Graz University of Technology

Institute for Computer Graphics and Vision

Dissertation

NONLINEAR INTRA-MODALITY REGISTRATION OF MEDICAL VOLUME DATA

Martin Urschler

Graz, Austria, November 2006

Thesis supervisors

Prof. Dr. Horst Bischof, Graz University of Technology

Prof. Dr. Daniel Rückert, Imperial College London

Abstract

The research interest in nonlinear image registration for medical intra-modality applications is a topic of constantly growing interest in the medical image analysis community. Areas of application include fusion of anatomical and functional data by angiography techniques, development of physiologic models of dynamic processes, investigation of organ/tumor growth or motion compensation in surgery. In the present thesis we review the large body of literature on nonlinear intra-modality registration with a special focus on CT image registration applications involving brain and thorax. Our contributions to the state-of-the-art reflect this variety by spanning from partially rigid over shape- and landmark-based feature registration using shape context and SIFT descriptors and full-grown anisotropic image-driven optic flow intensity-based methods to a hybrid nonlinear registration approach that effectively combines feature and intensity information. We also touch the important area of registration evaluation by proposing a framework to compare algorithms on synthetically deformed and clinical data. The necessary trade-off between algorithm efficiency and accuracy in the presence of important high-frequency information (vascular structures in thorax and brain) is thoroughly investigated, since we and our project partners from Siemens MED CT are interested in practically applicable algorithms. Our results show effective methods both from a theoretical and practical point of view.

Keywords. medical image analysis, nonlinear image registration, intra-modality applications, bone subtraction CT Angiography, partially rigid registration, thoracic perfusion imaging, surface-based matching & registration, landmark-based matching & registration, intensity-based nonlinear registration, hybrid landmark- and intensity-based registration, evaluation framework for nonlinear registration

Acknowledgments

Over the last few years a large number of people influenced or helped me on my ”quest” for achieving the PhD degree. I am aware of the great importance that discussions with project partners, colleagues, other researchers, friends and family have on the final outcome of a PhD thesis and on the personal development of a person. Here I finally have the opportunity to express my gratitude.

First, I would like to acknowledge the staff at the Institute for Computer Graphics and Vision. My thesis supervisor Prof. Horst Bischof supported me by leaving enough space to develop my own ways while at the same time always having an open ear for problems and questions and providing very useful hints, a balance which is not so easy to hold. For being the most important source of discussion on project- but also non-project-specific content I want to thank him. Another thank you goes to Prof. Franz Leberl who convinced me of starting the PhD study program, supported me in writing the FWF proposal and for introducing me to Prof. Milan Sonka. In this context I would also like to mention Prof. Rainer Rienmüller from the University Hospital Graz, Department of Radiology, who supported me during my diploma thesis and at the beginning of my PhD thesis, both financially and in terms of providing basic medical knowledge about human anatomy and physiology.

A large advantage of the Institute for Computer Graphics and Vision is its large and constantly growing number and diversity of people that are involved in the different projects. Over the last years I have seen some of my friends become colleagues and many colleagues who I consider friends now. A fruitful environment like this provides the basic prerequisite to perform exciting research by encouraging a discussion culture. At this spot I want to mention Alex, Andi, Ben, Bernhard, Christopher, Fritz, Georg, Helmut, Joachim, Konrad, Lukas, Markus, Martina, Martin, Mario, Matthias, Michael D., Michael G., Peter, Pierre, Roli, Tom and especially Reinhard who has laid the foundations of the growing medical imaging group at the Institute. Many thanks also to the staff at the secretary office, Renate, Christina and Manuela, and to Andi.

During my PhD project I also had the opportunity to work with a number of students who supported me and from whom I also had the pleasure to learn things. Stefan supported

my work by doing a master project and a diploma thesis. Student projects were successfully finished under my supervision or under my consulting by Alex, Bernhard, Christian, Marc, Martin, Michi, Tobias, Sasa and Werner.

From the academic side I would further like to thank Prof. Milan Sonka and Prof. Eric A. Hoffman from the University of Iowa for hosting me at their Institutes a total of four months during two visits in 2002 and 2003, stays which were financed by Graz University of Technology. Their support on the FWF proposal and the discussions with them and some people in Iowa City (Jürg, Mark, Steve, Andreas) had significant influence on my later work. I am also grateful that Prof. Daniel Rückert from Imperial College London agreed on being my second supervisor.

Most of the work of this PhD thesis would never have been possible without the financial support of Siemens MED SP, Erlangen, Siemens MED CT, Forchheim and Siemens PSE AS, Graz. I would like to acknowledge the support of Hendrik and Rainer in Germany, Herbert and Peter in Graz and the discussions with my Siemens colleagues Matthias, Michael and Ruxandra. I especially appreciate that Herbert and Peter enthusiastically supported and support me and my work.

I am truly convinced that the success of a project like a dissertation also crucially depends on ones private environment and the support from friends and family. Here I would like to thank my mother for her sacrifices to enable my education and my brother and all of my friends for being there and helping me to relax and forget work for some time during a week. Finally I want to express my sincere gratitude to my girlfriend Gerti who accompanies me on my journey through life and supports me with her love.

Graz, November 2006

Martin Urschler

”*Wehre nichts ab, was zu dir kommt, und halte nichts fest, was von dir geht.*”
(Buddhist saying)

Contents

Chapter 1

Introduction

Contents

Image registration is a prerequisite for a large number of applications in computer vision and image processing especially in medical image analysis [Maintz and Viergever, 1998]. Informally image registration is the process of aligning two images of the same or similar objects such that identical structures in both images are brought into *correspondence* thereby minimizing an application-dependent *dissimilarity* (or distance) measure under the assumption of a specific *transformation* (or deformation) model [Hajnal et al., 2001]. There exists a large body of research literature on image registration, the main distinction typically is made between registration using rigid/affine and nonlinear transformation models. While rigid registration is considered a solved problem in most application areas, research on nonlinear image registration techniques is still a *"hot topic"* and there are no signs that this will change in the near future [Crum et al., 2004].

There are various open questions and tasks in the area of nonlinear registration. One problem is a lack of general-purpose techniques which leads to a large number of slightly or significantly differing approaches and application-dependent solutions. However, there is currently no agreement in the research community if the problem-specific design of nonlinear registration algorithms might ever be replaced by a generic algorithm. Another issue is the large computational effort that is necessary to apply nonlinear registration techniques on real-world data especially in the medical domain. Efficient registration very often goes at the expense of tolerating lower accuracy for smaller (i.e. high frequency) information. The fact that evaluation of nonlinear registration is only possible in a rudimentary way due to the lack of ground truth data further complicates the situation. Consequently it is difficult to compare algorithms objectively in order to separate suitable from unsuitable methods. This has lead to an increasing interest in evaluation methods over the recent years [Crum et al., 2003, Schnabel et al., 2003]. Most of these challenges boil down to the notion of nonlinear registration being an ill-posed problem, where one has more unknowns in the transformation model than one has constraints.

Even if restricted to medical imaging, the research literature on nonlinear registration shows a vast and extensive amount of publications concerning both methods and applications [Maintz and Viergever, 1998]. This variety comes from the large number of medical applications needing nonlinear registration, and it also indicates that the research community is not yet fully satisfied with both the practical solutions and the theory behind it. In accordance with this diversity we also investigate a wide range of nonlinear registration techniques in the present thesis. The contributions of this thesis range from rigid registration as a building block for partially rigid bone registration over surface and landmark-based feature registration to intensity-based approaches. A hybrid approach combining feature and intensity-based registration tries to fuse the contrarious strengths of these disciplines. An evaluation that compares state-of-the-art methods with our contributions on our specific medical image data sets concludes this work.

The remainder of this introductory chapter defines the image registration problem in more detail, motivates a number of medical applications where rigid and especially nonlinear registration is essential, and describes the imaging modality which is used in this work. Afterwards formal definitions of the general and the nonlinear image registration problems are given to define a common mathematical language for the later chapters. Finally the goals and contributions of this thesis are summarized.

1.1 The Image Registration Problem

Image registration is a method to establish *spatial correspondence* between two images. Images to be registered may be taken at different times or with different imaging sensors (modalities). In traditional computer vision, image registration is needed e.g. for 3D reconstruction from two several camera images of a scene, tracking of objects in video sequences, model-based object recognition from databases or registration and fusion of different imaging sensors. General surveys can be found in [Brown, 1992] or [Zitova and Flusser, 2003].

Over the last 25 years especially medical image analysis, which has been selected one of the 11 most important medical developments over the last 1000 years by the *New England Journal of Medicine* [NEJM-Editors, 2000], was a driving force for image registration research [Hill et al., 2001]. The main reason for this progression was the growing number of imaging modalities complementing one another (X-ray Computed Tomography (`CT`), Magnetic Resonance Imaging (`MRI`), Positron Emission Tomography (`PET`), Single Photon Emission Computed Tomography (`SPECT`) or Ultrasound (`US`)) which are nowadays routinely used in clinical practice. Various demanding applications like fusion of anatomical and functional information, surgery and radiotherapy planning, studies of tumor growth, correction of motion artifacts, statistical investigation of population-dependent anatomical differences or segmentation and classification by atlas registration have emerged and are getting used by physicians for diagnosis. General overviews and seminal surveys of medical image registration can be found in [Maintz and Viergever, 1998], [Lester and Arridge, 1999], [Crum et al., 2004] or the book [Hajnal et al., 2001].

In medical image analysis one distinguishes intra- (mono-) and inter- (multi-) modality registration. *Intra-modality* registration is the case where two patient scans have to be aligned which were acquired with the same modality, e.g. `CT-CT` or `MRI-MRI` registration. In contrast *inter-modality* registration intends to align structures depicted with different imaging modalities, e.g. `CT-PET` or `CT-MRI`. This sub-discipline is more demanding since identical anatomical structures may or may not be present on both images. Another important distinction has to be made concerning intra- and inter-patient registration. While *intra-patient* registration deals with the same anatomy, *inter-patient* registration further introduces difficulties due to population-dependent shape variations of anatomical structures, an emerging sub-discipline which is also known as *Computational Anatomy* [Miller et al., 2002, Li et al., 2003].

The notion of *spatial correspondence*, which is at the heart of image registration (compare Figure 1.1), involves the search for a transformation that relates the information given in one image to the information in the other one. The number of parameters needed to define this transformation is denoted as its Degrees of Freedom (`DoF`).

In general a minimal amount of a priori knowledge about the registration problem has to be provided by selecting the kind of geometric transformation one wants to estimate during the alignment procedure [Hajnal et al., 2001]. We distinguish between affine and nonlinear transformations. The frequently encountered rigid transformations are a special case of affine transformations. In the rigid case the number of `DoF` is restricted to translational and rotational parameters, i.e. 6 `DoF` for 3D volumes, while the affine case leads to at most 12 `DoF`. On the other hand nonlinear transformations may easily involve thousands of unknowns (like e.g. when deriving a spline transformation given a

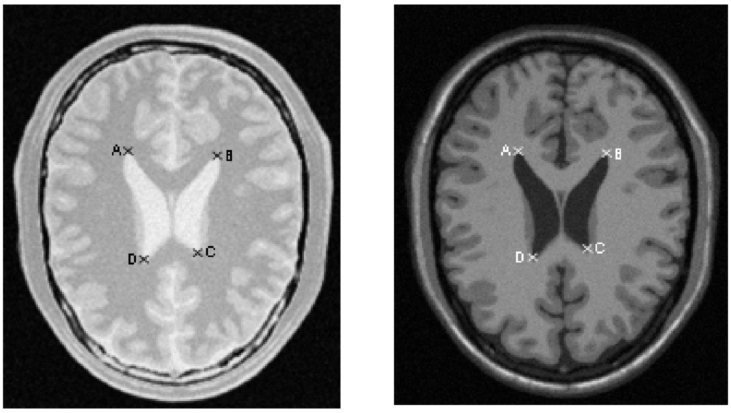
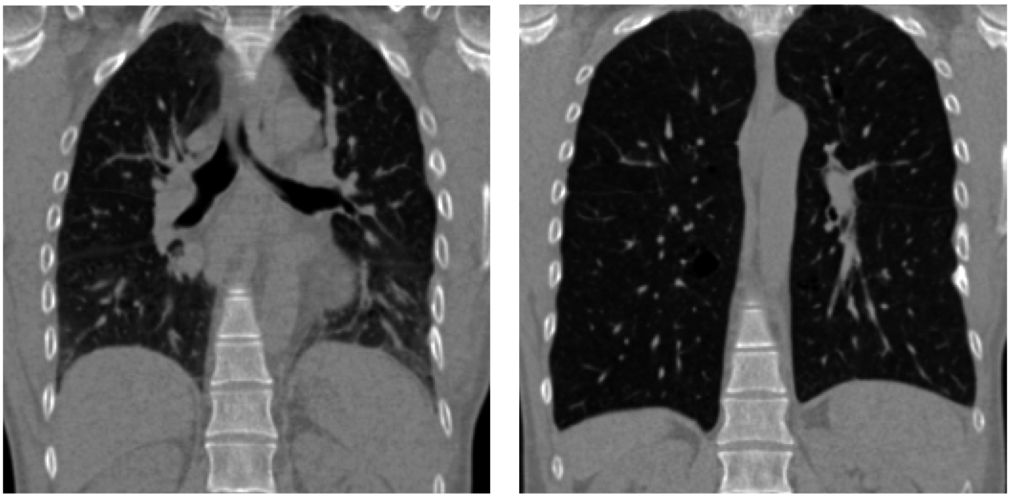


Figure 1.1: Illustration of spatial correspondence in image registration using two MR brain images with differing imaging protocols. Spatially corresponding structures of the anatomy, like the four landmarks A,B,C and D, should match each other after registration.

large number of corresponding landmarks) or several hundreds of thousands `DoF` (like e.g. when modeling the transformation with a regular B-spline grid). Obviously this leads to substantially differing registration methods, since general nonlinear registration is an ill-posed problem in the Hadamard sense [Hadamard, 1902]. Hadamard states the three properties of well-posed problem, the existence of a solution, the uniqueness of a solution and the continuous dependency of the solution on the data (in some reasonable topology), with the latter two properties generally being violated in nonlinear registration problems. If a problem is not well-posed one needs regularization (e.g. a smoothness assumption on the searched transformation) to derive stable numerical solution algorithms. In this case there exists a large number of possible nonlinear transformations which may lead to similar registration results. Some of these transformations will make no sense at all, while others will be meaningful. This has to be considered in algorithm design and during evaluation. Note that in literature the term *non-rigid* is often used for nonlinear registration as well, however we will denote every registration technique that models a transformation with more `DoF` as an affine transformation as *nonlinear*.

In the nonlinear registration case, the idea of *spatial correspondence* might be envisioned by thinking of a patients thorax who is imaged twice with a `CT` scanner under breath-hold, where the two images differ by the respiratory state. Consequently the process of nonlinear registration will establish which point on one image *corresponds* to a particular point on the other image, where *correspondence* means that points represent measurements at the same anatomical locations within the patient. The correspondence comes from a specific choice of the parameters of the transformation model, most often established by an optimization procedure defined on a dissimilarity (objective) function.

With this transformation model it is possible to *warp* one image into the coordinate frame of the other image, thereby enabling the direct comparison or the *fusion* of the information of the two input images. Figure 1.2 shows an example input for a nonlinear image registration problem involving coronal slices of two thoracic `CT` data sets which differ in their breathing state.



(a) (b)

Figure 1.2: Example input for an image registration problem. Corresponding coronal two-dimensional slices of two three-dimensional thoracic CT data sets without registration, (a) at exhalation and (b) at inhalation. Note that breathing motion occurs in three dimensions, thus not all anatomical structures correspond on these slices.

1.2 Medical Motivations

Medical image registration is necessary in a variety of application areas, the number of applications is constantly growing with the development and improvement of imaging modalities. Modalities depict shape, structure, size and spatial relationships of anatomical structures as well as information about function and pathology. Combination of these different types of information also contributes to the large number of application areas. From the large range of possible areas we are specifically interested in brain and soft-tissue CT Angiography (`CTA`) applications to fuse anatomical and functional data, therefore we will lay more emphasis on the description of these areas.

1.2.1 Rigid Body Image Registration Applications

When being confronted with patient images from one or several modalities, the physician's task is to build a combined model of this information to be able to derive a diagnosis. Here it becomes necessary to *mentally* compensate for changes in subject or object position with the help of computer-aided image registration. Nowadays, many clinical applications for image registration involve brain images due to its rigidity constraint, other areas like heart, thorax or abdomen are still more in the research stage due to the complexity of soft-tissue deformations. One such brain imaging example is the 3D imaging of perfused cerebral blood volume in acute stroke using `CT` [Ditt et al., 2003]. A good overview on further clinically relevant rigid body image registration applications can be found in Chapters 7 to 12 of [Hajnal et al., 2001]. Registration of serial `MRI` brain images are used to detect changes in the brain and allied structures for the purpose of disease progression analysis and monitoring of treatment response [Oatridge et al., 2001]. `MRI` or `CT` images (depicting morphological information) are combined with `PET` or `SPECT` images (showing functional information) to successfully locate tumors. Here the main registration problem is patient movement during the scans, this application area is also currently restricted to the brain in a clinical setting [Pietrzyk, 2001], however there is research going on for other areas like thorax and abdomen [Bailey, 2001] as well. Fusion of `MRI` and `CT` images is interesting for brain applications since `MRI` images depict the different soft tissue structures in the brain (white matter, gray matter, cerebrospinal fluid) with more contrast, while `CT` is useful since the cortical bone structures are invisible on `MRI` images. This is especially important in brain surgery planning [Hill and Jarosz, 2001].

The emergency evaluation of stroke and intracranial hemorrhage is nowadays primarily performed using multi-spiral `CT` [Lell et al., 2006a] and bone-subtraction `CTA` techniques. Patient movement during and between individual scans often lead to insufficient image quality due to motion artifacts. Rigid registration methods with high accuracy therefore are a basic building block in bone-subtraction `CTA` studies. Figure 1.3 shows an example of a Maximum Intensity Projection (`MIP`) visualization in a head `CTA` study.

1.2.2 Soft Tissue Image Registration Applications

Rigid body registration is very suited to applications in the brain, however, other organs are more prone to soft tissue deformations which have to be taken into account when relating images taken at different times. Sources of motion typically are respiration or heart-beat. Interestingly, even in brain imaging there is the necessity for nonlinear registration. In neurosurgery a *brain shift* occurs due to pressure changes, cerebrospinal fluid leakage, gravity effects and the mass of resected tumors consequently leading to nonlinear deformations [Clatz et al., 2005].

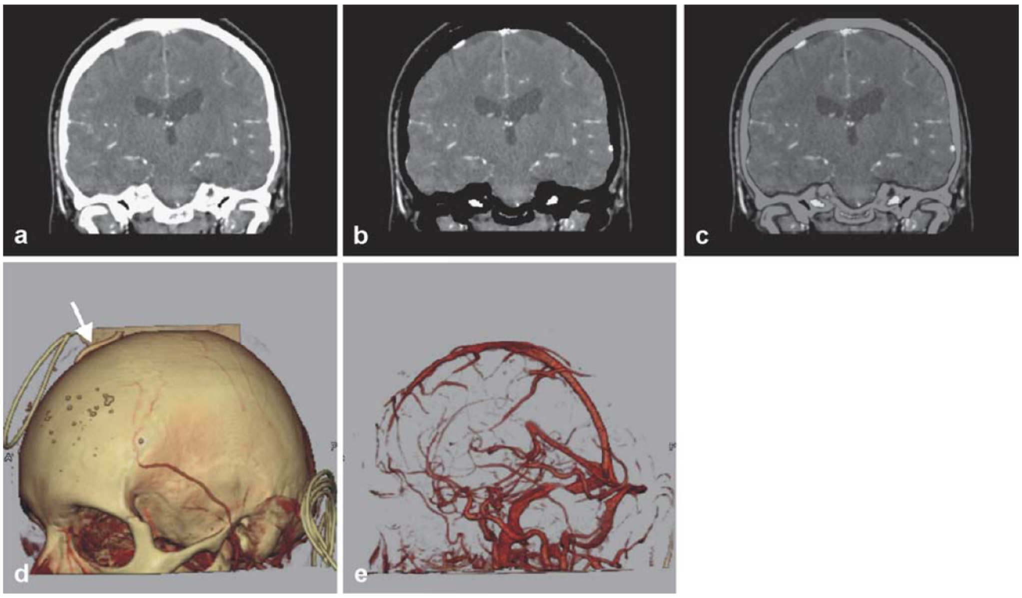


Figure 1.3: Example for a CTA head data set. a)-c) show the CT images with and without skull structures. d) is a MIP with and e) without skull structures. Taken from [Lell et al., 2006a]

One of the most important application of nonlinear image registration is the registration and fusion of anatomical and complementing functional data. The bone subtraction `CTA` application from the previous section might also be seen as a fusion application, where the contrast-enhanced scan is the functional image data. Many intra-modality fusion applications exist where the need to compensate for respiratory and cardiac motion artifacts is necessary, like registration of thoracic `PET` and `CT` images [Mattes et al., 2003] (see Figure 1.4). Also other time-consuming imaging modalities like `SPECT` require assessment of motion artifacts if one wants to relate these scans to `CT` or `MRI`.

Soft-tissue images from thoracic and abdominal data sets are the most prominent candidates for nonlinear registration applications since the organs of interest are highly deformable. According to the European Respiratory Society, lung diseases rank second behind cardiac diseases in terms of mortality and cost of treatment*. Computerized methods for objective, accurate and reproducible analysis of lung structure and function can provide important insights into these problems. However, due to the complexity of the breathing motion, investigations and applications working with thoracic but also abdominal images are often very complicated. There are several areas of application that could benefit from a proper treatment of problems due to breathing motion. Accurate fusion of highly-dynamic but spatially restricted functional (i.e. lung perfusion, lung ven-

---

[*] http://www.ersnet.org

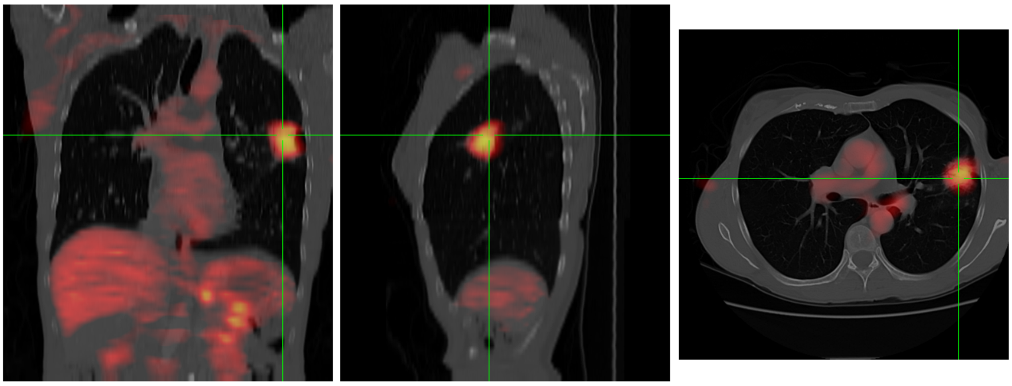


Figure 1.4: Example for a registration and color-coded fusion of a PET and a CT data set. Taken from the ITK public-domain courseware [ITK, 2006].

tilation) and anatomical data sets [Hoffman et al., 1995, Hoffman and McLennan, 1997, Maki et al., 1999a, Gefter, 2002] requires knowledge about the data sets' positions in the breathing cycle and about the possible soft-tissue deformations that are induced by diaphragm and rib cage movement. Anatomical data is here acquired over a single respiratory cycle, while functional data is acquired over several respiratory cycles with spatial resolution being restricted to just a small slab of a few slices. Based upon fusion of function and anatomy possible clinical applications using thorax data are studies of *Chronic Obstructive Pulmonary Disease (COPD)*, a group of diseases with fatal influence on gas exchange (e.g. Pulmonary Emphysema, Asthma), leading to a decrease in lung elasticity [Hoffman et al., 1995], or support studies for the clinical diagnosis of *Pulmonary Embolism* which can be investigated by detecting pulmonary perfusion disturbances [Wildberger et al., 2005].

In *radiotherapy planning*, which is an essential technique for the treatment and cure of thoracic or abdominal tumors and other lesions, the goal is to use different CT scans to track the location and growth of a tumor and to use these scans for very accurately determining the volume that confines it. After determining the tumor region high-energy radiation beams are focused precisely on the volume of interest to apply a high dose on the tumor only. Obviously it is very important to accurately compensate for physiological movement of the organs and the patient to prevent damaging healthy tissue [Weruaga et al., 2003]. Another method of dealing with cancer is surgical resection, which may or may not be effective depending on the cancer stage. Location of metastases and lymph nodes is an essential information for the accurate determination of the cancer stage, and might be used in surgery planning systems. Again the physiological movement of organs and patient involve complications in this planning approach [Tai et al., 1997]. Surgery planning systems often lead to decisions on the feasibility and risk of lung volume reduction surgery [Slone et al., 1997, Maki et al., 1999b] or liver surgery planning [Bornik et al., 2003]. Diagnosis and prognosis of cancer generally also depend upon growth assessment, repeated `CT` studies are used to assess the growth of pulmonary nodules. This leads to the issue of nonlinear nodule registration [Betke et al., 2003] to diagnose pulmonary metastases in oncology patients.

1.3 Imaging Modalities

While there is a large variety of imaging modalities in clinical practice, we are focusing on `CT` images. However, many of the developed algorithms are applicable to other imaging modalities as well with a slight adaptation effort. More specifically the `CT` modality delivering the image data is a Multi-slice Spiral CT (`MSCT`) scanner [Beutel et al., 2000]. `MSCT` is the current state-of-the-art in `CT` imaging, an imaging discipline that has its roots in the work of *Allan M. Cormack* and *Godfrey N. Hounsfield*, who were awarded with the Nobel Prize in Physiology and Medicine in 1979. `MSCT` scanners like the *Siemens SOMATOM Sensation*$^{\text{TM}}$are equipped with 64 x-ray detector rows, making it possible to scan the whole thorax in 20 to 30 seconds with an isotropic resolution of around 0.5 mm.

We are working with three kinds of medical image data. First, there are human brain images from 64-slice `MSCT`. The acquisition protocol of the brain images is a `CTA` protocol (in order to perform bone subtraction `CTA`) that involves a native image and an image with non-ionic iodinated contrast medium injected. Second, there are human thorax images at different breathing states between inhalation (Total Lung Capacity (`TLC`)) and exhalation (Functional Residual Capacity (`FRC`)), either with or without contrast medium injected. Both kinds of images were provided by Siemens MED CT, Forchheim, Germany. The last group of image data comes from `MSCT` sheep scans. The sheep `CT` data was provided by Prof. Eric Hoffman, University of Iowa, IA. The data was acquired with 64-slice scanners at five different breathing states between inhalation (`TLC`) and exhalation (`FRC`) by a protocol where breath is held at fixed inspiration levels during the 30 sec scan time leading to a static breathing scheme, which has to be considered for the interpretation of derived motion models from matched and registered shapes.

1.4 A Formal Definition of Image Registration

For later reference we now give a formal definition of the general image registration problem. We use a framework that was inspired by [Modersitzki, 2004]. Given two images, we refer to the target (reference) image as *fixed image* $I_F(\mathbf{x})$, representing the coordinate frame of reference, and to the source image (template) as *moving image* $I_M(\mathbf{x})$. $I_F(\mathbf{x})$ and

$I_M(\mathbf{x})$ map spatial coordinates $\mathbf{x}$, belonging to a certain set Ω, to intensity values. The dimension of the spatial domain is given by $d \in \mathbb{N}$, where $d = 3$ in this work.

$$I_F(\mathbf{x}) : \Omega_F \to \mathbb{R}, \qquad \Omega_F \subset \mathbb{R}^d$$

$$I_M(\mathbf{x}) : \Omega_M \to \mathbb{R}, \qquad \Omega_M \subset \mathbb{R}^d$$

We define $Img(d)$ as the space of d-dimensional images, with $I_F(\mathbf{x}), I_M(\mathbf{x}) \in Img(d)$. In image registration one is trying to determine a *spatial transformation* $\varphi(\mathbf{x}_F) : \Omega_F \to \Omega_M$, which defines the mapping between fixed and moving image coordinate frame (see Figure 1.5), such that the *dissimilarity* between fixed image $I_F(\mathbf{x})$ and *deformed* moving image $I_M(\varphi(\mathbf{x}))$ is minimized. To measure dissimilarity the notion of distance measures has to be introduced. A distance measure $\mathcal{D}$ takes two images as input and computes a dissimilarity score. The registration problem therefore translates to the following minimization problem:

Problem 1.1. Given a distance measure $\mathcal{D} : Img(d)^2 \to \mathbb{R}$ and two images $I_F(\mathbf{x}), I_M(\mathbf{x}) \in Img(d)$, find a mapping $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ such that $\mathcal{D}(I_F(\mathbf{x}), I_M(\mathbf{x})) = min$.

Very often this minimization problem has no direct solution, such that an optimization algorithm has to be applied as the numerical backbone of the registration procedure. Note that the transformation φ is defined in the fixed image coordinate frame, i.e. it maps points $\mathbf{x}_F$ to locations in the moving image coordinate system $\mathbf{x}_M$. From an implementation point of view the resampling of a warped moving image is often performed in the inverse direction, a so-called backward mapping to avoid holes and ambiguities in the warped image (see the excellent explanation of this concept in the ITK software guide [ITK, 2006]).

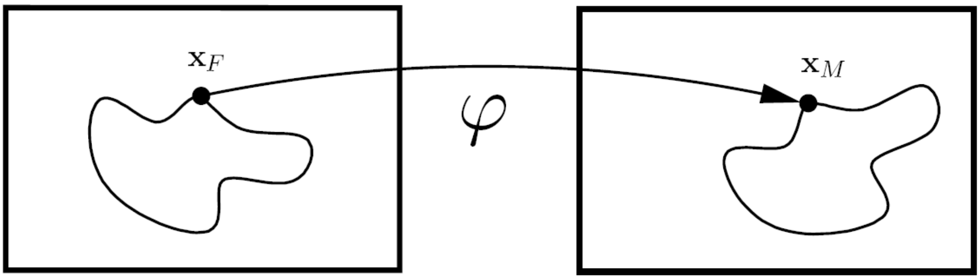


Figure 1.5: Image registration is the task of finding a spatial transform mapping the moving image into the fixed one. Taken from the ITK software guide [ITK, 2006].

We basically distinguish between affine and nonlinear transformations. In the following we will formally define the affine transformation and show simple extensions to a nonlinear transformation. In later sections of this thesis we will show further definitions of nonlinear transformations, which are more suitable in different areas of application. Affine transformations allow for rotation, translation, scaling and shearing. A 3D affine transformation $\varphi(\mathbf{x}_F)$ can be represented in homogeneous coordinates as:

$$\mathbf{x}_M = \begin{pmatrix} x_M \\ y_M \\ z_M \\ 1 \end{pmatrix} = \varphi(\mathbf{x}_F) = \mathbf{A} \cdot \mathbf{x}_F = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_F \\ y_F \\ z_F \\ 1 \end{pmatrix}$$

Here the elements a_{14}, a_{24}, a_{34} model the translation, while the remaining 3x3 matrix models rotation, scaling and shearing.

The obvious extension to the nonlinear case would be to model the nonlinear transformation with second- or higher-order terms, however these models are rarely used in medical imaging, since their ability to recover shape changes is very global, while in practice deformations occur rather localized, and higher order polynomials tend to oscillate. From now on we will also refer to a nonlinear transformation φ as a *deformation* or *displacement field*. Instead of using a polynomial as a linear combination of higher-order terms, one can use a linear combination of n basis functions θ_i (e.g. trigonometric functions, wavelets or spline functions) to represent the nonlinear transformation:

$$\begin{pmatrix} x_M \\ y_M \\ z_M \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3n} \\ 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \theta_1(x_F, y_F, z_F) \\ \vdots \\ \theta_n(x_F, y_F, z_F) \\ 1 \end{pmatrix}$$

So far the transformations have been defined in a *parametric* framework, note that regularization is implicitly incorporated by the choice of the basis functions in the example stated above [Rueckert et al., 1999]. However, in nonlinear registration one is often interested in *non-parametric* techniques. The basic idea is to come up with an appropriate measure both for the similarity of images as well as for the likelihood of a non-parametric transformation. This leads to a variational formulation [Lanczos, 1986] of the image registration problem consisting of a similarity measure term and a transformation regularization term. The minimizer of the functional boils down to a partial differential equation that stems from the Euler-Lagrange equations. Note that the displacement field φ now is treated independently for the whole image domain (no parameterization). In the variational setting [Modersitzki, 2004] the general nonlinear registration problem reads:

Problem 1.2. Given two images $I_F, I_M \in Img(d)$, a positive regularizing parameter $\alpha \in \mathbb{R}_+$, a distance measure $\mathcal{D}$ and a regularization method $\mathcal{S}$, find a deformation $\varphi : \mathbb{R}^d \to \mathbb{R}^d$

such that $\mathcal{J}[\varphi] := \mathcal{D}[I_F, I_M; \varphi] + \alpha\mathcal{S}[\varphi] = min.$

Note here that the terms are functionals instead of functions, i.e. $\mathcal{J}[\varphi]$ is an expression that depends itself on an unknown function φ. In later chapters we will present distance measures, regularization methods and numerical solution strategies for this general framework.

1.5 Aims of the Thesis

In this thesis we concentrate on nonlinear (deformable) intra-modality intra-patient registration of large volumetric data sets for `CTA` applications. If one considers the constantly growing size of volumetric data sets like e.g. modern `CT` scanners are producing, there clearly exists the need to come up with more efficient algorithms in practice, that still perform accurate nonlinear registration. State-of-the-art algorithms for nonlinear registration often show runtime and memory efforts that scale with the size of the input data sets, which is an unwanted behavior given constantly growing data set sizes. As a consequence the challenging memory and runtime requirements have to be considered very early in the design stage of registration algorithms.

More specifically we aim for the development of fast and accurate registration algorithms that range from applications in the brain to soft-tissue thorax applications. Our investigated transformation models range from rigid over partially rigid to fully nonlinear models. The most important application of our work is in the area of fast and accurate thoracic `CT` data registration, where one typically is confronted with large data-sets on the order of 512^3 voxels and there is the need for accurate registration of vessel structures in the lung (vascular and airway tree) in the presence of large breathing deformations. Further complications are due to contrast agent application, partial volume effect and disease structures in the images. Standard algorithms from the literature either take on the order of hours to perform accurate nonlinear registration or they fail in vessel registration if terminated too soon. In this case our goal is to investigate automated methods that scale with the number of features extracted from a data set. Another important issue we focus on is the need to create standardized evaluation frameworks and to compare registration algorithms given a specific task. This is an issue that is not yet dealt with in literature sufficiently.

There are three different application areas we are especially focusing on. First, the registration of head and neck `CTA` images that introduce several independent but rigid movements which have to be compensated for in order to allow `MIP` visualization of vascular structures. Second, the nonlinear registration of large thoracic `CT` data sets at several different breathing states in order to extract breathing motion of diaphragm, lung surface or airway tree. Third, the fast and memory-efficient nonlinear registration of large thoracic `CT` data sets in `CTA` applications for perfusion studies, where a large amount of breathing motion difference may exist in the two (native and contrast-enhanced) data sets. A compact summary of our aims looks as follows

From an application point of view our project aims are

Partially Rigid Bone Registration Develop a registration method for modeling the motion of bones in head and neck `CTA` studies. The task is to find a partially rigid registration that matches bones of one image with its counterpart in the other image even in the presence of individual bone movements. This application can be used to visualize vessel structures without disturbing bone structures in `MIP` or in orthopaedic surgery to register pre-operative `CT` to intra-operative C-Arm images.

Assessment of Breathing Motion Develop a nonlinear registration method for modeling the breathing motion of the lung from several different static breathing states. The input for this task consists of native `CT` thorax scans at two or more different breathing states between `TLC` and `FRC`. In this task segmentation methods will be used to locate thoracic structures. Registration methods will be developed to find the nonlinear deformation fields that occur due to breathing. This task is intended to provide insights on and create models of breathing motion.

Nonlinear Registration of Thorax CTA Images Develop a nonlinear registration method to be able to fuse lung images from different image acquisition protocols which were taken at distinct breathing states. The input for this application consists of native anatomical lung `CT` scans and lung `CTA` scans which have to be combined in an accurate way. Therefore, it is necessary to revert the influence that breathing has on the deformable lung tissue by using nonlinear registration techniques. This application is intended as an extension of previous anatomical and dynamic `CT` fusion work for diagnosing brain images for acute ischemic stroke patients [Ditt et al., 2003] or for diagnosing pulmonary embolism [Wildberger et al., 2005].

Nonlinear Registration Evaluation Framework The evaluation of the accuracy and the consistency of nonlinear registration algorithms is a hard problem. Typically this requires synthetically transformed data sets that are put into the registration algorithm to compare registration results with the synthetic ground truth. This can be performed by comparing warped images, therefore a standardized evaluation framework should be created to support this task. For evaluation the notion of ill-posedness has also to be kept in mind, nonlinear registration techniques often include implicit regularization assumptions, this makes it necessary to compare resulting transformations as well.

1.6 Structure & Contributions of the Thesis

The remaining chapters of this thesis are structured as follows. Chapter 2 investigates the current state-of-the-art in rigid and especially nonlinear image registration, and gives a general overview on widely used methods as well as a detailed description of related work on the specific problem domains we are focusing on. Chapter 3 describes an algorithm for partially rigid bone registration which is used to remove bone structures from brain `CTA` images. Further, an evaluation of the algorithm on several data-sets is presented. In Chapter 4 a surface-based registration algorithm using the shape context approach is presented and evaluated on thoracic `CT` data sets. Chapter 5 generalizes the surface-based registration approach to become independent of a required pre-segmentation step by directly extracting, matching and registering features from the images. The topic of Chapter 6 is a hybrid registration method that combines the advantages of the feature-based registration method with an intensity-based algorithm formulated in a variational framework. Here the best of both worlds (feature- and intensity-based registration) are effectively combined to form a flexible solution of our registration problem. Chapter 7 takes some of the proposed algorithms and compares them with each other and with several widely-used algorithms from the literature in an evaluation framework using several types of synthetic and real clinical data sets. Finally Chapter 8 discusses the results of the evaluations, summarizes the contributions of this thesis and concludes with an outlook on future work. A list of publication that arose from the work on this thesis can be found in Appendix D.

Chapter 2

State of the Art

Contents

To perform a full review of the state-of-the-art on image registration would be an enormous task, even the restriction to medical image registration leaves an extremely large number of publications from the last two decades. An important survey paper for image registration techniques in computer vision, image processing, medical imaging and remote sensing is [Brown, 1992]. A more recent survey on image registration methods in computer vision can be found in [Zitova and Flusser, 2003], where approaches relating features extracted from the images are dealt with more extensively than area- (i.e. intensity-) based methods.

Concerning medical imaging there are a few standard books that provide overviews on image registration for medical applications. The book [Hajnal et al., 2001] is a standard compendium focusing specifically on medical image registration, while [Sonka and Fitzpatrick, 2000] and [Bankman, 2000] focus more generally on medical image analysis techniques, however they also include several chapters on medical image registration. An important, frequently cited survey on medical image registration has been published in [Maintz and Viergever, 1998]. There one can also find references to older surveys on medical image registration. In this work a large number of relevant papers are reviewed and a classification of registration methods according to nine distinct criteria (dimensionality, nature of registration basis, nature of transformation, domain of transformation, interaction, optimization procedure, modalities, subject and object) is given, a scheme which was adapted in other review papers later-on.

There are other survey papers which focus more specifically on sub-disciplines in medical image registration. [Audette et al., 2000] give an algorithmic survey on automatic 3D surface registration techniques, i.e. methods to relate segmented 3D structures from different data sets. Surface registration is broken into three building blocks (choice of transformation, surface representation & similarity criterion, matching & optimization). Their most important distinction is according to the type of surface representation (sets of sparse, discriminative features, dense point clouds, physically deformable 3D models and global shape models). A survey on Mutual Information (`MI`) based registration methods can be found in [Pluim et al., 2003]. `MI` is an important similarity measure derived from information theory that recently gained lots of attention in the medical imaging community. `MI` methods are classified according to methodological aspects, implementation issues and applications. [Lester and Arridge, 1999] focus on the importance of hierarchical methods in nonlinear medical image registration. Coarse-to-fine strategies improve execution speed and robustness against local minima significantly. The classification of hierarchical methods shows three groups, hierarchical data representation (e.g. Gaussian pyramid), hierarchical deformation model representation (e.g. grid size of B-spline transformation) and hierarchical model complexity (e.g. rigid - affine - nonlinear strategy). A very recent, extensive overview paper of the state-of-the-art in nonlinear medical image registration can be found in [Crum et al., 2004]. A critical comment on validation strategies in nonlinear medical registration is given in [Crum et al., 2003] with the main conclusion that current clinical validation studies in the literature are not thorough enough. For more review articles dealing with the application of medical image registration methods in several areas one may follow the references in [Crum et al., 2004].

After reading these seminal survey papers we decided to classify methods according to their type of transformation into affine and nonlinear techniques. On the next finer classification level we distinguish between feature- and intensity-based methods, the difference being that intensity- (or voxel-) based methods always work on whole image volumes while in feature-based methods a pre-processing step either extracts point(s) of interest (`POI`) (fiducial markers, anatomical landmarks) or segments structures which are to be matched prior to registration.

2.1 Rigid and Affine Registration

Based upon the formal definition of image registration in Section 1.4 we now introduce the main concepts of 3D rigid and affine registration. See the book [Hajnal et al., 2001] for additional details about this topic. We split this overview into two parts, *feature-based* registration methods involving extracted points and surfaces and *intensity-based* registration methods involving whole volumes. Two additional sections deal with multi-resolution hierarchies in image registration and a literature overview on bone removal techniques in `CTA` applications.

2.1.1 Feature-Based Methods

Feature-based registration is usually dealing with structures of interest that have to be extracted from a volume. This might be *fiducial markers*, i.e. artificial objects attached to the patient which are well visible in a scan, *anatomical point landmarks*, i.e. manually or automatically derived points of interest representing prominent locations in the body, or *segmented surfaces*, i.e. an exact delineation of a structure of interest. Since rigid/affine feature-based registration is not of large interest for our later investigations, we will only present the basic ideas by describing a standard method in both categories, respectively.

2.1.1.1 Point-Based Methods - The Procrustes Alignment Problem

The simplest case of medical image registration uses *fiducial markers* which can easily be detected in 3D volumes from scanning modalities. A set of several corresponding 3D point pairs thus can be identified and used for a least-squares solution to estimate the unknown rigid or affine transformation parameters. So given a set of points $\{\mathbf{x}_F\}$ from the fixed image $I_F(\mathbf{x})$ and a set of points $\{\mathbf{x}_M\}$ from the moving image $I_M(\mathbf{x})$, one wants to estimate the unknown transformation $\varphi(\mathbf{x}_F)$ which is restricted to the space of rigid or affine transformations only. For the purpose of rigid registration this is called the *Procrustes Alignment Problem* and it leads to a least-squares solution giving an optimally fitting transformation that minimizes the cost function $G(\varphi) = |\varphi(\mathbf{P}_F) - \mathbf{P}_M|^2$ where $\mathbf{P}_F$ is a matrix containing row-wise the coordinates of the point set $\{\mathbf{x}_F\}$ and $\mathbf{P}_M$ the coordinates of $\{\mathbf{x}_M\}$. Procrustes alignment makes use of the special structure of a rigid transformation, as being composed of a rotational and a translational part. The translation is determined by computing the means of the point sets and the rotation is derived by a Singular Value Decomposition (`SVD`) in an instance of the orthogonal Procrustes problem. If φ is an affine transformation this kind of approach leads to a standard least square estimation. In principle the same approach can be used for *anatomical landmarks* instead of fiducial markers.

2.1.1.2 Surface-Based Methods - Iterative Closest Point

If the scanning modality allows for segmentation of the structures of interest a surface matching algorithm can be used for registration. Probably the most widely used surface matching algorithm for rigid/affine registration is the *Iterative Closest Point* algorithm [Besl and McKay, 1992]. This iterative algorithm has two stages. Given two sets of points $\{\mathbf{x}_F\}$ and $\{\mathbf{x}_M\}$ representing surfaces $\mathcal{S}_F$ and $\mathcal{S}_M$ respectively, the first stage of the algorithm involves identifying for each point $\mathbf{x}_{M,i}$ in $\{\mathbf{x}_M\}$ the closest point on surface $\mathcal{S}_F$. This is the point $\mathbf{x}_F$ for which the distance d between $\mathbf{x}_{M,i}$ and $\mathbf{x}_F$ is minimal.

$$d(\mathbf{x}_{M,i}, \mathcal{S}_F) = \min_{\mathbf{x}_F \in \mathcal{S}_F} \|\mathbf{x}_F - \mathbf{x}_{M,i}\|$$

As a result one gets a set of closest points $\mathbf{x}_{F,i}$. The second part of the algorithm estimates the transformation parameters φ by relating these two sets $\mathbf{x}_{M,i}$, $\mathbf{x}_{F,i}$ in a least squares manner. A standard way for this estimation is a Procrustes Alignment (see Section 2.1.1.1). After a transformation φ has been found, it is applied to $\{\mathbf{x}_M\}$ and the algorithm continues with the next iteration. A distance tolerance threshold is used such that the algorithm converges to a local minimum. A drawback of the method is its tendency to reach local minima. Initialization of the transformation therefore becomes an important issue. In practice the algorithm may be started multiple times with distinct initializations and the minimum of the distance threshold minima resembles the optimal solution.

2.1.2 Intensity-Based Methods

In contrast to feature-based registration, intensity-based methods do not extract corresponding points or structures of interest from volumes in a pre-processing stage. Instead they always work on the intensities of the whole volumes that are to be registered, therefore they are also referred to as *voxel similarity-based* registration techniques. By using all (or a large proportion of) the available voxel data, these methods have the advantage of averaging errors caused by noise fluctuations in the image data. Historically these methods have evolved from intra-modality applications to more demanding inter-modality applications, where in the first case a brightness constancy assumption between the volumes holds. Therefore, we will also present the different methods according to this development. Basically intensity-based methods have a structure which is composed of four main components, a *similarity metric*, an *optimization method*, a *transformation* and an *interpolation method*. Figure 2.1 shows the work-flow of these components.

The transformation is determined by the number of degrees of freedom that are to be estimated during registration. In rigid 3D registration, the transformation is modeled with 6 degrees of freedom, three to determine the unknown translation and three for rotation. An affine transformation is modeled with up to 12 degrees of freedom, adding scaling and shearing to the rigid case. The interpolation method is a crucial building block in terms of computational efficiency, since image interpolation due to changing transformation estimates is a very frequent task during the registration process. Nearest-

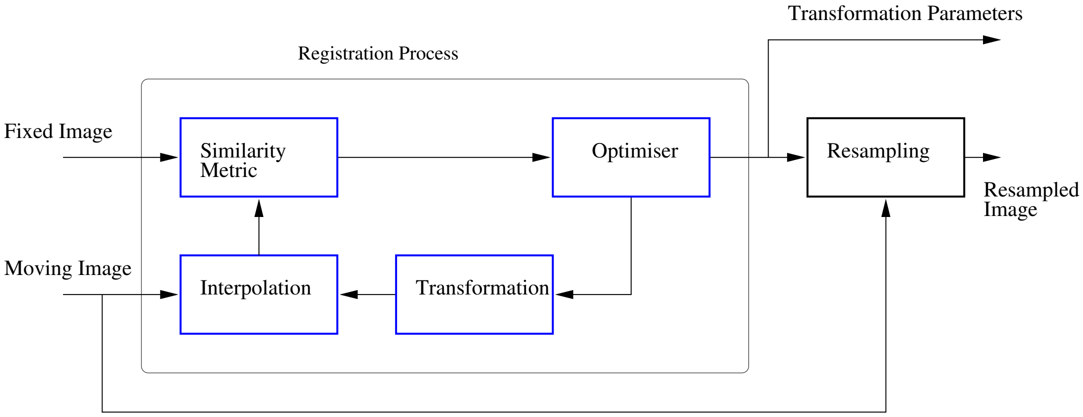


Figure 2.1: The basic image registration framework composed of the four components similarity metric, optimization method, transformation and interpolation method. Taken from the ITK public-domain courseware [ITK, 2006].

neighbor interpolation is simple and efficient but seldom used, while tri-linear optimization gives the best trade-off between accuracy and runtime. More elaborate interpolation methods (e.g. B-spline, windowed sinc) can be used but in practice are very inefficient in terms of runtime and memory consumption. Interpolation (i.e. resampling) and the direction of the estimated transformation are tightly connected. Figure 2.2 shows this connection more explicitly, there it can be seen that the "space transform" φ has to be performed in the physical fixed image coordinate system, as opposed to the discrete fixed image grid, and it leads to a point in the moving image coordinate system, where the interpolation has to be applied.

The choice of optimization method has to be carefully considered, very often standard gradient descent techniques are applied, however, these techniques are very slow in practice, more elaborate *Quasi-Newton* [Nocedal and Wright, 1999] approaches are much more appropriate. Finally the choice of similarity metric is very application-dependent, this choice will be the way how we distinguish methods in the following subsections.

2.1.2.1 Intensity Differences

A very simple similarity measure is the intensity-based Sum of Squared Differences (`SSD`) between images. During registration this measure is minimized. For voxels $\mathbf{x}_F$ and $\mathbf{x}_M = \varphi(\mathbf{x}_F)$ the `SSD` measure is:

$$SSD = \frac{1}{N} \sum_{\mathbf{x}_F} |I_F(\mathbf{x}_F) - I_M(\varphi(\mathbf{x}_F))|^2$$

The `SSD` needs to be normalized by the number of voxels N that lie in an overlap domain

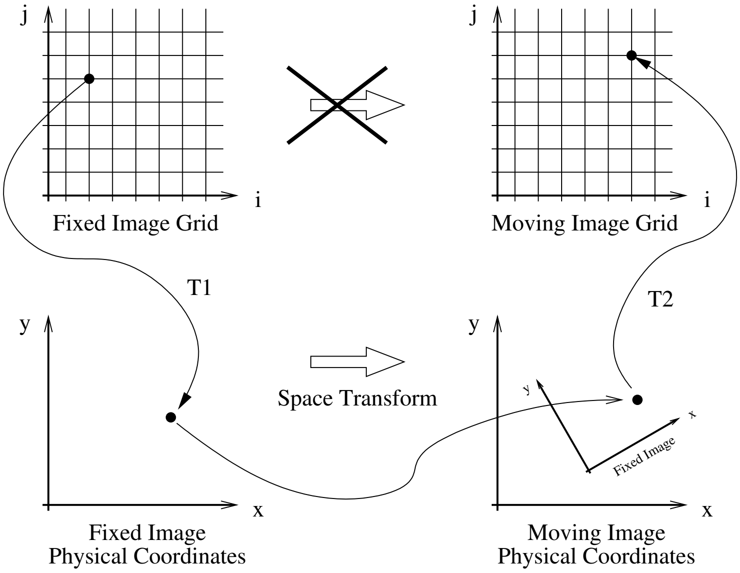


Figure 2.2: Different coordinate systems involved in the image registration process. The transformation being optimized is the one mapping from the physical space of the fixed image into the physical space of the moving image. Taken from the ITK public-domain courseware [ITK, 2006].

of I_F and I_M according to the transformation φ. It can be shown that this is the optimum measure when two images differ by Gaussian noise [Viola and Wells III, 1997]. However, in practice images do not simply differ solely by Gaussian noise, in case of injected contrast agents or due to metal artifacts there are many "outlier" intensity differences which the `SSD` is very sensitive to. Another measure which should be favored in theses cases is the Sum of Absolute Differences (`SAD`):

$$SAD = \frac{1}{N} \sum_{\mathbf{x}_F} |I_F(\mathbf{x}_F) - I_M(\varphi(\mathbf{x}_F))|$$

An example application that uses the `SSD` measure for clinical registration can be found in [Hajnal et al., 1995].

2.1.2.2 Normalized Cross Correlation

The `SSD` measure implicitly assume the images to differ by Gaussian noise, however, one can also think of less strict assumptions for intra-modality registration. If one only assumes a linear relationship between the image intensities in two images, the theoretically optimal similarity measure is the Normalized Cross Correlation (`NCC`):

$$NCC = \frac{\sum(I_F(\mathbf{x}_F) - \overline{I_F}) \cdot (I_M(\varphi(\mathbf{x}_F)) - \overline{I_M})}{\{\sum(I_F(\mathbf{x}_F) - \overline{I_F})^2 \cdot \sum(I_M(\varphi(\mathbf{x}_F)) - \overline{I_M})^2\}}$$

Here $\overline{I_F}$ is the mean voxel value in image I_F within the overlap region, and $\overline{I_M}$ is the mean of $I_M(\varphi(\mathbf{x}_F))$ within the overlap region. Example applications using `NCC` for registration are mentioned in [Hajnal et al., 2001], however, this measure has problems if there is partial occlusion of the structures of interest in both images.

2.1.2.3 Information Theory

With the emerge of functional imaging devices and the growing number of different scanning modalities used for clinical diagnosis, the need for inter-modality registration algorithms has been constantly growing over the recent years. The problem in inter-modality registration is, that the sensitivity to the same tissue might differ significantly between modalities, i.e. `CT` images show bones with very large intensity values while `MRI` images are not sensitive to bones at all. For this purpose [Woods et al., 1992] derived a similarity measure called *Ratio Image Uniformity* working with a derived ratio image calculated from I_F and I_M. Later, they refined this technique in the form of a voxel similarity measure that was used for `MRI-PET` registration [Woods et al., 1993]. This measure was called *Partitioned Intensity Uniformity* and it involved the distributions of the intensity values of the two modalities, assuming that "*all pixels with a particular MR pixel value represent the same tissue type so that values of corresponding PET pixels should also be similar to each other*". This measure was never widely used except for `MRI-PET` registration, however, it inspired follow-up research in the area of information-theoretic measures.

Information-theoretic similarity measures were independently proposed in the seminal papers of [Wells III et al., 1996] and [Maes et al., 1997]. An excellent review of these methods can be found in [Pluim et al., 2003]. The information-theoretic registration approach tries to *maximize the amount of shared information* in the two images. It can be thought of as reducing the amount of information in the *combined* image, leading to the use of *information* as a measure. A very commonly used measure of information in signal processing is the Shannon-Wiener [Shannon, 1948] entropy measure H which is the average information supplied by a set of N symbols whose probabilities are p_i. In other words, if an image I_F is regarded as consisting of a string of symbols, with each symbol having a certain probability of appearance. The expected amount of information $H(I_F)$ one can obtain from image I_F by probing the gray value of one (random) pixel is measured in *bits* and given by

$$H(I) = -\sum_{i=1}^{N} p_i ld(p_i).$$

In image registration two images I_F and I_M have to be aligned. There are two symbols at each corresponding voxel location linked by a transformation estimate φ. In this case joint entropy measures the amount of combined information. If the two images are totally unrelated, joint entropy is equal to the sum of the individual entropies of the images. The more similar the images are, the lower the joint entropy $H(I_F, I_M)$ compared to $H(I_F)$ and $H(I_M)$:

$$H(I_F, I_M) \leq H(I_F) + H(I_M)$$

The joint entropy $H(I_F, I_M)$ is defined as

$$H(I_F, I_M) = -\sum_{i=1}^{N}\sum_{j=1}^{N} p_{FM}^{\varphi}(i,j) ld(p_{FM}^{\varphi}(i,j))$$

and resembles the amount of information obtained from both images when probing pairs of gray values from the two images. $p_{FM}^{\varphi}(i,j)$ resembles the joint Probability Distribution Function (`PDF`) of the two images in their overlap domain. Joint entropy is not directly taken as a measure for image similarity in registration, since the estimated probabilities depend on the overlap of volumes which changes during registration. In the worst case it might happen that in medical images regions containing air are brought into correspondence without an overlap of image structures, since overlapping air minimizes the joint entropy. As a consequence, joint probability has to be measured in relation to the individual entropies. The mutual information (`MI`) measure

$$MI(I_F, I_M) = H(I_F) + H(I_M) - H(I_F, I_M) = \sum_{i=1}^{N}\sum_{j=1}^{N} p_{FM}^{\varphi}(i,j) ld\left(\frac{p_{FM}^{\varphi}(i,j)}{p_F(i) \cdot p_M^{\varphi}(j)}\right)$$

overcomes this problem. $p_F(i)$ here resembles the marginal `PDF` of the fixed image I_F and $p_M^{\varphi}(j)$ the marginal `PDF` of the moving image I_M. `MI` can qualitatively be thought of as measuring how well one image explains the other, it is maximized at optimal alignment.

The standard `MI` measure does not fully solve all overlap problems during registration. [Studholme et al., 1999] proposed the Normalized Mutual Information (`NMI`) measure to get rid of overlap problems. `NMI` is defined as

$$NMI(I_F, I_M) = \frac{H(I_F) + H(I_M)}{H(I_F, I_M)}$$

It is currently regarded as the state-of-the-art `MI` similarity measure for inter-modality registration.

2.1.3 Multi-Resolution Strategies for Practical Registration

Intensity based image registration often consists of a large number of parameters to optimize. Especially in nonlinear registration [Lester and Arridge, 1999] we will see that large parameter spaces lead to local extrema solutions which are undesired. As a consequence hierarchical multi-resolution approaches are always used in practical image registration algorithms, even in the comparably low-dimensional rigid/affine case [Pluim et al., 2003]. Input images are smoothed and down-sampled to coarser resolutions leading to Gaussian pyramids [Sonka et al., 1999]. Figure 2.3 shows the principle of a Gaussian pyramid as an input to a multi-resolution approach.

