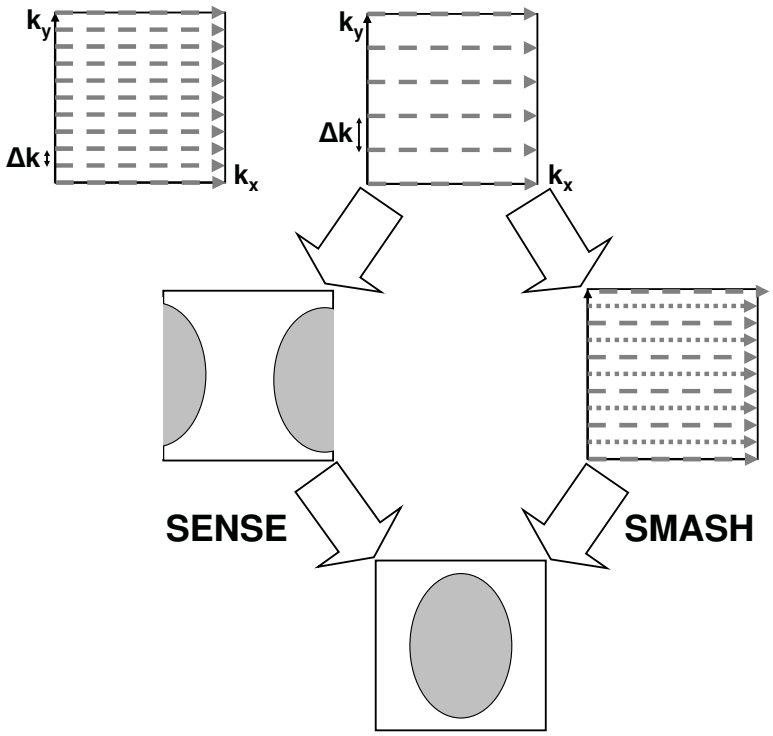


Figure 2.10: This figure shows the difference between parallel imaging using SMASH and SENSE. As the number of lines in readout direction is reduced, Δk increases. In SMASH methods the missing lines are reconstructed using spatial harmonics generated from the coil sensitivity profiles, so missing data are filled up in k-space. In SENSE methods aliased images are reconstructed which can be correctly unfolded to the original image of the oval phantom if the coil sensitivity profiles are known. In this case the missing data are reconstructed in image space.

2.5.5 Continuous Table Imaging

Developments like parallel imaging lead to a drastic decrease of scan time, which allows for performing a whole-body MR scan, as it is required for tumor screening for example, within 45 minutes. However, as the maximum FoV in MR scanners that are commercially available is about 50 cm because of limitations of the homogeneity of the magnet, scanning a complete body that measures about 1.80m has to be done in at least four or five stages. The number of stages is determined by the effective FoV that can be used for each scan and the overlaps that have to be considered to ensure complete coverage and to allow for a robust stitching and composing of the images acquired at the single stages.

The idea of continuous table imaging is to increase the FoV in transversal direction by moving the table and the patient during the scan. This was implemented by various groups, for

example by [KRGR02]. Doing a single scan instead of multiple scans at different stages has various advantages [Bla06]:

- The lacking need of repositioning reduces the source of motion artifacts
- As the data are not scanned in stages, no stitching and composing of the images is required, which simplifies the processing workflow.
- As all slices are scanned when they are located at the isocenter of the magnet, the scan can benefit from optimal homogeneity of the magnetic field and therefore provide optimal image quality.
- Furthermore there is a gain in scanning time in comparison to a multistage scan. The saved time results from the lacking need for repositioning and scanning of overlaps.

All the advantages mostly pay off in MR angiography scans, which was the main driver for the development of continuous table imaging. A slight shortcoming of this technique is that it is only feasible for very fast sequences, that are not affected by the continuous motion of the table, like coherent steady-state, EPI or HASTE sequences.

2.5.6 Artifacts in MR Imaging

In MR imaging there are many sources for image artifacts, i.e. image may show features which are not part of the anatomy but arise from shortcomings of the MR imaging technique. Some of them can be compensated for easily, some are tricky or unavoidable. As mentioned before, as a rule of thumb the following relation holds: the more acceleration techniques are used during the image acquisition, the higher the probability of artifacts in the image. This section summarizes some of the most important sources of artifacts.

Partial Volume Effects

Partial volume effects [HBTV99] arise from the fact that a voxel is so large that it contains two types of tissue, resp. its content is a linear combination of tissues. If e.g. a voxel v at the interface of tissue A and tissue B contains a fraction α of tissue A and a fraction of $1-\alpha$ of tissue B, then the total signal in this voxel will also be the same linear combination of signals $s(A)$ and $s(B)$.

$$s(v) = \alpha s(A) + (1-\alpha)s(B) \tag{2.23}$$

These partial volume effects reduce the contrast between two voxels.

Gibbs Ringing

Gibbs ringing occurs at interfaces between objects of high intensity and objects of low intensity, e.g. between the lung and surrounding tissue. An overshoot in the tissue of higher intensity and an undershoot in the tissue of low intensity can be observed. This is a direct consequence of

the signal being represented by a limited Fourier series, which restricts the possibility of representing high frequencies which appear at discontinuities as sharp edges. This kind of problem can be solved by increasing the number of data samples, which allows the representation of higher frequencies or by using apodizing filters like the Hanning filter [HBTV99], which is applied to the raw data and avoids overshoots to a certain extent.

Motion Artifacts

MR raw data is usually collected in chunks, i.e. that k-space data is filled by repeatedly sampling small fractions, e.g. in various phase encoding steps. The image is reconstructed after all fractions have been collected. The underlying assumption of this procedure is, that the data sampled this way are consistent, which means that the object does not move during the time of acquisition between any of the data chunks. If there is movement between two lines in k-space, this will have an effect on the reconstructed image. There are several artifacts which may be caused by motion.

Translational motion of the object being scanned will result in a general blur and a general loss of signal. Other artifacts may also occur depending on the type of motion and image reconstruction. The main reason for the signal loss is that due to the motion the phase of the time of the echo is not 0. For periodic motion with a period of the magnitude of the T_R of the acquisition, so called 'ghosting' occurs. Such motion can be breathing motion for example. These ghosts are periodic repetitions in phase encoding direction of the high-frequency content of the original image overlayed to the original image. This results in partial signal enhancement and partial signal nulling.

Bias Fields

MR images often show inhomogeneity of the grey values, which make equal tissues appear at different grey values. This inhomogeneity is caused by the spatially varying signal response in the coil that receives the MR signal. What is visible by the reconstruction is not the original signal of precessing magnetization $M(x)$ but a multiplicative overlay of the underlying true magnetic signal with a local gain field of the receiver coil. So, if $I(x)$ is the intensity of the image at voxel x, then

$$I(x) = M(x) \cdot B(x) \tag{2.24}$$

where $B(x)$ is the local gain at the voxel x or the so-called bias.

Several approaches exist for removing the effects of bias fields. Most of them, like homeomorphic filtering [GW02], rely on a high-pass filtering approach after taking the logarithm of the image, which converts the multiplicative bias into an additive bias. The additive bias is virtually removed by the high-pass filtering under the assumption that the spatial variance of B is small.

Fat/Water Shift

Protons bound in fat molecules resonate at a different frequency compared to water-bound protons. Because of the different chemical surroundings, protons in these two molecules expe-

rience different shielding from the basic magnetic field [HBTV99]. So the effective magnetic field B_{eff} is:

$$B_{eff} = B_0 \cdot (1 - \sigma) \tag{2.25}$$

with σ being the shift constant.

According to equation 2.3, this leads to a difference Δf in resonance frequencies of the protons. If for a certain scan protocol, the receiver bandwidth (resp. the frequency spread of a voxel) is smaller than $\frac{\Delta f}{2}$, parts of the fat signal caused by the fat content of a voxel can become misregistered to a neighbor voxel, which can lead to artifacts.

Gradient Distortion

In practice, the magnetic gradients, as described in section 2.3.2 are not perfectly linear. As a matter of fact, the non-linearity increases with distance from the magnetic bore. This non-linearity accounts for image distortion, as voxels in the non-linear tend to become stretched or compressed. By knowledge of the nature of the non-linearity, these effects can be compensated for by distortion correction algorithms, like [LDL99]. For morphological studies from MR images, as it is the case for anatomical modeling, this correction is essential.

3

Liver Segmentation from MR Image Slices Using An Active Shape Model

3.1 Introduction

3.1.1 Motivation

In recent years abdominal MR imaging and MR liver imaging have entered clinical routine, particularly with multibreathhold and free-breathing methods using navigators [RE89] for breathing motion compensation becoming available. Navigators allow for prospective motion compensation by taking special "navigator" echo images interleaved with the original sequence. The phase of the breathing cycle can be obtained from the navigator echoes. Positioning of the slices for a liver examination and of the navigator for breathing motion compensation is usually done on a set of stacked 2D localizer images, which are scanned using a fast prescan protocol before the examination. Large fields of view and fast acquisition time but moderate image quality and resolution are characteristic of these localizer images. There is also only sparse coverage of the organ of interest by the sliced localizers. Positioning based on these images can be tedious as the user working on 2D localizer images tends to forget to ensure a correct placement of the objects in all three dimensions. This chapter describes an algorithm for reconstructing position, shape and orientation of the liver from the series of stacked 2D localizer images. Knowing the position and the shape of the liver leverages an automatic, well-defined and reproducible placement of slices, navigators but also entails a lot of possible follow-up applications. The position and shape information of the reconstructed surface can be used as an initialization for following detailed organ segmentations. It may also be helpful to applications designed for assisting the image reading process by a physician, e.g. by providing a basis for functionality like automatically assigning findings and lesions to the correct organ regions resp. regions of the image or searching for corresponding findings in follow-up examinations.

The problem of reconstructing the position and the shape of the liver is closely related to the segmentation problem. Liver segmentation is an active field of research, with a lot of publications and a variety of algorithmic approaches. Most promising and best results are obtained by methods that incorporate prior knowledge and statistics by means of active shape models. Given the particular difficulties of MR imaging and the additional quest of reconstructing a 3D shape from sparsely covering 2D slices, which is an ill-posed problem, a model based approach is considered to be the best choice.

3.1.2 Model Based Approaches in Medical Imaging

The general goal of medical image analysis is to extract structural information from the medical images that may be used for interpretation or analysis by experts. As medical images and MR images even more tend to exhibit a large variety of quality in terms of noise and contrast, this task is exceptionally hard. Moreover, the structures in the images may be incomplete or show artifacts. This interferes with the huge variability of anatomical structures across the population and results in an enormous number of possible cases which are to be expected.

Model based approaches for image analysis make an underlying assumption of the structures that are expected in the images. Prior knowledge of the structures is incorporated into a model and algorithms searching for certain structures can make use this knowledge, limit the search spaces and thus decrease running times and reduce the likelihood of failures. This approach can be characterized as 'top-down' image analysis [CT99]. This means that with a clear concept or model of the structures to be found, the image is browsed for possible occurrences, while the standard or 'bottom-up' [CT99] approach would be to cluster the image by grouping local structures to objects which can be further examined if they fit the class one is looking for. In general, in medical images and in MR images even the more, the latter approach is harder and more error prone.

Hence, model based approaches leverage algorithms which are based on 'analysis by synthesis' approaches, which means the for the analysis of an image, first an instance of a reference of known structure, a model, is 'synthesized' and then compared to the actual image data. By maximizing the similarity between the model and the data, information on the image structures may be deduced.

A description of related work on model based medical image segmentation starts with Staib and Duncan [SD92] who proposed model based image segmentation using Fourier descriptors of closed 2D curves for the representation of the variability of the underlying object shapes, which are to be segmented from the image data. Due to the nature of Fourier descriptors an extension of this method to 3D objects and image data is not straightforward. Yuille *et al.* [YHC92] used deformable templates for a complex eye model which is used for face segmentation. Although being effective for the described problem, the modeling process is complex and requires a lot of work for each new object. Active contour models or snakes as proposed by Kass *et al.* [KWT88] are easy to implement in 3D but, in general, only consider image data and smoothness conditions but no anatomical knowledge for the segmentations.

The quality of a model can be classified by the two terms generalizability and specifity. As stated above, the model must be able to deal with a huge amount of variation. It is therefore essential that it captures a large amount of natural variation. This ability can be described by the term generalizability. There is a tradeoff to be made with the specifity of the model. On the one hand, the model should be able to represent all possible instances of a natural class. On the other hand, it should be specific enough to contain only valid instances of the structure. In data driven approaches, generating a model requires a training step, during which typical instances are presented to a learning algorithm. During the learning phase, the generalizability and specifity must be weighted against each other by a careful design of the learning algorithm.

If specifity dominates over generalizabilty too much, an effect which is called overfitting, can be observed: The model may be capable of representing all instances of the training data set but fail the generalize to instances which were not presented during the training phase. This has to be avoided by adapting the conditions of the learning algorithm.

3.1.3 Active Shape Models

Active shape models were proposed by Tim Cootes [CT92] as an extension to active contour models or snakes [KWT88] to overcome the shortcoming of not being able to consider prior anatomical knowledge in segmentations. Both active contour models and active shape models are techniques for model based image segmentation in the sense of section 3.1.2. Active contour models or snakes are generally based exclusively on image data and smoothness of their contours, although several extensions exist. Active shape models, however, contain statistical information on the geometry and structure of the objects that are to be segmented from sets of corresponding landmark points representing the contour of the object. From the corresponding landmark points, information about their statistical distributions can be inferred. This information is extracted from example data sets in a training step. Knowing about the statistical distribution of the shapes, valid instances and their corresponding probabilities can be calculated. The statistical knowledge may direct a guided active shape model search towards the most probable instance of the model.

Usually, the shapes or contours are represented by corresponding landmark points [CT92], which may be the support points of any interpolating or approximating surface or curve representation or the vertices of a discrete polygonal structure, etc. The statistics of the active shape model are calculated by first assigning all elementary data elements, i.e. coordinates of the landmarks, to a single vector. For instance, all items of all the dimensions of the corresponding landmark points l_j are assembled to one large vector.

$$x_{dim \cdot j+i} = l_{ji} \tag{3.1}$$

where x is the assembled vector, dim the dimensionality of the landmark points and l_j is the jth landmark point with l_{ji} being its component in the ith dimension.

Usually, the shapes are normalized by geometrical transformations to make sure that the statistical analysis includes only the actual variations of shape and not the variations imposed by translations, rotations, scales, etc. The geometrical transformation for the normalization may be rigid-body, similarity or affine transformations [CT99], depending on the application that the model is designed for. So usually the original landmarks l'_j are normalized by a transformation T before assembling them to a vector, i.e. $l_j \leftarrow T(l'_j)$. Algorithms for finding the normalization transformations between point sets with known correspondences are described in A.2 and A.3.

Statistical analysis of the data is done by principal component analysis (PCA) of the covariance matrix. PCA reveals the major modes of variation in descending order by computing an orthogonal system of eigenvectors and eigenvalues of the covariance matrix S of the assembled vectors x_i. A more efficient way of finding the principal components and the principal values is described in A.1. Computing the covariance matrix from a set of input data is well known.

$$S = \frac{1}{n-1} \sum_i (x_i - \bar{x}) \cdot (x_i - \bar{x})^T \tag{3.2}$$

S is the covariance matrix and $\bar{x}$ the mean value of all input data vectors x_i.

Computing the PCA from S corresponds to a transformation of the principal axes towards a coordinate system with minimal correlation. As the data of geometrical coordinates make up a point cloud in n-dimensional space, PCA computes an orthogonal system of main axes spanning the extent of the data cloud. The importance of the principal components decreases with increasing index, so by removing the last $n - m$ components, a reduction to a dimensionality of m can be achieved, while the variation covered by the m first principal components is maximal. An instance x of the model can then be created by a linear combination of its m principal components, which yields a linear model:

$$x = \bar{x} + \Phi \cdot b \tag{3.3}$$

where $\bar{x}$ is the mean instance, Φ is a matrix with the principal components aligned as columns and b is the feature vector containing the weights for the linear combination of the principal components.

As Φ is orthonormal, for a given instance x, the feature values b can be computed simply by

$$b = \Phi^T \cdot (x - \bar{x}) \tag{3.4}$$

For a sufficiently large number of samples, the data point cloud can be assumed to be a multi-variate Gaussian with the m principal components being the main directions of variation and the principal values resp. eigen values λ_i the variances for these directions. Having computed b, the probability $p(b)$ resp. its logarithm $\log p(b)$ of x being an instance of the model can be derived from the multivariate Gaussian distributions:

$$\log p(b) = \sum_i \left(-0.5 \cdot \frac{b_i^2}{\lambda_i} \right) \tag{3.5}$$

where b_i are the feature values of the instance and λ_i the principal values resp. variations of the principal components.

Cootes later extended active shape models to active appearance models [CET98]. Active appearance models also include statistics on the grey value distributions of the image data into the active shape model by sampling grey values at corresponding points in the image and adding them to the assembled data vector x. Hence, correlations between the grey value statistics and the shape statistics are established by PCA, which may lead to faster convergence of the segmentation algorithm using the active appearance model. In order to make sure that grey values of correct correspondence sample points are compared, the images have to be rewarped to the mean geometry, so that appearance samples can be taken in coordinates of the mean geometry. However, the assumption of having correlation between shape and grey values is not always valid or desired, and even may introduce biases to the model, so often active shape models are a better choice.

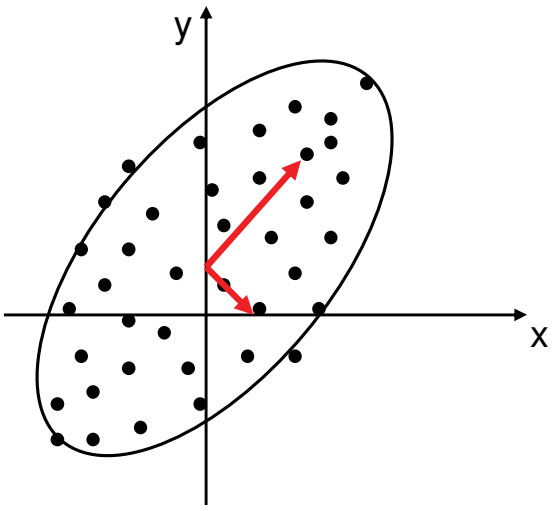


Figure 3.1: This figure shows the principle of principal component analysis. The data cloud is represented by the black dots. Principal component analysis computes an orthogonal system of vectors (red arrows) which represent the directions of main variation of the input data.

3.2 Related Work

This section will briefly cover related work on segmentation in medical images using active shape models or active appearance models. The first subsection will give a general overview, the second an overview of related work in the context of liver segmentation.

3.2.1 Statistical Models and Active Shape Models in Medical Imaging

Ever since the first publication by Cootes [CT92] active shape models have enjoyed large popularity in medical image analysis. Originating from the purpose of face segmentation, the huge potential of active shape models was extensively used in the context of medical image segmentation. This section will provide a brief overview of some of the most important papers.

Solloway *et al.* [STHW96] use an active shape model of the knee for cartilage segmentation. Duta and Sonka [DS98] use active shape models in the context of segmentation of structures of the human brain and a diagnostic interpretation of the shape segmentations which they obtain. van Ginneken *et al.* [vGFS+02] build a statistical shape model of the corpus callosum and also extend the active shape model search by using a different approach for finding optimal landmark displacements. They use a kNN-classifier instead of the Mahalanobis distance. This approach was also tested successfully for lung image segmentation.

A lot of work was published on building statistical shape models and active shape models for the heart or for single chambers of the heart. Ordas *et al.*[OOL+07] build a 3D statistical shape model of the complete heart and use it for segmentation of cardiac CT data. Lötjönen *et al.* [LKS+03] create a 3D statistical model of the four heart chambers from short-axis and long-axis MR images. A 4D cardiac active shape model is presented by Fritz *et al.* in [FRDS06]

and used for segmentation purposes. An interesting publication was made by Kohlberger *et al.* [KCR+06]. They come up with a 4D statistical shape model of the left ventricle using an implicit surface representation which avoids finding explicit surface representations and correspondences.

The abundance of active shape models in medical imaging literature and the variety of applications making use of them convey an impression of their power and their vast applicability.

3.2.2 Statistical Shape Models and Active Shape Models in Liver Segmentation

Lamecker *et al.* [LLS02] create a statistical shape model of the liver from a set of CT training data. In order to find corresponding points, the liver is partitioned into specific surface patches. The 3D correspondence problem is solved by minimizing distortion between two surface patches under the constraint of user defined feature points via a surface parameterization, which preserves arc-length and minimizes distortion. They apply their model to liver segmentation [LLS04] using an active shape model approach. This means that landmark vertex profiles are inferred from the training data as image data samples and used to calculate optimal landmark displacements. The profiles are sampled from the image data which are preprocessed using anisotropic diffusion filtering [WRV98] for smoothing the profiles. The model consists of 43 data sets and is tested with 20 principal modes, which account for 95% of the variance of their training data.

This approach was further tuned to a fully automatic shape segmentation of the liver constrained by a shape model and based on a heuristic intensity model [KLL07], which won the first prize at a competition for liver segmentation during the MICCAI [1] conference in 2007 [HvGS07] and can therefore certainly be considered as the best among current state of the art liver segmentation algorithms. The heuristic intensity model uses several histogram statistics to separate liver tissue from non-liver tissue and tumor tissue. The additional shape constraint keeps the segmentation within reasonable boundaries of the statistical shape model. The authors also introduce a robust initialization heuristic based on finding the relative position to the right lobe of the lung.

Heimann *et al.* published a paper [HWM06a] on liver segmentation which is similar to Lamecker's works [LLS04] but differs in detail. For finding 3D correspondences during the creation of the liver model, Heimann uses a method based on conformal mappings, remeshing and correspondence optimization using a measure based on minimum description length. During the segmentation process, they compute optimal landmark displacements and instead of producing a rigid projection to the model, they gradually relax the model from the displaced model instance to the model instance from the projection.

Florin *et al.* [FPFLW07] present an approach for using a set of 2D liver segmentations from dedicated slices and using a shape prior imposed by an active shape model to recover the full 3D shape from few slices. They make quantitative analysis on how many and which axial slices need to be segmented in order to compute a 3D shape segmentation with a defined error boundary. Despite the similarities, Florin's approach differs from the concepts described in this chapter by the fact that only axial slices are used whereas in case given in this thesis

[1] Conference on Medical Image Computing and Computer-Assisted Intervention

also coronal slices are considered for the segmentation and the slices can be located at arbitrary relative positions of the liver depending on the initial placement of the patient. Nevertheless, the fact proven by Florin *et al.* in his paper that the 3D shape of a liver can be recovered from 5 particular slices with a certain error threshold is remarkable.

3.3 Building the Liver Model

Building the liver model is a multi-step process. This section will describe the necessary steps and processes starting from the acquisition of the images to the later fine-tuning of the resulting model for the specific purpose. Some steps are based on related work by other authors. Whenever this is the case, the references are mentioned.

3.3.1 Liver Images

The images were acquired using a 3D FLASH-VIBE sequence with fat suppression [RLL+99]. This is a T1 weighted sequence with good morphological contrast, which is necessary for a precise segmentation of liver surface from the image data. The images were scanned from 23 volunteers during breathhold to avoid motion artifacts. Most of the images were scanned with 192 slices, each one with a thickness of 3.92 mm, an in-plane resolution of 512x512 pixels and a pixel size of 1.92 mm. As the breathhold interval of 35 seconds was too long for some volunteers to maintain, some of the data sets were taken with less slices or smaller in-plane resolution. A slice of an example data sets is shown in figure 3.2 a).

3.3.2 Manual Segmentations

High-quality gold-standard or expert segmentations are required to build a statistical shape model with sufficient precision. Semi-manual segmentation by an expert with algorithmic support remains the current gold standard. For 3D segmentation from CT image data, live wire [BM97] based segmentation approaches of one slice after the other provide good results within reasonable time. As segmentation of 3D objects from MR image data is challenging, high-level algorithmic support is needed.

The segmentation tool developed for performing the semi-manual segmentations of the liver was based on the following goals:

- Limit human interaction to a minimum: The less interaction is required from the user, the less time-consuming and exhausting the segmentation will be and the more efficient and exact the results will become. Therefore, different segmentation algorithms are provided from which the user can select the one that is locally most efficient. It must be possible to easily switch from one algorithm to another even within a single slice, if necessary.

- As MR images can be locally poor in contrast, high-level algorithmic support is needed for the semiautomatic segmentation: contour-based paradigms, i.e. live wire methods [BM97] including path cooling concepts, are combined with region-based segmentation algorithms based on minimal graph cuts [BJ00], and extended for the purposes in the scope of the example segmentations of the training images.

- Constrain user interaction to 2D, but augment his input making use of 3D image information: Interaction between a human and a machine by means of a mouse device can be more efficiently done in 2D. However, the software can use additional out-of-plane information to steer the user's input towards a more consistent result in 3D. Moreover, in regions of weak contrast, additional information can be drawn from neighbor slices.

- Support visual control and correction from different angles: As MR images can be locally poor in contrast it may be difficult to distinguish certain structures given only one viewing direction. Therefore the user can work in a convenient viewing direction. The segmentation results are kept independent of this direction.

- Support segmentation of disjoint parts. It must be possible to segment different parts of an organ or internal structures independently and later merge or split the segmentations in an arbitrary way.

- Support manual and algorithmic post-processing. The final segmentation result may still contain small errors like discretization artifacts or areas that have been segmented too sloppily, which may make the contour appear to be inconsistent from a different point of view in 3D. Several methods for enhancing 3D consistency are implemented.

The mentioned goals are realized in a segmentation software tool that allows for high-quality 3D segmentations of the liver from the image data described [FTS07]. The software tool supports different segmentation workflow modes, with one of those being the live wire mode, in which the user can repeatedly click points on a contour and the algorithm traces the path on the contour between those points. The second mode consists of a graph cut based foreground/background segmentation similar to [BJ00], in which the user is required to mark areas which are surely located inside the structure and points that are clearly outside the structure. The algorithm computes the maximum discontinuity between inside and outside, which can be verified by the user. By means of iterative refinement the user can gradually improve the segmentation result. The third work mode is the ribbon mode, in which the initializations of the graph cut model are made by drawing ribbons inside and outside the contour defined by the projections of contours segmented in neighbor slices.

A description of the tool can be found in [FTS07]. Figure 3.2 shows the workflow and the outcome of a liver segmentation done by use of the tool. The final surfaces are extracted as discrete polygonal meshes from the segmentation results by means of the marching cubes algorithm [LC87].

All liver training data sets were segmented and the initial surface representations were produced by means of the segmentation software tool.

