

Statistical Medical Image Registration with
Applications in Epilepsy Diagnosis and
Shape-Based Segmentation

Der Technischen Fakultät der
Universität Erlangen–Nürnberg

zur Erlangung des Grades

DOKTOR–INGENIEUR

vorgelegt von

Dieter Arnold Hahn

Erlangen — 2009

Deutscher Titel:
Statistische medizinische Bildregistrierung mit
Anwendungen in der Epilepsie-Diagnostik und der
formbasierten Segmentierung

Als Dissertation genehmigt von der
Technischen Fakultät der
Universität Erlangen-Nürnberg

Tag der Einreichung:	06.07.2009
Tag der Promotion:	20.11.2009
Dekan:	Prof. Dr.-Ing. R. German
Berichterstatter:	Prof. Dr.-Ing. J. Hornegger Prof. W. Wells

Acknowledgment

This thesis would not have come into existence without the help and contributions I received from others during the time I worked at the Pattern Recognition Lab of Prof. Hornegger in Erlangen.

First and foremost, I owe a tremendous debt of gratitude to my supervisor Prof. Joachim Hornegger for his faith in me, his understanding, for promoting my research since my student's thesis, and for opening up so many possibilities. If it had not been for him, I would not have had the chance to participate in this challenging but rewarding field of research. I am also very thankful to Prof. William Wells for co-supervising this thesis and his valuable feedback. My special thanks go to Prof. Torsten Kuwert, who not only agreed to participate in the committee, but also provided his medical expertise in many fruitful discussions.

I am very grateful to Volker Daum and Marcus Prümmer for encouraging discussions and a great time – not only at the lab. Among the students I have supervised, I want to thank Martin Spiegel and Eva Kollorz for their work on our journal publications. Furthermore, I am indebted to Volker Daum, Marcus Prümmer, and Gabriele Wolz for reading and correcting my thesis, and Simone Gaffling for her dedicated assistance in the Pattern Recognition exercises. At the lab, we are all very happy to have Prof. Elli Angelopoulou with us, and I want to thank her for always taking the time to thoroughly proof-read and improve the draft versions of our articles. In addition, I would like to acknowledge HipGraphics for providing the volume rendering software *InSpace*, which lets the images look so nice and provides such a powerful programming framework, and Stefan Lautenschläger from Siemens Healthcare for supplying AX phantom images.

Last, but certainly not least, I want to thank the most important people in my life – my family, and my girlfriend Gabi. Thank you for your great patience, your steady support, and the encouragement for all these years!

Dieter A. Hahn

Abstract

The advances in scanner technologies over the past years and a growing number of modalities in medical imaging result in an increased amount of patient data. The physicians are faced with an overwhelming amount of information when comparing different scans. Therefore, automatic image processing algorithms are necessary to facilitate everyday clinical workflows. The present work focuses on automatic, statistical image registration approaches and applications in epilepsy diagnosis and shape-based segmentation.

Registration algorithms based on image intensity statistics are currently state-of-the-art to automatically compute an alignment between multi-modal images. The parameters, however, are sensitive to the input data. In the present work, we study the mutual influences of these parameters on the intensity statistics and present data-driven estimation schemes to optimize them with respect to the input images. This is necessary to register large sets of images both accurately and reliably. The presented evaluation results, which are based on a database with an established gold standard, confirm that individually optimized parameters lead to improved results compared to standard settings found in literature.

Besides spatial accuracy, the reduction of the computation time for the registration is equally important. In this thesis, we present an approach to reduce the search space for the optimization of a rigid registration transform by a nonlinear projection scheme, which is closely related to the concept of marginalization of random variables. Within each projection, a disjoint subset of the transform parameters is optimized with greatly reduced computational complexity. With a good choice of the projection geometry, the search space can be separated into disjoint subsets. In the case of rigid 3-D image registrations, the nonlinear projection onto a cylinder surface allows for an optimization of the rotation around the cylinder axis and a translation along its direction without the need for a reprojection. Sub-volume registration problems are supported by fitting the projection geometry into the overlap domain of the input images. The required objective functions are constrained by systems of linear inequalities and solved by means of constrained, nonlinear optimization techniques.

A statistical framework is proposed to measure the accuracy of the registration algorithms with respect to manual segmentation results. The aforementioned concepts of the data-driven density estimators are adopted for the estimation of spatial densities of the segmented labels in order to model the observer reliability. The accuracy of the spatial registration transform is measured between the estimated distributions of the segmented labels in both input images using the Kullback-Leibler divergence.

The proposed algorithms are evaluated by a registration of a database of morphological and functional images with an established gold standard based on fiducial marker implants. Applications are presented for the subtraction of single emission computed tomography scans for epilepsy diagnosis, where the intensity distributions are estimated for both the task of the registration and the normalization of the images. Finally, the registration is utilized for shape-based image segmentation to establish a model for the variability within a collective of segmented training shapes.

Übersicht

Erzielte Fortschritte in der Technologie und den Aufnahmeverfahren medizinischer Geräte in den letzten Jahren führen zu einer merklichen Zunahme von Patientendaten. Die Ärzte werden mit einer Vielzahl an Informationen konfrontiert, wenn es darum geht, verschiedene Aufnahmen miteinander zu vergleichen. Deshalb ist es nötig, automatische Algorithmen zur Verarbeitung der Daten in die klinischen Abläufe zu integrieren. Die vorliegende Arbeit beschäftigt sich mit der automatischen, statistischen Bildregistrierung und Anwendungen in der Epilepsie Diagnostik und der formbasierten Segmentierung.

Registrierungsalgorithmen, die auf Bildstatistiken beruhen, sind derzeit aktueller Stand, um örtliche Zusammenhänge zwischen multi-modalen Bilddaten automatisch zu berechnen. Hierfür ist in einer praktischen Realisierung eine Vielzahl an Parametern nötig. Die entsprechenden Werte hängen von den Eingabedaten ab. In dieser Dissertation werden die Zusammenhänge und gegenseitigen Einflüsse dieser Werte auf die Bildstatistik untersucht und datengetriebene Verfahren vorgestellt, um optimale Werte bezüglich der Bilder zu ermitteln. Dies ist notwendig für die robuste und genaue Registrierung von großen Datenmengen. Die vorgestellte Evaluation der Methoden, basierend auf einer etablierten Datenbank mit vorhandenem Goldstandard, bestätigt, dass die automatisch ermittelten Parameterwerte zu einer Verbesserung führen verglichen mit Werten, wie sie standardmäßig in einschlägiger Literatur gefunden werden können.

Neben der räumlichen Genauigkeit ist auch die Laufzeit der Algorithmen von Bedeutung. In dieser Arbeit wird ein Ansatz vorgestellt, der den Suchraum für eine starre Registrierungstransformation durch nichtlineare Projektions schemata reduziert. Das Konzept ist eng verwandt mit der Marginalisierung von Zufallsvariablen. In jeder Projektion kann eine Untermenge der Transformationsparameter mit stark reduziertem Berechnungsaufwand optimiert werden. Die Wahl der Projektionsgeometrie spielt eine Rolle bei der Separierbarkeit des Suchraums. Im Falle von starren Bildtransformationen in 3-D erlaubt die nichtlineare Projektion auf eine Zylinderoberfläche die Trennung in eine Rotation um die Zylinderachse und eine Translationskomponente entlang des Zylinders. Während der Optimierung dieser Untermenge sind keine zusätzlichen Projektionen aus dem hochdimensionalen Raum nötig. Subvolumen Probleme lassen sich über eine geeignete Einpassung der Projektionsgeometrie in den Überlappungsbereich der Bilder erfassen. Die dazugehörigen Zielfunktionen werden durch eine Optimierung mit Nebenbedingungen gelöst.

Zusätzlich wird ein statistischer Ansatz vorgestellt, um die Genauigkeit von Registrierungsalgorithmen bezüglich manueller Segmentierungen von medizinischen Experten zu erfassen. Die erläuterten Konzepte der datengetriebenen Intensitätsstatistik können hierfür angepasst werden. Es werden örtliche Dichtefunktionen für die segmentierten Daten geschätzt, um die Variabilität der Beobachter zu integrieren. Die Genauigkeit der Registrierung kann anschließend durch ein Abstandsmaß zwischen den Dichten erfasst werden, das auf der Kullback-Leibler Divergenz beruht.

Vorgestellte Algorithmen werden anhand einer Datenbank mit morphologischen und funktionalen Bilddaten evaluiert. Die Genauigkeit der Transformation bestimmt sich hierbei aus dem Abstand von implantierten Markern. Eine Anwendungsmöglichkeit der Methoden findet sich in der Subtraktion von SPECT (Single

Emission Computed Tomography) Aufnahmen für die Epilepsie Diagnostik. Hierbei werden die aufgeführten, statistischen Verfahren im Rahmen der Registrierung und der Normierung der Bilddaten verwendet. Abschließend wird die Registrierung für die Generierung eines Formmodells zur Segmentierung eingesetzt.

Contents

List of Figures	v
List of Tables	ix
1 Introduction	1
1.1 Contributions	2
1.2 Document Overview	3
I Background on Medical Image Registration	5
2 Clinical Applications	7
2.1 Radiation Therapy	8
2.2 Follow-Up Patient Studies	8
2.3 Reconstruction	9
2.4 Challenges	9
3 Registration Methods	11
3.1 Parametric Registration	12
3.2 Non-Parametric Registration	13
3.3 Similarity Measures	15
3.3.1 Feature-Based Measures	15
3.3.2 Direct Voxel Intensity-Based Measures	18
3.3.3 Non-Statistical Indirect Measures	19
3.3.4 Statistical Measures	21
II Theory	25
4 Statistical Similarity Measures	27
4.1 Sampling Artifacts	30
4.2 Density Estimation	32
4.3 Automatic Data-Driven Parameter Selection	33
4.3.1 Leave-One-Out Cross-Validation	34
4.3.2 Discretization of the Parzen-Window Estimator	35
4.3.3 Histogram Layout	41
4.3.4 Coincidence Weighting	44

4.4	Implementational Aspects	47
4.5	Results	49
5	Projection-Based Registration	53
5.1	Decomposition of the Parameter Search Space	54
5.2	Orthogonal Projections	55
5.3	Nonlinear Projection Schemes	56
5.3.1	Circular Projections for 2-D Problems	57
5.3.2	Cylinder Projections for 3-D Problems	58
5.3.3	Fitting the Projection Geometry	59
5.3.4	Uniqueness of the Solution	66
5.4	Comparison to Marginal Space Learning	67
5.5	Results	68
6	An Evaluation Approach for Registration Algorithms	71
6.1	Statistical Evaluation Framework	72
6.2	Spatial Density Estimation	73
6.2.1	Parametric Models	73
6.2.2	Non-Parametric Models	74
6.3	Experiments	76
III	Applications	79
7	3-D/3-D Multi-Modal Rigid Registration of the RIRE Database	81
7.1	Significance Tests	82
7.2	Target Registration Errors	83
7.3	Acceptance Rates	85
7.4	Influence of MR Distortion Correction	86
8	SPECT Subtraction Imaging for Epilepsy Diagnoses	89
8.1	Related Work	90
8.2	Clinical Workflow	90
8.2.1	Rigid Registration	91
8.2.2	Intensity Normalization	91
8.2.3	Non-Parametric Registration	93
8.2.4	Visualization	94
8.3	Results	95
9	Registration Applied to Shape-Based Segmentation	99
9.1	Model Generation	100
9.2	Registered Point Correspondences	102
9.3	Results	105
9.3.1	Evaluation Measures	106
9.3.2	Experimental Results	107
10	Outlook	115

11 Summary	119
A Cylinder Projections on the GPU	125
A.1 Discretization of the Projection Image	125
A.2 Multi-Level Cylinder Projection	126
A.3 Coordinate Transformation	127
A.4 GPU Implementation	128
B Notation	133
C Abbreviations	137
Bibliography	139
Index	149

List of Figures

2.1	Example for an image fusion	7
3.1	Coarse classification of image similarity measures	16
4.1	Demonstration of the resulting mismatch due to wrong parameters . .	29
4.2	Illustration of sampling with and without jittering	30
4.3	Plots of the densely evaluated \mathcal{D}^{MI} objective function values for trans- lational parameter changes	31
4.4	Illustration of the approximation error between the continuous PDF estimator $p_{\lambda,n}$ and the discrete version $\hat{p}_{\lambda,n}$	36
4.5	Convolution kernels for the kernel width estimation using a cubic B- spline window function K_{λ}^{B}	39
4.6	Parzen-window log-likelihood function for the width of a cubic B-spline kernel	39
4.7	Optimal kernel width parameter values for various numbers of samples and bins	40
4.8	Automatic selection of the number of bins for a reference CT and a template PET image	44
4.9	Structured noise visible in the background	45
4.10	Intensity histograms for the images in Figure 4.8a	49
4.11	Results of the proposed PDF estimation for a PET-CT image pair . .	50
4.12	Registration of the images provided in Figure 4.1 using the proposed PDF estimation techniques	51
5.1	Illustration of a linear projection of features onto the x -axis	56
5.2	Projection schemes for 2-D images to separate (a) translational and (b) rotational parameters	57
5.3	The images contain (a) the illustration of a fitted cylinder geometry and (b) a visualization of the unfolded cylinder projection of a 3-D CT head image	59
5.4	Projection results for different overlap domains	60
5.5	Illustrations of the cylinder geometry	65
5.6	Schematic view from top onto two equal but shifted image domains and the solution space for the location of the cylinders	66
5.7	Slices taken from the 3-D checkerboard images resulting from a regis- tration between a pre- and post-therapy CT dataset	68

6.1	(a) Shows an example 2-D point set for a label within an image, (b) is the corresponding log-likelihood objective function	75
6.2	Estimated spatial distributions for a non-optimal, compared to an optimal B-spline kernel width	75
6.3	The images in the first row show (a) the plot of the segmentations for a label with a high variation	76
6.4	Label sets for a slice taken from an MR image of a human brain . . .	77
6.5	Variations in the segmented label data for the MRI brain slice	77
6.6	Artificial deformation defined by modification of (a) the control grid for a cubic B-spline transform in 2-D	78
6.7	The proposed evaluation criterion (6.1) applied to the labels in example 6.4 with varying values for α	78
7.1	Mean and standard deviation of the TRE for MP-Rage MR sequence image pairs	84
7.2	Mean and standard deviation of the TRE for CT-MR registrations . .	84
7.3	Mean and standard deviation of the TRE for PET-MR registrations .	85
7.4	Landmark acceptance rates for the various modality combinations . .	86
8.1	Input SPECT image dataset (a) before, and (b) after the rigid registration	91
8.2	The figures show the joint PDF of two input SPECT images before, after the registration, and after the intensity normalization	92
8.3	An example result of the proposed SPECT normalization approach .	93
8.4	Illustration of the bias of a rigid registration introduced by local variations in the image content	94
8.5	Illustration of the several steps during the evaluation of the SPECT difference images	96
8.6	A comparison between the results of the two workflow versions for an example dataset	97
8.7	Fusion visualization between an MRI and the subtraction result of the proposed workflow for two epilepsy patients	97
8.8	Two example patients taken from the collective. The images show the SPECT subtraction image fused with the corresponding MRI.	98
9.1	The components of a state-of-the-art ASM segmentation system . . .	100
9.2	Differences between deformation fields resulting from similarity measures \mathcal{D}^{SSD} and \mathcal{D}^{κ}	103
9.3	Directions of the registrations for the N shape images used for the training of the ASM	104
9.4	The three major principal modes of variation of a kidney shape model trained with 20 samples and created with the CSSD registration approach	106
9.5	Comparison of the three different ASMs based on 10 and 20 training samples for varying starting positions in segmentation of a right kidney	109
9.6	The images show the iterative refinement a right kidney segmentation example using the CSSD model trained with 20 samples	109

9.7	3-D visualization of the segmentation result for the example shown in Figure 9.6	110
9.8	A comparison of the sensitivity for the MDL, SSD and CSSD models generated by a different number of training samples	110
9.9	The MSE for the three compared model with respect to the number of training samples	111
9.10	A comparison of the mean shapes between the CSSD and the MDL models created from 20 training samples. Left: CSSD registration approach model. Right: MDL model.	112
9.11	Distances between the surfaces of a test shape to the resulting segmentations with the CSSD (left) and the MDL model (right)	112
A.1	Multi-level cylinder projection	127
A.2	The (a) node-based image grid is commonly used for image interpolation on the CPU	129
A.3	Example implementation for a GPU cylinder projection fragment program written in OpenGL Shading Language	131

List of Tables

3.1	Description of the PDFs required for the statistical image similarity measures.	21
7.1	Abbreviations used for the comparisons between the parameter selection methods	82
7.2	Median TRE values for the evaluation of CT-MR registrations	83
7.3	Median TRE values for the evaluation of PET-MR registrations	83
8.1	Comparison between the evaluation results of the two workflow versions	95
9.1	Characteristics of the clinical datasets for one side of the kidney pairs used for the MDL approach.	105
9.2	Results for the ASM segmentation of a left kidney for 13 different starting positions	108
9.3	Results for the ASM segmentation of a right kidney for 13 different starting positions	108
9.4	Mean values for the SE and MSE of the models with respect to varying numbers of training samples	111
A.1	Explanation of the variables used within the fragment program in Figure A.3	132

Chapter 1

Introduction

Scanning the human body with a variety of imaging devices has become a vital part of various clinical applications. Improvements of the acquisition technologies result in a growing amount of data. The physicians not only have to analyze thousands of image slices taken throughout the human body, they also have to combine information from different modalities. Each of these imaging devices has its specific field of application. Among the most important acquisition technologies that are used in the clinics nowadays are computed tomography (CT), C-arm computed tomography for X-ray angiography imaging (AX), fluoroscopy, magnetic resonance imaging (MRI), positron emission tomography (PET), single photon emission computed tomography (SPECT), and ultrasound (US). Patients may undergo several scans with a single or multiple imaging techniques throughout the treatment. Until a few years ago, the physicians had to compare the printout films using light boxes. This method requires the physician being able to mentally fuse the scans because the patient's position and orientation usually differs between the scans and the films are not registered.