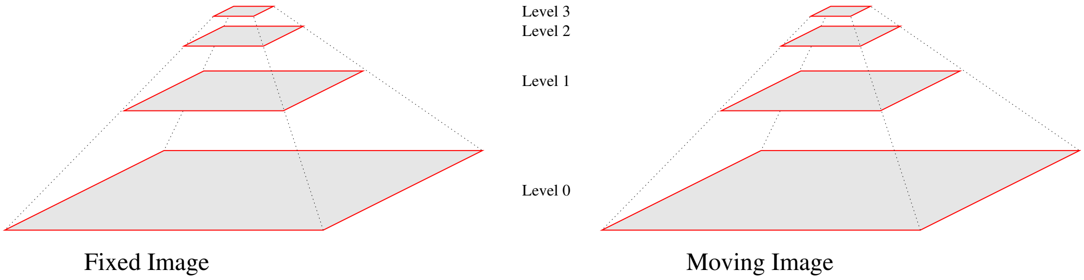


Figure 2.3: The principle of a hierarchical multi-resolution approach. It consists of Gaussian pyramids that resemble several resolution levels of the input images from coarse to fine. Registration starts with the lowest resolution and provides initial solutions for finer levels.

For each resolution level of the Gaussian pyramid, the registration process is repeated starting with the coarsest level. The estimated transformation parameters in a coarser level provide the input for the next finer level, while the coarsest level itself is initialized with a zero or identity transformation. Besides avoiding local extrema, an additional benefit of hierarchical approaches is the speed-up of the registration process since coarse levels provide low-frequency components of the desired transformation very efficiently due to the lower number of voxels to look at.

2.1.4 Bone Removal in Head CTA Applications

In this section we will summarize existing research work for highly accurate bone removal in head and neck `CTA` applications, which is a specialized area of medical imaging. `CTA` applications consist of native and contrast-enhanced `CT` images. In contrast-enhanced images vascular structures and bones have very similar intensities, such that simple segmentation algorithms like thresholding do not work which are frequently used for `CTA` images. The intended strategy for the removal of bone structures is to take a simple (threshold-based) bone segmentation taken from the native image and register it to the contrast-enhanced image. Registration is necessary instead of simple subtraction, since small patient movements may occur between the acquisition of both kinds of images. Especially in the neck area, these patient movements are too large for naive subtraction. Registration has to be very accurate in this area, since there are vessel structures that lie inside the bone structures as well. An accuracy that lies in the range of a half voxel is desired for this purpose. To perform accurate registration rigid intra-modal registration has to be used. It can be assumed that the bones themselves are rigid but the relative position of bones to each other may change due to small patient movements. Therefore each pair of bones are registered rigidly but the relative movements are taken into account, therefore leading to a partially rigid registration scheme.

A seminal overview on image registration techniques in Digital Subtraction Angiography (`DSA`) can be found in [Meijering et al., 1999]. In the literature there are mainly three directions to solve the problems in `CTA` studies where the accurate removal of bone structures is necessary to get rid of obstructed vessels. The first kind of approaches are rigid registration methods which are especially optimized for the bone removal application and may or may not incorporate nonlinear refinements. The second type of approaches are nonlinear registration techniques with or without rigidity constraints. The two kinds of approaches will be presented in the next sections in more detail. Finally, there are approaches which use segmentation algorithms in the contrast-enhanced and the native images to either detect bone structures and remove them or detect vascular structures and highlight them. In [Felkel et al., 2001] an extensive overview on vessel tracking algorithms is given which might be used for the segmentation of vessel structures in `CTA` images. [Alyassin and Avinash, 2001] propose a semi-automatic bone segmentation involving thresholding, region growing and morphological operations. [Kang and Shin, 2003] also show a bone segmentation involving region growing based on local adaptive thresholds and morphological operations followed by a boundary refinement. [Snel et al., 2002] and [Sebastian et al., 2003] show segmentation algorithms based on deformable models for wrist bones which might be adaptable for the problem at hand. However, all of these approaches are prone to typical difficulties inherent to segmentation algorithms like leaking, unstable threshold selection procedures and the need for parameter tuning. This is especially true if structures of interest have similar intensities and tend to intersect. Semi-automatic methods somehow overcome these problems but they add a lot of expert effort to each investigation. For these reasons we decided not to follow this direction any further.

2.1.4.1 Rigid registration approaches with or without nonlinear refinement

An influential paper in this area has been published by [Venema et al., 2001]. In this paper `CTA` was proposed for brain images by matching native and contrast-enhanced data sets and removing the bone structures by a rigid registration procedure. The Matched Mask Bone Elimination (`MMBE`) is a technique that finds for a native data set the corresponding bone structure voxels in a contrast-enhanced data set. The two 3D data sets therefore are matched by automatically registering the images with a rigid registration algorithm involving a gray value correlation measure and a downhill simplex optimizer. With this method, the mean of the squared differences between voxel values of native and contrast-enhanced image is minimized. Only pixels with a CT value between 600 and 800 Hounsfield Unit (`HU`) are used for registration. The matched images are combined to form a mask image. Finally the masked pixels in the contrast-enhanced image are set to a different gray-value, such that a `MIP` only shows the vessel structures.

[Jayakrishnan et al., 2003] propose a similar procedure like that from [Venema et al., 2001]. Yet, their setup is simpler than the Venema setup. Brain `CTA` is performed by simply subtracting native and contrast-enhanced studies and there is no software registration of the studies, since the head of the patient is held fixed physically. The authors argue that due to this fixation a registration post-processing is not necessary. Another difference is that here the images are subtracted, leading to a slight increase in noise levels in contrast to masking identified bone structures in the contrast enhanced image which leaves the noise level intact.

The paper of [Yeung et al., 1994] presented one of the first approaches for CT Angiography using 3D registration. Their proposed 3D registration algorithm consists of several steps. After isotropic resampling of the 2D slices, the resulting 3D volume is put into a feature identification and selection procedure (Moravec corner detector). This step produces a number of feature points in the images. Now, in a 3D image flow computation step, for each feature point a local correlation window is set up to find correspondences in the second image. The search window is searched for the voxel that minimizes the sum of squared differences. This feature point is regarded as matching point and the distance to this point is the flow vector. Then the algorithm tries to estimate the global motion parameters from the derived flow vectors. An iterative algorithm calculates a least squares solution for rotation and translation taking outliers into account by randomization. Final steps are applying the transformation and resampling the image to match the second one. No more discussion is given concerning local nonlinear movements that might occur between scans.

In [Luboldt et al., 1997] an "elastic" subtraction algorithm for the display of vessel structures in `CTA` is described. The paper does not go into detail concerning this algorithm, only a very general description is given. The authors state that their algorithm bases on a matching step that analyzes native and contrast-enhanced `CT` scan for common structures. Matching volumes and areas are given higher weights in the algorithm than matching edges and points. After identification of possible matching structures a rigid registration step is performed followed by a nonlinear ”elastic” registration step of sub-structures. The nonlinear step is performed iteratively, however no further information on this step is given. Due to the short computation times of 5 minutes (1997) stated in the paper it may be supposed that this elastic registration step is quite simple. However the need for manual post-processing is also mentioned, which takes approximately 25 minutes per data set, yet no further information is given on the kind of post-processing that is performed.

[Kwon et al., 2004] propose a 3D registration scheme based on normalized mutual information to remove bone structures from `CTA` images. The 3D registration step is followed by a subtraction step and a final refinement step. The registration is performed by using the normalized mutual information measure on both images. A speedup is proposed by (obviously) using a multi-resolution technique. After rigid registration, there are still some unwanted residuals of the bones which are removed in a post-processing refinement step. Refinement is performed using several masks of bones and air. The paper is very interesting, since it shows a registration based solution for the problem, yet the fact that single bones behave rigidly while all bones together may behave different (and in a nonlinear fashion) is not exploited.

[van Straten et al., 2004] extend the work of [Venema et al., 2001] by presenting a bone removal method that uses an additional, non-enhanced scan to create a mask of the bone by thresholding and dilation. After registration of the `CTA` scan and the additional scan, the bone in the `CTA` scan is masked. As the cervical area contains bones that can move with respect to each other, these bones are separated first using a watershed algorithm, and then registered individually. A phantom study was performed to evaluate and quantify the trade-off between the removal of the bone and the preservation of the arteries contiguous to the bone. The influence of algorithm parameters and scan parameters was studied. The method was clinically evaluated with data sets of 35 patients. Best results were obtained with a threshold of 150 `HU` and a dilation of 8 in-plane voxels and two out-of-plane voxels. The mean width of the soft tissue layer, which is also masked, was approximately 1 mm. The mAs value of the non-enhanced scan could be reduced from 250 mAs to 65 mAs without a loss of quality. In 32 cases the bones were registered correctly and removed completely. In three cases the bone separation was not completely successful, and consequently the bone was not completely removed. The piecewise matched mask bone elimination method proved to be able to obtain `MIP` images of the cervical arteries free from overprojecting bone in a fully automatic way and with only a slight increase of radiation dose.

[Maksimov et al., 2004] present a registration based bone removal technique applied to `CTA` images of the lower extremities. Two problems are specifically addressed in this work, separation of vessels near bones and removing calcifications from vessels. Their method starts with native and contrast-enhanced images. A 200 `HU` threshold segments roughly the objects of interest. A classification of objects due to their size gives bone structure candidates. An assumption of slow intra-scan patient movement is made, therefore stacks of 10 slices are built and compared to each other. A matching step matches classified bone structures of stacks to each other for native and contrast-enhanced scan. The details of the matching step are a surface extraction and a iterative closest point determination. After matching all objects of the native image are transformed into the contrast-enhanced image and establish a bone mask. The mask is locally refined by at most 3 pixels to adapt to the underlying segmented structure. Reported results show that the algorithm is still in a work in progress state, since there are still problems with vessel detection and calcifications.

2.1.4.2 Nonlinear registration approaches with or without rigidity constraints

In the work of [Little et al., 1997] a registration algorithm is presented that combines rigid and nonlinear registration. Structures that are known to be rigid are treated as rigid for the registration while all other structures are treated as nonlinearly deformable. The nonlinear registration is performed by using manually selected landmarks and radial basis functions for interpolation (e.g. the thin-plate spline as a special instance of a radial basis function). The nonlinear registration model is weighted by using a distance map that specifies the distance to the rigid anatomical structures. For points of the rigid structures, the influence of the nonlinear registration is zero, while the rigid registration model is fully performed there. This scheme is called an inverse distance weighted interpolation and it needs a predetermined rigid registration transform.

Separation of rigidly and nonlinearly transformable parts is performed by a segmentation of the rigid structure of interest. The algorithm is applied to MR images where the vertebra is segmented and rigidly registered while the surrounding tissues are treated in an elastic scheme. Disadvantages of the method are the need for manual selection of corresponding landmark points, the need for manual segmentation of the rigid structures and the choice of the elastic model that interpolates the elastic registration between landmarks. The approach is very interesting due to the weighting of the nonlinear registration with a distance map defined by the rigid structures. However, no information is given on the used rigid registration method or any registration accuracy, be it rigid or nonlinear.

[Pitiot et al., 2003] address the problem of accurately registering two images when the images consist of independent components subject to rigid or affine transforms embedded in a deformable medium. The algorithm rigidly registers two images, computes a dense similarity map and a displacement field by a block matching technique, applies a hierarchical clustering to partition the correspondence field into a number of classes and extracts independent pairs of subimages. Pairs of subimages are then rigidly registered. Finally the non-linear interpolation scheme that is also used in Little et al. is applied for deformable registration. The difference to Little et al. is, that no a priori information like landmarks, segmentations or transformations are required, these are derived in the clustering approach.

The work of [Rohlfing et al., 2003, Rohlfing and Maurer Jr, 2001] is an extension of the nonlinear registration algorithm from [Rueckert et al., 1999]. Nonlinear deformations are modeled by B-splines on a grid of equally distant points overlaid on the images to register. Local deformations are calculated by using a `NMI` similarity metric and optimizing a cost function over the B-pline parameters derived from the `NMI` measure. A special regularization constraint is used to incorporate the notion of incompressibility into the framework. This constraint is derived from the Jacobian of the displacement field, which is a second derivative measure and should stay close to 1 for incompressible objects. This method could be used on rigid structures that are slightly deformed between two scans but are known not to change its volume. However, the accuracy of this method will be smaller than a partially rigid approach since the model will presumably be too deformable for our application.

[Martin-Fernandez et al., 2005] describes an elastic registration scheme that includes rigid anatomical structures. Rigid structures like knee, hands or spine are defined using wire connecting landmarks. A distance transform that starts at the wires is used to constrain the elastic registration. At the wire locations an exact rigid registration is guaranteed, while the further away one gets from the wires the larger the influence of the elastic scheme is. The application of the algorithm is registration of hand radiographs. Manually selected landmarks of the bone joints are used to build a wire model as a combination of single wires which are defined between joints. Rigid registration incorporating scaling, translation and rotation is performed on the wires. A distance transform that is iteratively derived from the wires is used to weight the influences of each wire for the interpolation of the deformation field over the whole image domain.

In [Staring et al., 2005] the problem of rigid structures that get deformed in a nonlinear registration framework is tackled by filtering a calculated displacement field after the nonlinear registration. A b-spline deformation algorithm is applied and the resulting displacement field is smoothed with a smoothing filter that depends on a segmentation of the rigid structures. This smoothing filter is derived in a similar fashion like anisotropic diffusion, the rigid parts are incorporated into the stiffness term of the anisotropic diffusion partial differential equation. However, the nonlinear registration leads to a certain registration error that is smoothed after registration, but might still be quite large.

2.2 Nonlinear Registration

In medical image registration there are many applications that require a nonlinear transformation to be modeled in the registration process. Examples for such applications occur in brain images due to brain-shift, or in soft tissue thoracic or abdominal images due to heart-beat or breathing. Also growing structures like tumors which are tracked over long time-periods involve nonlinear tissue deformations. In contrast to rigid registration techniques, nonlinear methods are still a "hot" research topic, there are no state-of-the-art algorithms yet, only algorithms solving specific problems. A big difficulty in nonlinear image registration is the question of evaluation, it is very hard to define a ground truth for nonlinear registration algorithms.

Nonlinear registration techniques look back on a period of 25 years of research, one of the first publications is due to [Broit, 1981] who "invented" elasticity theory matching based on physical considerations of tissue deformation for medical image registration. In the seminal paper of [Bajcsy and Kovacic, 1989], this technique was later refined to form a multi-resolution elastic matching approach, we will describe this technique later in more detail.

In the following we again rely on the basic formal definition of image registration from Section 1.4. More general considerations were derived from the books [Hajnal et al., 2001], [Yoo, 2004] and [Modersitzki, 2004]. We divide the topic of nonlinear registration into three sections. First we distinguish feature- and intensity-based methods, which are the two families of classical registration approaches. Often feature-based methods are more accurate than intensity-based methods as long as the feature extraction or segmentation steps are reliable and accurate. Due to the reduction of the problem space, feature-based methods are also significantly faster to compute. On the other hand, segmentation of the organs of interest is not always an easy task and inaccuracies in the segmentation or feature extraction process have severe effects on a subsequent registration step, making the intensity based methods perform better in many practical applications. Afterwards we give a more specific survey on nonlinear registration algorithms that were designed for the registration of thoracic images at different breathing states. These techniques are often combinations of feature- and intensity-based approaches to combine advantages from both directions.

2.2.1 Feature-Based Methods

Similar to rigid registration, feature-based methods in nonlinear registration are divided into methods based on *anatomical landmarks* and methods based on *extracted surfaces* to find correspondences in the data sets to be registered. The latter approaches need (often quite sophisticated) surface segmentation approaches as a pre-processing step. These segmentation steps may be fully automatic, however, in practice some sort of user interaction often is required, leading to semi-automatic segmentation schemes. Segmentation and exact delineation of surfaces is a large research topic by itself, therefore we refer the reader to books like [Sonka and Fitzpatrick, 2000] or [Bankman, 2000].

The basic pipeline of feature-based methods consists of three steps, the extraction of the structures of interest, landmark or surface matching leading to sparse correspondences and correspondence interpolation/approximation leading to a dense displacement field.

2.2.1.1 Matching & Registration of Anatomical Landmarks

In the computer vision literature, registration by matching of feature points is a very prominent method that is also used in related topics like optical flow or stereo problems [Brown, 1992]. Very often features like corner points [Förstner, 1986, Harris and Stephens, 1988] or scale-space maxima [Lowe, 2004] are used to automatically find correspondences. The extraction of feature points in the medical imaging domain was studied in [Rohr, 2001] who compared several anatomical landmark extraction methods in terms of localization error. However, the direct matching of anatomical landmarks has, up to our knowledge, not been reported in the medical imaging literature. Very often only few landmarks are selected and manually matched, with matching results being the input to e.g. Thin-Plate-Spline (`TPS`) transformation models [Bookstein, 1989]. [Chui et al., 2003] have shown a unified nonlinear feature registration approach using a joint clustering and matching framework, however, their approach is based on segmented surfaces where point features are extracted and matched. There is a growing number of publications that incorporate feature matches into intensity-based registration approaches [Johnson and Christensen, 2002, Fischer and Modersitzki, 2003a], but again, the matches are considered to be derived in a manual fashion.

A Global Solution to Sparse Correspondence Problems

A theoretically very profound method to solve the feature correspondence problem can be found in [Maciel and Costeira, 2003]. They claim to perform correspondence estimation and outlier rejection in a single formulation, achieving a globally optimal solution in an integer optimization problem. The combinatorial space of possible solutions is relaxed to a smaller search domain where an efficient constrained optimization algorithm can find a solution within feasible time.

2.2.1.2 Matching & Registration of Shapes and Surfaces

A basic review of surface registration techniques in medical imaging can be found in [Audette et al., 2000]. According to this survey paper, "surfaces provide more redundancy than landmarks, and this redundancy may be particularly advantageous for characterizing non-rigid motion". However, we have to add that the cost of this redundancy is the highly complex process of surface segmentation that is involved in this approach.

Early research in surface registration is based on physically-based elastic contour models (a.k.a. "Snakes") [Kass et al., 1988], which were adapted to 3D problems and termed "Active Balloon Models" [Cohen, 1991]. These methods use a partial differential equation that models the equilibrium of external image-driven deformation forces vs. internal elasticity forces of a deformable contour or a deformable surface. These ideas were used by a large number of authors for medical image registration [Terzopoulos et al., 1988, Terzopoulos and Metaxas, 1991, McInerney and Terzopoulos, 1995, Thompson and Toga, 1996]. A problem of the active balloon model is its inherent restriction to the same topology involved in the formulation. A more recent research direction uses the "Level-Set Method" [Sethian, 1999, Osher and Fedkiw, 2003] for registration, an example in the domain of medical imaging can be found in [Malladi et al., 1995]. In some sense also the approach of [Bajcsy and Kovacic, 1989] should be mentioned here, since its original formulation involved a 3D surface representation of a brain atlas that was matched to a voxel data set, however, we will present these kinds of methods in the following section.

Currently there is a trend in medical surface registration research to use state-of-the-art shape matching approaches from the standard computer vision literature. The reader may refer to [Veltkamp and Hagedoorn, 1999] for a review of shape matching algorithms. In the following we will present some representative shape matching algorithms, however, due to the large number of existing methods this selection is far from being comprehensive.

Modal Matching for Correspondence and Recognition

[Sclaroff and Pentland, 1995] are dealing the correspondence problem by spectral analysis of point-sets. They describe deformable objects with their eigenmodes, shape similarity is defined as the amount of deformation necessary to align objects. Eigenmodes ordered from global (low frequencies) to local (high frequencies). The eigenmode representation is calculated from a finite-element model of the shape, it can be seen as an orthogonal object-centered coordinate system describing the object. The modal object description is used for determining correspondences by comparing mode shape vectors with minimal deformation energy. The limitation of spectral analysis approaches is the susceptibility to partial or additional, noisy data like spurious nodes and edges.

Robust Point Matching

[Chui and Rangarajan, 2003] have formulated feature-based nonlinear registration as point matching problem where the two unknowns *point correspondence* and *transformation* are modeled in an iterative optimization framework and solved for simultaneously. Their algorithm, which shows similarities with expectation-maximization schemes, consists of the soft-assign method to relax the binary correspondence problem and a thin-plate spline model of the transformation. The optimization occurs in a deterministic annealing framework. Later the authors have also proposed a generalization of this algorithm using an iterative joint point clustering and matching strategy. Results applied to nonlinear brain registration can be found in [Chui et al., 2003].

Shape Matching Using Shape Contexts

The shape context approach [Belongie et al., 2002] is a widely-used method for matching 2D shapes (especially hand-written digits and letters) and 2D object recognition without relying on extracted features combining global and local shape knowledge. Objects are treated as (possibly infinite) point sets and it is assumed that the shape of an object is described by a subset of points. The method derives a distribution of relative distances and orientations and uses this distribution as a feature descriptor in a one-to-one matching step.

Nonlinear Shape Matching Using Relaxation Labeling

[Zheng and Doermann, 2006] have recently proposed an algorithm related to the robust point matching algorithm of [Chui and Rangarajan, 2003]. They note that although due to nonlinear deformations the absolute distance of points will vary, the neighborhood structures will remain intact. This idea is used in a graph matching approach where graph edges represent the neighborhood information. Since the graph matching problem is a discrete NP-hard optimization problem, the matching is relaxed to be able to use a continuous optimization method (relaxation labeling). The initialization of this method is performed using shape context matching, the comparison to shape context and robust point matching shows slight advantages of the method, which come at a larger computational cost.

2.2.1.3 Estimation of Dense Displacement Fields from Sparse Correspondences

Many landmark or surface matching approaches only give sparse correspondences. These correspondences may by interpreted as an approximation of a displacement field φ'. The task now is to interpolate/approximate a dense displacement field φ from φ' to be able to warp the moving image I_M to I_F, thus finishing the registration process.

Spline-based techniques are a widely used method to perform dense displacement field estimation after corresponding structures in the images have been identified. The corresponding structures are often given in the form of two sets of control points. At these control points, spline-based transformations either interpolate or approximate the displacements such that corresponding points are mapped into each other. Between control points, a smoothly varying displacement field estimate will be generated. A famous family of splines are radial basis functions, with the Thin-Plate-Splines (`TPS`) being a specific instantiation of radial basis functions. The use of `TPS` in (medical) image registration dates back to the popular paper of [Bookstein, 1989] who was the first to show the utilization of the `TPS` transform in the computer vision literature.

The original `TPS` interpolation method has been developed in the context of aero-elastic calculations by [Harder and Desmarais, 1972], being termed *surface splines* in their work. The mathematical foundations in terms of generic nonlinear function interpolation have later been laid by [Duchon, 1976] and [Wahba, 1990]. Later in the work of [Rohr, 2001] the exact interpolation constraint has been replaced by an approximation scheme which makes interpolation results more numerically stable. We describe the `TPS` interpolation and approximation scheme in more detail in Appendix B.1 since this method is crucial for several parts of this thesis.

[Davis et al., 1997] argued that in the context of medical applications the Elastic Body Spline (`EBS`) is a more accurate model for displacement field calculation. It solely differs from the `TPS` kernel by the kernel function $U(r)$, which is derived from Navier's partial differential equations that model tissue deformation as the equilibrium of an elastic body subjected to forces. A main drawback of this method is its huge increase in terms of computation time. Recently another method was presented in [Kohlrausch et al., 2005]. They call it the Gaussian Elastic Body Spline (`GEBS`), in contrast to the `EBS` where polynomial rational forces are used to derive the deformation behavior, in this work Gaussian forces model local and global deformations. This increases both, accuracy as well as computation time.

2.2.2 Intensity-Based Methods

Voxel- or intensity-based methods do not rely on extracted image information, but take the whole volume data sets as input for registration. The nonlinearity may be introduced by putting a nonlinear transformation into the registration framework depicted in Figure 2.1. A widely-used instantiation of this technique is the B-spline deformable registration algorithm developed by [Rueckert et al., 1999]. Again the nonlinear transformation model aligns the fixed and moving images until a certain similarity criterion like normalized cross correlation or normalized mutual information reaches a minimum. The optimization algorithm in this case has to be chosen much more carefully, due to the large number of unknowns that come up in this scheme. However, most intensity-based nonlinear registration methods like the Demons algorithm [Thirion, 1998], elastic matching [Bajcsy and Kovacic, 1989] or fluid deformable matching [Christensen et al., 1996] work in a non-parametric framework (compare [Modersitzki, 2004]). In the following we will present the B-spline deformable registration algorithm as well as several non-parametric methods based on the framework we defined in Section 1.4.

2.2.2.1 B-spline Deformable Registration

In [Rueckert et al., 1999] a new approach for nonlinear registration of contrast enhanced breast MRI is presented. The registration method is based on a global and a local motion model. The global model describes the overall rigid motion of the breast, whereas the local model specifies local deformations, modeled by B-spline free-form deformations. The authors show that their algorithm achieves superior results on breast MRI compared to simple rigid or affine registration methods. Their influential contribution is the setup of the B-spline framework for nonlinear registration, which was also used and extended in many different application areas later-on [Rohlfing et al., 2003, Mattes et al., 2003, Rueckert et al., 2003, Schnabel et al., 2003, Xie and Farin, 2004, Papademetris et al., 2004, Rueckert et al., 2006]. An evaluation of this method can be found in [Denton et al., 1999]. We present a more detailed derivation in Appendix B.2.

2.2.2.2 Elastic Registration

[Broit, 1981] was one of the first who investigated *elastic registration*, a method based on the *linearized elastic potential* which serves as a regularizer in the generic non-parametric registration framework. [Bajcsy and Kovacic, 1989] adapted this technique and used it for a multi-resolution elastic matching scheme, to match explicit 3D brain atlas models to locally variant data. The goal of their work was measurement of local shape variations and segmentation. The basic ideas of this matching approach are very similar to the early work on active contour models (”Snakes”) as described in [Kass et al., 1988, Terzopoulos et al., 1988]. Informally the multi-resolution matching scheme consists of a global, rigid alignment of model and data, followed by local model deformations similar to deforming a piece of rubber. Citing [Bajcsy and Kovacic, 1989], ”when external forces are applied the object is deformed until an equilibrium state between external and internal forces resisting the deformation is achieved; the equilibrium state for an isotropic homogeneous body is described by the partial differential equations of Navier and serve as the constraint equations in the elastic matching”. Material parameters are incorporated into this matching approach via the Lamé constants λ and μ, since all material bodies are to some form deformable (elastic, plastic, local, global, small, large). The coarse-to-fine scheme guarantees global coherence and a gradual increase of local similarity. Appendix B.3 gives the details of this algorithm, together with a possible numerical scheme to implement it. Further sample applications of the elastic registration scheme can be found in [Davatzikos, 1996] or [Thompson and Toga, 1996].

2.2.2.3 Fluid Registration

A restriction of the elastic registration model is, that it is limited to highly localized deformations. In fluid registration this drawback is removed, the elastic regularization model is replaced by a *visco-elastic* model. Fluid registration goes back to [Christensen et al., 1996], a more recent treatment on fluid registration can be found in [Wollny and Kruggel, 2002]. The main difference to the elastic regularization is, that the spatial smoothing of the displacement field is replaced by a spatial smoothing of the velocity (i.e. displacement *update*) field. As a consequence one can obtain arbitrary displacements given enough time, however, this flexibility may also be a disadvantage. The main application area of fluid registration is inter-patient atlas matching. Due to population-dependent shape variations, especially in the brain, it is important to provide more flexible matching methods. Follow-up publications include the consistent image registration framework [Christensen and Johnson, 2001], and the investigation of diffeomorphic mappings in the context of fluid registration [Joshi and Miller, 2000]. These works are some of the seminal papers that lead to the development of a sub-area of research in medical image registration, i.e. *Computational Anatomy* [Miller et al., 2002]. We will not go into further detail here, since the fluid registration approach is not very suitable for intra-modality registration applications.

2.2.2.4 Optical Flow Matching - Demons Algorithm

One of the currently most popular intensity-based nonlinear registration algorithm is the "Demons" algorithm [Thirion, 1998]. In contrast to the complex, physically-based elastic and fluid registration methods, the "Demons" algorithm is solely based on the notion of smoothing the occurring displacement fields. General optical-flow based registration, where "Demons" is a special instance of, is also called *Diffusion Registration* in [Fischer and Modersitzki, 1999]. Optical-flow registration takes a `SSD` data term in a variational setting and regularizes the resulting ill-posed problem using Tikhonov regularization [Tikhonov and Arsenin, 1977]. This means that the norm of the gradient of the displacement field is penalized in the regularization term, thereby preventing large differences in neighboring displacement field vectors. Chapter 6 includes more detailed information about the derivation of the optical-flow method using standard Tikhonov regularization and an anisotropic variant.

The "Demons" algorithm approximates the optical flow registration scheme by two important simplifications. First the minimization of the `SSD` data term and the regularization term is decoupled. Regularization is performed by using a Gaussian filter applied to the displacement field components separately in approximation of the Laplacians that occur as a result of the Euler-Lagrange equations of the Tikhonov-regularized optical flow equation. Second, the optical flow data term is heuristically modified by a normalization term that stabilizes the solution in the presence of large image gradients. See [Thirion, 1998] and [Avants et al., 2004] and Appendix B.4 for additional details on both the classical "Demons" algorithm and a variant that uses a symmetric definition of the registrations driving forces.

2.2.2.5 Curvature Registration

[Fischer and Modersitzki, 2003b] have recently proposed an efficient nonlinear registration algorithm which they termed *curvature registration*. It is again based on the `SSD` distance measure, however, the regularizer is based on penalizing second order derivatives of the displacement field (i.e. the term curvature registration). An important practical advantage of this method is that it already includes affine registration as a special case, a fact that distinguishes it from elastic, fluid and diffusion registration. [Fischer and Modersitzki, 2003b] have shown that for a certain finite difference approximation and a certain choice of boundary conditions a solution for the discretized partial differential equation (`PDE`) can be found very efficiently by using a fast Fourier transform (`FFT`) technique which leads to a computation time complexity $O(N \log N)$ with N being the number of unknowns of the system. See Appendix B.5 for a more detailed explanation of this method.

2.2.2.6 Biomechanical FEM-Based Registration

There is a large body of work in nonlinear registration that makes use of biomechanic models to derive energy functionals that drive the registration process. One example was presented in [Edwards et al., 1998] who have shown a three-component model to simulate the properties of rigid, elastic and fluid structures. In the brain these structures would correspond to skull, white and gray matter and cerebral-spinal fluid. The discretization of the variational energy functionals underlying these processes is performed in a Finite-Element Mesh (FEM) model, an analysis technique that solves `PDE` describing complex systems and spatially distributed processes. The mesh is generated by a segmentation of the three tissue types and a triangularization to get the FEM nodes. By deforming the mesh iteratively the energy functional is minimized. Note, that the need for segmentation makes this method restricted to brain applications where segmentation can more easily be achieved. Further the method has the drawbacks that one has to choose proper elasticity constants for the different tissue types and that the complexity to implement such a FEM scheme efficiently is rather high.

2.2.2.7 Hybrid Feature- and Intensity-Based Methods

Recently a number of authors started to investigate the fusion of feature- and intensity-based nonlinear registration methods. [Gee, 1996] proposed a Bayesian framework to unify these contrarious approaches. [Johnson and Christensen, 2002] combined landmark `TPS` registration with intensity registration using a viscous fluid regularization in a consistent manner. In [Li et al., 2002] this consistent intensity registration was combined with airway tree branch point matching for deformable lung registration. [Fan and Chen, 1999a] initialized an optical flow based intensity registration method with a semi-automatic airway tree and lung segmentation. [Hellier and Barillot, 2003] proposed a unified brain registration framework combining a global intensity approach with sparse landmark constraints. [Liu et al., 2004] showed a combined volumetric and surface matching approach for inter-subject brain registration. [Fischer and Modersitzki, 2003a] proposed to use landmark constraints as an auxiliary condition in a variational framework, leading to a formulation where the landmark constraints are the Lagrange multipliers of the optimization problem. [Papademetris et al., 2004] recently proposed a hybrid registration combining B-spline free-form deformation intensity registration with the Robust Point Matching algorithm.

2.2.3 Thoracic CT Image Registration

While the publications concerning rigid and nonlinear image registration are numerous in application areas like e.g. the brain, the literature of the recent years does not show the same extensiveness regarding thoracic registration applications, i.e. lung or airway tree registration and warping. Possible reasons may be the large size of thoracic volume scans and the large degree of deformation that might occur on volumes at different breathing states. Some examples for rigid thoracic registration applications can be found in [Cai et al., 1999] who rely on the chamfer matching method based on segmented lung surfaces or in [Betke et al., 2003] who show a two-step procedure with template landmark matching and iterative closest point registration of extracted lung surfaces. One example for a nonlinear registration algorithm that relates `PET` and `CT` images of the thorax can be found in [Mattes et al., 2003]. There the localized cubic B-spline deformation model from [Rueckert et al., 1999] is used in combination with the mutual information measure. A model-matching approach for segmentation of thoracic volume scans using fuzzy implicit surfaces can be found in [Lelieveldt et al., 2000].

An early approach for elastic registration of chest `CT` and whole-body `PET` can be found in [Tai et al., 1997]. After a prior coarse registration using a similarity transform model (translation, rotation, scale), a nonlinear registration method based on a sub-volume model is introduced. In each iteration, the target image is divided into small subsets of equal volume, these subsets are moved into all three directions until a local similarity measure is minimized. Afterwards a Gaussian-weighted regularization step interpolates the displacements off the subvolume centers. After warping according to the found displacement field, the next algorithm iteration starts. The whole process is continued until a pre-defined convergence criterion is fulfilled. The whole procedure is very similar to the registration technique in [Likar and Pernus, 2001].

An interesting approach for nonlinear lung registration was proposed by [Fan and Chen, 1999a]. They identify bronchial points of airway trees and vessel structures as feature points and manually register them. These sparse correspondences are used as a priori knowledge in a 3D warping model based on continuum mechanics theory that results in a dense displacement field. The notions of mass-conservation and incompressibility and their inclusion in a divergence-free elastic model leads to a variational formulation which is regularized by an anisotropic smoothing constraint, which preserves motion discontinuities. The resulting partial differential equation is solved by a conjugate gradient method. One major drawback of this method is its dependency on accurate a priori feature correspondences, this issue is dealt with in a follow-up paper [Fan and Chen, 1999b] where segmented lung surfaces are used as an additional a priori constraint. Nonlinear registration of lung surfaces is performed in an optical-flow framework, again using anisotropic smoothing regularization. In [Fan et al., 2001] this registration approach is evaluated on several data sets in terms of average feature registration error and lung volume comparison. Further, the approach is applied to a ventilation study.

[Zhang and Reinhardt, 2000] show a 3D surface based registration approach of lung CT images to a standard lung atlas. The purpose is lung segmentation by atlas registration. The two-step procedure involves a global transformation and a local elastic transformation based on subvolumes, i.e. cubes. For the local cubes, two similarity matches, one from atlas to volume and one from volume to atlas, are constructed. Resulting displacements are smoothed using a dynamic model combining the two directions. The procedure iterates until a convergence criterion is met. Evaluation of the approach using average distances of pairs of deformed images give accuracies in the order of 7 ± 3 mm to 2.7 ± 2 mm.

[Weruaga et al., 2003] present a volumetric motion estimation technique for thoracic images with the intention to include a model of breathing motion in radiotherapy planning for tumor localization and radiation-dose planning. Their approach consists of two phases, a similarity-parameterization data analysis stage and a projection-regularization stage. Most of the reported methods for thorax motion estimation require either patient training/collaboration or expensive medical equipment. Therefore the spatial correspondence of two volumetric `CT` data sets is acquired at the extremes of the breathing cycle. The proposed method defines a similarity criterion based on cross-correlation and the sum of squared differences. This similarity coefficient is used in a block matching technique where blocks are parameterized differently, according to the underlying image data. Finally the set of parametric models are translated into motion vector fields by a fast convolution-based regularization approach with a dedicated convolution kernel. The method shows promising results, however, it was only evaluated on a single data set and there are a large number of parameters to adjust in addition to a necessary lung border segmentation.

[Li et al., 2002] propose an inter-subject thoracic `CT` image warping and registration scheme for standardized atlas building. They use a priori information in the form of a set of reproducible feature point correspondences derived from airway trees and combine this information with an inverse consistent intensity-based registration scheme according to [Johnson and Christensen, 2002]. They report accurate results on six volumetric `CT` data sets in terms of relative volume overlapping error and average landmark registration error. However, the method is very time-consuming, due to the necessary manual matching of corresponding feature points and the highly complex nonlinear inverse consistent registration scheme. The work was later successfully used to establish a normative atlas of the human lung which might be used for image segmentation and detection of abnormal lung structures [Li et al., 2003].

[Dougherty et al., 2003] have shown the application of a sum of squared difference based optical flow method for nonlinear intra-subject thoracic `CT` registration. Their method uses a local optical flow formulation including intensity-bias correction (due to e.g. contrast agents) in a coarse-to-fine manner. No further smoothness or regularization assumptions on the motion field are made. Evaluations on five cases are only performed using the cross-correlation of registered data sets. In [Dougherty et al., 2006] more evaluation experiments were performed to assess the algorithm behavior in a lung nodule study, a perfusion study and an air trapping study.

In [Tschirren et al., 2005b] an interesting approach for the matching of segmented airway trees is described. They are using the airway tree segmentation method from [Tschirren et al., 2005a] to robustly extract the airways, afterwards skeletonise the airway tree and derive a graph representation with branching points used as graph nodes. A maximum-clique graph matching algorithm is used to define a correspondence between airway trees.

2.3 Conclusion

Our extensive literature review has revealed the potential contribution areas for this work. Bone subtraction `CTA` techniques currently do not handle several independent rigid movements in a satisfying way, hence a novel method is of large interest. Although there exists a large body of work in soft-tissue deformable registration, only few methods exist that make use of shape- and feature-based semi- and fully-automatic matching algorithms (especially for thorax images), a topic which has already been thoroughly researched in the traditional computer vision community. By transporting several promising computer vision approaches into the medical imaging domain, a contribution to the search for efficient nonlinear registration algorithms might be performed. These algorithms may be applied to increasingly large data sets from clinical practice, while at the same time keeping small vessel structures (i.e. high frequency information) intact. The fusion of feature- and intensity based approaches is another topic from literature that promises to be useful for practical algorithms, since feature based constraints will speed up intensity based approaches significantly. An interesting fact we learned from the evaluation sections of the various nonlinear registration algorithms is, that there are currently no common frameworks to simplify and objectify the hard question of nonlinear registration evaluation. Consequently, the decision to investigate such a framework in the context of thoracic `CT` images seems to be very promising.

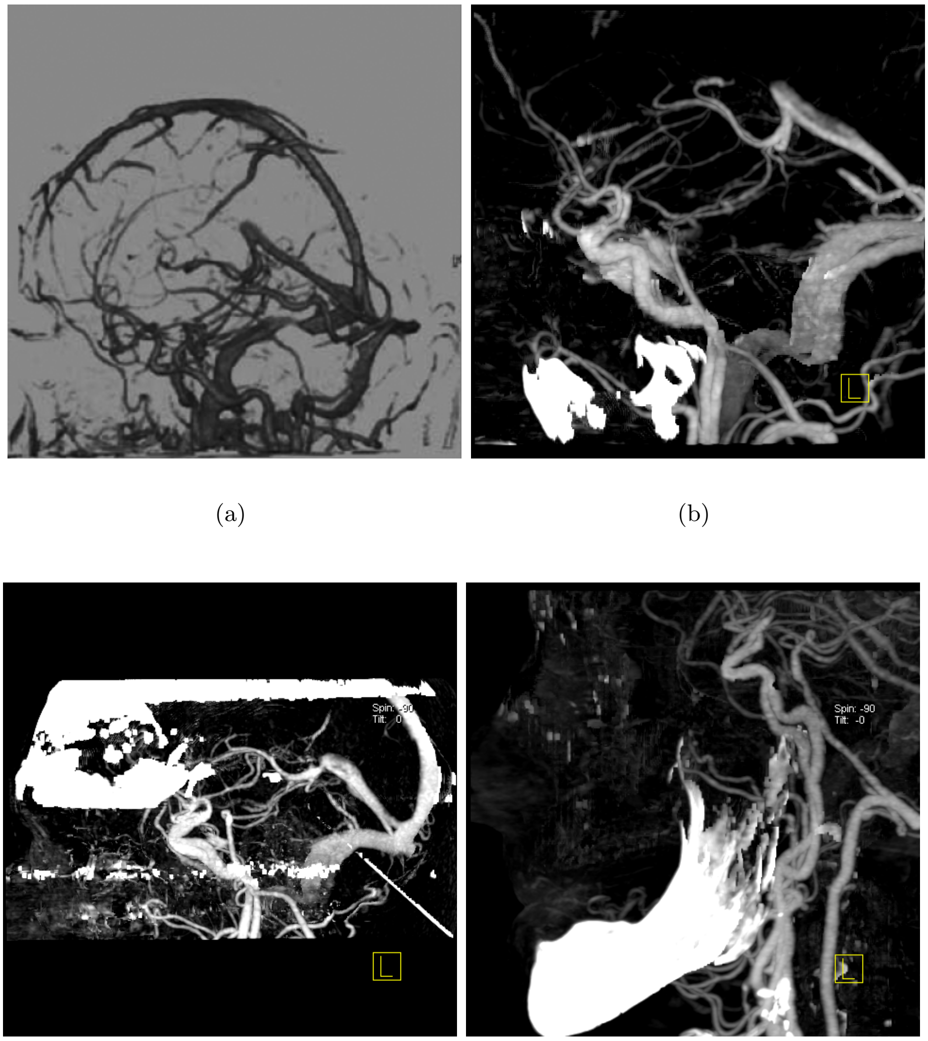
Chapter 3

Partially Rigid Registration for Bone Removal in CTA Applications

Contents

A Maximum Intensity Projection (`MIP`) study of CT Angiography (`CTA`) scans is a widely used imaging tool for artery and vein visualization especially in the brain. This method allows the detection of cerebral aneurysms, arterial stenosis, and other vascular brain anomalies which are important in the emergency evaluation of stroke and intracranial hemorrhage [Lell et al., 2006a, Tomandl et al., 2006] (see also Figure 1.3).

With the increase in temporal and spatial resolution, `CTA` has emerged as an alternative to Digital Subtraction Angiography (`DSA`). `CTA` studies are X-ray Computed Tomography (`CT`) scans where a contrast agent is applied via intravenous injection. Often the contrast-enhanced scan is accompanied by a native scan to be able to support vessel visualization by subtraction techniques. This is a large advantage compared to classical `DSA` where intra-arterial injection of contrast material has to be accomplished, being more invasive to a patient. More specifically, if one ignores intra- and inter-scan patient movement, `CTA` is performed by a subtraction of the native from the contrast-enhanced scan, leaving solely the vascular structures. The resulting data set is visualized in 3D using a `MIP`, where the maximum value in the `CT` volume data set is displayed along each ray through a pixel in the direction of the viewpoint projection (see Figure 3.1a for a `MIP` example).



(a) (b)

(c) (d)

Figure 3.1: Maximum intensity projections of several CTA data sets with and without bone artifacts obstructing vascular structures. (a) Example for MIP visualization showing vascular structures, (b) MIP of CTA data set with slight errors due to independently moving jawbone, (c) MIP of CTA data set with intra-scan artifacts in the skull region, (d) MIP of CTA data set with severe errors due to independently moving jawbone.

In this chapter we describe a novel algorithm for bone subtraction `CTA` under the influence of independent rigid movements occurring during or between scans. We have already given an overview of related work on bone removal techniques and partially rigid bone registration in Section 2.1.4. Section 3.1 gives a more detailed motivation of the problem. The partially rigid registration algorithm is described in Section 3.2. Section 3.3 introduces setup and results of the synthetic and real data experiments, while finally Sections 3.4 and 3.5 discuss the obtained results and conclude the chapter.

3.1 Bone-Subtraction CT Angiography

A drawback of `CTA` images is the similar `CT` intensity value of bone structures and calcifications compared to the contrast-enhanced vascular structures. Therefore, one acquires native and contrast-enhanced `CTA` images to remove bones and calcifications from the head and neck `CTA` images by subtraction techniques. In practice, patient movement during and between native and contrast-enhanced scans often is inevitable leading to insufficient image quality due to motion artifacts. Therefore, a subtraction algorithm has to deal with misregistered areas. Due to their high `CT` intensity especially misregistered bone structures lead to considerable visualization artifacts in `MIP` studies (see Figure 3.1b)-d) for examples). These artifacts have to be removed to get an undistorted view of the otherwise obstructed vascular structures. In many cases this removal is possible by a rigid registration of native to contrast-enhanced scan followed by a subtraction of a bone mask generated from the registered native scan. This technique is also referred to as "Matched Mask Bone Elimination" [Venema et al., 2001] (see also Section 2.1.4).

However, sometimes several independent patient movements occur during contrast-enhanced and native scans, thus complicating the subtraction and introducing the need for specialized registration methods. Although each of the independent movements can be regarded as being rigid, the combination of these inter- and intra-scan motion artifacts cannot be effectively removed by a single rigid registration step. In the remainder we will refer to the problem of independently moving bone structures as a "partially rigid registration" problem. An example of independent movements is a slight head rotation combined with a different position of the jawbone due to swallowing or yawning between scans. Further, one can easily imagine that the problem of independently moving bone structures becomes even more important as soon as not only the head is involved in the `CTA` study but the vascular structures of interest are extended into the neck and shoulder region [Lell et al., 2006a]. Independent shoulder movements due to e.g. an uncomfortable resting position of a patient additionally pose registration problems.

To address these problems we have developed a refinement of the "Matched Mask Bone Elimination" technique that incorporates a joint segmentation and registration method in an iterative fashion. Moreover our novel approach takes the large size of current routinely acquired CT scans into account. With volume data sets that consist of several hundred slices and an x-y resolution of 512×512 voxels, memory and run-time issues are a challenge in the development of medical imaging algorithms.

3.2 Partially Rigid Registration Algorithm

As already mentioned, this algorithm has to deal with several problems which may be summarized as follows:

- Accurate registration of native and contrast-enhanced scans taking independently moving bone structures like e.g. inter-scan skull and jawbone movement into account.
- Accurate registration despite local intra-scan errors like rapid skull movements during a single acquisition.
- Memory- and runtime-efficiency due to large volume data sets of several hundred slices.

We decided to follow a similar direction like [van Straten et al., 2004]. However, our approach replaces the error-prone watershed segmentation step by a more robust approach that uses joint segmentation and registration in an iterative fashion. Therefore, no high-level or semi-automatic segmentation step is necessary, the algorithm is based on the combination of low-level segmentations and rigid registrations. Our method guarantees that nonlinear deformations never occur at the bone structures but are solely used for tissue structures by a registration matrix interpolation step. The only important assumption that we have to make is that it is possible to rigidly register and remove independently moving bone structures in sequential steps, i.e. each registration step in the iterative loop has to remove a certain area of misregistration. If this assumption does not hold anymore the algorithm will terminate too soon. This situation might happen if two independent misregistration areas cancel each other out in terms of the registration metric.

The basic idea of the proposed algorithm is to iteratively perform rigid registration on areas where large misregistrations occur. The algorithm takes a native and a contrast-enhanced volume as input and starts with an initial mutual information based rigid registration that is restricted to bone structures segmented by a bone threshold. Calculating the misregistration error identifies areas where the registration has to be refined. This refinement is performed in an iterative manner as long as the number of misregistration errors is too large. Each iteration consists of calculating the misregistration area, a rigid registration step restricted to the current misregistered area and an interpolation step that combines the different registration results. Algorithm 1 shows this algorithm in pseudocode, while the following subsections explain its behaviour in more detail.

Algorithm 1 Partially Rigid Registration

```
Mutual information based rigid registration of native and contrast-enhanced scan
Initialize data structure of resulting registration matrices with initial transform
Calculate the misregistration area
while size of misregistration area larger threshold do
    Derive a bone mask from the misregistration error area
    Mutual information based rigid registration restricted to bone mask
    Update registration matrix data structure by checking if the new transform improves
    the error
    Smooth and interpolate registration matrices
    Update misregistration area
end while
```

3.2.1 Single Rigid Registration Step

Our technique consists of several rigid registration steps always using the same mutual information based matching method as its basic building block. Using Mutual Information (`MI`) as a similarity measure for image registration was already explained in Section 2.1.2.3. We decided to use the Normalized Mutual Information (`NMI`) metric due to its independence to volume overlap problems. According to the basic registration scheme (see Figure 2.1) a registration algorithm also consists of an optimization strategy, an interpolation method and a transformation representation that provides the parameters to be optimized. Our optimizer is a regular-step gradient descent optimizer that takes larger steps at the beginning of optimization and consecutively reduces its step size until a local minimum is reached. The interpolation method is tri-linear, a more complex method is impractical due to performance reasons. Finally our transformation is composed of six parameters, three representing 3D translation and the other three representing 3D rotation encoded as a unit quaternion. The unit quaternion or versor representation has the advantage that it natively models the 3 degrees of freedom of a rotation without the problem of gimbal lock. See [Ng and Ibanez, 2004] for a treatment of the unit quaternion representation in medical image registration. The regular-step gradient descent optimizer herein has to take into account the special structure of the parameter space which is not a vector space.

3.2.2 Partially Rigid Registration - Initial Stage

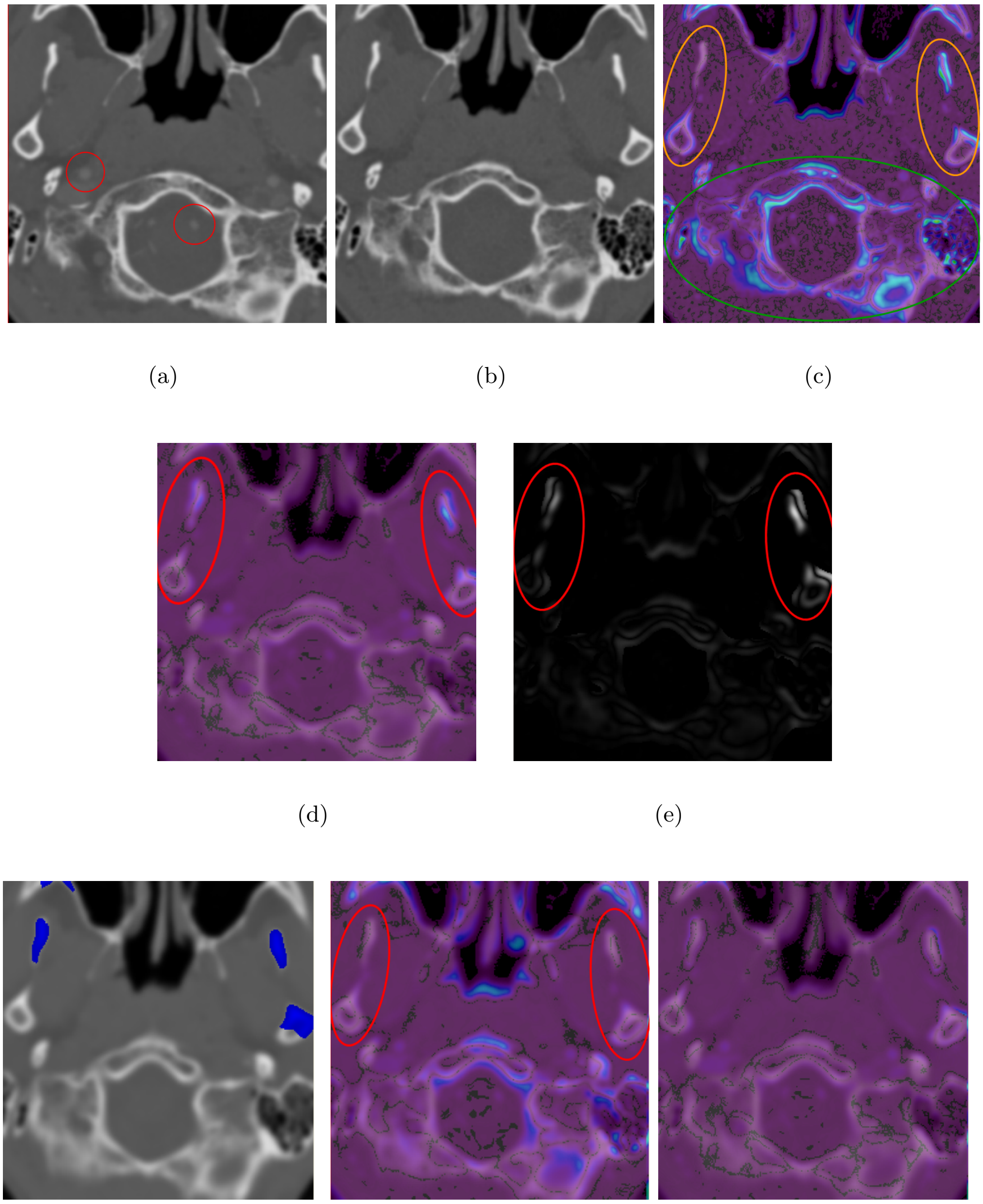
The partially rigid registration algorithm starts with a threshold based bone segmentation of the contrast-enhanced CTA image. This bone segmentation is used as a mask restricting the following initial mutual information based registration procedure to bone structures. The result is a transformation which is stored as the initial transformation in the result data structure at bone voxel locations only. With this initial transformation it is possible to calculate a misregistration area by subtracting the accordingly warped native image from the contrast-enhanced image. Now the iterative stage is entered.

3.2.3 Partially Rigid Registration - Iterative Stage

The iterative stage terminates if the misregistration area is smaller than a certain threshold. The first step in this loop is the creation of another bone segmentation mask to restrict the subsequent mutual information registration. The bone segmentation mask is derived from the misregistration area, which is located at the edges of misregistered bone structures, by a dilation into the bones nearby using a distance-constrained region growing method. This dilation step can be seen as a bone segmentation procedure, however this bone segmentation is only used to focus the following registration step on misregistered areas. The restricted registration step results in another transformation which overwrites the result data structure after checking if the current transformation is able to reduce the misregistration at each bone voxel location. Now the transformation results have to be smoothed in local neighborhoods to remove some noise and to prepare the following transformation interpolation. Since the rigid transformations are only stored at bone voxel locations, it is necessary to find transformation parameters for tissue voxels. This can be performed by a linear or a nearest-neighbor interpolation. Although a linear interpolation would be more accurate, we decided to use nearest-neighbor interpolation due to reasons of reduced computation time and its low memory consumption. The evaluation section will show that the accuracy of nearest-neighbor interpolation is sufficient for our application. The final step in the iterative stage is the warping of the native to the contrast-enhanced image according to the transformation result data structure. The result can be used to calculate another misregistration which gets checked if it is larger than a threshold by the loop termination condition. After the loop has terminated the Matched Mask Bone Elimination (`MMBE`) method is used to remove bone structures from the contrast-enhanced image. The whole algorithm work-flow is also depicted in Figure 3.2. Note that the color-coding represents intensity difference magnitudes where "cold colors" (magenta, blue, green) are assigned to small differences and with growing differences the colors get "warmer" (i.e. yellow, orange, red).

3.2.4 Memory and Runtime Efficiency Issues

As already mentioned above the large size of current routinely acquired volume data sets always poses restrictions on practically useful algorithms due to runtime and memory consumption issues. CT data sets the proposed algorithm is intended for easily have several hundred slices with x-y resolutions of 512 by 512 voxels respectively, thereby requiring around 250 MB in memory due to a 12 bit gray level resolution. Runtime efficiency re-



(a) (b) (c)

(d) (e)

(f) (g) (h)

Figure 3.2: Partially rigid registration work-flow. (a) contrast-enhanced image, (b) native image, (c) difference image showing two independent sources of movement (jawbone - orange, skull - green), (d) misregistration area after initial registration, (e) misregistration mask derived from (d), (f) distance-constrained region growing of misregistration mask, (g) result of another registration step focusing only on misregistration mask, (h) result of the nearest-neighbor fusion of the individual registration steps.

DS	Size	Problem Characteristics
A	512,512,231	independent head & jaw movement
B	512,512,344	independent head & jaw movement
C	512,512,429	head movement & teeth artifacts
D	512,512,233	intra- and inter-scan movements
E	512,512,269	swallowing and teeth artifacts

Table 3.1: Partially rigid registration. Evaluation data set characteristics.

quires the data sets to be fully held in memory, therefore it is important to reduce the need for intermediate data structures. The partially rigid registration algorithm only requires one additional volume data set of the same size as the two input images to store intermediate results. This can be achieved by representing bone segmentation results and registration error regions as single bits and by using indices into a list of possible transformation parameters to store the different registration results of the iterative algorithm. Since memory and time consumption always imposes some trade-off, we will show in the evaluation section that our runtime results are nevertheless acceptable.

3.3 Evaluation Experiments & Results

The presented approach was evaluated on several `CT` data sets showing problems of state-of-the-art bone removal techniques for CTA images based on maximum intensity projections. All of these data sets still have problems after one single registration step for Matched Mask Bone Elimination (compare first row of Figure 3.4). More specifically we used 5 clinical data sets whose characteristics are shown in Table 3.1. The images stem from a 64-slice Multi-slice Spiral CT (`MSCT`). The acquisition protocol is a `CTA` protocol that involves a native image and an image with non-ionic iodinated contrast medium injected. The scan delay is adapted using a bolus tracker. Data sets show an in-plane resolution of 0.5 mm and a z-resolution of up to 0.6 mm.

Most of the data sets show several independent movements typical for `CTA` acquisitions. Data sets C and E also show some artifacts in the tooth regions due to implanted gold teeth disturbing the CT scans. In our experiments we calculate two measures from the data sets. The first one is the progression of the number of misregistration voxels during our algorithm. The number of misregistration voxels is calculated as the number of voxels which is larger than a threshold from the difference between contrast-enhanced and (partially) rigidly warped native image. Table 3.2 depicts this measure, note that data sets A and E finished earlier due to additional termination conditions in the main loop. The first column specifies this measure before registration. To restrict runtime the loop was terminated after four iterations.

DS	0	1	2	3	4
A	306 705	37 078	26		
B	208 846	94 601	991	926	801
C	53 443	9 007	6 529	6 514	5 755
D	164 939	136 603	19 165	14 347	14 189
E	26 868	9 036	5 387		

Table 3.2: Partially rigid registration. Decrease of misregistration error voxels for evaluation data sets A-E.

The second measure is the progression of the sum of squared intensity differences (SSD) between the contrast-enhanced image I_{CE} and the warped native image $I_{N'}$ according to the partially rigid registration transformations. The sum of squared intensity difference is calculated as

$$SSD = \frac{1}{|\Omega|} \sum_{\Omega} (I_{CE}(\mathbf{x}) - I_{N'}(\mathbf{x}))^2$$

where Ω is the domain of the overlapping part of the images. One should note that the Sum of Squared Differences (`SSD`) will never decrease to zero, since there are always contrast differences in the images due to contrast agent injection. Figure 3.3 shows the decrease of the `SSD` for the five data sets.