3.3.3 Correspondences

The manual segmentations in section 3.3.2 produce a set of polygonal meshes with different numbers of vertices and triangles and with different connectivities. As mentioned in 3.1.3, for the principal component analysis, the geometric coordinates of the landmark vertices are assembled to a large single data vector. For capturing the correct variation of the surface

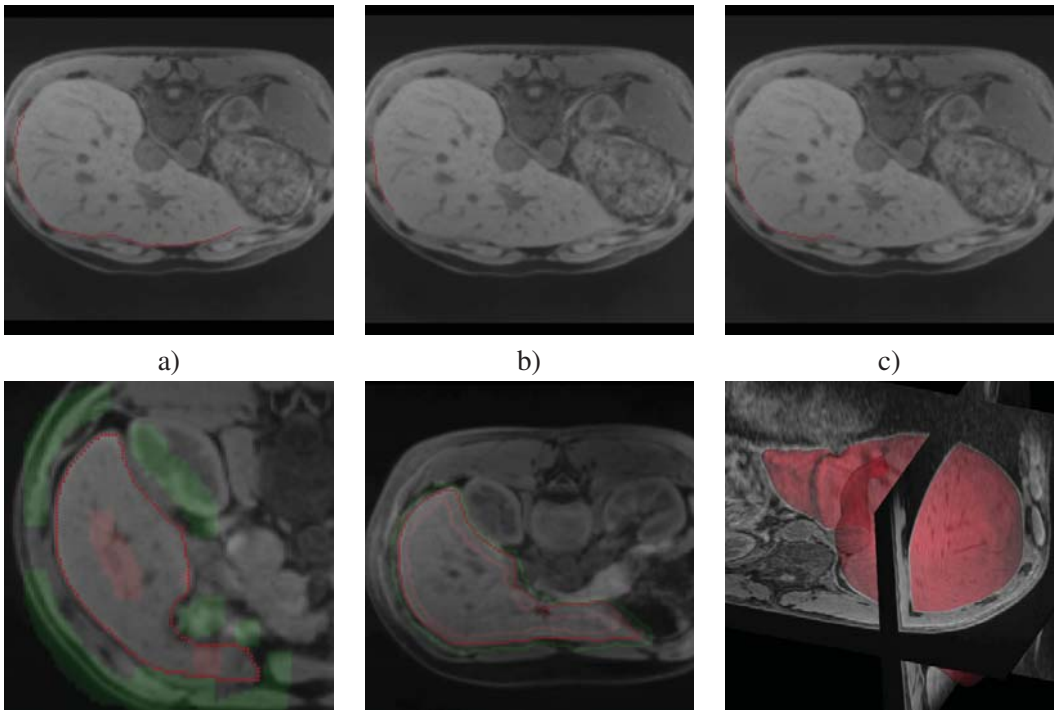


Figure 3.2: a), b) and c) show a live wire based contour following segmentation of the liver, d) shows a graph cut segmentation with the interior regions marked in red and the exterior regions marked in green, e) the initializations of a ribbon mode search and f) the resulting surface.

geometry, these landmark vertices must refer to corresponding points on the polygonal surface meshes. Finding correspondences among different surfaces is a non-trivial task. Even for a human operator it is often not evident in complex 3D structures to mark correspondences in different objects. Most promising approaches for building statistical models from medical image data are based on remeshing the surfaces using an ideal surface parameterization like Gu *et al.* [GWC+03] and optimizing the description length of the resulting active shape model as proposed by Thodberg *et al.* [Tho03]. This was successfully applied and implemented by Davies [DTC+02] and Heimann [HWWM05] and their groups.

Remeshing Using Conformal Mappings

Most anatomical structures and organs including the liver are so-called genus 0 manifolds. This means that they are homeomorphic to an object with 0 handles [PS95]. The number of handles denotes the number of holes inside the object. A sphere for example is an object with 0 handles. Consequently the liver is homeomorphic to a sphere. Being homeomorphic to another object means that it can be deformed to the other object by pure bending and stretching but without cutting or gluing [PS95].

The fact that the human liver is homeomorphic to a sphere and that it can be deformed to a spherical object leads to the idea of expressing all landmark point coordinates in spherical coordinates, i.e. one coordinate for longitude $[0..2\pi]$ and one coordinate for latitude $[0..\pi]$. Fixing one north pole point and the direction of the first longitude, as on the terrestrial globe with the north pole and the Greenwich median as references, produces a one-to-one mapping from the liver surface to spherical coordinates and vice versa. If all liver surfaces had a corresponding north pole and longitudinal degree, a common coordinate system could be defined that could be used for identifying and addressing corresponding points on the liver surfaces. Consequently, if M_i is the mapping from the liver surface S_i to spherical coordinates on the unit sphere S_0, a correspondence p_{S_j} of point p_{S_i} on surface S_j can be found by first transforming the point to the unit sphere and then back to the surface S_j.

$$p_{S_j} = M_j^{-1}(M_i(p_{S_i})) \tag{3.6}$$

Robustly identifying a north pole point and a longitude on the different shapes manually is not straightforward. However, as later the correspondences of the model are fine-tuned, a rough guess is enough for the initial parameterization.

The procedure of mapping a polygonal mesh surface to its homeomorphic primitive, the sphere in the case of genus 0 surfaces, is called a mesh parameterization. There is a survey by Alla Sheffer *et al.* [SPR06] providing a good coverage of mesh parameterization methods and applications. Mesh parameterizations can be classified into being length preserving, angle preserving or area preserving. The term preserving in this context means that the differences of this measure in the original surface and in the mapped surface should be minimal. While preservation of the triangular areas is the strongest requirement, it is sufficient for most applications including the one described in this thesis to require preservation of angles. Changes in the angles of a polygonal surface are often related to changes in quality of the triangulations of the surfaces, so keeping the angles constant corresponds to maintaining a degree of surface quality.

The method of choice for mesh parameterization is conformal mapping as proposed by Gu *et al.* [GWC+03] because it is well established in the context of parameterizing genus 0 surfaces of medical objects and because it is conceptionally simple and easy to implement. Conformal mapping is an angle preserving method for mesh parameterization. It is based on minimizing the harmonic energy of a mesh. The harmonic energy is defined as a function of the mesh connectivity C and the mapping function $\Omega(v)$, which maps the vertex positions v to points on the unit sphere. A consequence of the latter fact is that $||\Omega(v)|| = 1$ for all vertices [GWC+03]. Computing the mesh parameterization or conformal mapping Ω then boils down to minimizing the harmonic energy $E(C, \Omega)$ with respect to Ω

$$E(C, \Omega) = \sum_{[u,v] \in C} k_{u,v} ||\Omega(u) - \Omega(v)||^2 \tag{3.7}$$

where u and v are vertices of the mesh, $[u, v] \in C$ is an edge from connectivity C and $k_{u,v}$ is a weight for the edge between the vertices u and v.

Hence, Ω is computed as:

$$\Omega = \arg\min E(C, \Omega') \tag{3.8}$$

Setting the $k_{u,v}$ to 1 yields a Tuette or barycentric mapping where each vertex is located at the center of its neighbors. For the conformal mapping, the energy coefficients for the edges are given by the mean cotangent values of the two opposing angles $\alpha_{u,v}$ and $\beta_{u,v}$ of each edge.

$$k_{u,v}^{harmonic} = \frac{1}{2}(\cot \alpha_{u,v} + \cot \beta_{u,v}) \tag{3.9}$$

It can be shown that this function has a unique minimum, which can be found in a good approximation in a gradient descent approach if after each step, the vertices are projected back to the unit sphere $\Omega(v)' \leftarrow \frac{\Omega(v)}{||\Omega(v)||}$ [GWC+03].

Figure 3.3 shows an example of a simplified polygonal mesh surface of a liver and the resulting conformal mapping on a sphere.

Having the initial mesh parameterizations of all liver surfaces, and having all of them aligned by defining a north pole and a direction for the 0 longitude, an 'ideal' meshing can be designed for the surfaces on a unit sphere and transferred to all surfaces. This yields an ideal remeshing of all surfaces with all landmark vertices located at corresponding surface points.

The crucial step is designing an 'ideal' meshing. The ultimate goal is to obtain surfaces with uniform distributions of vertices. A good initial point is to start with a uniform meshing of the spherical parameter map and to adapt it locally so that a uniform distribution of vertices on the surfaces is achieved. It is commonly known that a good meshing of a unit sphere can be obtained by starting with the polygonal mesh of a icosahedron, which is shown in figure 3.3 c) and gradually subdividing this icosahedron [Zor00]. This leads to an equal distribution of the landmark vertices on the sphere and a good connectivity of the mesh. It is a reasonable starting point for the initial correspondences and later fine tuning.

MDL Optimization

The remeshing step is necessary but not sufficient for establishing robust corresponding landmarks in active shape models. If statistical shape models are generated from erroneous or unprecise landmarks, these errors are reflected in undesired additional variation of the statistical shape model. It is therefore recommended to run an additional 'tuning' step on the raw correspondences which are obtained from the remeshing step. The goal is to separate undesired random variation from systematic variation, which is necessary to describe the natural variation of this class of shapes.

While most authors like [LLS02] describe quality of correspondences by geometrical properties like minimum distances, distortions of landmarks, curvature, etc., Davies *et al.* [DTC+02] present an approach that tackles the problem from the point of view of the compactness of the resulting statistical shape model. Compactness can be quantified by the dimensionality of the model, which is defined as the number of modes, which are necessary to describe a certain percentile of the variation [LLS02]. A perfectly compact statistical model contains all systematic variation of the class of shapes but ignores all random variation from noise or imperfect correspondences.

Davies *et al.* introduce a different measure of correspondence quality of statistical shape models. This measure is called description length and refers to the compactness of the model.

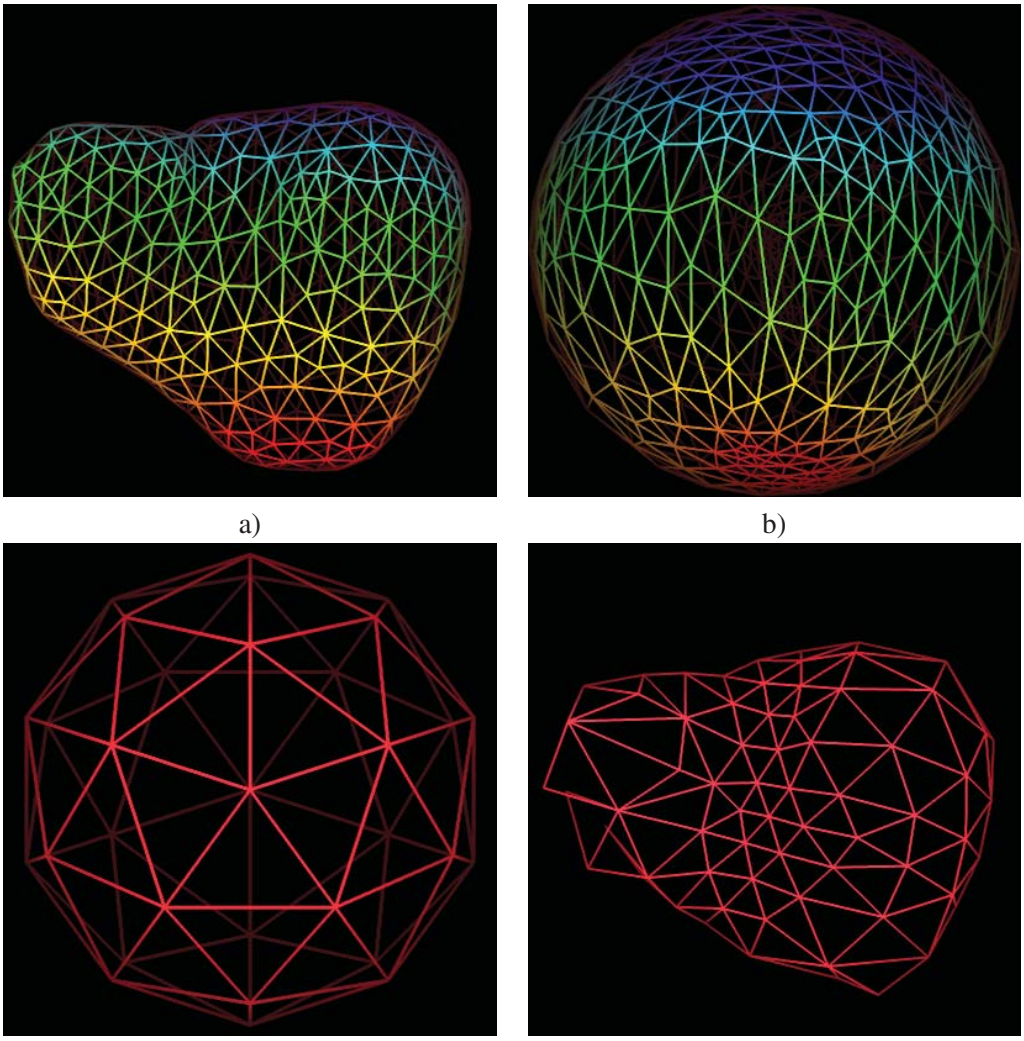


Figure 3.3: This figure shows a schematic liver surface model with low resolution for more clarity. Colors are used for better identification of corresponding vertices. a) shows the original surface, b) the conformal mapping on the unit sphere. c) shows a regular icosahedron after one subdivision step and d) the remeshed surface from a) using the subdivided icosahedron

In other words, one tries to find the most compact model among all models with the same specifity and generalization ability. While generalization ability refers to the capability of the model to not only describe instances which are present in the training data, but also other valid instances, as already explained in 3.1.2, the term compactness refers to the model's ability to describe the complete variance using as few modes resp. principal components as necessary.

This is motivated by the principle of Occam's razor [Tho18], which is basic to science in general, stating that of all possible theories the one that makes fewest assumption and postulates the fewest entities is the best one. Compactness or the description length in the case of Davies' work balances the complexity of the model, defined by the coding length of the model, against its ability to fit to the training data. The coding length can be expressed in terms of the model parameters, i.e. its principal values and a quantization constant.

Heimann [HWWM05] uses a simplified cost function based on minimum description length, which was first published by Thodberg [Tho03]. The cost function is defined as follows:

$$L_m(M) = \begin{cases} 1 + \log \frac{\lambda_m}{\lambda_{cut}} & \text{if } \lambda_m > \lambda_{cut} \\ \frac{\lambda_m}{\lambda_{cut}} & \text{else} \end{cases} \tag{3.10}$$

where the λ_m are the eigenvalues corresponding to the principal components of the statistical shape model M built from the surface parameterizations M_i. λ_{cut} is a predefined threshold .

He uses an initial parameterization Ω. This parameterization maps each landmark vertex v_i to spherical coordinates (θ_i, Φ_i). Ω is used for remeshing each surface S_i, resulting in remeshed surfaces $S_i(\Omega)$. The cost function 3.10 of the model $M(S_0(\Omega), S_1(\Omega), ..., S_{m-1}(\Omega))$ is minimized by optimizing $L_m(M)$ for Ω.

For modifying Ω during the optimization, a function Φ is used, which transforms Ω to $\Phi(\Omega)$. In their original publication, Davies *et al.* [DTC$^+$02] use theta functions, which spread the landmarks near the Cauchy kernel of the theta function and compress them in other areas. A combination of multiple kernels of different widths, amplitudes and positions are capable of generating arbitrary parameterizations. The main disadvantage of this approach is its global nature. Each transformation by a Cauchy kernel affects all of the landmarks. Heimann [HWWM05] tries to overcome this problem by using local transformation kernels, so-called Gaussian envelope functions. The transformation of the parameterization of each landmark is then given by the movement Δk of the envelope, the distance of the landmark from the center of the Gaussian kernel, and the kernel width.

$$\Delta k = (\Delta\theta_{kernel}, \Delta\Phi_{kernel}) \tag{3.11}$$

So each landmark parameterization is mapped as

$$(\Delta\theta_i, \Delta\Phi_i) = g(x, \sigma) \cdot \Delta k = g(x, \sigma) \cdot (\Delta\theta_{kernel}, \Delta\Phi_{kernel}) \tag{3.12}$$

with a Gaussian c, depending on the distance x of the landmark vertex from the center of the Gaussian kernel. The weighting function $g(x, \sigma)$ is defined as follows:

$$g(x, \sigma) = \begin{cases} e^{-\frac{x^2}{2\sigma^2}} - e^{-\frac{(3\sigma)^2}{2\sigma^2}} & for\ x < 3\sigma \\ 0 & for\ x \geq 3\sigma \end{cases} \tag{3.13}$$

So landmarks in the center of the Gaussian kernel move further, while landmarks outside a 3σ interval are not affected at all, which guarantees locality of the transformations.

As all landmarks should be treated equally, Heimann *et al.* [HWWM05] suggest to gradually increase the number of Gaussian kernels and randomly rotate the parameterizations relative to the Gaussian kernel positions to ensure equal transformation of all landmarks. In order to obtain the optimal movements of the Gaussian kernels Δk, a gradient descent method is chosen. The derivatives $\frac{\partial L_m}{\partial k}$ are calculated and

$$k_{t+1} = k_t + \frac{\partial L_m}{\partial k} \cdot \Delta t \tag{3.14}$$

is updated until the optimum is reached. Δt is a predefined step length.

This process is implemented in a multifold multi-scale sense: On the one hand, the number of kernels is gradually increased, on the other hand, the number of landmarks of the surfaces and their parameterizations is increased by a subdivision step after convergence of the MDL cost function as described in the preceding section 3.3.3. After convergence of each MDL optimization step, an additional subdivision step of particular triangles based on the variation of their areas within the space of valid shapes of the statistical shape model can be performed, which will be described in the following section. As a consequence, triangles with high variations of area undergo an additional local subdivision step. This results in a higher density of landmark vertices, i.e. a finer meshing of the surface in this area with smaller triangles and less variation of areas.

Reducing Triangle Area Distortions

After remapping the subdivided icosahedron optimized by minimum description length from the parameterization on the spherical surface to the original liver surfaces, mesh quality can be very poor, as figure 3.3 d) shows. Large triangles can cause large surface errors in areas of large surface variation and curvature. A simple idea would be to increase the number of landmarks, which, however, in turn also sacrifices computational efficiency.

Heimann *et al.* [HWM06b] solve this problem by measuring the average area distortions between the triangles of the original surfaces and the triangles in the resulting conformal mapping. The value of average distortion is then converted to a color, which is assigned to the respective triangle of the spherical map. This colored map is then used to produce a dithered map, with dither point density increasing with increasing brightness. All dither points are then considered to be vertices in the new mesh and a Delaunay triangulation is used to generate a connectivity. After remapping this connectivity to the original surfaces, an ideal meshing is obtained.

The approach used in this thesis is similar but not identical. The idea is to have an equal distribution of landmark vertices not only in the training liver surfaces, but in the statistical shape model built from the training surfaces. This is done by considering the variation of triangle area in the resulting model and a local subdivision strategy. The variation of triangle area is computed by generating different instances of the space of valid surface shapes, i.e. the 3σ interval of feature values from equation 3.3 and considering the different areas of the triangles.

More precisely, model instances are generated by repeatedly setting each of the feature values b_i in the linear system from equation 3.3 to ± 3. The statistics of all triangle areas in all model instances are calculated. All triangles that differ by more than a threshold t from the mean area, are subdivided by means of a locally smooth subdivision scheme. $\sqrt{3}$-subdivision as introduced by Kobbelt [Kob00] was used for this purpose, because this scheme supports local subdivision and because the number of triangles grows more slowly in $\sqrt{3}$-subdivision than in other subdivision schemes.

This additional subdivision step leads to a statistical shape model with a high-quality surface meshing and almost equally sized triangles over the entire space of valid shapes. It is also easier to implement than the dithering method by Heimann and integration into the hierarchical MDL based correspondence optimization algorithm is straightforward.

Putting it all together

The entire process of generating the statistical surface shape model is visualized in a flow-chart of figure 3.4. The loop starts with an initial number of 842 landmark vertices. The process is stopped after three subdivision steps and reduction of triangle area distortion steps, with a total number of 2563 landmarks.

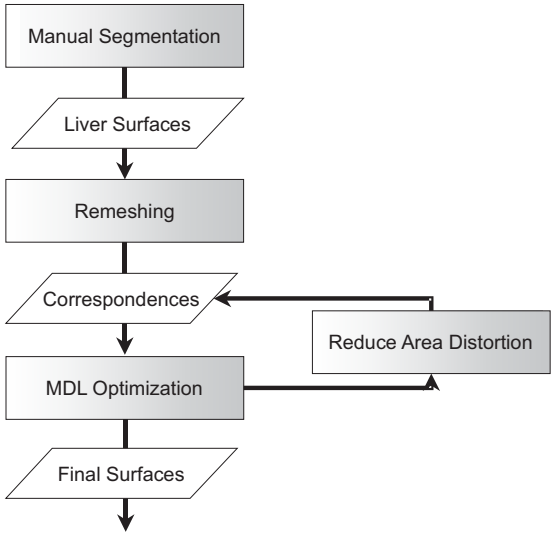


Figure 3.4: Flow-chart of the steps for the generation of the statistical surface shape model.

3.3.4 The Model and its Principal Modes of Variation

The final statistical surface shape model is created from the coordinates of the optimized correspondences. The coordinates of the corresponding landmarks are assembled to large vectors and principal component analysis (PCA) yields a linear model, as described in 3.1.3. Before PCA can be applied to the coordinate data, the shapes are preprocessed by a normalizing

geometric transformation, e.g. a similarity transformation, in order to remove the effects of rotation, shift and scaling from the model's actual variability in shape. The statistical variation of shape should be examined free of biases caused by different transformation offsets.

If robust point-to-point correspondences between two shapes S_1 and S_2 are available, a similarity transformation T with respect to a minimum least-squares-error $||T(S_1) - S_2||^2$ can be efficiently found as described by Zinßer *et al.*[ZSN05] or in section A.2. Normalization is done by selecting one of the shapes as a reference and transforming all other shapes to this reference. All coordinates are then expressed within the coordinate system of the reference shape. PCA computes the average shape and its principal modes of variation, the so called eigen shapes .

Figure 3.5 shows renderings of the final average liver model instance $\bar{x}$ and its eigen shapes, the principal components of shape variation v_i weighted by 3 times the roots of their principal values $\sqrt{\lambda_i}$, which corresponds to 3 times the standard deviation in this direction, resp. an interval of 99% in terms of the multivariate Gaussian.

3.4 Modeling Local Appearance

The statistical shape model can be used for generating valid instance of liver shapes and also for finding the shape instance from the model which is the closest representation of a given liver surface shape, but not the one which best fits given image data. Hence, the model is only capable of describing the surface itself because there is no correlation to image data. For image segmentation, however, this correlation is needed in order to assess how well the shape model 'fits' the image data. By optimizing the quality of this fit, the model instance that 'best' describes the image data is found.

Active appearance models [CET98] establish a correlation between the statistical surface model and the image data by extending the data vector on which the PCA is calculated from pure geometric landmark vertex coordinates to additional grey value samples from the image data taken at corresponding points. The differences between the samples in the image to be segmented and the samples assigned to the model are used to update the model parameters resp. feature values [CET99]. This procedure assumes that there is a statistical correlation between the geometric data and the grey values of the image, which may not always be the case. Moreover, for segmentation using active appearance models, the effects induced by geometry and the effects induced by grey value intensities must be weighted against each other carefully by finding robust empirical coefficients in an offline process.

Active shape models keep the statistics of the image data independent of the geometry by maintaining separate grey value profiles or statistics of the grey value neighborhoods for each landmark. The main difference to active appearance models is the local nature of the grey value sample data resp. profiles and their independence of the geometry statistics. The segmentation process consists of searching the best instance of the active shape model according to the local grey value profiles. During the active shape model search, the profiles are used to update the landmark vertex positions to the best matching displacements according to the image data. The shape model statistics restrict the search to valid instances of the shape. A detailed comparison of active shape models and active appearance models can be found in [CET99]

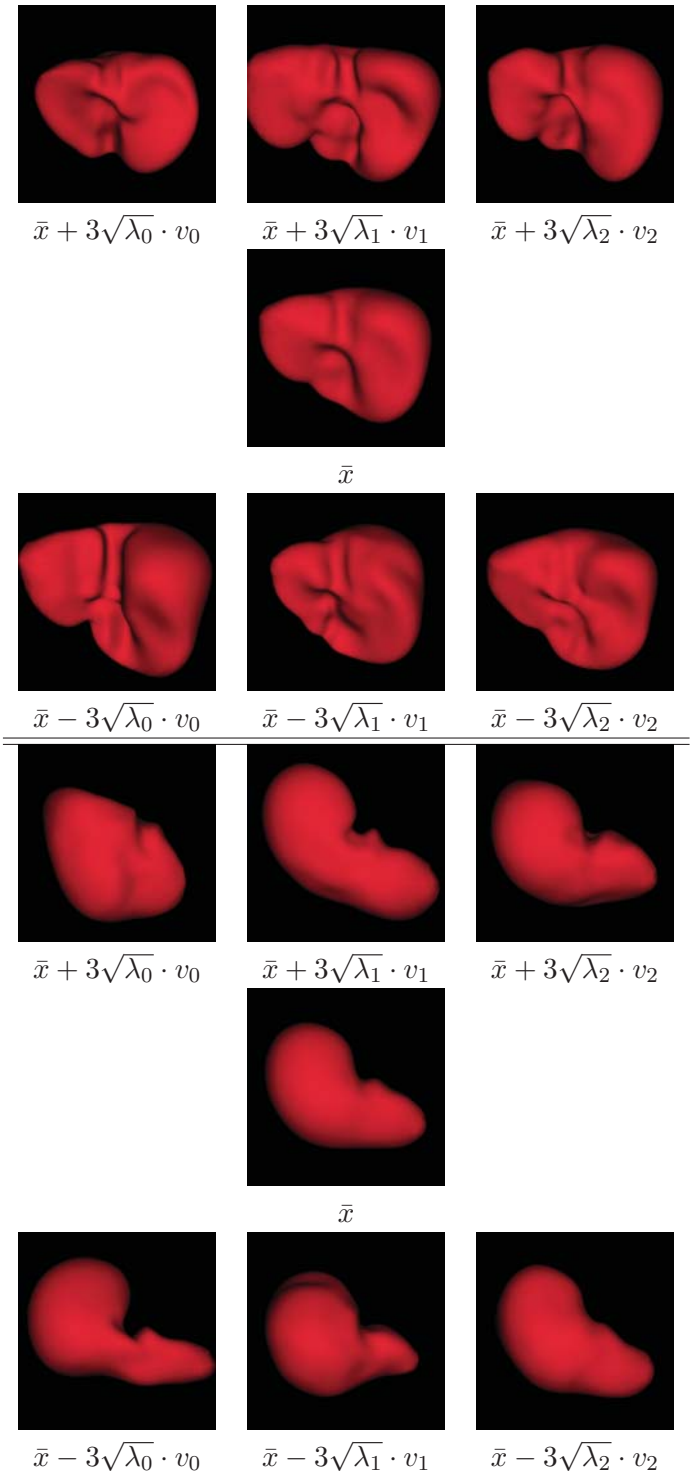


Figure 3.5: The liver model and its eigen shapes, i.e. its major modes of variation. All liver model instances are shown once in a coronal and once in an axial view.

In this work, the former approach is implemented, so image appearance is modeled by considering grey value statistics resp. profiles for each landmark vertex. The collection of all profiles is called the local model of appearance of the statistical surface model and must not be confused with an active appearance model. Two strategies of implementation of the local model of appearance were assessed which will be described in the following sections.

3.4.1 Offline Strategy

The offline strategy is similar to the techniques described by Cootes in [CET98], Cootes in [CET99], Heimann in [HWM06a] or Lamecker in [LLS04]. The name offline is chosen because all image data are collected and processed in an 'offline' process before the actual segmentations. For each landmark vertex v_i in each image data set I_j, a sampling ray is shot in normal direction to the surface $\vec{n_i}$ and image grey value samples p_{ji_k} are taken at equidistant positions along the ray. Instead of shooting just one ray, a bundle of rays is used and the profiles of all of them are collected weighted by a Gaussian kernel centered at the central ray. The weight for a given ray is calculated according to the distance of the ray to the center using a Gaussian kernel. Figure 3.6 a) shows a bundle of rays and the sample points for one landmark vertex. This has a low-pass filtering effect on the profiles and provides stability for subtle variations in the normal direction, which may occur.