In general, medical image registration has to support the physician by transforming the images into a common coordinate system. The computation of spatial relations between corresponding tissue provides the link between the different technologies and allows to make a diagnosis based on several acquisitions and modalities. At first glance, this theoretical statement of the registration problem may look rather simple, as we “just” have to compute the correspondences between two images. Every radiologist may well be able to succeed in this task, but in practice, finding a good solution without the need for manual interactions is very complex. Challenges arise, for instance, from differences in the multi-modal image content, noise, or patient movements.

The present thesis focuses on methods for statistical image registration and deals with applications of the techniques in 3-D/3-D multi-modal rigid registration, epilepsy diagnosis, and shape-based segmentation. The entire registration is based on three parts: the transformation, the optimization, and the similarity measure. The choice of the transform, whether parametric or non-parametric, has a high impact on the computational complexity. Within the optimization, the algorithm adapts the transform to achieve an optimal similarity measure between the images. Therefore, this measure is the driving force of the entire approach. Techniques based on intensity statistics are currently state-of-the-art for the automatic registration of multi-modal

images. We briefly summarize pitfalls when it comes to an implementation and propose data-driven schemes to estimate a set of optimal parameters that are necessary to achieve both robust and accurate measurements. Additionally, we introduce a novel scheme for the decomposition of the parameter space for a rigid registration and propose a model for the evaluation of registration techniques.

1.1 Contributions

The present thesis deals with methods for statistical image registration, more precisely with the estimation of intensity statistics for multi-modal similarity measures, optimization aspects and a statistical evaluation of both parametric and non-parametric algorithms. As we follow up on ideas that have been proposed earlier in this field of research, the following listing briefly summarizes our main contributions:

- We propose an automatic, data-driven approach for the estimation of a set of implementation parameters for statistical similarity measures. The parameters are optimally adapted to the registration data.
- Current projection-based optimization techniques for rigid transforms are limited to 3-D applications and lead to 2-D registration subproblems that cannot be further decomposed. We present a novel projection scheme that allows a complete decomposition of the parameter search space. We propose a non-linear projection scheme to split the high-dimensional parameter space into subproblems of only one parameter. The related optimization problems for these subspaces can be solved efficiently due to the low computational complexity. The results are iteratively combined to establish a solution for the high-dimensional parameters.
- There are currently only few established evaluation techniques for non-parametric registration algorithms. We present a framework that is based on manually selected, anatomical labels. It incorporates a statistical model for the observer reliability and yields a quality measure, which may be used as a registration benchmark for both parametric and non-parametric registration algorithms.
- We propose an application of rigid and non-parametric registration techniques in epilepsy diagnosis. The combination of the two types of algorithms reduces the registration error for SPECT brain images, even if there are significant variations in the blood flow measurements.
- Image segmentation based on prior knowledge requires a statistical model of the shape variation. We apply registration techniques to find correspondences within the training data that are used to compute the principal modes of variation. The standard sum of squared differences similarity measure is extended by a shape curvature term to incorporate surface properties into the registration. An application of active shape models for the segmentation of kidney from CT images is utilized to compare the registration to a standard approach for the computation of the shape statistics.

Some chapters of this thesis contain material that has been published or submitted to conference proceedings and journals. The following list specifies these articles:

- [1] D. A. Hahn, V. Daum, J. Hornegger, W. Bautz, and T. Kuwert. “Difference Imaging of Inter- and Intra-Ictal SPECT Images for the Localization of Seizure Onset in Epilepsy”. In: E. C. Frey, Ed., *IEEE Nuclear Science Symposium and Medical Imaging Conference*, pp. 4331–4335, IEEE Nuclear & Plasma Sciences Society, Honolulu, HI, USA, October 2007.
- [2] D. A. Hahn, V. Daum, and J. Hornegger. “Automatic Parameter Selection for Multi-Modal Image Registration”. 2009. submitted to *IEEE Transactions on Medical Imaging*.
- [3] D. A. Hahn, V. Daum, J. Hornegger, and T. Kuwert. “Data-Driven Density Estimation applied to SPECT Subtraction Imaging for Epilepsy Diagnosis”. 2009. submitted to *MICCAI 2009, Workshop on Probabilistic Models for Medical Image Analysis*.
- [4] V. Daum, D. A. Hahn, and J. Hornegger. “A Nonlinear Projection Scheme for Fast Rigid Registration”. In: E. C. Frey, Ed., *IEEE Nuclear Science Symposium and Medical Imaging Conference*, pp. 4022–4026, IEEE Nuclear & Plasma Sciences Society, Honolulu, HI, October 2007.
- [5] M. Spiegel, D. A. Hahn, V. Daum, J. Wasza, and J. Hornegger. “Segmentation of kidneys using a new active shape model generation technique based on non-rigid image registration”. *Computerized Medical Imaging and Graphics*, Vol. 33, No. 1, pp. 29–39, 2009.

1.2 Document Overview

The present thesis is structured into three parts: background information on medical image registration, theory, and applications. The first part presents clinical applications that require image registration and describes some of the main challenges when the algorithms have to be incorporated into a clinical workflow. It contains background information of the registration methods, the similarity measures, and the related mathematical formulations. The theoretical work is comprised of an automatic parameter estimation approach for statistical similarity measures, projection-based registration, and an evaluation scheme for registration results. Within the applications part, we present applications of the proposed methods to 3-D/3-D registration of multi-modal data, SPECT difference imaging for epilepsy diagnosis, and shape-based segmentation. The thesis concludes with an outlook and a summary. In the appendix, we describe an implementation of the projection scheme on graphics hardware.

Part I

Background on Medical Image
Registration

Chapter 2

Clinical Applications

When images are acquired from a patient using a single or multiple scanning modalities, image registration is required to enable the fusion of the data and, thus, provide the physician with additional information. In this context, we use the term *fusion* to indicate that the data is visualized within the same coordinate system after the registration, and an overlay is displayed to the user with variable transparency. Especially for functional modalities, a fusion with the corresponding CT or MR image may fill the morphological gaps and provide anatomical information. Figure 2.1 shows an example for a CT-PET image fusion for screening in lung cancer. Using only the CT, a small lung nodule with a high tracer uptake in the PET image could have easily been overseen. The combination of CT with PET or SPECT is commonly used in

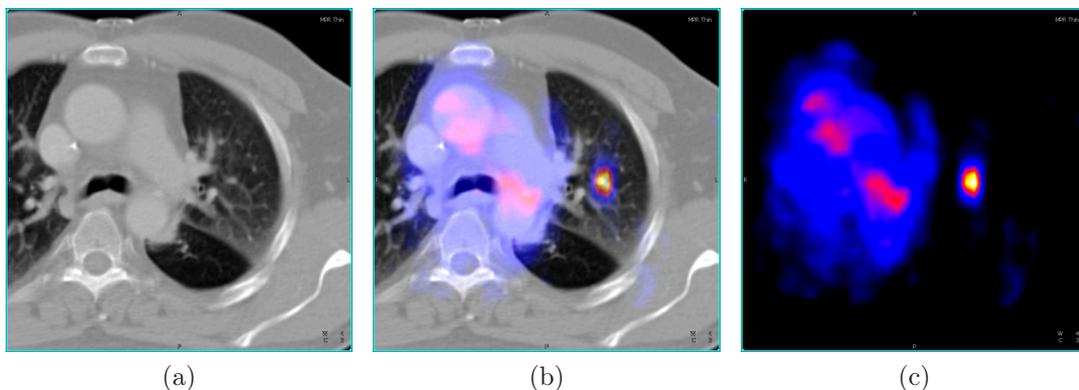


Figure 2.1: Example for an image fusion after registration of (a) a CT and (c) a PET image during a screening for metastases in lung cancer. The (b) fusion image provides the information about the function and the anatomical location of the lesion.

the clinics to locate lesions in cancer screening or also tissue inflammations. In addition, hardware registrations can be achieved by combining different modalities into one scanner. Recent innovations in hybrid scanning technologies led to PET-CT, SPECT-CT and PET-MR devices that are mitigating into clinical practice. With these, the patient does not have to be repositioned to acquire both scans, which basically eliminates the need for a retrospective, rigid registration. Studies report increased diagnostic accuracy using these devices, for example up to 30 percent for the nodal staging of thyroid carcinoma with SPECT-CT [Schm 08]. Results may lead

to the conclusion that the registration problem is solved by the invention of hybrid scanners, but retrospective registration techniques are still required. Although two modalities are combined within one device, the acquisition processes can be very different: PET, for instance, requires much longer scanning times than CT and effects due to respiration or other movements may lead to inaccuracies. The field of clinical usage is still limited for hybrid devices, and, additionally, there can hardly be a combination of devices for all necessary applications.

The following chapter presents some exemplary clinical applications that require different techniques of retrospective medical image registration, moreover, the usage of hybrid scanners does not yield substantial advantages. The requirements vary with the clinical purpose, therefore, it is necessary to determine a suitable transformation model and to choose an appropriate similarity measure with respect to the input images.

2.1 Radiation Therapy

Radiation therapy may be applied as part of a cancer treatment to reduce malignant tissue by the medical use of ionizing radiation. The goal is to destroy the cancer cells while sparing the healthy tissue surrounding the lesion. A clinical system has to provide a planning phase for dose calculations, support for radiation field adjustments, and an optimization of the distributed radiation. The calculations are largely based on 3-D CT data acquired before the treatment. During a treatment session, which may take up to 30 minutes, the patient and the lesions may move because of voluntary and involuntary actions, e. g. respiration or bowel motion [Bucc 05]. 2-D X-ray projection images can, for instance, be acquired with a C-arm system during the treatment and registered with the 3-D planning data to readjust the radiation device [Muac 06]. The registration result can then be used to compensate the motion and “follow” the lesion with the beam during the respiratory cycle [Muac 07]. Benefits for the patient are faster radiation sessions compared to gating, less damage to healthy tissue, and also a reduction of harmful radiation, as the malignant tissue can be destroyed more precisely.

2.2 Follow-Up Patient Studies

In clinical, longitudinal studies, in which patients are monitored to identify the progress or the influences of the treatment on the disease, follow-up image acquisitions are often used to conduct quantitative measurements. An example is the analysis of subsequent CT scans of the brain to study the ventricular change during Alzheimer’s disease [Leon 89]. Registration is necessary to relate the baseline with the follow-up images at the different acquisition times, which is not trivial for large-scale, longitudinal MR data sets. The number of required transformations grows exponentially with the number of images contained in the study, and the algorithms must be adapted to achieve results within reasonable computation times [Csap 07]. In order to achieve accurate measurements, the registration algorithm has to account for rigid move-

ments, deformations of anatomical structures, loss or gain of weight, different states of respiration, digestion, or also tissue that has been removed by surgery.

2.3 Reconstruction

The importance of 3-D and higher dimensional medical imaging is steadily increasing, which demands for suitable reconstruction techniques. In several imaging modalities, the reconstruction is based on series of 2-D projections that are acquired at different angles around the patient. Between subsequent scans, it is likely that the patient moves unintentionally, which is a major problem for slow imaging techniques, like PET or SPECT [Tsui 00]. In CT, the problem can be alleviated by accelerating the acquisitions with a large number of parallel slices being acquired in one rotational sweep, which already allows for detailed reconstructions of the beating heart. AX acquisitions are often required during interventions, but usually take longer than a CT scan. The motion of the heart and the lungs, due to breathing or other small movements, are large enough to result in reconstruction artifacts that may be visually distracting. For the reconstruction of 3-D cardiac images, multiple sweeps are usually performed and projections close to the desired heart phase are retrospectively selected for a reconstruction. Image registration provides an alternative, as it allows the incorporation of the deformation of the beating heart into the reconstruction to enhance the image quality by a substantial reduction of motion blurring [Prum 06]. If the spatial relations between projections from different heart phases are established by image registration, the number of required sweeps can be reduced while retaining the image quality. In addition, less dose has to be applied to the patient.

2.4 Challenges

The mentioned clinical applications all have specific requirements that a registration technique has to fulfill. In practice, challenges arise from the need to provide reliable and accurate registration results for various modalities, changes in image quality, different fields of view, varying resolutions, and a huge number of images. Implementations of registration algorithms usually need quite a few engineering variables that have to be specified. Unfortunately, these values are often dependent on the data and influence each other. We are convinced that this is a major reason why various research groups report partly contradictory, empirical proposals for these values. A single, empirically determined parameter configuration that is optimal for all input data cannot be found. Additionally, numerical issues regarding the discretization of the mathematical formulations have to be examined, otherwise the result may be biased towards a wrong solution and, therefore, unreliable. Another criterion, which is more important for clinical usage than research, is the time needed for the computation of the spatial transform. For a rigid registration, a runtime of more than five seconds is no longer tolerated, and non-parametric deformations have to be computed in less than a minute. These problems are discussed in more detail in the theoretical part of this thesis.

Chapter 3

Registration Methods

The main purpose of image registration is to find spatial relations between two or more images. These relations then provide the necessary information about correspondences between the images' content. They can be mathematically expressed using a spatial transform $\Phi : \mathbb{R}^n \mapsto \mathbb{R}^m$, where n is the dimension of the first image, which we call *reference image* $\mathcal{R} : \Omega(\mathcal{R}) \mapsto \mathbb{R}$, and m the dimension of the related image, the so-called *template image* $\mathcal{T} : \Omega(\mathcal{T}) \mapsto \mathbb{R}$. $\Omega(\mathcal{R}) \subset \mathbb{R}^n$ and $\Omega(\mathcal{T}) \subset \mathbb{R}^m$ are the bounded domains of the reference and the template image, respectively. In this thesis, we focus on registration problems of two images that have the same number of dimensions, i. e. $n = m$. The storage format of medical images – usually given in the *Digital Imaging and Communications in Medicine* (DICOM) standard [NEMA 09] – describes images as a lattice within the discrete image domain with intensity values provided at each lattice point together with spatial locations in the scanner's coordinate system. The application of spatial transformations requires means of interpolation to assess intensity values between the discrete image grid. We formulate the direction of the transformation from the reference to the template domain, the interpolated intensities at the corresponding locations are described using the term \mathcal{T}_Φ :

$$\mathcal{T}_\Phi(\mathbf{x}) = \mathcal{T}(\Phi(\mathbf{x})) . \quad (3.1)$$

During the registration, a similarity measure $\mathcal{D} : \Omega(\mathcal{R}, \mathcal{T}_\Phi) \mapsto \mathbb{R}$ between the two images is minimized with respect to the spatial transform Φ . Only the bounded region is taken into account that contains information from both images, i. e. $\Omega(\mathcal{R}, \mathcal{T}_\Phi) = \Omega(\mathcal{R}) \cap \Omega(\mathcal{T}_\Phi)$. The general objective function for the optimal transform $\hat{\Phi}$ is then:

$$\hat{\Phi} = \underset{\Phi}{\operatorname{argmin}} \mathcal{D}[\mathcal{R}, \mathcal{T}_\Phi] . \quad (3.2)$$

Depending on the type of transformation, registration algorithms are commonly divided into parametric and non-parametric classes [Mode04]. Some authors prefer the partitioning into rigid and non-rigid algorithms, however, this distinction is a bit vague, as there exist both parametric and non-parametric transformations that are non-rigid. As the formulation of the general objective function is basically the same for all parametric registrations, we favor the first type of distinction and briefly summarize some of the most popular parametric and non-parametric approaches in the following.

3.1 Parametric Registration

Several surveys presented in literature cover a wide range of parametric registration approaches. Examples can be found, for instance, in Brown [Brow 92], Elsen et al. [van 93], Hajnal et al. [Hajn 01], Hill et al. [Hill 01], and Maintz et al. [Main 98]. These parametric registration algorithms restrict the degrees of freedom (DOF) in the transformation by an inherent regularization, as only a restricted set of spatial relations is described by a parameter vector \mathbf{a} . The subject for all parametric registration optimizations is the parameter vector that defines the optimal transform:

$$\hat{\mathbf{a}} = \underset{\mathbf{a}}{\operatorname{argmin}} \mathcal{D}[\mathcal{R}, \mathcal{T}_{\Phi_{\mathbf{a}}}] . \quad (3.3)$$

Common types of parametric transforms that are applied in image registration are rigid, affine, and B-spline registrations. The rigid transform only allows for rotations $\mathbf{R} \in \mathbb{R}^{n \times n}$, with \mathbf{R} being an orthonormal matrix of $\det(\mathbf{R}) = 1$, and translations $\mathbf{t} \in \mathbb{R}^n$:

$$\Phi_{\mathbf{a}}^{\mathbf{R}}(\mathbf{x}) = \mathbf{R}\mathbf{x} + \mathbf{t} . \quad (3.4)$$

For clinical usage, the rigid registration has the advantage that the image content is not modified and results can be verified easily by the physicians. The accuracy of the algorithms is assessed using established evaluation strategies based on a gold standard of implanted fiducial markers (see Chapter 7).