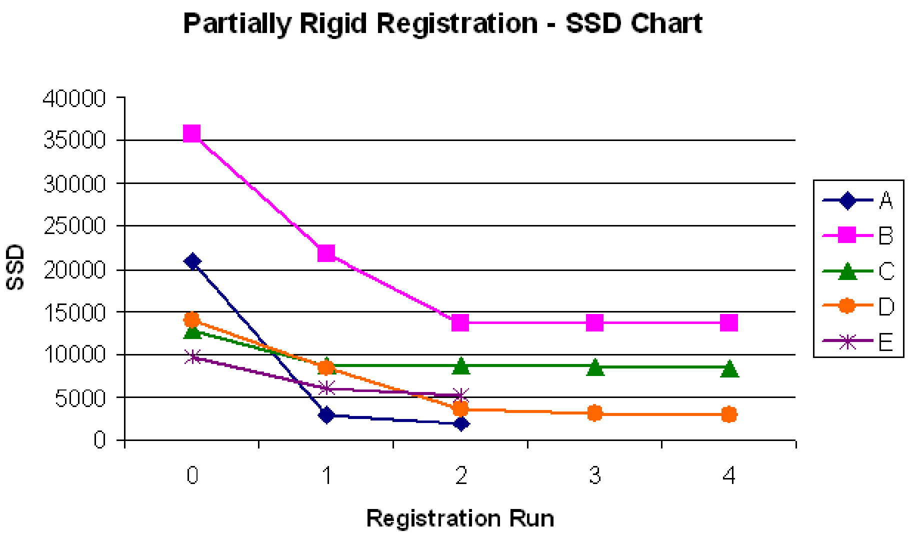


Figure 3.3: Partially rigid registration. Decrease of sum of squared difference measure for evaluation data sets A-E

Finally we give a table showing the execution times for the algorithm on each of the data sets A,B,C,D and E. The algorithm implementation was performed under Windows in C++ and evaluation experiments were executed on a Pentium M with 2.0 GHz and 1.5

GB RAM.

	A	B	C	D	E
time [s]	93,25	233,95	276,02	205,67	93,27

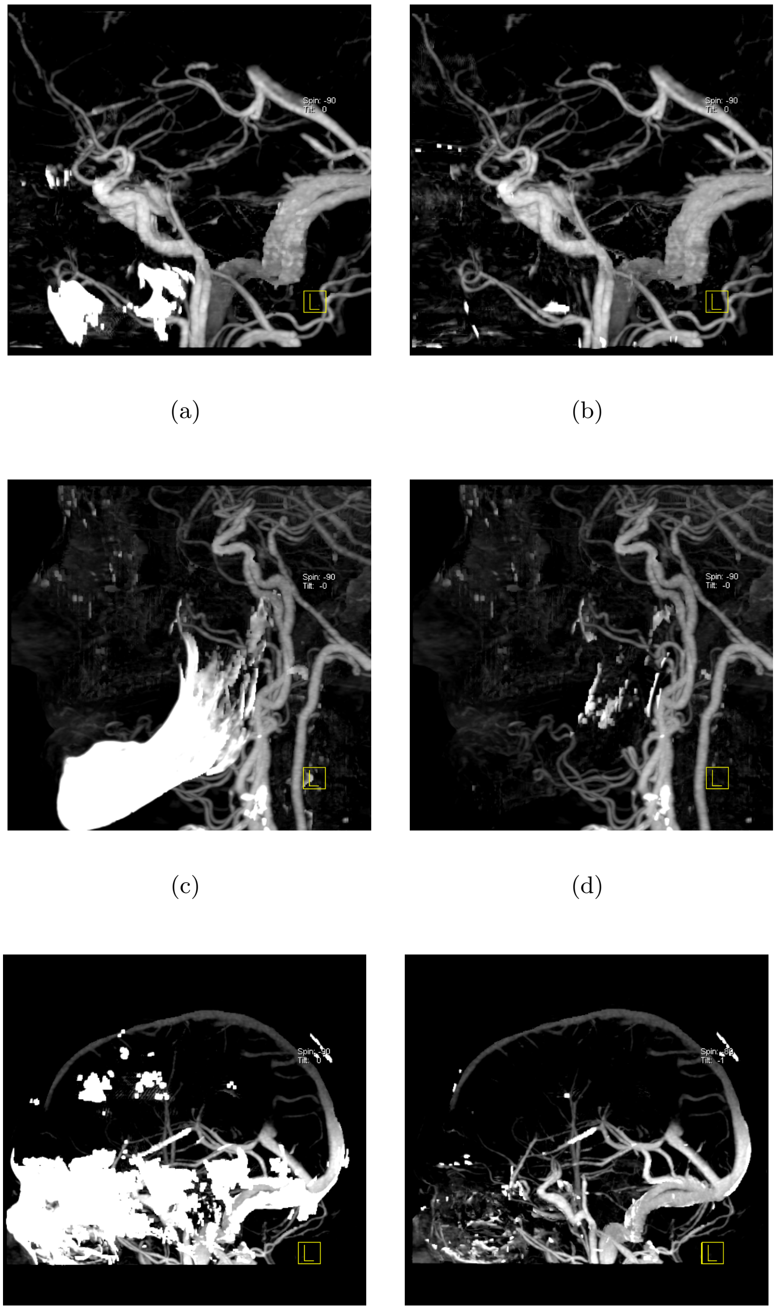
In Figure 3.4 the three data sets A,B and D are shown to give qualitative results as well. The first row is the `MIP` of each data set after a `MMBE` with a single rigid registration to take patient movement into account. One can clearly see the bright bone structures that obstruct several portions of the vascular structures. The second row shows the resulting `MIP` after a `MMBE` using the novel partially rigid registration procedure.

3.4 Discussion

Our experiments show very clearly that the proposed algorithm is capable to improve the `MMBE` method in those cases where several independent rigid movements occur during two scans in `CTA` studies. In all clinical test cases the sum of squared intensity differences and the number of registration error voxels is significantly reduced after one or two additional registration steps in our iterative algorithm. We observe that both measures are going into a converged state after a few iterations, therefore we decided to restrict the additional registration steps to at most three. Data sets A and B show the algorithms excellent behavior in the presence of independent head and jawbone movement. In data set A the obstructing structures have been completely removed, while data set B has improved a lot, although some very small regions still remain, where the whole algorithm has converged to a local minimum. Data sets C and E show the algorithms behavior in the presence of artifacts due to CT scan errors from gold teeth or due to a patient swallowing during scans. Both effects do not have a great impact on the result, all obstructions are successfully removed. Finally data set D has inter- and intra-slice scanning errors which are also removed, however in this case a larger number of obstructing bone structures remains. An important property of the algorithm is that it never worsens a result if the registration is already accurate enough after a single registration step. So it is very suitable as an additional refinement step if the classical MMBE method does not succeed. The fact that the algorithm runtime lies between two and five minutes on a standard notebook computer underlines that the additional computational effort is acceptable.

3.5 Conclusion

In a number of medical applications the removal of bone structures is of crucial importance for a high-quality `CTA` visualization using `MIP`. This chapter has presented a novel



(a) (b)

(c) (d)

(e) (f)

Figure 3.4: Partially rigid registration. Selected results comparing the classical MMBE method (first row) with the improved MMBE method (second row) (a) data set A after original MMBE, (b) data set A after partially rigid registration MMBE, (c) data set B after original MMBE, (d) data set B after partially rigid registration MMBE, (e) data set D after original MMBE, (f) data set D after partially rigid registration MMBE.

algorithm for bone removal addressing the problems of independent inter- and intra scan movements. It extends the classical `MMBE` algorithm by a joint segmentation and registration stage. The presented experimental results on clinical data sets show examples of intra- and inter-scan patient movements which were successfully registered with the proposed algorithm. The improved `MIP` visualization quality underlines the usefulness of our novel method, while the quantitative evaluation of the number of registration voxels and the progression of the sum of squared intensity differences proves the algorithms correct behavior.

The presented algorithm is in the process of clinical evaluation, an important topic for future work is to perform more evaluation experiments on these data-sets to show the clinical relevance of the refinement algorithm. A first result on a more thorough comparison of our partially rigid registration algorithm to the standard methods has recently been published in [Lell et al., 2006b]. Their promising results using our method show superior quality of `MIP` visualizations in clinical practice. Important topics to investigate further are the improvement of the algorithm accuracy by replacing the nearest-neighbor interpolation with a linear interpolation. A theoretical proof of convergence of the iterative algorithm would also be of great interest. Another possible direction for future work is to look into more sophisticated algorithms for joint segmentation and registration like techniques working in a variational framework [Yezzi et al., 2003] or compare it to the elastic registration scheme in [Bentoutou et al., 2002]. A further interesting approach that we recently found is [Arsigny et al., 2005] who propose a polyrigid/polyaffine transformation model as an intermediate representation between simple rigid and highly nonlinear transformations. Their transformation seems to be very suitable for our purpose and should be investigated more deeply.

Chapter 4

Surface-Based Nonlinear Registration

Contents

Studying the complex thorax breathing motion for fusion of functional and anatomical data and derivation of statistical breathing motion models has already been motivated in earlier sections of this thesis. In this chapter we investigate segmented thoracic `CT` structures like lung, airway tree and diaphragm surfaces at several different breathing states between inhalation (Total Lung Capacity (`TLC`)) and exhalation (Functional Residual Capacity (`FRC`)). We are working on thoracic sheep CT data sets provided by Prof. Eric Hoffman from the University of Iowa. Since the sheep data sets were acquired with high radiation doses, the signal-to-noise ratio of the images is very high, thus the segmentation of the structures of interest is possible without the need for highly complex, robust segmentation algorithms. Due to the simple surface extraction, the decision to investigate surface-based nonlinear registration methods has been made. We concentrate on efficient nonlinear registration based on a shape descriptor [Veltkamp and Hagedoorn, 1999], more specifically finding correspondences via shape matching and registering identified correspondences in the context of a Thin-Plate-Spline (`TPS`) framework [Bookstein, 1989]. In many cases shape-based methods are able to provide very accurate solutions with less computational effort since only the organ of interest is considered as opposed to intensity-based methods where the whole images are taken for matching and registration.

Generally, feature-based registration starts with a feature extraction step. After analyzing lung and diaphragm surfaces over the breathing cycle one will notice that the shapes undergo certain deformations which make it hard to robustly identify correspondences. Especially the deformations of the surface of the elastic lung tissue are hard to describe. The lung surface is attached to the rib cage by adhesive forces and the diaphragm-induced cranio-caudal forces move the lung tissue while at the same time the bones of the rib cage remain comparatively rigid. As a consequence the rib cage forces its shape on the moving elastic lung tissue and it is not necessarily the case that classical 3D features like ridges (or valleys) on the lung surface stay ridges (or valleys) during breathing. For this reason a shape matching algorithm has to be developed that is independent of points of interest like curvature maxima or similar 3D features. After reviewing the body of literature the shape context approach [Belongie et al., 2002] was identified as a reasonable and promising approach for this purpose. In the following this algorithm was extended to 3D and adapted for the task at hand. The idea of extending the shape context descriptor to 3D has been independently proposed by [Frome et al., 2004] for recognizing 3D objects in range scans. Given the correspondences from the shape matching, an estimate of the breathing motion in form of a displacement field can be calculated in a nonlinear registration step. [Belongie et al., 2002] propose the `TPS` framework [Bookstein, 1989] for this task.

The remainder of this chapter is organized as follows. Section 4.1 gives further motivation for studying shape-based registration. In Section 4.2 the basic techniques and their extensions which were used to identify shape correspondences and nonlinearly register the image data are presented. Section 4.3 introduces setup and results of the synthetic and real data experiments, while Sections 4.4 and 4.5 discuss the obtained results and conclude the chapter.

4.1 Understanding the Physiology of Breathing

Investigation of the physiology of breathing in order to model respiratory motion is an important research topic [Hoffman et al., 2004, Simon et al., 2005]. Respiration depends on the dynamic interaction of lung, rib-cage and diaphragm, topics which have already been extensively studied in human and animal studies in the physiology community [Wilson, 1988, Boriek et al., 1998, Angelillo et al., 2000]. Attempts for simulating and modeling respiration can be found in [Kaye et al., 1998], who combine static anatomical `CT` information with physics-based cardio-respiratory physiological models, or in [Segars et al., 2001], who have modeled respiratory motion in a spline-based mathematical torso phantom to reduce respiratory motion artifacts by using findings from respiratory mechanics. In the long run, investigation of breathing motion in thoracic and abdominal `CT` images should lead to statistical models of respiration, optionally using bio-mechanical constraints, which might be used for breathing motion compensation. Respiratory motion correction in coronary MR images based on 3D registration has already been studied by [Manke et al., 2002]. Since registration of pre- and intra-operative thoracic and abdominal images is nowadays being used more often by image-guided intervention and surgery systems, the accurate registration of thoracic and abdominal soft-tissue regions was also used for liver intervention systems taking breathing motion into account [Rohlfing et al., 2001, Blackall et al., 2004]. We consider the contributions made in this chapter as another component on the way to understand the physiology of breathing with the help of computer vision techniques.

4.2 Surface Registration using Shape Context

In this work the shape context approach is used to find correspondences on lung and diaphragm surfaces followed by a nonlinear `TPS` registration. Compared to other methods for registration of thoracic CT data (see Section 2.2.3), this is a more general approach, since basically all kinds of extracted surfaces (e.g. segmented airway or vessel tree surfaces as well as skeletonized trees) can be used in the same matching framework. Similar methods for shape matching have already been discussed in Section 2.2.1.

4.2.1 Shape Context Matching & Registration

The shape context approach [Belongie et al., 2002] was reported as a reasonable and promising method for matching 2D shapes (especially hand-written digits and letters) and 2D object recognition without relying on extracted features. It combines global (by regarding all points of a shape) and local (by storing information about the relation of all possible point pairs) shape knowledge in a clever way. Objects are treated as (possibly infinite) point sets and it is assumed that the shape of an object is captured by a finite subset of its points, giving a set $P = \{p_1, ..., p_n\}$. The points can be obtained as locations of edges from an edge detector or from another method to sample contour/surface points from a shape. The points need not and typically will not correspond to key points or structures such as maxima of curvature, inflection points or surface ridges. If one looks at the set of vectors emitted from one point p_k to all other points p_i of a shape with $i \neq k$, this set can be interpreted as a rich description of the shape configuration relative to p_i. Since this description is much too detailed, the *relative distribution* of this set of vectors is taken as a compact, yet highly discriminative descriptor instead. Therefore, for each point p_i a histogram h_i of the relative positions of the remaining points is calculated which is called the *shape context*. This histogram uses bins that are uniform in a three-dimensional spherical coordinate system (θ,ϕ,r). The r coordinate axis is logarithmically sampled, such that positions of nearby sample points have stronger influence on the descriptor than ones located farther away. When setting up the logarithmic bin edges for the radius, it is important to scale these bin edges accordingly to the data set. In order to achieve scale invariance for the descriptor, it is possible to calculate the median point-to-point distance of the shape and take this as a measure representing the scale of the shape. This scale may be used as the bin edge that discriminates the largest and the second-largest radius bin. To find the bin edge that distinguishes between smallest and second-smallest radius bin, a robust minimum point-to-point distance may be found by taking the median of the smallest one-percent of the point-to-point distances. Inbetween these two radius bin edges, the other logarithmically scaled bins are placed. Figure 4.1a shows the bin structure of 3D shape context histograms.

Now for each point p_i on the first shape, the "best" matching point q_j on the second shape has to be located. For a point tuple $< p_i, q_j >$, let

$$C_{ij} = C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^{K} \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)}$$

denote the cost of matching these two points. This cost function is simply based on the χ^2 test statistic which is a natural choice for comparing histograms. Given the set of costs C_{ij} between *all* pairs of points on first and second shape, one wants to relate each point from the first shape with *exactly* one point from the second shape under the constraint that the total cost of this mapping is minimized. This problem can be formulated in a graph framework by taking the sample points from both shapes as graph nodes in a bipartite graph, i.e. there are only graph edges *between* the two sets of sample points, no edges between sample points of a single set. Edges in the graph are weighted with the cost function C_{ij}. This graph setup is illustrated in Figure 4.1b. Finding a cost-minimizing mapping between point sets now transforms to a graph matching problem which is an instance of the weighted bipartite assignment problem. It can be solved in polynomial time, e.g. with the *Hungarian Algorithm* in $O(N^3)$ time, with N being the number of nodes in the bipartite graph. The final result of the graph matching step is a one-to-one mapping of corresponding points from the two shapes.

[Belongie et al., 2002] give some arguments about invariance and robustness of the shape context approach. It is implicitly invariant to translation, since all measurements are taken with respect to object points. Scale invariance can easily be achieved by normalizing the radial distances by the mean distance. In this work we replace the mean distance by the statistically more robust median distance. If it is desirable for an application, rotation

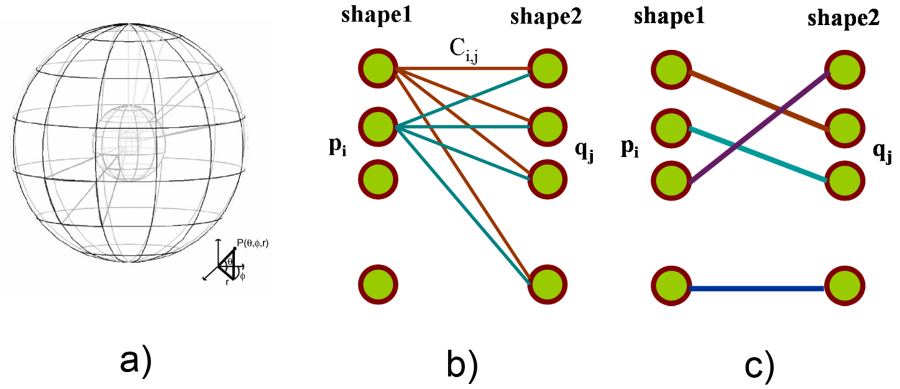


Figure 4.1: Components of the 3D shape context descriptor calculation. a) shows the histogram bin structure embedded in a spherical coordinate system (θ,ϕ,r). b) depicts the bipartite graph structure used for the graph matching algorithm, with nodes p_i and q_j representing all sample points from first and second shape, respectively, and edges weighted with $C_{i,j}$ representing the cost function to match a pair of points. c) is the resulting graph after bipartite matching.

invariance can also be achieved by incorporating local coordinate systems based on the tangent vector at each point instead of using a common global coordinate system. Rotation invariance is not an issue for this work, since data sets are either already in a similar position and orientation, or will be brought into similar orientation by a pre-processing step using rigid registration. Outlier handling can be introduced by using ”dummy” nodes attached to each point set, whose cost is chosen in a way that outliers have larger cost values.

After establishing the point correspondences [Belongie et al., 2002] make use of the `TPS` framework [Bookstein, 1989] for nonlinear registration. The `TPS` approach leads to a transformation that consists of an affine part and a nonlinear deformation part depending on the identified correspondences. The parameters of the `TPS` model are calculated from the constraints that the displacement at corresponding points is zero (exact interpolation of the displacement field) and that the spline model between corresponding points yields regular and smooth displacements tending to zero the further away from the corresponding points they are.

4.2.2 A Shape Matching & Registration Pipeline

A shape matching & registration pipeline was implemented to be able to derive a displacement field given two image data sets at two distinct states of the breathing cycle. A coarse overview of this pipeline is given in Figure 4.2. The ITK toolkit [ITK, 2006] was used as a very helpful basic software library for the implementation of this pipeline.

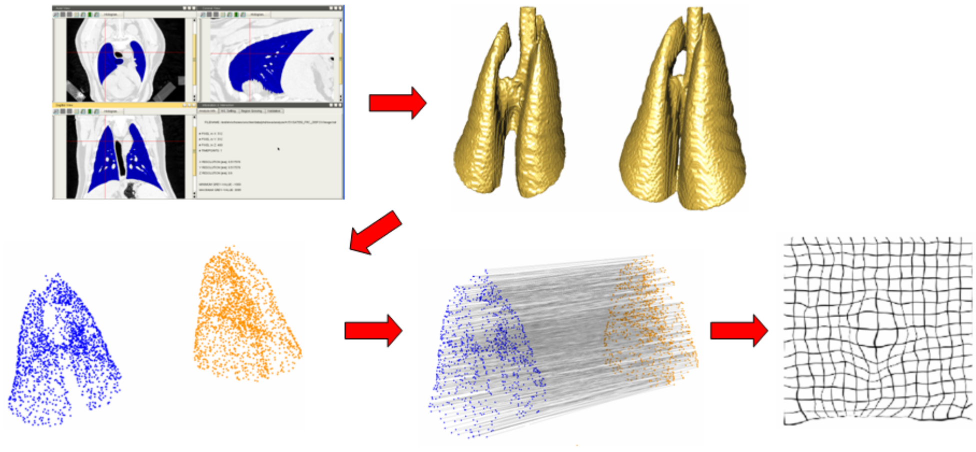


Figure 4.2: Shape Registration Pipeline. From left to right, volume segmentation, surface extraction, surface point sampling, shape matching to get point correspondences and nonlinear registration and volume warping based on the point correspondences are depicted.

4.2.2.1 Segmentation

The first step in the registration pipeline is an automatic segmentation of the lung, airway tree and diaphragm surfaces. A simple algorithm that incorporates a region grower for background removal, a gray value thresholding step and a connected component labeling gives a rough segmentation of the airway tree and the lung tissue. For accurate lung segmentation a region-growing algorithm with an increasing threshold interval [Boehm et al., 2000] is used. The seed point is located in the airway tree (trachea) of the topmost image. The lung segmentation result is morphologically closed to remove small vessel structures. Figure 4.3 shows an example for segmented lung surfaces at exhalation and inhalation respectively, visualized by using a polygonal mesh representation created with the Simplex Mesh algorithm [Delingette, 1999]. An airway tree segmentation algorithm that was developed in our group [Pock et al., 2005a] is used to get an accurate airway tree with the rough segmentation as a constraint. The lung segmentation is separated into left and right lobe by a labeling step based on the main branching point of the airway tree, using the same method that was presented in Beichel et al. [Beichel et al., 2004] for liver partitioning. This results in a binary volume for left and right lung lobe. From these binary volumes a diaphragm point set is extracted by regularily sampling points from the bottom part of the lung surface. It should be noted that this simple diaphragm segmentation scheme is working well for sheep CT data but can not so easily be used for human diaphragm segmentation, since the human heart sits on top of the diaphragm, while in sheep data the heart is surrounded by lung tissue.

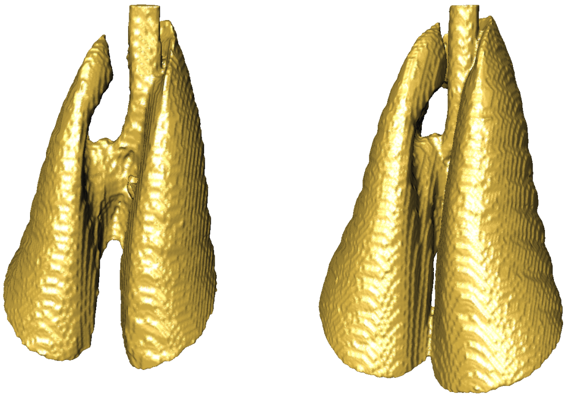


Figure 4.3: An example for segmented sheep lung surfaces. The left image shows the lung at exhalation (FRC), the right one at inhalation (TLC).

4.2.2.2 Polygonization & Sampling Point Extraction

The next step is the extraction of a point set from the segmented lung surfaces. In contrast to the original shape context paper, discretization is an important issue in this work since 3D discretizations easily produce very large numbers of points. Further, some methods might produce more points in areas of high curvature which is a disadvantage for calculating the shape context descriptor. In this case the descriptor would lose its ability to describe the small lung deformations due to breathing. To avoid these difficulties the binary volumes are triangulated using the Marching Cubes algorithm [Lorensen and Cline, 1987]. This gives a dense and regular triangle mesh, which is sampled in a *regular* way. Regular sampling is important for the shape context descriptor to form a rich description in a global and local manner and to be independent of small deformations due to lung movement at once. In all later experiments the number of sampling points per lung surface was varied between 200 and 3400 points.

4.2.2.3 Shape Context Matching

The calculation of the shape context descriptor was extended to incorporate local gray value information. At each extracted surface point a second term of the cost function based on the normalized cross correlation is calculated. The extended cost function C'_{ij} has the following form

$$C'_{ij} = C'(p_i, q_j) = \alpha C(p_i, q_j) + \beta NCC(p_i, q_j)$$

with $C(p_i, q_j)$ being the original shape context descriptor cost function and $NCC(p_i, q_j)$ being the normalized cross correlation evaluated at $5 \times 5 \times 5$ neighborhoods of the original gray value `CT` image centered around p_i and q_j respectively. This additional term adds robustness against segmentation errors and makes use of the otherwise neglected intensity information in the data sets.

Since the three-dimensional extension of the shape context approach requires a larger number of sample points than the simpler two-dimensional case, it is crucial to utilize an efficient algorithm for the bipartite graph matching problem. In this work the minimum weight assignment algorithm from the *LEDA*$^{\mathrm{TM}}$software library* was used, which solves the graph matching in $O(N(M + NlogN))$ time, with N being the number of nodes and M the number of edges in the bipartite graph. This algorithm proved itself as reasonably fast, since all practical experiments showed that the calculation of the cost functions at the N^2 points took more time than the graph matching step. After finding the one-to-one correspondences a certain percentage of matched points were removed. More specifically matched edges get sorted by their weights and the ones with largest weights get excluded. This adds some robustness against errors introduced in the preceding segmentation and point sampling steps. The subsequent registration step does not need the full number of correspondences since it interpolates the displacements between missing correspondences according to the chosen interpolation model. However, the percentage of correspondences to remove has to be chosen small enough to prevent entire loss of information in certain image regions.

4.2.2.4 Thin-Plate-Spline Interpolation

The final step in the registration pipeline is the nonlinear registration of the matched points in a `TPS` framework. A `TPS` interpolation function $f(x, y, z)$ has the form

$$f(x,y,z) = a_1 + a_x x + a_y y + a_z z + \sum_{i=1}^{n} w_i U(\parallel (x_i, y_i, z_i) - (x, y, z) \parallel)$$

with kernel $U(r) = r$. The reader may refer to Appendix B.1 for further details on the `TPS` method. Davis et al. [Davis et al., 1997] have reported that the elastic body spline (EBS), which solely differs from the `TPS` in the kernel function $U(r)$, is a more accurate model for medical imaging applications. They state that this kernel is actually better suited for modeling tissue deformations, since it is derived from Navier's partial differential equations that model the equilibrium of an elastic body subjected to forces. It incorporates a parameter that models the tissue elasticity derived from Poisson's ratio. In this implementation both types of spline kernels were used. In our modular framework the spline kernel is easy to modify. Another aspect in the registration framework is the large number of sample points that have to be extracted for a reasonable shape approximation. The

---

[*] http://www.algorithmic-solutions.com

TPS framework resembles an interpolation of the displacement field between corresponding points, therefore a large number of correspondences tend to introduce over-fitting to still fulfill the interpolation requirements. Besides over-fitting, it also allows non-diffeomorphic mappings, i.e. foldings in the displacement field. The group around Belongie published a method [Donato and Belongie, 2002] to deal with this problem. However, in this work the findings of [Rohr et al., 2001] to approximate thin-plate spline mappings were considered, since they are theoretically more profoundly justified. [Rohr et al., 2001] proposed to add a regularization term to the formulation, which is steered by a parameter λ weighting the trade-off between interpolation and smoothness of the solution. λ ranges between 0 for exact interpolation to 0.1 for an approximated purely affine transformation with hardly any local deformations.

4.3 Evaluation Experiments & Results

To assess the validity of the shape context matching and registration approach qualitative and quantitative evaluations were performed on synthetically transformed and real thorax data sets. Three different kinds of shapes were produced in the segmentation steps and taken as input for the evaluations, i.e. diaphragm, lung lobe and airway tree surfaces. Figure 4.4 shows examples for each kind of shape.

The image data that was used for these evaluations comes from MSCT sheep scans, provided by Prof. Eric Hoffman, University of Iowa. The data was acquired with 64-slice scanners at five different breathing states between inhalation (TLC) and exhalation (FRC) by a protocol where breath is held at fixed inspiration levels during the 30 sec scan time. This leads to a static breathing scheme, which has to be considered for the interpretation of derived motion models from matched and registered shapes. The image dimensions per breathing state are approximately $512 \times 512 \times 550$ with voxel dimensions of $0.52 \times 0.52 \times 0.6mm$.

4.3.1 Matching Experiments on Synthetically Transformed Data

The basic procedure for the synthetic matching experiments is to provide a data set $A(\mathbf{x})$ and a transformed version of the data set $B(\mathbf{x})$ as input to the shape matching algorithm. The synthetic transformation

$$\varphi : \{B(\mathbf{x}) = \varphi(A(\mathbf{x})), 1 \leq x \leq n_x, 1 \leq y \leq n_y, 1 \leq z \leq n_z\}$$

is defined by the evaluator and unknown to the matching algorithm, therefore B serves as a ”gold standard” data set for the shape matching algorithm. The matching algorithm

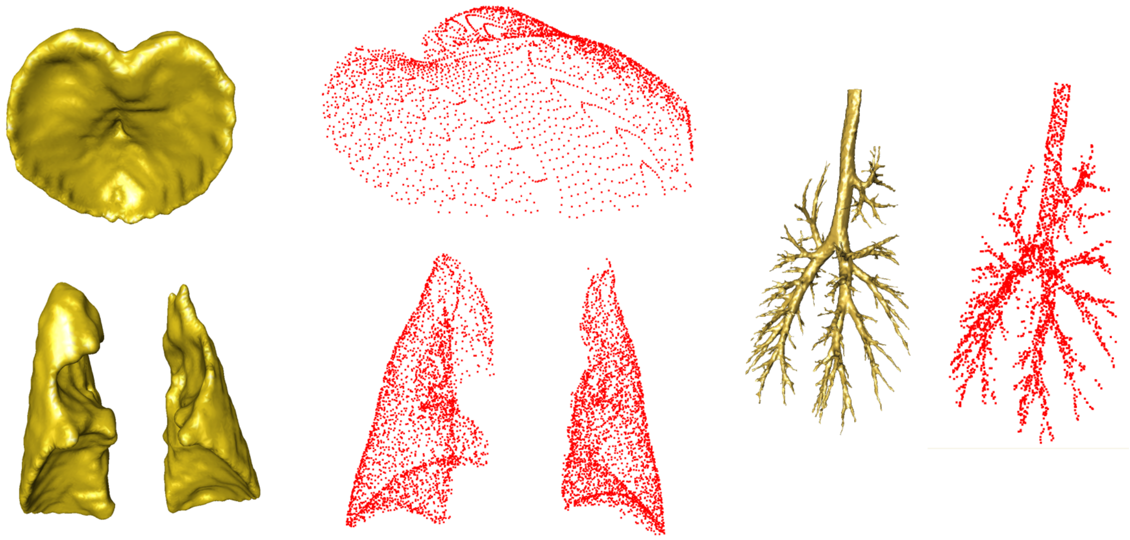


Figure 4.4: The different kinds of data used in the experiments. Segmented surfaces to the left and sampled points to the right, respectively. Right column shows airway tree data, while left column shows diaphragm at the top and lung lobes at the bottom.

computes from its inputs $A(\mathbf{x})$ and $B(\mathbf{x})$ a list of point correspondences $< p_i, q'_i >$, mapping a point set $p_i : \{p_i(\mathbf{x}) \in A(\mathbf{x}), 1 \leq i \leq n\}$ to a point set $q'_i : \{q'_i(\mathbf{x}) \in B(\mathbf{x}), 1 \leq i \leq n\}$ with n being the number of surface sample points. By applying the transformation φ on the point set p_i a set of synthetically transformed points $q_i : \{q_i(\mathbf{x}) = \varphi(p_i(\mathbf{x})), 1 \leq i \leq n\}$ is calculated. If the shape matching algorithm would be perfect q_i and q'_i would be identical for all values of i. Since in practice this is not the case, for each tuple of point sets q'_i and q_i minimum, mean and maximum distances are calculated and interpreted as an accuracy measure for the shape matching algorithm.

Some of the parameters involved in the matching and registration algorithm remain fixed during the experiments. The weighting parameters of shape context and normalized cross correlation cost function are set to $\alpha = 0.7$ and $\beta = 0.3$ to give more importance to the shape context cost function term. The percentage of point correspondences removed from the shape matching results is set to 20% and it was ensured that there are no regions lacking sample points after removal by visual inspection. The elastic body spline kernel is not used for the final evaluations. Compared to the `TPS` kernel we found that its need to tune the parameter resembling Poisson's ratio and its considerably higher computation time were not worth the only slightly better results. The regularization factor for the `TPS` interpolation is set to 0.01.

Two different kinds of transformations φ were applied to create synthetic data sets.

First, a simple rigid scaling transformation was applied to test the algorithms correct behavior. Second, a more elaborate synthetic nonlinear transformation was applied to derive error statistics on the matching performance.

4.3.1.1 Synthetic Scaling Transformations

In the simplest case, a rigid scaling transformation $q_i = T(p_i)$ with varying scale factors between 0.9 to 0.1 in steps of 0.05 was applied to the extracted points of an airway tree surface. The airway tree surface was sampled with 1000 and 2000 points, respectively. The outcome of this test case were percentages of correctly identified correspondences of 100% between scale factors of 0.9 and 0.6 which slowly start to decrease between scale factors. Figure 4.5 shows a plot of the percentages of correctly identified correspondences. Although the implementation of the shape context approach is not scale-invariant, the performance is very well as one might expect. The descriptors give very similar histograms if the points are simply scaled. Nevertheless if the scale factor becomes very small the points in the scaled point set tend to concentrate on few histogram bins in contrast to the points of the unscaled version. This could be prevented by making the algorithm scale-invariant, which was no issue for this work. This problem explains the performance drop for scale factors below 0.4. The data set with 2000 sample points reaches the state with histogram binning problems earlier, since there are more points involved.

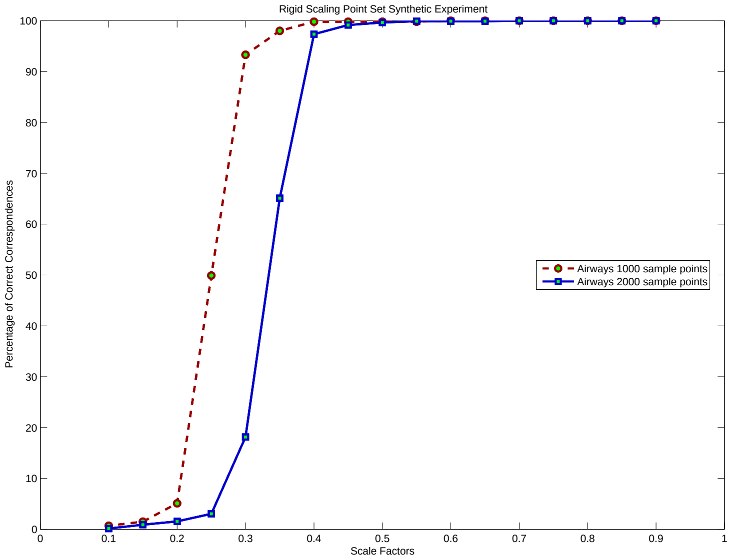


Figure 4.5: Percentages of correctly identified correspondences for scaled airway tree point sets.

The second synthetic test case took airway tree and lung lobe surfaces as its input and applied the rigid scaling transformations $q_i = T(p_i)$ with varying scale factors between 0.9 and 0.2 in steps of 0.1 again. Instead of applying the transformation on the already extracted surface points, this time it was applied on the segmented image data, in accordance to the real-world matching problem the algorithm is designed for. After scaling, 2000 point samples were picked *independently* from both data sets and presented to the shape context matching algorithm. As a consequence inaccuracies due to the sampling method have an effect on the matching outcome. Corresponding points were compared with the "gold standard" correspondences as described above. The results for airway tree and lung lobe surface are shown in Table 4.1. Similar to the results mentioned above a large scale factor induces larger distance errors, since the shape matching is scale dependent. However, the mean error distances for all surfaces remains in the range of a few mm, which is a good prerequisite for the subsequent registration task.

4.3.1.2 Synthetic Simulated Breathing Transformations

The shape context approach is used to find shape correspondences of thorax structures *over breathing*. Therefore a synthetic transformation has to simulate breathing behavior. A nonlinear transformation φ was designed to approximate a breathing-like deformation by simulating diaphragm and rib cage movement. According to the literature [Davies et al., 1994] diaphragm movement for deep breathing ranges between 25 to 40 mm in the vertical direction during a breathing cycle. The nonlinear deformation is calculated using two parameters $t_{vertical}$ and t_{inward}. The details of this synthetic transformation are given in Appendix B.6. An example for this transformation using $t_{vertical}$ of 25 mm and t_{inward} of 10 mm applied to a thorax data set is shown in Figure 4.6.

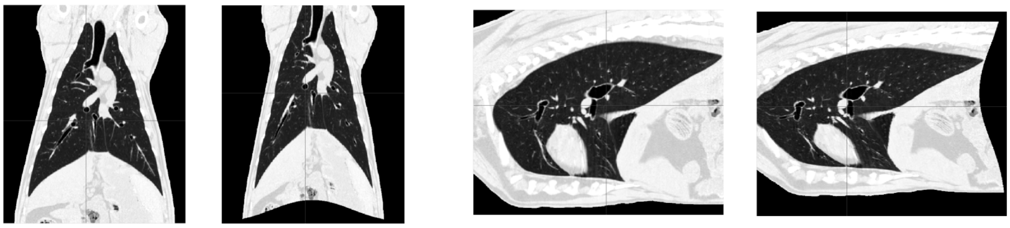


Figure 4.6: Sagittal and coronal views of a data set demonstrating the synthetic transformation that is used to simulate breathing behavior. Left column shows original, right column transformed data, respectively.

The first synthetic experiment operates on a segmented diaphragm. Diaphragm surface points are extracted from a segmented lung surface data set at TLC. The vertical translation $t_{vertical}$ and the inward translation t_{inward} remain fixed to 25 and 10 mm,

	Measure / Scale	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**
Airways	Min Distance [mm]	0.1223	0.1799	0.1441	0.1465	0.1310	0.1583	0.2133	0.3633
	Max Distance [mm]	17.1885	21.2219	19.9105	21.4395	22.4912	24.3155	27.6104	36.6797
	RMS Distance [mm]	3.7816	3.5720	3.6521	3.8789	3.9120	3.8108	5.8954	9.97706
	Mean Distance [mm]	4.5664	4.3826	4.5332	5.08678	5.1034	4.9803	7.6435	11.9974
	Std Dev [mm]	2.5606	2.5400	2.6865	3.2920	3.2785	3.2076	4.8666	6.6653
Lobe 1	Min Distance [mm]	0.2429	0.2671	0.2882	0.2929	0.2046	0.3526	0.4835	0.6547
	Max Distance [mm]	21.3413	23.8625	16.9326	24.0671	25.2402	37.3329	38.9801	48.4554
	RMS Distance [mm]	6.3142	6.9028	6.1249	5.6401	6.1177	7.3768	10.6431	14.1169
	Mean Distance [mm]	5.4949	5.9907	5.3347	4.8481	4.9199	5.3970	9.4445	11.5053
	Std Dev [mm]	3.1114	3.4304	3.0102	2.8832	3.6372	4.9087	5.0307	8.1832
Lobe 2	Min Distance [mm]	0.4233	0.2671	0.2416	0.2401	0.2929	0.3553	0.4972	0.7544
	Max Distance [mm]	17.4655	18.9396	22.9419	25.5128	28.9582	27.5386	38.8222	44.6533
	RMS Distance [mm]	5.9971	6.2529	5.9331	4.8263	5.9217	6.0102	14.6823	15.3978
	Mean Distance [mm]	5.2480	5.4102	5.1536	4.1630	4.8736	4.4864	11.9091	13.3015
	Std Dev [mm]	2.9033	3.1361	2.9408	2.4427	3.3649	4.0008	8.5904	7.7591

Table 4.1: Evaluation of shape context matching on rigidly scaled airway tree and lung lobe surfaces.

[%] / Samples	**400**	**800**	**1200**	**1600**	**2000**	**2400**	**2800**	**3200**	**3600**
Correspondences	100	100	100	100	99.6	99.1	97.3	94.5	89.2

Table 4.2: Percentage of correctly identified correspondences from synthetically transformed diaphragm point sets.

[mm]/Samples	**400**	**800**	**1200**	**1600**	**2000**	**2400**	**2800**	**3200**	**3600**
Min	0.309	0.270	0.216	0.128	0.100	0.135	0.063	0.103	0.058
Mean	8.991	5.867	6.027	5.500	5.202	4.983	4.761	4.817	4.441
Max	25.92	24.38	22.45	22.64	23.32	21.22	18.15	21.28	18.47

Table 4.3: Distance error measures for synthetically transformed diaphragm volumes over different sampling sizes.

respectively. The number of sample points used to represent the diaphragm surface are varied between 400 and 3600 sample points in steps of 400. To assess the validity of the shape matching method based on the shape context descriptor (without dependencies on segmentation and shape sampling errors) sampled point sets are transformed according to the simulated breathing transformation. In this case an exact correspondence of points is given initially and an ideal shape matching method would identify 100% correct correspondences. The percentage of correctly identified correspondences over different numbers of sampling points is given in Table 4.2. The expected 100% correspondence was achieved in most cases. The higher number of sampling points show a performance decrease, since a large number of shape context histogram bins would be necessary to prevent discretization artifacts.

To assess the dependencies on sampling inaccuracies the segmented volumes are then transformed instead of the sampled point sets. Each transformed volume is sampled independently. First of all the number of sampled points is varied while keeping the translation parameters at the same value as in the previous experiment. This demonstrates the matching algorithms behavior with increasing sampling rate, which clearly shows an increase using more sample points with a saturation at very high numbers. Table 4.3 shows these results.

From now on the number of diaphragm sample points is fixed to 2000 sample points. Following the previous results this is a good compromise between running time and matching quality, especially since the quality does not increase drastically for higher number of sample points. t_{inward} is set to 10 mm and $t_{vertical}$ is varied between 0 and 45 mm. Table 4.4 depicts these results, the same experiment with $t_{vertical}$ set to 25 mm and t_{inward} being varied show a similar result. Performance decreases when deformations get too large. Another algorithm run was performed with $t_{vertical}$ set to 25 mm and t_{inward} being varied between 0 and 25 mm. Table 4.5 states these results.

[mm] / t	0	5	10	15	20	25	30	35	40	45
Min	0.121	0.0500	0.135	0.034	0.153	0.184	0.143	0.104	0.097	0.262
Mean	3.406	3.453	3.549	3.895	4.506	4.903	5.367	6.029	6.506	7.404
Max	12.13	11.66	12.21	14.34	17.21	17.57	19.10	21.53	26.46	27.81

Table 4.4: Distance error measures for synthetically transformed diaphragm volumes over different transformations.

[mm] / t	0	5	10	15	20	25
Min	0.001	0.136	0.184	0.069	0.085	0.196
Mean	4.333	4.743	4.903	4.967	5.491	6.002
Max	17.35	19.28	17.57	21.53	24.16	25.23

Table 4.5: Distance error measures for synthetically transformed diaphragm volumes over different transformations.

	[mm]/Samples	200	600	1000	1400	1800	2200	2600	3000	3400
Airway	Min	1.705	0.371	0.355	0.132	0.169	0.164	0.192	0.131	0.069
	Mean	11.74	8.30	7.35	5.98	5.802	5.707	5.517	5.040	4.578
	Max	32.73	28.54	31.10	24.96	25.39	26.30	27.50	27.16	25.68
Lobe	Min	1.944	0.580	0.244	0.371	0.227	0.146	0.337	0.244	0.083
	Mean	15.84	9.48	7.93	7.14	6.70	6.44	5.72	5.98	5.99
	Max	38.95	26.57	28.36	24.08	19.16	21.74	22.18	27.26	26.15

Table 4.6: Distance error measures for synthetically transformed airway & lung volumes over different sampling sizes.

Another set of experiments is performed on airway tree and lung lobe surfaces. Airway tree and lung lobe segmentations from a data set at `TLC` are used as input to the shape matching algorithm. Again, the vertical translation $t_{vertical}$ and the inward translation t_{inward} remain fixed to 25 and 10 mm, respectively. The number of sample points is varied between 200 and 3400 sample points in steps of 400. Table 4.6 summarizes the achieved results which are similar to the results from the diaphragm experiments.

Figure 4.7a,b,c shows matching results for synthetically transformed diaphragm, lung lobe and airway tree surfaces transformed with $t_{vertical}$ of 25 mm, t_{inward} of 10 mm and sampled with 2000 sample points, respectively.

4.3.2 Registration Experiments on Synthetically Transformed Data

The synthetic registration experiments use a data set A and its transformed version B as input to the nonlinear registration algorithm which is composed of the shape matching and the displacement field interpolation component. Experiments are solely performed on lung lobe surface segmentations. Again a transformation φ is defined by the evaluator and unknown to the algorithm. The point correspondences $< p_i, q'_i >$ that are created

Measure/Samples	**200**	**600**	**1000**	**2000**
RVOE [%]	12.47	10.89	9.54	8.26
TRE [mm] Min/Mean/Max	1.81/7.39/14.03	1.70/6.06/11.09	1.79/5.37/10.08	1.61/5.16/9.89

Table 4.7: RVOE and TRE for synthetically transformed lung volume *T32* over different sampling sizes.

by shape matching of an original and a transformed data set are used to calculate a displacement field by regularized spline-based interpolation of the point correspondences within the thin-plate spline framework. This displacement field is utilized to warp the original data set A to a data set B'. By comparing B' with the synthetically transformed data set B the two quantitative error measurements relative volume overlapping error (`RVOE`) and target registration error (`TRE`) are calculated. `RVOE` is defined as $1 - \frac{B \cap B'}{B \cup B'}$, perfect registration would result in a value of 0 while 1 resembles total misregistration. `TRE` is calculated by manually identifying a number of corresponding points in the lung lobes. Therefore the original gray value data set is transformed according to the synthetic transformation φ and afterwards the airway tree is segmented in both the original and the transformed image. Both airway tree segmentations are skeletonised and 20 branch points are extracted, respectively [Tschirren et al., 2005b]. Figure 4.7f,g shows the graph-based airway tree representations and the manually labeled airway tree branch points. By warping identified airway tree branch points of data set A according to the displacement field from the registration step and calculating their distance from the corresponding branch points in synthetically transformed data set B, the minimum, mean and maximum target registration error is computed.

Five thorax data sets are used in this experiment. The first data set is examined at varying numbers of sample points between 600 and 2000, while the other four data sets are solely examined with 1000 sample points. `RVOE` (in %) and `TRE` results (in mm) for data set *T32* are given in Table 4.7. The mean `RVOE` of all 5 data sets is 10.166% while mean target registration errors over all 5 data sets are 1.576mm, 6.734mm and 16.502 for minimum, mean and maximum TRE, respectively. $t_{vertical}$ is 25 mm and t_{inward} is 10 mm. Figure 4.7 shows one lung lobe of data set *T32* before (d) and after (e) registration.

4.3.3 Registration Experiments on Real Thorax Data

The first evaluation experiment on the real thorax data makes use of a sheep thorax data set at five different breathing states between TLC and FRC, these data sets are called *T32*, *T24*, *T16*, *T8* and *F*. Each of them contains a lung lobe segmentation from which 1000 points are sampled. Four data subsets are built, with the first subset consisting of states

[mm]/Subset	**{T32,T24,T16}**	**{T24,T16,T8}**	**{T16,T8,F}**	**{T32,T16,F}**
Min Distance	0.1538	0.2732	0.2458	0.1997
Mean Distance	7.76223	7.9934	8.2067	8.2582
Max Distance	24.6284	24.2931	23.9196	30.1443

Table 4.8: Real data experiment showing accuracy of the interpolation of intermediate breathing states for five data sets.

[mm]/Data Set	**T32/T24**	**T32/T16**	**T32/T8**	**T32/F**
Min Distance	0.9465	0.3671	0.5833	2.4355
Mean Distance	5.6503	4.9746	4.8215	6.2446
Max Distance	22.2019	18.8932	14.1134	17.2176
[mm]/Data Set	**T24/T16**	**T24/T8**	**T24/F**	
Min Distance	1.4852	1.3285	1.8347	
Mean Distance	7.1077	7.4101	7.9520	
Max Distance	22.1003	22.5588	20.8391	

Table 4.9: Real data experiment showing target registration errors of pairs of data sets.

{*T32*, *T24*, *T16*}, the second subset consisting of {*T24*, *T16*, *T8*}, third subset {*T16*, *T8*, *F*} and fourth subset {*T32*, *T16*, *F*}. For each of these subsets the transformation T' relating first and second state of the subset and T'' relating second and third state of the subset is calculated by using the shape context matching approach. Further the transformation T''' is calculated relating first and third state of the subset. By comparing the results of applying T''' and $T''(T')$ on the points of the first state of each subset minimum, mean and maximum distances are calculated and shown in Table 4.8.

For the interpolation of the displacement fields these evaluations inherently assume a linear relationship between the different states in the breathing cycle, which is not necessarily the case for real thorax breathing motion. Therefore, a second evaluation experiment is performed on the real thorax data by using the airway tree branching points as corresponding landmarks again to calculate landmark `TRE`. The same procedure as described in the previous section is used. Registering two real data sets results in a displacement field which is applied to the manually identified airway tree branching points of the first data set and compared to the manually identified airway tree branching points of the second data set. In this way minimum, mean and maximum target registration errors are computed (see Table 4.9). Figure 4.7 shows the result of overlaying data sets *T8* and *T32* before (f) and after (g) registration.

4.4 Discussion

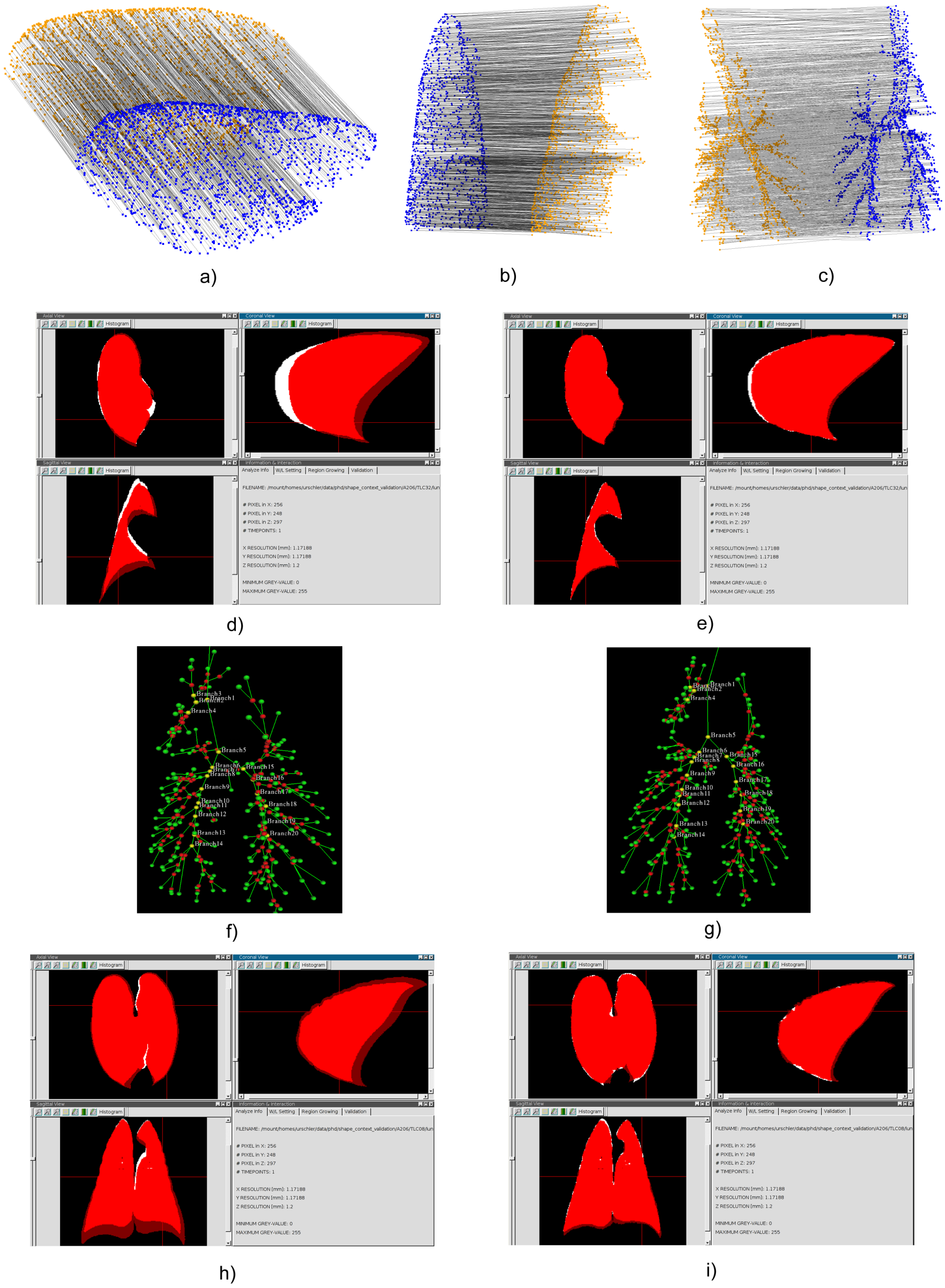
Experiments show that the nonlinear registration method based on the shape context matching approach is a well-suited method for a variety of soft tissue organ surfaces. Shape matching and subsequent volume registration and warping was successfully performed using airway tree, lung, lung lobe and diaphragm segmentations. Experiments on synthetic and real data sets show mean registration errors in the range of 5 to 8 mm. Since there is a sophisticated image processing pipeline necessary for the whole registration process, the final error is composed of several components. First of all segmentation and surface point sampling introduce errors that have an effect on the registration result. The matching evaluation experiments using synthetically transformed point sets assess the effect of segmentation errors. They show that the shape context descriptor based matching approach is a suitable method if there are no segmentation errors present, since the percentage of correctly identified correspondences stays stable at almost 100% over a wide range of different numbers of sample points. However, a very large number of sample points introduces inaccuracies due to the discrete nature of the shape context histogram. For a stable matching quality histogram size would have to be increased at larger numbers of sample points which leads to high computation times. The matching evaluation experiments using synthetically transformed input volumes and independently sampled point sets clearly show the dependency of matching quality on the number of sample points. In accordance to common insights from sampling theory an increase in the number of sample points increases matching quality until it stays stable between 4 and 5 mm starting around 2000 sample points. This is consistent with the former experiment, since at higher sampling rates histogram discretization issues come into play. The error of 4 to 5 mm can be explained by point sampling issues. The matching method uses independently sampled points of two shapes, the established correspondences implicitly carry an error in the sampling positions which becomes relevant in the graph-based optimization to find the one-to-one correspondences.

Evaluation experiments with varying synthetic deformation forces show that larger deformations induce a degrading matching accuracy, this indicates that the matching method is not suitable to model extremely large movements. Synthetic and real-data registration experiments also show acceptable behavior. Mean `TRE` are in the range of 5 to 9 mm. Here the effects of the preceding image processing pipeline steps have to be taken into account. Segmentation inaccuracies, surface sampling issues and discretization effects in the shape context matching all sum up to form an error which gets slightly smoothed by the thin-plate spline interpolation process at those points that are interpolated, but not at the correspondences. In addition, calculation of target registration errors is an error-prone process by itself, especially since for these evaluation studies the branchpoint labeling was not performed by a medical expert but by the author.

The outcome of this work is a method that allows to calculate deformation fields capturing organ motion. The focus lay on breathing motion but the basic concepts can be used for other kinds of motion as well. Displacement fields may be derived from organs like diaphragms or lung surfaces which enables one to build statistical models of organ motion, e.g. by using Active Shape Models [Cootes and Taylor, 2001] in subsequent stages. A statistical model of breathing motion resembles a very useful tool for segmentation and registration applications in medical imaging areas that suffer from motion artifacts.

4.5 Conclusion

In this chapter a three-dimensional extension of the shape context approach for matching and registering 3D surfaces was presented and experiments on lung, lung lobe, airway tree and diaphragm surfaces were shown. Shape context based nonlinear registration is a promising technique which has to be studied further to be able to lower the `TRE`. Future work will include algorithm fine-tuning on the one hand and more evaluation experiments on the other hand. Fine-tuning might be performed by using the Normalized Mutual Information measure instead of Normalized Correlation Coefficients for the intensity-based cost function and by finding a better way to get rid of segmentation and discretization errors and outliers. The thin-plate spline displacement field interpolation might be improved by using the weights from the matched points for regularization instead of simply having a single regularization parameter. Speed-up of the pipeline should be performed by approximating the `TPS`, since this is the most time-consuming part of the processing. A more appropriate evaluation study should include a large number of data sets and use manual correspondences identified by experts for the `TRE` calculation. Further, the method should be compared to other state-of-the-art techniques like e.g. [Chui et al., 2003]. Based on a large number of evaluated data sets, a statistical model of the displacement fields that describes breathing motion could be established.



a) b) c)

d) e)

```
FILENAME: /mount/homes/urschler/data/phd/shape_context_validation/A206/TLC32/lun

# PIXEL in X: 256
# PIXEL in Y: 248
# PIXEL in Z: 297
# TIMEPOINTS: 1

X RESOLUTION [mm]: 1.17188
Y RESOLUTION [mm]: 1.17188
Z RESOLUTION [mm]: 1.2

MINIMUM GREY-VALUE: 0
MAXIMUM GREY-VALUE: 255
```

f) g)

h) i)

```
FILENAME: /mount/homes/urschler/data/phd/shape_context_validation/A206/TLC08/lun

# PIXEL in X: 256
# PIXEL in Y: 248
# PIXEL in Z: 297
# TIMEPOINTS: 1

X RESOLUTION [mm]: 1.17188
Y RESOLUTION [mm]: 1.17188
Z RESOLUTION [mm]: 1.2

MINIMUM GREY-VALUE: 0
MAXIMUM GREY-VALUE: 255
```

Figure 4.7: Some shape matching results are depicted in a), b) & c). d) & e) show registration results of synthetically transformed data set *T32*. f) & g) give an example of corresponding airway branch points that were used to determine target registration error. h) & i) show results from registering data sets *T32* & *T8*. In d), e), h) and i) the white volumes are the original and dark volumes are the overlaid warped volumes.