The collection p_{ji_k} of all samples is called the profile p_{ji} of the landmark vertex v_i in image I_j. k denotes the index of the sample in the profile. Each profile is normalized by its maximum absolute value to reduce its dependence on global grey value scalings.

$$p_{ji} \leftarrow \frac{p_{ji}}{\max_k |p_{ji_k}|} \tag{3.15}$$

After normalization, for each landmark vertex a mean profile $\bar{p_i}$ and a covariance matrix S_p can be calculated.

$$\bar{p_i} = \frac{1}{n} \sum_j p_{ji} \tag{3.16}$$

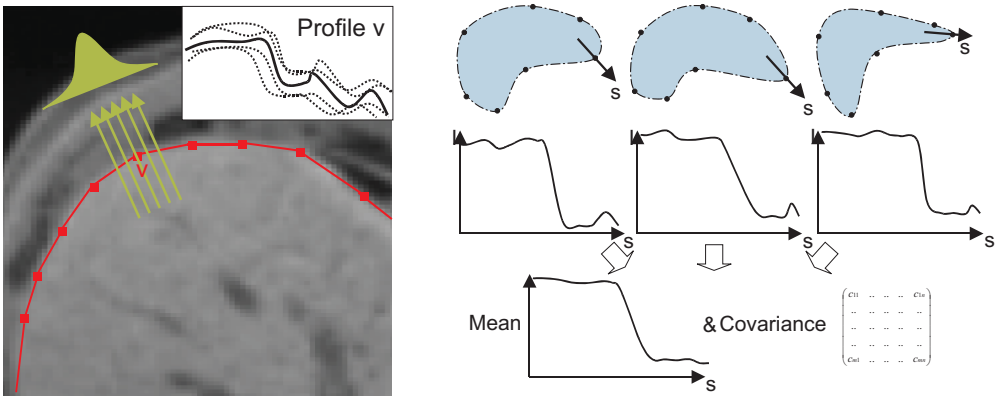
$$S_p = \frac{1}{n-1} \sum_j (p_{ji} - \bar{p_i})(p_{ji} - \bar{p_i})^T \tag{3.17}$$

Figure 3.6 shows a schematic overview of the offline strategy for creating a local model of appearance.

For each landmark vertex, the mean profile and the covariance matrix are stored for several image resolutions in a data structure which can be accessed easily during the active shape model search.

3.4.2 On-the-fly Strategy

In contrast to the offline case, the on-the-fly strategy preprocesses the training image data offline, but the actual local model of appearance resp. the grey value profiles are generated from the preprocessed image data during the segmentation process on-the-fly. The advantage



a) b)

Figure 3.6: a) shows a bundle of sampling rays for a landmark vertex with dashed lines, the samples and a resulting profile from the Gaussian weighting with the solid line. b) illustrates the offline strategy for modeling appearance by shooting sample rays, averaging and calculating the covariance of the image.

of this strategy is that arbitrary samples or profiles can be taken from the prepared image data, while in the offline case the sampling direction, the number of samples and the sampling distance have to be determined beforehand and are therefore fixed.

During the image preprocessing, statistical images are generated from the training data images, which can be used during the segmentation for on-the-fly lookup of arbitrary profiles and grey value statistics. For this purpose, each training image is warped to a reference image in a way that all the landmark vertices of the corresponding liver surfaces are aligned with the landmarks of the reference liver surface. The reference of choice is the mean liver surface $\bar{x}$ from equation 3.3 because choosing the mean ensures that minimal overall distortion is needed for the warps of all image data sets. See figure 3.7 for a clarification of the warping to the mean. A thin-plate-splines transformation [Boo89] is used for the image warps because this transformation fulfills the interpolation condition, ensuring that each transformed landmark vertex $T(l)$ coincides with each corresponding reference landmark vertex l'.

$$T(l) = l' \tag{3.18}$$

Besides, thin-plate-spline transformations provide smoothness between the interpolation points by minimizing the following energy functional [Boo89]:

$$E_{TPS} = \sum_{i}^{K} ||l'_i - T(l_i)|| + \lambda \int_{\Omega} \left[\sum_i \sum_j \left(\frac{\partial^2 T}{\partial x_i \partial x_j} \right)^2 \right] d\Omega \tag{3.19}$$

Consequently, the transformation T is the one which minimizes the given functional. Its computation is described in [Boo89].

$$T = \arg\min E_{TPS} \tag{3.20}$$

Smoothness is guaranteed because of the integral of norms of the second derivatives over the entire image domain are minimized. All these properties [Boo89] make thin-plate-splines a good choice for the warping transformation.

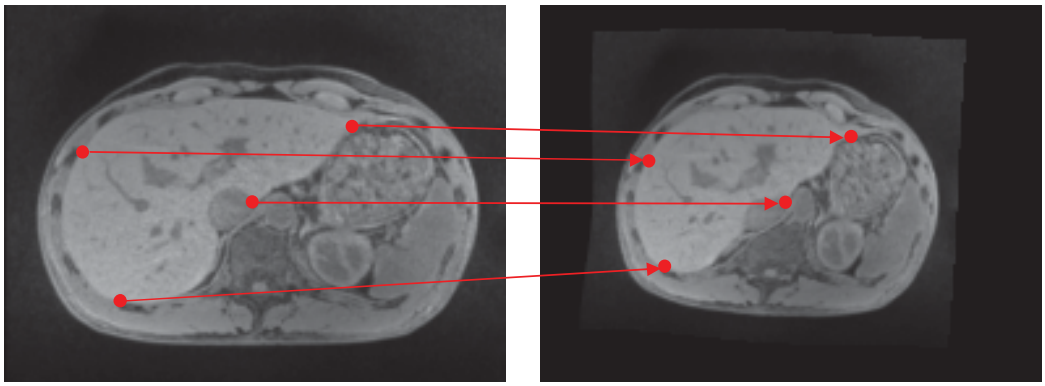


Figure 3.7: Example of a deformation warp using thin plate splines from the slice of a data set to the mean. Surface points are warped to their corresponding mean positions. The rest of the image is smoothly interpolated.

Warping all training data images produces n transformed images which can be averaged to a mean image, representing the pixelwise mean over all warped training images and a standard deviation image, which represents the pixelwise standard deviation over all training images. Choosing the thin-plate-splines warp transformation with the correspondences of the liver surfaces ensures pixelwise correspondences in the surrounding of the liver surface. So for each voxel j the mean image I_{mean} is computed as:

$$I_{mean}(x(j)) = \frac{1}{n} \sum_i I_i(T_i(x(j)) \tag{3.21}$$

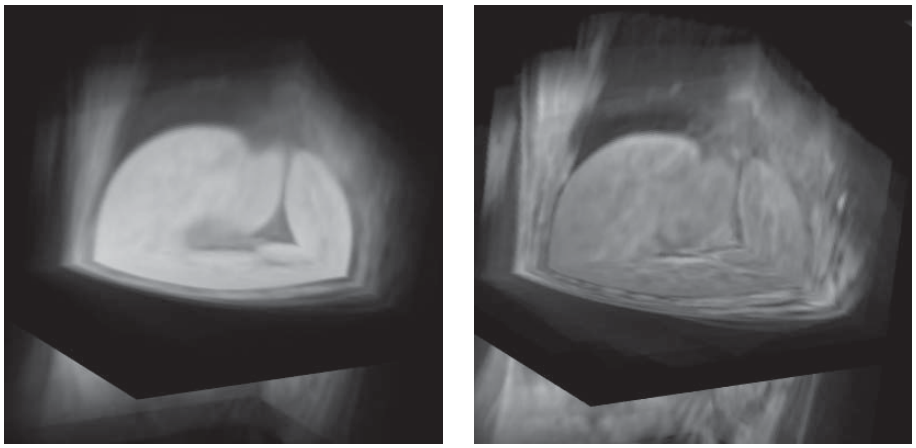
where $x(j)$ is the spatial coordinate of the voxel and T_i the transformation which transforms the i-th liver surface to the mean surface, as defined in equation 3.20. The standard deviation image I_{stddev} is computed in a similar way:

$$I_{stddev}(x(j)) = \frac{1}{n-1} \sum_i \sqrt{(I_i(T_i(x(j)) - I_{mean}(x(j)))^2} \tag{3.22}$$

Figure 3.8 shows the generated statistical mean and standard deviation images in high resolutions.

3.4.3 Samples Outside of the Field of View

Reference publications seldom describe how sample rays which leave the available field of view of the training images should be handled and how the missing samples should be treated. One way would be to reduce the number of samples for landmark vertices close to the borders of the FoV. However, this approach would artificially remove a lot of valuable information.



a) b)

Figure 3.8: a) shows the mean image in high resolution for the local on-the-fly model of appearance. b) shows the image of the pixelwise standard deviation of the appearance model in high resolution.

In this implementation a mixed approach is applied and found to be useful. For sample rays leaving the field of view in sagittal and in coronal directions, all samples outside are set to 0 making the valid assumption that the images are not truncated in these directions. For sample rays leaving the field of view in axial direction, the latter assumption does not hold. In this case, the sample values were set to values at the border and the variances of these samples were set to some maximum in order to appropriately consider the uncertainty about the sample values.

3.5 Active Shape Model Search

The actual segmentation or active shape model search consists of repeating the following steps until the optimum has been reached:

1. Optimal landmark vertex displacements based on the image data are calculated according to the profiles and the statistics of local surface appearance.

2. Normalization of the model shape with displaced vertices is done by the same transformation as during the model creation phase 3.1.3

3. The model instance which is closest to the model shape of displaced and normalized models is calculated by projecting the vertex coordinates of the normalized model to the linear system of the model in equation 3.3. Projections guide the model based search within reasonable bounds and are also used by [CET98], [CET99], [HWM06a] or [LLS04].

In most publications, the step of finding landmark displacements is done by comparing grey value profiles at different offset positions of the landmarks to the profile statistics contained in the local model of appearance. The landmark vertex is then shifted to the position with the highest 'similarity' to the model of local appearance. Figure 3.9 shows an example and a comparison of profiles at different vertex shifts. After the normalizing transformation, a projection to the statistical model, yields the most probable instance of the statistical shape model, given the image data profiles. The algorithm is also implemented in a hierarchical way, which means that a Gaussian pyramid of the image data is used both for building the local model of appearance and for the segmentation. The algorithm starts at a coarse resolution of image and profile data and continues with a finer resolution after convergence. Convergence is defined, if in three consecutive steps the vertex shape coordinates do not change by more than an experimental threshold.

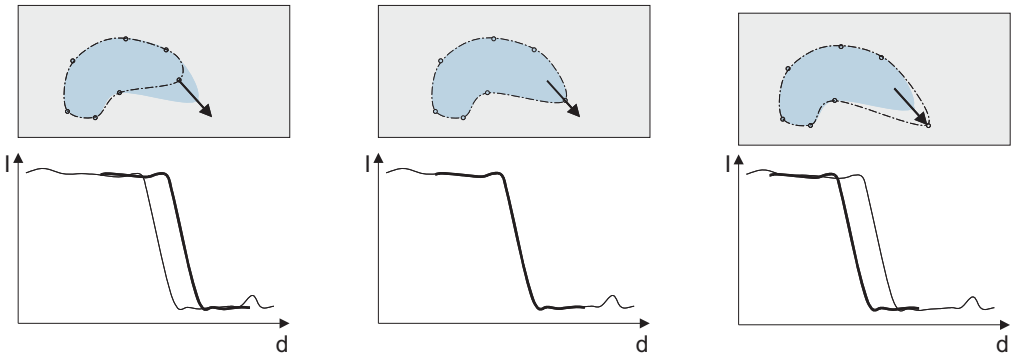


Figure 3.9: This figure shows a schematic overview of different profiles of a vertex at different offsets of the denoted landmark vertex. The profile of the local model of appearance is shown with the thin line, while the actual profile of the vertex at the respective offset is shown with a fat line. In b) similarity is maximal.

3.5.1 Profile Similarity Measures

For a quantitative assessment of the distance of a sample profile and a profile given by the local model of appearance, a similarity metric is needed. As stated in section 3.4.1, the profiles are low-pass filtered by sampling a Gaussian weighted bundle and normalized by the maximum absolute value. Similarity is quantified by comparing normalized profiles.

Under the assumption of multivariate Gaussian distributions of the profile values, a measure of how well a given sample profile p_i fits to the local model of appearance at landmark vertex i is given by the Mahalanobis distance $dist_M$ [Mah36], which is defined by:

$$dist_M(p_i) = (p_i - \overline{p}) \cdot C_{p_i}^{-1} \cdot (p_i - \overline{p}) \qquad (3.23)$$

where p_i is a given sample profile, $\overline{p}$ the mean appearance profile at landmark vertex i. C_{p_i} is the covariance matrix of the normalized profile values at the vertex p_i.

The Mahalanobis distance is directly related to the probability that a profile p_i is drawn from the distribution, i.e. the smaller the distance, the higher the probability.

The Gaussian distance $dist_G$ is a similar metric measure. It is directly obtained from the Mahalanobis distance if the correlation between different items of the sample profile is set to 0, i.e. only the diagonal elements of the covariance matrix are considered. Its computation is also more efficient.

$$dist_G(p_i) = \sum_j \frac{(p_{ij} - \overline{p_j})^2}{C_{p_i,jj}} \tag{3.24}$$

where $C_{p_i,jj}$ is the jth diagonal entry in the covariance matrix C.

3.5.2 Initializing the Model

Good initialization of the model is crucial because the better the initialization, the faster and the more robustly the model converges. The model is initialized at the average liver position and shape within a given torso, using a statistical approach by learning the average positions and scales from the training image data.

A coordinate frame is assigned to the torso of each 3D training image data set, which is called 'torso frame'. All other geometric calculations are made in 'torso coordinates'. The torso frame is defined by an origin $\vec{c}$ and a right-handed vector system $\{\overrightarrow{v}_{sag}, \overrightarrow{v}_{cor}, \overrightarrow{v}_{tra}\}$ which are computed from the image data. The coordinate frame is determined by first transforming the images to binary images by applying a noise threshold. The principal axes of the binary torso are calculated by computing the eigenvectors and eigenvalues from the covariance matrix of all pixel coordinates above the threshold. The principal components are assigned to the vectors $\overrightarrow{v}_{sag}$, $\overrightarrow{v}_{cor}$ resp. $\overrightarrow{v}_{tra}$. Scaling for $\overrightarrow{v}_{sag}$ and $\overrightarrow{v}_{cor}$ is done by the eigenvalues that correspond to the eigenvectors. As the image extent in transversal direction can be arbitrarily large or small, this strategy does not work for $\overrightarrow{v}_{tra}$. So the scale in axial direction is set to

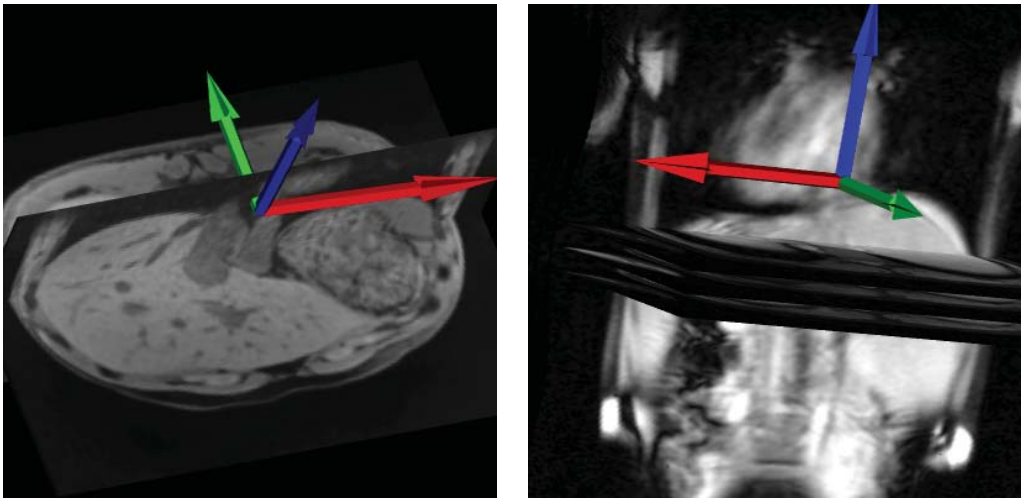
$$||\overrightarrow{v}_{tra}|| \leftarrow \frac{1}{2}(||\overrightarrow{v}_{sag}|| + ||\overrightarrow{v}_{cor}||) \tag{3.25}$$

The sagittal and coronal elements of $\vec{c}$ are taken from the center of gravity of the binary image. Again, the transversal component of $\vec{c}$ is harder to determine. A stable approach is to take the position of the diaphragm, which can be found by a simple threshold segmentation of the lung.

Figure 3.10 a) shows the torso frame in one of the 3D training data samples while b) shows a torso frame from a set of slice images.

Now the coordinates of each manually segmented liver l_i and its center of gravity $\bar{x_{l_i}}$, along with its scale can be expressed within its torso frame T_i. The scale of a liver shape is determined using the Frobenius norm, as suggested in [SG02].

$$s_i = \sqrt{\sum_j (x_j - x_{l_i})^2} \tag{3.26}$$



a) b)

Figure 3.10: In a) an example of a torso frame from a 3D volume image, in b) from a 2D torso frame image is shown with the arrows pointing in the directions of the principal components. The arrows have been scaled by the principal values.

The average scale $\bar{s}$ can be calculated from the scales s^{T_i} with respect to the torso frame T_i

$$\bar{s} = \frac{1}{n} \sum s_i^{T_i} \tag{3.27}$$

and center of gravity

$$\bar{x} = \frac{1}{n} \sum x_{l_i}^{T_i} \tag{3.28}$$

The model can be initialized using the average scale and center of gravity in the given torso frame for the following active shape model based search. The liver coordinates can be easily transformed from world coordinates $l_i^{T_w}$ to the corresponding torso frame coordinate system yielding $l_i^{T_i}$.

$$l_i^{T_i} = T_i^{-1} \cdot (l_i^{T_w} - c^{T_w}) + \bar{x} \tag{3.29}$$

In this equation, T_i is the matrix consisting of the right-handed vector system containing $\overrightarrow{v}_{sag}$, $\overrightarrow{v}_{cor}$ and $\overrightarrow{v}_{tra}$ as columns.

$$T_i \leftarrow \begin{pmatrix} v_{sag_x} & v_{cor_x} & v_{tra_x} \\ v_{sag_y} & v_{cor_y} & v_{tra_y} \\ v_{sag_z} & v_{cor_z} & v_{tra_z} \end{pmatrix} \tag{3.30}$$

where v_{sag_x} is the x component of the vector $\overrightarrow{v}_{sag}$ and so on.

In an unseen data set, first the torso frame T_i is calculated and each model point with coordinates x is subsequently initialized according to the average center of gravity and scale of this frame by calculating the corresponding world coordinates $l_i^{T_w}$.

$$l_i^{T_w} = T_i \cdot (\bar{s} \cdot (x - \bar{x})) + \bar{c}^{T_i} \quad (3.31)$$

For incomplete or sparse data like in the case of localizer images, the torso frame can be calculated in the same way and the initialization of the liver shape is identical to the equations 3.29 and 3.31.

3.5.3 Active Shape Model Search on 3D Data

Although the active shape model has been designed for the purpose of segmentation from 2D stacked localizer image slices, it can also be used for fully automatic segmentations of the liver from 3D MR image data scanned with same MR imaging protocol as the training image data.

Application to segmentation from 3D image data can also serve as a test of the performance and generalizability of the active shape model. Cross-validation experiments, as in [LLS04], in which the training data are divided into a training subset and a testing subset, can be set up in this way to check the quality of the active shape model.

The segmentation algorithm for 3D image data closely follows the standard 3D active shape model search as described by Cootes [CT99], which was successfully used by Lamecker *et al.*[LLS04] or Heimann *et al.* [HWM06a] for fully automatic liver segmentations from CT image data.

The algorithm consists of the following steps: As a starting point, the active shape model is initialized by computing the 3D torso frame and the segmentation runs as described for the generic active shape model search. During the iterative search, both the reference profile samples and profiles statistics of the appearance model are sampled in normal directions of the landmark vertices at the surface. The optimal landmark displacements are found by comparing the two and taking the landmark displacements with the minimal distance of the profiles according the Mahalanobis or Gaussian metric. The 3D active shape model search is later used for validation of the active shape model by computing cross-validation segmentations on the original data, from which the model has been created.

Since the cross-validation experiments showed good results (as described in the results section 3.6 of this chapter), the active shape model and the 3D segmentation algorithm could also be used for quantitative segmentation of 3D liver data in the clinical routine. Volumetric quantification of the liver is a useful measure in various clinical questions.

3.5.4 Active Shape Model Search on a Sparsely Covering Set of 2D Data

Differences to the 3D Case

The 3D active shape model search on sparsely covering 2D slices of image data requires a different way of sampling profiles from the image data and profile statistics from the training data because, for a given landmark vertex, the 2D image slice plane need not coincide with the normal direction of the liver surface. It would not be sufficient to take only those landmark vertices into account, which have normal directions that coincide with the image slice planes. However, holding statistics for an arbitrary number of image plane directions and vertex po-

sitions would not be very efficient, either. The following sections will show how to obtain reference profiles for the offline and the on-the-fly strategy in an efficient way.

Localizer Images

The main difference to the approach used by Heimann *et al.* [HWM06a] and Lamecker *et al.* [LLS04] and which was described in section 3.5.3, the basic input images for the liver segmentations in this case are not 3D volume images, but stacked slice images which sparsely cover the volume of interest. Sparsely covering stacks of slice images are common in the context of scan planning and also called localizer images or scouts. These images are scanned using fast sequences and are characterized by large fields of view (400 mm), short acquisition times (less than 30 s) and moderate image quality. Usually, fast FLASH sequences [HFM$^+$86] are used for abdominal scout images.

For the experiments of this chapter, between 3-6 coronal slices and 3-6 axial slices were used with a inter-slice distance of 15 mm and a slice thickness of 10 mm. The contrast of these images must be designed as similar as possible to the training image data from which the appearance model has been generated. For testing purposes, 'artificial' localizer images can be generated from the high resolution 3D training image data by computing slice images with the same parameters as in a localizer scan protocol.

Using the Offline Strategy

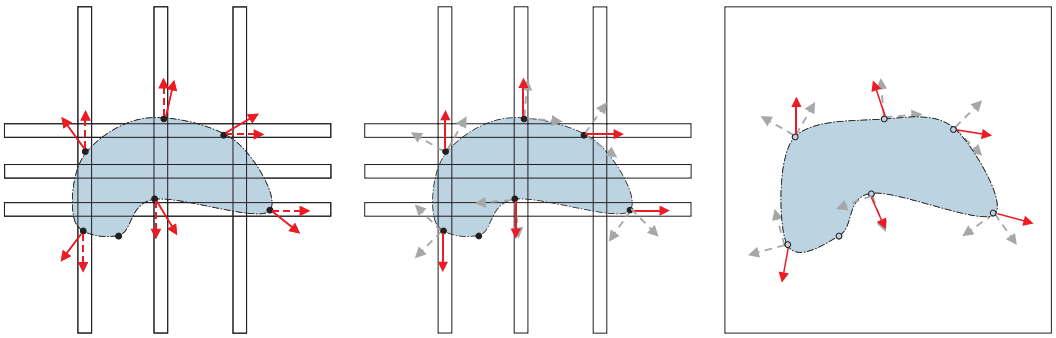
For the offline strategy where statistical reference sample profiles have been taken from the training image data sets using sample rays in normal directions beforehand, the profile sampling directions in the stacked 2D localizer image slices are projected from the normal directions to the image planes. In order to limit the error made by the projections, the accepted difference in angle is set to c_{angle} degrees. If the difference is larger, this landmark vertex is excluded from the computations in the current iteration because of the possible error induced by the projections. However, if c_{angle} was chosen too strictly, the number of landmark vertices with valid profiles would become too small. A threshold of $c_{angle} = 30°$ turns out to be a good compromise.

This procedure is shown in figure 3.11 a).

Using the On-the-fly Strategy

For the on-the-fly strategy, where all training liver images have first been warped to fit to the surface of the average liver shape, the sampling directions in the image planes can be transformed to the respective directions in the mean and variance image. This is implemented by setting up local coordinate frames for each landmark vertex in both the current shape and the average shape. The sampling direction can then be transformed by a simple linear transformation resp. matrix multiplication into the coordinate frame of the mean and variance image. Figure 3.11 b) and c) visualize how the sampling directions are transformed.

So, if $\vec{n_i^{T_w}}$ is the sampling direction at landmark vertex i with coordinates $v_i^{T_w}$ in the world coordinate frame T_w, then let T_i be the local coordinate frame at i. T_i is a matrix composed of 3 orthogonal vectors as its columns $\vec{t_x}$, $\vec{t_y}$ and $\vec{t_z}$.



a) b) c)

Figure 3.11: The three figures show a 2D scheme of a liver model and a set of stacked slice images indicated by the bars. a) illustrates the offline strategy: the sampling directions are projected to the image slice planes. The directions after the projections are shown with the dashed lines.
b) and c) show an example of the on-the-fly strategy: the sampling directions in the slice image planes are shown in b) along with the local coordinate frames (dashed grey arrows). c) shows how the sampling directions have been converted to the statistical images according to the local coordinate frames (again shown as dashed grey arrows). The converted sampling directions indicate the sampling directions for the reference profiles from the statistical images. The bounding box of the volume of the statistical images is delineated by the black rectangular frame.

$$T_i \leftarrow \begin{pmatrix} t_{xx} & t_{yx} & t_{zx} \\ t_{xy} & t_{yy} & t_{zy} \\ t_{xz} & t_{yz} & t_{zz} \end{pmatrix} \tag{3.32}$$

$\vec{t_z}$ is equal to $n_i^{\vec{T}_w}$. $\vec{t_x}$ and $\vec{t_y}$ are both located in the tangent plane D with the additional constraint that the projection of the edge vector connecting i to an anchoring neighbor $Nb(i)$ to the tangent plane is perpendicular to $\vec{t_x}$.

$$\vec{t_y'} \leftarrow (\vec{i} - \vec{Nb(i)}) - ((\vec{i} - \vec{Nb(i)}) \cdot n_i^{\vec{T}_w}) n_i^{\vec{T}_w}) \tag{3.33}$$

$$\vec{t_y} \leftarrow \frac{\vec{t_y'}}{||\vec{t_y'}||} \tag{3.34}$$

where $\vec{t_y'}$ is the connecting vector between i and $Nb(i)$ with all normal content removed and $\vec{t_y}$ the same vector with unit length.

$\vec{t_x}$ is chosen in a way that completes the right-handed orthonormal system:

$$\vec{t_x} \leftarrow \vec{t_y} \times \vec{t_z} \tag{3.35}$$

The sampling directions in the coordinate frame T_f of the mean and variance image $n_i^{\vec{T}_f}$ are then obtained according to

$$n_i^{\vec{T}_f} = T_f^{-1} \cdot T_w \cdot (n_i^{\vec{T}_w}) \tag{3.36}$$

Figure 3.12 shows an example of how a local frame is set up at a certain vertex.

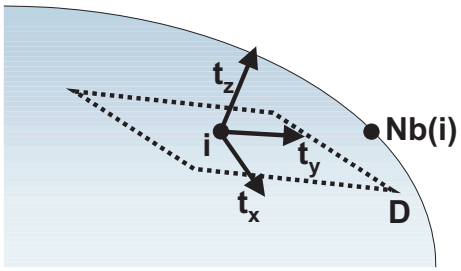


Figure 3.12: This image shows a local frame with the three orthogonal vectors $\vec{t_x}$, $\vec{t_y}$ and $\vec{t_z}$ and the tangent plane D. $Nb(i)$ is the anchoring neighbor vertex of i.