The affine transform extends the DOF by adding individual scalings and also shearing within the linear transformation matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, with $\det(\mathbf{A}) > 0$:

$$\Phi_{\mathbf{a}}^{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{t} . \quad (3.5)$$

Medical images typically contain information about the extent of the voxels, the image origin, and the orientation of the coordinate system within the storage format. An affine transform between images is, therefore, useful to correct scale or skew inaccuracies of the scanner geometry. For general clinical applications, affine registrations do not yield major advantages over rigid ones, as both types of transforms cannot account for more complex soft tissue deformations. However, it is used in practice for image corrections due to scanner-related inaccuracies, and for approximate alignments of brain images from different subjects [Hajn 01]. An additional application is the compensation of gantry tilt, which may be realized using an affine transform.

Both rigid and affine registrations are often not sufficient to describe motions within the human body between acquisitions or between different subjects. B-spline approaches provide more DOF, while maintaining a high amount of control due to their local support, which yields also a high degree of computational efficiency. The deformation may be modeled using a control grid that defines the B-spline coefficients. The result is a smooth and C^{d-1} -continuous transformation, with $d \geq 1$ being the spline degree [Ashb 99]. Following the notation of Rueckert et al. [Ruec 99] and Ino et al. [Ino 05], the cubic B-spline transformation in 3-D is defined by a discrete grid of control points \mathcal{C} of size $n_x \times n_y \times n_z$ and a uniform spacing of δ , which depends on the desired strength of the deformation. The physical spacings between the control points are usually chosen large in the beginning and reduced during the course of the iterative optimization scheme. Large spacings but few control grids allow for

larger deformations in the beginning of the optimization. The deformation becomes increasingly local with a reduction of the control grid spacing and more control points. The coordinate vector $\mathbf{x} = (x, y, z)^T$ specifies a position within the image grid and $(i, j, k)^T$ denotes the corresponding index of the control grid:

$$\Phi_{\mathbf{a}}^{\text{BS}}(\mathbf{x}) = \mathbf{x} + \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 B_l(u)B_m(v)B_n(w)\mathcal{C}(i+l, j+m, k+n) . \quad (3.6)$$

The relations between the indices are given by $i = \lfloor x/\delta \rfloor - 1$, $j = \lfloor y/\delta \rfloor - 1$, $k = \lfloor z/\delta \rfloor - 1$. The cubic B-spline basis functions $B_l, l = 0, 1, \dots, 3$, are evaluated at $u = x/\delta - \lfloor x/\delta \rfloor$, $v = y/\delta - \lfloor y/\delta \rfloor$, and $w = z/\delta - \lfloor z/\delta \rfloor$. The parameter vector \mathbf{a} for the transformation contains the points of the control grid:

$$\mathbf{a} = (\mathcal{C}(1, 1, 1), \mathcal{C}(2, 1, 1), \dots, \mathcal{C}(n_x, n_y, n_z)) . \quad (3.7)$$

For a uniformly distributed control grid, the basis functions are:

$$B_0(u) = \frac{1}{6}(1 - u^3) \quad (3.8)$$

$$B_1(u) = \frac{1}{6}(3u^3 - 6u^2 + 4) \quad (3.9)$$

$$B_2(u) = \frac{1}{6}(-3u^3 + 3u^2 + 3u + 1) \quad (3.10)$$

$$B_3(u) = \frac{1}{6}u^3 . \quad (3.11)$$

Medical applications include, for instance, breast image registration [Ruec 99], atlas generation [Kybi 03], and correction of respiratory motion [Bai 09].

3.2 Non-Parametric Registration

Compared to the previous class of registration algorithms, non-parametric techniques allow to incorporate even more DOF and can be used to model nonlinear soft tissue deformations. However, the results are far more difficult to analyze in a quantitative way. The overlay visualization is no longer sufficient, as also the deformation has to be investigated. Various approaches for non-parametric registration can be found in the works of Rumpf, Miller, Modersitzki and co-workers [Dros 04; Clar 06; Chri 97; Chri 96; Josh 00; Fisc 04; Mode 04]. This section briefly introduces the variational framework that we applied for non-parametric registration tasks. It is based on the works of Hermosillo et al. [Herm 02b], Clarenz et al. [Clar 06], and Modersitzki [Mode 04].

The main difference between parametric and non-parametric methods, in general, is that the transformation is not restricted by a set of parameters in the latter case. Instead, it is defined by a displacement field $\mathbf{u} : \Omega(\mathcal{R}) \mapsto \mathbb{R}^n$:

$$\Phi(\mathbf{x}) = \mathbf{x} - \mathbf{u}(\mathbf{x}) . \quad (3.12)$$

Again, the aim of the optimization is to compute a transform that is optimal regarding (3.2), but the minimization of only the measure without any constraints on the

deformation is an ill-posed problem. Especially for medical images, the transform has to be physically meaningful, for example cracks and foldings within a displacement field rarely have a physical correspondence within real soft tissue deformations. One way to address this problem is to add a regularization energy \mathcal{S} to (3.2) and minimize the resulting functional $\mathcal{J} : \mathcal{U} \mapsto \mathbb{R}$ with respect to an admissible displacement field from the set of all deformations: $\mathbf{u} \in \mathcal{U}$. \mathcal{S} is often called a *smoother* because it is a measure of the regularity – or vaguely the smoothness – of \mathbf{u} . The mathematical meaning of the term smoothness is eventually determined by its actual formulation as a regularization energy term.

There exist various ways to incorporate the regularization term into the registration functional [Clar 06]. In techniques related to classical Tikhonov regularization [Tikh 77], an additional energy is added to the functional. Thus, the smoother penalizes irregular displacement fields and enables a stable numeric optimization [Mode 04]. Another class of regularization is comprised of gradient flow approaches that use the regularization energy to compute a smooth path from the initial guess to the minimizer of the functional. Examples for such iterative relaxation techniques can be found, for instance, in the work of Thirion [Thir 98], Clarenz et al. [Clar 06], and Henn et al. [Henn 02]. Also mixtures between classical Tikhonov and iterative registrations are used in some applications [Dros 04].

In relation to classical Tikhonov regularization, the constrained energy functional yields:

$$\mathcal{J}[\mathbf{u}] = \mathcal{D}[\mathcal{R}, \mathcal{T}_\Phi] + \alpha \mathcal{S}[\mathbf{u}] , \quad (3.13)$$

with $\alpha \in \mathbb{R}_{>0}$ being a positive weighting factor for the smoothness constraint. The optimization problem becomes more complex, because we now have to find a minimizer for the functional in the space of deformations:

$$\hat{\mathbf{u}} = \underset{\mathbf{u} \in \mathcal{U}}{\operatorname{argmin}} (\mathcal{D}[\mathcal{R}, \mathcal{T}_\Phi] + \alpha \mathcal{S}[\mathbf{u}]) . \quad (3.14)$$

If we assume that \mathcal{J} is sufficiently smooth and differentiable, and with appropriate boundary conditions, we can apply the calculus of variations to find a minimizer $\hat{\mathbf{u}}$. For the direction $\mathbf{v} \in \mathcal{U}$ of the first variation, the Gâteaux derivative of (3.13) is defined as:

$$\delta \mathcal{J}[\mathbf{u}, \mathbf{v}] = \lim_{\epsilon \rightarrow 0} \frac{\mathcal{J}[\mathbf{u} + \epsilon \mathbf{v}] - \mathcal{J}[\mathbf{u}]}{\epsilon} = \left. \frac{d\mathcal{J}[\mathbf{u} + \epsilon \mathbf{v}]}{d\epsilon} \right|_{\epsilon=0} . \quad (3.15)$$

For the existence of $\hat{\mathbf{u}}$ it is necessary that the Gâteaux derivative vanishes for all variations \mathbf{v} : $\delta \mathcal{J}[\hat{\mathbf{u}}, \mathbf{v}] = 0$. If \mathcal{U} is assumed to be a Hilbert space that defines a scalar product, the gradient of the functional with respect to the optimal displacement vanishes, $\nabla_{\mathcal{U}} \mathcal{J}[\hat{\mathbf{u}}] = 0$, and the minimizer is a solution to the Euler-Lagrange equations associated with this problem. Due to the nonlinear nature of the similarity measure, finding a closed form solution for (3.14) is usually impossible. Therefore, the functional has to be minimized numerically with a suitable optimization scheme. In a gradient descent method, for instance, the displacement field can be iteratively refined by stepping in the direction of the negative gradient until a convergence criterion is met. The update rule for the displacement in the k -th iteration to minimize the functional is:

$$\mathbf{u}_{k+1} = \mathbf{u}_k - \tau_k (\nabla_{\mathcal{U}} \mathcal{D}[\mathcal{R}, \mathcal{T}_{\Phi_k}] + \alpha \nabla_{\mathcal{U}} \mathcal{S}[\mathbf{u}_k]) , \quad (3.16)$$

where τ_k is the stepsize and Φ_k the transformation using the deformation \mathbf{u}_k of the k -th iteration. The existence of $\hat{\mathbf{u}}$ depends on the nature of \mathcal{D} . It has to be continuous in \mathbf{u} and bounded below [Herm 02a]. The choice of a suitable smoother depends on the type of application. Common regularization techniques are based on Dirichlet, elasticity, fluidal, and higher order functionals. Among the latter ones, curvature regularization $\mathcal{S}^{\text{curv}}$ is an approach that features some advantages for medical image registration [Fisc 03b]:

$$\mathcal{S}^{\text{curv}}[\mathbf{u}] = \int_{\Omega(\mathcal{R})} |\Delta_{\mathbf{x}}\mathbf{u}|^2 \, d\mathbf{x} . \quad (3.17)$$

This regularization term does not penalize affine transformations and leads to smooth displacement fields. A non-parametric registration with curvature regularization is used to compute the displacement fields for the shape model generation in Chapter 9, and for the correction of the global bias in rigid registrations of SPECT brain images in Chapter 8.

We concentrate on the deformation model (3.12) through the rest of this thesis. This model is also called small deformation model, as it does not necessarily preserve the topology in case of larger deformations. The regularization term is used to control the degree of locality. In the case that larger deformations are expected, it may be more suitable to choose a diffeomorphic deformation model. Examples for diffeomorphic registration algorithms may be found, for instance, in Christensen [Chri 01], Beg et al. [Beg 05], Vercauteren et al. [Verc 07] or Ashburner [Ashb 07].

3.3 Similarity Measures

We have briefly summarized the theory of parametric and non-parametric registrations, but a key aspect remains to be discussed in the following: the measure that is used to determine the similarity between the images. Image similarity measures can be coarsely divided into feature-based, intensity-based, and mixtures of both. Figure 3.1 illustrates a classification scheme for commonly applied similarity measures. Feature-based measures require the specification of similar features within the two images, for example by manual or automatic detection, or segmentation. Implicit approaches are calculated from voxel intensities and do not require manual interactions. A mix of both strategies can be applied, for example if both the detected features and the surrounding image intensities have to be incorporated into the registration [Mode03]. We provide the formulations for some measures in the following. Note that the functions are denoted as similarities measures that have to be minimized for an optimal alignment.

3.3.1 Feature-Based Measures

Feature-based similarity measures shift the registration problem from the intensity to the geometrical domain. Points and surfaces are two examples for established types of geometrical features. In a preprocessing step, which may consist of manual interactions or an automatic segmentation, the features are selected based on the image content. Ideally, the correspondences are known a priori, which greatly simplifies the

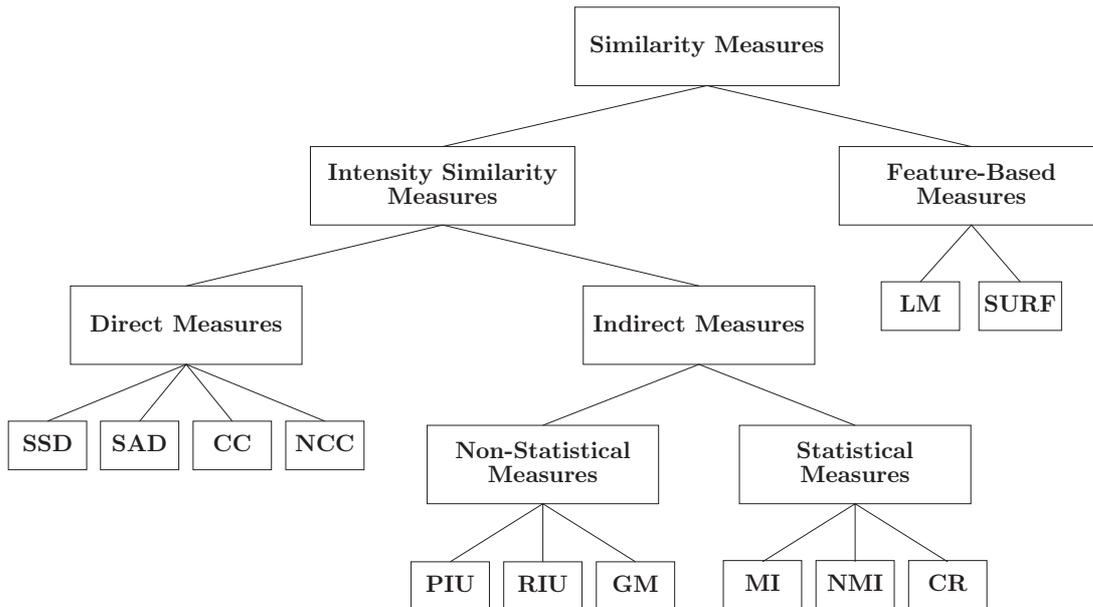


Figure 3.1: Coarse classification of image similarity measures into intensity and feature-based approaches. (See the description in the text or in Appendix C for the meaning of the abbreviations.)

optimization problem, as analytical solutions can be found for rigid and spline-based registrations. If the preprocessing consists of an automatic extraction algorithm, the problem of selecting corresponding features has to be solved in addition.

Landmark (LM) Registration

Anatomical landmarks are often used as point features for registration tasks. A common approach within current state-of-the-art registration applications requires the physician to specify the location and correspondences of the points. The mathematical formalism to establish a transform that yields a least squares distance between two sets of corresponding points is known as the orthogonal *Procrustes problem*. Solutions for the least squares fitting of two point sets using a rotation together with a scaling factor and a translation have been established by Schönemann [Scho66], Arun et al. [Arun87] and Horn [Horn87]. Fitzpatrick et al. [Fitz98] summarized the methodology for the registration of automatically detected fiducial markers, which have been implanted into the human skull, from images acquired with various modalities. As the markers can be detected with a high accuracy, the authors used their method as ground truth for comparisons of the target registration error (TRE) between rigid registration algorithms within their retrospective image registration evaluation project (RIRE) [West97].

For rigid registrations, a standard solution is based on a singular value decomposition (SVD) [Golu96]. Let $\mathbf{L}_{\mathcal{R}} = \{\mathbf{x}_{\mathcal{R}}\}_{i=1,\dots,N}$ be the set of N geometric landmarks in the reference and $\mathbf{L}_{\mathcal{T}} = \{\mathbf{x}_{\mathcal{T}}\}_{i=1,\dots,N}$ the corresponding points in the template

image coordinate space. The problem is to find a rigid transform with $\mathbf{a} = (\mathbf{R}, \mathbf{t})$ that results in the minimal squared distance between the landmarks:

$$\mathcal{D}^{\text{LM}}[\mathbf{L}_{\mathcal{R}}, \mathbf{L}_{\mathcal{T}, \Phi_a^{\mathbf{R}}}] = \frac{1}{N} \sum_{i=1}^N \|\Phi_a^{\mathbf{R}}({}^i\mathbf{x}_{\mathcal{R}}) - {}^i\mathbf{x}_{\mathcal{T}}\|^2. \quad (3.18)$$

Given the mean points $\bar{\mathbf{x}}_{\mathcal{R}}$ and $\bar{\mathbf{x}}_{\mathcal{T}}$ of the point sets, a solution for \mathbf{R} can be found by computing the SVD for the outer product of the matrices $\mathbf{X}_{\mathcal{R}}$ and $\mathbf{X}_{\mathcal{T}}$. The reference matrix, with $\mathbf{X}_{\mathcal{R}} = ({}^1\hat{\mathbf{x}}_{\mathcal{R}}, \dots, {}^N\hat{\mathbf{x}}_{\mathcal{R}})$, consists of the demeaned reference points ${}^i\hat{\mathbf{x}}_{\mathcal{R}} = {}^i\mathbf{x}_{\mathcal{R}} - \bar{\mathbf{x}}_{\mathcal{R}}$ denoted as column vectors. The corresponding matrix $\mathbf{X}_{\mathcal{T}} = ({}^1\hat{\mathbf{x}}_{\mathcal{T}}, \dots, {}^N\hat{\mathbf{x}}_{\mathcal{T}})$ contains the demeaned template image points ${}^i\hat{\mathbf{x}}_{\mathcal{T}} = {}^i\mathbf{x}_{\mathcal{T}} - \bar{\mathbf{x}}_{\mathcal{T}}$, respectively. The SVD yields:

$$\mathbf{M} = \mathbf{X}_{\mathcal{R}}\mathbf{X}_{\mathcal{T}}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (3.19)$$

$$\mathbf{R} = \mathbf{V}\mathbf{U}^T, \quad (3.20)$$

where (3.19) is the SVD of the measurement matrix \mathbf{M} . The translation is then determined by the displacement of the rotated reference features to the template mean:

$$\mathbf{t} = \bar{\mathbf{x}}_{\mathcal{T}} - \mathbf{R}\bar{\mathbf{x}}_{\mathcal{R}}. \quad (3.21)$$

If the correspondences between the point features are not known prior to the registration, or if the features have been automatically detected independently by an automatic extraction, an analytical solution cannot be found. Instead, iterative schemes are usually applied to minimize the least squares problem, for instance, using an iterative closest point algorithm (ICP) [Besl92], or detecting the similarity between feature points using neighboring image intensities [Huan04; Hahn06].