Chapter 5

Feature-Based Nonlinear Registration

Contents

This chapter presents a novel nonlinear registration approach based on automatically extracted and matched feature points. Although intensity-based approaches are getting more attention by the medical registration research community compared to feature-based approaches, they face two kinds of practical problems when applied to large thoracic data sets. First, due to their mathematical complexity they require large computational effort. Second, many approaches tend to misregister small structures in the lung like vessels and airways, due to similarity measures being dominated by homogeneous areas. Further, intensity variations that occur when comparing inhaled and exhaled lungs are not always modeled due to implicit brightness constancy assumptions. Figure 5.1 shows a nonlinear registration example on lung CT data where the widely-used Demons algorithm [Thirion, 1998] leads to misregistrations of vessel structures. Feature-based algorithms prevent such problems by explicitly concentrating on a subset of the provided image information, a subset which is derived from a segmentation or directly from the vascular structures of interest. This not only prevents misregistrations but also has a positive effect on computation time.

Consequently, an automatic feature matching and registration pipeline was established using state-of-the-art techniques from the computer vision community. This pipeline contains Förstner corner detection [Förstner, 1986], forward-backward matching using a 3D

Scale-Invariant Feature Transform (`SIFT`) descriptor [Lowe, 2004] and a global descriptor similar to shape context [Belongie et al., 2002] and dense displacement field estimation in the thin-plate spline (`TPS`) framework [Bookstein, 1989]. Especially `SIFT` and shape context have proved to be very powerful approaches in traditional computer vision applications like wide-baseline matching or object recognition. The main contributions of our work are the 3D extension and the runtime optimization of these stages and the setup of an automatic feature registration pipeline for application in medical imaging. Related work on feature-based matching and registration has already been presented in 2.2.1. Note that none of these works addresses the problem of fully automatic feature extraction, matching and registration.

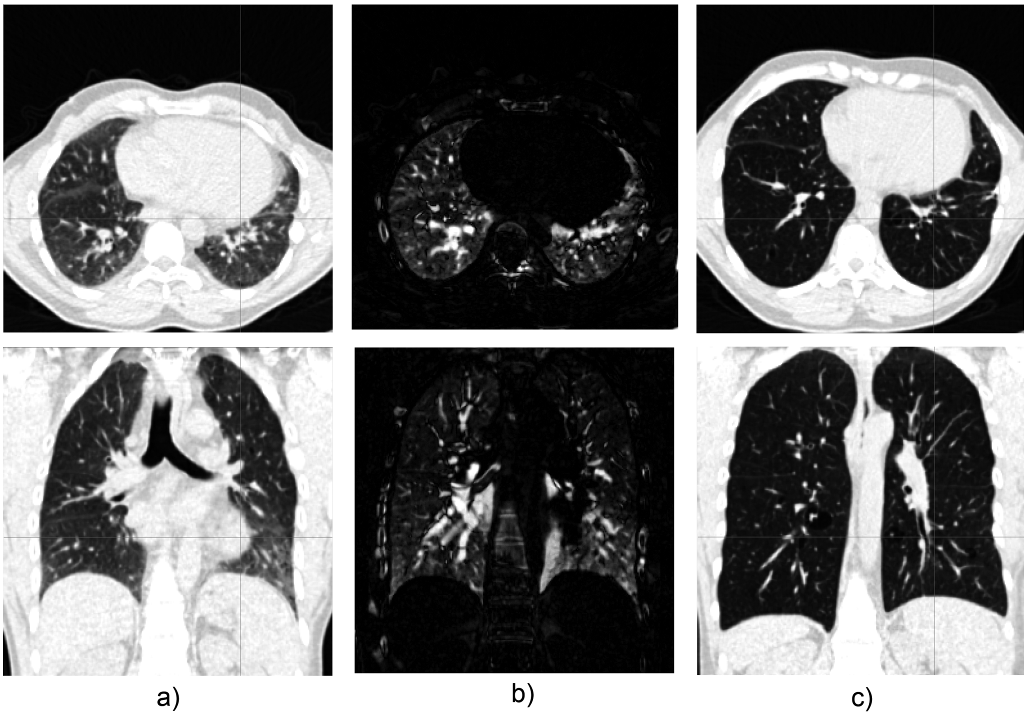


Figure 5.1: Example data set (data set F from evaluation data) with axial slices in top and coronal slices in bottom row. a) and c) show the differences in inspiration and exhalation. b) gives the difference image after Demons registration. Note the misregistered vessel structures.

This chapter is organized as follows. Section 5.1 introduces the components of the feature-based registration scheme. Section 5.2 shows setup and results of synthetic and real data experiments. Note that we only present a proof-of-concept evaluation in this chapter, since the developed algorithm will be evaluated more thoroughly in Chapter 7. Finally Section 5.3 discusses the obtained results and concludes the chapter.

5.1 An Automatic Feature-Based Registration Pipeline

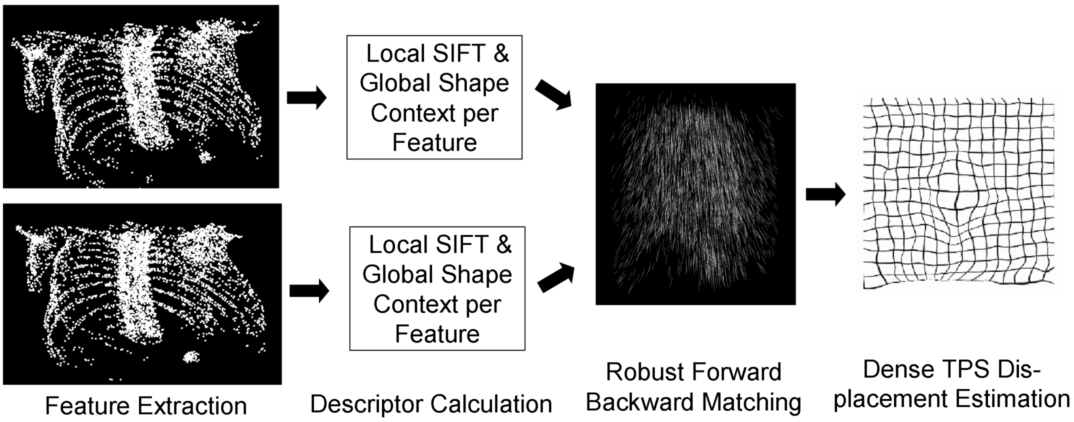


Figure 5.2: Nonlinear matching and registration pipeline. The feature extraction stage only shows extracted bone corners, while the method also extracts lung and tissue features.

Breathing motion mainly stems from two sources, the diaphragm and the rib cage muscles. Expected tissue deformations are not extremely large even in the case of matching full exhalation to full inhalation and they change smoothly over the image domain. The deformations are locally similar to translations with small tissue deformations, formally speaking the Jacobians of the displacements are nearly identity matrices. These considerations imply that a robust and reproducible feature extraction step producing large numbers of feature candidates, followed by the automatic matching of feature descriptors suited to the expected local deformations is a valid approach to find corresponding structures in the images (see Figure 5.2 for the matching and registration pipeline). Due to the large similarity of local neighborhoods in lung images (vessel and airway structures) it is important to not only look at local feature descriptors but also add a notion of global correspondence. [Mortensen et al., 2005] have recently proposed a combined local and global descriptor for the matching of repetitive patterns. The combined descriptor was applied to e.g. find correspondences from images of buildings with repetitive window structures. Their ideas were adapted to solve the ambiguities with locally similar structures. The final step after establishing corresponding features is the interpolation of a dense displacement field. Since the landmark publication of Bookstein [Bookstein, 1989] this is often performed in a `TPS` framework, however recent publications [Rohr, 2001, Donato and Belongie, 2002, Guo et al., 2004] have shown that the interpolating behavior of TPS is not desirable since it may lead to foldings (*non-diffeomorphic* mappings) of the interpolated displacement field. To overcome this problem the decision to use `TPS` approximation [Rohr, 2001] was made. Algorithm 2 gives the algorithmic overview of the matching and registration pipeline that was developed in this work.

Algorithm 2 Nonlinear Feature Matching and Registration Algorithm

```
for the two input images to register do
    Reproducible detection of 3D Förstner corners as image features
    Calculation of local 3D SIFT descriptors for each feature
    Calculation of global 3D shape context descriptors for each feature
    Robust forward-backward matching using Euclidean distance on SIFT and $\chi^2$ metric
        on shape context
end for
Nonlinearly register both images by establishing a dense TPS displacement field from
    the corresponding feature matches
```

The large size of current routinely acquired CT volume data sets always poses restrictions on practically useful algorithms due to runtime and memory consumption issues. This is especially true in the case of complex nonlinear registration methods. Acquired CT data sets easily have several hundred slices with x-y resolutions of 512 by 512 voxels respectively, thereby requiring hundreds of MB in memory due to their 12 bit gray level resolution. Therefore it is necessary to carefully consider computational and memory effort of registration algorithms during the design stage. In the presented algorithm two severe performance bottlenecks were identified due to the large number of detected feature points. First, the calculation of the shape context descriptor on large numbers of feature points is critical due to its internally used log-polar histogram structure and the $O(N^2)$ complexity of descriptor calculation (with N being the number of extracted feature points). Second, the direct solution of the `TPS` equation system and the calculation of the final dense displacement field is very expensive when confronted with a large number of matched points. These strategies were developed to overcome these performance issues:

- Approximate the shape context measure with a measure that uses axis-aligned histogram bins instead of bins in a log-polar coordinate frame. This makes it possible to use 3D integral images to significantly speed up descriptor computation.
- Approximate the global `TPS` transform with a locally restricted thin-plate-spline (`LRTPS`) using a k-d tree data structure.

The following subsections describe the different pipeline stages in more detail.

5.1.1 Feature Extraction

The first step in the nonlinear registration pipeline deals with feature extraction. The need for fast and reproducible features was already extensively investigated in [Rohr, 2001]. His evaluations of several different 3D anatomical point landmark detection operators and their localization errors resulted in the recommendation to use the Förstner operator [Förstner, 1986] for reproducible 3D feature detection. Compared to other presented operators like those that use local first and second order partial derivatives [Kitchen and Rosenfeld, 1982] or curvature extrema [Thirion and Gourdon, 1995] it has the advantage of only using first order partial derivatives in its formulation. This speeds up computation and makes the operator less prone to noise which gets enhanced by the discrete approximations of higher-order differentiation operators. Further, the reported experimental results show that the Förstner operator gives superior landmark detection and localization capabilities in terms of repeatability and reproducibility.

Structure Tensor Representation

The Förstner operator is based on the matrix $S_{\nabla I}$ which represents for each location the averaged dyadic product of the local intensity gradient:

$$S_{\nabla I} := \overline{\nabla I \nabla I^T} = \begin{pmatrix} \overline{(\frac{\partial I}{\partial x})^2} & \overline{(\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y})} & \overline{(\frac{\partial I}{\partial x})(\frac{\partial I}{\partial z})} \\ \overline{(\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y})} & \overline{(\frac{\partial I}{\partial y})^2} & \overline{(\frac{\partial I}{\partial y})(\frac{\partial I}{\partial z})} \\ \overline{(\frac{\partial I}{\partial x})(\frac{\partial I}{\partial z})} & \overline{(\frac{\partial I}{\partial y})(\frac{\partial I}{\partial z})} & \overline{(\frac{\partial I}{\partial z})^2} \end{pmatrix}$$

This symmetric, positive semi-definite matrix is also known as *structure tensor* [Jaehne, 1993], the overline represents averaging in a local neighborhood. The structure tensor is a measure that contains information on orientation and intensity of the surrounding structure around the location $I(\mathbf{x})$. It is derived from the tensor (outer) product $\nabla I \nabla I^T$ of the intensity gradients. Although this product contains no more information than the gradient itself, there is the advantage that it can be averaged without cancellation effects in areas where gradients have opposite signs. It is very common to perform the averaging by smoothing with a Gaussian kernel G_ρ with standard deviation ρ, leading to

$$S_{\nabla I} = G_\rho * (\nabla I \nabla I^T)$$

in a more compact notation. The smoothing over the local neighborhood has the effect that the structure tensor contains additional information about the homogeneity of orientations within the neighborhood. This information can be accessed via an eigendecomposition $S_{\nabla I} = \mathbf{U}\Phi\mathbf{U}^T$, where the eigenvectors (columns of U) denote the principal orientations and the eigenvalues (diagonal elements of Φ) denote the amount of intensity change along the principal orientations. The eigenvector to the smallest eigenvalue determines the dominant orientation, while the trace of $S_{\nabla I}$ (i.e. the sum of the eigenvalues,

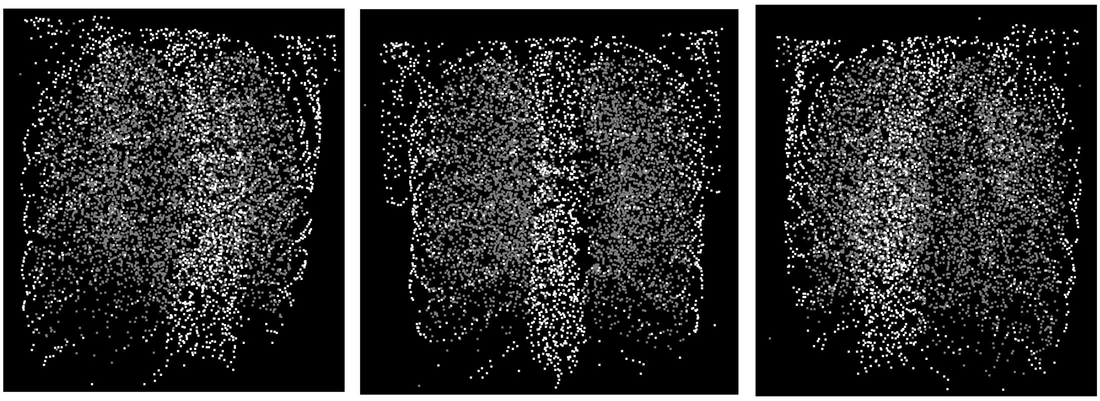


Figure 5.3: Three views of a 3D visualization of extracted lung and bone feature points from data set A.

or the sum of the diagonal elements of $S_{\nabla I}$) determines the magnitude of the dominant orientation.

Förstner Corner Detection

In two dimensions corners are defined as image structures where the eigenvalues of the structure tensor differ significantly. This can be described by the ratio of the determinant and the trace of the eigenvalues. A generalization to three dimensions has been presented in [Rohr, 2001], who also showed that this corner detector gives reliable responses for medical volumetric images. The Förstner corner response function is given as

$$F(x,y,z) = \frac{det(S_{\nabla I})}{(\frac{1}{3}tr(S_{\nabla I}))^3}.$$

The response of the operator lies between 0 and 1 and characterizes the similarity of the eigenvalues of $S_{\nabla I}$ without explicitly calculating them. Maxima of this response function have to be identified by thresholding to identify corners. Finally corners are labeled according to the gray-value of the underlying pixel, resulting in different corner classes like lung, tissue or bone. We refer to the set of extracted feature points from fixed and moving images as $\{\mathbf{x}_F\}$ and $\{\mathbf{x}_M\}$, respectively. Figure 5.3 shows extracted lung and bone features from an example data set, tissue features are not shown to achieve a better visualization.

5.1.2 Local SIFT Feature Descriptor

For each feature derived in the previous stage a distinctive local feature descriptor has to be built which is robust to deformations as well as feature localization errors. The `SIFT` descriptor [Lowe, 2004] which has been recently introduced in the computer vision community was identified as being suitable for this medical application. Performance evaluations show its excellent matching behavior on various kinds of transformations [Mikolajczyk and Schmid, 2005]. In this work only the descriptor representation of `SIFT` is used, since the keypoint localization is performed using Förstner corners instead of the computationally more intensive scale space localization approach. The original 2D `SIFT` descriptor representation is a histogram of gradient location and orientation in a local neighborhood. Location is quantized in a 4×4 grid and the gradient angle quantized into eight orientations, resulting in a descriptor of total size 128. For illumination invariance, the descriptor usually is normalized by its L_2-norm. An illustration of the gradient orientation histogram setup is given in Figure 5.4. We do not incorporate the notion of scale here, the feature points we are interested in are assumed to come from the original scale level.

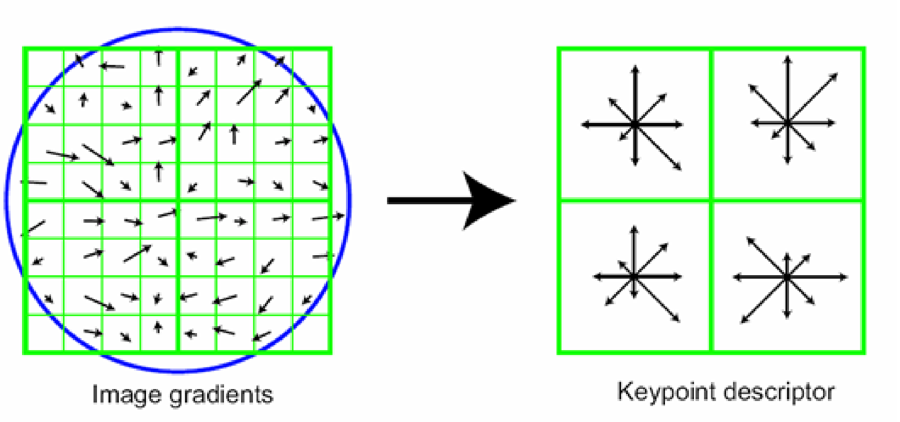


Figure 5.4: Illustration of the 2D SIFT gradient orientation histogram bins. Gradient orientations are sampled from the 2×2 location bins into 4 histograms. The blue circle indicates the Gaussian weighting of the gradient orientations.

To apply the `SIFT` descriptor representation on volume data an extension to 3D is necessary. The 3D `SIFT` descriptor quantizes gradient location in a $2 \times 2 \times 2$ grid while gradient orientation (consisting of two angles that comprise a spherical coordinate system) are quantized into two orientation histograms of 8 bins respectively. After putting the orientation histograms sequentially into the descriptor, the total descriptor size is 512 and it is again normalized by its L_2-norm. Note that the coordinate frame of the used location grid is aligned to the volume coordinate system, as opposed to the description of the original 2D `SIFT` descriptor. In the 2D case before descriptor construction a principal gradient orientation is derived and the local neighborhood is transformed into this coordinate frame to get rotation invariance. So the proposed 3D `SIFT` descriptor is not rotation-invariant. This saves computation time and moreover rotation invariance is not critical for the intended application, since breathing motion locally does not lead to strong rotations. Preventing an incorporation of different rotations further has the advantage of getting an improved discriminative behavior of the descriptor.

5.1.3 Global Shape Context Feature Descriptor

The 3D shape context descriptor [Belongie et al., 2002, Urschler and Bischof, 2005] treats objects as (possibly infinite) point sets and assumes that the shape of an object is captured by a finite subset of its points. We have already discussed this method in Section 4.2, the reader may refer there for more details. The result of the shape context descriptor calculation is a histogram of relative positions of all other points of a shape with respect to the point under consideration. This histogram uses bins that are uniform in a three-dimensional spherical coordinate system (θ,ϕ,r). The r coordinate axis is logarithmically sampled, such that positions of nearby sample points have stronger influence on the descriptor than ones located farther away. The log-polar histogram binning of this method is a performance problem due to the large number of feature points that are detected.

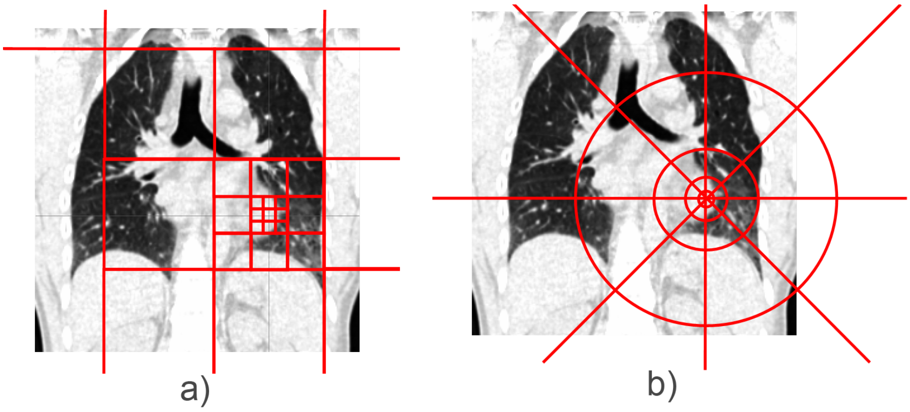


Figure 5.5: Approximated global context (a) vs shape context (b) histogram bin structure.

To overcome this performance problem an approximation of the shape context descriptor was used. This approach was recently developed in our group [Bauer, 2005] offering comparable matching behavior to the classical shape context approach. The descriptor replaces the log-polar histogram bin structure with a structure based on rectangular, axis-aligned image patches. Figure 5.5 shows the histogram bin structure of the classical and the approximated shape context descriptor. The patch size increases exponentially with increasing distance from the center (i.e. feature point). This strategy preserves some local information close to the feature point and generalizes at larger distance to a coarser quantization. A key advantage of this descriptor which is extremely important in case of medical applications with large 3D volumes is its low computational cost. By using 3D integral images [Viola and Jones, 2004] to store point locations the cost for computing the descriptor is linear in the number of layers since the features of interest in a rectangular, axis aligned patch can be computed in linear time from eight sampled values in a 3D integral image. Another advantage is that the size of the descriptor is linear in the number of layers, while the covered areas increase exponentially.

5.1.4 Robust Feature Matching

To find corresponding feature points a matching algorithm has to be used. The previous stages have established a local and a global descriptor for each feature point, the task of the matching stage now is to find those point pairs from two volumes that minimize a cost function derived from the descriptors. The simplest possible matching algorithm would be to relate each feature point in the first image to all feature points in the second image via the cost function. The point pair exhibiting the smallest value of the cost function is taken as the found correspondence. So for a point $\mathbf{x}_{F,1} \in \{\mathbf{x}_F\}$ from image I_F a corresponding point $\mathbf{x}_{M,1} \in \{\mathbf{x}_M\}$ from image I_M is identified. Note that it is possible that $\mathbf{x}_{M,1}$ is the best match for another point in $\{\mathbf{x}_F\}$ as well, the method therefore lacks in consistency. To make this approach more robust in terms of consistency, occlusions and erroneous feature point detection an idea from stereo matching is borrowed. Forward-backward matching in stereo vision (as presented in Fua [Fua, 1993]) takes care of the above mentioned problems. The basic idea is to perform the matching step twice by reversing the roles of the two feature point sets and considering as valid only those matches for which the corresponding points are identical when matching from $\{\mathbf{x}_F\}$ to $\{\mathbf{x}_M\}$ and from $\{\mathbf{x}_M\}$ to $\{\mathbf{x}_F\}$. So in addition to $\mathbf{x}_{F,1}$ being the best match for $\mathbf{x}_{M,1}$, $\mathbf{x}_{M,1}$ must also be the best match for $\mathbf{x}_{F,1}$. In practice the search area that is used for determining the correspondences is locally restricted according to the maximum amount of expected tissue deformations.

The cost function used in both the forward and the backward matching step is a weighted linear combination of distance metrics. The distance metric for the `SIFT` feature descriptor is Euclidean distance in the 512-dimensional feature space. Given two `SIFT` descriptors SF_i and SF_j the distance metric is defined as

$$d_{SIFT} := \sqrt{\sum_k (SF_{i,k} - SF_{j,k})^2}.$$

Shape context descriptors SC_i,SC_j are compared using a χ^2 statistic

$$d_{SC} := \chi^2 = \frac{1}{2} \sum_k \frac{(SC_{i,k} - SC_{j,k})^2}{SC_{i,k} + SC_{j,k}}$$

on the descriptors SC_i and SC_j. Both distance metrics are normalized between 0 and 1. The total cost function is given by the weighted linear combination

$$d = \omega d_{SIFT} + (1 - \omega) d_{SC}$$

where ω is a weighting factor that represents the relative importance of global and local matching cost term. Matches with a cost function value above some threshold T_d get discarded.

5.1.5 Dense Displacement Field Interpolation

The final step in the registration pipeline is the estimation of a dense displacement field from the sparse matching result. For this purpose the `TPS` framework is used [Bookstein, 1989, Rohr, 2001]. In its original formulation the interpolating behavior of the `TPS` often is too restrictive and might lead to over-fitting to the correspondences or folding of the displacement field in case of erroneous correspondences. Therefore, the findings of [Rohr, 2001] for approximating `TPS` mappings were considered. A regularization term is added to the formulation, which is steered by a parameter λ, weighting the trade-off between interpolation and smoothness of the solution. Further, a landmark error term is introduced to give each pair of corresponding features an uncertainty measure, which is directly derived from the matching costs of the feature matching stage. More details on `TPS` interpolation and approximation can be found in Appendix B.1.

It is desirable that the matching stage produces a large number of feature correspondences n_{match}. On typical volume data sets thousands of correspondences might be achieved. In this case the warping of large volume data sets is very costly, since for each voxel a multiplication with $O(n_{match})$ weighted landmarks including the calculation of $O(n_{match})$ vector norms is involved. Therefore a locally restricted version (`LRTPS`) of the global `TPS` transform is used. The source points of the feature correspondences are put into a k-d tree structure to give efficient access to its neighbor features. For each feature point a `TPS` transform restricted to a pre-defined number of neighbors is calculated. So only a subset of the total set of correspondences in the volume is taken for locally estimating the transform, however this subset is chosen as being near to the feature point. The dense displacement field approximation step now always looks for the nearest feature correspondence in the k-d tree and takes the stored local `TPS` transform to compute a displacement. Note that this approximation to the global `TPS` might lead to artifacts in the resulting displacement field, if the distribution of landmarks is not dense and even enough, or if the number of neighbors is too small. A possible solution to this problem might be a separation of the images into several regions, where each region has to contribute a pre-defined number of matches (so-called bucketing, see e.g. [Zhang et al., 1995]), however, this will also increase the number of noisy/outlier matches.

5.2 Evaluation Experiments & Results

To assess the validity of the feature-based registration approach qualitative and quantitative evaluations were performed on synthetically transformed and clinical thorax CT data sets. For the synthetic deformation experiments two different kinds of deformations were used. The first deformation model is a *Simulated Breathing Transformation* simulating rib-cage and diaphragm muscle behavior. The second synthetic transformation makes use of a regular grid of landmarks that are moved in random directions. The synthetic experiments give numbers on the root-mean-square (`RMS`) of the intensity differences before and after registration, compares the registered and the synthetic displacement fields and compares the method with the Demons algorithm. Real data experiments show the decrease in the `RMS` of the intensity differences, compare the `RMS` with the Demons algorithm and give qualitative difference images. All experiments were performed on a dual 2GHz AMD Opteron system with 8GB RAM running Linux.

5.2.1 Synthetic Deformation Experiments

Synthetic experiments were performed on seven test data sets (A,B,C,D,E,F,G) taken at inspiration, each of them having a volume size of $512 \times 512 \times 256$ voxels. The first synthetic transformation intends to model breathing behavior. The nonlinear transformation φ simulates diaphragm and rib cage movement and consists of two components $t_{vertical}$ and t_{inward}. The details of this synthetic transformation are given in Appendix B.6.

The seven test data sets were synthetically transformed with a small and a large deformation. The small deformation is defined by the translations $t_{vertical} = 25$mm and $t_{inward} = 10$mm while the large deformation is defined by $t_{vertical} = 55$mm and $t_{inward} = 25$mm. Figure 5.6 a)-c) shows the effects of these deformations on data set A. First, the feature matching produces corresponding points which can be compared to the ground truth simulated breathing transformation in terms of the `RMS` of the displacement difference vectors (RMS_{disp}) and the maximum of the lengths of the displacement difference vectors (MAX_{disp}) over all correspondences. Evaluations showed that the RMS_{disp} varies

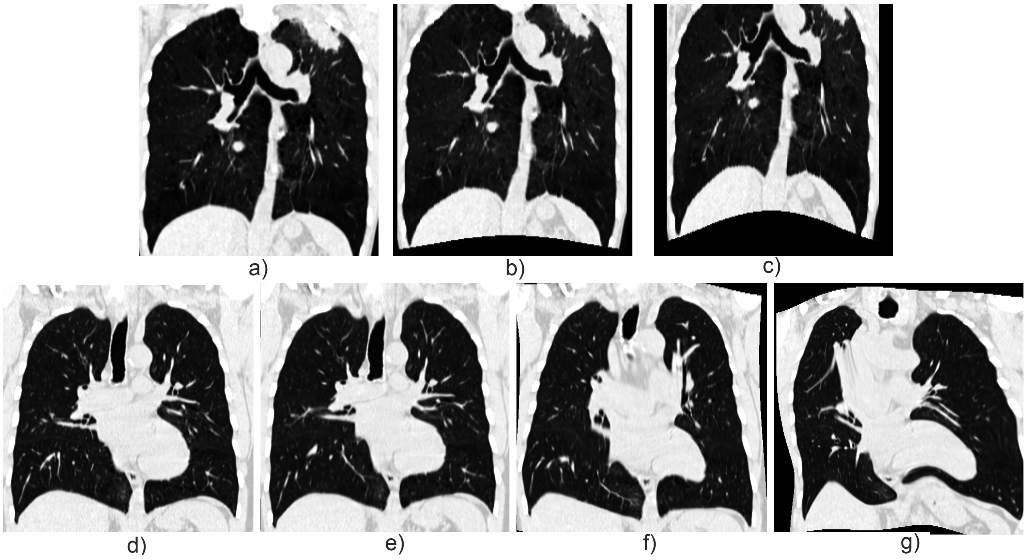


Figure 5.6: Synthetic transformations. a) original data set A, b) the small and c) the large simulated breathing deformation. d) original data set B, e) randomly displaced landmark transformation -8...+8, f) displacement -24...+24, g) displacement -48...+48.

	Measure		A	B	C	D	E	F	G	Mean
SimBreath 25-10	$RMSD_{initial}$	[HU]	385.99	327.25	303.92	359.49	318.73	316.58	316.45	332.63
	$RMSD_{demons}$	[HU]	114.44	115.56	92.05	100.06	96.33	90.21	105.95	102.08
	$RMSD_{feature}$	[HU]	45.18	45.71	43.43	50.78	46.09	41.64	47.91	45.82
	$RMS_{disp,demons}$	[mm]	4.564	5.961	5.412	4.842	4.756	4.931	5.741	5.172
	$RMS_{disp,feature}$	[mm]	0.662	0.769	0.59	0.66	0.601	1.338	1.025	0.806
	$MAX_{disp,demons}$	[mm]	34.55	38.56	38.06	29.95	28.45	35.26	37.59	34.63
	$MAX_{disp,feature}$	[mm]	8.59	9.56	8.02	8.52	8.87	12.64	11.79	9.71
	# matches		2678	2121	2330	2022	1805	5204	3714	2825.9
SimBreath 55-25	$RMSD_{initial}$	[HU]	549.02	477.81	446.84	521.65	477.07	472.65	442.95	484
	$RMSD_{demons}$	[HU]	134.46	135.69	128.57	181.32	145.77	153.67	153.11	147.51
	$RMSD_{feature}$	[HU]	69.35	82.24	71.09	96.98	70.11	74.78	78.93	77.64
	$RMS_{disp,demons}$	[mm]	6.844	7.384	6.912	5.822	5.113	7.012	6.992	6.583
	$RMS_{disp,feature}$	[mm]	1.059	1.382	1.327	1.331	1.252	2.256	2.03	1.519
	$MAX_{disp,demons}$	[mm]	39.45	38.12	39.99	41.72	38.09	43.01	42.95	40.48
	$MAX_{disp,feature}$	[mm]	15.22	21.44	19.84	23.49	18.24	24.14	21.66	20.58
	# matches		1940	1424	1709	1287	1277	2778	2856	1895.9

Table 5.1: Simulated breathing transformation. Registration results in terms of RMS intensity differences and displacement difference vectors.

Measure		-8...+8	-16...+16	-24...+24	-32...+32	-48...+48
$RMSD_{initial}$	[HU]	164.89	233.48	289.76	333.94	397.89
$RMSD_{demons}$	[HU]	161.13	188.67	193.75	184.12	253.81
$RMSD_{feature}$	[HU]	90.58	169.89	236.83	271.113	360.72
$RMS_{disp,demons}$	[mm]	5.834	9.374	13.874	17.099	19.933
$RMS_{disp,feature}$	[mm]	4.802	8.355	11.172	15.792	18.562
# matches		1729	1101	426	312	87

Table 5.2: Synthetic TPS transformation results. Registration results in terms of RMS intensity differences and displacement difference vectors.

between 0.265mm and 0.314mm for the small and between 0.558mm and 2.479mm for the large deformation over the data sets. Accordingly MAX_{DISP} varies between 2.62mm and 8.59mm for small and between 14.95mm and 28.37mm for large deformations respectively.

Table 5.1 gives the results of the synthetic registration experiments. All comparisons are always performed only on those regions which are present in both registered data sets. The `RMS` of the intensity differences before ($RMSD_{initial}$) and after ($RMSD_{feature}$) registration are calculated, as well as the difference of the resulting and the synthetic displacement fields in terms of the `RMS` of the displacement vector field RMS_{disp} and the maximum of the lengths of the displacement difference vectors MAX_{disp}. The algorithms performance is compared to the widely used Demons [Thirion, 1998] algorithm. Its implementation was taken from the Insight Segmentation and Registration Toolkit [ITK, 2006]. The Demons algorithm uses a five level multi-resolution framework to calculate a smooth displacement field with a fixed number of iterations per multi-resolution level (between 100 and 15 from coarse to fine) in a gradient-descent scheme. Figure 5.7 a)-c) shows difference images of data set D which had the worst behavior in terms of decreasing the `RMS` of the intensity differences using the large simulated breathing deformation.

The second synthetic transformation is calculated using a number of evenly distributed landmark points and randomly assigning displacements to these landmarks. The amount of the displacement is increased up to the sampling size of the landmark distribution. These displacements are not physically motivated and the larger the assigned displacements are the harder it is to correct them. The dense synthetic displacement field is calculated using a `TPS` interpolation. Note that we use `TPS` interpolation here not `TPS` approximation. For data sets size 512 × 512 × 256 every 64 voxels a landmark is placed in the original image. This leads to a grid of 7 × 7 × 3 landmarks. Now a random displacement is calculated for each landmark coordinate in the range from -8 to +8 voxels. This is repeated five times while always doubling the displacement range. Evaluations are solely performed on data set B, Figure 5.6 d)-g) shows the effect of these synthetic displacements. The main motivation of these experiments is to determine the degree of deformation where the algorithms are not capable to register the data anymore. Table 5.2 gives the results of these experiments.

5.2.2 Clinical Data

The algorithm was also evaluated on four clinical thoracic data sets consisting of two scans at different breathing states. The data sets show different problem characteristics. Data sets B and G differ by a small breathing deformation and intensity variations due to contrast agent application. Data set G additionally shows a lung disease in the upper lobe region. Data sets E and F differ by large breathing deformations and the images have intensity differences due to a lung disease making them very hard to register. For the clinical data no gold standard displacement was available for comparison, therefore solely the decrease in the `RMS` of the intensity differences before and after registration are calculated. Again the novel feature-based algorithm is compared with the Demons algorithm. The `RMS` of data set B was decreased from 201.57HU to 129HU (Demons) and to 104.83HU (feature-based). Data set E decreased from 403.49HU to 235.88HU and to 197.45HU, data set F from 413.62HU to 288.31HU and to 294.98HU and finally data set G from 367.66HU to 274.14HU and to 241.43HU. The numbers of found correspondences lies between 685 and 1632. For qualitative results difference images are shown in Figure 5.7 d)-f) for data set B and g)-i) for data set E.

5.3 Discussion & Conclusion

The different stages of the proposed feature-based algorithm require some parameters to be chosen. In the matching stage normalized local and global descriptors are used to form the matching cost function. We chose the weight ω that represents the influence of local and global descriptors to be 0.5, i.e. the influence of both descriptors is equal. This seems to be a reasonable choice in many cases, however if the probability of self-repeating structures becomes very high, this parameter should be modified in favor of the global descriptor. A meaningful threshold T_d has to be found to exclude bad matches which was empirically chosen at 0.3. The matching stage also produces some outliers. The MAX_{disp} measures reflect this fact, especially in the matching evaluation. However, the registration stage with the approximating `TPS` framework takes the magnitude of the matching cost into account, such that outlier matches have a low influence on the final dense displacement field. Parameter λ of the `TPS` displacement field approximation was selected as $\lambda = 0.005$ after experimenting with several data sets. The `LRTPS` implementation needs a choice on the number of neighboring points that defines a local `TPS`, this parameter was set to 200. None of these parameters is critical, i.e. small changes of these parameters lead to similar

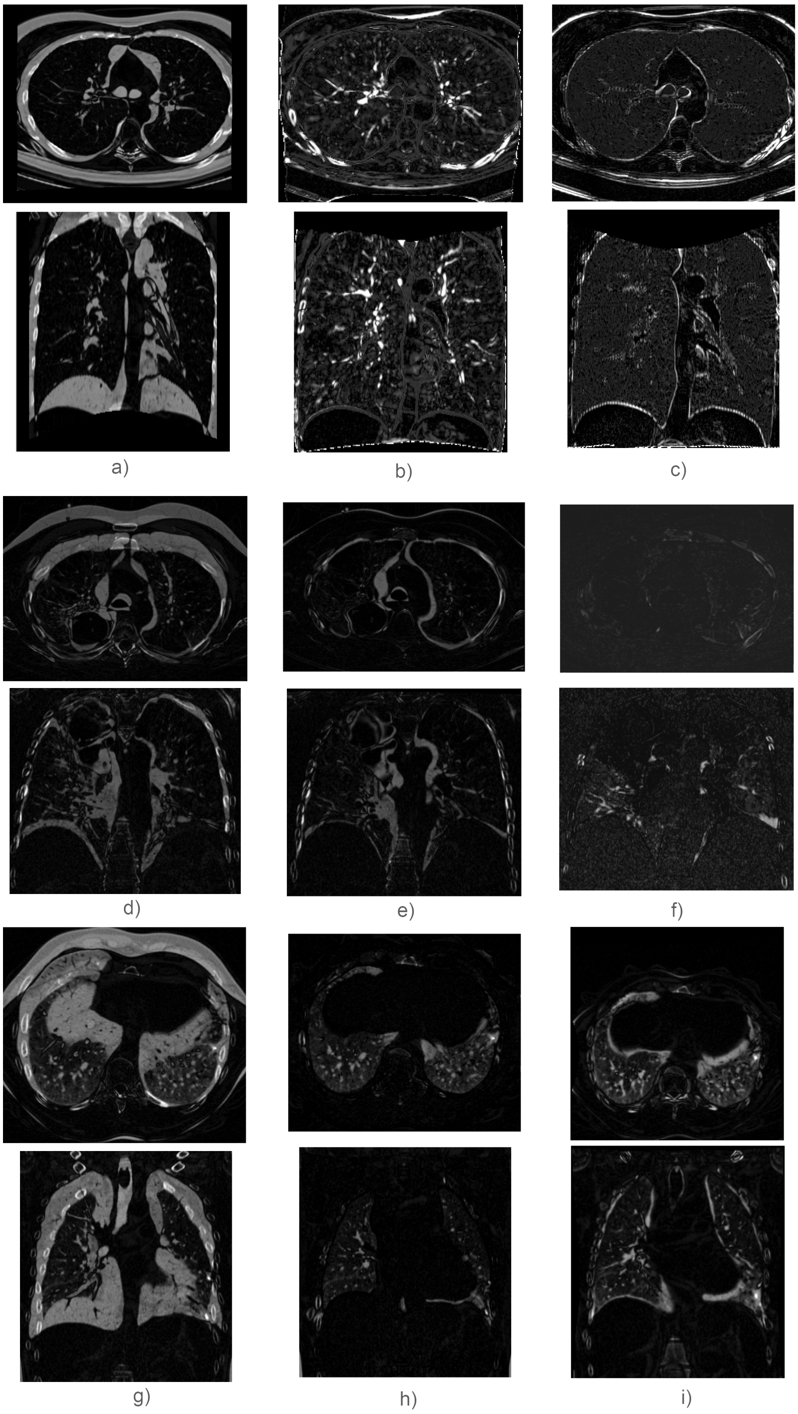


Figure 5.7: Selected Results. Top row always axial, bottom row always coronal slices. Left image shows difference image before, middle image after Demons and right image after feature-based registration. a)-c) shows synthetic results of data set D with a simulated breathing transformation of $t_{vertical} = 55$mm and $t_{inward} = 25$mm. d)-f) shows results on real data set B, g)-i) on real data set E.

results.

The two goals of the proposed algorithm, to be faster and similarily accurate as a state-of-the-art nonlinear intensity-based registration algorithm, were both met. Computation time of the feature-based algorithm on the 512 × 512 × 256 data sets ranges between 1632s and 2282s, depending on the number of identified correspondences. The largest part (more than 50%) of the algorithm runtime still goes into the calculation of the dense displacement field. The Demons algorithm takes on average around 2540s for registration of two 512 × 512 × 256 data sets. If one further increases data sets size a feature-based algorithm will be even more efficient due to its inherent reduction of matching complexity.

Registration accuracy of the feature-based algorithm exceeds the Demons algorithm in most of the synthetic examples. Demons only performs better on the evenly distributed landmark `TPS` experiments with a high degree of random deformation where the feature-based approach is not able to find enough correspondences. However, these deformations are physically implausible and not entirely representative. Especially the simulated breathing transformation was very accurately registered using the feature-based approach. This is reflected in the substantial decrease of the `RMS` of the intensity difference and the `RMS` of the displacement difference vector fields (in the order of 0.5mm to 2.0mm). The difference images of the simulated breathing experiment (Figure 5.7) illustrate the problems of the Demons approach with the vascular structures. The real data experiments show that the performance of the feature-based algorithm is comparable to Demons. Performance on data sets B and G was better, while the performance on the very difficult data sets E and F is similar. Although Demons shows lower `RMS` values, the difference images of the feature-based approach have the same quality. However, the feature-based approach also has some problems with registration of vessel structures on the difficult data sets. The largest disadvantage of the feature-based approach is to get a large number of robust, evenly distributed feature matches. This can not be guaranteed in the current implementation, which explains the registration problems and can also theoretically lead to artifacts at the edges of the local `TPS` regions, however in our set of images, these were not observed. A good direction for further work would therefore be the investigation of more sophisticated methods to approximate and speed-up the `TPS` like e.g. proposed in [Beatson and Newsam, 1998]. For another comparison of this method with more state-of-the-art nonlinear registration algorithms, the reader may refer to Chapter 7. However, a direct comparison to these results is difficult since we increase the number of iterations of the Demons algorithm in Chapter 7 and we modified the comparison measures. We also conclude that the influence of the feature extraction stage is the main accuracy bottle-neck, therefore, a comparison to other feature extraction methods would be interesting. All of the later stages in the matching and registration pipeline crucially depend on accurate and reproducible feature extraction, so one might think of coming up with a sub-voxel accurate extraction algorithm. Sub-voxel accuracy would have to be used later as well for the local descriptor calculation and one could also think of a way how to locally refine matches during the matching stage, i.e. find a more probable sub-voxel match after a voxel-accurate match has been identififed. On the other hand, one should be aware of the fact that these procedures will increase the algorithm run-time dramatically.

Chapter 6

Hybrid Feature and Intensity Based Nonlinear Registration

Contents

This chapter describes a novel hybrid nonlinear registration algorithm that combines feature- and intensity-based registration in a variational framework. Feature-based registration often uses manually or automatically derived point landmarks [Rohr, 2001]. After matching the resulting sparse displacement field from the point correspondences is interpolated/approximated in the `TPS` [Bookstein, 1989, Rohr, 2001] framework which crucially depends on an even distribution of point correspondences over the data set. However, this even distribution can rarely be guaranteed by automatic matching algorithms. Consequently we have decided to combine feature and intensity based methods, leading to a very flexible registration scheme that still provides good accuracy in the vessel structures. Speed up of the convergence of intensity based methods will also be increased this way.

Building upon the automatic point landmark extraction, matching and registration from Chapter 5 the goal of this chapter is to replace the `TPS` displacement field estimation by intensity-based registration, thus establishing the hybrid nonlinear registration scheme. The assumption that the coupling of landmark- and intensity-based registration is more powerful than the individual steps has already been investigated by other authors. We have already presented a number of hybrid approaches in Section 2.2.2.7. Many of these hybrid schemes need a large computational effort during registration or during automated segmentation. Others require manual landmark selection or tedious semi-automatic segmentation methods for pre-processing. In this work we present a novel fully-automatic hybrid nonlinear registration framework that is both efficient and independent of tedious or time-consuming pre-processing. This framework uses a flexible variational formulation [Modersitzki, 2004] where different data and regularization terms can be used and the corresponding landmarks from the automatic landmark matching of the previous chapter act as an additional regularization constraint.

The work we present in this chapter is closely related to [Hellier and Barillot, 2003] and extends [Fan and Chen, 1999a] to be less dependent on a pre-segmentation. The proposed algorithm is based on the fully automatic feature extraction and matching pipeline from Chapter 5 and combines this feature-based step with optical flow based intensity registration in a generic variational framework [Modersitzki, 2004] consisting of a data and a regularization term. This combination leads to a flexible formulation where different data and regularization terms can be used and the matched features always form an additional regularization constraint. In the proposed algorithm the variational framework uses an $L_2 - norm$ data term and an anisotropic image-driven regularization term [Weickert and Schnörr, 2001a] leading to an optical flow formulation similar to [Fan and Chen, 1999a, Hellier and Barillot, 2003]. In contrast to [Hellier and Barillot, 2003] feature-based registration is not restricted to segmented surfaces and the derivation of the combined registration energy term incorporates image information for the regularization and weighs the feature matches according to their matching accuracy. In [Fan and Chen, 1999a] the derivation of the intensity-based registration is very similar to the formulation used in this work, however, they also rely on several pre-segmentation steps making the algorithm very time-consuming and they only use the feature information as a hard initialization for the intensity-based registration without taking matching costs into account.

This chapter is organized as follows. Sections 6.1 and 6.2 introduce the components of the hybrid registration scheme. Section 6.3 presents the proposed novel hybrid registration algorithm. Section 6.5 shows setup and results of synthetic and real data experiments, while Section 6.6 discusses the obtained results and gives an outlook on future work.

6.1 Landmark-Based Registration

In Chapter 5 an automatic landmark extraction, matching and registration approach consisting of a four-stage pipeline was constructed, inspired by state-of-the-art computer vision matching techniques [Mikolajczyk and Schmid, 2005]. We used Förstner corner detection, local `SIFT` and global shape context descriptors, forward-backward matching and and a `TPS` approximation for estimating the dense displacement field. However, the algorithm, while being very efficient in the first three stages, still needs large computational effort for the final dense displacement field estimation and shows a crucial dependence on the even distribution of matched landmarks over the input volume. For a quick overview of the algorithm see Algorithm 2 and Figure 5.2.

6.2 Variational Intensity-Based Registration

As presented in Section 1.4 the general nonlinear registration problem can be formulated as a minimization process of a cost functional $\mathcal{J}$ that depends on a similarity function $\mathcal{D}$ and a regularization constraint $\mathcal{S}$ [Modersitzki, 2004]. Fixed image $I_F(\mathbf{x})$ and moving image $I_M(\mathbf{x})$ are considered as intensity functions $I_F(\mathbf{x}) : \Omega_F \to \mathbb{R}$ and $I_M(\mathbf{x}) : \Omega_M \to \mathbb{R}$ over the domains $\Omega_F, \Omega_M \subset \mathbb{R}^3$. The formalized registration problem is stated according to Definition 1.2 as the problem of finding a transformation φ, represented as a displacement field $\varphi(\mathbf{x}_F) = (u(\mathbf{x}_F), v(\mathbf{x}_F), w(\mathbf{x}_F))^T : \Omega_F \to \Omega_M$, that minimizes (for a weight $\alpha > 0$) the continuous cost functional

$$\mathcal{J}[\varphi] := \mathcal{D}[I_F, I_M; \varphi] + \alpha\mathcal{S}[\varphi] = min. \tag{6.1}$$

6.2.1 Choice of a Similarity Measure

A frequently used choice for a similarity measure $\mathcal{D}$ is the $L_2 - norm$ of the intensity differences. This is especially useful in intra-modality applications where a constant brightness assumption is holding. The `SSD` similarity measure is well-known in the motion estimation and computer vision literature, where it is used for *optical flow* [Lucas and Kanade, 1981, Barron et al., 1994, Weickert and Schnörr, 2001b] estimation, and it is also a wide-spread similarity measure in the medical image registration community [Fan and Chen, 1999a, Hellier and Barillot, 2003]. The following formulation characterizes the `SSD` measure:

$$\mathcal{D}[I_F, I_M; \varphi] := \mathcal{D}^{SSD}[I_F, I_M; \varphi] = \iiint\limits_{\Omega} (I_F(\mathbf{x}) - I_M(\varphi(\mathbf{x})))^2 \mathrm{d}\mathbf{x} \tag{6.2}$$

Since the `SSD` measure is quite sensitive to outliers due to squaring the intensity differences, a more robust alternative is given by the Sum of Absolute Differences (`SAD`). It is defined as the $L_1 - norm$ of the intensity differences:

$$\mathcal{D}^{SAD}[I_F, I_M; \varphi] = \iiint\limits_{\Omega} |I_F(\mathbf{x}) - I_M(\varphi(\mathbf{x}))| \, \mathrm{d}\mathbf{x} \tag{6.3}$$

Obviously it is also possible to define different similarity measures in this framework, examples of further measures (correlation and mutual information-based) can be found in [Modersitzki, 2004]. There the correlation-based similarity measure is derived as a generalization of the `SSD` measure and they also show how to define a `MI`-based measure, which nevertheless has the major drawback of being costly to compute and difficult to approximate the derivatives of the similarity measure. A further interesting similarity measure based on normalized intensity gradients has been presented in [Haber and Modersitzki, 2006].

6.2.2 Choice of a Regularizer

The minimization term $\mathcal{S}$ classically minimizes the squared norms of the gradient vectors of the displacement field components (u, v, w):

$$\mathcal{S}[\varphi] := \mathcal{S}[\varphi]^H = \iiint\limits_{\Omega} |\nabla u|^2 + |\nabla v|^2 + |\nabla w|^2 \, \mathrm{d}\mathbf{x} \tag{6.4}$$

This enables a certain amount of smoothing of the displacement field by penalizing large gradient differences in neighboring displacement vectors. According to the taxonomy of [Weickert and Schnörr, 2001a] we refer to this regularizer as a *homogeneous regularizer*, a formulation which first appeared in [Horn and Schunck, 1981] for solving the optical flow problem globally in a variational framework.

A major drawback of this choice of regularizer is its inherent capability to regularize isotropically over the domain of interest. However, in the case of practical image registration applications, deformations of soft tissue organs during breathing or cardiac motion are investigated. Since these deformations never occur smoothly across the whole body, this regularization model is far from being realistic. For example the reader may think of the diaphragm and rib cage muscles which exert forces on the lung due to contraction and expansion. As a result the lung moves independent of the rest of the thorax due to the pleural interface that connects lung and rib cage. Fortunately the bordering structures between lungs and rib cage can easily be depicted on `CT` images due to a large intensity gradient occurring at this anatomical border. Consequently we propose to use an anisotropic regularization of the deformation field, which specifically takes large intensity gradients into account. According to the taxonomy of [Weickert and Schnörr, 2001a] again, we define the *anisotropic image-driven regularization* as

$$\mathcal{S}[\varphi] := \mathcal{S}[\varphi]^{AI} = \iiint_{\Omega} (\nabla u)^T D_{\nabla I_F} \nabla u + (\nabla v)^T D_{\nabla I_F} \nabla v + (\nabla w)^T D_{\nabla I_F} \nabla w \mathrm{d}\mathbf{x}, \qquad (6.5)$$

where $D_{\nabla I_F}$ is a tensor which represents how much regularization one wants in specific directions. These directions are encoded by a set of orthogonal basis vectors, which are derived from the local intensity gradient directions. Note that these considerations were pioneered by [Nagel, 1987]. The $D_{\nabla I_F}$ is often referred to as the regularized projection matrix. For the two-dimensional case (the classical optical flow formulation), this diffusion tensor might be easily derived by using the local intensity gradient and its normal vector $\nabla I_F^{\perp}$ [Alvarez et al., 2000]. In this case it is defined as

$$D_{\nabla I_F} := \frac{1}{|\nabla I_F|^2 + 2\epsilon^2} \left(\nabla I_F^{\perp} \nabla I_F^{\perp T} + \epsilon^2 \mathbf{I} \right),$$

where ϵ denotes a small constant value to prevent division by zero which occurs in completely homogeneous regions.

In the three-dimensional case, the derivation of the diffusion tensor has to be done differently, since the normal to the local intensity gradient is not well-defined, i.e. there exists an infinite amount of normals for a given three-dimensional vector. To overcome this problem the structure tensor [Jaehne, 1993] might be used, which is a different representation of the local intensity gradients. We have already come across the structure tensor in Section 5.1 where its eigenvalues were used to determine corner-like situations. Here we use the structure tensor $S_{\nabla I_F}$ to define the local directions where the regularization should be allowed/restricted. Therefore we employ a model suggested by [Scharr and Uttenweiler, 2001]. They derived the diffusion tensor from the structure tensor by performing an eigen-decomposition of $S_{\nabla I_F}$

$$S_{\nabla I_F} = \mathbf{U} \Phi \mathbf{U}^T.$$

The columns of $\mathbf{U}$ hold the eigenvectors of the structure tensor, while Φ is the diagonal matrix of its eigenvalues ϕ_i. The eigenvalues give a measure of the degree of contrast change along the respective gradient direction represented by the eigenvectors. Therefore to transform the structure tensor to a diffusion tensor, it is obvious that large eigenvalues have to be mapped to small ones, while the eigenvectors remain unchanged. In [Scharr and Uttenweiler, 2001] the following mapping function is proposed

$$\Lambda_{ii} := \lambda_i = \begin{cases} 1 - (1-c)\exp\left(\frac{-d}{(\phi_i - \sigma)^2}\right) & \phi_i \geq \sigma \\ 1 & \text{else} \end{cases}$$

We adopt this definition for our derivation of the anisotropic regularization scheme. c is a constant that prevents this function from becoming equal to one, d represents the distribution of eigenvalues and depends on the intensity range of the input images, ϕ_i is the eigenvalue of the structure tensor and σ is a parameter that represents the noise level of the input images. With this definition the diffusion tensor can be constructed

$$D_{\nabla I_F} = \mathbf{U}\Lambda\mathbf{U}^T.$$

and used for the anisotropic image-driven regularization term.

6.3 Hybrid Landmark and Intensity Registration

A registration approach that is solely depending on landmark correspondences, like the approach that was presented in Chapter 5, works well as long as an evenly distributed set of correspondences can be found. To avoid this shortcoming and to properly combine the advantages of landmark- and intensity-based registration such that both approaches concurrently contribute to the optimal solution, we propose to use the sparse set of matched landmarks as an additional regularization constraint $\mathcal{Q}$ in the general nonlinear registration functional that was presented above. This leads to the continuous cost functional (for weights $\alpha > 0, \beta > 0$)

$$\mathcal{J}[\varphi]_{hyb} := \mathcal{D}[I_F, I_M; \varphi] + \alpha\mathcal{S}[\varphi] + \beta\mathcal{Q}[\varphi] = min. \tag{6.6}$$

We define the $L_2 - norm$ optical flow data term as

$$\mathcal{D}[I_F, I_M; \varphi] := \iiint_{\Omega} (1 - W(\mathbf{x}))(I_F(\mathbf{x}) - I_M(\varphi(\mathbf{x})))^2 \mathrm{d}\mathbf{x}, \tag{6.7}$$

where the factor $(1 - W(\mathbf{x}))$ penalizes the data term near landmark correspondences. The weight factor $W(\mathbf{x})$ can be imagined as a binary mask image that is set to 1 at locations where a landmark match has been identified and 0 otherwise. Now for those locations where we have a landmark match, we decide to reduce the influence of the `SSD` data term by multiplying $(1 - W(\mathbf{x}))$. This way the feature- and intensity-information is fused. Note that the coupling to other locations in the volume is performed via the regularization term, which propagates landmark-matching information to the neighbors of the matches. In practice we also applied the matching uncertainties from the landmark matching stage into $W(\mathbf{x})$).

Anisotropic image-driven regularization is given as $\mathcal{S}[\varphi]^{AI}$, where according to the previous section $D_{\nabla I_F}$ is the diffusion tensor that creates a dependency between the estimated displacement field and the image boundaries of the fixed image I_F such that regularization across image boundaries is penalized. Finally, with $\varphi = (u(\mathbf{x}), v(\mathbf{x}), w(\mathbf{x}))^T$ the landmark matching constraint is

$$\mathcal{Q}[\varphi] := \iiint\limits_{\Omega} W(\mathbf{x}) \left\| (u(\mathbf{x}), v(\mathbf{x}), w(\mathbf{x}))^T - (u_q(\mathbf{x}), v_q(\mathbf{x}), w_q(\mathbf{x}))^T \right\|^2 \mathrm{d}\mathbf{x}. \tag{6.8}$$

$W(\mathbf{x})$ is the same weighting function incorporating landmark correspondences and its matching uncertainties as given above and $(u_q(\mathbf{x}), v_q(\mathbf{x}), w_q(\mathbf{x}))^T$ is the sparse displacement field derived from the correspondences of the automatic landmark matching step. Note that the additional landmark matching regularization term is independent of the explicit choice of data or regularization terms.