With this strategy, it is also possible to extend the search direction perpendicular to the image plane. Additional profiles can be sampled by moving the sampling ray perpendicular to the image plane. This means that the movement of a landmark may also obtain an out-of-plane component, which in some cases makes the model search converge faster.

3.5.5 Subspace Shape Model Projection

After moving the landmark vertices to the optimal displacement positions, a projection to the statistical model has to be performed to ensure a valid shape instance. In the standard active shape model search [CT99], the closest instance of the class of shapes of the statistical shape model is found after normalizing the shape by the parameters of a geometrical normalization transformation [CT99]. The shape normalization step is identical to the one during the creation of the statistical shape model in section 3.1.3. The feature vector b of the closest instance is then given by

$$b = \Phi^t \cdot (x - \overline{x}) \tag{3.37}$$

where x are the coordinates of the shape after shape normalization by the normalizing transformation. We obtain equation 3.37 by solving for b in equation 3.3. This corresponds to an orthogonal projection to the space of valid shapes and yields the closest instance of the model according to a least squares error metric of the coordinates [CT99]. After the projection, the values b_i are confined to a 3σ-interval of the implicit multi-variate Gaussian distribution of the model.

$$b_i \leftarrow \min(3\sigma_i, \max(-3\sigma_i, b_i)) = \min(3\sqrt{\lambda_i}, \max(-3\sqrt{\lambda_i}, b_i)) \tag{3.38}$$

where $\sigma_i = \sqrt{\lambda_i}$ is the standard deviation corresponding to the i-th principal vector, with λ_i being its principal value.

Often, the dimensionality of the model is reduced by removing columns from Φ. For example, if only the n most important modes of variation are considered, then Φ contains only n columns and b only n entries. This corresponds to an abstraction reducing the complexity of the model to the n degrees of freedom, which account for the maximum variance of any set of n modes of variation.

For sparsely covering 2D slices of image data, the standard projection has to be adapted. Due to the sparse coverage of the volume by the image slices, not all vertices of the normalized model are located within at least one slice during the active shape model search but only a subset. Hence, the set of landmark vertices V is partitioned into active landmark vertices U located within at least one slice and having at least one sample profile available to be compared to a reference and inactive landmark vertices $\bar{U}$.

Figure 3.11 a) illustrates the partition into active vertices, containing shift arrows and inactive vertices without arrows. The adapted projection may only include active landmark vertices. This projection can be called subspace shape model projection.

The active shape model that corresponds to the subset of active landmark vertices is called subspace shape model. As the slice images are 2D, only active vertices provide image samples and profiles which can be used for the active shape model search, so equation 3.3 is no longer sufficient. Instead it must be adapted to include a projection Π from the space of coordinates of the vertices of the complete model V to the space of coordinates of the subspace shape model U.

$$\Pi : x^* = P \cdot x \tag{3.39}$$

$$x^* = P \cdot \overline{x} + P \cdot \Phi \cdot b = \overline{x}^* + \Phi^* \cdot b \tag{3.40}$$

where x^* is a vector containing only the coordinates of the active landmark vertices, i.e. $v_i \in U$ and P is a projection matrix, transforming x to x^*.

In contrast to Φ, Φ^* is not orthonormal. However, as only the n most important eigenvectors are considered and as the number of active landmark vertices $|U|$ is usually large enough, Φ^* has full rank, and equation 3.40 can be solved for b. The pseudo-inverse Φ^{*+} can then be calculated by means of a singular value decomposition [VF02]. So the subspace shape model projection is calculated as follows:

$$b = \Phi^{*+} \cdot (x^* - \overline{x}^*) \tag{3.41}$$

Since by adapting the model, the partition of V into U and $\bar{U}$ can change, Φ^* has to be updated and its pseudo-inverse Φ^{*+} has to be recomputed whenever the set of active landmarks changes. In practice this makes the method a little bit less efficient than the 3D active shape search, where because of Φ being orthonormal computing the inverse is straightforward and only has to be done once for all steps.

3.6 Validation and Results

The method was validated on two levels:

First, the soundness of the active shape model was evaluated by performing cross-validation tests on the 3D training image data using a 3D active shape model search as described in 3.5.3, based on the works published by Lamecker *et al.* [LLS04] or Heimann *et al.* [HWM06a]. This yields information on the overall performance of the active shape model itself and its local model of appearance.

Then the method was validated by generating slice images from the 3D data and comparing the results of the algorithm to the results obtained using the 3D active shape model search.

Both leave-all-in and leave-one-out tests were performed on the data sets available in all scenarios. A leave-all-in test consists of building the shape model and its local model of appearance from all training data sets available. Then the algorithm is run on each of these data sets. This yields information on the overall soundness of the method. A leave-one-out test consists of building the active shape model and its model of appearance of all but one data set. Then the algorithm is run on the data set left out. This yields information on the generalizability of the model.

The manual segmentations of the input data were taken as the ground truths and compared to the results of the surface reconstructions computed by the algorithm. Four error metrics were chosen to measure the difference between a model m and the ground truth t:

- The average Euclidean distance between the landmark vertex positions in the ground truth segmentations and the positions in the segmentation result.
- The volumetric error $2|V_m - V_t|/(V_m + V_t)$, which corresponds to the relative difference in volume, with V_m being the volume of the model instance and V_t being the volume of the ground truth surface.
- The RMS surface distance error, similar to [HWM06a], defining the fraction of the surface area, which is further than 7.5 mm, i.e. three quarters of the thickness of a slice away from the ground truth surface.
- The Tanimoto error [Tan58], which is defined as the fractional volume being part of the combined volume but not of the overlap volume: $1.0 - V_{m \bigcap t}/V_{m \bigcup t}$. Here $V_{m \bigcap t}$ is the jointly covered volume and $V_{m \bigcup t}$ is the combined volume.

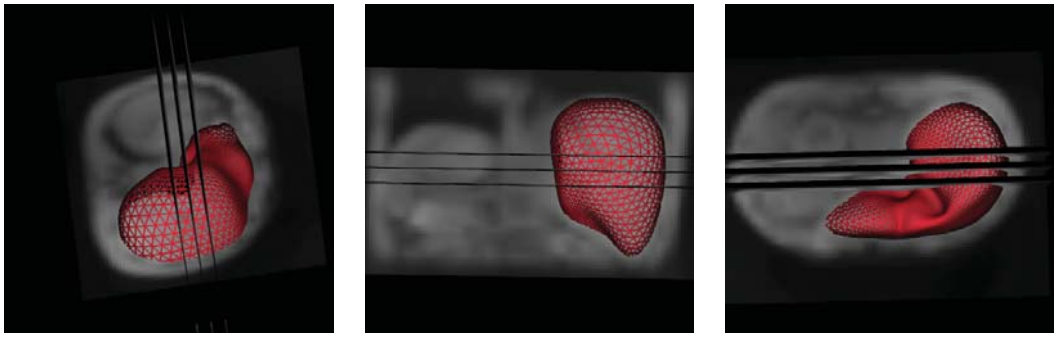
The validation experiments were done using the 12 major modes of the active shape model. The 2D slice adaptation examples were performed once with very sparse slice placement (3 coronal slices and 3 transversal slices of the abdomen with 15 mm slice distance) and once with a denser placement of the slices (6 coronal slices and 6 transversal slices also with 15 mm slice distance).

Table 3.1 summarizes the results. Given the fact that MR images provide lower resolution and less morphological contrast than CT scans, the overall results are competitive to the ones published by Heimann *et al.* [HWM06a]. One can see that, although the model is not the most comprehensive one, consisting of 23 data sets, it generalizes well, since there is only a small difference between the results of the leave-all-in and the leave-one-out tests.

One can also state that, if slice coverage is very sparse, as in the case with only 6 slices, i.e. 3 coronal and 3 axial slices, the on-the-fly strategy becomes superior. The number of slices also has an influence on the speed of convergence: As more slices provide more information, the number of iterations and the overall running time of the algorithm is reduced if more slices are used. Altogether, a complete hierarchical segmentation and reconstruction of a surface from a set of 2D image slices takes about 20 seconds on an Intel Core Duo CPU with 3 GB RAM.

Data	Test	Strategy	Slices	Avg err. (mm)	RMS err. (%)	Vol. err. (%)	Tanimoto err. (%)
3D	all-in	offline	-	3.2 ± 1.0	3.3 ± 4.5	2.1 ± 1.9	8.7 ± 2.9
3D	one-out	offline	-	5.9 ± 3.2	22.4 ± 28.6	2.6 ± 2.4	14.0 ± 4.0
3D	all-in	on-the-fly	-	4.3 ± 3.0	12.4 ± 22.3	1.9 ± 1.7	9.8 ± 3.8
3D	one-out	on-the-fly	-	6.6 ± 4.7	26.8 ± 32.4	3.1 ± 2.9	14.7 ± 4.9
2D	all-in	offline	12	3.8 ± 1.4	6.4 ± 13.2	1.8 ± 2.1	9.8 ± 3.7
2D	one-out	offline	12	6.6 ± 3.5	31.5 ± 24.8	2.7 ± 3.4	14.8 ± 4.2
2D	all-in	on-the-fly	12	4.8 ± 3.0	12.4 ± 22.0	1.7 ± 1.4	9.8 ± 3.2
2D	one-out	on-the-fly	12	6.2 ± 2.2	25.5 ± 22.9	2.3 ± 2.0	14.3 ± 2.7
2D	all-in	offline	6	7.9 ± 7.5	31.9 ± 33.7	2.7 ± 2.4	12.8 ± 5.0
2D	one-out	offline	6	10.1 ± 7.6	54.3 ± 15.6	3.9 ± 4.7	18.8 ± 7.9
2D	all-in	on-the-fly	6	6.8 ± 3.2	33.2 ± 28.3	3.0 ± 3.0	13.0 ± 5.6
2D	one-out	on-the-fly	6	8.7 ± 3.5	51.8 ± 27.2	4.0 ± 3.8	17.9 ± 6.3

Table 3.1: Validation Results: Average errors and standard deviations. All tests were done with 12 model modes and 12 resp. 6 image slices, half of them coronal the other half transversal. The column 'Test' specifies the kind of the respective validation test, which is either leave-all-in or leave-one-out. The errors are given in mm resp. percent. The calculation times range from 20 to 30 seconds on a Intel 3.2 Ghz dual core machine with 3 GB RAM.



a) b) c)

Figure 3.13: a), b) and c) show the result of the active shape model search with 6 image slices and 12 modes. The segmentation result resp. the surface reconstructed is delineated by the wire frame model.

The method and its results have been published in [FTS08b].

3.7 Discussion

From table 3.1, one can conclude that the overall performance of the method is competitive to the results presented by Heimann *et al.*[HWM06a], given the lower resolution and weaker morphological contrasts of the training MR images, compared to an average CT scan. For 3D

image data and not so sparse 2D image slices, the offline strategy is superior and also faster, while for very sparse 2D data, the on-the-fly strategy yields slightly better results. However, this conclusion would have to be confirmed by further experiments, e.g. with a different model on different input images.

The method is well suited for the problem of reconstructing the 3D position, orientation, shape and extent of the liver from a set of sparsely covering image slices, within the described error boundaries. Its results provide a good basis for follow-up automatic applications like scan planning on localizer images, etc. An example application and prototype implementation for fully automatic liver scan planning including slice and navigator placement for a MR liver exam can for example be found in [FTS06].

For high-quality clinical liver segmentations of the liver, however, more generalization and extension abilities of the model towards unseen data sets during the search would have to be implemented and integrated. Relaxing the model towards the model space after each iteration step, as mentioned in [HWM06a] instead of simply performing a strict subspace shape model projection could provide further generalization ability. However, this could interfere with the subspace projection calculations of the model as mentioned, which are necessary in the 2D case. A more comprehensive model with more generalizability created by extending the data basis and including more training shapes would certainly help in terms of precision. Nevertheless, the presented results can be considered as good pre-segmentations, e.g. for volumetric estimations, since the overall volume error is quite small. Still a lot of applications can be developed based on the given statistical shape model and the active shape model algorithm, including pre-segmentation, localization and positioning tasks.

Further algorithmic improvement, though, may still enhance the results e.g. by integrating ideas like optimal appearance features, as published by van Ginneken *et al.* [vGFS$^+$02]. There are also plans to implement a different algorithm for the initialization of the model using robust image features and integrating them into the shape model. Doing so would also allow for not starting with the average shape but with an initial shape guess based on these features. This could be achieved by joining the feature values with the statistical shape parameters prior to the PCA analysis and exploiting statistical correlations between the feature positions and the shape parameters resp. feature values.

4

Anatomical Labeling of FastView Images Using a Statistical Atlas

4.1 Introduction and Motivation

MR examination workflows are complex and require a lot of human interaction, like positioning, coil selection, sequence tuning, etc. Advanced knowledge and a lot of experience is necessary to acquire high quality image data in every case. Reducing human interaction e.g. by automation of certain workflow steps or suggestion of reasonable scan parameters would help inexperienced operators to achieve high quality examinations. For this reason there is growing need for more automation and standardization. This would also lead to more reproducibility, better comparability of the examinations and to shorter examination slot times and hence higher scanner workloads, as well. Methods that use image information from fast non-diagnostic pre-scan acquisitions for the automation of otherwise manual steps in the MR examination workflow are highly appreciated. To name but a few of these methods, van der Kouwe *et al.* [vdKBF+05] presented a solution for reproducible MR head scan planning using an atlas of the human brain. Peschl [PESH04] describe a method for fast and virtually automatic scan planning for the spine. The previous chapter delineated a method that would allow for fully automatic liver scan planning from fast localizer scans, which had been published in [FTS06] and [FTS08a]. All the methods mentioned share the drawback that they all depend on special acquisition protocols, which require adequate coil selections and pre-positioning of the patient in a way that the region including the organ which is to be scanned and for which the scan range is to be planned is within the field of view of the scanner. As the homogeneous volume of the magnet is usually in the range of 40-50 cm, it is not possible to position for various organs within a single scan.

Recent developments in the hardware of MR scanners have brought up continuous table imaging, as described in 2.5.5. Whereas in multi-stage MR image acquisitions, whole-body images are scanned as blocks of 3D volume which are retrospectively stitched together, the idea of continuous table imaging is to scan a whole body image by moving the patient table continuously in axial direction during the acquisition. This concept has been realized by various groups like Barkhausen *et al.* [BQL+01] and Kinner *et al.* [KZZ+06]. In continuous table imaging, the slice which is currently scanned is always located at the isocenter of the magnet. This allows for acquiring whole-body images within a single scan without repositioning or scanning in stages. This means that the complete image of a patient or of a torso of a patient is

available without composing or stitching of image stages, which also eliminates the need for overlaps, i.e. scanning certain regions twice for robust stitching. The advantages of continuous table imaging have already been summarized in section 2.5.5. By scanning 2D slices during continuous movement of the table, full 3D image information is available. If the scanning protocol is fast enough, the image data can be used as a 3D localizer.

One of the protocols making use of continuous table imaging is FastView [Bla06]. FastView is a FLASH [HFM+86] based (as described in subsection 2.5.3), proton-density weighted, 2D axial acquisition technique with the opposed phase condition of fat and water signal. Opposed phase [HBTV99] means, that the complex signals generated by fat and water have a phase shift of $180°$ at the time of the rephasing of the precessing protons, which means that their vectors point in opposite directions. This results in a complete extinction of signal for voxels containing equal amounts of fat and water, which in turn causes strong discernible boundaries of various tissues. Strong boundary information and contrast can be helpful for image processing applications, so FastView images, provide good input for postprocessing applications, although their resolution is only moderate.

The FastView protocol can be perfectly used as a 3D localizer, as it provides full 3D information, large field of view, sufficient resolution (5 mm isotropic) and fast acquisition (about 20 seconds). This allows for using FastView image data for image-based automatic labeling of multiple organs or anatomical structures from the entire body or torso from one single fast non-diagnostic prescan acquisition. Labeling of anatomical structures is equivalent to a multi-class segmentation of the anatomical structures from the image data. This information can be used for scan automation like automatic slice positioning, navigator positioning, SAR value estimations, or adaptations of specific sequence and scan protocol parameter for a patient (fat saturation, resolution), etc. Other applications of the method could include automatic pre-segmentations, croppings for visualization, volumetric estimations, etc. In short terms, the FastView localizer opens up vast opportunities for image-based fully automatic whole-body scan planning, reporting and processing algorithms.

The approach, which was developed in this thesis for the multi-organ segmentation problem is based on the creation of an atlas of the human torso from a set of FastView training images. An atlas contains information about the complete spatial relationships of pixel values of certain anatomical structures [RBM+05]. The atlas therefore represents an anatomical template which can be used to align an unseen image. In the majority of cases, the template is an average instance, created from a group of individuals. However, the atlas can also be a special individual instance depending on the application. The spatial alignment can then be used to infer information on the structures of the unseen image from the well-known structures of the atlas. In particular for the given application, anatomical structures are labeled manually in the atlas of the human torso. Registering an unseen FastView image to the atlas and propagating the anatomical labels then yields the anatomical labeling or multi-class segmentation of the unseen FastView image.

The atlas of the human torso used in this thesis is created as a statistical atlas. This means that it does not only contain an average instance of the human torso but also its main modes of statistical variation, which are obtained from principal component analysis. Statistical variation of the atlas refers to both variation of the shapes and spatial relations which are represented by deformation fields, but also grey value variation. The deformation field statistics capture

naturally occurring variations in shapes and spatial relations in comparison to the average torso represented by deformation fields. Grey value statistics can be used to describe the variation of grey value appearance in different data sets caused by effects like different proton densities resp. relaxation times of the organs among other effects.

4.2 Related Work

Atlas based methods are popular throughout medical image segmentation literature. A thorough overview of existing techniques and publications can be found in a review by Rohlfing *et al.* [RBM+05]. Most publications dealing with anatomical atlases have been made in the context of MR brain imaging and diagnosis. Talairach presents an atlas and complete coordinate system for the human brain [TT88] defined on an atlas created from a brain image of 60-year old left-handed female. Gee *et al.*[GRB93] use a human brain atlas from MR images for matching deformed instances by elastic warpings. The on-line positioning method in [vdKBF+05] is also based on an atlas of the human head. Cootes introduces active shape models and active appearance models and published their application to atlas based matching in [CBET99]. Rueckert proposes statistical analysis of non-rigid deformation fields for the creation of a statistical atlas [RFS03] of the brain, which can be used for brain segmentation.

In the field of whole-body atlas based segmentations, Park *et al.* [PBM03] use a probabilistic abdominal atlas for supervised multi-organ segmentation from 32 abdominal CT-scans. They use thin-plate-splines transformations for the deformable transformations between the different data sets. A Bayesian framework is used for assigning segmentation class labels to the voxels of unseen image data. Zhou and Bai [ZB05] create an abdominal atlas from CT scans. They use atlas-based registration for an initial segmentation of the abdominal organs. The initial segmentations are then fine-tuned by means of a novel connectedness approach based on fuzzy functions.

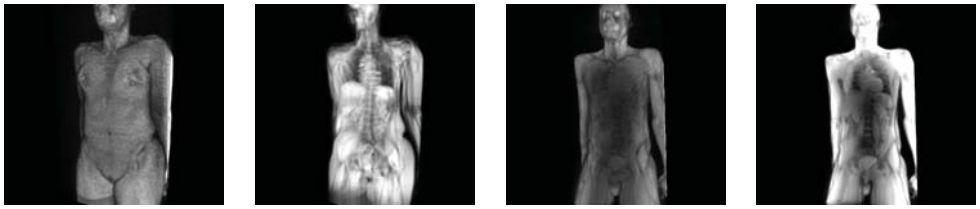
4.3 Building the Statistical Atlas

This section covers the various steps which are necessary for creating the statistical atlas of the human torso from FastView training images. The method presented in this section is closely related to the concepts presented by Rueckert *et al.*[RFS03]. First, a representative group of individuals is scanned using the FastView protocol. Then one data set of this group is selected as the reference image. By performing non-rigid registrations of all data sets to the reference image, a set of deformation fields is obtained. Applying principal component analysis on the deformation fields yields the average deformation field and its eigen modes of variation. Performing principal component analysis on the pixel values of the registered FastView images of the individuals produces the average image and its eigen modes of variation. Backwarping the average image and its eigen modes by the inverse average deformation then yields the actual atlas in its natural coordinate system. For reasons of simplicity and generalizability, a single atlas containing both male and female individuals is created. The selection of individuals for the creation of the atlas was done in a way that a possibly vast variety of statures, size and ages are taken into account in order to obtain representative statistics.

4.3.1 FastView Image Data

The data basis of the atlas is set up from the scans of a group of representative volunteers using the FastView protocol. FastView, as mentioned above, is a fast FLASH-based protocol that produces 3D images with a proton density weighting during continuous movement of the patient table. The FoV was fixed at 1000 mm, so that a huge portion of the volunteers' bodies, including head, torso, pelvis and a part of the upper legs is covered. Running the scan protocol with these parameters takes about 20 seconds per individual.

Selection of the individuals is done in way to cover a preferably huge variety of height, stature, age and sex. The image data are also acquired on different scanner devices with different basic magnetic field strengths B_0, $1.5T$ resp. $3T$, different bore dimensions, different coils, etc in order to capture all potential variations of the image data. Figure 4.1 shows two example FastView data sets in direct volume rendering.



a) b) c) d)

Figure 4.1: FastView images from two individuals in direct volume rendering. a) and c) with large opacity values, b) and d) with a small opacity values.

4.3.2 Arm Stripping

As the images are acquired with the patients in supine position and no additional coils apart from the body coil [HBTV99] are used which might constrain the patient positions, the variation of arm configurations is large. Since differing arm positions can affect the following registration process in an undefined way and since this would add additional undesired variation and complexity to the image data, pure torso images are created by stripping off the arms using an automatic algorithm.

The arm stripping algorithm is motivated by [GPAK08] but more intricate in its implementation. The first step is to make a segmentation of the torso including the arms using inverse region growing from outside with an upper noise threshold, which yields a binary image. Each axial slice of the binary image is analyzed and clustered into connected components. The components of the slices containing 3 components are divided into left arm component, torso component and right arm component.

The components of successive slices are then linked to the components of the predecessor slices in a way that a metric of distance and difference in size is minimized. Missing components are added by interpolating the components between adjacent slices. The clustered and interpolated components of left and right arm are connected to tubular structures. The surfaces of tubular structures are used as the initialization of a 3D active contour model or

3D snake [KWT88], that snaps to the nearest edges of the image data, while maintaining a smoothness condition. The snake is encoded as a level set [CKS95]. Zeroing the voxels inside the snake then results in a FastView image without arms.

See figure 4.2 for an explanation of the concept of the arm stripping algorithm.

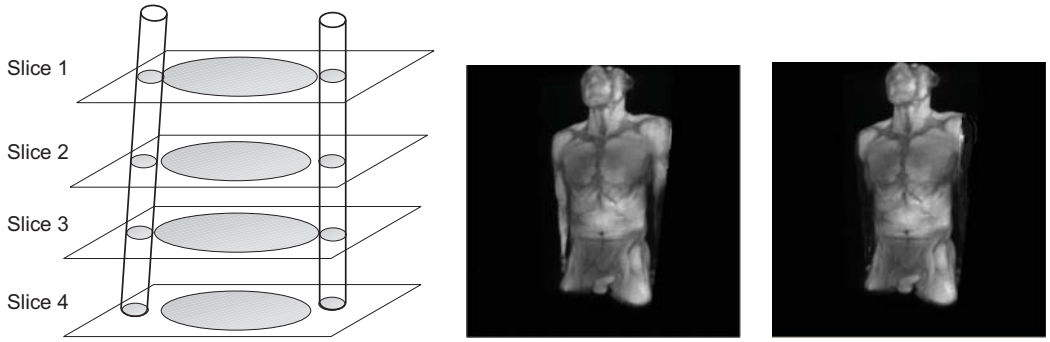


Figure 4.2: a) shows the clustering step of the arm stripping algorithm. In slice 4 and slice 2 all components are separated. The tubular structure results from connecting the components which are truly separated and interpolating between. b) and c) show an example of a FastView image with arms and a FastView image without arms after application of the arm stripping.

4.3.3 Image Histogram Normalization

As grey values in MR images have no immediate physical meaning such as Hounsfield units have for CT, which was already explained in section 2.2, the grey value appearance of the images can depend on a lot of parameters, like magnets, coils, signal encoding, etc. and thus cause undesired variation of the image data. For statistical considerations, it is therefore important to exclude some of those parameters by an initial image normalization.

A popular normalization method is the one published by Nyul *et al.* [NU99] that maps the histograms based on normalization parameters which are learned from input histograms. These normalization parameters are then used for a piecewise linear scaling of the image histogram to the ideal histogram.

For the FastView image atlas, a simplified normalization method is used, which works by scaling the image grey values in a way that the peak of the non-background component of the two-modal histograms coincide. This is motivated by the fact that the statistical distributions of the foreground voxels of the torsos should be equalized. Normalizing the mean values of the foreground voxels reduces a large amount of variation of the grey value data. The intensities between the peaks are adjusted by a linear scaling defined by the peak values which are to be matched. The algorithm requires the histograms to be segmented into foreground and background components which can e.g. be done by expectation-maximization techniques [GW02] or simpler and sufficient for this case by thresholding the histograms using the mean intensity of the voxels [NU99].

Figure 4.3.3 shows how the normalization is done.

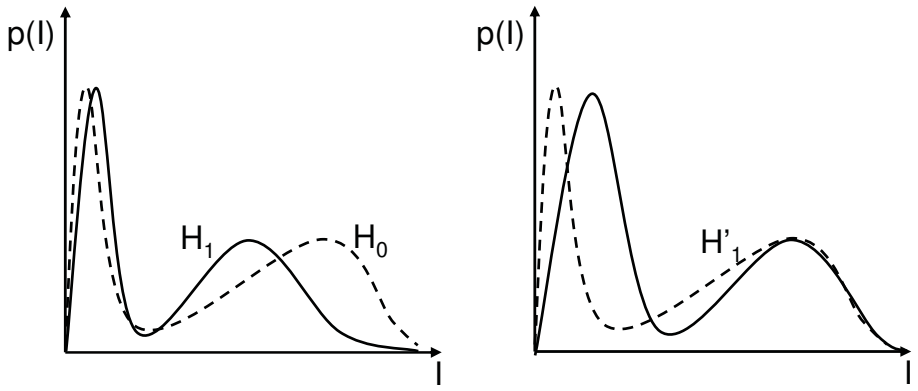


Figure 4.3: The histogram H_1 shown with solid lines is normalized by a scaling factor according to the reference histogram H_0 shown with dashed lines, such that $max(p(I_1')) = max(p(I_0))$. The x-axes of the histograms represent the intensity values I, the y-axes show the corresponding probabilities $p(I)$

4.3.4 Image Data Registration

The crucial step for building the atlas is finding spatially corresponding points for each pair of training images across the collection of volunteer image data. This is done implicitly by performing registrations of all data sets to a reference data set. The reference data set is a randomly selected image data set with good overall image quality of an individual with preferably average anatomy. Image registration produces spatial correspondence by maximizing the similarity between two images. Image registration is an optimization procedure in which, one image is defined to be the fixed image, while the moving image is transformed and resampled in a way that maximizes similarity between the two. This is typically done by iteratively warping the moving image according to a geometrical transformation, computing a similarity value between the images, then updating the transformation in a way that the similarity is increased and so on. The final transformation which maximizes the similarity of the moving image to the fixed image is called the registration result.