Point Features in Non-Parametric Registration

The problem of finding point-to-point correlations using non-parametric registration algorithms requires, again, the solution of a least squares energy functional with additional constraints. Rohr [Rohr01] applied thin-plate splines (TPS) and presented an analytical solution that also incorporates an uncertainty measurement for the landmarks. His energy functional is composed of a quadratic landmark registration error together with a bending energy that is dictated by the spline. The according partial differential equation (PDE) can be solved using a system of linear equations, which yields the TPS-regularized deformation field. Modersitzki and Fischer [Fisc03a] directly incorporated the quadratic landmark distance into their registration functional and provided the corresponding Euler-Lagrange equations. If just the landmark energy is applied, the system is linear with respect to the smoother. The authors mixed landmark constraints with an additional similarity term that is based on image intensities to improve the registration quality in the neighborhood regions. As the similarity measure between the intensities is nonlinear, this mix between point distances and an intensity similarities results in a PDE that requires a nonlinear optimization.

Surface Registration

Surfaces within medical images are usually the result of the extraction of boundaries between organs or different types of tissue using segmentation approaches. Compared to landmarks, surfaces provide more information, however, the direct correspondences between points located on the two related surfaces are commonly not known. Therefore, the solution can no longer be computed by an analytical, linear system of equations. Approaches for rigid surface registration usually make use of an ICP algorithm between sets of discretely sampled points on the reference and the template surface [Hill01].

For non-parametric registrations, a level set shape representation may be applied, which implicitly describes the surfaces and avoids the computation of explicit correspondences between the surfaces. For example, Han et al. [Han05] proposed a similarity energy $\mathcal{D}^{\text{SURF}}$ based on the signed Euclidean distances $\Gamma_{\mathcal{R}}$ to the surface in the reference and $\Gamma_{\mathcal{T}}$ to the surface in the template domain, respectively:

$$\mathcal{D}^{\text{SURF}}[\Gamma_{\mathcal{R}}, \Gamma_{\mathcal{T}, \Phi}] = \frac{1}{|\Omega(\mathcal{R}, \mathcal{T}_{\Phi})|} \int_{\Omega(\mathcal{R}, \mathcal{T}_{\Phi})} (\Gamma_{\mathcal{R}}(\mathbf{x}) - \Gamma_{\mathcal{T}}(\Phi(\mathbf{x})))^2 \cdot d\mathbf{x} \quad (3.22)$$

The surfaces are implicitly represented as level set surfaces in the discrete image domain, where each intensity in the curve image specifies the signed distance to the zero level set.

3.3.2 Direct Voxel Intensity-Based Measures

Similarity measures that are based on the intensity values of the voxels are especially suitable for fully automatic registration tasks. For each measure, a specific relation between the image intensity values is assumed in order to define the maximal similarity. In the following, we distinguish between two subclasses: direct and indirect similarity measures. The direct measures require a functional relation between the image intensities, whereas the indirect measures are calculated on basis of statistics or preprocessed image content.

Sum of Squared Differences (SSD)

The SSD measure is based on the assumption that the intensities of corresponding tissue within the two images are equal, or differ by noise at the utmost:

$$\mathcal{D}^{\text{SSD}}[\mathcal{R}, \mathcal{T}_{\Phi}] = \frac{1}{|\Omega(\mathcal{R}, \mathcal{T}_{\Phi})|} \int_{\Omega(\mathcal{R}, \mathcal{T}_{\Phi})} (\mathcal{R}(\mathbf{x}) - \mathcal{T}_{\Phi}(\mathbf{x}))^2 d\mathbf{x} . \quad (3.23)$$

Sum of Absolute Differences (SAD)

Outliers in the intensity domain have a quadratic influence on the SSD measure and may, therefore, lead to inaccurate alignments. This may be the case, for example if beam artifacts of metal implants are present [Hill01]. These influences are reduced if the absolute instead of the squared distance is applied, however, the function is no

longer differentiable at the optimal position. Again, an equality between intensities of corresponding tissue within the two images is assumed:

$$\mathcal{D}^{\text{SAD}}[\mathcal{R}, \mathcal{T}_\Phi] = \frac{1}{|\Omega(\mathcal{R}, \mathcal{T}_\Phi)|} \int_{\Omega(\mathcal{R}, \mathcal{T}_\Phi)} |\mathcal{R}(\mathbf{x}) - \mathcal{T}_\Phi(\mathbf{x})| \, d\mathbf{x} . \quad (3.24)$$

Cross Correlation (CC)

The rather strict assumption of equal iso-intensity surfaces, or level sets, within both images is loosened by assuming a linear relation between the intensities. The position of the optimum of the CC measure with respect to Φ is invariant to a scaling factor between the image values:

$$\mathcal{D}^{\text{CC}}[\mathcal{R}, \mathcal{T}_\Phi] = - \int_{\Omega(\mathcal{R}, \mathcal{T}_\Phi)} \mathcal{R}(\mathbf{x}) \mathcal{T}_\Phi(\mathbf{x}) \, d\mathbf{x} . \quad (3.25)$$

Normalized Cross Correlation (NCC)

An extension of the CC is achieved by modeling the intensity relation by an affine transform, which yields the NCC measure:

$$\begin{aligned} \mathcal{D}^{\text{NCC}}[\mathcal{R}, \mathcal{T}, \Phi] &= \\ &= - \frac{\int_{\Omega(\mathcal{R}, \mathcal{T}_\Phi)} (\mathcal{R}(\mathbf{x}) - \mu_{\mathcal{R}}) (\mathcal{T}_\Phi(\mathbf{x}) - \mu_{\mathcal{T}_\Phi}) \, d\mathbf{x}}{\sqrt{\int_{\Omega(\mathcal{R}, \mathcal{T}_\Phi)} (\mathcal{R}(\mathbf{x}) - \mu_{\mathcal{R}})^2 \, d\mathbf{x} \int_{\Omega(\mathcal{R}, \mathcal{T}_\Phi)} (\mathcal{T}_\Phi(\mathbf{x}) - \mu_{\mathcal{T}_\Phi})^2 \, d\mathbf{x}}} \\ \mu_{\mathcal{R}} &= \int_{\mathbb{R}} r p_{\mathcal{R}}(r) \, dr \\ \mu_{\mathcal{T}_\Phi} &= \int_{\mathbb{R}} t p_{\mathcal{T}_\Phi}(t) \, dt , \end{aligned} \quad (3.26)$$

with $p_{\mathcal{R}}(r)$ and $p_{\mathcal{T}_\Phi}(t)$ being the probability density functions (PDFs) on the intensities in the reference and the transformed template image. The mean values of the reference and template image are defined by $\mu_{\mathcal{R}}$ and $\mu_{\mathcal{T}_\Phi}$, respectively. The intensities r of the reference and t of the template image are regarded as sample values of random variables associated with the intensity information of the images within the overlap domain $\Omega(\mathcal{R}, \mathcal{T}_\Phi)$ (see also Section 3.3.4).

3.3.3 Non-Statistical Indirect Measures

For some combinations of imaging modalities, direct similarity measures may often be sufficient. For example in CT images, the intensity values are normalized to Hounsfield units (HU), and a direct correspondence between image content and intensity values is established. Registrations between other modalities are more difficult, as the intensities are not normalized. Images from different MR sequences, for instance, cannot be related to each other by linear functions, therefore, the similarity is calculated on basis of derived intensity characteristics, such as ratio images, partitions, or normal fields.

Ratio Image Uniformity (RIU)

The RIU [Wood 92] measures the standard deviation of the ratio between the intensities of the reference image and the transformed template image within the overlap domain. The ratio between the images at a specific position within the overlap domain is:

$$\mathcal{I}_{\text{RIU}}(\mathbf{x}) = \frac{\mathcal{R}(\mathbf{x})}{\mathcal{T}_{\Phi}(\mathbf{x})}, \text{ with } \mathbf{x} \in \Omega(\mathcal{R}, \mathcal{T}_{\Phi}) \text{ and } \mathcal{T}_{\Phi}(\mathbf{x}) \neq 0. \quad (3.27)$$

The mean ratio μ_{RIU} is computed using the PDF p_{RIU} of the intensity ratios i :

$$\mu_{\text{RIU}} = \int_{\mathbb{R}} i p_{\text{RIU}}(i) \, di \quad (3.28)$$

$$\mathcal{D}^{\text{RIU}}[\mathcal{R}, \mathcal{T}_{\Phi}] = \frac{\sqrt{\frac{1}{|\Omega(\mathcal{R}, \mathcal{T}_{\Phi})|} \int_{\Omega(\mathcal{R}, \mathcal{T}_{\Phi})} (\mathcal{I}_{\text{RIU}}(\mathbf{x}) - \mu_{\text{RIU}})^2 \, d\mathbf{x}}}{\mu_{\text{RIU}}}, \quad (3.29)$$

where $\mu_{\text{RIU}} \neq 0$.

Partition Image Uniformity (PIU)

So far, the presented measures are only suitable for intra-modality or intra-sequence registrations. In order to support also multi-modal data, the concept of ratios between image intensities has been extended to ratios between iso-intensity sets (level sets) by Woods et al. [Wood 93]. The PIU measures the deviation of locations on the level sets of the reference image to corresponding intensities in the transformed template image. It has been successfully applied to MR-PET registrations, however, the scalp in the MR images has to be removed in a preprocessing step to adhere to the assumptions. The definition requires the PDF $p_{\mathcal{T}_{\Phi}|\mathcal{R}(\mathbf{x})=r}$ of the template image at a reference iso-value r , and the corresponding mean value $\mu_{\mathcal{T}_{\Phi}|\mathcal{R}(\mathbf{x})=r}$:

$$\mathcal{D}_{\mathcal{R}}^{\text{PIU}}[\mathcal{R}, \mathcal{T}_{\Phi}] = \int_{\mathbb{R}} \frac{\sqrt{\int_{\Omega(\mathcal{R}, \mathcal{T}_{\Phi})|\mathcal{R}(\mathbf{x})=r} (\mathcal{T}_{\Phi}(\mathbf{x}) - \mu_{\mathcal{T}_{\Phi}|\mathcal{R}(\mathbf{x})=r}(r))^2 \, d\mathbf{x}}}{\mu_{\mathcal{T}_{\Phi}|\mathcal{R}(\mathbf{x})=r}(r)} p_{\mathcal{R}}(r) \, dr. \quad (3.30)$$

Gaussian Maps (GM)

The idea of using level sets within two images has also been applied by Droske et al. [Dros03; Dros05] who assume that the morphology between the input images is very similar. They applied differential geometry to define a measure for the deviation between corresponding level set surfaces. The distance between the morphology is defined as the Lebesgue-measure between the image gradients:

$$\mathcal{D}^{\text{GM}}[\mathcal{R}, \mathcal{T}_{\Phi}] = \int_{\Omega(\mathcal{R}, \mathcal{T}_{\Phi})} g_0(\nabla_{\mathbf{x}}\mathcal{R}(\mathbf{x}), \nabla_{\mathbf{x}}\mathcal{T}_{\Phi}(\mathbf{x}), \text{Cof}(J_{\Phi}(\mathbf{x}))) \, d\mathbf{x}, \quad (3.31)$$

where $\text{Cof}(J_{\Phi})$ is the cofactor matrix of the Jacobian of the transform and g_0 is an extension of the function g to avoid singularities that occur if normalizing zero-length gradients directly, i. e. $g_0(\mathbf{x}, \mathbf{y}, \mathbf{A}) = 0$ if $\|\mathbf{x}\| = 0$ or $\|\mathbf{y}\| = 0$, and $g_0(\mathbf{x}, \mathbf{y}, \mathbf{A}) =$

$g(\frac{\mathbf{x}}{\|\mathbf{x}\|}, \frac{\mathbf{y}}{\|\mathbf{y}\|}, \mathbf{A})$ otherwise. An example function is given by $g(\mathbf{x}, \mathbf{y}, \mathbf{A}) = \|(\mathbb{I} - \mathbf{x}\mathbf{x}^T) \cdot \mathbf{A}\mathbf{y}\|^\gamma$, with $\gamma \geq 1$. \mathcal{D}^{GM} can also be thought of a measure for the length of the normal vectors at the reference level sets projected onto the tangent space of the transformed template level sets.

A very similar approach was presented later by Haber and Modersitzki [Habe06]. Instead of a function extensions, they applied a modified vector norm that yields a differentiable distance between the normalized gradient fields. Both groups presented multi-modal registration examples, however, the examples were all taken from morphological modalities.

3.3.4 Statistical Measures

Nowadays, similarities based on image intensity statistics are widely used for multi-modal registration tasks. Unlike the direct and non-statistical examples mentioned above, statistics of image intensities can also be used to measure similarities between morphological and functional modalities. The statistics are defined by the marginal and joint PDFs described in Table 3.1. In the statistical framework, the intensity

PDF	Description
$p_{\mathcal{R}}(r)$	1-D PDF of reference image intensities r
$p_{\mathcal{T}_\Phi}(t)$	1-D PDF of template image intensities t
$\mathbf{p}_{\mathcal{R}, \mathcal{T}_\Phi}(\mathbf{i})$	2-D PDF of joint image intensities $\mathbf{i} = (r, t)$

Table 3.1: Description of the PDFs required for the statistical image similarity measures.

values r and t are associated with probabilities on the collection of the intensities of the images. r and t are regarded as random values with the associated PDFs $p_{\mathcal{R}}(r)$, $p_{\mathcal{T}_\Phi}(t)$, and $\mathbf{p}_{\mathcal{R}, \mathcal{T}_\Phi}(\mathbf{i})$. The PDFs all depend on the transformation Φ , as it determines the amount of overlap between the spatial domains.

Correlation Ratio (CR)

Roche et al. [Roch98] generalized the concept of NCC by measuring the functional dependency between the two images. They decomposed the variance $\text{VAR}[\mathcal{T}_\Phi]$ of the transformed template image into two opposite parts:

$$\text{VAR}[\mathcal{T}_\Phi] = \text{VAR}[\text{E}[\mathcal{T}_\Phi|\mathcal{R}]] + \text{VAR}[\mathcal{T}_\Phi - \text{E}[\mathcal{T}_\Phi|\mathcal{R}]] . \quad (3.32)$$

The first part consists of the variance of the conditional expected value and measures the prediction of \mathcal{T}_Φ by \mathcal{R} . The second variance is a measure of the amount in \mathcal{T}_Φ that is functionally independent from \mathcal{R} . There are two reasons why the second part may be low compared to the first one. First, there may be a strong functional relationship between the two random variables, which yields a large variance of the conditional expectation. In the second case, the variance in the random variable \mathcal{T}_Φ may already be very small, e. g. if \mathcal{T}_Φ consists of only background. As the distributions of the two random variables depend on the overlap domain, minimizing only the second part of (3.32) would tend to solutions with a small overlap domain and, therefore, to a

spatial disconnection of the two images. Instead, the authors suggested to minimize it with respect to the total variance:

$$\mathcal{D}^{\text{CR}}[\mathcal{R}, \mathcal{T}_\Phi] = \frac{\text{VAR}[\mathcal{T}_\Phi - \text{E}[\mathcal{T}_\Phi|\mathcal{R}]]}{\text{VAR}[\mathcal{T}_\Phi]} . \quad (3.33)$$

The following standard function definitions are used to compute the CR measure [Herm 02a]:

$$\mu_{\mathcal{T}_\Phi} = \int_{\mathbb{R}} t p_{\mathcal{T}_\Phi}(t) dt \quad (3.34)$$

$$\text{VAR}[\mathcal{T}_\Phi] = \int_{\mathbb{R}} t^2 p_{\mathcal{T}_\Phi}(t) dt - \mu_{\mathcal{T}_\Phi}^2 \quad (3.35)$$

$$\mathbf{p}_{\mathcal{R}, \mathcal{T}_\Phi}(t|r) = \frac{\mathbf{p}_{\mathcal{R}, \mathcal{T}_\Phi}(\mathbf{i})}{p_{\mathcal{R}}(r)} \quad (3.36)$$

$$\mu_{\mathcal{T}_\Phi|\mathcal{R}}(r) = \int_{\mathbb{R}} t \mathbf{p}_{\mathcal{R}, \mathcal{T}_\Phi}(t|r) dt \quad (3.37)$$

$$\text{VAR}[\mathcal{T}_\Phi|\mathcal{R}](r) = \int_{\mathbb{R}} t^2 \mathbf{p}_{\mathcal{R}, \mathcal{T}_\Phi}(t|r) dt - \mu_{\mathcal{T}_\Phi|\mathcal{R}}(r)^2 \quad (3.38)$$

$$\text{VAR}[\mathcal{T}_\Phi - \text{E}[\mathcal{T}_\Phi|\mathcal{R}]] = \int_{\mathbb{R}} \text{VAR}[\mathcal{T}_\Phi|\mathcal{R}](r) p_{\mathcal{R}}(r) dr . \quad (3.39)$$

Mutual Information (MI)

Based on Shannon's theory [Shan 48], the information content within the images can be measured using the following entropies that require the marginal PDFs $p_{\mathcal{R}}$ and $p_{\mathcal{T}_\Phi}$, and the joint PDF $\mathbf{p}_{\mathcal{R}, \mathcal{T}_\Phi}$:

$$\mathcal{H}(\mathcal{R}) = - \int_{\mathbb{R}} p_{\mathcal{R}}(r) \log p_{\mathcal{R}}(r) dr \quad (3.40)$$

$$\mathcal{H}(\mathcal{T}_\Phi) = - \int_{\mathbb{R}} p_{\mathcal{T}_\Phi}(t) \log p_{\mathcal{T}_\Phi}(t) dt \quad (3.41)$$

$$\mathcal{H}(\mathcal{R}, \mathcal{T}_\Phi) = - \int_{\mathbb{R}^2} \mathbf{p}_{\mathcal{R}, \mathcal{T}_\Phi}(\mathbf{i}) \log \mathbf{p}_{\mathcal{R}, \mathcal{T}_\Phi}(\mathbf{i}) d\mathbf{i} . \quad (3.42)$$

Wells et al. [Well 96] and Maes et al. [Maes 97] applied the mutual information, which was originally developed as a measure between the input and output of a communication channel, to determine the similarity between two images:

$$\mathcal{D}^{\text{MI}}[\mathcal{R}, \mathcal{T}_\Phi] = - (\mathcal{H}(\mathcal{R}) + \mathcal{H}(\mathcal{T}_\Phi) - \mathcal{H}(\mathcal{R}, \mathcal{T}_\Phi)) . \quad (3.43)$$

The formulation in (3.43) incorporates the overlap domain between the images. Another formulation without the restriction to the overlap domain is used in Wells et al. [Well 96], who embedded the images content within an unlimited background domain and circumvented an explicit overlap requirement.