6.4 Numerical Aspects of the Minimization of Energy Functionals

The calculus of variations gives solutions for the minimization of energy functionals like Equations 6.1 or 6.6 by the notion of the Gâteaux derivatives of a functional, which have to be equal to 0 similar to the case of minimizing functions [Lanczos, 1986]. Setting the functional derivatives $\delta\mathcal{J}_{hyb}$ from Equation 6.6 equal to 0 results in a coupled Euler-Lagrange partial differential equation (`PDE`) system. The solution of the `PDE` system gives the stationary value, i.e. the desired deformation field φ.

The `PDE` system derived from the Gâteaux derivatives reads

$$\begin{aligned}
&1. \quad (1-W)\left(I_F(\mathbf{x}) - I_M\left(\varphi\left(\mathbf{x}\right)\right)\right)\frac{\partial I_M}{\partial x} + \alpha \mathrm{div}\left(D_{\nabla I_F}\nabla u\right) - \beta W\left(u - u_q\right) = 0\\
&2. \quad (1-W)\left(I_F(\mathbf{x}) - I_M\left(\varphi\left(\mathbf{x}\right)\right)\right)\frac{\partial I_M}{\partial y} + \alpha \mathrm{div}\left(D_{\nabla I_F}\nabla v\right) - \beta W\left(v - v_q\right) = 0\\
&3. \quad (1-W)\left(I_F(\mathbf{x}) - I_M\left(\varphi\left(\mathbf{x}\right)\right)\right)\frac{\partial I_M}{\partial z} + \alpha \mathrm{div}\left(D_{\nabla I_F}\nabla w\right) - \beta W\left(w - w_q\right) = 0
\end{aligned} \tag{6.9}$$

This boundary value problem assumes a Neumann boundary condition. An explicit scheme could be derived for this `PDE` system by introducing an artificial time dependency in u, v, w and setting the temporal derivatives equal to 0. This would lead to a gradient-descent minimization scheme with an explicit time-step. However, we decided to use a more robust semi-implicit scheme for solving the Euler-Lagrange equations. It makes use of a first-order Taylor approximation of the nonlinear expression $I_M(\varphi(\mathbf{x}))$. After discretization using standard finite difference stencils [Strikwerda, 1989] a huge but sparse system of linear equations has to be solved. The $3N$ unknowns of this equation system are the deformation field coordinates $(u(\mathbf{x}), v(\mathbf{x}), w(\mathbf{x}))^T$ for each voxel (supposing a total of N voxels). To solve this linear equation system we make use of a Gauss-Seidel scheme [Press et al., 1992]. Gauss-Seidel is an instance of relaxation methods for solving huge but sparse linear systems, working with *in-place* updates, thus removing the need for storing multiple displacement fields to calculate the iterative solution. The whole registration algorithm is performed in a *multi-resolution* manner to further speed-up computation, to avoid local minima and to be able to handle large displacements. The implementation of the multi-resolution coarse-to-fine strategy uses a Gaussian pyramid, where the first-order derivative stencils are taken from [Farid and Simoncelli, 1997].

We now go into the details of the Taylor approximation of $I_M(\varphi(\mathbf{x}))$ and the choice of the finite difference stencils for a single resolution level. The first-order Taylor approximation introduces the unknown displacement field updates $\Delta u, \Delta v, \Delta w$ which have to be solved for in the iterative relaxation step. However, since we also linearize the nonlinearity $I_M(\varphi(\mathbf{x}))$, it is also necessary to establish an outer iteration loop for the displacement field components. The relation between inner and outer loop displacements is

$$\Delta u = u^i - u^{i-1}, \Delta v = v^i - v^{i-1}, \Delta w = w^i - w^{i-1}.$$

Inserting these expressions into Equation 6.9 and putting unknowns on one side of the equations, one gets the following `PDE` system in the unknowns $\Delta u, \Delta v, \Delta w$:

$$\begin{aligned}
1.\quad (1-W)\left(\Delta u \frac{\partial I_M}{\partial x} + \Delta v \frac{\partial I_M}{\partial y} + \Delta w \frac{\partial I_M}{\partial z}\right) \frac{\partial I_M}{\partial x} - \alpha D_{\nabla I_F} \mathcal{L}(\Delta u) + \beta W \Delta u \quad &= \\
(1-W)(I_F - I_M) \frac{\partial I_M}{\partial x} + \alpha D_{\nabla I_F} \mathcal{L}\left(u^{i-1}\right) - \beta W \left(u^{i-1} - u_q\right) & \\
2.\quad (1-W)\left(\Delta u \frac{\partial I_M}{\partial x} + \Delta v \frac{\partial I_M}{\partial y} + \Delta w \frac{\partial I_M}{\partial z}\right) \frac{\partial I_M}{\partial y} - \alpha D_{\nabla I_F} \mathcal{L}(\Delta v) + \beta W \Delta v \quad &= \\
(1-W)(I_F - I_M) \frac{\partial I_M}{\partial y} + \alpha D_{\nabla I_F} \mathcal{L}\left(v^{i-1}\right) - \beta W \left(v^{i-1} - v_q\right) & \\
3.\quad (1-W)\left(\Delta u \frac{\partial I_M}{\partial x} + \Delta v \frac{\partial I_M}{\partial y} + \Delta w \frac{\partial I_M}{\partial z}\right) \frac{\partial I_M}{\partial z} - \alpha D_{\nabla I_F} \mathcal{L}(\Delta w) + \beta W \Delta w \quad &= \\
(1-W)(I_F - I_M) \frac{\partial I_M}{\partial z} + \alpha D_{\nabla I_F} \mathcal{L}\left(w^{i-1}\right) - \beta W \left(w^{i-1} - w_q\right) & \qquad (6.10)
\end{aligned}$$

The divergence operator is discretized using standard stencils for second derivatives symbolized as linear operator $\mathcal{L}$. Therefore, the diffusion tensor $D_{\nabla I_F}$, which may be regarded as the following symmetric matrix

$$D_{\nabla I_F} = \begin{pmatrix} d_1 & d_4 & d_5 \\ d_4 & d_2 & d_6 \\ d_5 & d_6 & d_3 \end{pmatrix}$$

has to be multiplied with the divergence of the gradient vectors. For the left hand side term of the first equation this leads to the expression

$$\begin{aligned} D_{\nabla I_F}\mathcal{L}(\Delta u) = d_1 \left(\Delta u\right)_{xx} + d_2 \left(\Delta u\right)_{yy} + d_3 \left(\Delta u\right)_{zz} + \\ 2d_4 \left(\Delta u\right)_{xy} + 2d_5 \left(\Delta u\right)_{xz} + 2d_6 \left(\Delta u\right)_{yz} \end{aligned}$$

Equivalent expressions follow for the right hand side and the other two Euler-Lagrange equations. The three-dimensional stencils used for the discretization of the second-order derivatives $u_{xx}, u_{yy}, u_{zz}, u_{xy}, u_{xz}, u_{yz}$ are derived from the optimally rotation equivariant derivative stencils of [Farid and Simoncelli, 1997].

Now it is possible to derive the relaxation scheme from these final discretized equations. For each of the N voxels there are three discretized Euler-Lagrange equations, the total system has $3N$ unknowns, the system matrix A is sparse and block-diagonal. In relaxation schemes it is not necessary to explicitly construct the system matrix, instead one goes over all voxels of the fixed image and calculates updates to the current deformation field entry that corresponds to this voxel. The linear system $\mathbf{Ax} = \mathbf{b}$ is the starting point for the derivation of the Gauss-Seidel scheme. The iteration scheme can be expressed as

$$x^k = \frac{1}{a_{ii}} \left(b_i - \sum_{j<i} a_{ij} x_j^k - \sum_{j>i} a_{ij} x_j^{k-1} \right), \quad i = 0, 2, \cdots, 3N-1$$

where k is the iteration count, $\mathbf{a}_{ij}$ are the components of system matrix $\mathbf{A}$ and b_i the components of $\mathbf{b}$.

Another important numerical aspect is the proper choice of the regularization weight α. We decided to split α into two components, a global weight factor α_0 and a factor $\alpha_i, \cdots, i \in \{x, y, z\}$, that depends on the direction of interest. The second component α_i has two responsibilities. First, it balances the energy data term and the regularization term with respect to their measuring units. The `SSD` data term subtracts intensities and squares them, i.e. its measuring unit is $[HU^2]$. The regularization term works on squared displacement field gradients, i.e. its measuring unit is $[mm^2]$. Consequently the measuring unit of α is equal to $\frac{[HU^2]}{[mm^2]}$. Second, α_i takes care of the anisotropic spacing that might be given in the input data sets. We calculate the α_i as

$$\alpha_i = \frac{(max(max_{\mathbf{x}}(I_F(\mathbf{x})), max_{\mathbf{x}}(I_M(\mathbf{x}))) - min(min_{\mathbf{x}}(I_F(\mathbf{x})), min_{\mathbf{x}}(I_M(\mathbf{x}))))^2}{\text{spacing}_i}$$

Now that the anisotropic spacing and the mismatch in measuring units have been taken care of it is possible to choose α_0 as a global,abstract weighting parameter that takes values in the range of 0.1 to 0.001.

6.5 Evaluation Experiments & Results

To assess the validity of the hybrid registration approach qualitative and quantitative evaluations were performed on synthetically transformed and clinical thorax `CT` data sets showing breathing motion. All experiments were executed on a standard Windows notebook computer with 2GHz and 2GB RAM.

6.5.1 Synthetic Deformations

Synthetic experiments were performed on four test data sets (A,B,C,D) taken at inspiration, each of them having a volume size of $256 \times 256 \times 256$ voxels. The applied transformation intends to model a simple displacement field similar to exhalation. A synthetic nonlinear transformation φ that simulates diaphragm and rib cage movement can be found in B.6. φ is defined using two parameters $t_{vertical}$ and t_{inward}. The four test data sets were synthetically transformed with a small ($t_{vertical}$ = 25mm, t_{inward} = 10mm) and a large ($t_{vertical}$ = 55mm, t_{inward} = 25mm) deformation. Figure 6.1 a)-c) shows the effects of these deformations on data set A. Table 6.1 gives the results of these experiments by comparing the hybrid approach ("*hyb*") with the landmark-based approach ("*lm*") from Section 6.1 and the optical flow intensity registration ("*of*") from Section 6.2. All comparisons are always performed only on those regions which are present in both registered data sets. The `RMS` of the intensity differences INT_{rms} before ("*init*") and after registration are calculated, as well as the difference of the resulting and the synthetic displacement fields in terms of the `RMS` of the whole displacement vector field DF_{rms} and the maximum of the displacement difference vector components DF_{max}. Qualitative results are shown in terms of difference images of data set D in Figure 6.3 a)-c).

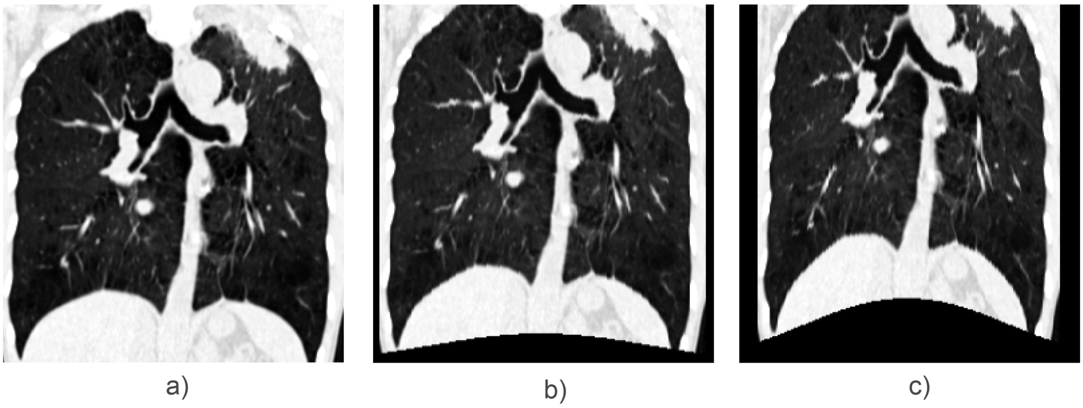


Figure 6.1: Simulated breathing deformation. a) original data set A, b) the small and c) the large deformation.

			Simulated Breathing 25-10				**Simulated Breathing 55-25**			
Measure			A	B	C	D	A	B	C	D
INT_{rms}	*init*	[HU]	373.85	313.70	291.08	348.11	539.98	468.57	438.07	513.42
	of	[HU]	77.045	75.084	64.972	75.529	123.43	106.01	102.27	116.022
	lm	[HU]	61.888	61.891	55.351	69.595	149.69	137.94	105.22	178.82
	hyb	[HU]	57.719	45.187	48.769	44.441	111.47	103.32	101.64	108.74
DF_{rms}	*of*	[mm]	4.4201	6.0393	4.1328	4.6852	9.4332	12.428	8.3835	10.039
	lm	[mm]	1.2700	1.6068	1.2759	1.3829	4.3041	4.7178	3.9078	4.6794
	hyb	[mm]	1.6826	2.4241	1.7129	2.1525	4.7948	5.7478	5.0956	5.7385
DF_{max}	*of*	[mm]	19.926	22.976	16.925	18.366	40.826	48.667	40.868	43.429
	lm	[mm]	22.712	21.981	21.586	22.219	49.008	50.426	47.969	49.375
	hyb	[mm]	17.488	19.770	16.489	17.909	35.090	36.737	32.985	41.322

Table 6.1: Simulated breathing transformation. Registration results in terms of the RMS of the intensity differences INT_{rms}, the RMS of the displacement field DF_{rms} and the maximum of the displacement difference vector components DF_{max}.

6.5.2 Clinical Data

The algorithm was also evaluated on four clinical thoracic data sets (A,B,C,D), each of them consisting of two scans at different breathing states. The data sets show different problem characteristics. Data sets A and D differ by small breathing deformations. Data sets B and C differ by large breathing deformations and also show intensity variations due to diseases making them very hard to register. For the clinical data no gold standard displacement was available for comparison, therefore solely the decrease in the `RMS` of the intensity differences before and after registration are calculated and presented in Figure 6.2. For qualitative results difference images are shown in Figure 6.3 d)-e) for data sets C and D.

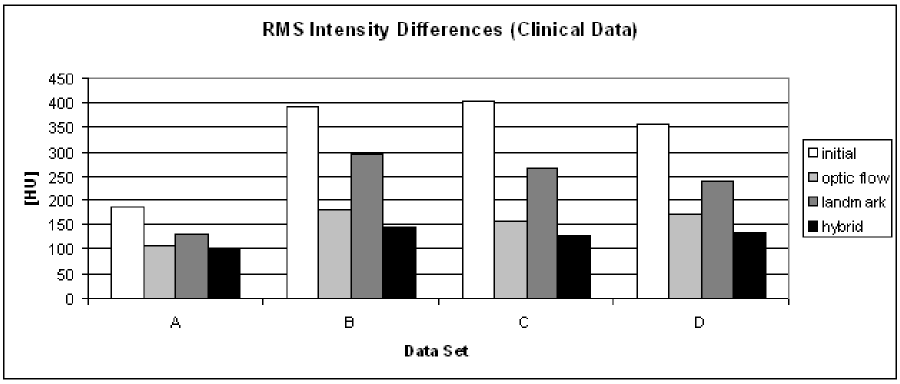


Figure 6.2: Result chart showing the RMS of the intensity differences on the clinical data.

6.6 Discussion & Conclusion

The quantitative evaluation on the simulated data shows that the presented hybrid approach is an effective way to combine the advantages of landmark- and intensity-based registration. The `RMS` of the intensity differences on the small deformation data is slightly better than the individual methods, however all three methods are able to register these deformations very well. The displacement field outliers (DF_{max}) are also improved while the `RMS` of the displacement field (DF_{rms}) is slightly worse due to the influence of the intensity registration. However with an absolute value of DF_{rms} lying around $2mm$ this is already a very accurate result. The large deformation data is more demanding to register. Here the landmark-based approach has problems generating a dense set of correspondences. On the other hand, due to the large imposed deformation, the intensity-based registration has problems to align small vascular structures. With the hybrid approach the best of both approaches is achieved which can be seen in the decrease of INT_{rms} and on the qualitative results in Figure 6.3 a)-c). The `RMS` of the displacement fields again decrease to a similar level as in the landmark registration and the displacement field outliers are reduced. The large absolute values of the displacement field outliers are due to problems at the border of the data sets. The clinical data also shows good correspondence after registration. However, registration of the vascular structures still can be improved. An important thing to consider is the proper choice of the algorithm parameters in the hybrid approach. Careful choice of parameters λ and μ are crucial. In our experiments, after appropriate pre-normalization of data and regularization terms, suitable values were determined empirically in a set of initial experiments and remained fixed during the evaluations. In our case we chose $\lambda = 0.05$ and $\mu = 2$. For solving the Euler-Lagrange equations the linearizing Taylor approximation leads to the choice of a number of outer iterations. A trade-off of computation time and accuracy leads to more iterations on coarser resolu-

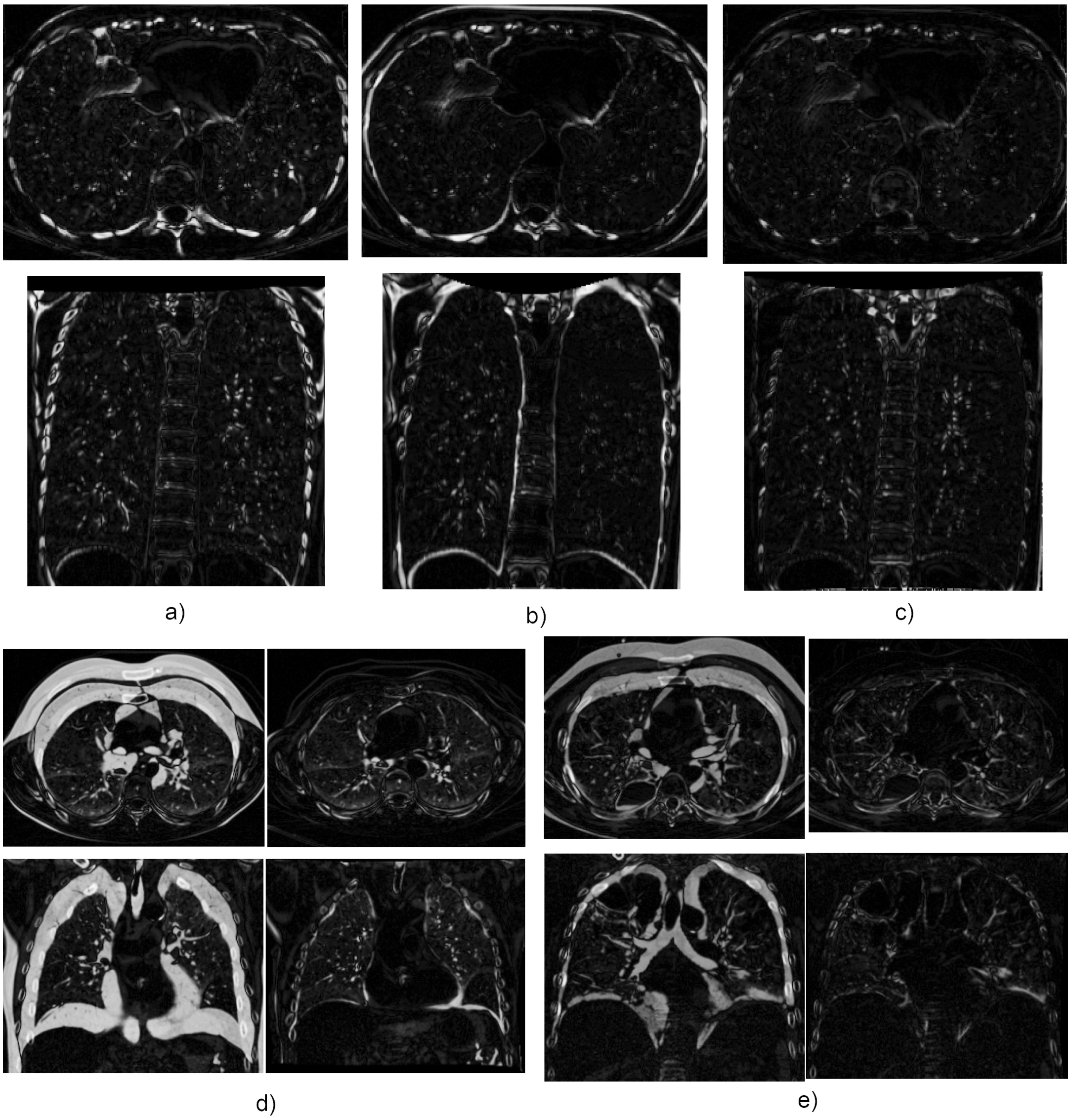


Figure 6.3: Selected Results. a)-c) Simulated Breathing results for the large deformation on data set D. a) shows difference images after optical flow, b) after landmark-based and c) after hybrid registration. d) depicts difference images of clinical data set C before and after hybrid registration, e) shows clinical data set D before and after hybrid registration. (Image contrast was enhanced to improve visibility.)

tions compared to the finer ones. For the Gauss-Seidel stage three internal iterations were chosen. The whole algorithm obviously has the drawback that the computation time is rather high with most of the time spent in solving the `PDE`.

In this chapter a hybrid feature- and intensity-based nonlinear registration algorithm was presented that consists of a fully automatic feature matching stage that acts as an additional regularization constraint on a generic intensity-based nonlinear registration method formulated in a variational framework. An evaluation of the algorithm on four synthetic and four real data sets was performed and the method was also compared to the feature- and intensity-based algorithms alone. Results show that the hybrid approach outperforms the single approaches, while its runtime is similar to the intensity-based approach.

Future work will investigate methods to speed-up the time-consuming steps of the algorithm. Another interesting perspective is to find suitable inter-modality similarity terms to derive a multi-modality nonlinear registration scheme. Finally an automatic way to determine the regularization weighting parameters will be of great interest.

Chapter 7

Comparing Nonlinear Registration Methods on Thoracic CT Images

Contents

In this chapter we describe our evaluation framework which is consequently used to compare a number of nonlinear registration algorithms applied to thoracic `CT` images. Evaluation of nonlinear registration algorithms is a hard task, since there generally is no ground-truth data to compare to. In contrast to validation we only refer to evaluation, since our definition of validation is the measurement of the "absolute" performance of an algorithm. However, all quantitative measures are inherently influenced by noise, partial volume effect, limited numerical precision, interpolation schemes, etc. [Hellier et al., 2003] Therefore we restrict ourselves to measuring "relative" algorithm performance for evaluation. Our evaluation consists of two parts, one is working with synthetically transformed data, while the second part uses clinical data that includes a number of real-world problems like small to large breathing differences, contrast agents present in one of the images or cases with lung diseases. Only the synthetic experiments can be assessed using ground-truth data, the clinical data is evaluated by some representative quantitative measures based on the comparison of intensities and by visual inspection.

The choice of synthetic deformations is crucial in an evaluation procedure. If the chosen transformation is too simple or restrictive, then the derived quantitative measures are restricted to a very coarse approximation of the originally investigated problem. This fact especially complicates the task of nonlinear registration. One of the few publications that use a complex, physically plausible model of synthetic deformation is [Schnabel et al., 2003]. The synthetic transformation is generated by a bio-mechanical tissue deformation based on the Finite-Element method. As a consequence, deformations from the registration can be compared to plausible ground-truth deformation data. However, the drawback of this evaluation method is that it is very complicated to implement, since it consists of a segmentation step that has to be very accurate, a meshing step to create the Finite-Element nodes and the solution of the FEM equations, where in [Schnabel et al., 2003] commercial products were used for all of these stages. Due to this complicated (and expensive) set-up we decided not to investigate this direction further, however, we keep it in mind as a direction for future work.

Our approach is to use a combination of simpler synthetic transformation methods to derive quantitative measures on the performance of several nonlinear registration algorithms. Up to our knowledge we are the first to provide such an extensive comparison of state-of-the-art registration methods from the literature. Especially in the area of registration of thoracic data in the presence of breathing deformations, there are no reports on the comparison of algorithms. We also compare our proposed algorithms from the previous chapters in this evaluation procedure. We assume that the calculation of many different synthetic deformations, with each of them testing different behavior, along with a large number of different evaluation measures leads to a method for thorough evaluation and comparison.

7.1 Evaluation Setup

An evaluation framework has been built to assist in the quantitative comparison of different algorithms over a variety of synthetic transformations and for clinical-data evaluation. This framework uses Python as high-level scripting language to call the C++ methods that perform the synthetic transformation calculation, the registration and the computation of the quantitative evaluation measures. See [Kluckner, 2006] for further details. We have selected three different images from human and sheep data sets as the input to the synthetic part of the evaluation framework. These images are synthetically transformed with the help of a configuration file. A list of nonlinear registration algorithms along with its parameterizations may be specified to perform the computations. Finally the synthetic ground-truth deformation fields and the original moving images are compared with the outcomes of the registration algorithms and log files with the quantitative evaluation results are written. Note, that all of these processes are performed in an easily customizable fully automatic fashion. Another interesting aspect this evaluation framework allows, is to automatically test algorithms with different parameterizations to determine optimal parameters with respect to the evaluation measures. We did not yet look further into this capability of the framework, but we intend to do that in future. Up to now we chose the parameters for the algorithms under investigation by performing some tests on representative data sets and held them fixed after finding suitable choices.

For the clinical part we selected eight different thoracic data sets, which are (similar to the synthetic experiments) registered and evaluated in an automated fashion. However, since there is no ground-truth data on the deformation fields available we restrict the evaluation to image-related comparison measures in this case. The following subsections will present more details on the parts of the evaluation framework.

All of the evaluations were performed on AMD Opteron dual-processor machines running a 64-bit Linux with 8 GB of memory.

7.1.1 Input Images

We selected three different images of size 256 × 256 × 256 for our synthetic evaluation experiments. The first image is a human thorax scan from the free NLM data collection (*nlm*) [NLM, 2006]. The second one is a sheep data set at inspiration which was provided by Prof. Eric Hoffman from the University of Iowa (*sheep*). The third one is a human lung data set at inspiration (*human*). This data set was provided by Siemens MED CT, Forchheim, Germany. The scans were performed using 64-slice `MSCT`, the x-, y- and z-resolution range of 0.5-0.6 mm lead to nearly isotropic original data set sizes of around 512×512×400 slices. All of these data sets were resampled to 256 × 256 × 256 voxels from their original size using the windowed-sinc resampling approach from [Meijering et al., 2001]. Since we were only interested in nonlinear registration we assume all our input data sets to be rigidly pre-registered. To assess algorithm behavior on larger data sets, we decided to use one data set with a size of 512×512×512 voxels during our evaluations. This data set was resampled from the *sheep* data set with its original resolution of 512×512×590 voxels. We resampled the z-dimension to simplify algorithm setup. Figure 7.1 shows example slices of the used data sets.

For the evaluation of the clinical data we used seven different data sets (A,B,C,D,E,F,G), which stem from human thorax scans provided by Siemens MED CT, Forchheim, Germany. The specific characteristics of these data sets are given in Table 7.1. For our evaluations the data sets were resampled to 256 × 256 × 256 voxels to simplify algorithm setup. Again the resampling approach from [Meijering et al., 2001] was used. We depicted sample slices from data sets B and C in Figure 7.2 and from data sets E and G in Figure 7.3. Note that we are using a color coded representation to visualize *intensity differences* of corresponding locations. This color code is inspired

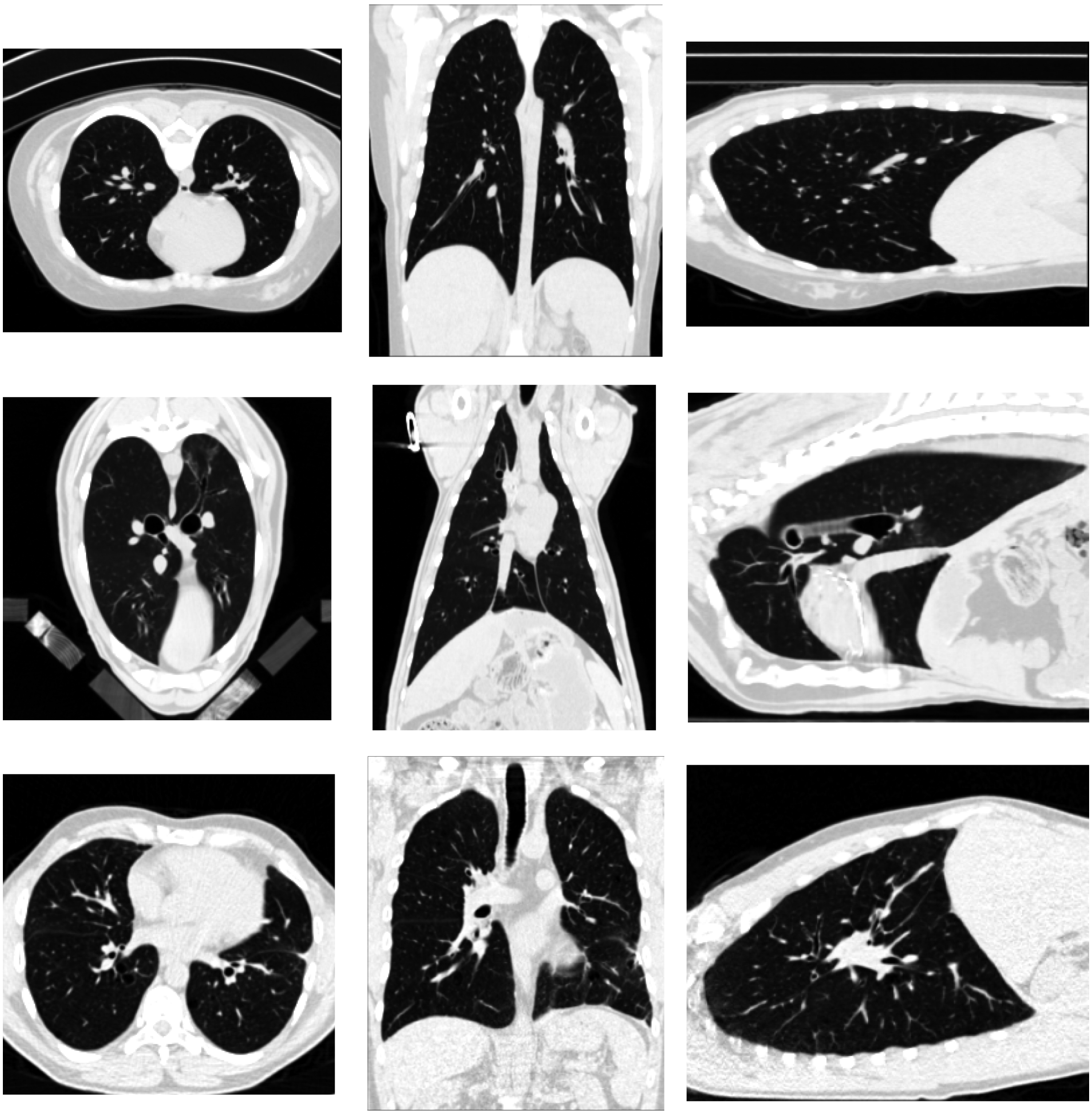


Figure 7.1: Input images used for the synthetic deformation experiments. First row shows data set *nlm*, second row *sheep*, third row *human*. First column shows axial, second column coronal, and third column sagittal view plane.

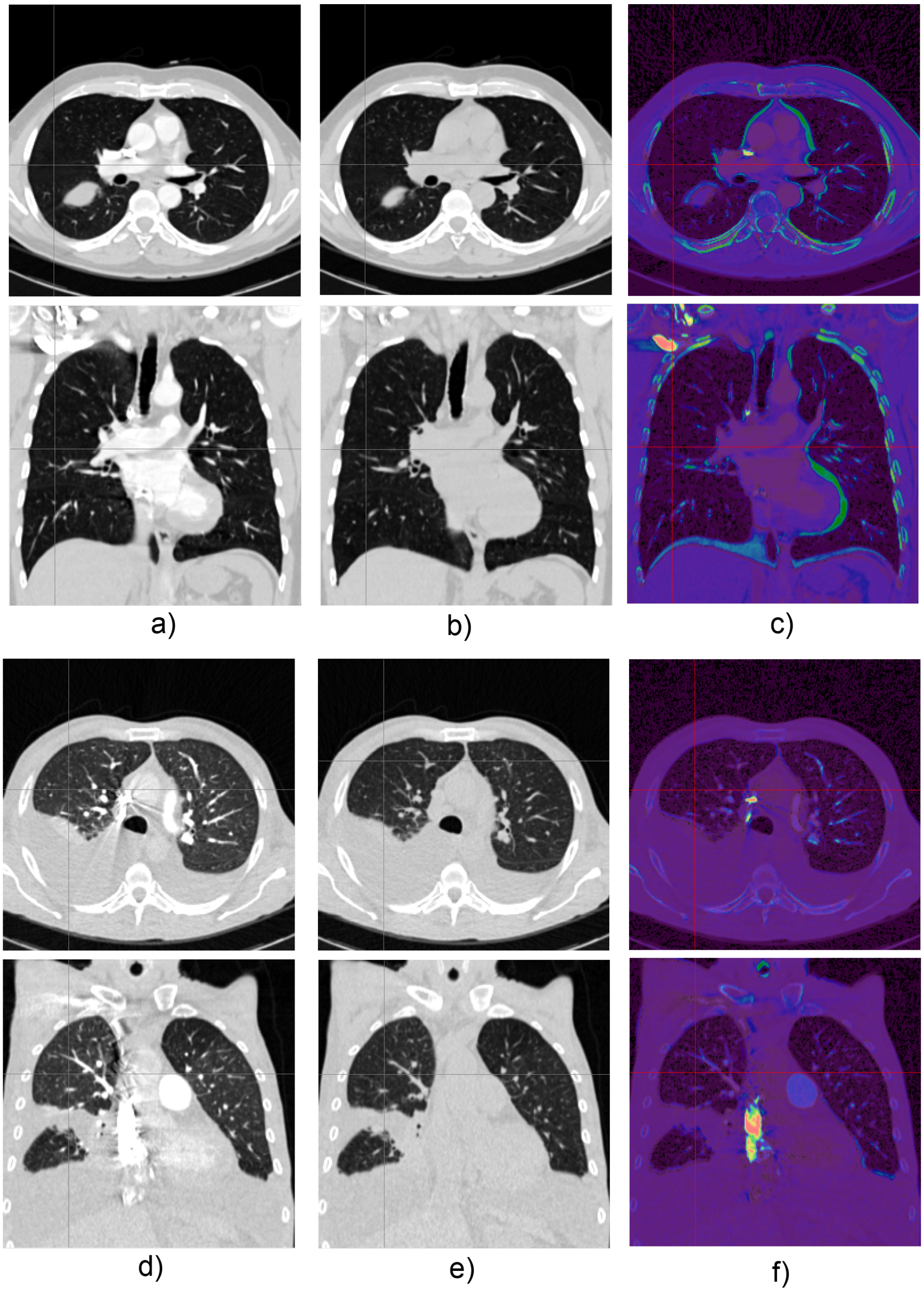


Figure 7.2: Sample slices from data sets B (top two rows) and C (bottom two rows). Axial and coronal slices respectively. a),d) scan with contrast agent. b),e) native scan. c),f) shows the differences using a color code representation.

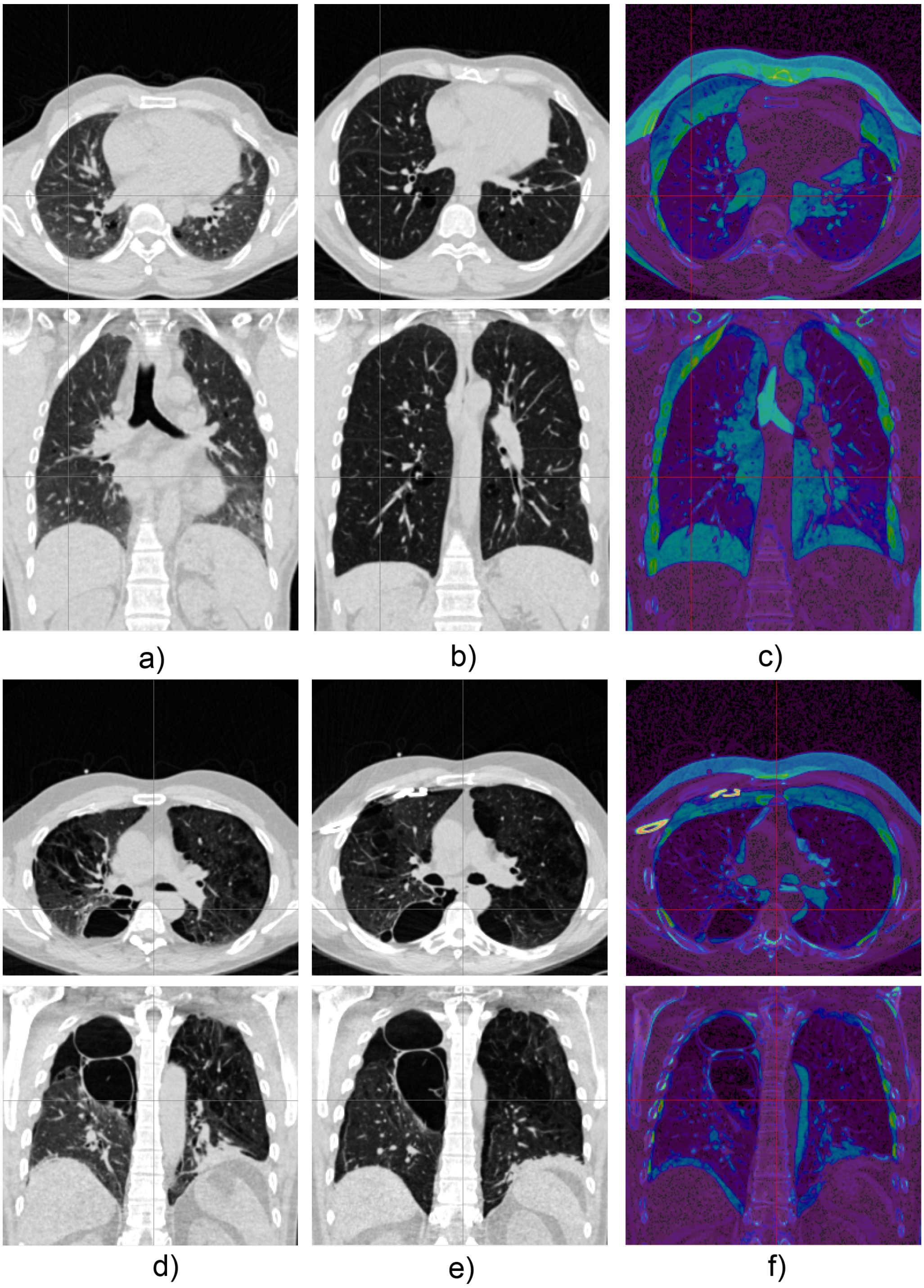


Figure 7.3: Sample slices from data sets E (top two rows) and G (bottom two rows). Axial and coronal slices respectively. a),d) exhalation scan. b),e) inhalation scan. c),f) shows the differences using a color code representation. Note the severe lung disease in data set G.

DataSet	orig. size	Problem Characteristics
A	512,512,306	native & contrast-enhanced, small breath. deformation
B	512,512,329	native & contrast-enhanced, medium breath. deformation
C	512,512,307	native & contrast-enhanced, small breath. def., lung disease
D	512,512,316	native & contr.-enh., small breath. def., lung pathology
E	512,512,438	large breath. deformation, slight lung disease
F	512,512,438	large breath. deformation, slight lung disease
G	512,512,421	medium breath. deformation, severe lung disease

Table 7.1: Nonlinear registration evaluation clinical data set characteristics.

by the electro-magnetic spectrum of visible light, i.e. it ranges from violet to red, where small intensity differences are mapped to colors near violet and large intensity differences are mapped to colors near red. Figure 7.4 shows this color-coding.

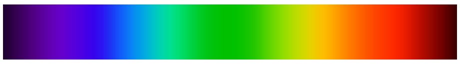


Figure 7.4: Intensity differences color coding scheme. Small differences are mapped to colors near violet, large differences mapped to colors near red of the spectrum.

7.1.2 Synthetic Transformations

We are using three different kinds of synthetic transformations. The first is a very simple simulated breathing transformation. It depends on two parameters $t_{vertical}$ and t_{inward} which provide a means for a simple approximation of a diaphragm and rib-cage movement. Further, a slight intensity variation is applied to the interior of the lung, to simulate the partial volume effect. A more detailed description of this transformation is given in Appendix B.6, we will denote it *simbr*. Examples for this transformation are given in Figure 7.5, for the experiments we use four instances of this transformation (*simbr-10-5*, *simbr-25-10*, *simbr-55-25*, *simbr-70-30*).

The second synthetic transformation consists of a regular grid with random deformations, interpolated by an Elastic Body Spline (`EBS`) [Davis et al., 1997]. There are two parameters that may be varied, the grid size of the points to be placed and the possible extent of the random deformation that is applied per grid-point (see Appendix B.7 for further details). It will be denoted *grid* from now on. The parameters of this transformation are *gridSize* and *maxDeviation*. Figure 7.6 gives some examples for this transformation, later on we will be using two instances of this transformation (*grid-32-2*, *grid-32-4*).

The third synthetic transformation is the most complex and most realistic one. We take one of our data sets (*sheep*) where we have a scan at full inhalation (`TLC`) and one at full exhalation (`FRC`). From these two data sets we derive a synthetic ground-truth deformation in the following way. From both data sets the airway trees are segmented

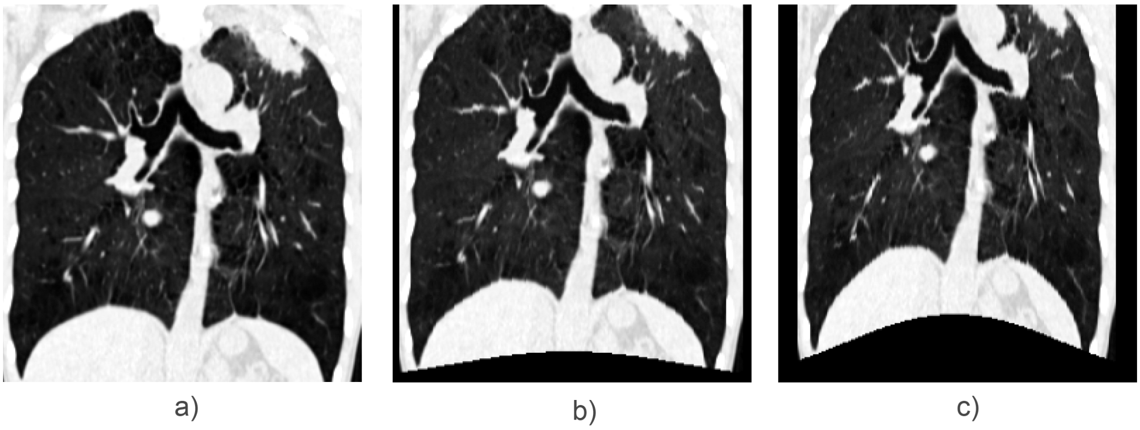


Figure 7.5: Synthetic transformation - Simulated breathing (*simbr*). a) original data set, b) effect of $t_{vertical} = 25mm$ and $t_{inward} = 10mm$ and c) effect of $t_{vertical} = 55mm$ and $t_{inward} = 25mm$.

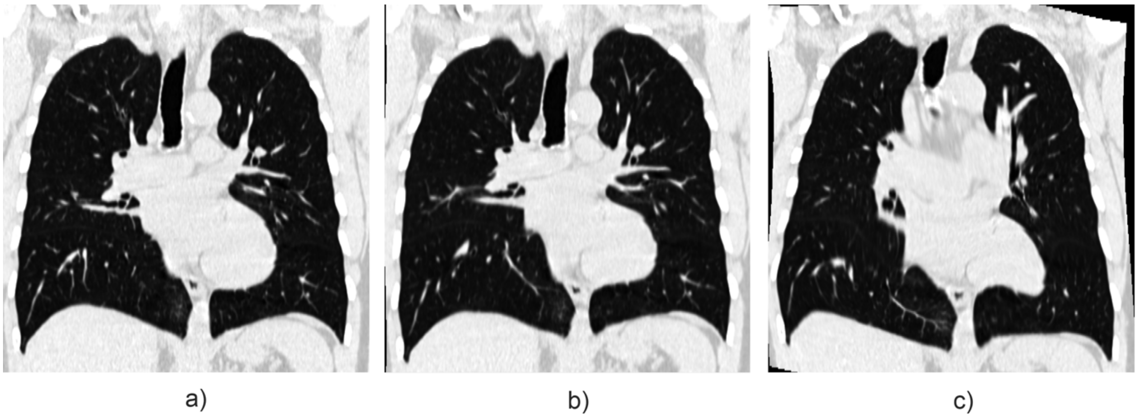


Figure 7.6: Synthetic transformation - grid based (*grid*). a) original data set, b) effect of $gridSize = 32$ and $maxDeviation = 2$ and c) effect of $gridSize = 32$ and $maxDeviation = 4$.

using the vessel segmentation approach from [Pock et al., 2005b]. These two binary segmentation masks are skeletonized and the branching points are extracted with the method of [Palagyi et al., 2003]. Afterwards we manually identify and label 20 corresponding branching points from the `TLC` and `FRC` states using a tool described in [Tschirren et al., 2005b]. We interpret these correspondences as a sparse displacement field. Finally we use an `EBS` transformation [Davis et al., 1997] to acquire a dense displacement field. This displacement field now is our ground truth deformation which represents a breathing behavior from `TLC` to `FRC` (exhalation). By applying this displacement field to the `TLC` image, we get a warped `TLC` image that resembles `FRC`. We denote this synthetic transformation as *airway*. Since this whole work-flow of producing the *airway* transformation is very tedious we only applied it to one data set, and we chose the sheep data set due to its high imaging quality, which makes the segmentation and branch-point extraction steps easier. Figure 7.7 shows the input images and the work-flow of creating this synthetic transformation. We chose the large $512 \times 512 \times 512$ sheep data set as input for the *airway* experiment, since we also are interested in the computational effort and memory consumption the different algorithms show on full-scale `CT` volume data sets.

Our basic strategy for the evaluation is in all cases identical. We take an original image, apply a synthetic transformation and store the synthetically warped image and the resulting displacement field. The displacement field will be our ground truth to compare to. Now each of the investigated nonlinear registration algorithms gets the synthetically warped image as fixed input and the original image as moving input, i.e. we try to find a displacement field that warps the original image to the synthetically transformed. In this way we can finally compare the warped image with the synthetically warped image and the ground truth displacement field with the calculated displacement field. Note that only this way it can be guaranteed that the displacement fields represent transformations in the same directions (from fixed to moving image). Compare Figure 7.8 for an overview of this strategy. This figure also clearly shows that the nonlinear registration algorithms have no insights on the working of the synthetic transformations, making their behavior completely independent from these transformations.

7.1.3 Compared Algorithms

In this section we list the nonlinear registration algorithms that are used for the comparison. We shortly describe the method, how we acquired the implementation and we state all of the important parameters that were chosen for the evaluation procedure. Note that all of these parameters remained fixed during our evaluations. The algorithms presented below are implemented in a straight-forward fashion without using parallelisms and multi-threading strategies. Further, we used standard implementations of the memory-consuming data structures like deformation fields (memory buffers of 12 32-bit floating point data types), more thorough investigation of more efficient storage data structures for the deformation fields will definitely be useful for practical applications.

Symmetric Demons Registration

The Symmetric Demons registration method is explained in Appendix B.4. We will refer to this algorithm with the abbreviation *symdemons*. We are using the publicly available

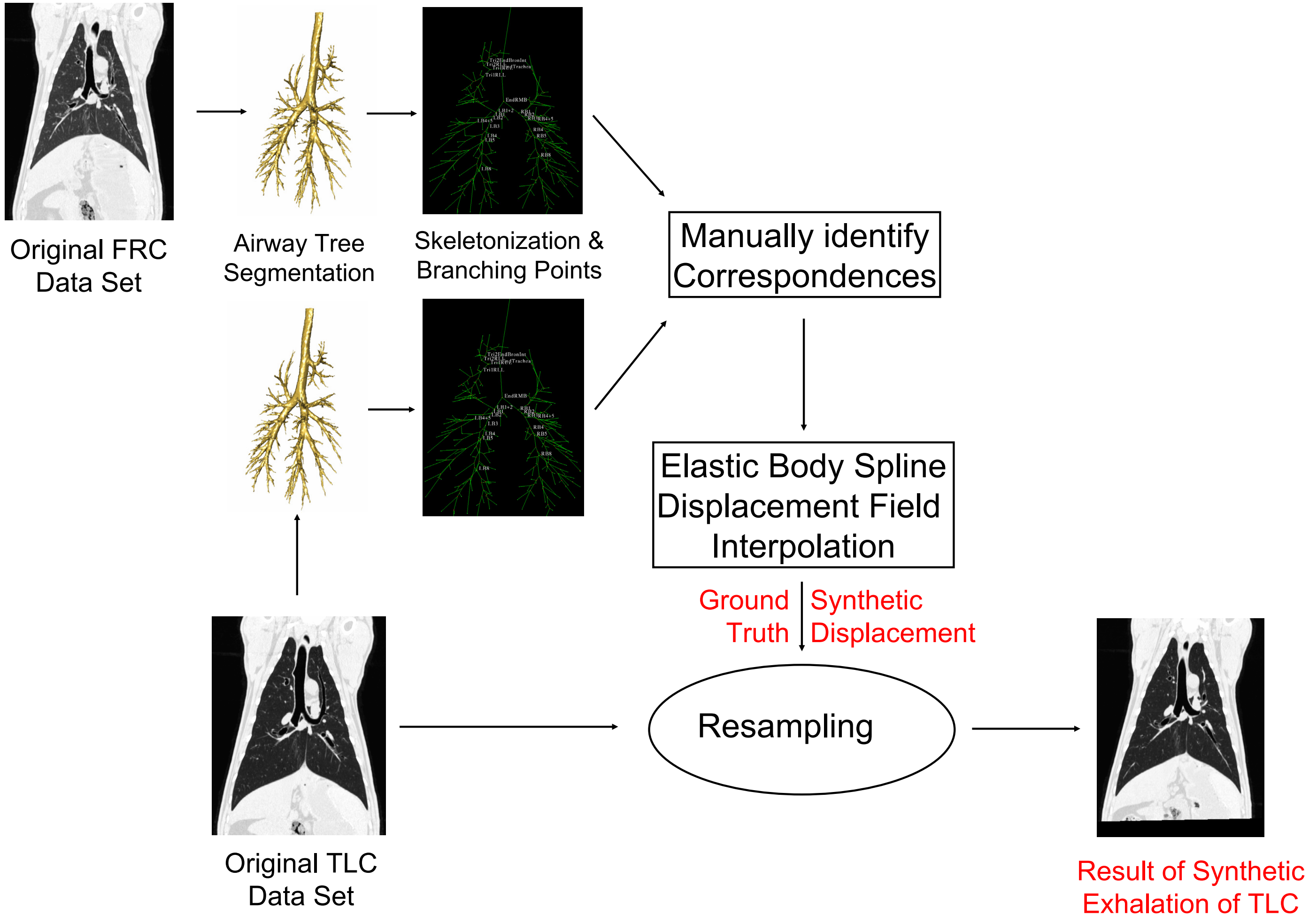


Figure 7.7: Synthetic transformation - airway matching (*airway*). This diagram shows the work-flow to derive the synthetic airway transformation. Starting from a TLC and an FRC data set, the airways are segmented and skeletonized, branch points are extracted and manually labeled. Correspondences are interpolated using elastic body spline interpolation. Finally the synthetic displacement field is used to simulate exhalation on the TLC data set.

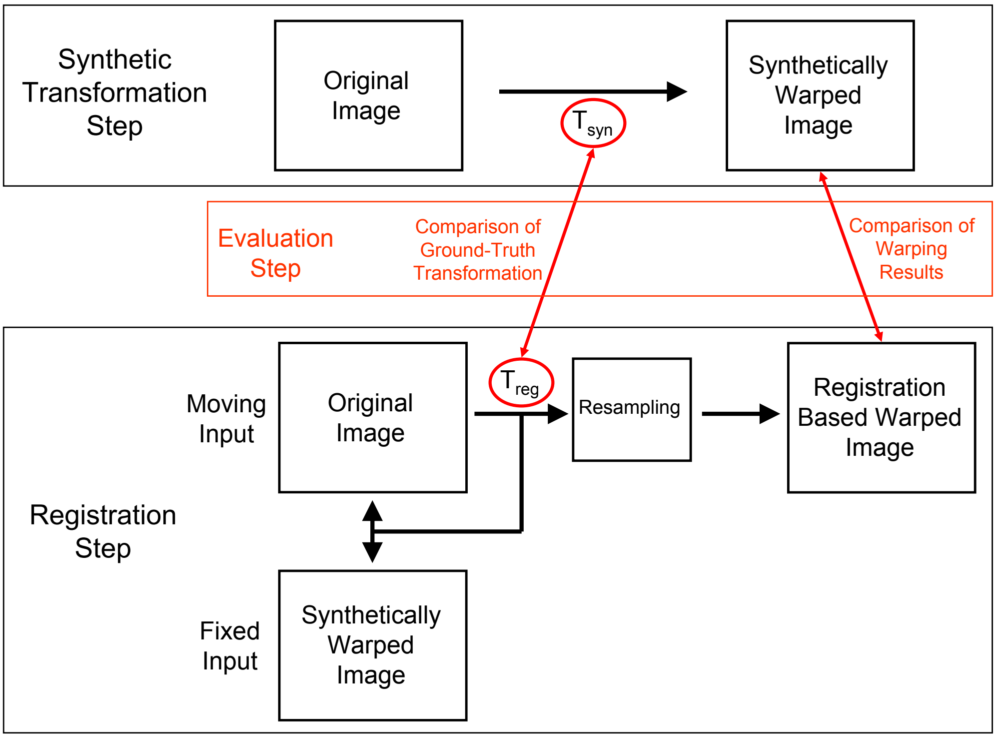


Figure 7.8: The basic setup for the synthetic transformation evaluation experiments. The synthetic transformations are applied to the original image, the nonlinear registration of original and synthetically transformed image leads to a transformation and a warped image which can be compared.

implementation of this method from the Insight Segmentation and Registration Toolkit ITK [ITK, 2006]. Symmetric Demons is very similar to the Demons algorithm with an additional increase in robustness since image gradients in both fixed and moving image are used. The necessary parameters that have to be chosen are the number of multi-resolution levels in the Gaussian pyramid, the number of iterations per level and the standard deviation of the regularization Gauss-filter. These parameters are given in Table 7.2. The total number of iterations per level is given by the expression $nrIter1 * shrinkLevel + nrIter2$ with $shrinkLevel = 0, \cdots, nrLevels$ and $shrinkLevel = 0$ resembling the original resolution in the Gaussian pyramid. This formula reflects the idea that low resolution levels use many iterations since they are cheap to compute while when going up in the multi-resolution hierarchy, a compromise has to be found between accuracy and computation speed.

Method	nrLevels	nrIter1	nrIter2	sigma [voxel]
symdemons	4	35	20	1

Table 7.2: Symmetric Demons registration evaluation parameters setup

Curvature Registration

The Curvature registration method (*curvature*) is explained in Appendix B.5. We are using the publicly available implementation of this method from the ITK [ITK, 2006]. The necessary parameters that have to be chosen are the number of multi-resolution levels in the Gaussian pyramid, the number of iterations per level, the time-step of the semi-implicit scheme τ and the regularization weight α. These parameters are given in Table 7.3. The total number of iterations per level is given by the expression $nrIter1 * shrinkLevel + nrIter2$ with $shrinkLevel = 0, \cdots, nrLevels$ and $shrinkLevel = 0$ resembling the original resolution in the Gaussian pyramid.

Method	nrLevels	nrIter1	nrIter2	τ	α
curvature	4	35	20	1	1

Table 7.3: Curvature registration evaluation parameters setup

B-spline Mutual Information

The B-spline Mutual Information registration method (*bspline*) is explained in Appendix B.2. We are using an implementation of this scheme that was taken from the ITK [ITK, 2006]. The necessary parameters of this method are the number of multi-resolution levels in the Gaussian pyramid, the number of maximum iterations per level, the number of histogram bins for `MI`, the convergence tolerance of the LBFGS optimizer and the grid spacing of the regular B-spline grid. These parameters are given in Table 7.4. The maximum number of iterations per level is given by the expression $nrIter1 * shrinkLevel + nrIter2$ with $shrinkLevel = 0, \cdots, nrLevels$ and $shrinkLevel = 0$ resembling the original resolution in the Gaussian pyramid.

Method	nrLevels	nrIter1	nrIter2	histogram bins	optimizer tolerance	gridSpacing [mm]
bspline	4	250	150	70	0.01	12.0

Table 7.4: B-spline Mutual Information registration evaluation parameters setup

Optical Flow Registration

The Optical Flow registration method (*opticflow*) is explained in Chapter 6. Although the derivation is not completely given, it can be seen as a simpler version of the landmark constrained anisotropic image-driven diffusion registration. We are using an implementation of this scheme that makes use of a Gauss-Seidel approximation of the `PDE` system that resembles the Gâteaux derivative of the energy functional. Necessary parameters of this algorithm are the number of multi-resolution levels in the Gaussian pyramid, the number of maximum iterations per level, the regularization weight α and the number of Gauss-Seidel iterations per iteration. A summary of the parameters is given in Table 7.5.

Method	nrLevels	nrIter1	nrIter2	nrGaussSeidelIters	α
opticflow	4	25	15	5	0.08

Table 7.5: Optical flow registration evaluation parameters setup

Elastic Registration

The Elastic registration method (*elastic*) is explained in Appendix B.5. We are using an implementation of this scheme that makes use of a Gauss-Seidel approximation of the `PDE` system that resembles the Gâteaux derivative of the energy functional, very similar to the optical flow registration. Since this method is a classic method from literature, we decided to apply it similarly as done in other publications (see e.g. [Modersitzki, 2004]). This means that we set the elasticity parameters λ and μ to 0 and 1 respectively, such that λ has no influence at all, and the influence of μ gets included into the global regularization weight α. The other parameters are similar to the methods stated above and summarized in Table 7.6.