If T_i is a transformation that registers image I_i to image I_{ref} using backward warping, then $T_i(x)$ and x are corresponding points or pseudo-landmarks in image I_i and I_{ref}. Backward warping of the moving image means that the grey value for each target pixel in the image is computed by transforming its spatial coordinates according to the warping transformation and looking 'back' at the grey value at the original position in the unwarped image using an interpolation scheme. The advantage of using backward warping is that no undersampling or holes can occur in the warped image, although it is conceptually less straightforward.

Depending on the degrees of freedom that the expected transformations can have, the transformation of the registration can be set up in different ways. Because of the huge variation of inter-individual whole-body anatomy, the transformation for the registrations in this context are

assumed to be free-form non-rigid transformations with reasonable constraints. The complete registration transformation T_i is therefore set up by a global affine transformation T_{global} and a local free-form transformation T_{local}. The complete transformation T_i is then a concatenation of the two: $T_i \leftarrow T_{local} \circ T_{global}$.

Affine Preregistration

Before the actual non-rigid registrations can be computed, all data sets have to be transformed to a common coordinate frame to compensate for different coordinate systems of the acquired data sets and systematic offsets. This helps to obtain reasonable starting points for the registrations.

To this end, manual landmark points are placed at discernible anatomical correspondence points on the surfaces of the torso data sets. The discernible points, for example comprise the tip of the nose, the hips, etc. Each data set is then warped to the common coordinate frame of the reference image by applying an affine transformation T_A. T_A is a transformation which minimizes the sum of squared Euclidean distances between the transformed landmarks l_i and the landmarks in the reference data set l_i^r.

$$T_A = \arg\min_{T_A} \sum_i ||l_i^r - T_A(l_i)||^2 \tag{4.1}$$

The affine transformation is calculated from the correspondence points as described in section A.3.

Non-rigid Transformation

There are many ways and methods for computing non-rigid registration and warping transformations to maximize similarity between images. A collection of approaches can be found in a thorough coverage by Modersitzki [Mod04]. This publication describes free-form deformations with dense deformation vector fields and regularizers which are based on physical diffusion, curvature, elastic or flow models. Dense deformation vector fields means that for each voxel in the image, a transformation vector is computed which maps to a spatial position in the reference image. Since this is an ill-posed problem, a regularization has to be done which actually makes the difference between the several methods. In short terms, the regularizer constrains the problem by imposing physically motivated conditions, e.g. by punishing large differences in the vectors of neighboring voxels.

Another kind of free-form deformation transformations is based on B-Spline tensor product deformation fields. This is a powerful tool for modeling 3D deformable objects. The key idea of this method is to superimpose a regular grid of $M \times N \times K$ control points onto an object of dimensions $(x_{max}, y_{max}, z_{max})$, e.g. the volume of a 3D image, and to deform the object by moving the underlying control points. So each control point (i, j, k) is assigned a deformation vector $d_{i,j,k}$ and the deformation $T_{local}(x, y, z)$ at an arbitrary point (x, y, z) can be described as the tensor product of the 1D B-Spline curves that weight the deformation of the surrounding control points [RFS03].

$$T_{local}(x,y,z) = \sum_{l=0}^{2}\sum_{m=0}^{2}\sum_{n=0}^{2} B_l(u)B_m(v)B_n(w)d_{i+l,j+m,k+n} \tag{4.2}$$

where

$$u = \frac{x \cdot M}{x_{max}} - \lfloor \frac{x \cdot M}{x_{max}} \rfloor \tag{4.3}$$

$$v = \frac{y \cdot N}{y_{max}} - \lfloor \frac{y \cdot N}{y_{max}} \rfloor \tag{4.4}$$

$$w = \frac{z \cdot K}{z_{max}} - \lfloor \frac{z \cdot K}{z_{max}} \rfloor \tag{4.5}$$

and

$$i = \lfloor \frac{x \cdot M}{x_{max}} \rfloor \tag{4.6}$$

$$j = \lfloor \frac{y \cdot N}{y_{max}} \rfloor \tag{4.7}$$

$$k = \lfloor \frac{z \cdot K}{z_{max}} \rfloor \tag{4.8}$$

and the B-Spline coefficients for a cubic B-Spline are defined as follows

$$B_l(u) = B(u-l-1) = B(x) = \begin{cases} \frac{(2-|x|)^3}{6} & 2 > |x| \geq 1 \\ \frac{(4-6|x|^2+3|x|^3)}{6} & 1 > |x| \geq 0 \\ 0 & elsewhere \end{cases} \tag{4.9}$$

The B-Spline tensor based transformation ensures both smoothness of the transformation and strict local support of the parameters. Both of these properties are required for modeling variations of anatomy. In addition the number of degrees of freedom can be directly controlled by the number of control points and is not related to the resolution of the images, in contrast to registration transformations using dense vector fields. The feature of strict local support of B-Spline tensor based transformations is due to the fact, that changing one coefficient of the deformation of an arbitrary control grid point only affects a $4 \times 4 \times 4$ area of grid points around that control grid point. See figure 4.4 for an example of a 3D control grid superimposed on a 3D image volume and a warping example of an image based on the underlying grid.

Similarity Measure

In order to assess the quality of an image registration, a similarity measure is needed that makes a quantitative statement on how well two images are aligned and how well the spatial correspondence is. An overview of common similarity measures can be found in [HHH01], including summed squares of differences (SSD), cross-correlation or histogram based measures.

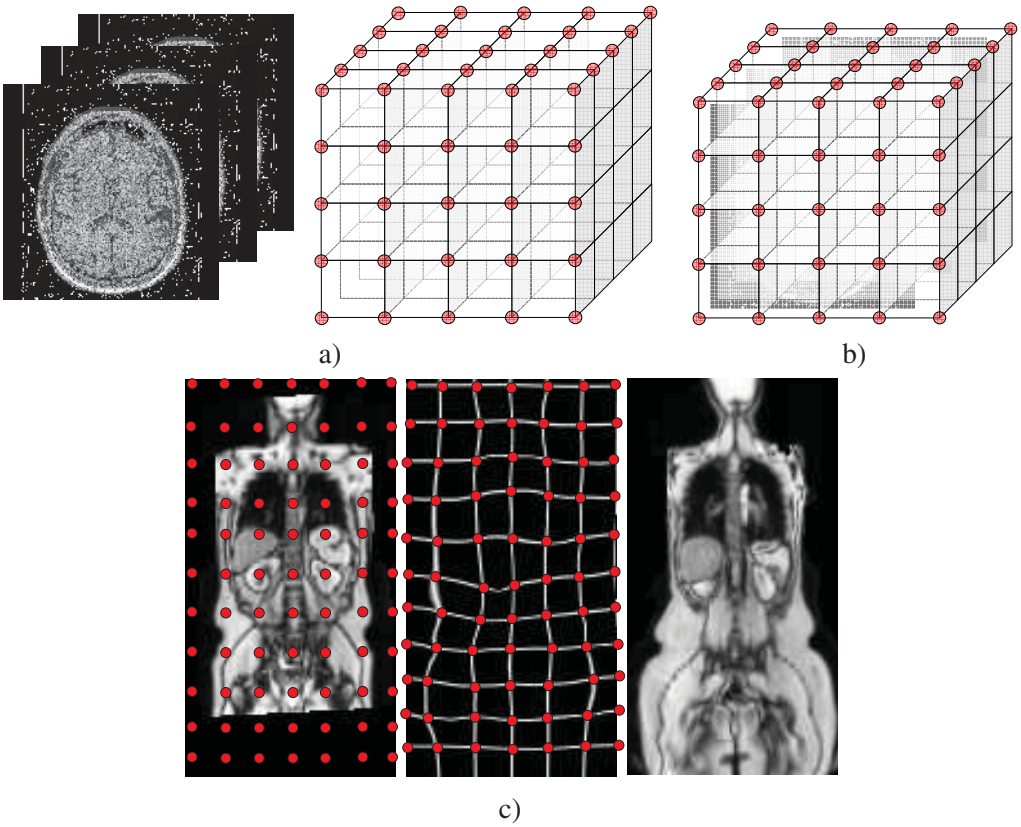


Figure 4.4: a) shows an image volume represented by a stack of slices and the regular grid of control points. b) shows how the grid is superimposed onto the image volume. c) shows a warping example of a 2D slice, with the left showing one slice with superimposed grid points, the center image showing a warped regular grid with a reference grid image for more clarity and the right the original image warped by the deformation field.

A common histogram based measure is mutual information [PMV03] which is frequently used in multi-modality and multi-individual registration applications. While SSD and cross-correlation are calculated directly in image space, i.e. directly on the grey values of the image, mutual information is calculated from a feature space defined on the two images. For mutual information, the feature space is the joint histogram of the two images, which is subject to change as the alignment of the images changes. The joint histogram can be calculated by counting the joint occurrence frequencies $p(i, j)$ of the each pair of pixel values i and j. Mutual information is now calculated from the joint histogram based on Shannon's entropy [Sha48] which was introduced in information theory and describes the amount of uncertainty of a given piece of information.

More specifically, if $p(i, j)$ are the frequencies in the joint histogram of grey value i in image A being aligned with grey value j in image B, then Shannon's entropy H of this joint distribution is:

$$H(A,B) = -\sum_{i,j} p(i,j) \log p(i,j) \tag{4.10}$$

The marginal entropies $H(A)$ resp. $H(B)$ are computed by summing up each column and each row of the joint histogram and calculating the entropy on those.

$$H(A) = -\sum_{i} (\sum_{j} p(i,j)) \log (\sum_{j} p(i,j)) \tag{4.11}$$

Mutual information between two images A and B is defined as the difference of the joint entropy and the sum of the marginal entropies. The formula is:

$$I(A,B) = H(B) - H(B|A) = H(A) + H(B) - H(A,B) \tag{4.12}$$

This finally boils down to [PMV03]:

$$I(A,B) = \sum_{i,j} p(i,j) \log \frac{p(i,j)}{p(i)p(j)} \tag{4.13}$$

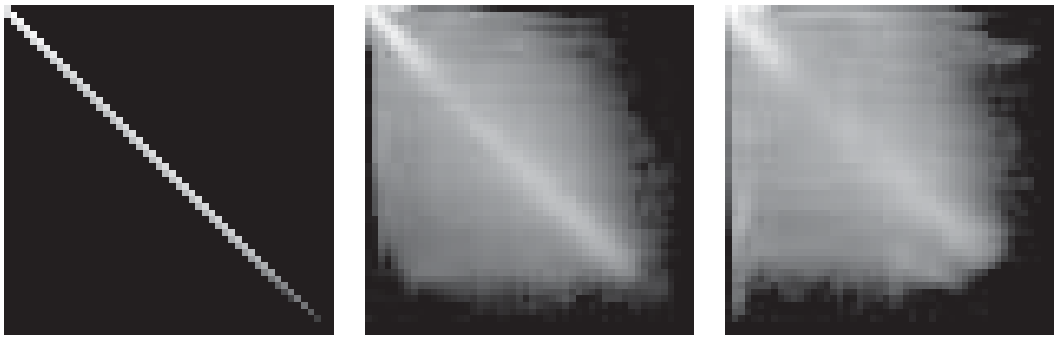
In short terms, mutual information makes a quantitative statement on the amount of certainty one has about the grey values in image B, given the grey values of image A under the given alignment. Hence, maximizing mutual information means sharpening the distributions in the joint histogram, because the sharper the clusters in the histogram, the more certainty there is about the joint intensity pairs. Figure 4.5 shows examples of sharp and unsharp joint histograms and the relation to changing similarity of moving and fixed images.

The huge advantage of using mutual information which makes it superior to other similarity measures for multi-modality or inter-subject registrations is that it allows for non-linear variation of contrast between different subjects which is particularly important in cross-subject MR image registration, where different relaxation times or proton densities make similar structures appear at different grey values. The drawback is that its expensive computation. However, since the registrations are only done once during the creation of the atlas, this shortcoming can be accepted.

Since the mutual information term is sensitive to the size of the image overlap, usually a normalized mutual information term is used, as explained by [PMV03].

$$NMI(A,B) = \frac{H(A) + H(B)}{H(A,B)} \tag{4.14}$$

The implementation of the mutual information similarity function for the creation of the atlas is motivated by the paper by Thevenaz and Unser [TU00]. They use Parzen windowing [Par62] for the calculation of the joint histograms. This results in smoother and continuous estimations of the joint probabilities and facilitates the computation of direct derivatives of the metric [TU00]. The idea behind Parzen windowing is not to assign binary values to the frequency bins in the joint histogram but to spread the information over neighboring bins using a kernel function. This avoids discontinuities in the case of subtle changes of grey values, which



a) b) c)

Figure 4.5: a) shows the joint histogram of two perfectly aligned volumes, b) after a rotation of 10 degrees and c) after a rotation of 20 degrees. The larger the rotation and the spatial misalignment between the two volumes, the unsharper the joint histogram becomes and the smaller the value of mutual information.

is important to the computation of derivatives. See figure 4.6 for a visualization of Parzen windowing. [TU00] suggest to use B-Spline kernel functions for the Parzen windows. B-Splines have some useful properties, one of them being the partition of unity, which means that the binary sample is partitioned into fractions which are spread over the bins and the values of which sum up to 1.

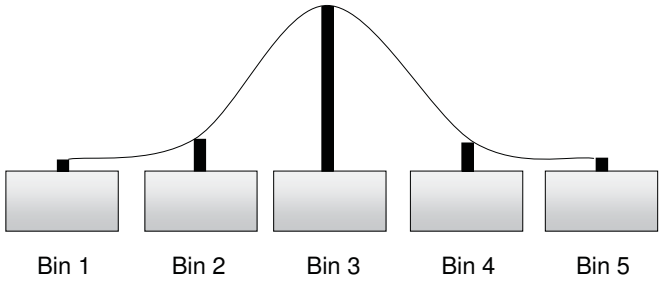


Figure 4.6: The sample value is distributed to the different bins according to their distances to the bins. A spline kernel function is used to obtain the weights of the fractional frequencies, which all sum up to 1.

The mutual information term is integrated into a combined cost function. Minimization of this cost function drives the registration process. For the optimal registration, the cost function will reach its minimum value. In order to avoid unnatural or rupturing deformations an additional regularization term sums up the squared norm of differences of adjacent control point shifts v_i. This regularization term prevents neighboring control grid points from moving too close or too far from each other. It also avoids overfolding of the control points, which can happen in theory if no regularization is used. The regularization term is similar to the so-called diffusion regularizer [Mod04] in free-form non-rigid registration. Computing its derivative is

straightforward and can be implemented efficiently. The cost function also contains a term for considering predefined correspondences, so called manual landmarks LM. If certain structures need to be aligned beforehand to steer the registration process towards a desired solution, this can be done by placing landmarks at the corresponding structures in the moving image and the fixed image. So the combined cost function reads as follows with the empirical weighting constants λ and γ:

$$c(T) = -NMI + \lambda \sum_{(i,j)\in NB} (v_i - v_j)^2 + \gamma \sum_{x_i \in LM} \left(x_i' - T(x_i)\right)^2 \tag{4.15}$$

where T is the non-rigid deformation, NB is the 27-neighborhood on the grid of control points v_i is the movement of control point i, LM is the set of all manual landmarks, x_i' is the landmark in the reference image and x_i the corresponding landmark in the moving image. The term for normalized mutual information NMI is subtracted because, thus, increasing mutual information is correlated to decreasing the cost function and vice versa. λ and γ are empirical coefficients that weight the different terms. The coefficients are determined experimentally and globally for all data sets so that the effects by the different terms have appropriate relative magnitudes for the cost function.

Optimization

The optimal similarity between the two images according to the given criteria is found when the upper cost function is minimized. That is the case if and only if the following two conditions hold:

$$\nabla c(T_{opt}) = 0 \tag{4.16}$$

$$c(T_{opt}) \leq c(T)\ \forall T \tag{4.17}$$

A value which fulfills condition 4.16 can be computed using a gradient descent approach [VF02]. Starting from an initial transformation T_0, the gradient of the cost function $\nabla c(T_0)$ is calculated at this position and then the transformation is updated by moving into the negative direction of the gradient using a step size Δs.

$$T \leftarrow \begin{cases} T - \nabla c(T) \cdot \Delta s & if\ c(T - \nabla c(T) \cdot \Delta s)) < c(T) \\ T & else \end{cases} \tag{4.18}$$

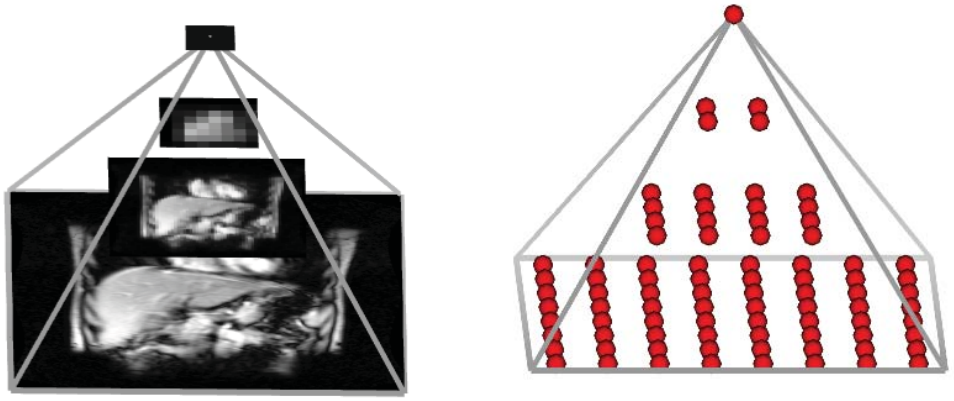
In order to increase the speed of convergence, a strategy for adaptive step size is used. This means that Δs is a function of the iteration number resp. time t, i.e. $s \leftarrow \Delta s(t)$.

$$\Delta s(t) \leftarrow \begin{cases} f \cdot \Delta s(t-1) & if\ c(T - \nabla c(T) \cdot \Delta s) < c(T) \\ \frac{1}{f} \cdot \Delta s(t-1) & else \end{cases} \tag{4.19}$$

with f being a constant and $f > 1$.

This procedure increases the step length as long as the cost function value decreases steadily, but decreases its step length if it does not.

The second condition 4.17 is harder to fulfill as the cost function is multi-modal and has many local side optima. In order to avoid becoming stuck in local optima, a hierarchical multi-scale approach is implemented. A Gaussian pyramid [GW02] is built from the input image by successive low-pass filtering and downsampling. The control grid is also implemented as a pyramid of control grids with different resolutions. Figure 4.7 shows the resolution pyramids of the images and the control grid. Hence, the registration starts with the coarsest image resolution of the pyramid and the coarsest grid from the pyramid of control points. After convergence of the registration optimization, the next finer resolution of the control grid is used. As soon as this registration setup has converged, the next finer image resolution is taken and so on. The final registration result is obtained from both the finest image and finest grid resolution. In the finest grid resolution, there is one control point for every four voxels. This means that the maximum grid resolution is 20 mm.



a) b)

Figure 4.7: Visualization of the Gaussian image pyramid (a)) and the pyramid of control grid points (b)).

Results

The images in figure 4.8 show the results of the registration process for three different data sets after checkerboard image fusion. Checkerboard fusion combines two images to one by selecting image patches in an alternating fashion from the two inputs like the black and white fields on a checkerboard. In a perfect registration, no discontinuities of the structures, except for different grey value levels of the individuals, should be visible at the borders of the checkerboard fields.

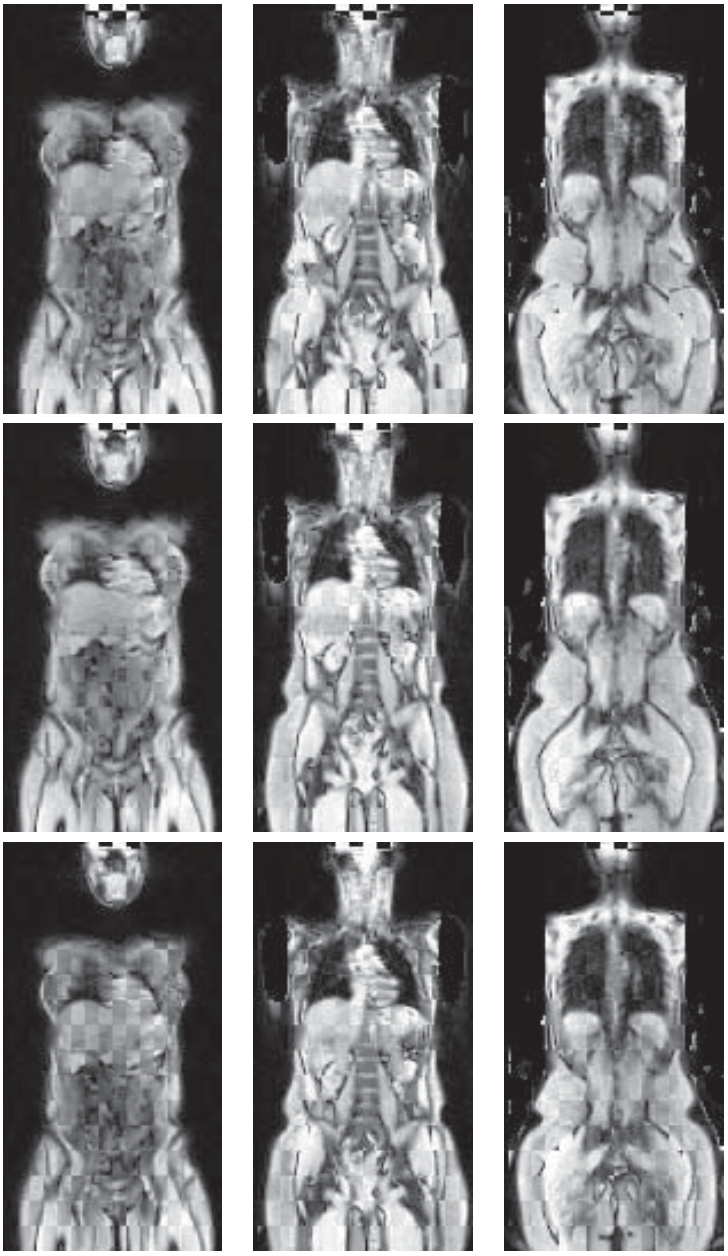


Figure 4.8: This figure shows the results of the registrations to the selected reference image instance after checkerboard fusion, which means that alternating patches from the moving image and the fixed image are combined to a single image. The first column shows slice 20, the second slice 30 and the third column slice 40 of the registered image. Each row shows the result of a different data set.

4.3.5 Statistical Model of Deformation

Statistical models of deformation [RFS03] are an extension to standard active shape models as proposed by [CT92] and as described in chapter 3. Instead of doing statistical analysis on the large vector of the coordinates of corresponding points, principal component analysis is directly applied to the underlying deformation fields. So, let d_i be the large vector of the movements of all control grid points, which set up the transformation that warps image I_i to the reference image I_r. The average deformation field can then be calculated according to:

$$\bar{d} = \frac{1}{n} \sum_{i}^{n} d_i \tag{4.20}$$

and the covariance matrix S

$$S = \frac{1}{n-1} \sum_{i}^{n} (d_i - \bar{d})(d_i - \bar{d})^T \tag{4.21}$$

Applying principal component analysis (PCA) to the covariance matrix yields the principal components of deformation. Its linear model can be written as follows:

$$d = \bar{d} + \Phi_d \cdot b_d \tag{4.22}$$

where Φ_d is the matrix of principal components and b_d the feature values of deformation.

Varying the parameters of the feature vector b_d then generates all valid instances of deformation fields with the underlying assumption of the space of deformation fields being linear.

Before the deformation fields can be analyzed, it has to be made sure that any global resp. affine content is removed so that only true variation between the subjects is taken into account and that biases like different positions, orientations and sizes of the subjects are excluded from the considerations. To this end, all affine content is removed from the deformation fields by performing a least squares fit of an affine transformation to the deformation field. This content is removed from the B-Spline transformations and the statistical analysis is made on those deformation fields which have been 'cleaned' from affine biases.

The atlas image a is then generated by pixelwise averaging the registered and warped images I_i. However, the atlas which has been generated this way depends on the choice of the reference image, which introduces an undesired bias. Even though the reference image may seem normal, it may represent some abnormal anatomy. See figure 4.9 for an explanation. The atlas is supposed to be independent of the choice of the reference. So in an additional step, the atlas has to be computed in its natural coordinates, which are the coordinates that minimize the overall sum of deformations over all registrations [RFS03]. The natural coordinates represent the true average of the data sets instead of the average of all images registered to an arbitrary reference.

There are different approaches for computing the true average resp. the atlas in natural coordinates in medical imaging literature dealing with anatomical atlases.

Rohlfing *et al.*[RBJM01] suggest a method of iteratively registering all images to the average image. This gradually converges to the unbiased average in its natural coordinate system. In the first iteration, one data set is chosen as the reference but only affine transformations

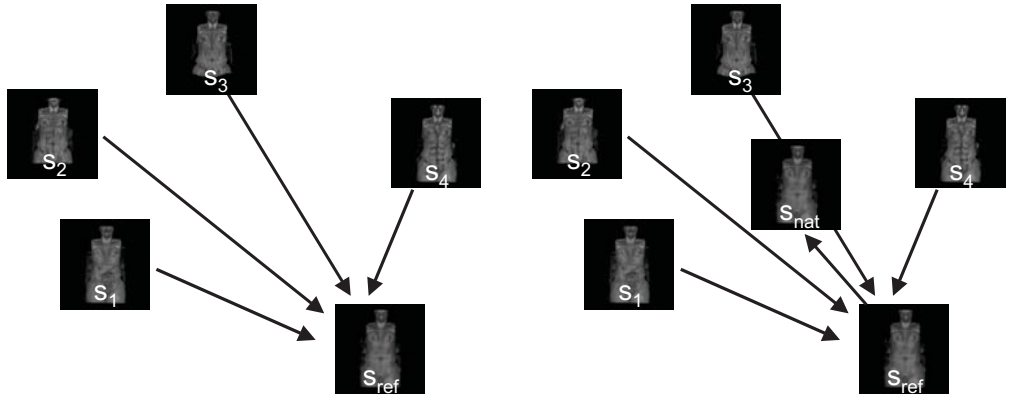


Figure 4.9: a) shows a couple of data sets s_i which are registered to the reference s_{ref}. After the registration, the atlas is in the coordinate system of s_{ref}. b) After backwarping by the average deformation, the atlas is in its natural coordinate systems s_{nat}, which is located at the center of all data sets. Hence, the sum of deformations is minimized.

are used to warp all other data sets to the reference. The average is computed by pixelwise averaging all warped data sets. During the iteration, all data sets are repeatedly registered to the average using affine and non-rigid transformations. This procedure converges to the actual mean atlas in its natural coordinates.

Studholme [Stu03] comes up with a population based registration approach, that makes use of simultaneous group-wise registration under the constraint of keeping the sum of all deformations zero. This means that all data sets are registered and warped at the same time while the sum of all deformations is fixed at 0. This leads to a registration of all data sets to the natural coordinate system without the need to select a special reference or template data set.

Rueckert *et al.* [RFS03] use a reference template instead and register all images to this template. The natural coordinate system is then obtained by a backwarp from the reference coordinate system by the average transform. Therefore, the atlas must be transformed to its natural coordinates x', which are defined by each point x being shifted by the average deformation field $\bar{d}$.

$$x' = x + \bar{d}(x) \tag{4.23}$$

This corresponds to a backwarp of the atlas image with the inverse average deformation field.