MI measures the lack of statistical dependence between the intensity distributions $p_{\mathcal{R}}$ and $p_{\mathcal{T}_\Phi}$. In the case of an ideal alignment, the statistical dependence between the two distributions is maximal. Note that, as well as in the case of the other

measures, \mathcal{D}^{MI} is a similarity measure that has to be minimized for a better alignment. Interestingly, the joint entropy $\mathcal{H}(\mathcal{R}, \mathcal{I}_\Phi)$ alone does not provide sufficient information for a correct alignment, similar to the variance measure for the functional relation in the CR. The global optimum found by optimizing only the joint entropy can be far from the desired position and may refer to, for instance, a good overlap between the background regions. The marginal entropies are necessary to avoid registration results where only background parts are aligned and the remaining image content is mapped outside the overlap domain.

Normalized Mutual Information (NMI)

Since its introduction into the field of medical image registration, MI has been successfully applied in various applications and is proposed in a large number of articles about multi-modal registration. However, it can sometimes be strongly dependent on the overlap domain between the images. If the background is extended with respect to the object, the probability for object elements in the image domain decreases. The joint entropy increases accordingly, but the peak of the MI at the position of correct alignment is flattened because no new object information is gained. The NMI helps to compensate this effect by dividing the marginal entropies by the joint entropy [Stud 99]:

$$\mathcal{D}^{\text{NMI}}[\mathcal{R}, \mathcal{I}_\Phi] = -\frac{\mathcal{H}(\mathcal{R}) + \mathcal{H}(\mathcal{I}_\Phi)}{\mathcal{H}(\mathcal{R}, \mathcal{I}_\Phi)}. \quad (3.44)$$

This does not mean that the NMI is invariant to the overlap domain, but it is less sensitive to changes of the amount of overlap than the original MI.

Part II

Theory

Chapter 4

Statistical Similarity Measures

The field of applications for non-statistical similarity measures in medical applications is limited. The physicians usually have tough time schedules and the clinical workflows need algorithms that require very few user interactions. Nonetheless, the registration results have to be reliable, accurate, and computed within a very short time – usually in just a few seconds. Fully automatic approaches are preferred to registrations requiring user interactions. Basically all of the aforementioned, non-statistical approaches have limitations when applied to multiple modalities, however, an implementation for clinical usage has to support a plurality of images from various modalities. For example, the computation of the GM measure within a CT-MR registration makes sense, however, if applied to PET-MR, it will likely fail because of the entirely different nature of the images. Thus, statistical similarity measures, namely MI and NMI, have become state-of-the-art in automatic, multi-modal image registration over the past ten years. Their computation requires PDFs of the reference, template, and joint intensities. The density functions are not known a priori, so they have to be estimated from intensity samples that are assumed to be independent of each other and identically distributed (i.i.d.) random measures. Although there exist approaches to approximate the PDFs with parametric or semi-parametric models, certain assumptions about the PDF shapes soon become unreliable for different modalities. The models would also have to be adapted to varying fields of views or changes in the reconstruction settings. Another technique for this task is non-parametric Parzen-window estimation that only requires a kernel PDF of appropriate width(s), a property that makes it very attractive.

The discretization of the statistical measure is not as straightforward as one might initially think. Various parameter values for an implementation have to be chosen with care to achieve a both robust and accurate measurement. Unfortunately, a generally applicable parameter configuration for an optimal discretization cannot be found, if the approach has to reliably work for large sets of images. The input data has a great influence on the values of the parameters. In our opinion, this explains why various research groups propose partly contradictory, empirical values for these variables.

An implementation of this PDF estimator requires variables for the kernel width and / or the number of histogram bins. Challenges arise from the fact that these values are dependent on the data and mutually influence each other. Furthermore,

to keep the computation times within the runtime requirements, it may be necessary to restrict the sampling to only a subset of the overlap domain. The number of samples also interferes with the previous parameters, i. e. it affects the width of the Parzen-windows and also the adaptive layout of the histogram. Some of the problems have been presented in related articles, but the empirical results are contradictory. There is ambiguity, for instance, about the optimal number of bins used for the discrete PDF representations, and whether or not it is necessary to automatically determine the kernel size for the Parzen-windowing. Additional problems arise from structured noise within the background region, especially in medical images. The artificial structures account for inaccuracies within the measurements, and there exist proposals in literature to treat the background of the images differently when evaluating the similarity. Tailoring all the necessary parameter values manually to the application problems may be possible for sets containing only few images. However, large-scale applications require the registration of hundreds, if not thousands, of images. Thus, the manual adjustment of the values soon becomes impossible. Wrong parameters can have a rather destructive impact on the result of the registration, as it is demonstrated in Figure 4.1 for an intra-patient CT-PET image pair. From two different starting positions, a rigid registration using NMI with default parameters (see Chapter 7 for a detailed description of the values) delivered wrong results. The outlines of the lesion, for instance, are clearly mismatched, which is also the case for the outlines of the brain.

The focus of this chapter lies on crucial numerical aspects of the joint PDF estimation, which is based on sampled random measures. A measure, in this context, consists of the image intensity values at related spatial positions within the overlap domain between the reference and the template image. From a numerical point of view, the sampling technique is very important, as the results are prone to grid effects that are dependent on the utilized pattern. We briefly summarize the steps to reduce the sampling artifacts using a jittering technique. The discretization of the PDF estimates based on these samples requires some parameters. Instead of choosing the values by empirical adjustment, data-driven approaches are used to adapt them to the input data. We extend ideas proposed by Viola [Viol95] and Hermosillo et al. [Herm02b] to present an efficient kernel width estimation algorithm that uses histogram binning. The commonly applied method of linear binning in the discrete histograms, the so-called isotropic binning, does not account for the variability within the structure of the PDF [Bish97]. Research results of Knops et al. [Knop06] or Katkovic and Shumulevich [Katk02] show that isotropic binning is outperformed by adaptive techniques. Our experiments also lead to the conclusion that adaptive histogram layouts improve the registration accuracy, as the structure of the estimated PDF plays an important role in the entire registration. Unfortunately, adaptive kernel sizes applied in the Parzen-windowing yield a high complexity and prevent fast computation techniques. Therefore, we propose a novel, quasi-adaptive scheme for the kernel width selection that is comprised of a combination of an isotropic estimator with an adaptive histogram binning layout. The adaptive computation of the histogram layout brings about an additional benefit in form of a measure for the quantization error that is used as a criterion to determine a suitable number of bins. Additionally, we propose a novel technique to tackle the aforementioned problem of

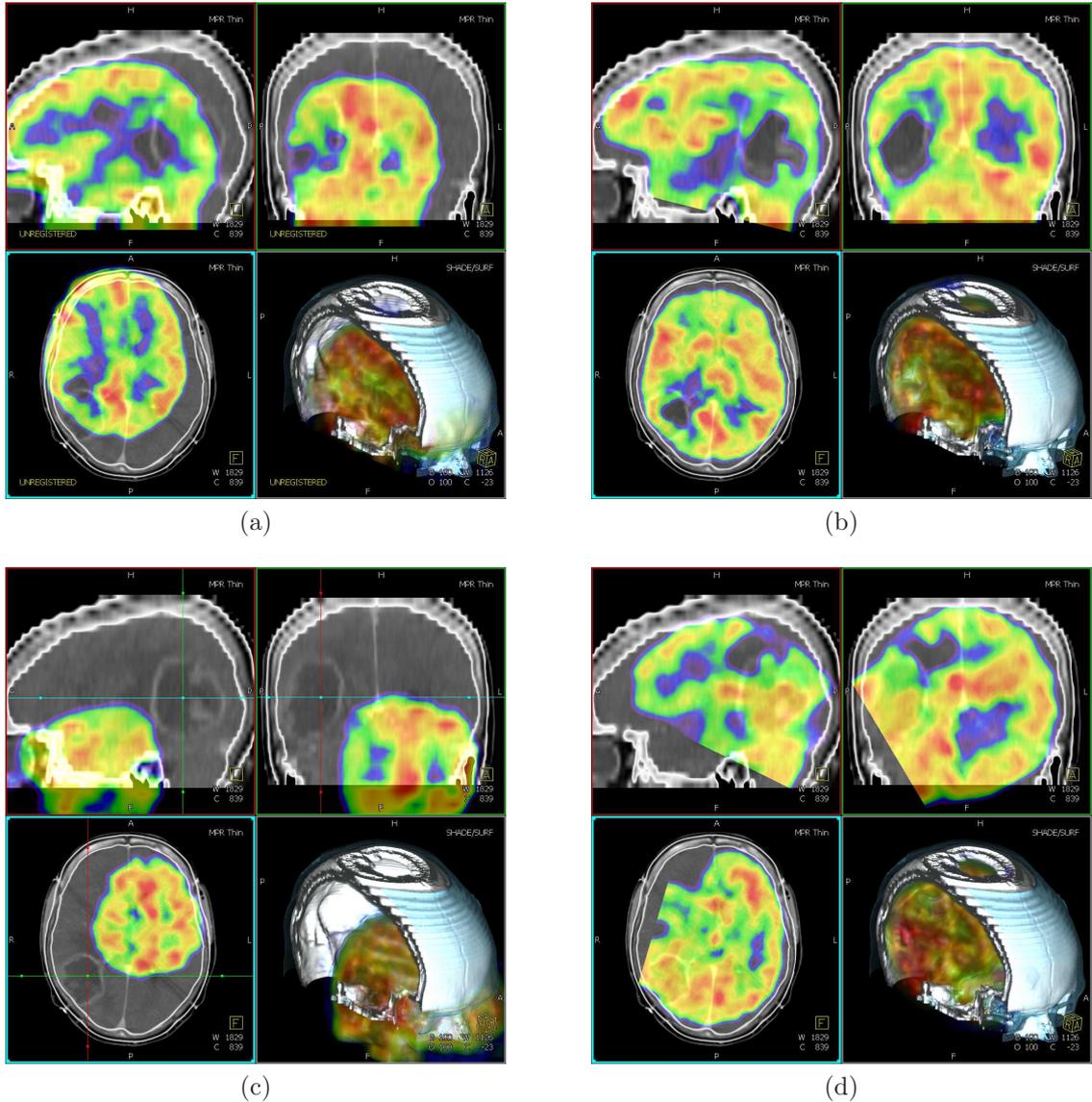


Figure 4.1: Demonstration of the resulting mismatch due to wrong parameters applied to a rigid registration using NMI. For two initial starting positions shown as overlay visualizations (a), (c), the results are provided in the right column (b), (d). For comparison, results using optimally adapted parameters are shown in Figure 4.12.

structured noise within the background. A weighting scheme between the probabilities for object-object and background-background correspondences is applied to the joint PDF to ensure that the background, and especially the noise within, does not dominate the similarity measure. A manual specification of implementation variables is no longer necessary, because the data-driven schemes lead to parameter values that are optimal with respect to the given input images. The number of samples may be automatically selected as well, based on the runtime requirements of the application.

4.1 Sampling Artifacts

Medical images are usually represented by a discrete lattice together with information about the position, orientation, and spacing. As the true density functions are not known, they have to be estimated from the discrete observations stored at the vertices of the lattice. Intensity values for positions between the grid points are interpolated. Various research groups reported about accuracy and benefits of interpolation methods in the field of image registration. Most frequently, nearest neighbor, linear, and partial volume methods are applied [Maes98]. Higher order approaches are described, for instance, by Thévenaz et al. who introduced continuous models of the images based on cubic B-Splines. This image representation allows for subpixel accurate registrations [Thev98] and high quality multi-resolution schemes [Thev00]. Independent of the employed interpolation method, a regular sampling scheme of gridded data leads to serious numerical problems within the subpixel range. Maes [Maes98] and Pluim et al. [Plui00] were among the first who described these effects for the MI measure. They showed that, due to these numerical effects, local extrema are introduced into the objective function and lead to inaccurate registration results. The optimal transform is biased and no longer subpixel-accurate. The authors described these numerical problems as *interpolation artifacts* and proposed to resample the images in order to avoid grid-aligning positions. It is known from rendering techniques in computer graphics algorithms that uniform sampling produces highly visible noise when the regular structure in the images correlates with the sampling patterns. The disturbing effects are perceived as aliasing, especially at lines that are non-orthogonal to the image grid [Mitt87]. The aliasing in computer visualization is closely related to the numerical effects introduced by regular sampling in image registration. The problems can be resolved almost completely by non-uniform sampling, for example by adding a random jitter to the sampling positions. The jittering is motivated by the idea to reduce systematic effects that occur if the sampling – and also the interpolation coefficients – are based on regular patterns throughout the entire grid. It is schematically illustrated in Figure 4.2. The sampling positions (dots)

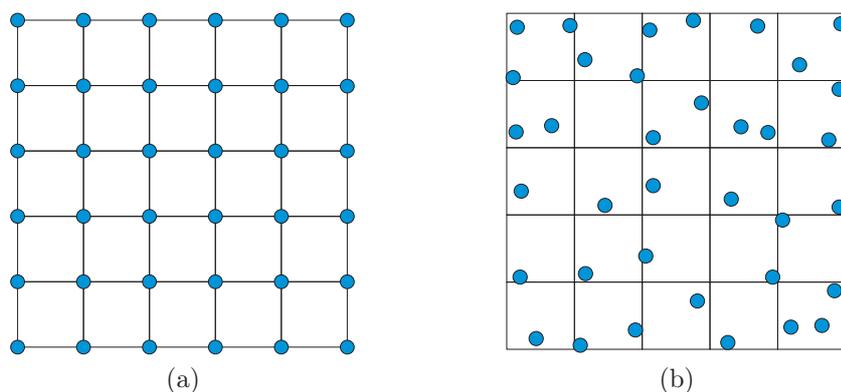


Figure 4.2: Illustration of sampling with and without jittering: (a) regular sampling pattern at the grid positions, (b) jittering with randomly distributed offsets that are used to destroy regular patterns.

in a regular scheme spatially coincide with the grid vertices (Figure 4.2a), whereas

with the usage of jittering, the sample positions are achieved by adding random offsets to the grid coordinates (Figure 4.2b). Tsao [Tsao 03] has thoroughly evaluated jittering of the sampling coordinates and smoothing of the discrete histograms as additional methods to cope with the numerical problems. The effects of jittering are demonstrated on the densely sampled objective function for the case of \mathcal{D}^{MI} in Figure 4.3. It shows the \mathcal{D}^{MI} values for translational changes of the transform param-

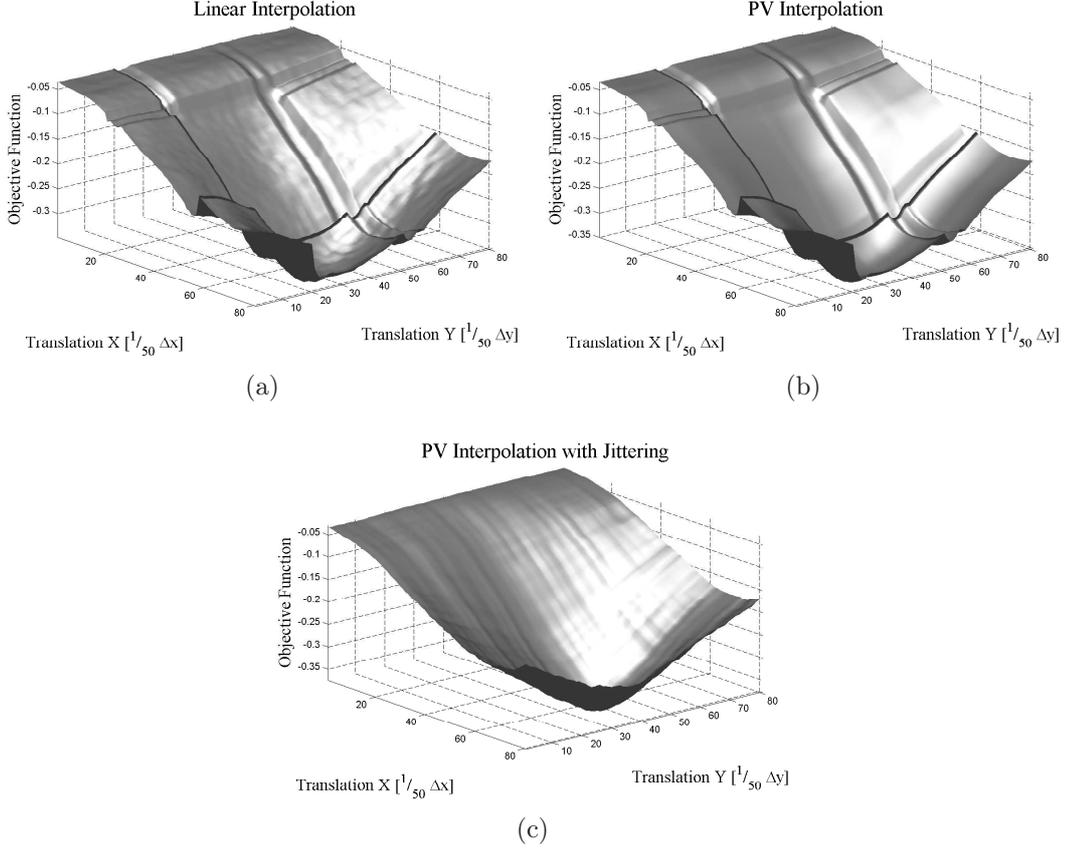


Figure 4.3: Plots of the densely evaluated \mathcal{D}^{MI} objective function values for translational parameter changes with a rate of $\frac{1}{50}$ of the voxel spacings Δx in x - and Δy in y -direction around the ground truth transform of a 3-D registration. We applied (a) linear interpolation, (b) PVI, and (c) a combination of both PVI and jittering.