Method	nrLevels	nrIter1	nrIter2	nrGaussSeidelIters	α	λ	mu
elastic	4	25	15	5	0.08	0	1

Table 7.6: Elastic registration evaluation parameters setup

Feature Based Registration

The Feature-based registration method (*feature*) is presented in detail in Chapter 5. The required parameters to set for this method are the maximum number of features to extract from each of the data sets, the corner noise threshold that determines which corner responses are removed, the matching cost threshold that resembles a hard threshold on the matching cost in the robust forward-backward matching step, some histogram discretization parameters for `SIFT` and shape context, the global `TPS` regularization weight and the number of neighbors for the locally reduced `TPS`. Table 7.7 states these parameters as they were used in our experiments.

Method	nr Corner	corner NoiseThr	match CostThr	sift HistoSize	scHisto SizeRadius	scHisto SizeOrient	tps Weight	tps Nghbrs
feature	25000	140	0.3	8	8	6	0.005	300

Table 7.7: Feature-based registration evaluation parameters setup

Hybrid Registration

The Hybrid registration method (*hybrid*) is presented in detail in Chapter 6. The required parameters to set for this method are a combination of the parameters used in the feature-based method and the optical flow method, additionally one has to define the feature constraint weight β, while α now is the global anisotropic image-driven regularization weight. Table 7.8 states these parameters as they were used in our experiments.

Method	nr Levels	nr Iter1	nr Iter2	nrGauss SeidelIters	α	β	nr Corner	corner NoiseThr	match CostThr	histo Sizes
hybrid	4	25	15	5	0.08	1.0	25000	140	0.3	(8,8,6)

Table 7.8: Hybrid registration evaluation parameters setup

7.1.4 Quantitative Evaluation Measures

Two kinds of measures are used in our evaluation framework. The first one is used to compare displacement field ground truth data with displacement field results calculated during the registration. The second kind of measures compares two sets of images, before and after registration. This comparison step is performed on fixed and moving input image and on fixed and warped moving image. After successful registration, warping the moving image should result in an increase of similarity with the fixed image. We also give the computation time and the peak memory effort for each of the compared algorithms.

We always compute the evaluation measures only on the overlapping regions of the input image data sets and the deformation fields. This is especially important in the case of synthetic transformations, where certain areas might vanish due to a shrinking like behavior, we always mark these regions with special values during synthetic transformation and we omit those marked regions from the evaluation process.

We have to remark here that all of these measures possess one inherent shortcoming. Data sets with a size of 256 × 256 × 256 consist of a total around 16 million voxels and each measure of interest will try to reduce this huge amount of information to one single number. Although some of our measures are inspired by robust statistics, their information value has to be at least discussed and also questioned. One should always keep this in mind when looking at quantitative evaluation results.

7.1.4.1 Displacement Field Measures

To assess the similarity of a synthetic ground-truth deformation field φ_{syn} and a deformation field computed by a nonlinear registration algorithm φ_{reg}, we use:

Root-Mean-Square of Displacement Field RMS_{disp} The `RMS` of the displacement field interprets the two displacement fields of interest as feature vectors of dimension $3 \times N_1 \times N_2 \times N_3$, where N_i is the number of voxels on the input image grid in the according dimension. The measure reads

$$RMS_{disp} := \sqrt{\frac{1}{3N_1N_2N_3} \sum_{\mathbf{x}\in\Omega} \sum_{i=1}^{3} \left(\varphi_{syn}(\mathbf{x}_i) - \varphi_{reg}(\mathbf{x}_i)\right)^2}$$

Median Absolute Deviation of Displacement Field MAD_{disp} This measure is similar to the `RMS`, however it is a more robust variant in the presence of outliers. The median absolute deviation is defined as

$$MAD_{disp} := \text{Median}\left(|m_i - \text{Median}\left(m_i\right)|\right)$$

with $m_i = |\varphi_{syn}(\mathbf{x}_i) - \varphi_{reg}(\mathbf{x}_i)|$ and $i = 1 \cdots 3$.

Maximum Deviation of Displacement Field MAX_{disp} The maximum deviation of the displacement field states the maximal difference of all components of the two displacement fields $\varphi_{syn}$and φ_{reg}. We use a robust maximum, i.e. our maximum difference is defined as the difference that is larger than 95% of all other values.

7.1.4.2 Image Similarity Measures

To assess the similarity of images before and after registration we use:

Root-Mean-Square of Intensity Differences RMS_{int} The `RMS` of the intensity differences needs as its input the pixel-wise intensity differences over the overlapping region of the image domain. It is defined as

$$RMS_{int} := \sqrt{\frac{1}{N_1 N_2 N_3} \sum_{\mathbf{x}\in\Omega} d(\mathbf{x})^2}$$

with

$$d(\mathbf{x})^2 := \begin{cases} 100 & \text{if } (I_F(\mathbf{x}) - I_M(\mathbf{x}))^2 > 100 \\ (I_F(\mathbf{x}) - I_M(\mathbf{x}))^2 & \text{otherwise} \end{cases}$$

We decided to clamp intensity differences that are larger than 100 Hounsfield Unit (`HU`) and assign the absolute difference of 100 `HU` to all of these differences. This makes the measure more sensitive to lower intensity differences, the unmodified `RMS` measure would weight outliers too strong. Note that we perform this clamping consistently for comparing all image pairs, either original fixed and original moving or original fixed and warped moving.

Median Absolute Deviation of Intensity Differences MAD_{int} This measure is similar to the `RMS`, however it is a more robust variant in the presence of outliers. The median absolute deviation is defined as

$$MAD_{int} := \text{Median}\left(|d(\mathbf{x}) - \text{Median}\left(d(\mathbf{x})\right)|\right)$$

with $d(\mathbf{x}) = |I_F(\mathbf{x}) - I_M(\mathbf{x})|$.

Maximum Intensity Difference MAX_{int} The maximal intensity differences on the overlap region of the image grid. We use a robust maximum, i.e. our maximum intensity difference is defined as the intensity difference that is larger than 95% of all other values. Here of course no clamping (like for the `RMS` measure) of the intensity differences is performed.

Normalized Mutual Information NMI_{int} The normalized mutual information is a measure from information theory, which relates the information content of two images by probability distributions of the gray values. See Chapter 2.1.2.3 for further details.

Edge Overlap $EDGE_{int}$ We are using a simple scheme to extract strong gradients from both images, based on the Canny edge detector [Canny, 1986]. The Canny edge detector is used with a σ of 3 times the voxel spacing, the result is a binary image with edges marked with a 1. The sum of the absolute differences divided by the

	256 × 256 × 256		512 × 512 × 512	
Algorithm	Runtime	Peak-Memory	Runtime	Peak-Memory
symdemons	2500 s	741 MB	9100 s	4.2 GB
curvature	2450 s	803 MB	failed	-
bspline	11400 s	1500 MB	failed	-
elastic	5100 s	741 MB	38000 s	5.6 GB
feature	900 s	785 MB	2700 s	3.2 GB
opticflow	2400 s	741 MB	18400 s	5.6 GB
hybrid	5500 s	741 MB	38000 s	5.6 GB

Table 7.9: Algorithms computational efforts and peak-memory consumptions. All values are only approximate or mean values.

number of voxels in the overlap region is used to compare the binary images. This function lies between 0 and 1 with 0 denoting optimal overlap.

7.2 Experiments on Synthetic Data Sets

According to the setup derived in the previous sections we have performed a large number of synthetic evaluations using three data sets (*nlm*, *sheep*, *human*), three different synthetic transformations (*simbr*, *grid*, *airway*) and seven algorithms (*symdemons*, *curvature*, *feature*, *hybrid*, *bspline*, *elastic*, *opticflow*). Tables C.1, C.2, C.3, C.4, C.5, and C.6 present the results of these evaluations given all of our quantitative measures. These tables are listed in Appendix C. We treated the *airway* transformation a little different, since it was applied only to one data set and using a larger volume size. The results of this experiment are shown in Table 7.10. With our chosen parameters for each of the algorithms we get the approximate algorithm run-times (approximate mean values of all algorithm runs) and peak-memory consumptions as shown in Table 7.9.

The $256 \times 256 \times 256$ data sets have been used in most of our experiments to derive the quantitative evaluation measures. Additionally we have performed the *airway* experiment on the sheep data set which was resampled to $512 \times 512 \times 512$ voxels. In this case the computational and memory requirements are of course more demanding, which can be seen from Table 7.9. Two of the algorithms (*curvature* and *bspline*) did not terminate on these large data sets, due to unresolvable problems with the ITK implementation. They were omitted from the efficiency comparison and we give no quantitative results for the *airway* experiment.

7.3 Experiments on Clinical Data Sets

We have also performed a number of experiments using seven clinical thorax data sets of the human (A,B,C,D,E,F,G). We evaluated the algorithms performance in terms of intensity-based quantitative measures only, since there is no deformation field ground-truth data available. Table 7.11 gives the results of these evaluations. As was already mentioned before, the intensity-based measures should not be overrated, because the corresponding deformation measures are unknown. Therefore we will also show detailed visual results for qualitative comparison.

7.4 Discussion

Our automatic nonlinear registration framework produced a large number of quantitative and qualitative results. We decided to present the conclusions that we drew from the evaluation experiments in several sub-sections, i.e. *computational effort*, *small- and large-scale synthetic deformation experiments*, *realistic synthetic airway deformation experiments* and *clinical data experiments*. The key messages of these parts will be underlined by images and evaluation results.

We use a number of differing quantitative measures, so we start by giving a categorization of the significance of these measures. Most important are the displacement field statistics which are derived from a comparison with ground-truth deformation field data. However, this is only possible for the synthetic experiments, for the clinical data there is no ground-truth available. The `RMS` of the displacement field differences (RMS_{disp}) and its median absolute deviation (MAD_{disp}) are very similar, however, the latter is more robust in the presence of outliers. They provide a means to compare the quality of registration methods in an absolute manner, under the assumption that the synthetic deformation is meaningful. These two measures are the most important ones for assessing algorithms, however, one has to always bear in mind that they try to describe the behavior of millions of voxels with a single number. We expect that RMS_{disp} and MAD_{disp} will point out which algorithms show the best registration accuracy, thereby separating suitable from unsuitable algorithms. The maximal deviation of the displacement field differences provides information about the behavior of an algorithm at image borders where discontinuities might occur in practice. Its relevance should not be overrated, due to the large number of voxels that are matched, while this measure only makes a statement about the outliers that often occur due to occlusions.

The second group of quantitative measures (intensity based measures) has to be interpreted more carefully. In addition to the problem that a large number of intensity differences is described with a single number representing a ”standard-deviation”-like measure, the direct comparison of a warped moving image with the fixed (target) image is dangerous, since this measure does not take into account the underlying deformation field. There are many deformations which lead to the same warped image result, therefore these measures always have to be regarded in combination with the displacement field comparison measures. As a consequence the relevance of these measures on clinical data is questionable, however, up to our knowledge, it is the only possible way to quantitatively describe real data registration results. The `RMS` of the intensity differences RMS_{int} and the median absolute deviation MAD_{int} make statements about the similarity of the fixed target image and the moving image which gets warped according to the deformation field. Due to the nature of the `RMS` calculation, this measure is more prone to outliers, so we expect that the median absolute deviation will give the most reliable statistics of the intensity differences. In order to compare clinical image data we also decided to use the normalized mutual information as a similarity measure. The influence of contrast agents or lung diseases on the intensity differences leads to outliers that restrict the suitability of the previous two measures. Since we are highly interested in accurate registration of vascular structures we also calculate a measure describing an edge overlap statistic. We expect this measure to be rather independent from intensity changes due to its reduction to binary edge images which get compared. A further advantage of the binarization of the edges is that the magnitude of the gradients do not influence it. We expect this measure to distinguish between algorithms that match the vascular structures well and algorithms that have problems with vascular registration. Finally the maximum intensity difference will let us draw conclusions about the intensity difference outliers which might occur at border regions.

7.4.1 Computational Effort

The computational effort of nonlinear registration algorithms is quite high, since we are in general dealing with an ill-posed problem that has to be solved by (regularized) iterative methods, i.e. optimization algorithms. What we can see from Table 7.9 is that most of the investigated methods require computation times on the order of hours for full-sized ($512\times512\times512$) clinical data sets. The main reason for long computation times is the large number of iterations that are needed in the underlying optimization algorithms. The trade-off between algorithm convergence and run-time has to be chosen carefully in practice. For the smaller data sets one gets computation times around 40 minutes if one uses the more efficient of the intensity-based methods (*demons*, *curvature*). These efficient methods often are using certain approximations (e.g. in the *demons* method the regularization is performed by convolution of the displacement field with Gaussian kernels) or they make use of sophisticated methods for solving the underlying `PDE` system (e.g. the *curvature* method uses a fast Fourier transform (`FFT`) to speed-up the solution). Theoretically more profound methods like *elastic* registration and the *hybrid* approach take a long time to compute since they have to solve large `PDE` systems. The *bspline* algorithm needs the highest computational effort due to its formulation with a large number of Degrees of Freedom (`DoF`) modeled in the B-spline framework, where the large number of B-spline coefficients are the unknowns in an optimization framework.

Since the *feature* based registration algorithm uses only a subset of the image information compared to all of the intensity-based approaches, it obviously is the fastest method of all. If one could further speed-up the dense displacement field estimation, it would even be possible to compute in around 10 minutes on the full-sized data sets, a time-span that would be acceptable for applications in clinical practice.

The memory consumption of the algorithms has also to be taken into account. Current 32-bit machines still have limitations on the size of the virtual memory a process can acquire, a value that e.g. lies at 2GB on Windows machines. Most of the presented algorithms do need more than this size if one looks at the processing of the full-sized data sets. So, an obvious solution to this problem is a splitting of the data sets into chunks, which are registered by separate processes. Another solution is to use memory-mapped data structures. However, these techniques are much more complicated to implement and there is a trade-off that has to be made concerning the processing speed, since a lot of computational work would go into the synchronization and communication of the processes or the disk caching. The simpler way to go will be to simply wait until 64-bit systems have conquered the work-stations, making the memory consumption problems more or less obsolete.

We conclude that without further optimizations some of the algorithms like *bspline*, *hybrid* or *elastic* are definitely not suitable for the registration of large CT volume data sets due to their high computational efforts and memory consumptions. The only way to use these algorithms is to apply them to downsampled data sets and to perform a less costly algorithm on the original size using the result from the downsampled step as initial solution. A very promising approach for fast registration is the *feature* based approach, which should be investigated further for possibilities to speed-up. Also a very efficiently implemented variant of *demons* could be useful in clinical practice.

7.4.2 Small-Scale Synthetic Deformation Experiments

The first type of synthetic deformation experiment is the simulated breathing deformation which uses small magnitudes of artificial deformation (*simbr-10-5* and *simbr-25-10*). The measures that describe displacement field differences compared to the ground-truth show that *curvature*, *symdemons*, *feature* and *hybrid* perform best here (see Figure 7.9). These results are consistent with the intensity-based measures which show similar decreases, as an example we give the RMS_{int} in Figure 7.10. However, after looking at the qualitative results of these methods (data set *human*), we found that the quantitative measures do

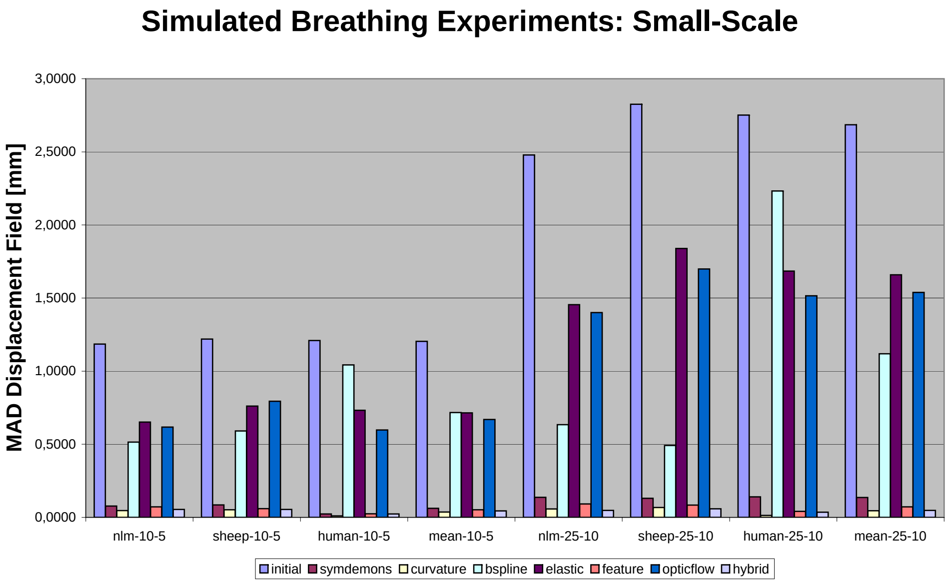


Figure 7.9: Median absolute deviation of the displacement fields for small simulated breathing experiments *simbr-10-5* and *simbr-25-10*.

hide some problems on the one hand while making e.g. the *bspline* algorithm look much worse than it actually is. Figure 7.11 shows a selection of result images. We can see that the *curvature* registration (c) aligns the interior of the images best, however, it also artificially shrinks the warped image, which is not desirable. The *feature* approach (b) gives an appealing visual result, note that the result of the hybrid approach looks visually nearly identical, so in this case the full-grown hybrid model does not bring significant advantages. The *elastic* approach (d) shows problems with the registration of the vascular structures, this is consistent with the quantitative measures and a behavior that also occurs for the *opticflow* algorithm. Finally, we find that the *bspline* algorithm (e) also leads to a very accurate registration result, a fact which is not supported by the quantitative measures. The reason for this discrepancy is a rather large mismatch region that can be seen on the lower part of the axial slice. Since most of this occlusion area is included into the quantitative evaluation the numbers drop significantly. We did not find the reason for this mismatch in the ITK implementation, however, we conclude here that the observation of the visual quality of the registration result is equally important as looking at the numbers.

The second group of small-scale synthetic deformation experiments are *grid-32-2* and *grid-32-4*. Here we can see a very interesting behavior. Although the RMS_{disp} measure

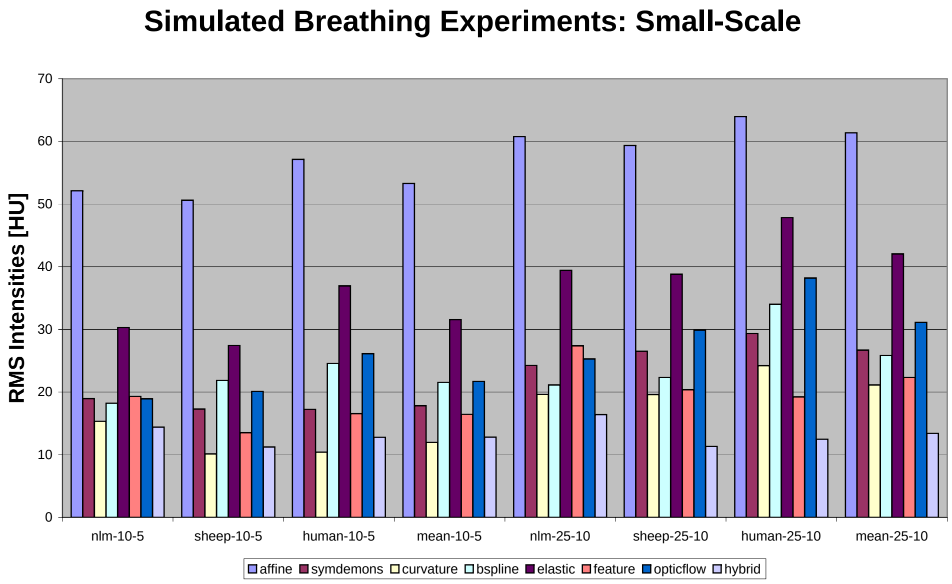


Figure 7.10: RMS of intensity differences for small simulated breathing experiments *simbr-10-5* and *simbr-25-10*.

sees the *curvature* and *symdemons* approaches as the worst methods, a closer look at MAD_{disp} reveals that the `RMS` is influenced by outliers. By looking at the robust measure these two algorithms perform better by an order of magnitude. The *feature* approach does not deal well with this kind of deformation. This is due to the fact that the *grid* experiment makes use of a regular grid of random displacement vectors, i.e. the warping is performed rather locally and similar to local rotations. It can be assumed that this creates severe problems for the repeatability of the feature extraction. The *hybrid* approach is not able to use the feature information, it performs similar to the rest of the methods, the median absolute deviations of the displacement fields are shown in Figure 7.12. The visual results of this set of experiments on data set *nlm* using deformation *grid-32-4* also clearly show that the *curvature* and the *symdemons* approaches (c) perform best. The *opticflow* and the *hybrid* approaches perform very similar, with problems at some of the vascular structures, here the number of iterations presumably was too low. The low quality of the *feature* approach underlines the short-comings of this method and supports the quantitative results. Only a small number of feature correspondences were found here. Finally it should be noted that this kind of synthetic deformation is not physically justified, but was chosen since we found some publications in literature that use similar transformations.

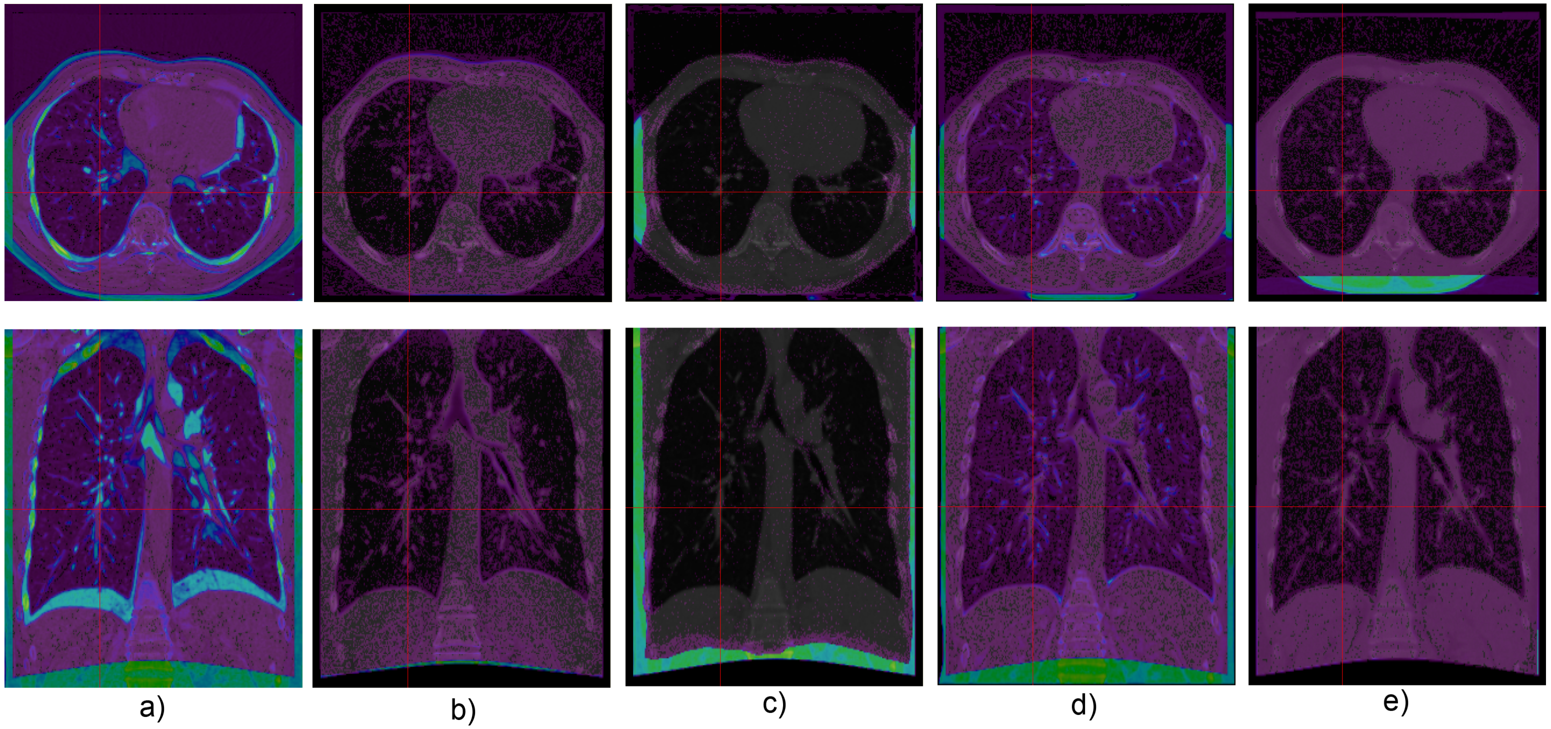


Figure 7.11: Synthetic result images for small-scale simulated breathing experiment *simbr-25-10* on data set *human*. a) affine registration, b) feature registration, c) curvature registration, d) elastic registration, e) bspline registration. Top row axial, bottom row coronal slices.

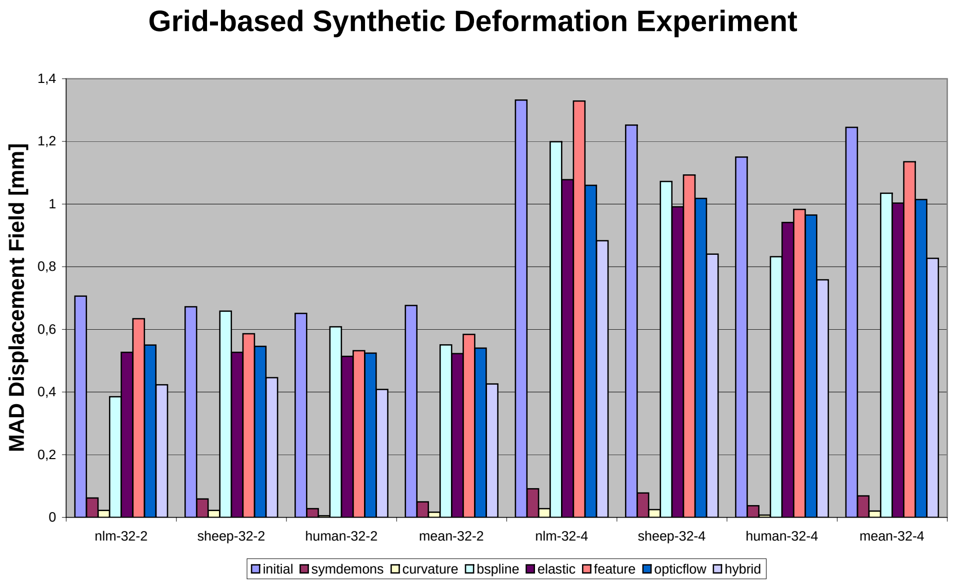


Figure 7.12: Median absolute deviation of the displacement fields for small grid based synthetic experiments *grid-32-2* and *grid-32-4*.

7.4.3 Large-Scale Synthetic Deformation Experiments

We now investigate the performance on the large-scale simulated breathing experiments (*simbr-55-25* and *simbr-70-30*). While the *simbr-55-25* experiment shows quite similar results to the small-scale experiments, we will focus on the extremal *simbr-70-30* experiment, which was designed to show the limitations of the algorithms given a deformation that is larger than a real diaphragm or rib cage movement will be during normal breathing. Note that we already have a significant occlusion at the top part of the volume. Figure 7.14 shows the distribution of the `RMS` of the displacement field differences for this synthetic deformation. The quantitative results imply that the *curvature* registration gives the best results. This is also supported by the edge overlap measure as depicted in Figure 7.15. The *feature* registration is not able to find enough correspondences. The *hybrid* approach is able to register the data sets quite well, however, there are too few iterations to converge. Similar problems apply to the *elastic* and *opticflow* methods. The *bspline* method again shows excellent behavior which can also be seen from the qualitative results in Figure 7.16d.

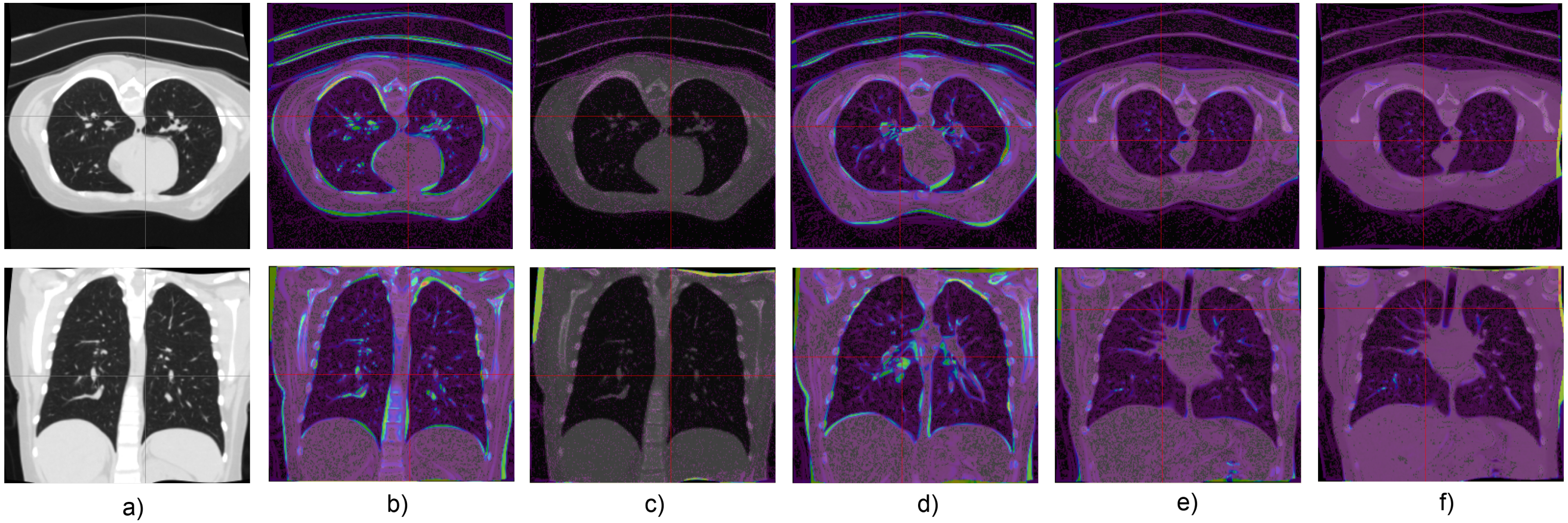


Figure 7.13: Synthetic result images for grid-based synthetic deformation experiment *grid-32-4* on data set *nlm*. a) synthetically warped image, b) affine registration, c) symmetric demons registration, d) feature registration, e) optic flow registration, f) hybrid registration. Top row axial, bottom row coronal slices.

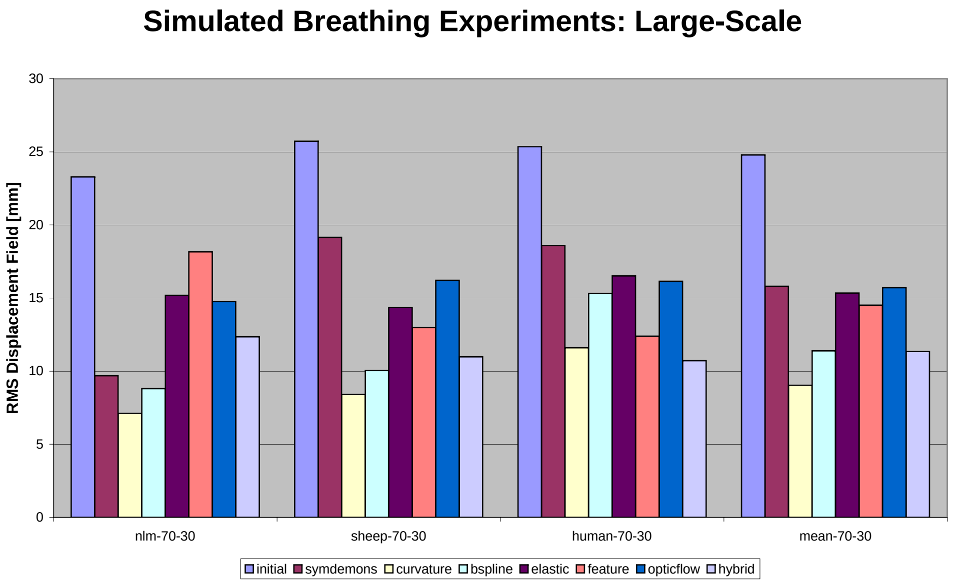


Figure 7.14: RMS of displacement fields for large-scale simulated breathing experiment *simbr-70-30*.

7.4.4 Synthetic Airway Deformation Experiment

The physically most realistic synthetic experiment (*airway*) is performed on the large volume data sets of size $512 \times 512 \times 512$. It uses manually identified airway tree landmarks as its basic building block and an `EBS` interpolation to get the dense synthetic displacement field. Unfortunately two of the methods (*bspline* and *curvature*) did not successfully finish the registration, we were not able to recover these problems of the ITK implementations. The `RMS` and the median absolute deviation of the displacement field differences (compare Table 7.10) show that the *feature* registration, the *symdemons* registration and the *hybrid* registration perform very well, while the *elastic* and the *opticflow* methods are not able to fully register the data. This outcome is supported by the intensity difference measures and the edge overlap measure. An interesting fact is that the *feature* approach out-performed the *hybrid* algorithm. The reason is, that the *hybrid* approach did not converge to a minimum fast enough given only a restricted number of iterations. This restriction had to be introduced due to the high computational effort of this method on the large data sets. Figure 7.17 shows the result images for qualitative comparison. We can see the result images of the *feature* approach in c) and the *symdemons* approach in d). Both are visually nearly indistinguishable, there is only a slight misregistration of a bone in the

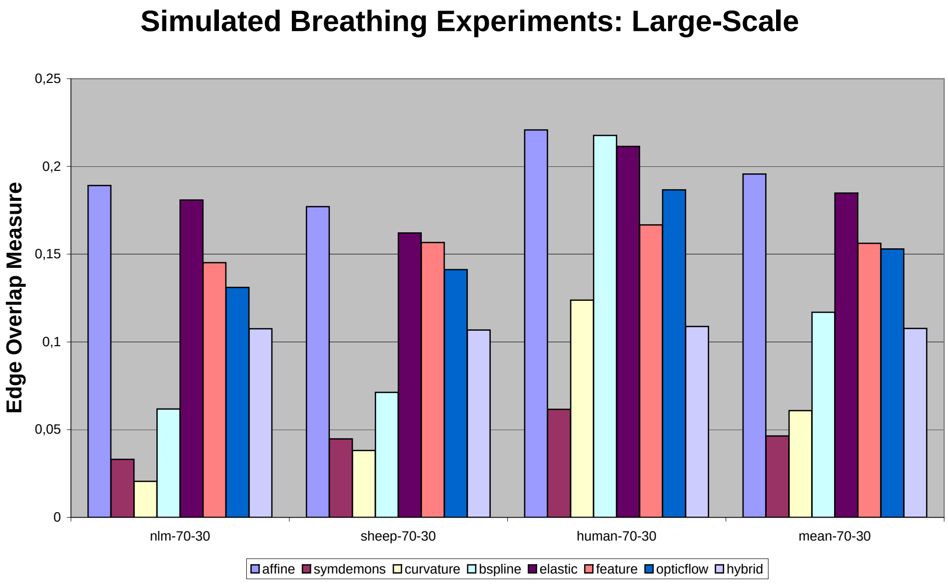


Figure 7.15: Edge overlap measure for large-scale simulated breathing experiment *simbr-70-30*.

upper right part of the coronal image for the *feature* approach. Finally we also show the *elastic* approach in e) which has problems with some of the vascular structures and the bones.

7.4.5 Clinical Data Experiments

Our seven clinical data sets comprise of three different application scenarios. First, there are three data sets (A,C,D) that only show very small breathing deformation. However, these three data sets are composed of a native and a contrast-enhanced scan, i.e. we are confronted with a violation of the brightness constancy assumption. Data set B also shows an intensity change due to contrast-agent, but in addition there is also a certain amount of breathing deformation. Data sets E,F and G are composed of two native scans, where E and F both show a significant breathing deformation and G a medium-sized breathing deformation. These latter three data sets are all very challenging, since in addition to the intensity changes from partial volume effects due to breathing, all of them show more or less severe kinds of lung diseases. We expect data sets E,F and G to be very hard to register. In the following we concentrate on the four data sets B,C,E and G, as representatives for the rest.

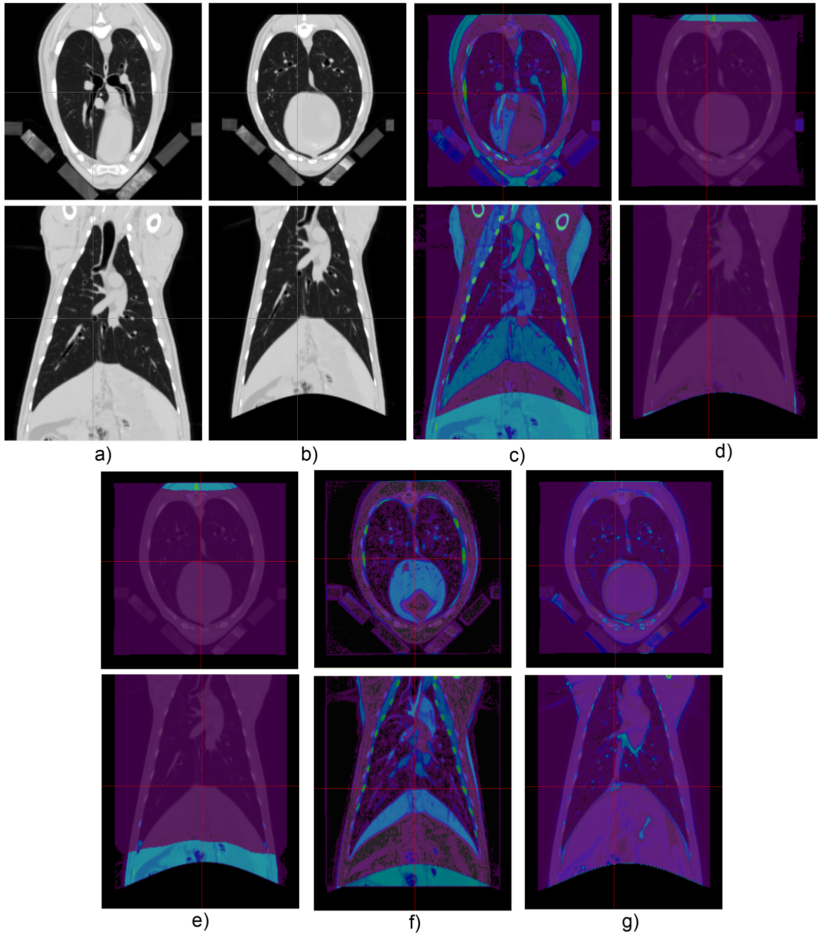


Figure 7.16: Synthetic result images for large-scale simulated breathing experiment on data set *sheep*. a) original image, b) synthetically warped image, c) affine registration, d) bspline registration, e) symmetric demons registration, f) feature registration, g) hybrid registration. Top row axial, bottom row coronal slices.

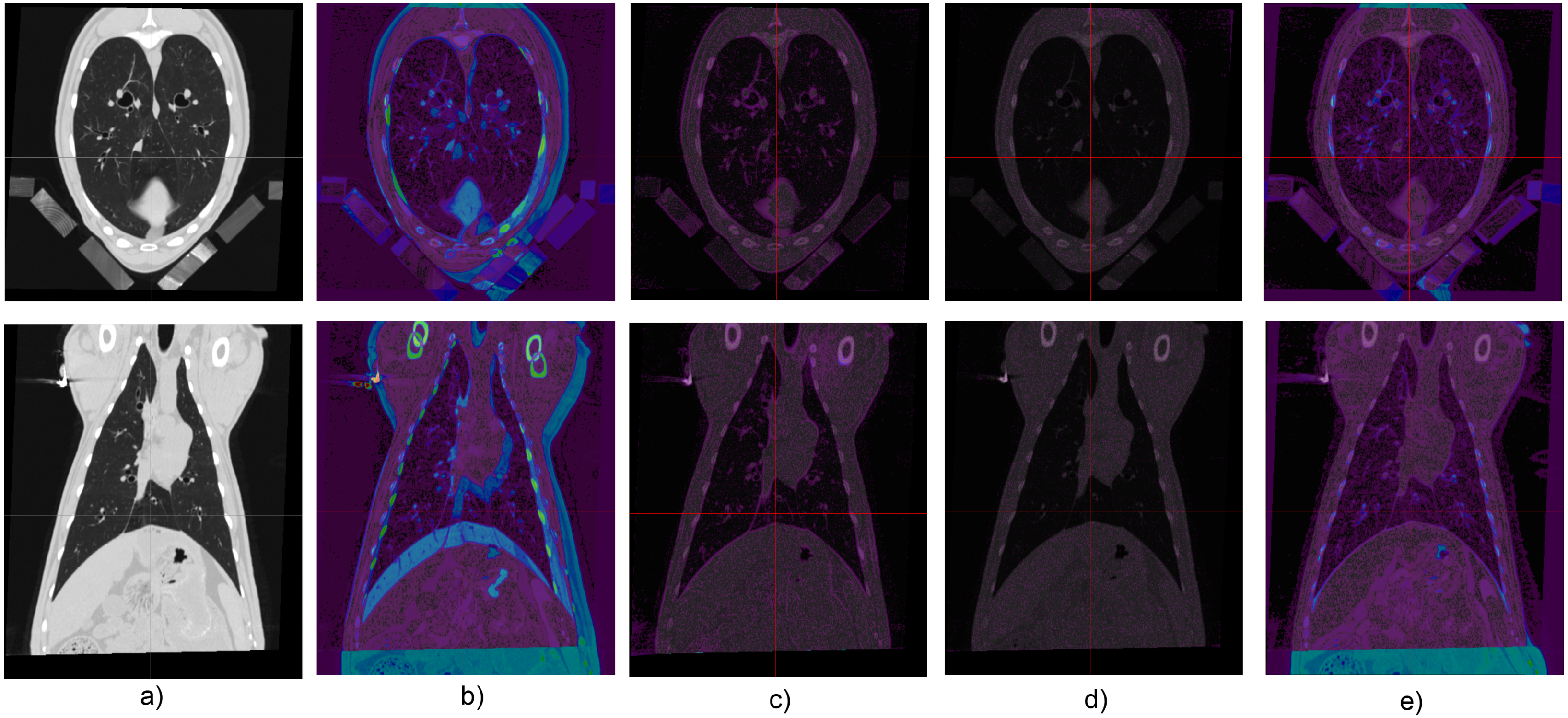


Figure 7.17: Synthetic result images for synthetic deformation experiment *airway* on data set *sheep*. a) synthetically warped image, b) affine registration, c) feature registration, d) symmetric demons registration, e) elastic registration. Top row axial, bottom row coronal slices.

Intensity Measures			**airway - sheep**	Deformation Measures			**airway - sheep**
RMS_{int}	*affine*	[HU]	63.428	RMS_{disp}	*initial*	[mm]	13.405
	symdemons	[HU]	5.096		*symdemons*	[mm]	0.402
	elastic	[HU]	43.148		*elastic*	[mm]	9.678
	feature	[HU]	4.532		*feature*	[mm]	0.476
	opticflow	[HU]	46.351		*opticflow*	[mm]	10.204
	hybrid	[HU]	8.596		*hybrid*	[mm]	1.401
MAD_{int}	*affine*	[HU]	30	MAD_{disp}	*initial*	[mm]	5.253
	symdemons	[HU]	1		*symdemons*	[mm]	0.043
	elastic	[HU]	10		*elastic*	[mm]	2.928
	feature	[HU]	1		*feature*	[mm]	0.096
	opticflow	[HU]	12		*opticflow*	[mm]	3.287
	hybrid	[HU]	2		*hybrid*	[mm]	0.241
NMI_{int}	*affine*		0.1237	MAX_{disp}	*initial*	[mm]	27.664
	symdemons		0.8871		*symdemons*	[mm]	10.912
	elastic		0.2916		*elastic*	[mm]	22.258
	feature		0.7691		*feature*	[mm]	12.254
	opticflow		0.2649		*opticflow*	[mm]	22.929
	hybrid		0.6734		*hybrid*	[mm]	16.736
$EDGE_{int}$	*affine*		0.2361				
	symdemons		0.0230				
	elastic		0.2141				
	feature		0.0595				
	opticflow		0.2203				
	hybrid		0.1124				
MAX_{int}	*affine*	[HU]	923				
	symdemons	[HU]	12				
	elastic	[HU]	214				
	feature	[HU]	20				
	opticflow	[HU]	300				
	hybrid	[HU]	42				

Table 7.10: Airway synthetic transformation results in terms of intensities and displacements.

Data set D only shows a minor breathing deformation. All of the algorithms are able to recover this small deformation, however, the measures on the `RMS` and the median absolute deviation of the intensity differences imply a different behavior. Of course here one has to take into account the influence of the contrast agent, which misleads these two measures. Therefore, we rely on the `NMI` measure for comparing the algorithms performance (see Table 7.11). Here, the performance is very similar for the compared algorithms, which is consistent with the selected visual results depicted in Figure 7.18. As a consequence of these similar results, we restrict ourselves to the *feature* and *bspline* approaches here. Note that the intensity differences in the vascular structures of the registered images are due to the contrast agent. The feature-based approach shows a slight misregistration at the extremal regions of the diaphragm which comes from a lack of nearby extracted feature information.

The second data set we want to focus on is data set B, here a medium breathing de-

Measure			A	B	C	D	E	F	G
RMS_{int}	*affine*	[HU]	41.249	48.086	44.938	43.220	68.921	64.466	66.854
	symdemons	[HU]	25.750	24.814	30.134	25.454	36.830	36.178	32.269
	curvature	[HU]	31.811	29.313	37.628	30.160	42.171	39.529	40.705
	bspline	[HU]	41.357	42.700	49.362	38.714	69.941	68.740	64.824
	elastic	[HU]	36.390	39.400	41.684	34.487	70.836	70.322	65.865
	feature	[HU]	32.061	39.382	41.165	32.388	45.836	46.322	43.865
	opticflow	[HU]	34.373	33.632	39.826	31.634	50.014	46.937	48.023
	hybrid	[HU]	30.365	36.192	37.255	32.104	40.246	41.371	39.127
MAD_{int}	*affine*	[HU]	12	12	16	10	47	30	41
	symdemons	[HU]	3	2	6	2	6	4	5
	curvature	[HU]	8	6	11	6	11	9	11
	bspline	[HU]	12	9	15	9	24	20	21
	elastic	[HU]	10	9	13	8	21	15	19
	feature	[HU]	9	8	11	8	32	27	26
	opticflow	[HU]	9	7	13	7	15	12	15
	hybrid	[HU]	9	7	10	8	12	13	13
NMI_{int}	*affine*		0.4263	**0.3574**	0.3520	**0.4024**	**0.1574**	0.1807	**0.1634**
	symdemons		0.6269	**0.6452**	0.5174	**0.6342**	**0.4889**	0.4946	**0.5289**
	curvature		0.5204	**0.5442**	0.4085	**0.5331**	**0.4029**	0.4249	**0.4251**
	bspline		0.4463	**0.4370**	0.3591	**0.4588**	**0.2511**	0.2617	**0.2640**
	elastic		0.4811	**0.4541**	0.3834	**0.4876**	**0.2787**	0.2991	**0.2988**
	feature		0.5509	**0.5009**	0.4278	**0.5564**	**0.2200**	0.2338	**0.2379**
	opticflow		0.5028	**0.5017**	0.4002	**0.5180**	**0.3489**	0.3625	**0.3652**
	hybrid		0.5728	**0.5284**	0.4567	**0.5837**	**0.3964**	0.3892	**0.3732**
$EDGE_{int}$	*affine*		0.1524	0.2020	0.1935	0.1893	0.2487	0.2346	0.2380
	symdemons		0.0905	0.1031	0.1478	0.1020	0.1656	0.1654	0.1454
	curvature		0.1083	0.1190	0.1646	0.1198	0.1844	0.1777	0.1656
	bspline		0.1273	0.1530	0.1757	0.1394	0.2162	0.2013	0.2074
	elastic		0.1193	0.1557	0.1723	0.1396	0.2263	0.2164	0.2140
	feature		0.1015	0.1235	0.1286	0.1498	0.2232	0.2208	0.2195
	opticflow		0.1149	0.1367	0.1677	0.1308	0.2041	0.1984	0.1921
	hybrid		0.1036	0.1036	0.1136	0.1347	0.1632	0.1723	0.1590
MAX_{int}	*affine*	[HU]	162	383	168	243	958	981	905
	symdemons	[HU]	60	50	82	60	118	99	78
	curvature	[HU]	83	76	108	84	141	115	109
	bspline	[HU]	125	143	139	126	355	342	298
	feature	[HU]	129	168	122	141	885	861	734
	elastic	[HU]	93	139	111	97	380	307	283
	opticflow	[HU]	80	86	100	77	159	149	142
	hybrid	[HU]	75	84	99	73	158	140	140

Table 7.11: Clinical data evaluation results in terms of intensity-based measures.

formation in combination with an applied contrast agent is the problem domain. Again, this registration task is performed very well by most of the algorithms, see also Table 7.11. Selected visual results are shown in Figure 7.19. Here we can see the drawbacks of the *feature* registration (c). Near the diaphragm borders there was not enough reliable feature information extracted, therefore the approach fails to align these contours. This problem was effectively solved by the *hybrid* approach (f). However, one will note that the vascular structures already show an excellent overlap for the efficient *feature* algorithm. The

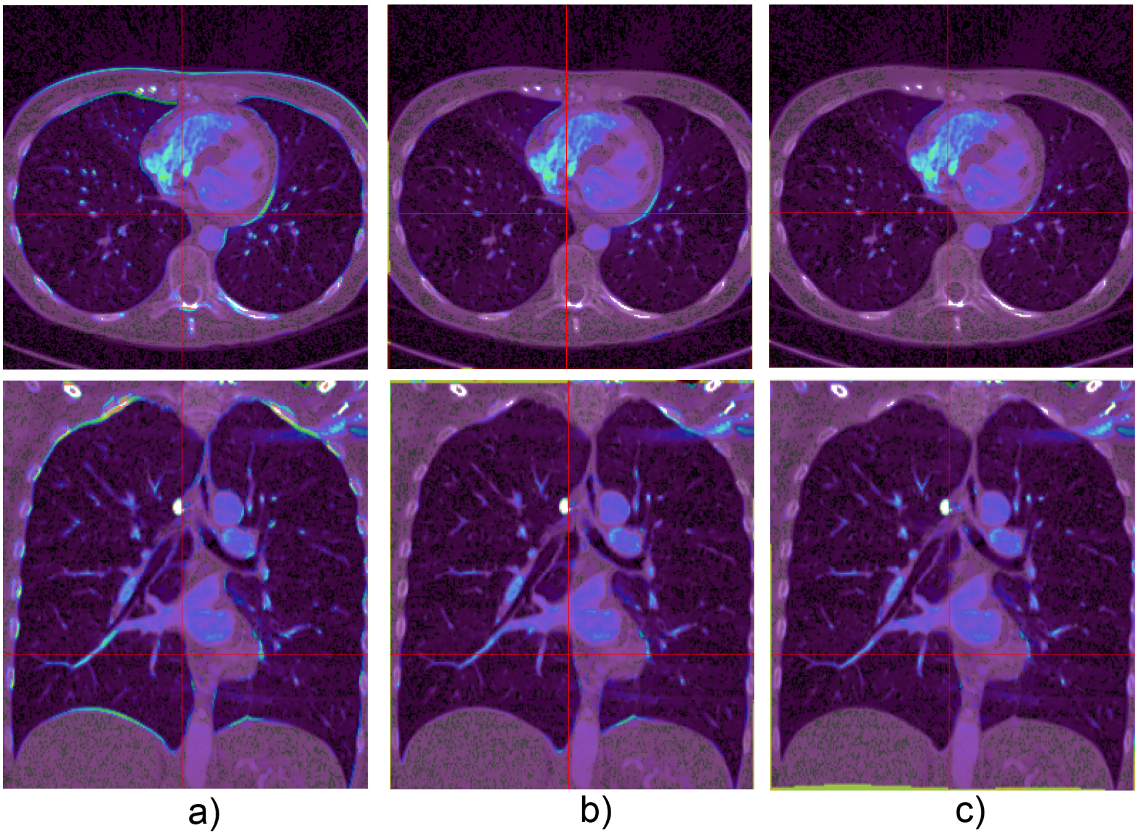


Figure 7.18: Selected clinical result images for data set D. a) affine registration, b) feature registration, c) bspline registration.

symdemons and *bspline* algorithms also show very good registration behavior.

Data set E combines a large breathing deformation with a slight lung disease. In this case now, we have reached the point where most of the investigated algorithms fail. The quantitative results do not clearly show this behavior, the `RMS` and median absolute deviations of the intensity differences as well as the `NMI` measure show that *symdemons* should lead to accurate registration while the *bspline* method gives quite poor numbers. In reality, *symdemons* and *curvature* show severe artifacts in the interior of the lung, destroying the vascular structures. This can be seen on Figure 7.20c,d. We assume that the regularization is not strong enough here, the calculated deformation field is physically implausible. Without proper regularization the `SSD` similarity measure is allowed to match intensities without restriction, and due to the partial volume effect the brightness constancy assumption is violated. The combination of these behaviors leads to the visible problems. On the other hand, the *bspline* algorithm performs a much more meaningful registration. No vascular structures are destroyed and the rough lung and bone structures

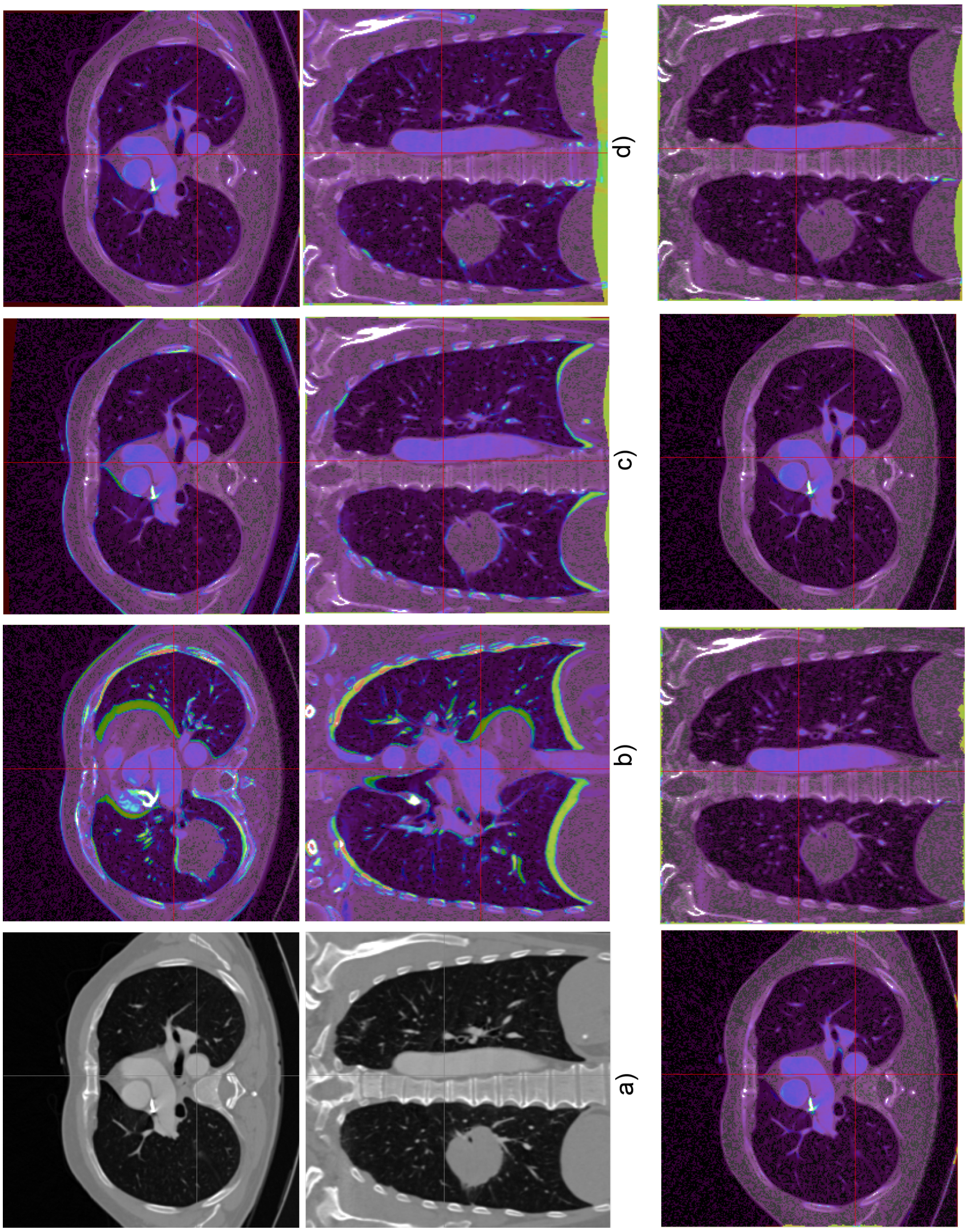


Figure 7.19: Selected clinical result images for data set B. a) contrast-enhanced image, b) affine registration, c) feature registration, d) bspline registration, e) symmetric demons registration, f) hybrid registration.

are aligned accordingly, however, the alignment of the vascular structures is still far from perfect. Also the *hybrid* (Figure 7.20f) approach shows physically plausible behavior, here one has to keep in mind that the hybrid approach falls back to intensity registration only, since there is nearly no feature information that could be found in the *feature* matching algorithm. This is the reason why the *feature* result shows behavior that is even worse than the original moving image.