The latter approach is used for the implementation of the statistical atlas in this thesis. Since, as mentioned above, a linear model of the deformation field is calculated, the average deformation field is easy to compute. However, in this implementation, the atlas is not explicitly transformed to the natural coordinate system but the offset given by the average deformation is implicitly considered during the atlas based registration process. This has the advantage that no real backwarp has to performed on the image data, which either requires a,

possibly unstable, inversion of a deformation field or a forward warping step with the hazard of undersampling artifacts.

Figure 4.10 shows an overview of the statistical model of deformation and its major modes of variation. This version of the statistical atlas is created from 31 representative data sets. Although 31 individuals are a sound number for a preliminary study, it should be mentioned that adding an additional number of data sets could improve the generalization ability of the atlas.

4.3.6 Statistical Model of Appearance

In order to be able to describe the variation of the appearance of the grey value intensities of different organs in different subjects, a statistical model of grey value appearance is added to the atlas. Principal component analysis (PCA) is applied to the voxel intensity values of the registered data sets. Let a_i be the assembled vector containing the normalized intensity values, according to section 4.3.3, of all voxels of the input image I_i. Hence, the image of average voxel intensity values is given by

$$\bar{a} = \frac{1}{n} \sum_{i}^{n} a_i \tag{4.24}$$

and the covariance matrix S by

$$S = \frac{1}{n-1} \sum_{i}^{n} (a_i - \bar{a})(a_i - \bar{a})^T \tag{4.25}$$

Applying PCA to the covariance matrix yields the principal components of appearance and the linear model of appearance similar to the linear model of deformation:

$$a = \bar{a} + \Phi_a \cdot b_a \tag{4.26}$$

Varying the parameters of the feature vector b_a then generates valid instances of grey value intensity distributions of the linear model.

4.3.7 The Statistical Atlas

Combining the statistical model of deformation and the statistical model of appearance results in a powerful statistical atlas, which allows for generating arbitrary instances of FastView images. Linear combinations of the deformation parameters and the appearance parameters are capable of representing the natural variability of training individuals scanned with the FastView protocol. The more representative the training group, the more capable is the atlas to generalize to all possibly occurring instances across the population.

In contrast to active appearance models [CET98], however, it was decided to keep the appearance of the data independent of the shape and geometry resp. deformation model and not to correlate the two. This is motivated by the fact that the appearance of intensities of an image may be strongly influenced by external effects like the actual MR scanner device, its basic magnetic field, hardware, etc. and not only by the proton densities of the actual

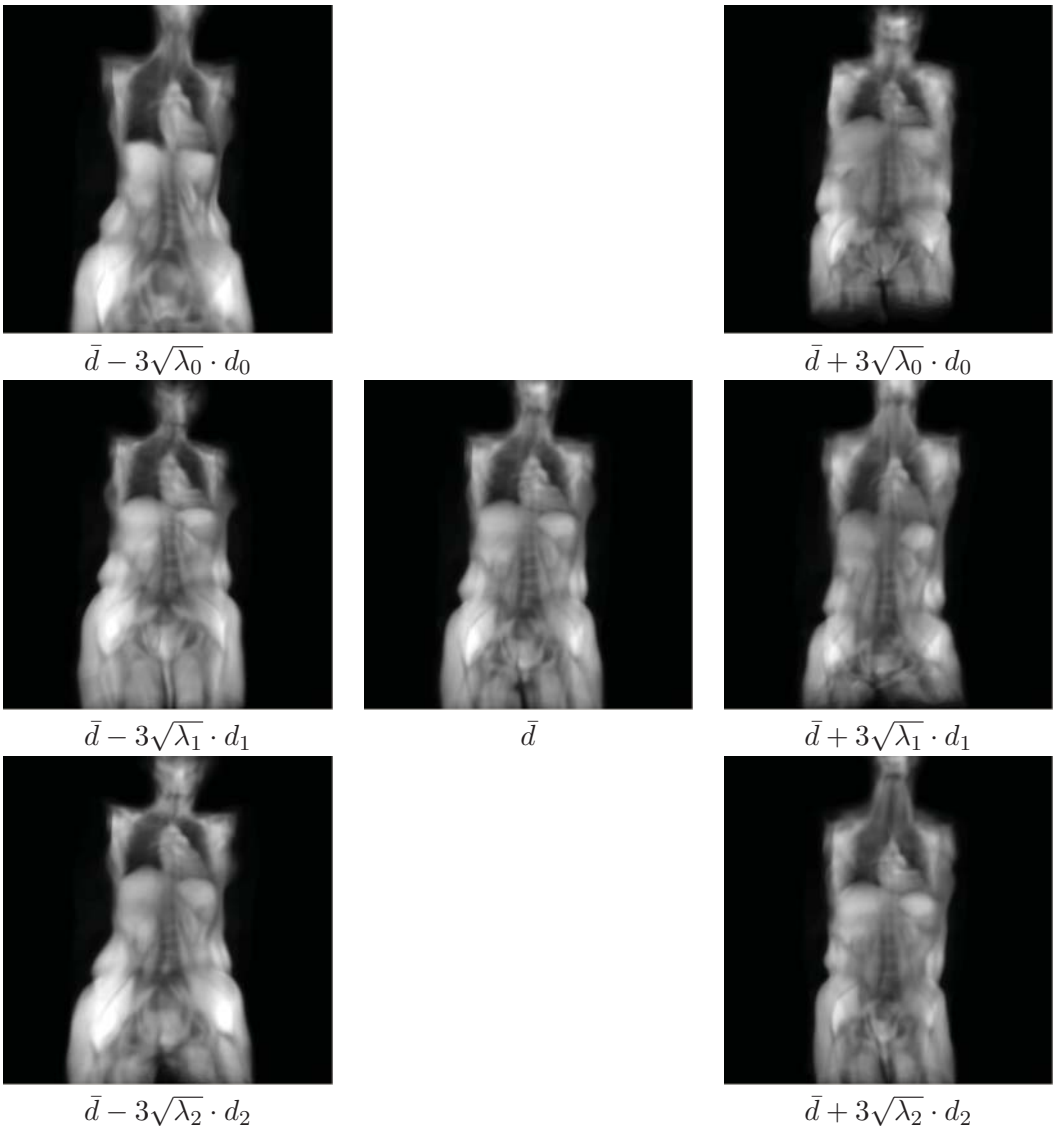


Figure 4.10: This figure presents the statistical model of deformation of the atlas and its major modes of variation. The center image shows the average instance of deformation. Each row shows a different mode of the deformation, with the left and right column being the outer extrema of the 3σ-interval of the underlying implicit multi-variate Gaussian.

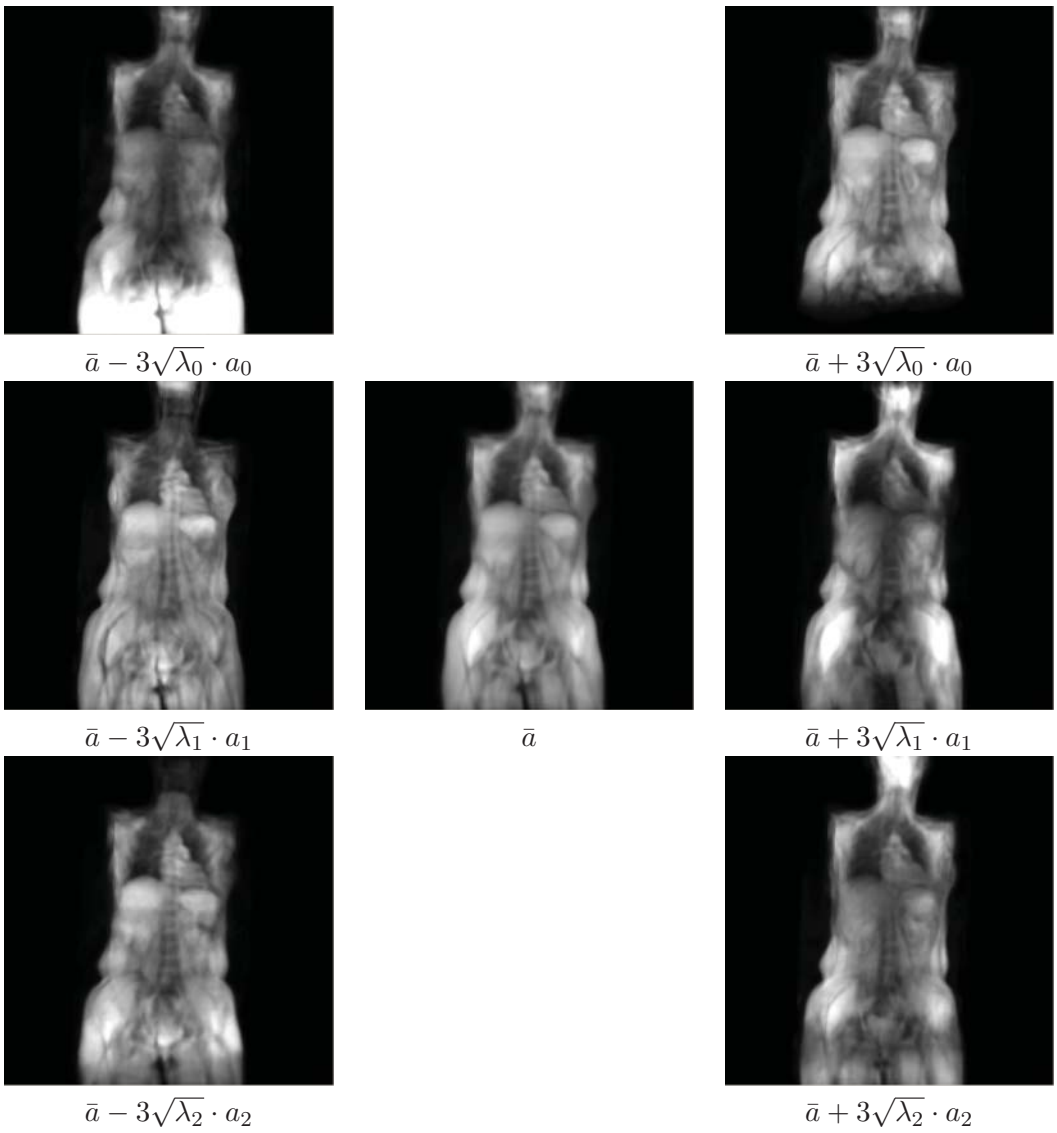


Figure 4.11: This figure presents the statistical model of appearance of the atlas with the average instance of intensity in the center and its major modes of variation. Each row shows a different mode of the appearance, with the left and right column being the outer extrema of the 3σ-interval of the underlying implicit multi-variate Gaussian.

organs. For example in $3T$ systems, B_1-effects can have a strong influence on the overall signal intensity with an overall or local loss of signal [Ros04]. Since many of these effects are obviously independent of the deformation fields, they should be treated differently in order to keep the atlas free from biases.

The quality of the atlas can be visually assessed by inspecting the sharpness of edges and boundaries. A high-quality atlas should not appear more blurred than the input images, although a slight blur results from the trilinear interpolation of the images. Visual inspection of the overall quality of the computed atlas shows that its edges and boundaries compare well to the quality of the registered input images.

Labeling of the anatomical structures in the atlas is done by means of a semi-manual segmentation tool like [FTS07]. For the prototype, 4 organs were selected which are most interesting in terms of fully automatic examination and MR slice positioning purposes: heart, spine, liver and kidneys, which account for a large fraction of diagnostic examinations within the human torso and which can be potentially tedious and tricky for manual scan position planning.

4.4 Registration of FastView Image Data to the Atlas

As described before, atlas based segmentation of an unseen FastView image data set I_{uns} resp. its anatomical labeling is done by first registering the unseen image to the atlas and then propagating the labels to the unseen data set. In other words, one tries to find the instance of the statistical atlas that best explains the unseen image data and then transform the anatomical knowledge from the atlas to the unseen template. This can be considered as an analysis-by-synthesis approach, which makes the method more robust against outliers of poor image quality.

4.4.1 Registration Algorithm

Registering an unseen FastView image I_{uns} to the atlas requires finding a non-rigid transformation, i.e. an affine offset and a deformation field, and the normalized intensity appearance of the atlas which maximize the similarity between the two. Considering the linear models from equations 4.26 and 4.22 which define the atlas, the registration process computes the feature values b_d and b_a plus an affine offset transformation A which was removed from the linear model of deformation fields on purpose. As the deformation fields have been computed by registering all data sets to the reference data set via backwarping, the atlas based registration is also done by warping the unseen data set to the atlas instead of vice versa. This allows for directly using the deformation fields from the atlas avoids the need for computing potentially inconsistent inverse deformation fields. Furthermore, it allows for easier propagation of spatial correspondence among the data sets. Hence, in terms of image registration, I_{uns} is the so-called moving image and the instance of the atlas is the fixed image.

The complete registration process can be divided into a preprocessing step, which includes a histogram normalization as described in 4.3.3, the arm stripping as explained in 4.3.2 and an affine preregistration. The affine preregistration roughly aligns the unseen image I_{uns} to the atlas by computing the bounding boxes of binary thresholded images and an affine transformation that maps the coordinates of the bounding boxes. The affine transformation is calculated

according to equation 4.1 by using the corners of the bounding boxes as landmarks. This transformation is used as the initial affine offset of the registration. Details for computing the affine transform are explained in section A.3.

The non-rigid part of the registration finally computes the precise affine offset and the instance of the deformation field and grey value appearance that best map the warped image I_{uns} to an instance of the atlas. The number of degrees of freedom is reduced to a previously selected number n_d of modes of the statistical model of deformation and a number n_a of modes of the statistical model of grey value appearance. Including the degrees of freedom of the affine offset transformation, the registration is driven by $12 + n_d + n_a$ degrees of freedom.

The actual registration algorithm repeatedly generates an instance of the intensity appearance of the atlas and of the deformation field from the current optimization parameters. The unseen input image is warped according to the affine offset and the deformation field. A similarity score is computed between the instance of the appearance model of the atlas and the warped image. This score is used by an optimizer to steer the registration towards the optimum. Figure 4.12 shows the computational components and the data flow of the registration algorithm.

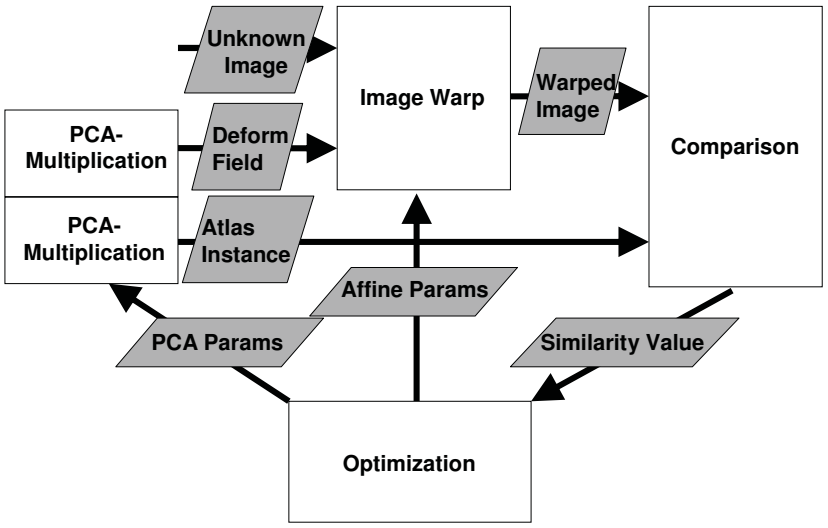


Figure 4.12: Data flow and computational components of the registration algorithm.

4.4.2 Similarity Measure

The registration is steered by a similarity measure which quantifies the amount of similarity between the two images. The similarity between the warped image I_{uns} and the atlas instance image is expressed by a distance function d. d itself is a function of the instance of the atlas I_{atlas} and the warped image I_{uns}. I_{atlas} is a function of the appearance parameters b_a, while I_{uns} is a function of the transformation T_{A,b_d}, which in turn is set up of an affine offset transformation A and the deformation parameters b_d. So, in total, d is a function of A, b_d and b_a, i.e. $d \leftarrow d(b_a, b_d, A)$.

Minimizing $d(b_a, b_d, A)$ is equal to maximizing the similarity and leads to an optimal registration. Two types of distances are implemented and tested. One of them is the pixelwise sum of squared differences between the instance of the atlas and the other one a pixelwise weighted sum of squared differences.

The unweighted distance is computed as:

$$d(b_a, b_d, A) = \frac{\sum_{x_i \in \Omega} \left(I_{atlas}(b_a, x_i) - I_{uns}(T_{A,b_d}(x_i))\right)^2}{|\Omega|} \tag{4.27}$$

where Ω is the overlapping region of the two images. x_i are the world coordinates of pixel i and $T(x_i)$ its transformation according to the concatenation of the global affine and local non-rigid transformation. The final score is divided by the size of the overlap resp. the number of pixels in the overlap Ω.

The weighted distance extends the formula to:

$$d(b_a, b_d, A) = \frac{\sum_{x_i \in \Omega} w_i \left(I_{atlas}(b_a, x_i) - I_{uns}(T_{A,b_d}(x_i))\right)^2}{\sum_{i \in \Omega} w_i} \tag{4.28}$$

where w_i a specific weight encoding the confidence one has in the value at x_i. The final score is divided by the total sum of weights over the image overlap domain.

If w_i is set to

$$w_i \leftarrow \frac{1}{\sigma(x_i)^2} \tag{4.29}$$

the pixelwise variance in the training images, a Gaussian-like distance between the atlas and the image I_{uns} is obtained. Final division by the total sum of weights provides a normalization with respect to the overlap region of I_{atlas} and I_{uns}. Weighting the contributions of different pixels by the pixelwise squared standard deviations $\sigma(x_i)$ allows for considering the degree of uncertainty that a given voxel may have. This is particularly important since certain structures with high variation, like the head, the legs or the stomach should not affect the registration of organs with small uncertainties.

The parameters that minimize the cost function are then the registration parameters of the optimum resp. of the atlas instance with minimum distance or maximum similarity:

$$(A_{opt}, b_{d_{opt}}, b_{a_{opt}}) = \arg \min_{A, b_d, b_a} d(b_a, b_d, A) \tag{4.30}$$

4.4.3 Optimization

Gradient Descent

The simplest way of optimizing the registration cost function is by using a gradient descent algorithm. This can be combined with an adaptive step size as described in section 4.3.4. This

requires a computation of the gradient or the derivative of the cost function at every step. The gradient ∇d is approximated as a first order difference:

$$\nabla d = \frac{d(b_a + \epsilon_a, b_d + \epsilon_d, A + \epsilon_A) - d(b_a - \epsilon_a, b_d - \epsilon_d, A - \epsilon_A)}{2\epsilon} \quad (4.31)$$

Although the gradient descent method with adaptive step size converges slowly and requires many function evaluations, it allows for coupling the step sizes ϵ for the computation of the numerical gradient to the current global step size of the gradient descent algorithm. Hence, there is no need for additionally computing optimal step sizes for the calculation of the gradients.

However, since the computation of the gradient is expensive and requires 2 times the number of degrees of freedom of evaluations of the cost function and since gradient descent algorithms converge slowly, it is desirable to have an alternative optimization algorithm.

Gauss-Newton

The optimum of the registration cost function can be calculated by means of the Gauss-Newton algorithm [Bjö96]. The Gauss-Newton algorithm is suitable for optimizing multi-valued least squares problems with multiple input parameters.

This means that this algorithm finds a solution to the following problem: Given a number of n observations y_i and a n-valued non-linear function $f_i(x)$, find a parameter setting x such that $\sum_i (f_i(x) - y_i)^2$ becomes minimal. The number of observations resp. the number of values has to exceed the number of input parameters.

The upper distance functions 4.27 and 4.28 have multiple input parameters but a scalar output, which means they are 1-valued in the terms of the upper description of the Gauss-Newton algorithm. In order to obtain multiple values, the image has to be partitioned into multiple blocks. See figure 4.13 for an example of the partitioning into blocks. The distance function is evaluated independently in each of these blocks. Each block then provides a small contribution to the overall optimization. Partitioning the images and the distance function in n blocks consequently yields an n-valued distance function.

The basic idea behind the Gauss-Newton approach is a linearization of the problem by making a first-order Taylor expansion. Let p be the union of all relevant parameters $p = \{A, b_d, b_a\}$ and p_0 the starting parameters.

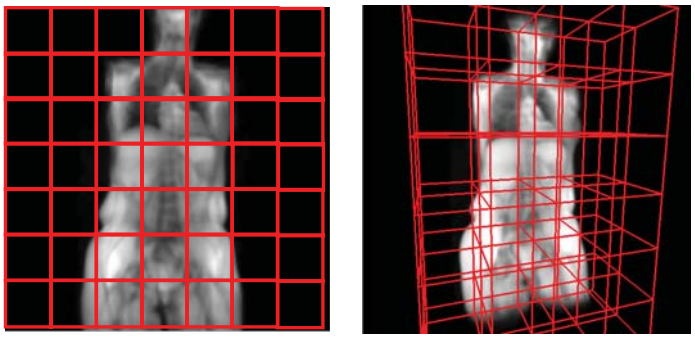
$$d(p) = d(p_0) + J_d \cdot (p - p_0) + O(p)^2 \quad (4.32)$$

where J_d is the Jacobi matrix of the distance function.

From equation 4.32 one can derive that minimizing $d(p)$ means minimizing the zero and first-order terms of the right hand side. The square term can be neglected if p is close enough to the optimum. Consequently,

$$\min ||d(p)||^2 \iff \min ||d(p_0) + J_d \cdot (p - p_0)||^2 = \min ||d(p_0) + J_d \cdot r||^2 \quad (4.33)$$

with $r \leftarrow p - p_0$ being the residual. Expanding the norm yields



a) b)

Figure 4.13: a) shows a 2D block partitioning of a FastView data set, b) shows a 3D rendering of the block partitions.

$$\min(r^T \cdot J_d^T \cdot J_d \cdot r + 2d(p_0)^T \cdot J_d \cdot r^T + d(p_0) \cdot d(p_0)) \tag{4.34}$$

The cost function is minimized by computing the derivative of the scalar function, setting the derivative resp. gradient to zero and solving the resulting equation for the residual r_{min}:

$$2J_d^T \cdot J_d \cdot r_{min} + 2J_d^T \cdot d(p_0) = 0 \tag{4.35}$$

$$J_d^T \cdot J_d \cdot r_{min} = -J_d^T \cdot d(p_0) \tag{4.36}$$

This can be used to update the current parameters p_0:

$$p_0^{t+1} = p_0^t + r_{min}(t) \tag{4.37}$$

Since the matrix product $J_d^T \cdot J_d$ is symmetric and positive definite, equation 4.36 can be solved using a Cholesky decomposition [VF02], which is faster and more robust than a matrix inversion. In some cases $J_d^T \cdot J_d$ may be ill-conditioned and therefore require prior preconditioning before the Cholesky decomposition can be computed [VF02]. A simple linear preconditioning is sufficient [Saa03].

After each step of the optimization, the deformation and appearance parameters are confined to a 3σ interval of the multi-variate Gaussian in order restrict the search to the space of valid instances:

$$p_i \leftarrow \min(3\sigma_i, \max(-3\sigma_i, p_i)) = \min(3\sqrt{\lambda_i}, \max(-3\sqrt{\lambda_i}, p_i)) \tag{4.38}$$

where λ_i is the principal value, that corresponds to the $i-th$ principal component.

The Gauss-Newton optimization algorithm is sensitive to the quality of the Jacobian J_d resp. of the derivatives of the cost function. It is therefore essential to use symmetric differences for the estimation of the derivatives instead of forward differences. The step size for the computation of the symmetric derivatives has an important effect on the performance of the

optimization, too. If chosen too small, the derivative may be not robust, if chosen too large, the derivative may be incorrect. For this reason, the step size for the derivative is determined dynamically and individually for each parameter.

The individual step sizes for the geometric parameters, i.e. for the affine and the deformation parameters, are determined by considering their effects on all voxels in the image domain Ω. Let $T_{\Delta p_i}$ be the transformation T with parameter i changed by the value Δp_i. Then the step size Δs_i for the computation of the symmetric differences of parameter i is set to the value that the maximum coordinate shift of any voxel in Ω is at most the edge length e of one voxel.

$$\Delta s_i = \arg_{\Delta p_i} \sup_{x \in \Omega} ||T_{\Delta p_i}(x) - x|| \leq e \tag{4.39}$$

As the geometric parameters are all linear, the actual values of Δs_i can be determined easily by pure linear algebra.

The step sizes for the parameters of the linear appearance model are fixed at an empirical constant.

Again, the registration of an unseen FastView image to the atlas is computed in a hierarchical top-down multi-scale fashion. The algorithm starts with coarse image and atlas resolutions from Gaussian pyramids, and refines the resolutions after the optimizer converges. The number of degrees of freedom is also gradually increased. In the final stage of the algorithm with the finest image and atlas resolutions, 12 affine, n_d deformable and n_a appearance parameters are used for the optimization. For the ith of n stages, this number is reduced to $\frac{i \cdot n_d}{n}$ deformation and $\frac{i \cdot n_d}{n}$ appearance parameters. The number of affine parameters remains constant, of course. This leads to better performance through faster convergence during the most time-consuming high-resolution stages of the algorithm.

4.4.4 GPU Based Hardware Acceleration

Motivation

A straight implementation of the registration algorithm described is rather slow and requires runtimes of more than one hour, which makes it infeasible for practical usage in clinical routine. A common way of acceleration is to make use of the massive computation power available in modern graphics card hardware. Modern GPUs [1] provide gigantic floating point computation and parallelization capabilities off the shelf at reasonable prices. This development is driven by the huge economic market of computer game and hardware industry which strives for ever higher frame rates, complexity of scenes and more texture memory towards more and more realistic rendering.

Due to this development, GPUs have been outperforming CPUs for some considerable time already. The differences in computation performance are remarkable: While a common CPU like the Pentium 4 3GHz supports up to 6 GFLOPS, the GPU which was used for this thesis, an NVIDIA 8800 GTS, has a potential of 345 GFLOPS in theory [NVI06]. This gives an idea of the possible performance boost, which can be exploited by porting applications to the GPU. This enormous advantage in performance of GPUs in comparison to CPUs is due

[1] graphics processing units

to their SIMD [2] architecture, their inherent parallelism, their specialization on floating point computations and special hardware components for interpolation computations.

The intrinsic parallelism and floating point capabilities make GPUs also interesting for various problems in scientific computing. The activities of using the computation power of GPUs for purposes other than rendering scenes are grouped under the term general purpose graphics processing unit (GPGPU). A good overview of works on GPGPU can be found in [Pha04] or in an introduction by Mark Harris [Har04]. The applications range from image filtering, over flow simulation to fast Fourier transformations [Pha04]. A lot of these developments were leveraged by the development of the programmable rendering pipeline and the development of high level shader languages like GLSL, HLSL or CG. These programming languages greatly reduce the complexity of shader programming. Writing a shader in these high-level languages is almost as easy as writing a common piece of software.