eters in a 3-D registration. The measure is evaluated at densely sampled positions around the ground truth optimum with variations along two coordinates of the transform parameter space. The PVI within the histogram space reduces the noise in the measure, but not the additional local ridges due to numerical effects. The jittering, in combination with the PVI, delivered the best results, as it greatly reduced the aliasing errors – in fact, they are no longer visible in the plot for this example. The sampling scheme has a major influence on the smoothness of the objective function for accurate alignments within the subpixel range. From our experiments we conclude that the random offsets of the jittering do not necessarily have to be normally distributed around the grid knot coordinates, as proposed by Tsao. Instead, any random placement of the samples that avoids regular grid patterns can be used. For

example, Thévenaz et al. [Thev08] proposed quasi-random sampling based on Halton sequences. This sampling strategy has a low discrepancy regarding the resulting coordinates, which means that the sampled space is covered more regularly than *true* random sampling and avoids clustering. Even if the number of samples is low, as it may be the case for the highest resolutions in a multi-resolution scheme to reduce the costs for the measure computation, the additional variance in the objective function is smaller than in true random sampling. Therefore, the authors argued for using non-stochastic optimization techniques even if sparse quasi-random sampling is applied.

4.2 Density Estimation

Once the samples have been determined with respect to the current transform parameters, the joint PDF has to be estimated. Non-parametric estimation techniques like simple, discrete histogram methods, k -nearest-neighbors (k -NN), and kernel-based approaches [Bish97] are usually applied. A histogram is obtained by partitioning the domain of the random measures into a number of bins. The discrete PDF is then estimated by the fraction of samples that fall inside the bins. Although the number of bins acts as a smoothing parameter, the histogram suffers from discontinuities at the bin boundaries. The requirement of differential measures is very important for complex registration transforms with either many parameters, or entirely non-parametric approaches. At least first-order optimization techniques are required to achieve results in reasonable computation times. The differentiability of stochastic measures, hence, depends on the density estimator. On the one hand, standard histograms with few bins are usually less accurate compared to histograms that use a larger number of bins for a representation of the density. On the other hand, the estimate should be adequately smooth for various numbers of samples, which is not the case for histograms with a large number of bins. As initially mentioned, the bin size acts as a smoothing parameter, however, reducing the number of bins also leads to a loss of structural information within the PDF estimate, if the measured distribution has a high local variance. Although histograms are not differentiable at the bin boundaries, in practice, finite difference schemes between the bin center positions are used as an approximation.

The k -NN and kernel approaches are related to each other. In the k -NN method, one assumes that k random measures fall inside some region of the domain. The volume of this region depends on the chosen value of k and is determined by the data. The variable k acts as a smoothing parameter that is independent of the position. However, the estimated density has discontinuities between data points and the integral over all x -space diverges. Alternatively, one can keep the volume fixed and determine the number of random measures that fall into the region, which leads to a kernel-based estimator. In the one-dimensional case, given n random samples x_1, x_2, \dots, x_n , the Parzen-window PDF estimator [Parz62; Duda01] is:

$$p_{\lambda,n}(x) = \frac{1}{n} \sum_{i=1}^n K_{\lambda}(x - x_i) , \quad (4.1)$$

where K_λ is the kernel PDF with a width of λ , the smoothing parameter for this method. Equation (4.1) yields the sample mean of the kernel functions centered at the sampling positions. For i.i.d. random samples, the mean estimator $\bar{p}_{\lambda,n}(x)$ for an infinitely large number of samples is an unbiased estimate of the true PDF $p(x)$ if the following equation is fulfilled [Duda 01; Viol 95]:

$$\begin{aligned} \lim_{n \rightarrow \infty} \bar{p}_{\lambda,n}(x) &= \lim_{n \rightarrow \infty} \text{E}[p_{\lambda,n}(x)] \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n K_\lambda(x - x_i) p(x_i) \\ &= (K_\lambda \star p)(x) = p(x) . \end{aligned} \quad (4.2)$$

The convolution in (4.2), denoted by the operator “ \star ”, yields a blurred version of the true density unless the number of samples becomes infinitely large and the kernel width reaches zero. In this case, the kernel K_λ converges into a Dirac delta peak that is centered at each random measure. The true density can also be recovered if $p(x)$ has bounded frequency content and K_λ is a perfect low pass filter with an appropriate cut-off frequency. In practice, this means that, although the number of samples is finite, the PDF can be well approximated if $p(x)$ is a smooth function and K_λ a low pass filter [Well 96]. We assume in the following that these conditions hold true for our registration problem and apply this PDF estimator instead of k -NN or clustering techniques.

There are non-parametric PDF estimation methods particularly designed for multi-modal image registration. Maes et al. [Maes 97] applied a discrete representation of the PDF with equidistantly-spaced histograms, whereas Wells et al. [Well 96] suggested a continuous Parzen-window estimator to model the joint PDF. Although the latter approach does not require a binning scheme for the discrete PDF, it yields a relatively high computational complexity, i. e. $\mathcal{O}(nm)$ for n random samples and m evaluations of the estimator. Some effort has been made to reduce the computational load by transforming the problem into the frequency domain [Held 04], which reduces the costs to $\mathcal{O}(n + m)$. A compromise is suggested by Thévenaz et al. [Thev 00] who combine the smoothing properties of B-spline kernels with an equidistant binning scheme. Except for the first approach, a value for the kernel width parameter has to be determined for all of the described methods.

4.3 Automatic Data-Driven Parameter Selection

Some authors argue that it is rather simple to determine suitable values for the kernel size by empirical adjustment [Xu 08], while others state that data-driven parameter estimations are possible, but the entropy estimation is not very sensitive to empirical values [Well 96]. Unfortunately, the kernel width λ is dependent on the intensity distribution of the image, the discrete representation of the PDF, and the number of random measures. For a fixed, finite number n , the estimator $p_{\lambda,n}(x)$ is sensitive to λ . If on the one hand the chosen value for λ is too large, the estimated density is over-smoothed and the accuracy within the result is lost. On the other hand, too small values for λ insert artificial structure that may not be present in the real

data. A general registration algorithm for clinical applications should handle various imaging modalities and different field of views. Finding a good overall kernel width empirically for each modality combination is quite cumbersome and, as we will show in the results, a non-optimal value leads to mis-registrations and less accurate results.

Finding appropriate values for λ is further complicated by multi-level techniques that are employed for the optimization of a registration transform [Thev 98; Thev 00]. The result for a lower level of the resolution pyramid is applied as a starting point for the next one, which contributes to the robustness of the entire registration through enlarging the attraction range for the correct optimum. The ratio between the number of samples and the information contained in the images of the pyramid, however, varies between the levels, and in turn, influences the parameter λ .

After this summary of Parzen-windowing, we will concentrate on its discretization that is required for an implementation in practice. We gradually present the steps that are necessary to achieve an estimate that adapts to the structure of the true PDF, but still has the advantages of a discrete Parzen-window approach with a single kernel width. The basis for the kernel width estimation in the discrete case is a maximum likelihood formulation. The resulting log-likelihood objective function to determine the optimal kernel width parameter, with respect to the random samples, has a high computational complexity for the continuous Parzen-window approach. Therefore, we modify the PDF estimator to use discrete histograms, and present the derivatives of the related log-likelihood function. Finally, we propose an extension of the isotropic estimator by applying an adaptive binning scheme for the histograms. In this novel approach, we still use a single kernel width for the entire Parzen-window estimation and avoid spatially varying, anisotropic kernel widths. Therefore, we can represent local variations of the PDF structure, while – and in our opinion this is the major advantage – keeping the computational costs at a low level.

4.3.1 Leave-One-Out Cross-Validation

To determine appropriate values for the kernel width λ , the observations themselves are used. In this data-driven approach, care has to be taken to use disjoint sample sets for estimating the kernel width parameter and the objective test function. A common technique to resolve this problem of overfitting is cross-validation [Hast 01].

For a leave-one-out cross-validation strategy, let $p_{\lambda, n-1}^j$ be the estimator after deleting the j -th sample:

$$p_{\lambda, n-1}^j(x) = \frac{1}{n-1} \sum_{i=1, i \neq j}^n K_{\lambda}(x - x_i) . \quad (4.3)$$

This estimator is independent of the value x_j . The probability $p_{\lambda, n-1}^j(x_j)$ may, therefore, be used as a measure for how well the estimator fits to x_j with respect to the parameter λ . The resulting log-likelihood objective function is [Chow 83]:

$$\mathcal{L}(\lambda) = \sum_{j=1}^n \log p_{\lambda, n-1}^j(x_j) . \quad (4.4)$$

An optimal value for the kernel width yields a maximum log-likelihood:

$$\hat{\lambda} = \underset{\lambda}{\operatorname{argmax}} \mathcal{L}(\lambda) . \quad (4.5)$$

In practice, such a data-driven approach is well known to deliver reliable results. However, its efficiency drastically decreases with increasing sizes of the sample set. Typically, a leave-one-out cross-validation method to determine \mathcal{L} for a specific kernel size has a relatively high complexity of $\mathcal{O}(n(n-1))$. In the following section, we propose an approximation, which leads to a binned version of the discrete PDF with a major reduction in complexity.

4.3.2 Discretization of the Parzen-Window Estimator

The theory of Parzen-window estimation is based on a continuous representation of the data, and the estimated PDF is optimal with respect to the correct kernel width. As described above, this approach has a high computational complexity, which can be reduced by the usage of discrete histograms. Applying histograms, the non-parametric estimator resembles the behavior of a mixture model with as many components as bins. The n samples are stored in a histogram h_n with b bins ($b > 1$). Here, $h_n(x_i)$ denotes the fraction of samples that fall into the bin containing x_i . The bin width for an equidistantly-spaced histogram is given by $w = (x_{\max} - x_{\min})/(b-1)$, with x_{\max} being the maximal and x_{\min} the minimal image intensity value, respectively. A discretization like this results in an error because the correct location of the random measures is no longer continuous, but a discrete bin index, and the estimated PDF value is assumed to be constant for the entire bin. Let $\hat{p}_{\lambda,n}$ be the discrete PDF estimator that is based on a binning scheme using a histogram and defined only at discrete bin center positions c_j . Unlike its continuous counterpart, the discrete estimator results in a piecewise constant PDF estimate given by:

$$\hat{p}_{\lambda,n}(c_j) = \sum_{i=1}^b h_n(c_i) K_{\lambda}(c_j - c_i) \approx p_{\lambda,n}(c_j) , \quad (4.6)$$

where c_j is the intensity value corresponding to the center of the j -th bin. The normalized approximation error e in (4.6) can, for instance, be defined similar to the quantization noise, or noise power in the context of signal quantization (see also Chapter 4.3.3):

$$e = \frac{1}{(x_{\max} - x_{\min})^2} \sum_{j=1}^b e(c_j) , \quad \text{where} \quad (4.7)$$

$$e(c_j) = \int_{c_j - \frac{w}{2}}^{c_j + \frac{w}{2}} (c_j - x)^2 p_{\lambda,n}(x) dx .$$

The normalization with respect to the maximal range of the input values, given by the largest value x_{\max} and the minimal value x_{\min} , allows to specify an error threshold that is invariant to the input sample range. An illustration of this error for the j -th bin is given in Figure 4.4. If the approximation errors are neglected, equation (4.6)

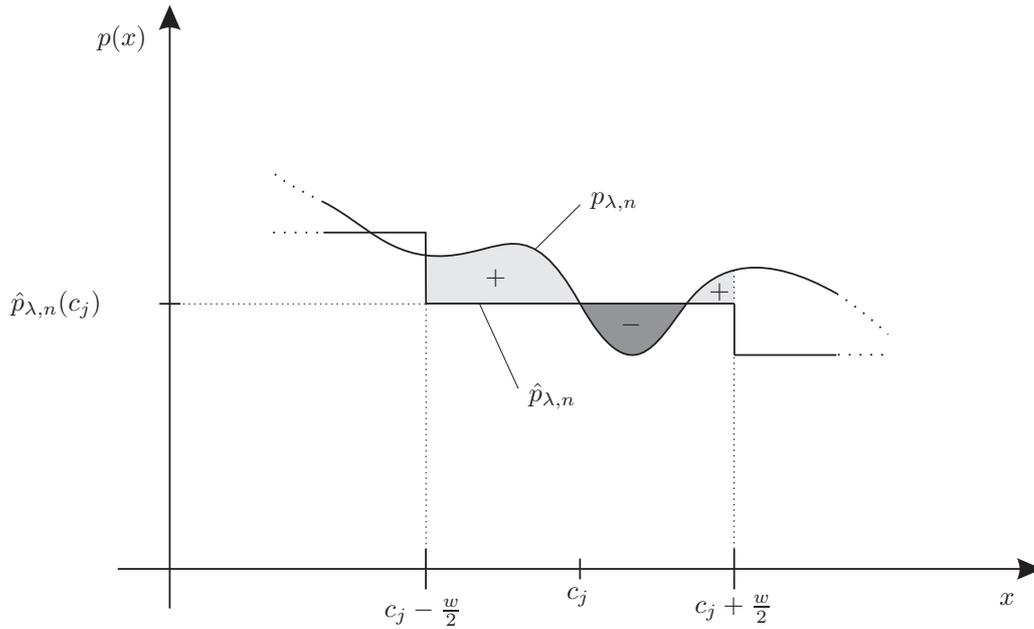


Figure 4.4: Illustration of the approximation error between the continuous PDF estimator $p_{\lambda,n}$ and the discrete version $\hat{p}_{\lambda,n}$ utilizing a binning scheme. The variation between the two graphs in the j -th bin defines the error $e(c_j)$. Graphically, this is denoted by the “+” and “-” marked areas representing under- and over-estimations, respectively.

yields a complexity of $\mathcal{O}(n \cdot b)$ with $b \ll n$, which allows for a substantially faster computation of the discrete PDF estimate compared to a continuous approach. Of course, the number of bins and the binning scheme used for the histogram affects the accuracy of the estimation.

Very common choices for the kernel PDF K_λ are the Gaussian g_λ with variance λ or cubic B-splines [Unse 93a; Unse 93b]. Using a cubic B-spline B yields the following Parzen-window kernel K_λ^B :

$$K_\lambda^B(x) = \frac{1}{\lambda} B\left(\frac{x}{\lambda}\right). \quad (4.8)$$

The B-spline function is usually defined recursively by the Cox-de Boor recursion formula, however, in the case of a third-degree B-spline, the recursion scheme yields the following kernel function:

$$K_\lambda^B(x) = \begin{cases} \frac{1}{\lambda} \left(\frac{4}{3} - 2\frac{|x|}{\lambda} + \frac{x^2}{\lambda^2} - \frac{|x|^3}{6\lambda^3} \right) & , \text{ if } \frac{|x|}{\lambda} \in [1, 2[\\ \frac{1}{\lambda} \left(\frac{2}{3} - \frac{x^2}{\lambda^2} + \frac{|x|^3}{2\lambda^3} \right) & , \text{ if } \frac{|x|}{\lambda} \in [0, 1[\\ 0 & , \text{ otherwise .} \end{cases} \quad (4.9)$$

K_λ^B can either be discretized by sampling the kernel values with a distance of w or by recursive filtering. In practice, however, the discretization of the Gaussian g_λ raises problems due to its infinite impulse response, which is also a reason why it violates

the partition of unity constraint [Thev 00]. A kernel function K satisfies the partition of unity, if the following equation holds true:

$$\sum_{x \in \mathbb{Z}} K(x + f) = 1, \quad \forall f \in \mathbb{R}. \quad (4.10)$$

The condition (4.10) can only be fulfilled by kernel functions with local support. Therefore, the application of a Gaussian kernel function leads to numerical errors in the estimates, and in multi-variate density estimation, as the marginal along one dimension is not totally invariant to the corresponding random variable. A kernel that fulfills the partition of unity constraint is, for instance, the cubic B-spline kernel K_λ^B . In practice, using a Gaussian may still be attractive, as it can be implemented with cascaded uniform-coefficient finite-impulse-response digital filters [Well 86] or with recursive filtering schemes that can also be efficiently evaluated [Deri 90]. The computational complexity of the recursive Gaussian filter is invariant to the variance of the Gaussian. In cases where the partition of unity constraint [Thev 00] is required for the density estimation, especially where the marginals of joint PDFs have to be computed, the Gaussian may be replaced.