Finally data set G poses a challenging problem since it contains a pathologic lung. On the other hand, the breathing deformation between the inhalation and exhalation images is only medium-sized. Now the registration results are much better, we assume that the medium-sized breathing deformation is still in the capture range of the algorithms as opposed to the large-size deformation from data set E. The *feature* approach again does not find enough correspondences to be very accurate, especially some large-scale structures like the diaphragm are not aligned well. Furthermore, it has to be mentioned that the diseased regions in the upper parts of the lung do not exhibit structure, which prevents features from being extracted there. The *symdemons* and *curvature* approaches are prone to smearing effects again in the lower part of the right lung. The *bspline* method shows a good registration result, while the quantitative evaluations from Table 7.11 suggests the contrary. We note that the resulting warped image from *bspline* appears more smoothed than the other methods, presumably due to the internal B-spline interpolations. The *hybrid* approach shows a reasonable result, however, a converged state has not been reached.

7.4.6 Conclusions

In this extensive evaluation study, we found that the simple *Symmetric Demons* algorithm often out-performs the other compared methods while still needing moderate run-times. This is especially true for the synthetic deformation experiments. Also the *curvature* algorithm leads to appealing results, however, there are some cases where it leads to border artifacts that shrink the warped volume. Both approaches have their difficulties when it comes to clinical data registration of data sets that differ by large breathing motion. Interestingly, the registration of small-to-medium sized breathing differences in the presence of contrast agents does not prevent these methods from successful registration, although the important constant brightness assumption is violated. At this point we also want to remark that, compared to the evaluations in Chapter 5 we have chosen a larger number of iterations for the Symmetric Demons algorithm, which explains the performance improve (at the cost of run-time).

The trade-off between run-time and accuracy inevitably leads to the *feature* algorithm. It is the fastest algorithm of our study due to its inherent reduction of the problem

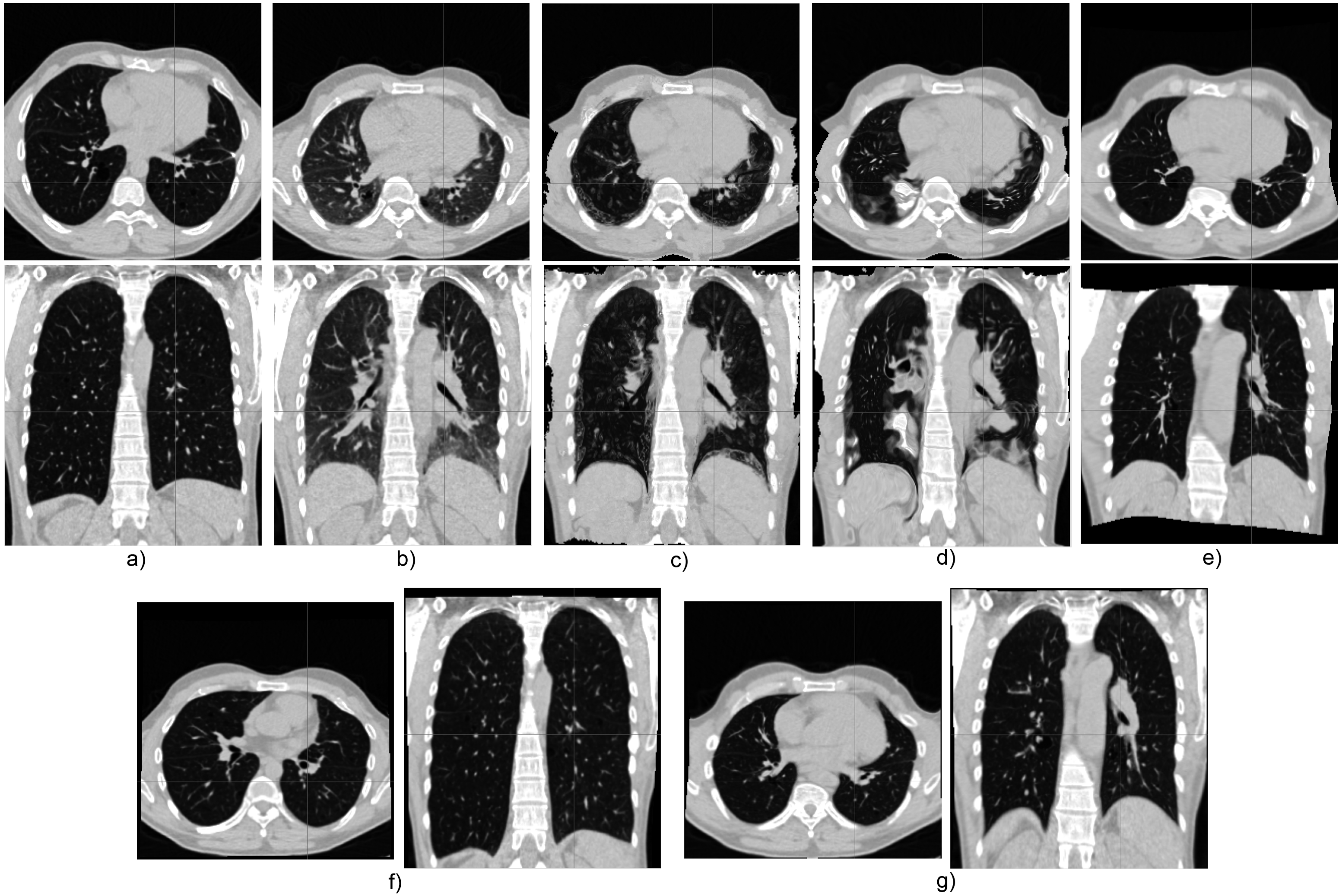


Figure 7.20: Selected clinical result images for data set E. The moving inhalation image b) is registered to the fixed exhalation image a). We show the warping results for the different methods. c) symmetric demons, d) curvature, e) bspline, f) feature, g) hybrid registration.

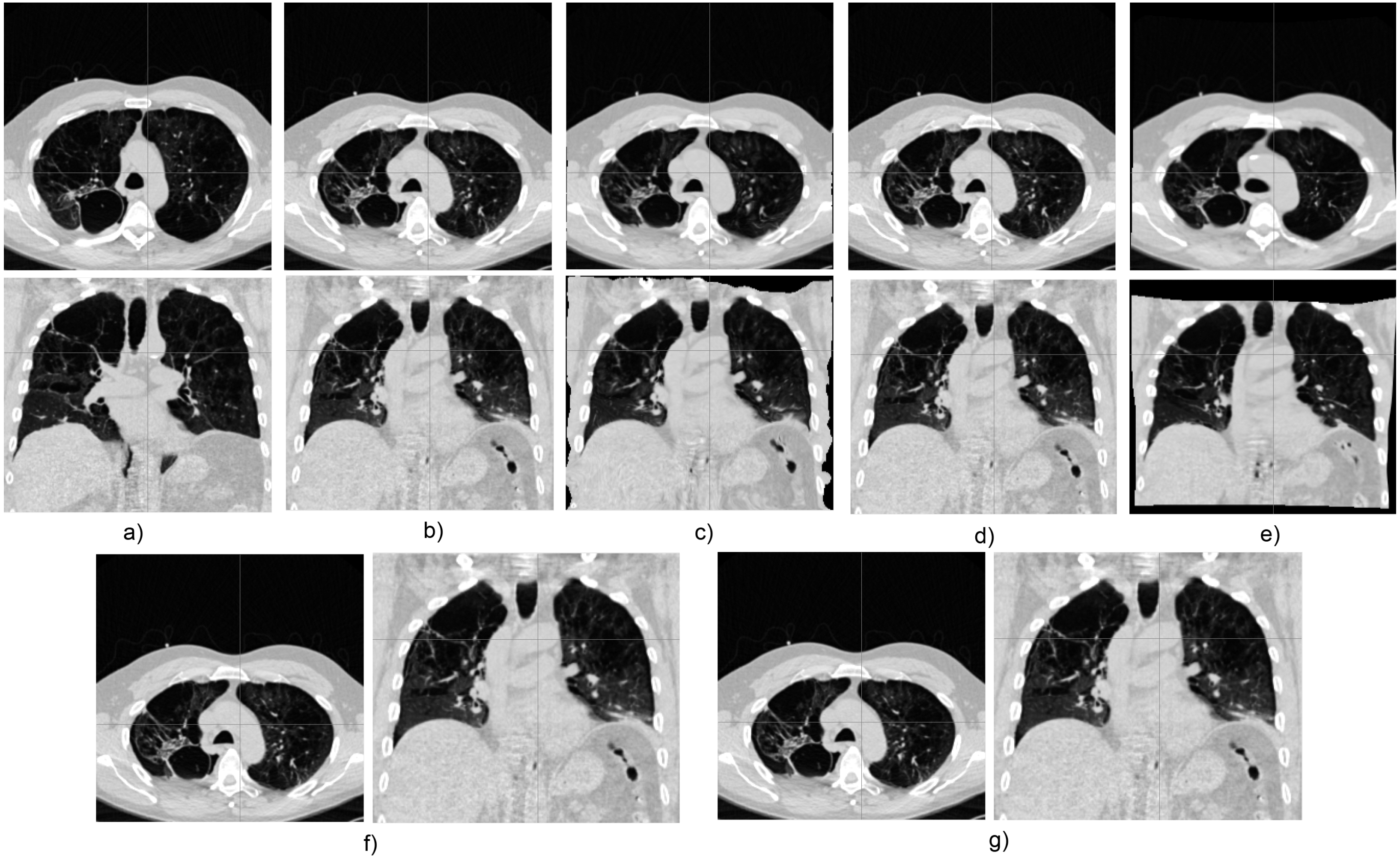


Figure 7.21: Selected clinical result images for data set G. The moving inhalation image b) is registered to the fixed exhalation image a). We show the warping results for the different methods. c) symmetric demons, d) curvature, e) bspline, f) feature, g) hybrid registration.

domain to a feature subset and it shows acceptable performance on a number of synthetic experiments. It also masters the clinical data sets very well, however, there are some cases where it totally fails due to a too small number of extracted feature correspondences or their uneven distribution. These cases are all experiments that involve large-scale deformations. The largest potential of this method is that unlike the intensity-based methods it scales very well when further increasing volume size. We also want to note here, that the downsampling of volumes (which involves smoothing) is a severe disadvantage for the feature extraction steps. In our experience (and this was also demonstrated in the synthetic *airway* experiment) the feature extraction step is much more reliable when performed on the original volume sizes, since the features we are interested in mostly are located on this original scale. Yet, to be able to perform a fair comparison with the other methods we decided to perform this method to the downsampled images. We conclude that further investigation of feature-based approaches are definitely a promising direction to follow.

The shortcomings of the *feature* method were effectively prevented by the *hybrid* algorithm. This method is very appealing from a theoretical point of view, however, the practical suitability is largely reduced due to its high computational cost. Further investigation of speeding up this method should definitely be performed.

The *elastic* and the *opticflow* methods did not convince in our experimental study. Both methods use a regularization scheme that is too restrictive for matching of rather large breathing deformations. An increase in the number of iterations per multi-resolution level would be necessary to get more accurate results. This was also shown in other publications (e.g. [Modersitzki, 2004] for elastic registration). In combination with the rather high computational effort that is necessary for these methods (especially the *elastic* scheme) we do not intend to investigate them further.

B-spline mutual information registration depends on a regularly-sized grid defining the B-spline control points. This has two severe drawbacks. First, the granularity of the grid can not be adapted to the underlying image information. Background or large homogeneous tissue regions get more control points than necessary, while regions that show much high-frequency information (lung tissue including airways and vascular structures) get fewer. Therefore it is very hard to determine the necessary grid spacing in practice. The second drawback is its large number of parameters that have to be tuned to be able to get a decent registration. From the ITK algorithms that we investigated, we spent most of the time tuning the *bspline* parameters. Despite these efforts we were unable to register the *airway* experiment. However, if meaningful parameters can be found, the performance of the registration is very well on synthetic and clinical data. Note that in practice, the large amount of parameters that have to be optimized in this method pose a restriction on the B-spline grid size to use, so we conclude that this algorithm has the worst scaling behavior when volume sizes increase.

Another interesting fact the *bspline* algorithm showed, is that one can not always trust the quantitative results that we computed from the experiments. Apart from the necessity to look at intensity measures only in combination with the ground-truth deformation field measures, we also found that it is very important to look at the result images as well. We conclude that our automated evaluation framework is not yet capable of presenting visual results in connection with the quantitative result, so we have to think of a way how to tackle this problem. Nevertheless, we are convinced that the only way for thorough evaluation of nonlinear registration is to use an automated framework, since the computation and the sighting of all these experiments is a time-consuming, tedious process.

Chapter 8

Summary and Conclusion

This thesis is concerned with the investigation of a variety of nonlinear registration algorithms for application to intra-modality medical image `CT` data sets. The investigation lead to a number of contributions to the current state-of-the-art in intra-modality registration. Special concern was laid on the topics of `CTA` registration concerning brain and soft tissue thoracic image data sets. The topic of nonlinear registration is considered as an important research area in the medical imaging community, a fact that is reflected in our extensive literature review presented in Chapter 2. Algorithms make use of anatomical landmarks, segmented structures or whole intensity volumes, their complexity ranges from simple block-matching techniques to full-grown frameworks based on energy functionals that require systems of partial differential equations to be solved. There is a large space of methods inbetween these complexity extrema, the most promising techniques provide an approximation to the full-grown functional minimizers as a natural trade-off between computational effort and theoretical performance. Many researchers also begin to fuse anatomical landmark and intensity-based methods to hybrid approaches. Evaluation of nonlinear registration is far from being a straight-forward topic, this is reflected in the small number of existing evaluation studies for thoracic soft-tissue registration algorithms. Among these topic areas we have contributed in several parts, summarized in the following section.

8.1 Summary of our Contributions

There are situations in clinical practice when brain `CTA` images have to be registered in the presence of several independently moving bone structures. Current methods from the literature either ignore this fact or try to solve it using highly complex segmentation steps as a pre-processing step for registration. In Chapter 3 we have proposed a novel approach to deal with the problem of several rigid movements occurring during a single or between native and contrast-enhanced scans of a brain `CTA` study. Our method makes use of a joint segmentation and registration algorithm that iteratively refines the registration while adding more and more rigid model assumptions to the independently moving bone structures. We have shown a comparison with a standard approach from the literature that neglects independent movements and have proved that our algorithm performs more accurately by the means of quantitative and qualitative comparisons. We also showed that the proposed method is efficient to compute which is very important giving the size of current clinical brain `CTA` input data. A first clinical evaluation study on a large number of data sets has already been published in [Lell et al., 2006b], who show promising results of the `MIP` visualizations compared to the standard methods.

Despite its large diversity the literature on nonlinear soft tissue registration lacks in automatic landmark extraction, matching and registration methods, a work-flow that is very popular in the more traditional computer vision literature, where it is used for topics like two- and multi-view 3D reconstruction or optical flow motion computation. In this thesis we investigated several popular computer vision standard techniques for its applicability to medical image data. First, we focused on the use of the shape context descriptor for point matching, leading to a pipeline consisting of shape segmentation from thoracic image data (lungs, airway tree, diaphragm), 3D surface approximation, regular point sampling, graph-based matching of shape points using a combined cost function from the shape context descriptor and the normalized cross correlation, and dense displacement field interpolation using the `TPS` framework. This work-flow was described in Chapter 4. We found that given the segmentation of organs of interest this method can be a very useful approach for efficient feature-based nonlinear registration. The shape context registration was used to register several sheep thorax data sets to study the static behavior of breathing. The research interest from the medical point of view is the better understanding of the physiology of breathing. Our evaluations showed that it is possible to predict a simple breathing model with this approach.

One of the drawbacks of the shape context based approach was the need for segmentation as a pre-processing step. On the high-contrast and low Signal to Noise Ratio (`SNR`) sheep data sets the segmentation was rather easy, however, on real-world clinical patient data segmentation is a much more complex task. To get away from this need for segmentation we established a feature-based registration pipeline (compare Chapter 5), consisting of a Förstner point(s) of interest (`POI`) detection scheme, a combined local and global `POI` descriptor calculation using `SIFT` and shape context, a robust forward-backward matching and a displacement field densification using approximating `TPS`. This approach was designed with computational efficiency and accurate matching of small vessel structures in mind. We showed that the computational effort of the automatic pipeline scales with the number of extracted features as opposed to the number of voxels like in intensity-based methods. This becomes more and more important if one considers the increasing sizes of clinical volume data sets. Our quantitative evaluations showed comparable behavior to the Demons algorithm, with lower computational effort. However, we also identified the need for evenly distributed feature matches over the volume as a slight disadvantage of the method.

In Chapter 6 we investigated a hybrid nonlinear registration framework that unifies feature- and intensity-information. Our general framework makes use of a formulation as a variational energy functional. We created a special instance of our registration framework using an optical flow `SSD` data term, an anisotropic image-driven diffusion regularization term and a feature constraint term derived from the feature-matching algorithm of Chapter 5. This model is a full-grown nonlinear registration scheme that integrates the theory of regularization of ill-posed inverse problems and the wide-spread optical flow concept with our novel feature-based registration. The minimization of the energy functional leads to solving the Euler-Lagrange partial differential equation system in order to get to a solution of the desired nonlinear transformation. However, the obvious drawback of this full-grown model is its increased computational effort. Despite the feature-based solution being able to speed up the convergence of the `PDE` solver, the runtime is in our opinion not yet low enough to apply the method to real-world clinical application areas. Our quantitative and qualitative evaluations of the method have shown the excellent behavior on synthetic and real-world data.

Our final contribution was a comparison study of several state-of-the-art nonlinear registration algorithms with our proposed algorithms using a number of synthetic transformations and evaluation measures in addition to real-world clinical data. For the synthetic transformations we knew the ground-truth, so an exact quantitative comparison was possible, however, the findings should be interpreted with care. Especially in the area of nonlinear registration, the definition of meaningful evaluation methods is a very hard problem. We argued that a combination of several different synthetic transformations in combination with a large number of measures describing different quantitative information and with a fully-automated evaluation framework may be used to derive meaningful statistics. This study is presented in Chapter 7. The outcome of the study showed that the efficient feature-based algorithm is competitive with standard algorithms like Symmetric Demons. It also showed that for extremal deformations, the performance of the feature-based approach decreases. By putting the feature matches into a hybrid variational framework approach it was possible to make it robust to feature mismatches. The cost was a dramatic increase in computation time. We also found that, although we tested a large number of different algorithms, it was not possible to accurately register our more difficult clinical data sets. This leaves more space for future developments.

8.2 Directions for Future Work

From the findings of this thesis a number of directions for future work have emerged, which have already been discussed in the final sections of the chapters of this thesis. Here we would like to summarize these directions and present them in a compact manner in addition to further more general ideas.

The partially rigid registration could be improved by investigating joint segmentation and registration frameworks like the one presented in [Yezzi et al., 2003]. A variational formulation could e.g. minimize the similarity measure while at the same time modeling the segmentation of the bone structures using the Mumford-Shah functional for segmentation. An additional constraint would keep the number of different rigid registration models low and implicitly perform the interpolation of adjacent registration matrices. However, it is to be expected that such a model would not be efficient to compute anymore. The approach could also benefit from a comparison to elastic registration approaches like in [Bentoutou et al., 2002] or similar to [Little et al., 1997], however, the need for manual interaction has to be reduced. An interesting approach to look into is [Arsigny et al., 2005], where a novel class of geometrical transformations is presented (polyrigid/polyaffine transformations). These transformations have few `DoF`, however, by proper weighting of few rigid/affine transformations a model suitable for our registration purpose could be built. Also the algorithm of [Keeling and Ring, 2005] which provides a nonlinear registration method with maximal rigidity constraints should be considered for future work.

Future work on the surface-based registration approach should analyze the various pipeline stages for their respective contribution to the overall registration error. We assume that the shape discretization and the shape matching stage have the largest effect on the error. Therefore, fine-tuning these stages by a more dense point sampling and by adding an outlier robustness term to the matching should be performed. Another way of using the results of the surface-based registration is, by incorporating them in a more generic feature registration framework, where correspondences from landmarks and surfaces are combined. A comparison of the shape context method with other standard methods like robust point matching [Chui et al., 2003] or relaxation labeling [Zheng and Doermann, 2006] should prove our assumption (which was derived from the two-dimensional equivalents) that the shape context is most suitable for our purpose.

The landmark-based feature registration approach is a very promising way for efficient nonlinear image registration. However, we still see some ways of improvement, since we only started the investigation of dense feature based registration methods. The method could benefit from a comparison of further landmark feature extraction methods like those presented in [Kitchen and Rosenfeld, 1982, Thirion and Gourdon, 1995, Loy and Zelinsky, 2003]. Based on previous work of our group presented in [Pock et al., 2005a] we also think of using the centerlines of the vascular structures or their branching points [Tschirren et al., 2005b] as features for matching. Centerline matching could be performed by a relaxation labeling approach [Zheng and Doermann, 2006]. In the long run, it would be interesting to define a framework using a large number of different features that were extracted using several methods. Unifying these features in a common mathematical formulation, that also incorporates a robustness term, should be a primary goal for future contributions. A large number of publications from the traditional computer vision literature deals with learning-based feature definition/detection. To achieve more robustness in terms of lung diseases this could be the way to go.

Apart from the feature extraction we would also improve the feature matching by tiling the images into sub-volumes and extracting a certain number of features in each of these sub-volumes. This way it could be guaranteed that we have an even distribution of feature for the `TPS` registration step. However, one has to keep in mind that the nonlinear deformations might move features from one sub-volume to the other and that the matching cost threshold determining a valid match will have to be reduced, thereby leading to more outliers. An idea for outlier removal is the incorporation of a robust median filtering procedure on the displacement field vectors, which would remove those correspondences that do not fit into the overall thoracic movement model. Finally, the `TPS` approximation is still the bottle-neck of the method, this could be improved by using e.g. a moment-based method for fast radial basis function approximation [Beatson and Newsam, 1998], which would improve the displacement field calculation by an order of magnitude.

The hybrid landmark and intensity-based registration approach is very appealing from a theoretical point of view. Obviously its most important drawback is the larger computation time compared to other methods. Since we are solving a large system of partial differential equations and need an outer iteration stage to linearize the inherent nonlinearity of the energy functional, there is the need to come up with more efficient schemes. One idea would be similar to the ideas used in the Demons [Thirion, 1998] algorithm, where the homogeneous optical flow equation system is approximated by an alternating scheme of optical flow constraint equation calculation and Gauss-filtering to perform the displacement field regularization. Since the Gauss-filter approximates the Laplacian, this is a valid assumption [Bro-Nielsen and Gramkow, 1996]. In our case we use an anisotropic image-driven regularization, therefore an obvious way for approximation would be an anisotropic diffusion filtering of the displacement field [Perona and Malik, 1990, Weickert et al., 1998]. Extensions of these methods on vector fields have been shown e.g. in [Weickert and Brox, 2002]. Another important research direction is the multi-grid method [Briggs et al., 2000], more specifically a Full Multi-Grid Cycle with pre- and post-smoothing, for efficiently solving the elliptic partial differential equation system. This method is the current state-of-the-art in solving `PDE` systems, however, its implementation is far from straight-forward.

Another important concept that has to be researched further is the incorporation of more robust similarity measures into the intensity-based registration framework. The `SSD` measure is not very suitable for non-Gaussian noise cases and if outliers have to be taken into account. In our case partial volume effect due to large breathing differences, contrast agent application and lung disease lead to geometrical and intensity outliers that currently are not modeled in our approach. A promising research direction for this problem is to use robust M-estimators on the terms (see [Hellier et al., 2001] for details).

Concerning the evaluation framework we would like to include more types of synthetic deformations. One method that is especially interesting is the finite-element method based approach from [Schnabel et al., 2003]. We also think of using our interactive segmentation refinement tools [Bornik et al., 2004] to generate simulated deformations. Experimenting with the framework's capability to add another dimension, different parameterizations of the algorithms that are investigated, seems to be another promising direction, however, one starts to run into problems due to the combinatorial explosion of necessary algorithm runs, distributed grid-based computer networks would be a possible way to still compute results in acceptable times (see [Glatard et al., 2006]). Incorporating further similarity measures as well as nonlinear registration methods obviously would contribute to the significance of evaluation studies. This is a task where feedback from the research community would be most welcome. In combination with the open-source Insight Segmentation and Registration Toolkit [ITK, 2006], one could think of a web-based version of our evaluation framework, where open-source data sets (like ones taken from the NLM database [NLM, 2006]) could be transformed with pre-defined, adjustable synthetic deformations, and one could upload his nonlinear registration algorithm and perform the tests comparing it with other methods and with the known ground-truth data. Such an effort would need a community that supports this idea, similar to the stereo evaluation project from Middlesbury College [Scharstein and Szeliski, 2002].

Appendix A

Acronyms and Symbols

List of Acronyms

`BFGS`	Broyden-Fletcher-Goldfarb-Shanno
`CT`	X-ray Computed Tomography
`CTA`	CT Angiography
`DoG`	Difference of Gaussian
`DSA`	Digital Subtraction Angiography
`EBS`	Elastic Body Spline
`FRC`	Functional Residual Capacity
`GEBS`	Gaussian Elastic Body Spline
`HU`	Hounsfield Unit
`LRTPS`	locally restricted thin-plate-spline
`MIP`	Maximum Intensity Projection
`MI`	Mutual Information
`MMBE`	Matched Mask Bone Elimination
`MRI`	Magnetic Resonance Imaging
`MSCT`	Multi-slice Spiral CT
`NCC`	Normalized Cross Correlation
`NMI`	Normalized Mutual Information
`PDE`	partial differential equation
`PDF`	Probability Distribution Function
`PET`	Positron Emission Tomography
`POI`	point(s) of interest
`RBF`	radial basis function
`RMS`	root-mean-square

`RVOE`	relative volume overlapping error
`SNR`	Signal to Noise Ratio
`SPECT`	Single Photon Emission Computed Tomography
`SAD`	Sum of Absolute Differences
`SIFT`	Scale-Invariant Feature Transform
`SSD`	Sum of Squared Differences
`SVD`	Singular Value Decomposition
`TLC`	Total Lung Capacity
`TRE`	target registration error
`TPS`	Thin-Plate-Spline
`US`	Ultrasound

Appendix B

Implementation Details

B.1 Thin-Plate-Spline Interpolation/Approximation

The `TPS` interpolation is a method for nonlinear function interpolation from a set of given points. It is a multi-dimensional generalization of the one-dimensional cubic splines, which can be used for function interpolation using an arbitrary number of data points. Cubic splines are low-order representations that are joined together at the interpolation points and twice continuously differentiable everywhere, thus preventing the interpolation function from oscillating like high-order polynomials would do. `TPS` interpolation now generalizes this concept in two or more dimensions by decomposing the multi-dimensional deformations into several coupled one-dimensional interpolation problems.

[Bookstein, 1989] was the first who proposed `TPS` for point-based image registration. `TPS` interpolation can be seen as the problem of finding a continuous transformation $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ that minimizes a given energy functional subject to the interpolation conditions (assuming the number of landmark correspondences as n)

$$\mathbf{x}_{F,i} = \varphi\left(\mathbf{x}_{M,i}\right), \quad i = 1, \cdots, n$$

i.e. the transformed points of one image have to be located *exactly* at the corresponding points of the other image. This condition clearly resembles the step from estimating a dense displacement field from sparse correspondences. In addition to the interpolation conditions, the choice of the energy functional also defines the behavior of this estimation. To arrive at the `TPS` interpolation scheme this energy functional represents the bending energy of a thin plate separately for each dimension component of φ. In the following we will assume $d = 3$, i.e. 3D `TPS` interpolation. The problem now can be decomposed into three sub-problems according to the components of φ, $(u(\mathbf{x}), v(\mathbf{x}), w(\mathbf{x}))^T$. For the first sub-problem in u, the energy functional reads

$$\mathcal{J}[u] = \iiint\limits_{\mathbb{R}^3} u_{xx}^2 + u_{yy}^2 + u_{zz}^2 + 2\left(u_{xy}^2 + u_{xz}^2 + u_{yz}^2\right) \mathrm{dxdydz}$$

with the other two being defined equivalently. Under some very general conditions concerning the minimum number of landmark correspondences and uniqueness of least squares regression of the nullspace of the energy functional [Duchon, 1976, Wahba, 1990], the minimization problem has a unique solution that can be calculated *analytically*

$$u(\mathbf{x}) = a_0 + a_1 x + a_2 y + a_3 z + \sum_{i=1}^{n} w_i U\left(\mathbf{x}, \mathbf{x}_{M,i}\right).$$

Here a_0, a_1, a_2, a_3, w_i is a set of weights and $U(r)$ is the `TPS` basis function which is defined as $U(r) = -\frac{1}{8\pi} r$, with r the distance of a given point to a landmark point. If one summarizes the weights by $\mathbf{a} = (a_0, \cdots, a_3)^T$ and $\mathbf{w} = (w_0, \cdots, w_{n-1})^T$ their derivation can be posed as the solution of the linear equation system

$$\begin{aligned} \mathbf{Pa} + \mathbf{Kw} &= \mathbf{v} \\ \mathbf{P}^T \mathbf{w} &= \mathbf{0} \end{aligned}$$

where $\mathbf{K} = (K_{ij}), K_{ij} = U(r_i)$, $\mathbf{P} = (P_{ij}), P_{ij} = \phi_j(\mathbf{x}_{M,i})$ and $\mathbf{v}$ is the column vector of one component of the coordinates of the target landmark correspondences $\mathbf{x}_{F,i}$. ϕ is a compact representation of the vector $(1, x, y, z)^T$. This linear equation system can be solved by standard methods, e.g. using an LU decomposition [Press et al., 1992]. For more information on `TPS` interpolation the reader may refer to [Rohr, 2001], where the relation of this functional to elasticity theory and to the Navier equation are given.

B.1.1 Approximating Thin-Plate-Splines using Landmark Errors

A generalization of the interpolating `TPS` weakens the hard interpolation constraint to take into account localization and matching errors that occur in practice during landmark localization [Rohr, 2001]. The approximation condition now reads

$$\mathbf{x}_{F,i} \approx \varphi\left(\mathbf{x}_{M,i}\right), \quad i = 1, \cdots, n$$

A similar derivation compared to the interpolating case leads to an energy functional that consists of s `SSD` term of the landmark correspondences and the same regularization term as above. Interestingly this regularization problem also has a direct, *analytical* solution that uses the same basis functions as the interpolating `TPS`. The final linear equation system is given as

$$\begin{aligned}\mathbf{P}\mathbf{a} + \left(\mathbf{K} + n\lambda\mathbf{W}^{-1}\right)\mathbf{w} &= \mathbf{v}\\ \mathbf{P}^T\mathbf{w} &= \mathbf{0}\end{aligned}$$

where λ is the global regularization weight and $\mathbf{W}^{-1}$ denotes the diagonal matrix of variances representing the landmark localization/matching uncertainties. The whole system is very similar to the equation system from the interpolating `TPS` scheme. An additional benefit is, that the uncertainty term makes the system more numerically stable, since the diagonal entries of $\mathbf{K}$ now become non-zero. With this derivation it is possible to prevent folding of the estimated dense displacement field, thus better approximating a diffeomorphic mapping.

B.2 B-spline Deformable Registration

The B-spline nonlinear registration method defines the transformation $\varphi : \mathbb{R}^d \to \mathbb{R}^d$, which relates fixed and moving images $I_F(\mathbf{x})$,$I_M(\mathbf{x})$ as

$$\varphi(\mathbf{x}_F) = \varphi_{global}(\mathbf{x}) + \varphi_{local}(\mathbf{x})$$

The global transformation is parameterized as a 12 degree of freedom affine transformation. The local motion model is described by means of a free form deformation model based on the tensor product of the one-dimensional cubic B-splines. An underlying mesh of uniformly-arranged control points allows to deform the moving image. Denoting the image domain $\Omega = \{(x, y, z)|0 \leq x < X, 0 \leq y < Y, 0 \leq z < Z\}$, a grid resolution $n_x \times n_y \times n_y$, a uniform grid spacing δ and the control points $\phi_{i,j,k}$, the definition of a free form deformation is:

$$\varphi_{local}(x, y, z) = \sum_{l=0}^{3}\sum_{m=0}^{3}\sum_{n=0}^{3} B_l(u)B_m(v)B_n(w)\phi_{i+l,j+m,k+n}$$

with

$$\begin{gathered} i = \left\lfloor \tfrac{x}{n_x} \right\rfloor - 1,\ j = \left\lfloor \tfrac{y}{n_y} \right\rfloor - 1,\ k = \left\lfloor \tfrac{z}{n_z} \right\rfloor - 1,\\ u = \tfrac{x}{n_x} - \left\lfloor \tfrac{x}{n_x} \right\rfloor,\ v = \tfrac{y}{n_y} - \left\lfloor \tfrac{y}{n_y} \right\rfloor,\ w = \tfrac{z}{n_z} - \left\lfloor \tfrac{z}{n_z} \right\rfloor \end{gathered}$$

and where $B_l(u), B_m(v), B_n(u)$ represent the {l,m,n}-th basis functions of the B-splines

$$B_0(u) = \tfrac{(1-u)^3}{6}$$

$$\begin{aligned} B_1(u) &= \tfrac{3u^3-6u^2+4}{6} \\ B_2(u) &= \tfrac{-3u^3+3u^2+3u+1}{6} \\ B_3(u) &= \tfrac{u^3}{6} \end{aligned}$$

The optimization algorithm has to minimize the free parameters of the free form deformation model. In order to make optimization more time-efficient and robust to local minima, a multi-resolution approach is used. The resolution of control points is increased in each multi-resolution step. Optimization starts with a low resolution to rapidly obtain a rough estimate of the overall local deformation. For each resolution an increasing number of control points is defined to refine the estimated deformation from the previous level. The resulting additional transformation parameters are calculated by a B-spline subdivision algorithm. For each of the L resolutions a new φ^l_{local} is calculated. The overall local transformation is defined as the sum of the individual transformations:

$$\varphi_{local}(\mathbf{x}) = \sum_{l=1}^{L} \varphi^l_{local}(\mathbf{x})$$

The similarity measure of the algorithm is the normalized mutual information as presented in Section 2.1.2.3. [Rueckert et al., 1999] propose a simple gradient descent technique for the optimization of the similarity cost functions for the local and global transformation models, respectively. However a far more suitable optimization method in case of a large number of free parameters is the Broyden-Fletcher-Goldfarb-Shanno (`BFGS`) approach [Nocedal and Wright, 1999], which is a quasi-Newton optimization technique. The pseudo code in Algorithm 3 outlines the main steps of the deformable B-spline algorithm.

Algorithm 3 B-spline Nonlinear Registration Algorithm

```
calculate the optimal affine transformation parameters $\Theta$
initialize the control points $\phi$
repeat
  calculate the gradient vector of the cost function $C$ with respect to the nonrigid
  transformation parameters $\phi$:
      $\nabla C = \frac{\partial(\Theta,\phi^l)}{\partial\phi^l}$
  while $\|\nabla C\| > \epsilon$ do
    recalculate the control points $\phi = \phi + \mu\frac{\nabla C}{\|\nabla C\|}$
    recalculate the the gradient vector $\|\nabla C\|$
  end while
  increase the control point resolution $\phi^l \rightarrow \phi^{l+1}$
  increase the image resolution
until finest level of resolution is reached
```

A disadvantage of nonlinear registration using free form deformations is the high computational effort, mainly due to the need for optimization of a large number of parameters. In [Klein et al., 2005] an improvement is proposed and the influence of different gradient estimation techniques which are needed to find a search direction in the optimization process is investigated. Three different gradient estimation methods are compared and its effects on the gradient descent optimization strategy are shown. The most promising method is the one using an analytic approach for cost function gradient derivation. Mutual information is assumed a continuous and differentiable function, thus the method by [Thevenaz and Unser, 2000] can be used for computing the gradient vector. Another important improvement proposed by [Klein et al., 2005] is a random selection of samples for the mutual information computation.

B.3 Elastic Intensity-Based Deformable Registration

We describe the elastic registration scheme [Bajcsy and Kovacic, 1989, Modersitzki, 2004] according to the variational framework from Section 1.4. In terms of $\mathcal{J}[\varphi]$, the similarity metric used is the `SSD` measure:

$$\mathcal{D}^{SSD}[I_F, I_M; \varphi] = \iiint\limits_{\Omega} (I_F(\mathbf{x}) - I_M(\varphi(\mathbf{x})))^2 \mathrm{d}\mathbf{x}.$$

In the elastic registration scheme, the regularization term is

$$\mathcal{S}[\varphi]^E := \iiint\limits_{\Omega} \frac{\lambda}{2}(u_x + v_y + w_z)^2 + \mu\left\{(u_x)^2 + (v_y)^2 + (w_z)^2\right\} + \frac{\mu}{2}\left\{(v_x + u_y)^2 + (w_x + u_z)^2 + (w_y + v_z)^2\right\} \mathrm{d}\mathbf{x}.$$

The derivation of a solution algorithm for this variational problem now involves the Gâteaux derivative of the functional $\mathcal{J}[\varphi]$. A necessary condition for a minimizer φ of $\mathcal{J}[\varphi]$ is that the Gâteaux derivative of $\mathcal{J}[\varphi]$ vanishes for all suitable perturbations. This analysis of the first variation of $\mathcal{J}[\varphi]$ leads to the Euler-Lagrange equations, a system of partial differential equations. Its numerical solution determines a minimizer φ. The Euler-Lagrange equations of the elastic regularization term are also known as the Navier-Lamé equations

$$f = \mu\nabla^2\varphi + (\lambda + \mu)\nabla \mathrm{div}\varphi$$

where f is an external force working against the internal elasticity constraints. For this specific regularization term and the `SSD` similarity measure the final system of Euler-Lagrange equations reads:

$$1.\quad 2\left(I_F(\mathbf{x}) - I_M\left(\varphi\left(\mathbf{x}\right)\right)\right)\frac{\partial I_M}{\partial x} + \left(\lambda + 2\mu\right)u_{xx} + \mu u_{yy} + \mu u_{zz} + \left(\lambda + \mu\right)\left(v_{xy} + w_{xz}\right) = 0$$

$$2.\quad 2\left(I_F(\mathbf{x}) - I_M\left(\varphi\left(\mathbf{x}\right)\right)\right)\frac{\partial I_M}{\partial y} + \mu v_{xx} + \left(\lambda + 2\mu\right)v_{yy} + \mu v_{zz} + \left(\lambda + \mu\right)\left(u_{xy} + w_{yz}\right) = 0$$

$$3.\quad 2\left(I_F(\mathbf{x}) - I_M\left(\varphi\left(\mathbf{x}\right)\right)\right)\frac{\partial I_M}{\partial z} + \mu w_{xx} + \mu w_{yy} + \left(\lambda + 2\mu\right)w_{zz} + \left(\lambda + \mu\right)\left(u_{xz} + v_{yz}\right) = 0$$

For the numerical solution of this functional a semi-implicit fixed-point approach can be used. A Taylor approximation neglecting the second- and the higher order terms linearizes the nonlinearity $I_M\left(\varphi\left(\mathbf{x}\right)\right)$ leading to a number of outer iterations to approximate the equation system. Each iteration consists of an iterative relaxation scheme (e.g. Gauss-Seidel or Successive Over-Relaxation (`SOR`)) to solve the resulting linear equation system with all of the displacements on the image grid as its unknown. The discretization of the `PDE` is commonly performed using finite differences. Other numerical solution methods for this approach include the finite-element model [Gee et al., 1994] or an `FFT` technique [Fischer and Modersitzki, 1999].

B.4 Demons Registration

The "Demons" registration approach [Thirion, 1998] has its origin in the work on deformable model matching. The novel idea of this method is that instead of attracting a deformable model defined in the moving image to a given fixed image by internal and external forces, it provides a regular grid of "Demons" (in accordance with Maxwell's demons introduced in physics) defined on the fixed image that pushes the iso-intensity surfaces of the moving image in the normal direction of the iso-intensity surfaces of the fixed image. The direction of these pushes depends on the current grid point being inside or outside the model iso-intensity surface. Intuitively, this tends to push the content of the moving image inside the object shape, and rejects the background of the moving image outside the object shape, where an object is regarded as an iso-intensity surface (but may also be a segmented surface).

A 3D grid of "Demons", which can be used for nonlinear intensity registration, is composed of the iso-intensity surfaces that pass through each voxel where the norm of the intensity gradient does not vanish. In that case we have an oriented normal $\nabla I_F(\mathbf{x})$. The demon in each voxel $\mathbf{x}$ pushes the moving image according to $\nabla I_F(\mathbf{x})$ if $I_F(\mathbf{x}) < I_M(\mathbf{x})$ and according to $-\nabla I_F(\mathbf{x})$ if $I_F(\mathbf{x}) > I_M(\mathbf{x})$. The standard demons algorithm now uses the following transformation update equation derived from the optical flow constraint

$$\varphi_{update} = \frac{I_M(\varphi(\mathbf{x})) - I_F(\mathbf{x})}{|\nabla I_F(\mathbf{x})|^2} \nabla I_F(\mathbf{x}).$$

One has to take care in situations when $\nabla I_F(\mathbf{x})$ vanishes, in this case the transformation update goes to infinity. Therefore a numerically stable version makes use of an additional stability term in the denumerator of the equation. In order to avoid an additional parameter this stability term is the square of the intensity difference leading to the following stable Demons update equation

$$\varphi_{update} = \frac{I_M(\varphi(\mathbf{x})) - I_F(\mathbf{x})}{|\nabla I_F(\mathbf{x})|^2 + (I_M(\varphi(\mathbf{x})) - I_F(\mathbf{x}))^2} \nabla I_F(\mathbf{x}).$$

The algorithm now iterates in the following way. Starting from an initial transformation φ_0 at each iteration the transformation update is calculated for every grid point $\mathbf{x}$, then a Gauss-filtering step is applied to the components of the deformation field to approximate the global regularization. In [Bro-Nielsen and Gramkow, 1996] it is proven that the Gauss-filter approximates the Laplacian of the displacement field, i.e. Tikhonov regularization.

In practice one often uses another formulation of the "Demons" algorithm, which is more symmetric in the role of the fixed and the moving image [Thirion, 1995]. We will refer to this method as "Symmetric Demons" in the following. The only difference compared to the original formulation is its transformation update equation, which reads

$$\varphi_{update} = \frac{2\left(I_M(\varphi(\mathbf{x})) - I_F(\mathbf{x})\right)}{\left(|\nabla I_F(\mathbf{x})| + |\nabla I_M(\mathbf{x})|\right)^2 + (I_M(\varphi(\mathbf{x})) - I_F(\mathbf{x}))^2} \left(\nabla I_F(\mathbf{x}) + \nabla I_M(\mathbf{x})\right).$$

Here one can easily see, that the gradients of fixed and moving image are considered concurrently, which makes the scheme more stable to outliers that appear only in one of the images. However, this advantage comes at the cost of higher computational and memory efforts in the implementation.

The relation of the "Demons" algorithm to standard optical-flow methods has been shown in [Modersitzki, 2004], further attempts to understand the relationships and implications of this method are given in [Bro-Nielsen and Gramkow, 1996, Pennec et al., 9999].

B.5 Curvature Registration

The curvature registration scheme [Fischer and Modersitzki, 2003b] is described according to the variational framework from Section 1.4. The similarity metric used is the `SSD` measure:

$$\mathcal{D}^{SSD}[I_F, I_M; \varphi] = \iiint\limits_{\Omega} (I_F(\mathbf{x}) - I_M(\varphi(\mathbf{x})))^2 \mathrm{d}\mathbf{x}.$$

The regularization term is defined as

$$\mathcal{S}[\varphi]^C := \iiint\limits_{\Omega} \frac{1}{2} \left(\left(\nabla^2 u\right)^2 + \left(\nabla^2 v\right)^2 + \left(\nabla^2 w\right)^2 \right) \mathrm{d}\mathbf{x}$$

and therefore quite similar to Tikhonov regularization, however, here the second derivatives of the components of the displacement field are penalized, instead of the first derivatives. This prevents oscillations and another interesting aspect is, that this regularizer has a non-trivial kernel that contains affine linear transformations. As a consequence, it does not penalize affine linear transformations and there is no theoretical need for affine pre-registration. Another interesting fact about curvature registration is its similarity to the `TPS` energy functional that describes the minimization of the bending energy of a function.

The resulting Euler-Lagrange equations incorporate the squared Laplacians of the displacement field components. This fourth order `PDE` system (a biharmonic equation) under Neumann boundary conditions can be solved by introducing an artificial time variable and computing the steady state solution in a semi-implicit scheme. The `PDE` is discretized using finite differences on the whole voxel grid, the system is only coupled via the external forces (similarity measure) and not via the internal energies, therefore it can be solved independently for each component. This leads to a fast and stable way of solution by discrete cosine transformations in an $O(N \log N)$ scheme. The final decoupled equation system reads

$$\begin{aligned} 1. \quad \left(\mathbf{I}_N + \tau\alpha A^2\right) u^{t+1} &= u^t + \tau F_u^t \\ 2. \quad \left(\mathbf{I}_N + \tau\alpha A^2\right) v^{t+1} &= v^t + \tau F_v^t \\ 3. \quad \left(\mathbf{I}_N + \tau\alpha A^2\right) w^{t+1} &= w^t + \tau F_w^t \end{aligned}$$

where τ is the time-step, t is the iteration variable that proceeds along the time axis towards infinity, $\mathbf{I}_N$ is an N-dimensional identity matrix, A is the system matrix derived from the finite difference approximation of the Laplacian on the voxel grid given Neumann boundary conditions and $F_l, \quad l \in \{u, v, w\}$ are the force components derived from the `SSD` measure.

B.6 Synthetic Simulated Breathing Transformation

The synthetic breathing transformation φ is simulated by applying a translational force to the diaphragm surface in the data sets negative z direction. A nonlinear force is calculated by weighting the constant translation $t_{vertical}$ with a two-dimensional Gaussian distribution that depends on the $x-$ and $y-$ coordinates of the data set, i.e. the further away from the center of the diaphragm surface a point is, the smaller is the negative z translation. Mathematically a displacement vector $\boldsymbol{d_1} = (0, 0, z')$ is applied to each point $(x, y, z)^T$ that maps it to $(x, y, z')^T$ with

$$z' = z - t_{vertical} e^{-\frac{(x-\mu_x)^2 + (y-\mu_y)^2}{2\sigma^2}}$$

where (μ_x, μ_y) corresponds to the $x-$ and $y-$ coordinates of the center of gravity of the diaphragm points and σ is chosen such that points lying at the exterior of the diaphragm surface (where the diaphragm is attached to the rib cage) nearly remain fixed.

In a similar fashion, simulation of rib cage behavior during breathing leads to the development of another displacement force component. A radial, center-directed translation t_{inward} is used to form a second displacement $\boldsymbol{d_2} = (x', y', 0)$ that maps points $(x, y, z)^T$ to $(x', y', z)^T$ with

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} + t' * \frac{\boldsymbol{c}}{|\boldsymbol{c}|}$$

where

$$\boldsymbol{c} = \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix} \qquad \textit{and} \qquad t' = |\boldsymbol{c}| - t_{inward} * (1 - e^{-\frac{(x-\mu_x)^2 + (y-\mu_y)^2}{2\sigma^2}})$$

Combining displacements $\boldsymbol{d_1}$ and $\boldsymbol{d_2}$ gives a total displacement $\boldsymbol{d}$ that is equivalent to a non-linear transformation $\varphi = \mathbf{x} + \boldsymbol{d}$.

Finally, to simulate a change in lung gray-values due to inhalation, all gray values smaller than -800 Hounsfield Units (HU) are increased by a random number drawn from a normal distribution centered at 25 HU with a standard deviation of 3 HU.

B.7 Grid-based Random Synthetic Transformation

This synthetic transformation is calculated using a number of evenly distributed landmark points and randomly assigning displacements to these landmarks. The amount of the displacement is increased towards the sampling grid inter-point distance of the landmark distribution. These displacements are not exactly physically motivated due to their regular placing, and the larger the assigned displacements are the harder it is to correct them. The dense synthetic displacement field is calculated using an `EBS` interpolation. This method goes back to [Davis et al., 1997], who proposed it as an alternative method to `TPS` interpolation (see Appendix B.1). The `EBS` method also uses a spline kernel transform framework, however, the kernel basis function is derived from the Navier partial differential equations describing elastic body deformations. [Davis et al., 1997] motivate their method with an improved realism for deforming tissues in medical imaging data sets compared to the `TPS` method. A drawback of this method is the increased computational effort, since the interpolation functions for three dimensions do not decouple as in the `TPS` case.

As an example for the grid-based synthetic transform, consider a data sets size of $256 \times 256 \times 256$. If we choose a grid size of 32, then every 32 voxels a grid point is placed. The maximal deviation is the second important parameter of this transformation, it resembles the maximal amount of deformation that may be calculated per grid point. The three-dimensional displacement vector is randomly drawn from a normal distribution between 0 and the maximal amount of deformation. The main motivation for this synthetic transformation is to determine the degree of deformation where algorithms are not capable to register the data anymore, in some sense this can be seen as an algorithms breakdown point.

Appendix C

Evaluation Framework Quantitative Results

			simbr10-5			**simbr25-10**		
Measure			nlm	sheep	human	nlm	sheep	human
RMS_{disp}	*initial*	[mm]	3.349	3.594	3.543	7.937	8.762	8.604
	symdemons	[mm]	3.202	2.313	1.958	6.710	7.130	5.368
	curvature	[mm]	2.975	2.063	0.872	5.266	3.810	3.696
	bspline	[mm]	1.598	4.566	4.166	1.607	2.122	8.617
	elastic	[mm]	1.623	1.656	1.690	3.771	4.156	4.139
	feature	[mm]	1.033	0.471	0.539	2.653	1.237	1.430
	opticflow	[mm]	1.930	1.630	1.418	3.312	3.871	3.559
	hybrid	[mm]	1.023	0.422	0.59	1.950	1.237	1.430
MAD_{disp}	*initial*	[mm]	1.185	1.219	1.209	2.479	2.826	2.752
	symdemons	[mm]	0.077	0.085	0.023	0.137	0.130	0.140
	curvature	[mm]	0.046	0.052	0.010	0.057	0.067	0.013
	bspline	[mm]	0.515	0.591	1.043	0.634	0.492	2.232
	elastic	[mm]	0.652	0.761	0.732	1.455	1.839	1.684
	feature	[mm]	0.072	0.060	0.024	0.092	0.084	0.041
	opticflow	[mm]	0.617	0.794	0.598	1.401	1.699	1.515
	hybrid	[mm]	0.054	0.054	0.023	0.047	0.059	0.035
MAX_{disp}	*initial*	[mm]	7.293	7.991	7.874	18.340	20.114	19.827
	symdemons	[mm]	4.546	3.664	1.959	12.501	8.153	7.156
	curvature	[mm]	5.169	4.093	0.443	11.502	5.848	2.787
	bspline	[mm]	2.762	6.343	5.800	2.373	2.673	12.708
	elastic	[mm]	3.137	3.094	3.206	6.950	7.901	8.078
	feature	[mm]	2.625	1.191	1.450	7.085	3.164	3.976
	opticflow	[mm]	2.967	3.177	2.861	6.606	8.045	7.096
	hybrid	[mm]	2.335	2.547	2.314	4.715	5.149	4.844

Table C.1: Simulated breathing transformation *simbr10-5* and *simbr25-10*. Registration results in terms of displacement field measures.

			simbr10-5			**simbr25-10**		
Measure			nlm	sheep	human	nlm	sheep	human
RMS_{int}	*affine*	[HU]	52.112	50.621	57.147	60.7764	59.357	63.981
	symdemons	[HU]	18.933	17.287	17.237	24.2617	26.529	29.335
	curvature	[HU]	15.339	10.118	10.413	19.6199	19.598	24.201
	bspline	[HU]	18.215	21.861	24.562	21.1404	22.313	34.013
	elastic	[HU]	30.271	27.410	36.925	39.4448	38.803	47.847
	feature	[HU]	19.304	13.506	16.551	27.3814	20.354	19.225
	opticflow	[HU]	18.914	20.102	26.103	25.2813	29.892	38.193
	hybrid	[HU]	14.408	11.240	12.772	16.3982	11.316	12.464
MAD_{int}	*affine*	[HU]	11	12	23	22	21	34
	symdemons	[HU]	1	1	1	1	1	1
	curvature	[HU]	1	1	1	1	1	1
	bspline	[HU]	1	2	3	2	2	4
	elastic	[HU]	5	4	10	6	7	16
	feature	[HU]	2	1	2	2	2	3
	opticflow	[HU]	3	3	8	4	6	12
	hybrid	[HU]	1	1	1	2	1	1
NMI_{int}	*affine*		0.293	0.339	0.250	0.212	0.242	0.182
	symdemons		0.777	0.882	0.860	0.764	0.823	0.901
	curvature		0.764	0.879	0.881	0.758	0.837	0.795
	bspline		0.705	0.726	0.653	0.671	0.739	0.619
	elastic		0.489	0.583	0.427	0.388	0.446	0.326
	feature		0.647	0.825	0.712	0.634	0.750	0.705
	opticflow		0.577	0.638	0.505	0.485	0.513	0.391
	hybrid		0.721	0.794	0.750	0.690	0.794	0.736
$EDGE_{int}$	*affine*		0.197	0.178	0.223	0.199	0.183	0.226
	symdemons		0.014	0.025	0.015	0.020	0.030	0.041
	curvature		0.015	0.031	0.012	0.016	0.033	0.014
	bspline		0.046	0.058	0.066	0.058	0.058	0.082
	elastic		0.120	0.117	0.166	0.151	0.144	0.197
	feature		0.072	0.048	0.052	0.077	0.061	0.067
	opticflow		0.092	0.105	0.140	0.108	0.123	0.170
	hybrid		0.041	0.036	0.047	0.069	0.052	0.070
MAX_{int}	*affine*	[HU]	602	515	555	835	846	871
	symdemons	[HU]	8	4	6	10	6	7
	curvature	[HU]	15	8	7	16	9	13
	bspline	[HU]	32	35	34	37	32	39
	elastic	[HU]	88	76	102	172	148	192
	feature	[HU]	53	20	28	74	42	39
	opticflow	[HU]	38	44	58	60	81	101
	hybrid	[HU]	18	22	33	26	21	39

Table C.2: Simulated breathing transformation *simbr10-5* and *simbr25-10*. Registration results in terms of intensity measures.

			simbr55-25			**simbr70-30**		
Measure			nlm	sheep	human	nlm	sheep	human
RMS_{disp}	*initial*	[mm]	18.408	20.173	19.897	23.280	25.727	25.349
	symdemons	[mm]	8.264	19.230	15.963	9.689	19.151	18.588
	curvature	[mm]	6.771	7.798	7.702	7.112	8.401	11.601
	bspline	[mm]	10.407	6.000	27.739	8.808	10.039	15.318
	elastic	[mm]	11.176	11.092	11.792	15.183	14.346	16.512
	feature	[mm]	7.889	5.302	6.748	18.163	12.975	12.393
	opticflow	[mm]	11.332	12.607	11.750	14.754	16.213	16.148
	hybrid	[mm]	7.202	4.849	5.817	12.350	10.974	10.711
MAD_{disp}	*initial*	[mm]	5.945	6.974	6.805	7.280	8.743	8.458
	symdemons	[mm]	0.215	0.590	0.233	0.674	0.894	1.117
	curvature	[mm]	0.075	0.104	0.019	0.094	0.118	0.023
	bspline	[mm]	2.242	1.713	11.87	1.973	3.490	6.290
	elastic	[mm]	4.738	4.763	4.736	6.128	6.127	6.167
	feature	[mm]	1.468	1.107	1.223	4.329	3.398	3.924
	opticflow	[mm]	4.504	4.913	4.474	5.859	6.201	5.703
	hybrid	[mm]	1.229	1.008	2.336	3.160	3.994	3.463
MAX_{disp}	*initial*	[mm]	41.399	45.119	44.639	52.978	57.687	57.133
	symdemons	[mm]	21.368	43.101	37.229	23.461	44.392	44.838
	curvature	[mm]	13.595	10.348	11.225	15.729	11.517	17.044
	bspline	[mm]	15.342	8.529	53.223	13.147	15.132	42.291
	elastic	[mm]	22.602	21.846	24.447	31.842	28.065	35.510
	feature	[mm]	21.012	13.864	12.582	29.726	28.441	26.362
	opticflow	[mm]	23.747	25.010	23.807	30.526	31.615	31.833
	hybrid	[mm]	15.142	17.525	17.205	25.639	24.732	21.341

Table C.3: Simulated breathing transformation *simbr55-25* and *simbr70-30*. Registration results in terms of displacement field measures.

			simbr55-25			**simbr70-30**		
Measure			nlm	sheep	human	nlm	sheep	human
RMS_{int}	*affine*	[HU]	73.827	70.898	72.889	77.131	73.868	75.375
	symdemons	[HU]	28.641	41.269	46.390	30.635	41.868	50.348
	curvature	[HU]	24.172	27.381	33.361	25.012	28.395	35.692
	bspline	[HU]	34.785	29.118	71.843	29.459	34.498	52.456
	elastic	[HU]	54.821	49.949	56.644	58.915	52.855	59.101
	feature	[HU]	37.758	24.000	29.485	48.573	38.746	45.461
	opticflow	[HU]	36.274	43.121	49.196	39.412	45.967	51.384
	hybrid	[HU]	27.118	16.956	22.305	32.383	27.935	26.306
MAD_{int}	*affine*	[HU]	76	62	67	95	87	87
	symdemons	[HU]	1	1	1	1	1	1
	curvature	[HU]	1	1	0	1	1	1
	bspline	[HU]	2	2	52	3	2	13
	elastic	[HU]	15	13	23	22	16	26
	feature	[HU]	3	2	4	7	4	12
	opticflow	[HU]	7	11	17	8	13	19
	hybrid	[HU]	2	2	2	6	7	9
NMI_{int}	*affine*		0.1315	0.1480	0.1129	0.1183	0.1289	0.0986
	symdemons		0.7114	0.6893	0.5742	0.6614	0.6794	0.5343
	curvature		0.6912	0.7614	0.6846	0.6501	0.7396	0.6570
	bspline		0.6365	0.7236	0.1646	0.6475	0.6787	0.4936
	elastic		0.2463	0.3142	0.2314	0.2123	0.2753	0.2047
	feature		0.6333	0.7202	0.6061	0.5181	0.5693	0.5569
	opticflow		0.3641	0.3651	0.2795	0.3249	0.3244	0.2539
	hybrid		0.6383	0.7345	0.6905	0.6803	0.6904	0.6906
$EDGE_{int}$	*affine*		0.1924	0.1792	0.2223	0.1892	0.1771	0.2208
	symdemons		0.0203	0.0444	0.0453	0.0330	0.0447	0.0616
	curvature		0.0182	0.0368	0.0195	0.0205	0.0381	0.1238
	bspline		0.0599	0.0642	0.2118	0.0618	0.0712	0.2177
	elastic		0.1774	0.1599	0.2103	0.1809	0.1621	0.2114
	feature		0.1048	0.1088	0.1326	0.1452	0.1567	0.1667
	opticflow		0.1286	0.1388	0.1860	0.1311	0.1412	0.1867
	hybrid		0.0761	0.0801	0.0876	0.1075	0.1068	0.1088
MAX_{int}	*affine*	[HU]	923	963	988	939	994	1013
	symdemons	[HU]	12	15	41	14	17	62
	curvature	[HU]	22	14	26	23	15	30
	bspline	[HU]	39	37	975	41	46	93
	elastic	[HU]	435	320	353	555	381	447
	feature	[HU]	235	126	151	414	192	233
	opticflow	[HU]	116	157	180	138	194	212
	hybrid	[HU]	31	41	40	33	48	49

Table C.4: Simulated breathing transformation *simbr55-25* and *simbr70-30*. Registration results in terms of intensity measures.

			grid-32-2			**grid-32-4**		
Measure			nlm	sheep	human	nlm	sheep	human
RMS_{disp}	*initial*	[mm]	1.850	1.772	1.723	3.517	3.244	3.090
	symdemons	[mm]	2.722	1.382	2.228	2.478	1.649	1.816
	curvature	[mm]	2.219	0.716	0.641	2.964	1.073	1.877
	bspline	[mm]	1.543	2.044	2.569	5.638	4.097	2.512
	elastic	[mm]	1.456	1.534	1.466	3.064	2.792	2.680
	feature	[mm]	1.751	1.671	1.503	3.510	3.047	2.815
	opticflow	[mm]	1.484	1.550	1.479	3.021	2.823	2.682
	hybrid	[mm]	1.304	1.410	1.361	2.836	2.614	2.450
MAD_{disp}	*initial*	[mm]	0.706	0.672	0.651	1.332	1.252	1.150
	symdemons	[mm]	0.062	0.059	0.028	0.091	0.078	0.037
	curvature	[mm]	0.022	0.022	0.005	0.028	0.025	0.007
	bspline	[mm]	0.385	0.658	0.608	1.199	1.072	0.832
	elastic	[mm]	0.527	0.527	0.514	1.078	0.991	0.941
	feature	[mm]	0.634	0.586	0.532	1.329	1.093	0.983
	opticflow	[mm]	0.550	0.546	0.524	1.060	1.018	0.965
	hybrid	[mm]	0.423	0.446	0.408	0.883	0.840	0.758
MAX_{disp}	*initial*	[mm]	3.439	3.395	3.268	6.795	6.353	5.892
	symdemons	[mm]	4.616	2.167	4.435	5.466	3.084	3.350
	curvature	[mm]	3.800	0.886	0.358	4.407	1.149	1.321
	bspline	[mm]	3.407	4.184	4.110	9.626	7.470	5.087
	elastic	[mm]	2.869	3.121	2.925	6.185	5.754	5.279
	feature	[mm]	3.485	3.377	2.968	6.937	6.312	5.613
	opticflow	[mm]	2.944	3.139	2.964	6.116	5.747	5.358
	hybrid	[mm]	2.702	2.940	2.856	5.981	5.561	5.055

Table C.5: Grid synthetic transformation *grid-32-2* and *grid-32-4*. Registration results in terms of displacement field measures.

			grid-32-2			**grid-32-4**		
Measure			nlm	sheep	human	nlm	sheep	human
RMS_{int}	*affine*	[HU]	42.492	40.722	47.046	48.793	48.405	53.232
	symdemons	[HU]	14.357	15.364	14.542	13.888	14.585	13.808
	curvature	[HU]	13.868	12.307	11.721	14.676	12.545	15.716
	bspline	[HU]	22.607	24.664	27.700	36.087	35.543	35.557
	elastic	[HU]	27.529	27.549	33.022	36.031	36.162	41.953
	feature	[HU]	20.123	18.553	21.520	27.505	26.332	29.633
	opticflow	[HU]	19.097	23.722	28.879	25.795	32.874	37.964
	hybrid	[HU]	15.068	16.248	19.561	20.092	20.450	21.015
MAD_{int}	*affine*	[HU]	6	6	15	9	10	19
	symdemons	[HU]	1	1	1	1	1	1
	curvature	[HU]	1	1	1	1	1	1
	bspline	[HU]	3	3	6	5	5	9
	elastic	[HU]	4	4	9	6	6	12
	feature	[HU]	2	2	3	2	2	4
	opticflow	[HU]	3	4	8	4	6	11
	hybrid	[HU]	2	2	3	2	2	3
NMI_{int}	*affine*		0.3797	0.4455	0.3365	0.3092	0.3500	0.2764
	symdemons		0.8085	0.8913	0.7816	0.8095	0.8855	0.8215
	curvature		0.7793	0.8784	0.8246	0.7750	0.8715	0.8270
	bspline		0.6186	0.6483	0.5616	0.4975	0.5440	0.4626
	elastic		0.5242	0.5903	0.4651	0.4365	0.4939	0.3896
	feature		0.6999	0.7859	0.6954	0.7343	0.8033	0.8143
	opticflow		0.5812	0.6142	0.4950	0.5044	0.5184	0.4228
	hybrid		0.7519	0.8195	0.7102	0.7529	0.8166	0.8315
$EDGE_{int}$	*affine*		0.1663	0.1523	0.1950	0.1844	0.1713	0.2158
	symdemons		0.0191	0.0260	0.0268	0.0214	0.0264	0.0223
	curvature		0.0202	0.0245	0.0130	0.0221	0.0248	0.0205
	bspline		0.0790	0.0945	0.1165	0.1182	0.1255	0.1623
	elastic		0.1111	0.1120	0.1498	0.1391	0.1374	0.1844
	feature		0.1594	0.1423	0.1801	0.1783	0.1694	0.2119
	opticflow		0.0928	0.1028	0.1393	0.1103	0.1237	0.1676
	hybrid		0.0758	0.0865	0.1153	0.0913	0.1062	0.1433
MAX_{int}	*affine*	[HU]	325	249	271	483	454	428
	symdemons	[HU]	8	4	10	8	4	9
	curvature	[HU]	15	8	8	16	8	14
	bspline	[HU]	47	57	52	93	105	97
	elastic	[HU]	72	71	85	120	122	135
	feature	[HU]	272	199	187	422	371	343
	opticflow	[HU]	37	53	63	58	87	98
	hybrid	[HU]	22	31	43	37	53	64

Table C.6: Grid synthetic transformation *grid-32-2* and *grid-32-4*. Registration results in terms of intensity measures.