Shader programs can be integrated into the rendering pipeline of the GPU. The rendering pipeline can be roughly divided into a geometric part and a rasterizing part. See [Ros06] for more details. The geometric part includes all vertex operations like transformations, lighting, projections, clipping etc. The rasterization part does all pixel operations like texturing, depth test, etc. Programmable shaders, i.e. programs which are executed on the GPU, can either be plugged into the geometric part of the rendering pipeline, which means that they operate on each vertex or, into the rasterizing part of the pipeline, which means that they are integrated into the rasterization process and executed for each fragment, which is a pixel with additional information of depth. The former shaders are called vertex shaders, while the latter ones, which are more suitable for GPGPU computations are called fragment shaders. A more detailed introduction into programmable graphics hardware can be found in [Ros06].

Despite the simplicity of actually writing and executing shader programs on the GPU, there is one drawback about using fragment and vertex shaders for GPGPU applications. The drawback is the fact, that almost always the algorithms have to be redesigned completely to fit to the standard framework imposed by the rendering pipeline of a GPU. In many cases, this generates huge efforts and sometimes it is not possible at all. Image processing algorithms, however, can often be adapted smoothly to fit to the rendering pipeline framework which makes this type of algorithms, e.g. registration tasks well suited for GPGPU adaptations.

Recent developments like CUDA[3][Sil07] try to overcome the obstacle of the need of algorithm and software redesign for GPGPU applications by providing tools and a special compiler that supports particular language extensions to the C programming language. This allows the user to program his GPGPU application in standard C and specify which parts are to be executed on the GPU and which ones on the CPU. The compiled code is managed by the CUDA runtime, which handles the management of components which are executed on the GPU and components which are executed on the CPU.

For the reason of backward compatibility, however, it was decided to implement the GPGPU based atlas registration described in this thesis using fragment shaders instead of CUDA. This allows for running the algorithm also on non NVIDIA graphic cards or non-recent graphic cards generations, as those which are used in current MR image scanner hosts and reading

[2] single instruction multiple data

[3] Compute Unified Device Architecture

workstations, in order to faciliate practical implementations on current hardware and their usage in prototypes in clinical research or routine.

The main bottleneck in GPGPU applications is the transfer between RAM and GPU memory via the AGP or PCI express bus. Although the bandwidths for the upload, i.e. for the transfer of data from RAM to the GPU, in modern systems are high with rates of about 3.2 GB/s, slow readbacks and latencies may often slow down computation if a lot of data transfer is necessary [Pha04]. For this reason, this GPGPU based atlas registration algorithm is implemented with the goal of bringing as many components of the calculation as possible to the GPU and reducing the number of the upload and download operations. Figure 4.14 shows an overview of the data flow and the computation components of the registration in the GPGPU implementation.

As figure 4.14 shows, the statistical atlas and its components, the statistical model of deformation and the statistical model of appearance reside completely as float textures on the GPU. This reduces the amount of data transfer between RAM and GPU memory to a minimum, i.e. the parameters of the optimization. All computations apart from the optimization itself can be done using fragment shaders, which exploits the parallelity and floating point capabilities of the GPU to a maximum. The following sections describe the fragment shaders resp main computational components which are necessary for the registration in detail.

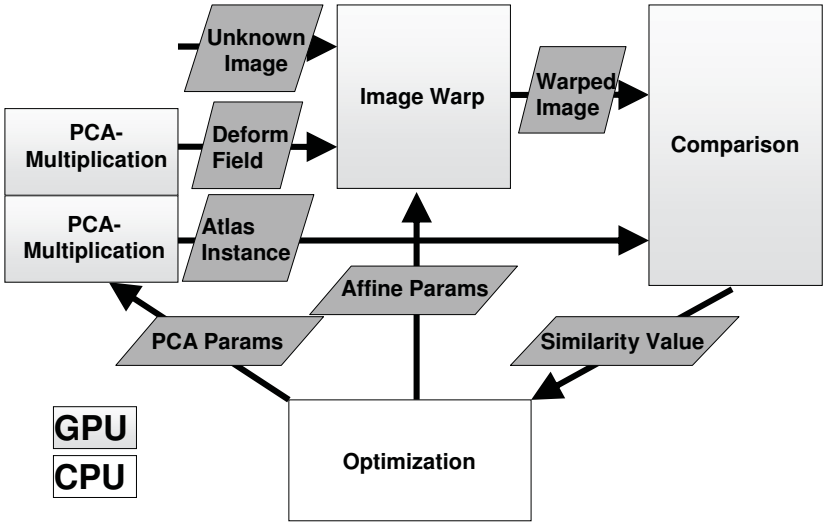


Figure 4.14: Data flow and computational components of the GPGPU based atlas registration algorithm. The white box indicates the computational components which are based on the CPU, while the boxes shaded in grey indicate GPU based computational components.

Generation of an Atlas Instance on GPU

Generating an instance of appearance of the statistical atlas mainly consists of evaluating equation 4.26 and converting the resulting single vector of intensity values to an image of the correct

dimensions. $\bar{a}$ and Φ_a are encoded as float textures which have been uploaded to the texture memory of the GPU. Generating the instance of appearance boils down to a matrix multiplication, a vector addition and writing the result into another float texture. This target texture of the atlas instance can be implemented using a frame buffer object [Gre05]. [Pha04] already describes a GPU based matrix multiplication, but in the case of producing an instance of appearance from the statistical atlas, it is slightly more complicated, as the input data are encoded as 3D float textures and not as 2D texture structures like an actual matrix in linear algebra.

The shader basically has to do a reformatting of the data by transforming the 3D texture indeces to 2D matrix indeces. After reformatting, the shader does the multiplications and additions of the correct values according to matrix algebra and transforms the indeces back to 3D. Special care has to be taken not to violate the hard constraints of maximum 3D texture dimensions. Especially, matrix Φ has to be reformatted such that the hard constraints are respected. See figure 4.15 for a clarification of the transformation of indeces.

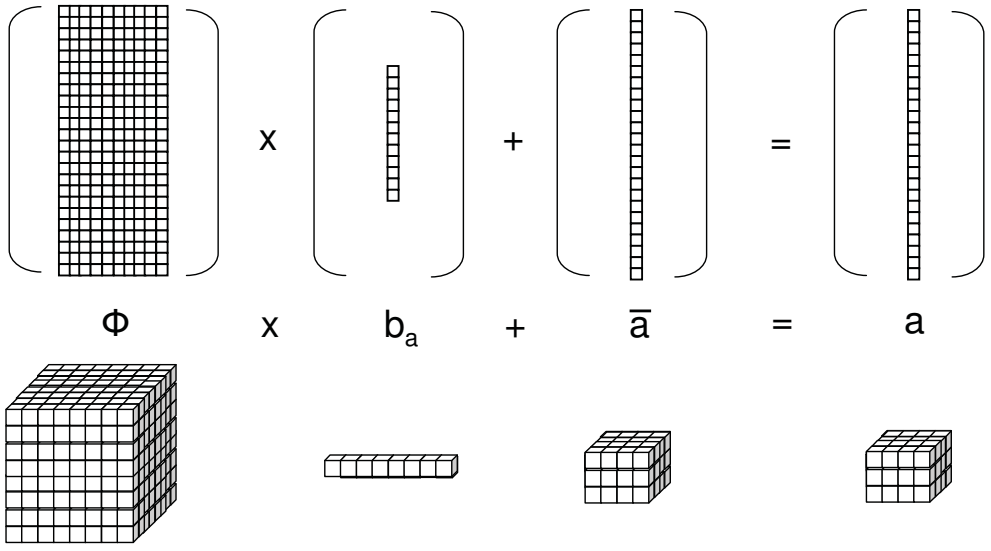


Figure 4.15: This figure shows the reformatting of the textures by converting the indeces from 3D to 2D and back. Linear algebra requires a memory layout as in the upper row. The data structures used are represented as 3D textures as in the lower row. The shader has to convert the coordinates.

Image Warping on GPU

Warping the unseen image I_{uns} on the GPU is done in several steps: First, I_{uns} is loaded offline as a texture into the texture memory of the GPU. This can be done once before the actual registration starts. Next, the deformation field has to be calculated from the deformation parameters b_d. This requires evaluating equation 4.22, which is similar to generating the atlas instance on the GPU: a matrix multiplication of the deformation field modes of variation followed by a vector addition of the average deformation field. The only difference stems from the fact that

the data structures that are associated with the deformation fields are encoded as RGB textures while the data structures related to the scalar atlas image data are pure luminance textures. The resulting deformation field is rendered to an RGB frame buffer object [Gre05].

The warped image is produced by having a warping fragment shader render to a frame buffer object. This shader first transforms each fragment resp. voxel to world coordinates, applies the affine offset transformation and computes the B-spline tensor product of the deformation vector. The B-spline tensor product is obtained by evaluating equation 4.2. The resulting deformation vector is added to the transformed world coordinates. The world coordinates are then transformed back to texture coordinates and can be used for an interpolation lookup at the respective texture coordinate in the original image texture of I_{uns}. Figure 4.16 shows a figure of the GPU based image warping. All linear algebra operations can be efficiently implemented on the GPU. The interpolation operations also greatly benefit from hardware accelerations.

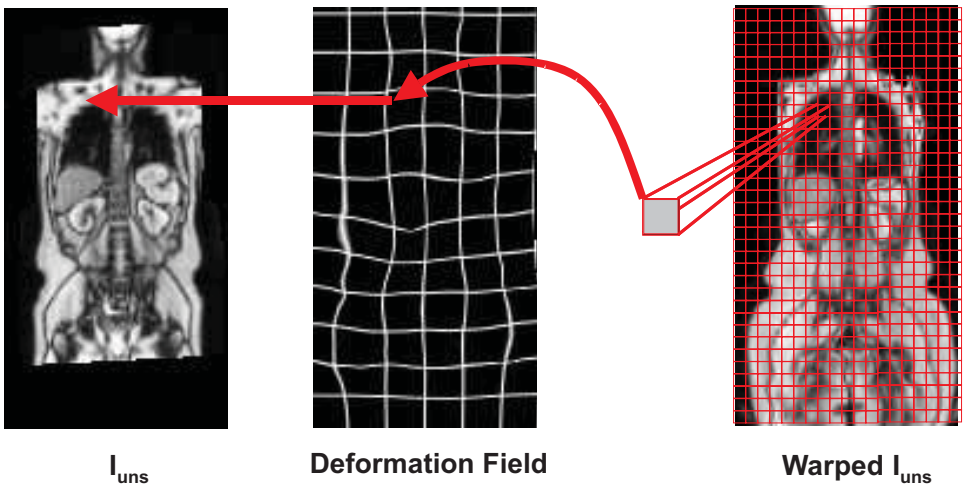


Figure 4.16: This figure shows GPU fragment shader based warping. The warped output image is rendered, by looking up the grey values in the original image after application of the deformation field and the affine offset transformation to the world coordinates of the voxel.

Warping the image on the GPU results in a huge performance boost. By means of the GPU based implementation, an image warp can be done in real time, i.e. in less than 0.02 seconds, which is almost 100 times faster than the native implementation.

Similarity Measure on GPU

Porting the calculation of the similarity measure to the GPU requires a fragment shader which evaluates the equations 4.27 and 4.28 by means of the GPU. This is done in a two step process.

First, the pixelwise contributions to the similarity measure, e.g. squared differences between the atlas instance texture and the warped image texture, are calculated and rendered to an intermediate texture. For the weighted distance, another input texture is used which contains

the weights, i.e. a texture which contains the pixelwise standard deviations of the image. In a second step, all pixels in the intermediate texture are summed up by another fragment shader to produce the final value. The most primitive approach for summing up all values would be to render an output image with the dimensions of a single voxel by means of a fragment shader, which iterates over all input voxels to calculate the final sum. The sum of the weights can be cropped in the same way using the second output channel of the texture. However, this would not be very efficient because huge amounts of the parallelization capabilities of the GPU would be left idle. A better approach is to half the image dimensions iteratively by summing all voxels in a local neighborhood at each step and replacing this neighborhood by a single voxel with the value of the sum of its neighborhood. See figure 4.17 for details. This strategy best exploits the parallelization capabilities of the GPU. The number of necessary summing steps is logarithmic in the dimensions of the image. This value can be read back from the GPU and used to steer the optimization process.

For the Gauss-Newton optimization approach, as described in 4.4.3, not a single similarity value but a value for each block of the image has to be calculated. This can be easily achieved by stopping the process of summing values and halving the dimensions at a level i at which the number of voxels corresponds to the number of blocks.

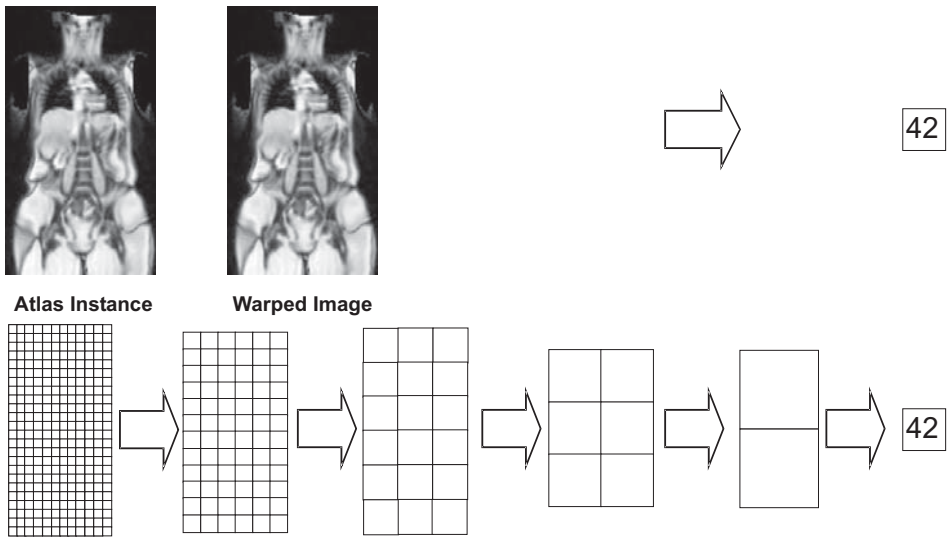


Figure 4.17: This figure shows how the similarity measure is calculated on the GPU. First, the voxelwise distance is computed. Then the dimensions are iteratively halved and neighboring voxels are summed up. This makes use of the parallelization capabilities.

4.5 Validation and Results

Validation of the atlas and the statistical models of deformation and appearance is done by letting a human expert place manual anatomical landmarks at corresponding positions in all

data sets. The corresponding landmarks are selected at anatomical points of interest, which are discernible throughout all data sets. Three landmarks are placed at discernible points of the liver, two on the heart, one at the top and the bottom of left and right kidney and two at different vertebrae of the spine. The manual landmarks are placed using a software tool that uses a combination of volume picking and marking of points in two perpendicular MPR[4] slice views based on the initial picking ray. This ensures consistent placement of landmarks within the 3D volume by ensuring that the landmark point is confirmed in two perpendicular slices.

Reproducibility is studied by repeating the placements multiple times and with different operators. The mean positions of these landmarks are then considered to be the ground truth. The standard deviations were about 10.4 mm. The performance of the atlas based registrations and labelings are assessed by calculating the displacements of the propagated landmark positions to the ground truth. All tests are performed in leave-all-in and leave-one-out scenarios. Leave-all-in means that the atlas is built from all data sets and then compared to the result of the registration of one of those data sets. In a leave-one-out test, which indicates the generalization ability, the atlas is built from all data sets but one and tested against this data set. Table 4.1 gives an overview of the results. The tests are repeated with different numbers of modes of the statistical atlas.

The average calculation times for the Gauss-Newton based optimization were about 10-30 seconds depending on the number of modes being used. For the gradient descent approach, the running times are in the range of a minute. Although the results are a little bit more accurate, this approach is not feasible in practice, so a practical application would be based on the Gauss-Newton algorithm. The given results were published in [FTS08a]

4.6 Discussion and Future Work

This chapter presents a method for fast anatomical labeling in FastView MR localizer images. The results are promising and suggest applicability of the algorithm to fully automatic positioning applications. Evidently, the results improve with an increasing number of modes. Overall, the deviations converge to a σ resp. 2σ interval of the ground truth for the leave-all-in resp. leave-one-out case. A obvious and non-surprising observation is that, using more modes results in more exactness of the output of the algorithm. More modes, however, also lead to an increase in running times, so a tradeoff between exactness and performance has to be made.

It should be noted, however, that the atlas would certainly benefit from additional training data and an increase of generalizability. This can be stated by considering its leave-one-out performance and its dimensionality. The dimensionality of a statistical model of deformation is defined by the number of deformation modes necessary to explain a certain percentile (e.g. 95 percent) of the variation of deformation. Adding image data sets still has an increasing effect on the dimensionality of the atlas. In a comprehensive and complete statistical atlas or model, this incremental increase should be marginal and gradually converge to 0. As a rough estimate from the numbers, another 20-30 data sets should be acquired in order to reach a state of convergence and a asymptotic level of dimensionality. Similar observations hold for the statistical model of grey value appearance. In a complete statistical atlas, there should be

[4] multi-planar reformatting

# Deformation modes	# Appearance modes	Optimizer Type	SM	LM displacements (mm)	
				leave-all-in	leave-one-out
0	5	GN	U	23.7	28.5
10	5	GN	U	17.4	25.4
20	10	GN	U	13.7	24.0
30	15	GN	U	11.7	22.4
0	5	GD	U	30.5	29.1
10	5	GD	U	14.9	25.9
20	10	GD	U	12.8	23.8
30	15	GD	U	10.7	23.0
0	5	GN	W	27.4	27.4
10	5	GN	W	17.3	23.6
20	10	GN	W	14.1	23.0
30	15	GN	W	13.3	23.0
0	5	GD	W	25.7	27.6
10	5	GD	W	22.6	25.6
20	10	GD	W	21.7	25.5
30	15	GD	W	20.6	24.7

Table 4.1: The experimental results of the displacements between the manual landmarks and the landmarks determined by the atlas based registration. The four first columns indicate the parameter settings of the test case, the number of deformation modes, the number of appearance modes, the optimizer type, which is either Gauss-Newton (GN) or adaptive step gradient descent (GD). The similarity measure (SM) is either the weighted (W) or the unweighted cost function (U).

no significant difference between leave-all-in and leave-one-out test cases. However, specifity may decrease in turn, if additional data are added which would slightly reduce accuracy in the leave-all-in tests. Building a separate atlas for male and female, children and adults and each specific scanner system could also be evaluated to increase the specifity of the atlas.

As for further applications, more exactness may be needed, future work will focus on extending and improving the data basis of the atlas and further tuning the performance by switching from numerical to analytical derivatives. Using analytical derivatives would significantly reduce the running times of the algorithm but require implementing analytical derivatives of the equations 4.28 and 4.27.

As mentioned, the search space of the registration is constrained to the space of valid deformation instances of the atlas. In future implementations, this could be extended to free-form non-rigid registrations guided by deformation priors imposed by the statistical model of deformation according to a Bayesian model. The registration process would then seek to optimize the a-posteriori probability of the deformation field, given by the product of the likelihood of the current image registration and the prior probability from the multi-variate Gaussian of the statistical model of deformation similar to the works by Chen [CKPS99]. This would allow the algorithm to leave the search space of the linear deformation model with its hard constraints of

a 3σ-interval. The deformation prior, however, helps to avoid deformation fields that strongly differ from the training data.

Future work will also evaluate if exactness can be improved by additional advanced registration techniques like the one by Park *et al.*[PBM03] who mask out certain pixels that do not belong to the organs of interest. This forces the registration process to more exactness in the regions of interest, while effects in other regions are neglected. This could help to prevent the registrations of small organs from becoming dominated by registrations of large organs and thus enhance the exactness of details of the segmentations. Exactness of the kidney registrations and segmentations could benefit from this adaptation.

Another improvement could be a final multi-class fine-tuning segmentation of the initial labelings provided by the atlas based segmentation. Fine-tuning would be done by using region based active contours [CKS95], which consider local histogram based measures and shape priors. Shape priors would be calculated from active shape models as described in chapter 3. Histogram based measures could either be global models of appearance or local grey value profiles as in the case of active shape models.

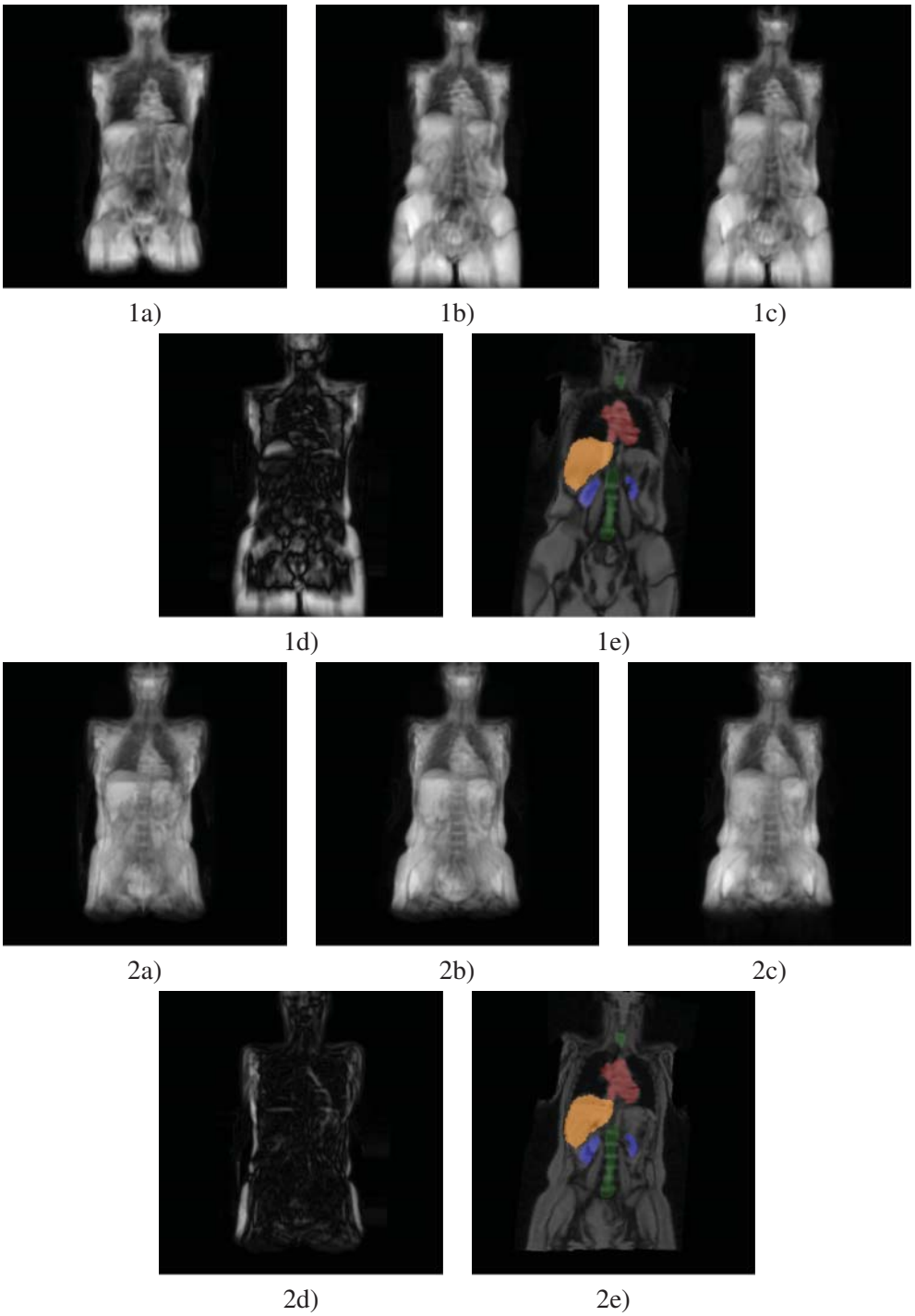


Figure 4.18: This figure shows two results of the atlas based registrations: the original images (a), the warped image (b), the best fitting instances of the atlas (c), difference images (d) and slice images of propagated labels for heart, kidneys, spine and liver (e).

5 Conclusion

5.1 Summary and Discussion

This thesis presented two algorithms for segmentation of anatomical structures from MR localizer images: A special technique for liver segmentation from a sparsely covering stack of 2D localizer images and a more general technique for multi-organ segmentation from a FastView 3D localizer. Both methods are based on statistical models which are generated from training data images. The main advantage of using statistical models is that the algorithms incorporate prior knowledge of the structures which are to be expected by capturing their average shape and their natural variability. This reduces the search space of possible solutions and is highly beneficial if image quality is poor, which often is the case if localizer images are analyzed. The segmentation process is steered by the statistics inferred from the training data towards the most probable solution in the given input data. As shown by the results of the methods, this approach is powerful enough for segmentations from low resolution, low quality image data.

However, it should be mentioned that the creation of the statistical model requires some effort including image data collection, selecting a representative group of representatives that covers the natural variability, preprocessing of the data and generation of a statistical model. All the steps described are necessary for the creation of a compact model of high quality. Although the model may be extended quite easily by adding additional data sets, substantial changes in the image acquisition protocols or sequences may require collecting new data sets and creating a new model. This is particularly important when the statistical models of appearance or the statistical atlases are taken into account for finding the best solution, as those are directly based on statistics of the grey value distributions of the images. This problem may be partly solved by using image normalization strategies and collecting training data that cover possibly large portions of the natural variability and the variation related to the MR image acquisition itself. Normalization strategies can help to separate undesired variance or biases from statistical variance and improve the generalization abilities of a statistical model. Still significant changes in the image contrast may require regenerating a statistical model. In practice, however, this limitation is of minor importance as the localizer protocol can be adapted and consolidated, within some basic constraints to optimally satisfy the needs of a given automatic localization and alignment algorithm, because its images are not used for later diagnostic purposes. The FastView protocol, for example, is a specialized protocol for obtaining a fast 3D overview, with limited diagnostic value.

5.2 Outlook

The described algorithms leverage a lot of possible applications, which can be created based on these methods. Scan automation like slice positioning, sequence adaptations, placement of saturation bands, SAR value estimations etc. can be done using anatomical information inferred from localizer image data. For this reason, robust segmentation algorithms which extract relevant anatomical information from localizer images are highly appreciated and beneficial for more automatic, or ideally, completely automatic MR applications.

The method for the liver segmentation from stacked 2D slice images could also be extended to other organs, like the kidneys or the heart. Since cardiac MR examinations require complex positioning tasks, a more detailed model and segmentation method are needed. For this purpose high quality localizer images with higher resolutions but a smaller fields of view (FoV) should be used, which requires a prior rough estimate of the position and orientation of the heart for the acquisition of high-resolution localizers with a smaller field of view.

A fascinating idea would be to combine the atlas based initial anatomical labeling of FastView images with subsequent precise organ segmentations from high-resolution localizers and detailed active shape models. This means setting up a hierarchical scanning and segmentation workflow, which could be used for fully automatic scan planning: First, when a new patient undergoes an MR scan of certain organs or regions of interest, a FastView protocol is scanned which captures a 3D localizer image of the patient. The FastView image is submitted to the automatic organ labeling algorithm. Estimations for the examination of the desired organs are obtained from the anatomical labeling, as described. These rough positionings could be used to set up high-resolution and high-quality localizers which are subsequently scanned in the regions of interest, that have been determined. The latter scans can then be used for a precise segmentation, e.g. based on a active shape model, like the one described in this thesis for the liver.

The detailed localizer scans could also be 3D localizer acquisitions, e.g. from adequate fast 3D protocols in the regions given by the initial segmentations and ROIs from the atlas based automatic labeling of the FastView images. 3D localizer images would allow for more high-performance and precise active shape model based segmentations although the acquisition time in general would be longer. For this reason, a tradeoff between scan time and precision of the segmentations and the results has to be made.