The discrete Parzen-window estimator is written in terms of a convolution of the histogram with the discrete representation of the kernel function:

$$\hat{p}_{\lambda,n}(c_j) = \sum_{i=1}^b h_n(c_i) K_\lambda(c_j - c_i) = (h_n \star K_\lambda)(c_j). \quad (4.11)$$

Nonetheless, it is still required to optimize the log-likelihood (4.5) to achieve a value for the kernel width that yields the estimator with the best resemblance to the data. The problem can be solved using an iterative, nonlinear optimization scheme, e. g. Newton's method. Therefore, the derivatives of \mathcal{L} with respect to λ are required:

$$\begin{aligned} \frac{\partial}{\partial \lambda} \mathcal{L}(\lambda) &= \sum_{j=1}^n \frac{1}{\hat{p}_{\lambda,n-1}^j(x_j)} \frac{\partial}{\partial \lambda} \hat{p}_{\lambda,n-1}^j(x_j) \\ \frac{\partial^2}{\partial^2 \lambda} \mathcal{L}(\lambda) &= \sum_{j=1}^n -\frac{1}{(\hat{p}_{\lambda,n-1}^j(x_j))^2} \left(\frac{\partial}{\partial \lambda} \hat{p}_{\lambda,n-1}^j(x_j) \right)^2 \\ &\quad + \frac{1}{\hat{p}_{\lambda,n-1}^j(x_j)} \frac{\partial^2}{\partial^2 \lambda} \hat{p}_{\lambda,n-1}^j(x_j). \end{aligned} \quad (4.12)$$

Similar to the leave-one-out notation for the PDF estimator introduced above, h_{n-1}^j refers to the histogram without the j -th sample. For Gaussian kernel functions, the partial derivatives of the Parzen-window estimator yield:

$$\begin{aligned} \frac{\partial}{\partial \lambda} \hat{p}_{\lambda,n-1}^j(x_j) &= \sum_{i=1}^b \frac{h_{n-1}^j(c_i)}{2\lambda} g_\lambda(x_j - c_i) \cdot \left[\frac{(x_j - c_i)^2}{\lambda} - 2 \right] \\ \frac{\partial^2}{\partial^2 \lambda} \hat{p}_{\lambda,n-1}^j(x_j) &= \sum_{i=1}^b \frac{h_{n-1}^j(c_i)}{4\lambda^2} g_\lambda(x_j - c_i) \cdot \left[\frac{(x_j - c_i)^4}{\lambda^2} - \frac{8(x_j - c_i)^2}{\lambda} + 8 \right]. \end{aligned} \quad (4.13)$$

The partial derivatives can be formulated accordingly, if a Parzen-window estimator is based on cubic B-splines:

$$\begin{aligned}\frac{\partial}{\partial \lambda} \hat{p}_{\lambda, n-1}^j(x_j) &= \sum_{i=1}^b h_{n-1}^j(c_i) \frac{\partial}{\partial \lambda} K_{\lambda}^{\text{B}}(x_j - c_i) \\ \frac{\partial^2}{\partial^2 \lambda} \hat{p}_{\lambda, n-1}^j(x_j) &= \sum_{i=1}^b h_{n-1}^j(c_i) \frac{\partial^2}{\partial^2 \lambda} K_{\lambda}^{\text{B}}(x_j - c_i).\end{aligned}\tag{4.14}$$

Similar to the kernel function (4.9), we provide the derivatives of K_{λ}^{B} for three cases.

Case 1: $\frac{|x|}{\lambda} \in [1, 2[$

$$\begin{aligned}K_{\lambda}^{\text{B}}(x) &= \frac{1}{\lambda} \left(\frac{4}{3} - 2\frac{|x|}{\lambda} + \frac{x^2}{\lambda^2} - \frac{|x|^3}{6\lambda^3} \right) \\ \frac{\partial}{\partial \lambda} K_{\lambda}^{\text{B}}(x) &= \frac{1}{\lambda^2} \left[\frac{2|x|}{\lambda} - \frac{2x^2}{\lambda^2} + \frac{|x|^3}{2\lambda^3} - \lambda K_{\lambda}^{\text{B}}(x) \right] \\ \frac{\partial^2}{\partial^2 \lambda} K_{\lambda}^{\text{B}}(x) &= \frac{2}{\lambda^3} \left[-\frac{2|x|}{\lambda} + \frac{3x^2}{\lambda^2} - \frac{|x|^3}{\lambda^3} - \lambda^2 \frac{\partial}{\partial \lambda} K_{\lambda}^{\text{B}}(x) \right].\end{aligned}\tag{4.15}$$

Case 2: $\frac{|x|}{\lambda} \in [0, 1[$

$$\begin{aligned}K_{\lambda}^{\text{B}}(x) &= \frac{1}{\lambda} \left(\frac{2}{3} - \frac{x^2}{\lambda^2} + \frac{|x|^3}{2\lambda^3} \right) \\ \frac{\partial}{\partial \lambda} K_{\lambda}^{\text{B}}(x) &= \frac{1}{\lambda^2} \left[\frac{2x^2}{\lambda^2} - \frac{3|x|^3}{2\lambda^3} - \lambda K_{\lambda}^{\text{B}}(x) \right] \\ \frac{\partial^2}{\partial^2 \lambda} K_{\lambda}^{\text{B}}(x) &= \frac{2}{\lambda^3} \left[\frac{3|x|^3}{\lambda^3} - \frac{3x^2}{\lambda^2} - \lambda^2 \frac{\partial}{\partial \lambda} K_{\lambda}^{\text{B}}(x) \right].\end{aligned}\tag{4.16}$$

Case 3: $\frac{|x|}{\lambda} \notin [0, 2[$

$$\begin{aligned}K_{\lambda}^{\text{B}}(x) &= 0 \\ \frac{\partial}{\partial \lambda} K_{\lambda}^{\text{B}}(x) &= 0 \\ \frac{\partial^2}{\partial^2 \lambda} K_{\lambda}^{\text{B}}(x) &= 0.\end{aligned}\tag{4.17}$$

Equations (4.13) and (4.14) are again convolutions with the partial derivatives for the kernel width of the Parzen-window kernel. Due to its local support, the spline kernel is very suitable for an implementation of the log-likelihood optimization using a discrete convolution operator and the kernel functions shown in Figure 4.5. The multi-variate kernel width estimation is realized analogue to the 1-D case, because both the Gaussian kernel PDF with diagonal covariance matrix and the B-spline kernels are separable. The 1-D convolution kernels have to be subsequently applied to each dimension of the histogram. Figure 4.6 shows example curves that can typically be observed for the log-likelihood optimization of the kernel width. In this example, the width of a cubic B-spline kernel has been adapted to 100 sample values drawn

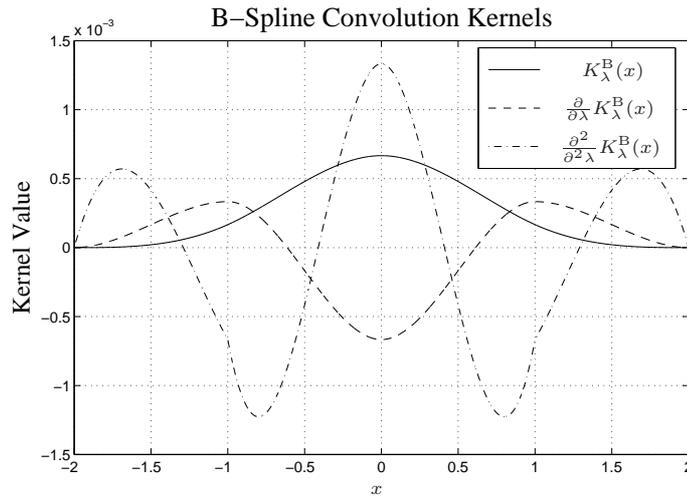


Figure 4.5: Convolution kernels for the kernel width estimation using a cubic B-spline window function K_λ^B printed for a kernel width $\lambda = 1$. The kernel function, the first derivative, and the second derivative are plotted within the locally supported region with respect to λ .

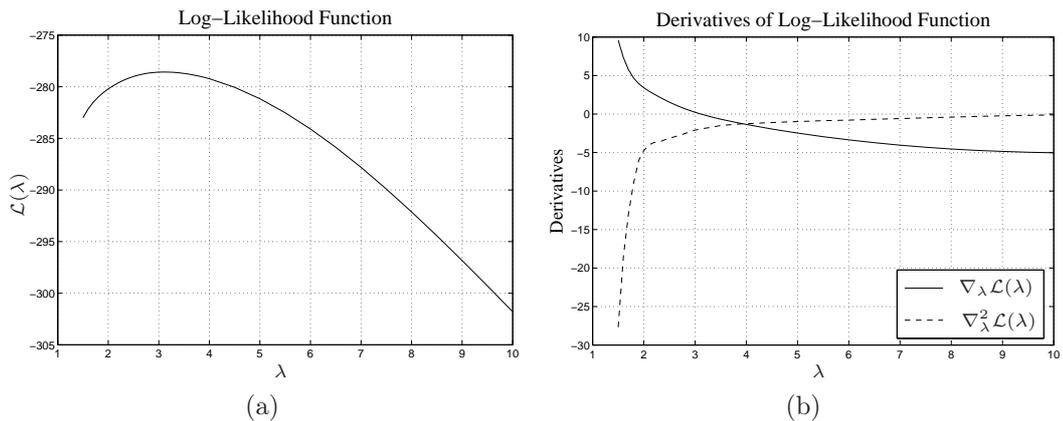


Figure 4.6: (a) Parzen-window log-likelihood function for the width of a cubic B-spline kernel in an estimation for 100 samples drawn from a normal distribution with $\lambda = 4$, and (b) the corresponding first and second order derivatives with respect to λ .

from a Gaussian distribution with mean zero and a variance of 4. The objective function shown in Figure 4.6a has a convex shape, which provides a good basis for the nonlinear optimization. In some cases, we observed a change in the sign of the Hessian for larger values of λ , however, this can be detected quite easily and the update is computed by stepping into the direction of the gradient instead of applying Newton's rule. The convergence for initial starting position may, in the worst case, be linear instead of superlinear.

Note that, from a numerical point of view, the domain of the intensity random variable is rather important. In order to achieve numerically stable results, the density transform theorem [Brem 87] can be applied. According to this theorem, the kernel

width parameter for the discrete estimator is invariant to constant offsets applied to all random values but not to linear scalings. Let $s \in \mathbb{R}^+$ be a scaling factor for the intensities and $t \in \mathbb{R}$ an offset. An affine transform of the random measure x is given by:

$$y = f(x) = sx + t . \quad (4.18)$$

Applying, for instance, the Gaussian as density function, the kernel PDF $K_{\lambda'}$ of the transformed samples can be expressed using the determinant of the Jacobian J_f of (4.18):

$$\begin{aligned} K_{\lambda'}(y) &= K_{\lambda}(f^{-1}(y)) \frac{1}{\det J_f(f^{-1}(y))} \\ &= \frac{1}{\sqrt{2\pi s^2 \lambda}} e^{-\frac{1}{2} \frac{(y-t-\mu)^2}{s^2 \lambda}} . \end{aligned} \quad (4.19)$$

The kernel width λ' for the transformed domain is therefore:

$$\lambda' = s^2 \lambda . \quad (4.20)$$

This property is very convenient as both the convolution (4.11) and the optimization of the kernel width parameter (4.5) can be performed in an affine-transformed domain. The sampled histograms may, therefore, be directly convolved with a Gaussian of adapted kernel size λ' without having to account for the bin spacings – a technique which has also been applied similarly by Hermosillo et al. [Herm02b].

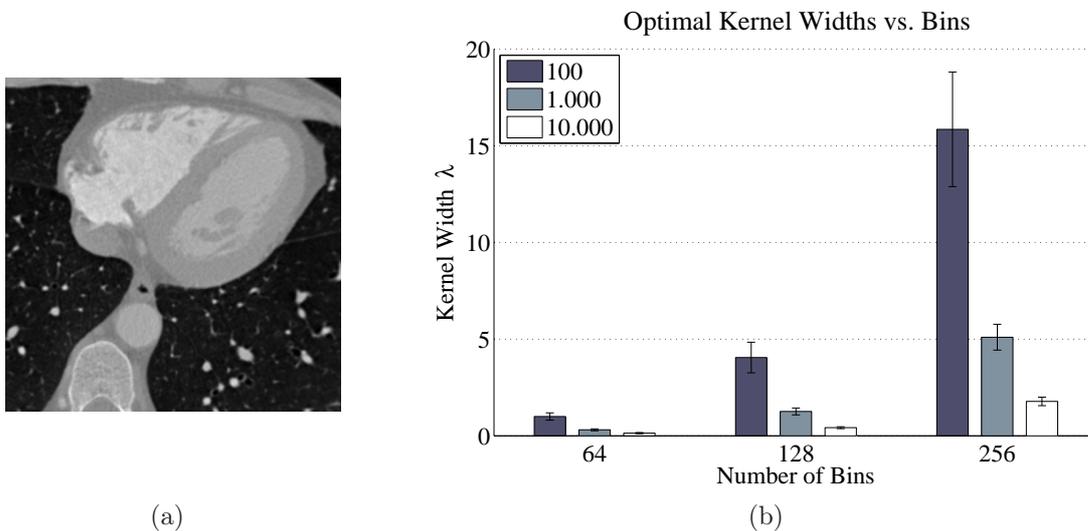


Figure 4.7: Optimal kernel width parameter values for various numbers of samples and bins. For (a) a slice taken from a 3-D AX scan, the (b) vertical bars show the corresponding optimal values for the kernel widths and their standard deviations due to random sampling with different sampling numbers (100, 1.000 and 10.000). The results are averaged from 100 subsequent runs with different random seeds.

4.3.3 Histogram Layout

In data-driven approaches for estimating the optimal kernel width, one can observe that the result is directly related to the uncertainty within the data, i. e. the number of samples provided for an intensity value. Due to the discrete nature of histograms, this uncertainty is reflected by a varying smoothness or degenerations. Estimators using constant kernel widths cannot distinguish between regions of high and low certainty within one histogram. Therefore, several authors suggest making this parameter spatially variant. In medical images, the PDF of the intensity values is often degraded, as the background yields a strong and dominating peak. A convolution with a low pass filter, as proposed in Section 4.3.2, smears this peak over the neighboring bins, which may overshadow information on low intensity image content. The ability to adapt the estimator to PDFs of varying smoothness is, hence, an important feature for image registration. An additional drawback of estimating the discrete PDFs using equidistantly-spaced histograms is that intensities of a single tissue class may end up in different bins. Intensities measured in medical imaging rarely follow a uniform distribution because the probabilities for all tissue classes would have to be equal, which is obviously not the case. The dimensions of the organs inside the human body vary between individuals, and the extent of the background region depends on the field of view. Adapting the bin sizes to the structure of the PDF, therefore, leads to a better representation with respect to a smaller quantization error. This approach has been proposed previously by Knops et al. [Knop06] who applied intensity clustering. In the following, we use the term *adaptive, anisotropic kernel width* in the context of PDF estimation to express the property that the estimator is spatially adapted to the structure of the underlying PDF using varying kernel sizes.

Given kernel widths $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)$, an adaptive Parzen-window estimator reads:

$$p_{\boldsymbol{\lambda},n}(x) = \frac{1}{n} \sum_{i=1}^n K_{\lambda_i}(x - x_i) . \quad (4.21)$$

This estimator has recently been applied in human motion tracking for the modeling of position and orientation priors [Brox07]. The adaptive estimator focuses better on the structure of the PDF by allowing smaller kernel sizes in areas with many training samples. Sparsely sampled areas of the PDF can still be approximated by larger kernel widths. The authors suggested a linear combination of local covariance matrices and a scaled identity matrix to determine the i -th kernel width. Katkovnik et al. [Katk02] computed confidence intervals of the random variable domain using a pilot density from an estimation with a constant kernel size and knowledge about the sample variance. The intersections of these intervals determine the adaptive kernel sizes. The method results in small widths in regions with high variance compared to areas with low variance, where larger values of the kernel size tend to decrease the mean squared error (MSE) between the estimate and the true PDF. Their approach has some similarities to an adaptive k -NN method. The authors show that an estimator using adaptive kernel widths produces estimates with less variance in the MSE.

Quasi-Adaptive Kernel Widths

A disadvantage of adaptive, anisotropic kernel widths, if applied in image registration, is the increased computational complexity for both the estimator and the formulation of its derivatives. The efficient evaluation scheme that was presented in Section 4.3.2 cannot be applied to estimators with varying kernel sizes. We propose a novel combination of an adaptive binning scheme with isotropic kernel sizes for the PDF estimation as a trade-off. Instead of determining different kernel widths for an equidistantly-spaced histogram, we propose to approximate the PDF using a histogram with varying bin sizes. The corresponding bin centroids define a quantization characteristic, which can be used to map the input intensities to requantized output values. These, in turn, can be represented with an equidistantly-spaced histogram. A density estimation on this requantized intensity space does not have to account for different bin widths of the histogram and the proposed estimation scheme of Section 4.3.2 can be applied. The nonlinear mapping may be computed as a preprocessing step in the beginning, which means that it has to be computed only once for each image. Of course, the registration of the images containing these processed values requires a distance measure that is robust to this type of intensity transform, a property that is fulfilled by stochastic measures.