Appendix D

List of Publications

In this appendix we give a list of relevant publications that arose from the work on this thesis:

Partially Rigid Bone Registration in CT Angiography

In this work we have presented a partially rigid bone registration algorithm for `CTA` images and evaluated this algorithm quantitatively and qualitatively on several head data-sets using `MIP` visualization. [Urschler et al., 2006b]

Matching 3D Lung Surfaces with the Shape Context Approach

This publication presented our first results on shape matching approaches based on a 3D extension of the shape context method. Segmented lung surfaces were matched automatically and compared to a synthetic ground-truth. [Urschler and Bischof, 2004a]

Registering 3D Lung Surfaces with the Shape Context Approach

The shape context matching approach was put into a `TPS` registration framework in this publication, which showed promising results on real and synthetic data sets. [Urschler and Bischof, 2004b]

Assessing breathing motion by shape matching of lung and diaphragm surfaces

In this work the shape context matching and registration was finalized and evaluated on a large number of different shapes extracted from `CT` data sets. These shapes were airway tree, diaphragm and lung surface segmentations. [Urschler and Bischof, 2005]

SIFT and Shape Context for Feature-Based Nonlinear Registration of Thoracic CT Images

A feature-based automatic nonlinear registration pipeline for thoracic `CT` data sets was presented that consisted of feature extraction, matching and `TPS` registration. Evaluation on real and synthetic data showed that feasibility of this approach. [Urschler et al., 2006a]

Automatic Point Landmark Matching for Regularizing Nonlinear Intensity Registration: Application to Thoracic CT Images

Tackling the drawbacks of the feature-based registration approach was the purpose of this publication. Herein the feature extraction and matching steps are used to determine correspondences, while a hybrid variational intensity registration incorporates an optical-flow based data term with an anisotropic image-driven regularization and a feature constraint term. [Urschler et al., 2006c]

List of Figures

List of Tables

Bibliography

[Alvarez et al., 2000] Alvarez, L., Weickert, J., and Sanchez, J. (2000). Reliable Estimation of Dense Optical Flow Fields with Large Displacements. *International Journal of Computer Vision*, 39(1):41–56.

[Alyassin and Avinash, 2001] Alyassin, A. and Avinash, G. (2001). Semi-Automatic Bone Removal Technique from CTA Angiography Data. In Sonka, M. and Hanson, K. M., editors, *Proc SPIE Conf on Medical Imaging: Image Processing*, volume 4322, pages 1273–1283.

[Angelillo et al., 2000] Angelillo, M., Boriek, A. M., Rodarte, J. R., and Wilson, T. A. (2000). Shape of the canine diaphragm. *Journal of Applied Physiology*, 89(5):15–20.

[Arsigny et al., 2005] Arsigny, V., Pennec, X., and Ayache, N. (2005). Polyrigid and polyaffine transformations: A novel geometrical tool to deal with non-rigid deformations - Application to the registration of histological slices. *Medical Image Analysis*, 9(6):507–523.

[Audette et al., 2000] Audette, M. A., Ferrie, F. P., and Peters, T. M. (2000). An algorithmic overview of surface registration techniques for medical imaging. *Medical Image Analysis*, 4(3):201–217.

[Avants et al., 2004] Avants, B., Sundaram, T., Duda, J. T., Gee, J. C., and Ng, L. (2004). Non-Rigid Image Registration. In Yoo, T. S., editor, *Insight into Images*, chapter 11, pages 307–348. A K Peters, Ltd.

[Bailey, 2001] Bailey, D. L. (2001). Image Registration in Nuclear Medicine. In Hajnal, J., Hill, D., and Hawkes, D. J., editors, *Medical Image Registration*, chapter 11, pages 233–252. CRC Press, Boca Raton.

[Bajcsy and Kovacic, 1989] Bajcsy, R. and Kovacic, S. (1989). Multiresolution Elastic Matching. *Computer Vision, Graphics and Image Processing*, 46(1):1–21.

[Bankman, 2000] Bankman, I. N., editor (2000). *Handbook of Medical Imaging: Processing and Analysis*. Academic Press Series in Biomedical Engineering.

[Barron et al., 1994] Barron, J. L., Fleet, D. J., and Beauchemin, S. S. (1994). Performance of Optical Flow Techniques. *International Journal of Computer Vision*, 12(1):43–77.

[Bauer, 2005] Bauer, J. (2005). Evaluation of Global Descriptors for Automatic Image Orientation. Technical report, VRVis Research Centre, Graz, Austria.

[Beatson and Newsam, 1998] Beatson, R. K. and Newsam, G. N. (1998). Fast evaluation of radial basis functions: Moment-based methods. *SIAM J. Sci. Computing*, 19(5):1428–1449.

[Beichel et al., 2004] Beichel, R., Pock, T., Janko, C., Zotter, R. B., Reitinger, B., Bornik, A., Palagyi, K., Sorantin, E., Werkgartner, G., Bischof, H., and Sonka, M. (2004). Liver segment approximation in CT data for surgical resection planning. In Fitzpatrick, J. M. and Sonka, M., editors, *Proc SPIE Conf on Medical Imaging: Image Processing*, volume 5370, pages 1435–1446.

[Belongie et al., 2002] Belongie, S., Malik, J., and Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):509–522.

[Bentoutou et al., 2002] Bentoutou, Y., Taleb, N., Chikr El Mezouar, M., Taleb, M., and Jetto, L. (2002). An invariant approach for image registration in digital subtraction angiography. *Pattern Recognition*, 35(12):2853–2865.

[Besl and McKay, 1992] Besl, P. J. and McKay, N. D. (1992). A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256.

[Betke et al., 2003] Betke, M., Hong, H., Thomas, D., Prince, C., and Ko, J. P. (2003). Landmark detection in the chest and registration of lung surfaces with an application to nodule registration. *Medical Image Analysis*, 7(3):265–281.

[Beutel et al., 2000] Beutel, J., Kundel, H. L., and Van Metter, R. L., editors (2000). *Handbook of Medical Imaging: Volume 1 - Physics and Psychophysics*. SPIE Press.

[Blackall et al., 2004] Blackall, J. M., Penney, G. P., King, A. P., Adam, A. N., and Hawkes, D. J. (2004). Tracking alignment of sparse ultrasound with preoperative images of the liver and an interventional plan using models of respiratory motion and deformation. In *Proc SPIE Conf on Medical Imaging: Visualization, Image-Guided Procedures, and Display*, volume 5637, pages 218–227.

[Boehm et al., 2000] Boehm, D., Krass, S., Kriete, A., Rau, W., Selle, D., Jend, H.-H., and Peitgen, H.-O. (2000). Segmentbestimmung im Computertomogramm der Lunge. In *Bildverarbeitung fuer die Medizin*, pages 168–172. Springer.

[Bookstein, 1989] Bookstein, F. L. (1989). Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585.

[Boriek et al., 1998] Boriek, A. M., Miller III, C. C., and Rodarte, J. R. (1998). Muscle fiber architecture of the dog diaphragm. *Journal of Applied Physiology*, 84(1):318–326.

[Bornik et al., 2003] Bornik, A., Beichel, R., Reitinger, B., Gotschuli, G., Sorantin, E., Leberl, F. W., and Sonka, M. (2003). Computer Aided Liver Surgery Planning: An Augmented Reality Approach. In Galloway, R. L., editor, *Proc SPIE Conf on Medical Imaging: Visualization, Image-Guided Procedures, and Display*, volume 5029, pages 395–406.

[Bornik et al., 2004] Bornik, A., Reitinger, B., Beichel, R., Sorantin, E., and Werkgartner, G. (2004). Augmented-reality-based segmentation refinement. In Galloway, R. L., editor, *Proc SPIE Conf on Medical Imaging: Visualization, Image-Guided Procedures, and Display*, volume 5367, pages 77–87.

[Briggs et al., 2000] Briggs, W. L., Henson, V. E., and MacCormick, S. F. (2000). *A multigrid tutorial.* SIAM - Society for Industrial and Applied Mathematics, 2nd edition.

[Bro-Nielsen and Gramkow, 1996] Bro-Nielsen, M. and Gramkow, C. (1996). Fast fluid registration of medical images. In *Proc. 4th Int. Conf. on Visualization in Biomedical Computing*, volume 1131 of *LNCS*, pages 267–276. Springer.

[Broit, 1981] Broit, C. (1981). *Optimal Registration of Deformed Images.* Doctoral dissertation, University of Pennsylvania, Philadelphia.

[Brown, 1992] Brown, L. G. (1992). A Survey of Image Registration Techniques. *ACM Computing Surveys*, 24(4):325–376.

[Cai et al., 1999] Cai, J., Chu, J. C. H., Recine, D., Sharma, M., Nguyen, C., Rodebaugh, R., Saxena, V. A., and Ali, A. (1999). CT and PET lung image registration and fusion in radiotherapy treatment planning using the chamfer-matching method. *Int J Radiation Oncology Biology Physics*, 43(4):883–891.

[Canny, 1986] Canny, J. (1986). A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698.

[Christensen and Johnson, 2001] Christensen, G. E. and Johnson, H. J. (2001). Consistent Image Registration. *IEEE Transactions on Medical Imaging*, 20(7):568–582.

[Christensen et al., 1996] Christensen, G. E., Rabbitt, R. D., and Miller, M. I. (1996). Deformable Templates Using Large Deformation Kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447.

[Chui and Rangarajan, 2003] Chui, H. and Rangarajan, A. (2003). A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3):114–141.

[Chui et al., 2003] Chui, H., Win, L., Schultz, R., Duncan, J. S., and Rangarajan, A. (2003). A unified non-rigid feature registration method for brain mapping. *Medical Image Analysis*, 7(2):113–130.

[Clatz et al., 2005] Clatz, O., Delingette, H., Talos, I.-F., Golby, A. J., Jolesz, F. A., Ayache, N., and Warfield, S. K. (2005). Robust Non-Rigid Registration to Capture Brain

Shift from Intra-Operative MRI. *IEEE Transactions on Medical Imaging*, 24(11):1417–1427.

[Cohen, 1991] Cohen, L. (1991). On Active Contour Models and Balloons. *Computer Vision, Graphics and Image Processing: Image Understanding*, 53(2):211–218.

[Cootes and Taylor, 2001] Cootes, T. F. and Taylor, C. J. (2001). Statistical models of appearance for medical image analysis and computer vision. In *Proc SPIE Conf on Medical Imaging: Image Processing*, volume 4322, pages 236–248.

[Crum et al., 2003] Crum, W. R., Griffin, L. D., Hill, D. L. G., and Hawkes, D. J. (2003). Zen and the Art of Medical Image Registration: Correspondence, Homology, and Quality. *NeuroImage*, 20(3):1425–1437.

[Crum et al., 2004] Crum, W. R., Hartkens, T., and Hill, D. L. G. (2004). Non-rigid image registration: theory and practice. *The British Journal of Radiology - Imaging Processing Special Issue*, 77:S140–S153.

[Davatzikos, 1996] Davatzikos, C. (1996). Image Registration Based on Boundary Mapping. *IEEE Transactions on Medical Imaging*, 15(1):112–115.

[Davies et al., 1994] Davies, S. C., Hill, A. L., Holmes, R. B., Halliwell, M., and Jackson, P. (1994). Ultrasound Quantitation of Respiratory Organ Motion in the Upper Abdomen. *British Journal of Radiology*, 67(11):1096–1102.

[Davis et al., 1997] Davis, M. H., Khotanzad, A., Flamig, D. P., and Harms, S. E. (1997). A Physics-Based Coordinate Transform for 3-D Image Matching. *IEEE Transactions on Medical Imaging*, 16(3):317–328.

[Delingette, 1999] Delingette, H. (1999). General Object Reconstruction Based on Simplex Meshes. *International Journal of Computer Vision*, 32(2):111–146.

[Denton et al., 1999] Denton, E. R. E., Sonoda, L. I., Rueckert, D., Rankin, S. C., Hayes, C., Leach, M., Hill, D. L. G., and Hawkes, D. J. (1999). Comparison and evaluation of rigid and non-rigid registration of breast MR images. *Journal of Computer Assisted Tomography*, 23:800–805.

[Ditt et al., 2003] Ditt, H., Klotz, E., Raupach, R., Tomandl, B. F., Koenig, M., Schramm, P., and Kloska, S. P. (2003). 3D-imaging of perfused cerebral blood volume in acute stroke using multislice CT. In *Proc European Congress of Radiology*, volume 13, page 515, Vienna.

[Donato and Belongie, 2002] Donato, G. and Belongie, S. (2002). Approximate Thin Plate Spline Mappings. In *Proc European Conf on Computer Vision (ECCV)*, volume 2352 of *LNCS*, pages 21–31. Springer.

[Dougherty et al., 2003] Dougherty, L., Asmuth, J. C., and Gefter, W. B. (2003). Alignment of CT Lung Volumes with an Optical Flow Method. *Academic Radiology*, 10(3):249–254.

[Dougherty et al., 2006] Dougherty, L., Torigian, D. A., Affusso, J. D., Asmuth, J. C., and Gefter, W. B. (2006). Use of an Optical Flow Method for the Analysis of Serial CT Lung Images. *Academic Radiology*, 13(1):14–23.

[Duchon, 1976] Duchon, J. (1976). Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces. *R.A.I.R.O. Analyse Numerique*, 10(12):5–12.

[Edwards et al., 1998] Edwards, P. J., Hill, D. L. G., Little, J. A., and Hawkes, D. J. (1998). A Three-Component Deformation Model for Image-Guided Surgery. *Medical Image Analysis*, 2(4):355–367.

[Fan and Chen, 1999a] Fan, L. and Chen, C. W. (1999a). 3D Warping and Registration from Lung Images. In *Proc SPIE Conf on Medical Imaging: Physiology and Function from Multidimensional Images*, volume 3660, pages 459–470.

[Fan and Chen, 1999b] Fan, L. and Chen, C. W. (1999b). Integrated Approach to 3D Warping and Registration from Lung Images. In *Proc of SPIE Conf on Developments in X-Ray Tomography II*, volume 3772, pages 24–35, Denver, CO.

[Fan et al., 2001] Fan, L., Chen, C. W., Reinhardt, J. M., and Hoffman, E. A. (2001). Evaluation and Application of 3D Lung Warping and Registration Model Using HRCT Images. In *Proc SPIE Conf on Medical Imaging: Physiology and Function from Multidimensional Images*, volume 4321, pages 234–243.

[Farid and Simoncelli, 1997] Farid, H. and Simoncelli, E. P. (1997). Optimally Rotation-Equivariant Directional Derivative Kernels. In *Proc. of the 7th Int'l Conf Computer Analysis of Images and Patterns*, pages 207–214, Kiel, Germany. Springer.

[Felkel et al., 2001] Felkel, P., Wegenkittl, R., and Kanitsar, A. (2001). Vessel Tracking in Peripheral CTA Datasets - An Overview. In *IEEE Spring Conference on Computer Graphics (SCCG 2001)*, pages 232–239.

[Fischer and Modersitzki, 1999] Fischer, B. and Modersitzki, J. (1999). Fast inversion of matrices arising in image processing. *Numerical Algorithms*, 22(1):1–11.

[Fischer and Modersitzki, 2003a] Fischer, B. and Modersitzki, J. (2003a). Combination of automatic non-rigid and landmark based registration: the best of both worlds. In *Proc SPIE Conf on Medical Imaging: Image Processing*, volume 5032, pages 1037–1048.

[Fischer and Modersitzki, 2003b] Fischer, B. and Modersitzki, J. (2003b). Curvature based image registration. *Journal of Mathematical Imaging and Vision*, 18(1):81–85.

[Förstner, 1986] Förstner, W. (1986). A Feature Based Correspondence Algorithm for Image Matching. *Int. Arch. of Photogrammetry and Remote Sensing*, 26(3):150–166.

[Frome et al., 2004] Frome, A., Huber, D., Kolluri, R., Buelow, T., and Malik, J. (2004). Recognizing Objects in Range Data Using Regional Point Descriptors. In *Proc European Conf on Computer Vision (ECCV)*, volume 3023 of *LNCS*, pages 224–237. Springer.

[Fua, 1993] Fua, P. (1993). A Parallel Stereo Algorithm that Produces Dense Depth Maps and Preserves Image Features. *Machine Vision and Applications*, 6(1):35–49.

[Gee, 1996] Gee, J. C. (1996). *Probabilistic Matching of Deformed Images.* PhD thesis, Univ. Pennsylvania, Philadelphia.

[Gee et al., 1994] Gee, J. C., Haynor, D. R., Reikvich, M., and Bajcsy, R. (1994). Finite element approach to warping of brain images. In *Proc SPIE Conf on Medical Imaging: Image Processing*, volume 2167, pages 18–27.

[Gefter, 2002] Gefter, W. B. (2002). Functional CT Imaging of the Lungs: The pulmonary function test of the new millenium? *Academic Radiology*, 9(2):127–129.

[Glatard et al., 2006] Glatard, T., Pennec, X., and Montagnat, J. (2006). Performance Evaluation of Grid-Enabled Registration Algorithms Using Bronze-Standards. In Larsen, R., Nielsen, M., and Sporring, J., editors, *Proc Intern Conf on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 4191 of *LNCS*, Copenhagen, Denmark. Springer.

[Guo et al., 2004] Guo, H., Rangarajan, A., and Joshi, S. (2004). Non-Rigid Registration of Shapes Via Diffeomorphic Point Matching. In *Proceedings of the 2004 IEEE International Symposium on Biomedical Imaging (ISBI): From Nano to Macro*, pages 924–927.

[Haber and Modersitzki, 2006] Haber, E. and Modersitzki, J. (2006). Intensity Gradient Based Registration and Fusion of Multi-modal Images. In Larsen, R., Nielsen, M., and Sporring, J., editors, *Proc Intern Conf on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 4191 of *LNCS*, pages 726–733, Copenhagen, Denmark. Springer.

[Hadamard, 1902] Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52.

[Hajnal et al., 2001] Hajnal, J., Hill, D., and Hawkes, D. J., editors (2001). *Medical Image Registration.* CRC Press, Boca Raton.

[Hajnal et al., 1995] Hajnal, J. V., Saeed, N., Soar, E. J., Oatridge, A., Young, I. R., and Bydder, G. M. (1995). A Registration and Interpolation Procedure for Subvoxel Matching of Serially Acquired MR Images. *Journal of Computer Assisted Tomography*, 19(2):289–296.

[Harder and Desmarais, 1972] Harder, R. L. and Desmarais, R. N. (1972). Interpolation Using Surface Splines. *Journal of Aircraft*, pages 189–191.

[Harris and Stephens, 1988] Harris, C. and Stephens, M. J. (1988). A combined corner and edge detector. In *Alvey Vision Conference*, pages 147–152.

[Hellier and Barillot, 2003] Hellier, P. and Barillot, C. (2003). Coupling Dense and Landmark Based Approaches for Nonrigid Registration. *IEEE Transactions on Medical Imaging*, 22(2):217–227.

[Hellier et al., 2003] Hellier, P., Barillot, C., Corouge, I., Gibaud, B., Le Goualher, G., Collins, D. L., Evans, A., Malandain, G., Ayache, N., Christensen, G. E., and Johnson, H. J. (2003). Retrospective Evaluation of Intersubject Brain Registration. *IEEE Transactions on Medical Imaging*, 22(9):1120–1130.

[Hellier et al., 2001] Hellier, P., Barillot, C., Memin, E., and Perez, P. (2001). Hierarchical Estimation of a Dense Deformation Field for 3D Robust Registration. *IEEE Transactions on Medical Imaging*, 20(5):388–402.

[Hill et al., 2001] Hill, D., Batchelor, P., Holden, M., and Hawkes, D. (2001). Medical image registration. *Physics in Medicine and Biology*, 46(3):R1–R45.

[Hill and Jarosz, 2001] Hill, D. and Jarosz, J. (2001). Registration of MR and CT images for clinical applications. In Hajnal, J., Hill, D., and Hawkes, D. J., editors, *Medical Image Registration*, chapter 10, pages 217–232. CRC Press, Boca Raton.

[Hoffman et al., 2004] Hoffman, E. A., Clough, A. V., Christensen, G. E., Lin, C.-I., McLennan, G., Reinhardt, J. M., Simon, B. A., Sonka, M., Tawhai, M. H., van Beek, J. R., and Wang, G. (2004). The Comprehensive Imaging-Based Analysis of the Lung: A Forum for Team Science. *Academic Radiology*, 11:1370–1380.

[Hoffman and McLennan, 1997] Hoffman, E. A. and McLennan, G. (1997). Assessment of the Pulmonary Structure-Function Relationship and Clinical Outcomes Measures: Quantitative Volumetric CT of the Lung. *Academic Radiology*, 4:758–776.

[Hoffman et al., 1995] Hoffman, E. A., Tajik, J. K., and Kugelmass, S. D. (1995). Matching Pulmonary Structure and Perfusion via combined Dynamic Multislice CT and Thin-Slice High-Resolution CT. *Computerized Medical Imaging and Graphics*, 19(1):101–112.

[Horn and Schunck, 1981] Horn, B. K. P. and Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence*, 17(1-3):185–203.

[ITK, 2006] ITK (2006). Insight Software Consortium Segmentation and Registration Toolkit. http://www.itk.org.

[Jaehne, 1993] Jaehne, B. (1993). *Spatio-Temporal Image Processing: Theory and Scientific Applications*. Springer.

[Jayakrishnan et al., 2003] Jayakrishnan, V. K., White, P. M., Aitken, D., Crane, P., McMahon, A. D., and Teasdale, E. M. (2003). Subtraction Helical CT Angiography of Intra- and Extracranial Vessels: Technical Considerations and Preliminary Experience. *American Journal of Neuroradiology*, 24:451–455.

[Johnson and Christensen, 2002] Johnson, H. J. and Christensen, G. E. (2002). Consistent Landmark and Intensity-Based Image Registration. *IEEE Transactions on Medical Imaging*, 21(5):450–461.

[Joshi and Miller, 2000] Joshi, S. C. and Miller, M. I. (2000). Landmark Matching via Large Deformation Diffeomorphisms. *IEEE Transactions on Image Processing*, 9(8):1357–1370.

[Kang and Shin, 2003] Kang, H. W. and Shin, S. Y. (2003). Enhanced Lane: Interactive Image segmentation by incremental path map construction. *Graphical Models*, 64(5):282–303.

[Kass et al., 1988] Kass, M., Witkin, A. P., and Terzopoulos, D. (1988). Snakes: Active Contour Models. *International Journal of Computer Vision*, 1(4):321–331.

[Kaye et al., 1998] Kaye, J., Primiano, F. P., and Metaxas, D. (1998). A three dimensional virtual environment for modeling mechanical cardiopulmonary interactions. *Medical Image Analysis*, 2(2):169–195.

[Keeling and Ring, 2005] Keeling, S. L. and Ring, W. (2005). Medical Image Registration and Interpolation by Optical Flow with Maximal Rigidity. *Journal of Mathematical Imaging and Vision*, 23(1):47–65.

[Kitchen and Rosenfeld, 1982] Kitchen, L. and Rosenfeld, A. (1982). Gray-level corner detection. *Pattern Recognition Letters*, 1(2):95–102.

[Klein et al., 2005] Klein, S., Staring, M., and Pluim, J. (2005). Comparison of gradient approximation techniques for optimisation of mutual information in nonrigid registration. In Fitzpatrick, J. M. and Reinhardt, J., editors, *Proc SPIE Conf on Medical Imaging: Image Processing*, volume 5747, pages 192–203.

[Kluckner, 2006] Kluckner, S. (2006). Non-linear 3D-Volume Registration using a local affine Model. Master's thesis, Institute for Machine Vision and Graphics, Graz University of Technology.

[Kohlrausch et al., 2005] Kohlrausch, J., Rohr, K., and Stiehl, H. S. (2005). A New Class of Elastic Body Splines for Nonrigid Registration of Medical Images. *Journal of Mathematical Imaging and Vision*, 23(3):253–280.

[Kwon et al., 2004] Kwon, S. M., Kim, Y. S., Kim, T.-S., and Ra, J. B. (2004). Digital subtraction CT angiography based on efficient 3D registration and refinement. *Computerized Medical Imaging and Graphics*, 28(7):391–400.

[Lanczos, 1986] Lanczos, C. (1986). *The Variational Principles of Mechanics*. Dover Publications, New York, 4th edition.

[Lelieveldt et al., 2000] Lelieveldt, B. P. F., Sonka, M., Bolinger, L., Scholz, T. D., Kayser, H., van der Geest, R., and Reiber, J. H. C. (2000). Anatomical Modeling with Fuzzy Implicit Surface Templates: Application to Automated Localization of the Heart and Lungs in Thoracic MR Volumes. *Computer Vision and Image Understanding*, 80(1):1–20.

[Lell et al., 2006a] Lell, M., Anders, K., Klotz, E., Ditt, H., Bautz, W., and Tomandl, B. F. (2006a). Clinical evaluation of bone-subtraction CT angiography (BSCTA) in head and neck imaging. *European Radiology*, 16(4):889–897.

[Lell et al., 2006b] Lell, M., Ditt, H., Panknin, C., Sayre, J. W., Ruehm, S. G., Klotz, E., Tomandl, B. F., and Villablanca, J. P. (2006b). Bone-subtraction CTA: Evaluation of Two Different Fully Automated Image Registration Procedures for Inter-scan Motion Compensation. *American Journal of Neuroradiology*. in press.

[Lester and Arridge, 1999] Lester, H. and Arridge, S. R. (1999). A Survey of Hierarchical Non-Linear Medical Image Registration. *Pattern Recognition*, 32(1):1–36.

[Li et al., 2002] Li, B., Christensen, G. E., Fill, J., Hoffman, E. A., and Reinhardt, J. M. (2002). 3D Inter-Subject Warping and Registration of Pulmonary CT Images for a Human Lung Model. In *Proc SPIE Conf on Medical Imaging: Physiology and Function from Multidimensional Images*, volume 4683, pages 324–335.

[Li et al., 2003] Li, B., Christensen, G. E., Hoffman, E. A., McLennan, G., and Reinhardt, J. M. (2003). Establishing a Normative Atlas of the Human Lung: Intersubject Warping and Registration of Volumetric CT Images. *Academic Radiology*, 10(3):255–265.

[Likar and Pernus, 2001] Likar, B. and Pernus, F. (2001). A hierarchical approach to elastic registration based on mutual information. *Image and Vision Computing*, 19:33–44.

[Little et al., 1997] Little, J. A., Hill, D. L. G., and Hawkes, D. J. (1997). Deformations Incorporating Rigid Structures. *Computer Vision and Image Understanding*, 66(2):223–232.

[Liu et al., 2004] Liu, T., Shen, D., and Davatzikos, C. (2004). Deformable registration of cortical structures via hybrid volumetric and surface warping. *NeuroImage*, 22:1790–1801.

[Lorensen and Cline, 1987] Lorensen, W. E. and Cline, H. E. (1987). Marching Cubes: A High Resolution 3D Surface Construction Algorithm. *Computer Graphics (Proc. SIGGRAPH '87)*, 21(4):163–169.

[Lowe, 2004] Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110.

[Loy and Zelinsky, 2003] Loy, G. and Zelinsky, A. (2003). Fast Radial Symmetry for Detecting Points of Interest. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):959–973.

[Luboldt et al., 1997] Luboldt, W., Stehling, M., Pearlman, J., Seemann, M., and Raptopoulos, V. (1997). CT-Subtraktionsangiographie (CTSA). *Radiologe*, 37:89–93.

[Lucas and Kanade, 1981] Lucas, B. D. and Kanade, T. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision. In *Proc 7th Intl Conf on Artificial Intelligence*, pages 674–679.

[Maciel and Costeira, 2003] Maciel, J. and Costeira, J. P. (2003). A Global Solution to Sparse Correspondence Problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):187–199.

[Maes et al., 1997] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P. (1997). Multimodality Image Registration by Maximization of Mutual Information. *IEEE Transactions on Medical Imaging*, 16(2):187–198.

[Maintz and Viergever, 1998] Maintz, J. B. A. and Viergever, M. A. (1998). A Survey of Medical Image Registration. *Medical Image Analysis*, 2(1):1–36.

[Maki et al., 1999a] Maki, D. D., Gefter, W. B., and Alavi, A. (1999a). Recent Advances in Pulmonary Imaging. *Chest*, 116(5):1388–1402.

[Maki et al., 1999b] Maki, D. D., Miller, W. T., and Gefter, W. B. (1999b). Advanced emphysema: preoperative chest radiographic findings as predictors of outcome following lung volume reduction surgery. *Radiology*, 212(1):49–55.

[Maksimov et al., 2004] Maksimov, D., Finkel, F., Dietz, T., Dueber, C., Rebholz, P., Hesser, J., and Maenner, R. (2004). An Interactive Application for Removal of Bone Information in CT Angiography. In *Proc 17th Symposium on Computer-Based Medical Systems (CBMS04)*, pages 396–401.

[Malladi et al., 1995] Malladi, R., Sethian, J. A., and Vemuri, B. C. (1995). Shape Modeling with Front Propagation: A Level Set Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175.

[Manke et al., 2002] Manke, D., Rösch, P., Nehrke, K., Börnert, P., and Dössel, O. (2002). Model evaluation and calibration for prospective respiratory motion correction in coronary MR angiography based on 3-D image registration. *IEEE Transactions on Medical Imaging*, 21(9):1132–1141.

[Martin-Fernandez et al., 2005] Martin-Fernandez, M. A., Munoz-Moreno, E., Martin-Fernandez, M., and Alberola-Lopez, C. (2005). Articulated registration: elastic registration based on a wire model. In Fitzpatrick, J. M. and Reinhardt, J. M., editors, *Proc SPIE Conf on Medical Imaging: Image Processing*, volume 5747, pages 182–191.

[Mattes et al., 2003] Mattes, D., Haynor, D. R., Vesselle, H., Lewellen, T. K., and Eubank, W. (2003). PET-CT Image Registration in the Chest Using Free-form Deformations. *IEEE Transactions on Medical Imaging*, 22(1):120–128.

[McInerney and Terzopoulos, 1995] McInerney, T. and Terzopoulos, D. (1995). A Dynamic Finite Element Surface Model for Segmentation and Tracking in Multidimensional Medical Images with Application to Cardiac 4D Image Analysis. *Computerized Medical Imaging and Graphics*, 19(1):69–83.

[Meijering et al., 1999] Meijering, E. H. W., Niessen, W. J., and Viergever, M. A. (1999). Retrospective motion correction in digital subtraction angiography: A Review. *IEEE Transactions on Medical Imaging*, 18(1):2–21.

[Meijering et al., 2001] Meijering, E. H. W., Niessen, W. J., and Viergever, M. A. (2001). Quantitative evaluation of convolution-based methods for medical image interpolation. *Medical Image Analysis*, 5(2):111–126.

[Mikolajczyk and Schmid, 2005] Mikolajczyk, K. and Schmid, C. (2005). Performance Evaluation of Local Descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10):1615–1630.

[Miller et al., 2002] Miller, M. I., Trouve, A., and Younes, L. (2002). On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, 4:375–405.

[Modersitzki, 2004] Modersitzki, J. (2004). *Numerical Methods for Image Registration*. Oxford University Press.

[Mortensen et al., 2005] Mortensen, E. N., Deng, H., and Shapiro, L. (2005). A SIFT Descriptor with Global Context. In *Proc Conf on Computer Vision and Pattern Recognition (CVPR)*, pages 184–190.

[Nagel, 1987] Nagel, H.-H. (1987). On the estimation of optical flow: Relations between different approaches and some new results. *Artifical Intelligence*, 33(3):299–324.

[NEJM-Editors, 2000] NEJM-Editors (2000). Looking Back on the Millennium in Medicine. *New England Journal of Medicine*, 342:42–49.

[Ng and Ibanez, 2004] Ng, L. and Ibanez, L. (2004). Medical Image Registration: Concepts and Implementation. In Yoo, T. S., editor, *Insight into Images*, chapter 10, pages 239–306. A K Peters, Ltd.

[NLM, 2006] NLM (2006). National Library of Medicine Image Data Collection Project. http://nova.nlm.nih.gov/Mayo/.

[Nocedal and Wright, 1999] Nocedal, J. and Wright, S. J. (1999). *Numerical Optimization*. Springer, 2nd edition.

[Oatridge et al., 2001] Oatridge, A., Hajnal, J. V., and Bydder, G. M. (2001). Registration and Subtraction of Serial Magnetic Resonance Images of the Brain: Image Interpretation and Clinical Applications. In Hajnal, J., Hill, D., and Hawkes, D. J., editors, *Medical Image Registration*, chapter 7, pages 143–182. CRC Press, Boca Raton.

[Osher and Fedkiw, 2003] Osher, S. and Fedkiw, R. (2003). *Level Set Methods and Dynamic Implicit Surfaces*. Applied Mathematical Sciences. Springer Verlag.

[Palagyi et al., 2003] Palagyi, K., Tschirren, J., and Sonka, M. (2003). Quantitative Analysis of Intrathoracic Airway Trees: Methods and Validation. In *Proc Information Processing in Medical Imaging*, volume 2732 of *LNCS*, pages 222–233. Springer.

[Papademetris et al., 2004] Papademetris, X., Jackowski, A. P., Schultz, R. T., Staib, L. H., and S., D. J. (2004). Integrated Intensity and Point-Feature Nonrigid Registration. In Barillot, C., Haynor, D. R., and Hellier, P., editors, *Proc Intern Conf on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 3216 of *LNCS*, pages 763–770, Rennes, France. Springer.

[Pennec et al., 9999] Pennec, X., Cachier, P., and Ayache, N. (19999). Understanding the Demon's algorithm. In Taylor, C. and Colchester, A., editors, *Proc Intern Conf on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 1679 of *LNCS*, pages 597–605, Cambridge, England. Springer.

[Perona and Malik, 1990] Perona, P. and Malik, J. (1990). Scale-Space and Edge Detection by Anisotropic Diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639.

[Pietrzyk, 2001] Pietrzyk, U. (2001). Registration of MRI and PET images for clinical applications. In Hajnal, J., Hill, D., and Hawkes, D. J., editors, *Medical Image Registration*, chapter 9, pages 199–216. CRC Press, Boca Raton.

[Pitiot et al., 2003] Pitiot, A., Malandain, G., Bardinet, E., and Thompson, P. M. (2003). Piecewise Affine Registration of Biological Images. In Gee, J. C., Maintz, J. B. A., and Vannier, M. W., editors, *Second International Workshop on Biomedical Image Registration WBIR'03*, volume 2717 of *Lecture Notes in Computer Science*, pages 91–101. Springer Verlag.

[Pluim et al., 2003] Pluim, J. P. W., Maintz, J. B. A., and Viergever, M. A. (2003). Mutual-Information-Based Registration of Medical Images: A Survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004.

[Pock et al., 2005a] Pock, T., Beichel, R., and Bischof, H. (2005a). A Novel Robust Tube Detection Filter for 3D Centerline Extraction. In *Proc Scandinavian Conference on Image Analysis (SCIA)*, volume 3540 of *LNCS*, pages 481–490. Springer.

[Pock et al., 2005b] Pock, T., Janko, C., Beichel, R., and Bischof, H. (2005b). Multiscale Medialness for Robust Segmentation of 3D Tubular Structures. In *Proceedings of the Computer Vision Winter Workshop 2005*.

[Press et al., 1992] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in C. The Art of Scientific Computing*. Cambridge University Press, Cambridge, England, 2nd edition.

[Rohlfing and Maurer Jr, 2001] Rohlfing, T. and Maurer Jr, C. R. (2001). Intensity-Based Non-rigid Registration Using Adaptive Multilevel Free-Form Deformation with an Incompressibility Constraint. In Niessen, W. and Viergever, M., editors, *Proc Intern Conf on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 2208 of *LNCS*, pages 111–119, Utrecht, The Netherlands. Springer.

[Rohlfing et al., 2003] Rohlfing, T., Maurer Jr, C. R., Bluemke, D. A., and Jacobs, M. A. (2003). Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. *IEEE Transactions on Medical Imaging*, 22(6):730–741.

[Rohlfing et al., 2001] Rohlfing, T., Maurer Jr, C. R., O'Dell, W. G., and Zhong, J. (2001). Modeling liver motion and deformation during the respiratory cycle using intensity-based free-form registration of gated MR images. In Mun, S. K., editor, *Proc SPIE Conf* *on Medical Imaging: Visualization, Image-Guided Procedures, and Display*, volume 4319, pages 337–348.

[Rohr, 2001] Rohr, K. (2001). *Landmark-Based Image Analysis Using Geometric and Intensity Models*. Computational Imaging and Vision. Kluwer Academic Publishers.

[Rohr et al., 2001] Rohr, K., Stiehl, H. S., Sprengel, R., Buzug, T. M., Weese, J., and Kuhn, M. H. (2001). Landmark-Based Elastic Registration Using Approximating Thin-Plate Splines. *IEEE Transactions on Medical Imaging*, 20(6):526–534.

[Rueckert et al., 2006] Rueckert, D., Aljabar, P., Heckemann, R. A., Hajnal, J. V., and Hammers, A. (2006). Diffeomorphic Registration Using B-Splines. In Larsen, R., Nielsen, M., and Sporring, J., editors, *Proc Intern Conf on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 4191 of *LNCS*, pages 702–709, Copenhagen, Denmark. Springer.

[Rueckert et al., 2003] Rueckert, D., Frangi, A. F., and Schnabel, J. A. (2003). Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. *IEEE Transactions on Medical Imaging*, 22(8):1014–1025.

[Rueckert et al., 1999] Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L. G., Leach, M. O., and Hawkes, D. J. (1999). Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images. *IEEE Transactions on Medical Imaging*, 18(8):712–721.

[Scharr and Uttenweiler, 2001] Scharr, H. and Uttenweiler, D. (2001). 3D Anisotropic Diffusion Filtering for Enhancing Noisy Actin Filament Fluorescence Images. In Radig, B. and Florczyk, S., editors, *Proc Pattern Recognition - DAGM Symposium*, volume 2191 of *LNCS*, pages 69–75, Munich. Springer.

[Scharstein and Szeliski, 2002] Scharstein, D. and Szeliski, R. (2002). A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision*, 47(1/2/3):7–42.

[Schnabel et al., 2003] Schnabel, J. A., Tanner, C., Castellano-Smith, A. D., Degenhard, A., Leach, M. O., Hose, D. R., Hill, D. L. G., and Hawkes, D. J. (2003). Validation of Nonrigid Image Registration Using Finite-Element Methods: Application to Breast MR Images. *IEEE Transactions on Medical Imaging*, 22(2):238–247.

[Sclaroff and Pentland, 1995] Sclaroff, S. and Pentland, A. (1995). Modal Matching for Correspondence and Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(6):545–561.

[Sebastian et al., 2003] Sebastian, T. B., Tek, H., Crisco, J. J., and Kimia, B. B. (2003). Segmentation of carpal bones from CT images using skeletally coupled deformable models. *Medical Image Analysis*, 7(1):21–45.

[Segars et al., 2001] Segars, W. P., Lalush, D. S., and Tsui, B. M. W. (2001). Modeling Respiratory Mechanics in the MCAT and Spline-Based MCAT Phantoms. *IEEE Transactions on Nuclear Science*, 48(1):89–97.

[Sethian, 1999] Sethian, J. A. (1999). *Level Set Methods and Fast Marching Methods.* Cambridge University Press, 2nd edition.

[Shannon, 1948] Shannon, C. (1948). The mathematical theory of communication (parts 1 and 2). *Bell Syst. Tech. J.*, 27:379–423 and 623–656. Reprint available from http://www.lucent.com.

[Simon et al., 2005] Simon, B. A., Christensen, G. E., Low, D. A., and Reinhardt, J. M. (2005). Computed Tomography Studies of Lung Mechanics. *Proc American Thoracic Society*, 2:517–521.

[Slone et al., 1997] Slone, R. M., K., P. T., and Gierada, D. S. (1997). Lung volume reduction surgery: Comparison of radiologic features and clinical outcome. *Radiology*, 204(3):685–693.

[Snel et al., 2002] Snel, J. G., Venema, H. W., and Grimbergen, C. A. (2002). Deformable triangular surfaces using fast 1-D radial lagrangian dynamics-segmentation of 3-D MR and CT images of the wrist. *IEEE Transactions on Medical Imaging*, 21(8):888–903.

[Sonka and Fitzpatrick, 2000] Sonka, M. and Fitzpatrick, J. M., editors (2000). *Handbook of Medical Imaging: Volume 2 - Medical Image Processing and Analysis.* SPIE Press.

[Sonka et al., 1999] Sonka, M., Hlavac, V., and Boyle, R. (1999). *Image Processing, Analysis and Machine Vision.* Brooks/Cole Publishing Company, Pacific Grove, CA, USA, 2nd edition.

[Staring et al., 2005] Staring, M., Klein, S., and Pluim, J. P. W. (2005). Nonrigid Registration with Adaptive Content-Based Filtering of the Deformation Field. In Fitzpatrick, J. M. and Reinhardt, J., editors, *Proc SPIE Conf on Medical Imaging: Image Processing*, volume 5747, pages 212–221.

[Strikwerda, 1989] Strikwerda, J. C. (1989). *Finite Difference Schemes and Partial Differential Equations.* International Thomson Science.

[Studholme et al., 1999] Studholme, C., Hill, D. L. G., and Hawkes, D. J. (1999). An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32(1):71–86.

[Tai et al., 1997] Tai, Y.-C., Lin, K. P., Hoh, C. K., Huang, S. C. H., and Hoffman, E. J. (1997). Utilization of 3D Elastic Transformation in the Registration of Chest X-Ray CT and Whole Body PET. *IEEE Transactions on Nuclear Science*, 44(4):1606–1612.

[Terzopoulos and Metaxas, 1991] Terzopoulos, D. and Metaxas, D. (1991). Dynamic 3D Models with Local and Global Deformations: Deformable Superquadrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):703–714.

[Terzopoulos et al., 1988] Terzopoulos, D., Witkin, A., and Kass, M. (1988). Constraints on Deformable Models: Recovering 3D Shape and Nonrigid Motion. *Artificial Intelligence*, 36(1):91–123.

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

[Thirion, 1995] Thirion, J.-P. (1995). Fast Non-Rigid Matching of 3D Medical Images. Rapport de recherche 2547, Unite de recherche INRIA Sophia-Antipolis - Projet Epidaure.

[Thirion, 1998] Thirion, J.-P. (1998). Image matching as a diffusion process: An analogy with Maxwell's demons. *Medical Image Analysis*, 2(3):243–260.

[Thirion and Gourdon, 1995] Thirion, J.-P. and Gourdon, A. (1995). Computing the Differential Characteristics of Isointensity Surfaces. *Computer Vision and Image Understanding*, 61(2):190–202.

[Thompson and Toga, 1996] Thompson, P. and Toga, A. W. (1996). A Surface-Based Technique for Warping Three-Dimensional Images of the Brain. *IEEE Transactions on Medical Imaging*, 15(4):402–417.

[Tikhonov and Arsenin, 1977] Tikhonov, A. and Arsenin, V. (1977). *Solutions of Ill-Posed Problems*. Winston, Washington D.C.

[Tomandl et al., 2006] Tomandl, B. F., Hammen, T., Klotz, E., Ditt, H., Stemper, B., and Lell, M. (2006). Bone-Subtraction CT Angiography for the Evaluation of Intracranial Aneurysms. *American Journal of Neuroradiology*, 27(1):55–59.

[Tschirren et al., 2005a] Tschirren, J., Hoffman, E. A., McLennan, G., and Sonka, M. (2005a). Intrathoracic Airway Trees: Segmentation and Airway Morphology Analysis From Low-Dose CT Scans. *IEEE Transactions on Medical Imaging*, 24(12):1529–1539.

[Tschirren et al., 2005b] Tschirren, J., McLennan, G., Palagyi, K., Hoffman, E. A., and Sonka, M. (2005b). Matching and Anatomical Labeling of Human Airway Tree. *IEEE Transactions on Medical Imaging*, 24(12):1540–1547.

[Urschler et al., 2006a] Urschler, M., Bauer, J., Ditt, H., and Bischof, H. (2006a). SIFT and Shape Context for Feature-Based Nonlinear Registration of Thoracic CT Images. In Beichel, R. and Sonka, M., editors, *Proc Computer Vision Applications to Medical Image Analysis*, volume 4241 of *LNCS*, pages 73–84. Springer.

[Urschler and Bischof, 2004a] Urschler, M. and Bischof, H. (2004a). Matching 3D Lung Surfaces with the Shape Context Approach. In Wilhelm Burger, J. S., editor, *28th Workshop of the Austrian Association for Pattern Recognition*, pages 133–140. Oesterreichische Computer Gesellschaft.

[Urschler and Bischof, 2004b] Urschler, M. and Bischof, H. (2004b). Registering 3D Lung Surfaces Using the Shape Context Approach. In *Proc Medical Image Understanding and Analysis (MIUA)*, pages 212–215, Imperial College London.

[Urschler and Bischof, 2005] Urschler, M. and Bischof, H. (2005). Assessing breathing motion by shape matching of lung and diaphragm surfaces. In *Proc SPIE Conf on Medical Imaging: Physiology and Function from Multidimensional Images*, volume 5746, pages 440–452, San Diego.

[Urschler et al., 2006b] Urschler, M., Ditt, H., and Bischof, H. (2006b). Partially Rigid Bone Registration in CT Angiography. In *Proc of the Computer Vision Winter Workshop (CVWW)*, pages 34–39, Telc, Czech Republic.

[Urschler et al., 2006c] Urschler, M., Zach, C., Ditt, H., and Bischof, H. (2006c). Automatic Point Landmark Matching for Regularizing Nonlinear Intensity Registration: Application to Thoracic CT Images. In Larsen, R., Nielsen, M., and Sporring, J., editors, *Proc Intern Conf on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 4191 of *LNCS*, pages 710–717, Copenhagen, Denmark. Springer.

[van Straten et al., 2004] van Straten, M., Venema, H. W., Streekstra, G. J., Majoie, C., den Heeten, G. J., and Grimbergen, C. A. (2004). Removal of bone in CT angiography of the cervical arteries by piecewise matched mask bone elimination. *Medical Physics*, 31(10):2924–2933.

[Veltkamp and Hagedoorn, 1999] Veltkamp, R. C. and Hagedoorn, M. (1999). State of the Art in Shape Matching. Technical Report UU-CS-1999-27, Utrecht University, Netherlands, Utrecht.

[Venema et al., 2001] Venema, H. W., Hulsmans, F. J. H., and den Heeten, G. J. (2001). CT Angiography of the Circle of Willis and Intracranial Internal Carotid Arteries: Maximum Intensity Projection with Matched Mask Bone Elimination - Feasibility Study. *Radiology*, 218(3):893–898.

[Viola and Jones, 2004] Viola, P. and Jones, M. J. (2004). Robust Real-Time Face Detection. *International Journal of Computer Vision*, 57(2):137–154.

[Viola and Wells III, 1997] Viola, P. and Wells III, W. M. (1997). Alignment by Maximization of Mutual Information. *International Journal of Computer Vision*, 24(2):137–154.

[Wahba, 1990] Wahba, G. (1990). *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

[Weickert and Brox, 2002] Weickert, J. and Brox, T. (2002). Diffusion and Regularization of Vector- and Matrix-Valued Images. Technical Report 58, University Saarbrücken, Saarland University, Faculty of Mathematics and Computer Science.

[Weickert and Schnörr, 2001a] Weickert, J. and Schnörr, C. (2001a). A Theoretical Framework for Convex Regularizers in PDE-Based Computation of Image Motion. *International Journal of Computer Vision*, 45(3):245–264.

[Weickert and Schnörr, 2001b] Weickert, J. and Schnörr, C. (2001b). Variational Optic Flow Computation with a Spatio-Temporal Smoothness Constraint. *Journal of Mathematical Imaging and Vision*, 14(3):245–255.

[Weickert et al., 1998] Weickert, J., ter Haar Romeny, B. M., and Viergever, M. A. (1998). Efficient and Reliable Schemes for Nonlinear Diffusion Filtering. *IEEE Transactions on Image Processing*, 7(3):398–410.

[Wells III et al., 1996] Wells III, W. M., Viola, P., Atsumi, H., Nakajima, S., and Kikinis, R. (1996). Multi-Modal Volume Registration by Maximization of Mutual Information. *Medical Image Analysis*, 1(1):35–51.

[Weruaga et al., 2003] Weruaga, L., Morales, J., Nunez, L., and Verdu, R. (2003). Estimating volumetric motion in human thorax with parametric matching constraints. *IEEE Transactions on Medical Imaging*, 22(6):766–772.

[Wildberger et al., 2005] Wildberger, J. E., Klotz, E., Ditt, H., Mahnken, A. H., Spüntrup, E., and Günther, R. W. (2005). Multi-slice CT for Visualization of Acute Pulmonary Embolism: Single Breath-hold Subtraction Technique. *Fortschritte auf dem Gebiet der Roentgenstrahlen und der bildgebenden Verfahren*, 177(1):17–23.

[Wilson, 1988] Wilson, T. A. (1988). Mechanics of compartmental models of the chest wall. *Journal of Applied Physiology*, 65(5):2261–2264.

[Wollny and Kruggel, 2002] Wollny, G. and Kruggel, F. (2002). Computational Cost of Nonrigid Registration Algorithms Based on Fluid Dynamics. *IEEE Transactions on Medical Imaging*, 21(8):946–952.

[Woods et al., 1992] Woods, R. P., Cherry, S. R., and Mazziotta, J. C. (1992). Rapid automated algorithm for aligning and reslicing PET images. *Journal of Computer Assisted Tomography*, 16(4):620–633.

[Woods et al., 1993] Woods, R. P., Mazziotta, J. C., and Cherry, S. R. (1993). MRI-PET Registration with Automated Algorithm. *Journal of Computer Assisted Tomography*, 17(4):536–546.

[Xie and Farin, 2004] Xie, Z. and Farin, G. E. (2004). Image Registration Using Hierarchical B-Splines. *IEEE Transactions on Visualization and Computer Graphics*, 10(1):85–94.

[Yeung et al., 1994] Yeung, M. M., Yeo, B.-L., Liou, S.-P., and Bani-Hashemi, A. (1994). Three-Dimensional Image Registration for Spiral CT Angiography. In *Proc Conf on Computer Vision and Pattern Recognition (CVPR)*, pages 423 – 429.

[Yezzi et al., 2003] Yezzi, A., Zöllei, L., and Kapur, T. (2003). A variational framework for integrating segmentation and registration through active contours. *Medical Image Analysis*, 7(2):171–185.

[Yoo, 2004] Yoo, T. (2004). *Insight into Images*. A K Peters, Ltd.

[Zhang and Reinhardt, 2000] Zhang, L. and Reinhardt, J. M. (2000). 3D Pulmonary CT Image Registration with a Standard Lung Atlas. In *Proc SPIE Conf on Medical Imaging: Physiology and Function from Multidimensional Images*, volume 3978, pages 67–77.

[Zhang et al., 1995] Zhang, Z., Deriche, R., O., F., and Luong, Q. T. (1995). A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78:87–119.

[Zheng and Doermann, 2006] Zheng, Y. and Doermann, D. (2006). Robust Point Matching for Nonrigid Shapes by Preserving Local Neighborhood Structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4):643–649.

[Zitova and Flusser, 2003] Zitova, B. and Flusser, J. (2003). Image registration methods: A survey. *Image and Vision Computing*, 21(11):977–1000.

Statutory Declaration

Herewith I affirm that the presented dissertation was made autonomous without illegal help. Further I declare that I did not use other sources than those cited.

Graz, November 2006

Martin Urschler