In the described scenario, the diagnostic protocols could then be scanned based on the detailed segmentations in the regions of interest and , ideally, with optimized parameter settings that could be derived from these segmentations. This hierarchical procedure would allow for detailed, robust and reproducible positioning of slices with respect to the individual anatomy. All these steps could be done completely automatic.

As an example and proof of concept, this thesis presented two segmentation algorithms from localizer images. Possible extensions and applications and their value for the MR exam workflow were shown. A combination of these two methods, as explained, along with extensions to one or two additional organs would account for a large portion of scans in clinical practice, which could be run fully automatic. The individual chapters already discussed a number of possible improvements and extensions of the methods themselves, which could be implemented in order to increase precision an robustness. In fact, the prototypes in this thesis

were developed by virtually taking practical applicability into account. Concepts and tests of integrations of these methods into actual MR product software and hardware for basic automatic positioning applications have been worked out and studied [FTS06]. Future work would further intensify this aspect of the implementations. The thesis showed a proof of concept of using statistical models like active shape models or statistical atlases for challenging segmentation topics like organ segmentation from fast localizer image and the applicability in practice. The value of these methods in terms of increasing quality and reproducibility and reducing time and costs of an exam would have to be assessed by clinical studies in the future.

A Mathematical Details

A.1 Fast Computation of the Principal Components

Usually, principal component analysis (PCA) is computed on the covariance matrix of the input data. The principal components and their corresponding principal values can be found for example by a singular value decomposition (SVD) [VF02]. A SVD is a decomposition of a matrix A into three matrices U, D and V, with

$$A = UDV^T \tag{A.1}$$

It is commonly known that D^2 is a diagonal matrix, that holds the eigenvalues of AA^T and U its eigenvectors. The computational complexity of a SVD on a $k \times l$ matrix is $O(\min(k^2 l, kl^2))$. Let n now be the number of items in each data vector and m the number of data vectors, on which the PCA is computed. For large data vectors, i.e. $n \gg m$, it is inefficient to compute the principal components from the covariance matrix S with

$$S = \frac{1}{m-1} \sum_{i=1}^{m} (x - \bar{x})(x - \bar{x})^T \tag{A.2}$$

. As S is a $n \times n$ matrix, the cost of this SVD would be $O(n^3)$.

Let now S' be defined as follows:

$$S' = \frac{1}{\sqrt{m-1}} (X - \bar{X}) \tag{A.3}$$

X is a $n \times m$ matrix containing the vector data as columns and $\bar{X}$ a $n \times m$ matrix that contains m copies of the average data vector as columns.

If the SVD is computed on S' instead, the eigenvalues of S are directly obtained as principal components along with eigenvectors of S' [HWWM05]. The computational complexity is thus reduced to $O(n)$, if $n > m$, which in general is a huge advantage since in most cases $n \gg m$. This method also benefits from better accuracy.

A.2 Similarity Transformation Estimation

This section is based on the works by Zinsser *et al.* in [ZSN05]. Given two point sets $\{a_i\}$ and $\{b_i\}$ with n points and fixed correspondences given by the indices, one would like to find

a transformation given by a translation, rotation and scaling, which best maps the $\{a_i\}$ to the $\{b_i\}$.

So the task is to find an orthonormal rotation matrix M, a translation vector v and a scale factor s, so that that the following score is minimized:

$$\sum_i \|s \cdot M \cdot a_i + v - b_i\|^2 \tag{A.4}$$

First, the centers of mass are calculated $\bar{a} = \sum_i a_i/n$ and $\bar{b} = \sum_i b_i/n$ Then the rotation matrix is found by calculating the singular value decomposition of K:

$$K = \sum_i (b_i - \bar{b})(a_i - \bar{a})^T = UDV^T \tag{A.5}$$

Then $M = UV^T$, if UV^T is a pure rotation, which is the case if $det(UV^T) \geq 0.0$. In the other case, multiplying the third column of U by -1 then yields a rotation. The scaling part s is then computed by minimizing the following energy function:

$$s = \arg\min_s \sum_i \left\|(b_i - \bar{b}) - s \cdot M(a_i - \bar{a})\right\| \tag{A.6}$$

$$s = \frac{\sum_i (b_i - \bar{b})^T M(a_i - \bar{a})}{\sum_i (a_i - \bar{a})^T \cdot M^T \cdot M(a_i - \bar{a})} \tag{A.7}$$

The translational part v is then:

$$t = \bar{b} - s \cdot M\bar{a} \tag{A.8}$$

A.3 Affine Transformation Estimation

Given two point sets $\{a_i\}$ and $\{b_i\}$ with n points and fixed correspondences given by the indices, one would like to find a transformation given by a matrix M and a translation t, which best maps $\{a_i\}$ to the $\{b_i\}$. Let without loss of generaliy $\{a_i\}$ and $\{b_i\}$ be centered at the origin. Otherwise, the point sets are translated to the origin and the resulting translations are later integrated into t.

More specifically, the following sum is to be minimized:

$$E = \sum_i \|M \cdot a_i + t - b_i\|^2 \tag{A.9}$$

with the affine matrix

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \tag{A.10}$$

The following abbreviations are used:

$$S_{xx} = \sum_i a_{xi} a_{xi} \tag{A.11}$$

$$S_{xy} = \sum_i a_{xi} a_{yi} \tag{A.12}$$

$$S_{xz} = \sum_i a_{xi} a_{zi} \tag{A.13}$$

$$S_{yy} = \sum_i a_{yi} a_{yi} \tag{A.14}$$

$$S_{yz} = \sum_i a_{yi} a_{zi} \tag{A.15}$$

$$S_{zz} = \sum_i a_{zi} a_{zi} \tag{A.16}$$

$$S_{xx'} = \sum_i a_{xi} b_{xi} \tag{A.17}$$

$$S_{yy'} = \sum_i a_{yi} b_{yi} \tag{A.18}$$

$$S_{zz'} = \sum_i a_{zi} b_{zi} \tag{A.19}$$

$$S_{x'} = \sum_i b_{xi} \tag{A.20}$$

$$S_{y'} = \sum_i b_{yi} \tag{A.21}$$

$$S_{z'} = \sum_i b_{zi} \tag{A.22}$$

(A.23)

Deriving with respect to the single parameters m_{ij} and t_i and setting the derivatives to 0 then yields the following equations:

$$t_x = S_{x'} \tag{A.24}$$

$$t_y = S_{y'} \tag{A.25}$$

$$t_z = S_{z'} \tag{A.26}$$

and

$$M = S^{-1} \cdot S' \tag{A.27}$$

with

$$S = \begin{pmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{pmatrix} \tag{A.28}$$

and

$$S' = \begin{pmatrix} S_{xx'} & S_{xy'} & S_{xz'} \\ S_{yx'} & S_{yy'} & S_{yz'} \\ S_{zx'} & S_{zy'} & S_{zz'} \end{pmatrix} \tag{A.29}$$

Bibliography

[BHP46] F. Bloch, W. W. Hansen, and Martin Packard. Nuclear induction. *Phys. Rev.*, 69(3-4):127, Feb 1946.

[BJ00] Yuri Boykov and Marie-Pierre Jolly. Interactive organ segmentation using graph cuts. In *Proceedings Medical Image Computing and Computer-Assisted Intervention*, pages 276–286, 2000.

[Bjö96] Ake Björk. *Numerical Methods for Least Squares Problems*. SIAM, 1996.

[Bla06] Mathias Blasche. syngo TimCT - continuous table move, powered by Tim. *Magnetom Flash, Special Issue RSNA Edition*, (3):1–22, 2006.

[BM97] W. Barrett and E. Mortensen. Interactive live-wire boundary extraction. In *Proceedings Conference on Computer Visision and Pattern Recognition*, pages 6–7, 1997.

[Boo89] F.L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[BQL+01] J. Barkhausen, H.H. Quick, T. Lauenstein, M. Goyen, S.G. Ruehm, G. Laub, J.F. Debatin, and M.E. Ladd. Whole-body MR imaging in 30 seconds with real-time TRUE-FISP and a continuously rolling table platform: feasibility study. *Radiology*, 220:252–256, 2001.

[CBET99] T. F. Cootes, C. Beeston, G. J. Edwards, and C. J. Taylor. A unified framework for atlas matching using active appearance models. *Lecture Notes in Computer Science*, 1613:322–330, 1999.

[CET98] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. In *Proceedings 5th European Conference on Computer Vision*, volume 2, pages 484–498, 1998.

[CET99] T. Cootes, G. Edwards, and C. Taylor. Comparing active shape models with active appearance models, 1999.

[CKPS99] Mei Chen, Takeo Kanade, Dean Pomerleau, and Jeff G. Schneider. 3-D deformable registration of medical images using a statistical atlas. In *MICCAI '99: Proceedings of the Second International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 621–630, London, UK, 1999. Springer-Verlag.

[CKS95] Vincent Caselles, Ron Kimmel, and Guillermo Sapiro. Geodesic active contours. In *ICCV*, pages 694–699, 1995.

[CT92] T. Cootes and C. Taylor. Active shape models - smart snakes. In *Proceedings of British Machine Vision Conference*, pages 266–275. Springer Verlag, 1992.

[CT99] T. Cootes and C. Taylor. Statistical models of appearance for computer vision. Technical report, University of Manchester, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, September 1999.

[Dam71] Raymond Damadian. Tumor Detection by Nuclear Magnetic Resonance. *Science*, 171(3976):1151–1153, 1971.

[DHL+89] M.D. Deck, C. Henschke, B.C. Lee, R.D. Zimmerman, R.A. Hyman, J. Edwards, L.A. Saint Louis, P.T. Cahill, H. Stein, and J.P. Whalen. Computed tomography versus magnetic resonance imaging of the brain. a collaborative interinstitutional study. *Clinical Imaging*, 1(13):2–15, March 1989.

[DS98] Nicolae Duta and Milan Sonka. Segmentation and interpretation of MR brain images: An improved active shape model. *IEEE Trans. Med. Imaging*, 17(6):1049–1062, 1998.

[DTC+02] Rhodri H. Davies, Carole J. Twining, Timothy F. Cootes, John C. Waterton, and Christopher J. Taylor. 3D statistical shape models using direct optimisation of description length. In *ECCV*, pages 3–20, 2002.

[DWRB02] D. Diehl, U. Weinert, R. Röckelein, and G. Brinkner. SAR-calculations of a realistic, scalable human model in an MRI-resonator. In *10th International IGTE Symposium on Numerical Field Calculation in Electrical Engineering*, 2002.

[Ess00] D. Essex. The many layers of workflow automation. *Healthcare Informatics*, pages 128–135, June 2000.

[FPFLW07] Charles Florin, Nikos Paragios, Gareth Funka-Lea, and James Williams. Liver segmentation using sparse 3D prior models with optimal data support. In Nico Karssemeijer and Boudewijn P. F. Lelieveldt, editors, *IPMI*, volume 4584 of *Lecture Notes in Computer Science*, pages 38–49. Springer, 2007.

[FRDS06] D. Fritz, D. Rinck, R. Dillmann, and M. Scheuering. Segmentation of the left and right cardiac ventricle using a combined bi-temporal statistical model. In *Proceedings of the SPIE Medical Imaging*, volume 6141, 2006.

[FTS06] Matthias Fenchel, Stefan Thesen, and Andreas Schilling. Fully automatic liver scan planning - slice and navigator positioning from stacked 2D localizer scans. In *Proceedings of the 14th Annual Meeting of ISMRM 2006*, page 467, 2006.

[FTS07] Matthias Fenchel, Stefan Thesen, and Andreas Schilling. An interactive tool for efficient segmentation of volumetric MR images - combining livewire and graph cut paradigms. In *Proceedings Computer Assisted Radiology and Surgery*, 2007.

[FTS08a] Matthias Fenchel, Stefan Thesen, and Andreas Schilling. Automatic labeling of anatomical structures in MR FastView images using a statistical atlas. In *MICCAI 2008: Proceedings of Medical Image Computing and Computer Aided Interventions*, pages 576–584, 2008.

[FTS08b] Matthias Fenchel, Stefan Thesen, and Andreas Schilling. Reconstructing liver shape and position from MR image slices using an active shape model. In *SPIE Medical Imaging*, pages 4N-1–4N-11, 2008.

[GPAK08] Lixin Gong, Sayan Pathak, Adam Alessio, and Paul Kinahan. Automatic arm removal in PET and CT images for deformable registration. *Computerized Medical Imaging and Graphics*, 30, 2008.

[GRB93] J. Gee, M. Reivich, and R. Bajcsy. Elastically deforming a three-dimensional atlas to match anatomical brain images. *Journal of Computer Assisted Tomography*, 17:225–236, 1993.

[Gre05] Simon Green. The OpenGL framebuffer object extension, 2005.

[GW02] Rafael Gonzales and Richard Woods. *Digital Image Processing*. Prentice Hall International, 2002.

[GWC+03] X. Gu, Y. Wang, T. Chan, P. Thompson, and S. Yau. Genus zero surface conformal mapping and its application to brain surface mapping. In *Proceedings of Information Processing in Medical Imaging*, pages 172–184, 2003.

[Har04] Mark Harris. GPGPU: General purpose computation on GPUs. In *Eurographics*, 2004.

[HBTV99] E. Mark Haacke, Robert W. Brown, Michael R. Thompson, and Ramesh Venkatesan. *Magnetic Resonance Imaging: physical principles and sequence design*. Wiley, New York, 1999.

[HFM+86] A. Haase, J. Frahm, D. Matthei, W. Hänicke, and K.-D. Merboldt. FLASH imaging: Rapid NMR imaging using low flip angle pulses. *Journal of Magnetic Resonance*, 67(2):258–266, April 1986.

[HHH01] Joseph Hajnal, Derek Hill, and David Hawkes. *Medical Image Registration*. CRC Press Inc, 2001.

[HvGS07] T. Heimann, B. van Ginneken, and M. Styner. Workshop on 3d segmentation in the clinic: - a grand challenge. http://mbi.dkfz-heidelberg.de/grand-challenge2007, October 2007.

[HWM06a] Tobias Heimann, Ivo Wolf, and Hans-Peter Meinzer. Active shape models for a fully automated 3d segmentation of the liver - an evaluation on clinical data. In Rasmus Larsen, Mads Nielsen, and Jon Sporring, editors, *MICCAI (2)*, volume 4191 of *Lecture Notes in Computer Science*, pages 41–48. Springer, 2006.

[HWM06b] Tobias Heimann, Ivo Wolf, and Hans-Peter Meinzer. Optimal landmark distributions for statistical shape model construction. In *SPIE Medical Imaging*, pages 518–528, 2006.

[HWWM05] Tobias Heimann, Ivo Wolf, Tomos G. Williams, and Hans-Peter Meinzer. 3D active shape models using gradient descent optimization of description length. In Gary E. Christensen and Milan Sonka, editors, *IPMI*, volume 3565 of *Lecture Notes in Computer Science*, pages 566–577. Springer, 2005.

[JDBH07] Sam Joffe, Donna Drew, Manju Bansal, and Michael Hase. Radiology operations: What you dont know could be costing you millions. *Radiology Management*, July 2007.

[KCR+06] Timo Kohlberger, Daniel Cremers, Mikaël Rousson, Ramamani Ramaraj, and Gareth Funka-Lea. 4D shape priors for a level set segmentation of the left myocardium in spect sequences. In Rasmus Larsen, Mads Nielsen, and Jon Sporring, editors, *MICCAI (1)*, volume 4190 of *Lecture Notes in Computer Science*, pages 92–100. Springer, 2006.

[KLL07] D. Kainmüller, T. Lange, and H. Lamecker. Shape constrained automatic segmentation of the liver based on a heuristic intensity model. In *Proceedings MICCAI, Workshop on 3D Segmentation in the Clinic*, pages 109–116, 2007.

[Kob00] Leif Kobbelt. $\sqrt{3}$ subdivision. In *SIGGRAPH*, pages 103–112. ACM, 2000.

[KRGR02] David G. Kruger, Stephen J. Riederer, Roger C. Grimm, and Phillip J. Rossman. Continuously moving table data acquisition method for long FOV contrast-enhanced MRA and whole-body MRI. *Magnetic Resonance in Medicine*, 47(2):224–231, 2002.

[KWT88] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

[KZZ+06] S. Kinner, A. Zielonka, M. O. Zenge, S.C. Ladd, and M.E. Ladd. Whole-body MR imaging with continuously moving table and multiplanar reformations: toward parameter optimization for SSFP imaging in patient examinations. In *Proceedings of the 14th Annual Meeting of ISMRM 2006*, page 2421, 2006.

[Lau73] P.C. Lauterbur. Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature*, 242:190–191, 3 1973.

[LC87] W.E. Lorensen and H.E. Cline. Marching cube: A high resolution 3D surface reconstruction algorithm. In *Computer Graphics*, volume 21, pages 163–169, 1987.

[LDL99] S. Langlois, M. Desvignes, and Serge Langlois. MRI geometric distortion: a simple approach to correcting the effects of non-linear gradient fields. *Magnetic Resonance Imaging*, 9:821–831, 1999.

[LKS+03] Jyrki Lötjönen, Juha Koikkalainen, Daniel Smutek, Sari Kivistö, and Kirsi Lauerma. Four-chamber 3-D statistical shape model from cardiac short-axis

and long-axis MR images. In Randy E. Ellis and Terry M. Peters, editors, *MICCAI (1)*, volume 2878 of *Lecture Notes in Computer Science*, pages 459–466. Springer, 2003.

[LLS02] Hans Lamecker, Thomas Lange, and Martin Seebass. A statistical shape model for the liver. In Takeyoshi Dohi and Ron Kikinis, editors, *MICCAI (2)*, volume 2489 of *Lecture Notes in Computer Science*, pages 421–427. Springer, 2002.

[LLS04] Hans Lamecker, Thomas Lange, and Martin Seebaß. Segmentation of the liver using a 3D statistical shape model. ZIB Preprint 04-09, 2004.

[Mah36] P.C. Mahalanobis. On the generalised distance in statistics. In *Proceedings of the National Institute of Science of India*, volume 12, pages 49–55, 1936.

[Mic27] A.A. Michelson. Studies in optics. *University of Chicago Press*, 1927.

[MK73] P. Mansfield and Grannell P. K. NMR 'diffraction' in solids? *Journal of Physics C: Solid State Physics*, 6(22):L422–L426, 1973.

[Mod04] Jan Modersitzki. *Numerical Methods for Image Registration*. Oxford University Press, 2004.

[NU99] Laszlo G. Nyul and Jayaram K. Udupa. On standardizing the MR image intensity scale. *Magnetic Resonance in Medicine*, 42(6):1072–1081, 1999.

[NVI06] NVIDIA. *Technical Brief - NVIDIA GeForce 8800 GPU Architecture Overview*, November 2006.

[OOL+07] Sebastian Ordas, Estanislao Oubel, Ruben Leta, Francesc Carreras, and Alejandro F. Frangi. A statistical shape model of the heart and its application to model-based segmentation. In *Proceedings of the SPIE Medical Imaging*, volume 6511, 2007.

[Par62] E. Parzen. On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 33:1065–1076, 1962.

[PBM03] Hyunjin Park, Peyton H. Bland, and Charles R. Meyer. Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Trans. Med. Imaging*, 22(4):483–492, 2003.

[PESH04] S. Peschl, T. Ernst, O. Speck, and J. Hennig. Auto alignment of intervertebral disk. In *Proceedings of the 12th Annual Meeting of the ISMRM*, number 2216, 5 2004.

[Pha04] Matt Pharr. *GPUGems 2: Programming techniques for High-Performance Graphics and General-Purpose Computation*. Addison Wesley, 2004.

[PMV03] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual-information-based registration of medical images: a survey. *Medical Imaging, IEEE Transactions on*, 22(8):986–1004, 2003.

[PS95] V. V. Prasolov and A. Sossinsky. *Intuitive Topology*. American Mathematical Society, 1995.

[PTP46] E. M. Purcell, H. C. Torrey, and R. V. Pound. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.*, 69(1-2):37–38, Jan 1946.

[PWSB99] Klaas P. Pruessmann, Markus Weiger, Markus B. Scheidegger, and Peter Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.

[RBJM01] Torsten Rohlfing, Robert Brandt, Calvin R. Maurer Jr, and Randof Menzel. Bee brains, B-splines and computational democracy: Generating an average shape atlas. In *MMBIA '01: Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA'01)*, pages 187–194, Washington, DC, USA, 2001. IEEE Computer Society.

[RBM+05] Torsten Rohlfing, Robert Brandt, Randolf Menzel, Daniel B. Russakoff, and Calvin R. Maurer, Jr. Quo vadis, atlas-based segmentation? In Jasjit Suri, David L. Wilson, and Swamy Laxminarayan, editors, *The Handbook of Medical Image Analysis – Volume III: Registration Models*, chapter 11, pages 435–486. Kluwer Academic / Plenum Publishers, New York, NY, August 2005.

[RE89] J.P. Felmlee R.L. Ehman. Adaptive technique for high-definition imaging of moving structures. *Radiology*, 1(173):255–263, Oct 1989.

[RFS03] Daniel Rueckert, Alejandro F. Frangi, and Julia A. Schnabel. Automatic construction of 3D statistical deformation models of the brain using non-rigid registration. *IEEE Trans. Med. Imaging*, 22(8):1014–1025, 2003.

[RLL+99] N.M. Rofsky, V.S. Lee, G. Laub, M.A. Pollack, G.A. Krinsky, D. Thomasson, M.M. Ambrosino, and J.C. Weinreb. Abdominal MR imaging with a volumetric interpolated breath-hold examination. *Radiology*, 3(212):876–884, Sep 1999.

[Ros04] Jeffrey S Ross. The high-field-strength curmudgeon. *American journal of neuroradiology*, pages 168–177, 2004.

[Ros06] Randi J. Rost. *OpenGL Shading Language*. Addison-Wesley, 2 edition, 2006.

[RSC02] B. Reiner, E. Siegel, and J.A. Carrino. Workflow optimization: current trends and future directions. *Digital Imaging*, 15(3):141–152, September 2002.

[Saa03] Yousef Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, 2nd edition, 2003.

[SD92] L.H. Staib and J.S. Duncan. Boundary finding with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11):1061–1075, November 1992.

[SG02] M. B. Stegmann and D. D. Gomez. A brief introduction to statistical shape analysis, March 2002.

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell system technical journal*, 27, 1948.

[Sil07] Mark Silberstein. High performance computing on GPUs using NVIDIA CUDA. In *GPGPU Tutorial at Siggraph*, 2007.

[SM97] Daniel K. Sodickson and Warren J. Manning. Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays. *Magnetic Resonance in Medicine*, 38(4):591–603, 1997.

[SPR06] Alla Sheffer, Emil Praun, and Kenneth Rose. Mesh parameterization methods and their applications. *Foundations and Trends in Computer Graphics and Vision*, 2(2):105–171, 2006.

[STHW96] Stuart Solloway, Christopher J. Taylor, Charles E. Hutchinson, and John C. Waterton. Quantification of articular cartilage from MR images using active shape models. In *ECCV (2)*, pages 400–411, 1996.

[Stu03] Colin Studholme. Simultaneous population based image alignment for template free spatial normalisation of brain anatomy. In James C. Gee, J. B. Antoine Maintz, and Michael W. Vannier, editors, *WBIR*, volume 2717 of *Lecture Notes in Computer Science*, pages 81–90. Springer, 2003.

[Tan58] T. T. Tanimoto. An elementary mathematical theory of classification. Technical report, I.B.M. Research, 1958.

[Tho18] William Thorburn. The myth of occam's razor. *Mind 27*, pages 345–353, 1918.

[Tho03] Hans Henrik Thodberg. Minimum description length shape and appearance models. In Christopher J. Taylor and J. Alison Noble, editors, *IPMI*, volume 2732 of *Lecture Notes in Computer Science*, pages 51–62. Springer, 2003.

[TT88] J. Talairach and P. Tournoux. *Co-planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System - an Approach to Cerebral Imaging*. Thieme Medical Publishers, New York, 1988.

[TU00] P. Thevenaz and M. Unser. Optimization of mutual information for multiresolution image registration. *IEEE Transactions on Image Processing*, 9(12):2083–2099, 2000.

[vdKBF+05] A. van der Kouwe, T. Benner, B. Fischl, F. Schmitt, D. Salat, M. Harder, A. Sorensen, and A. Dale. On-line automatic slice positioning for brain MR imaging. *NeuroImage*, 27(1):222–230, 2005.

[VF02] William T. Vetterling and Brian P. Flannery. *Numerical Recipes in C++: The Art of Scientific Computing*. Cambridge University Press, February 2002.

[vGFS+02] B. van Ginneken, A. Frangi, J. Staal, B. Haar, and R. Viergever. Active shape model segmentation with optimal features. *IEEE Transactions on Medical Imaging*, 21(8):924–933, 2002.

[WL01] T. Wendler and C. Loef. Workflow management - integration technology for efficient radiology. *Medicamundi*, 45(4):41–48, 2001.

[WRV98] J. Weickert, B. Romeny, and M. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3):398–410, March 1998.

[Yea07] David Yeager. Workflow efficiency - maybe the best answer to radiologys growing volume problem. *Radiology Today*, 8(16):20, August 2007.

[YHC92] A.L. Yuille, P.W. Hallinan, and D.S. Cohen. Feature extraction from faces using deformable templates. *International Journal of Computer Vision*, 8(2):99–111, 1992.

[ZB05] Y. Zhou and J. Bai. Atlas based automatic identification of abdominal organs. In *Proceedings of the SPIE Medical Imaging*, volume 5747, pages 1804–1812, April 2005.

[Zor00] Denis Zorin. Subdivision zoo. subdivision for modeling and animation. In *ACM SIGGRAPH 2000 Conference Course Notes*, number 23, pages 65–102, 2000.

[ZSN05] T. Zinßer, J. Schmidt, and H. Niemann. Point Set Registration with Integrated Scale Estimation . In R. Sadykhov, S. Ablameiko, A. Doudkin, and L. Podenok, editors, *Proceedings of the Eighth International Conference on Pattern Recognition and Image Processing* , pages 116–119, 2005.

Lebens- und Bildungsgang

13. Februar 1977	geboren in Tübingen
1983 - 1987	Grundschule Hirschau
1987 - 1996	Wildermuth-Gymnasium, Tübingen, Abschluß Abitur
1996 - 1997	Zivildienst Mobiler Sozialdienst, DRK Tübingen
1997 - 2000	Studium der Biochemie an der Eberhard-Karls-Universität Tübingen, Vordiplom
2000 - 2004	Studium der Bioinformatik an der Eberhard-Karls-Universität Tübingen, Abschluss Diplom
2000 - 2004	Werkstudent Software-Entwicklung, itdesign GmbH, Tübingen
2002 - 2004	Wissenschaftlich Hilfskraft am Lehrstuhl für Graphisch Interaktive Systeme am Wilhelm-Schickard-Institut für Informatik der Eberhard-Karls-Universität Tübingen (Prof. Straßer)
2004 - 2005	Research Associate, Max-Planck-Institut für Informatik, Saarbrücken, Abteilung für Computergrafik (Prof. Seidel)
2005 - 2008	Promotionsstipendium (Siemens AG, Healthcare Sector, Magnetic Resonance Division, Erlangen)
2005 - 2008	Promotionsstudent Wilhelm-Schickard-Institut für Informatik der Eberhard-Karls-Universität Tübingen
seit Mai 2008	Application Engineer/Scientific Lead, Siemens AG, Healthcare Sector, Magnetic Resonance Division, Erlangen