Given the number of bins, we want to distribute the location and the sizes of the bins over the domain of the random variable, such that the discrete estimator accounts for local variances of the PDF. Problems like this are known from quantization theory, where a quantization characteristic is optimized for a signal with respect to a minimal quantization noise [Niem 83]. We applied an approach that was introduced by Lloyd [Lloy 82] and Max [Max 60]. It minimizes the noise power \mathcal{N} for a specific number of bins by an iterative refinement of the bin center locations. The spatial region of the i -th bin within the domain of the random variable is defined by the interval $[l_{i-1}; l_i]$ with the centroid c_i . The noise power of the requantization with respect to the signal PDF $p(x)$ is:

$$\mathcal{N} = \sum_{i=1}^b \int_{l_{i-1}}^{l_i} (c_i - x)^2 p(x) dx, \quad (4.22)$$

which we also used in (4.7) to measure the approximation error for the discrete PDF estimator. According to Lloyd [Lloy 82], a fixed point iteration scheme can be applied to optimize an adaptive layout of the bins that minimizes (4.22). The updates in each iteration are:

$$\begin{aligned} c_i &= \frac{\int_{l_{i-1}}^{l_i} xp(x) dx}{\int_{l_{i-1}}^{l_i} p(x) dx} \\ l_i &= \frac{c_i + c_{i+1}}{2}, i \in \{1, 2, \dots, b\}. \end{aligned} \quad (4.23)$$

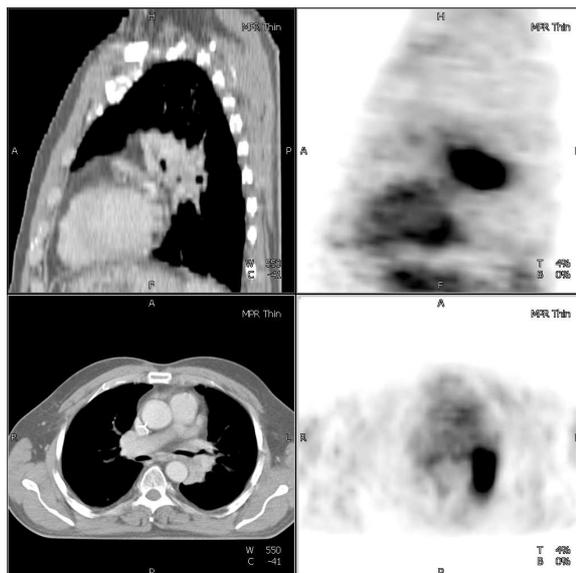
If applied to the quantization of images, l_0 and l_b can simply be chosen as the minimal and maximal image intensity value, respectively. Again, p denotes the unknown true PDF that, of course, is not known. For an implementation, we propose to use a discrete Parzen-window estimator instead, but with large values for both the number of bins and the samples. During the iterative solution of the registration

problem, such an estimator is prohibitive, as the computational benefit from the discrete histogram compared to the continuous representation is lost. However, we only need to determine the nonlinear distribution of the bins once for each input image. In practice, medical images are currently stored with 16 bit accuracy or less, which, for instance, allows for a discrete representation of the estimator for p with a full resolution of 2^{16} bins. This computationally rather expensive estimator is then plugged into (4.22) to determine the nonlinear intensity mapping in a preprocessing step of the algorithm. After the images have been requantized based on the nonlinear intensity mapping, the estimators used within the main iteration of the registration algorithm make use of the fast discretization presented in Section 4.3.2. The whole approach is described for 1-D PDFs, however, it can be easily extended also to 2-D joint densities.

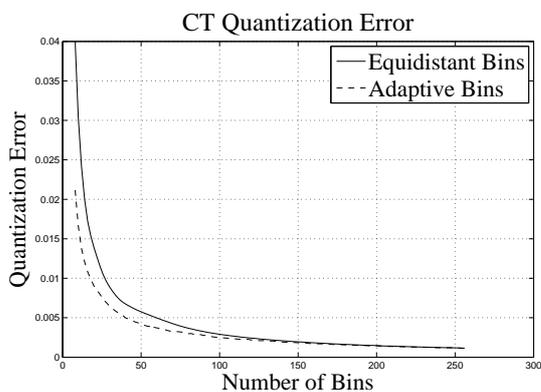
Number of Histogram Bins Selection

The histogram binning introduces quantization errors and, therefore, accuracy is lost compared to the optimal density. The question is how many bins are a good compromise between efficiency and accuracy. An empirical result states that the optimal number of bins for a joint histogram used in NMI is 64 for each dimension [Knop06]. From a theoretical point of view, there is no explanation why 64 bins should be the best possible choice. As the final quantization error is also dependent, for instance, on the image content and the number of samples, the general nature of this result is questionable. Therefore, we propose to estimate this parameter for each pair of input images individually.

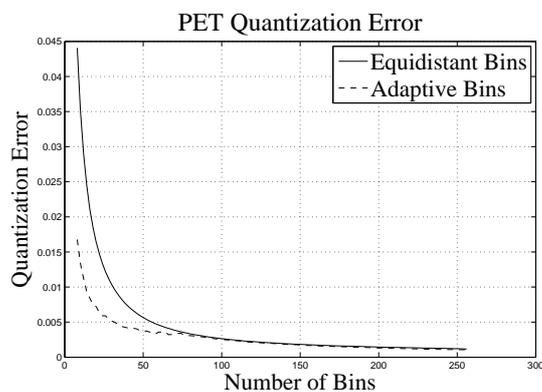
The quantization error criterion (4.7) and (4.22), which is used for the adaptive histogram binning, can also be utilized to automatically determine a suitable number of bins. The continuous density estimate may, for this purpose, be approximated by a discrete histogram using the aforementioned methods. It can be generated with the full intensity resolution, i. e. 2^{16} bins for a 16-bit quantized image, and estimated from all intensity values stored in the image grid. Such a huge number of bins is undesirable for a computationally efficient distance measure computation. Thus, we suggest to define a lower threshold for the discretization error e between this high-resolution and the target estimate $\hat{p}_{\lambda,n}$. The number of bins can now be computed in an iterative procedure starting from an initial, minimal value, e.g. 16 bins. Equation (4.7) is solved for each discrete estimator built with this number of histogram bins. The iteration stops if e falls below the defined threshold. As the integrals over the discrete density estimates are normalized to one, the threshold is invariant of the image content and a universally applicable value can be chosen. In the performed experiments, a value of 0.005 has been used successfully. Figure 4.8 shows an example for the computation of the number of bins needed for the reference (CT) and template (PET) image. The quantization error drops heavily in the beginning of the iteration and converges slowly towards zero as the number of bins is increased. The adaptive binning proposed in 4.3.3 yields a smaller number of bins compared to equidistant bin spacings at the same error level.



(a)



(b)



(c)

Figure 4.8: Automatic selection of the number of bins for (a) a reference CT and a template PET image. The quantization error curves (b) for the CT and (c) the PET have been computed with respect to a high resolution density estimate. The plots show that the adaptive histogram layout results in a smaller quantization error compared to uniform binning. A threshold level of 0.005 in this case yields values between 40 and 50 for the number of bins.

4.3.4 Coincidence Weighting

In the previous sections, we have established an efficient PDF estimator by a discrete Parzen-window approach. It can be directly applied in the formulations of the stochastic similarity measures, however, there is still an important aspect of medical, tomographic images that often leads to inaccurate registration results. More precisely, problems arise due to structured noise contained in the background regions of the reconstructed images. Tomographic images are the result of discrete, modality-specific reconstruction methods that are based on physical measurements. In practice, these measurements are affected by detector noise and physical effects, for example beam hardening or scattering in CT. This noise is propagated through the reconstruction

chain. Structured noise is, unfortunately, not only dependent on the reconstructed object itself, but also partly on the acquisition geometry. Although the noise may occur in all regions of an image, the major problems are primarily caused by clusters of artificial noise in the background, which consists of a homogeneous region in an ideal case. Examples for this type of background noise are shown in Figure 4.9. If two

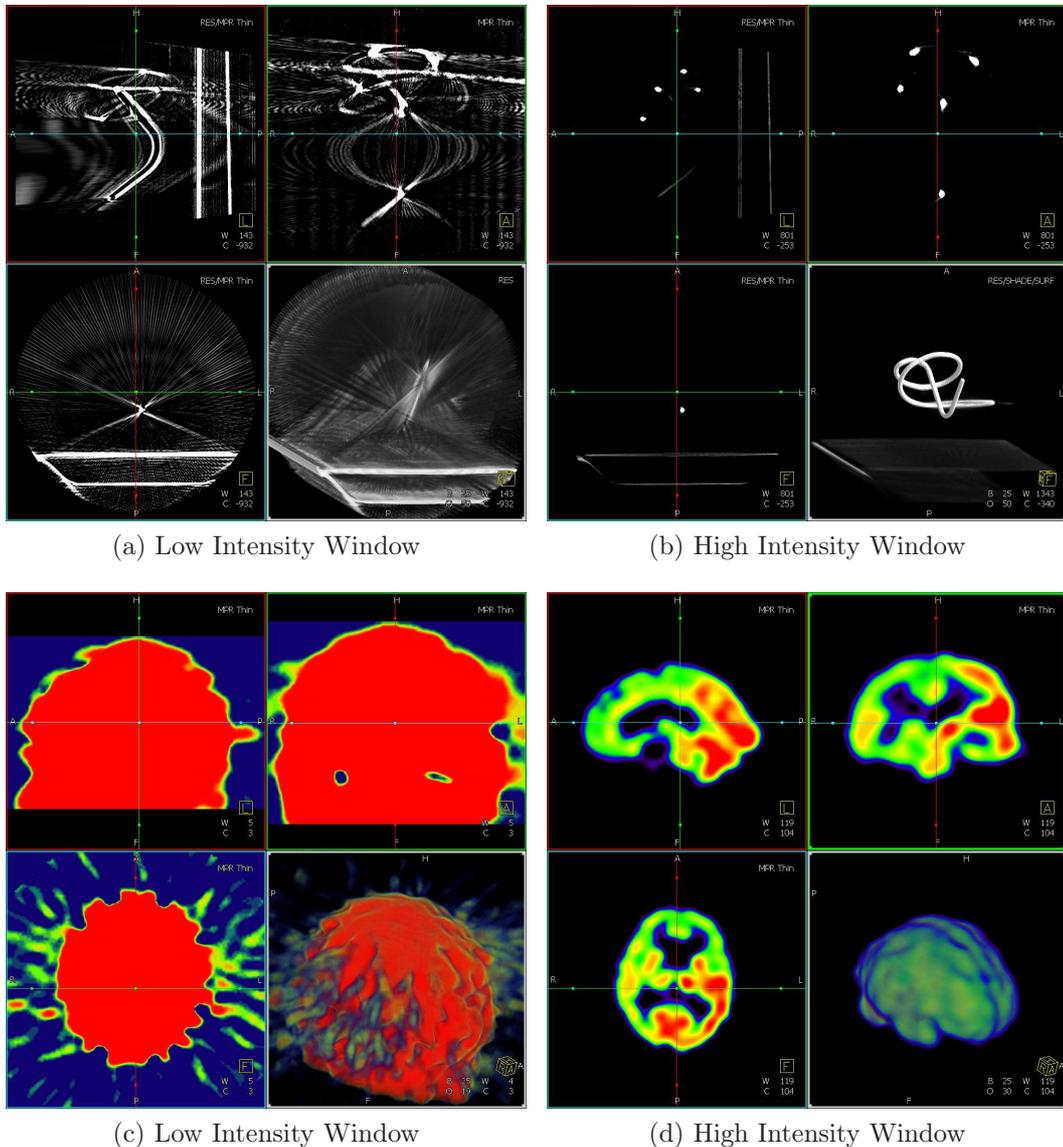


Figure 4.9: Structured noise visible in the low intensity range of the background region of the images. The images relate to (a), (b) two different transfer function settings (called *window*, with center and width defining the visible intensity range) for an AX image of a vessel phantom, and (c), (d) similar visualizations of a SPECT image of a human brain.

images differ only in the structured noise in the background, the resulting transformation between the objects becomes less accurate. These artificial structures impair the similarity measure, because the registration algorithm tends to optimize their alignment as well. For the case of medical images, one typically assumes that inten-

sities belonging to the background are at the lower end of the intensity range. Some authors have tried to eliminate this problem by using intensity thresholds within the joint histogram for a standard and coincidence thresholding [Rohl 99; Rohl 00], or by masking the background region of the images [Thev 98]. The standard thresholding is similar to applying a mask on the background. Only those joint intensities are taken into account with both the reference and the template intensity value being above the corresponding background threshold. Thus, only the overlap between the object parts is incorporated into the distance measure and combinations with the background regions are disregarded. The robustness of the registration, therefore, decreases, as large initial misalignments with relatively small overlaps between the objects cannot be recovered.

To account for these blind spots in the object-background relations, Rohlfing and Beier [Rohl 99] have proposed a thresholding, which only affects alignments between background parts of the images. Elements in the joint histogram are discarded, if they correspond to intensities below both the reference and the template background threshold. They called this technique *coincidence thresholding* and reported a reduction of the maximum registration error without loss of accuracy. For an implementation of this technique, the threshold values for the reference and the template background intensities have to be specified. The authors provided experimentally determined values. However, these thresholds are very modality- and image-specific. Even if the intensities are related to a specific type of tissue (e.g. HU in CT) the images may still differ in content and contrast. For example, cardiac CT images have a different intensity distribution than whole-body scans. Depending on the field of view, the background may not even be included within the image. For modalities that lack an intensity normalization, a fixed threshold could be too restrictive in some cases, which yields results that are computed without valuable information about the objects.

Instead, we propose an alternative to coincidence thresholding that does not require fixed threshold values. The disadvantage of manually determined settings is that they do not apply to all the data that is used as input to the algorithm. The specification of a background threshold can, again, be solved by data-driven parameter estimation techniques. In addition, we do not want to completely discard the background-background (B-B) region in the joint PDF. We rely on a robust, automatic detection of the background value, which is afterwards used to distinguish between object-object (O-O) and B-B alignment regions within the joint PDF. Thus, we are able to determine a weighting factor between the two regions to ensure that the B-B does not dominate the similarity measure, but still contributes to the overall measure value. This strategy is less restrictive than the coincidence thresholding. We call it *coincidence weighting* in the following.

Thévenaz et al. [Thev 98] presented a robust technique to distinguish between the object and background region within an image. They used the aforementioned Max-Lloyd quantization algorithm on a low pass filtered version of the image. Together with the filtering, the algorithm computes the bin widths for a discrete histogram containing two bins. The boundary between the two bins is assumed to separate intensities in the background from object values. The authors used the resulting threshold to determine a mask for the background region within PET images to

get rid of structured noise, but at the expense of losing the ability to detect false object-background (O-B) alignments. Especially in the early iteration steps of the registration, it is likely that the initial alignment contains large O-B parts. As the background is masked within the images, these pixels are discarded within the similarity measure computation. In other words, the similarity measure is “blind” to misalignments where object content of the template image is transformed to spatial locations of the background in the reference image, and vice versa.

The determined threshold may, for instance, be used as a substitute for the manually defined value in the coincidence thresholding scheme, but we relax the strict thresholding constraint and propose a weighting instead. In some acquisitions, the field of view is placed totally inside the boundaries of the patient’s body. If coincidence thresholding is applied to the joint density with the resulting Max-Lloyd threshold, the algorithm loses information about low-intensity structures between the objects. We, therefore, propose a trade-off, which does not clamp the coincidence region in the joint density to zero, but rather applies a weighting to the B-B probabilities. The estimated probabilities \hat{p}_{0-0} for O-O and \hat{p}_{B-B} for B-B intensity combinations are calculated from the joint histogram $\mathbf{h}_{\mathcal{R},\mathcal{T}_\Phi}$:

$$\begin{aligned}\hat{p}_{0-0} &= \sum_{i=\psi_{\mathcal{R}}+1}^{b_{\mathcal{R}}} \sum_{j=\psi_{\mathcal{T}}+1}^{b_{\mathcal{T}}} \mathbf{h}_{\mathcal{R},\mathcal{T}_\Phi}(c_i, c_j) \\ \hat{p}_{B-B} &= \sum_{i=1}^{\psi_{\mathcal{R}}} \sum_{j=1}^{\psi_{\mathcal{T}}} \mathbf{h}_{\mathcal{R},\mathcal{T}_\Phi}(c_i, c_j).\end{aligned}\tag{4.24}$$

Here, $\psi_{\mathcal{R}}$ and $\psi_{\mathcal{T}}$ denote the bin indices that contain the corresponding Max-Lloyd threshold values for the reference and the template images, respectively. The number of bins used for the reference image is given by $b_{\mathcal{R}}$, the corresponding number of bins for the template image by $b_{\mathcal{T}}$. The following equation provides the formula for the coincidence weighting factor w_h , which is then applied to all joint histogram entries with indices in the range of $[1, 1]$ to $[\psi_{\mathcal{R}}, \psi_{\mathcal{T}}]$:

$$w_h = \begin{cases} 1 & , \text{ if } \hat{p}_{B-B} \leq \hat{p}_{0-0} \\ \frac{\hat{p}_{0-0}}{\hat{p}_{B-B}} & , \text{ otherwise .} \end{cases}\tag{4.25}$$

As mentioned above, the weighting with w_h ensures that the background combinations do not dominate the joint PDF estimate. As we modify the relative frequency of the samples in this weighting step, the histogram has to be normalized to sum up to one before the discrete Parzen-window estimation 4.11 is applied (see Section 4.3.2).

4.4 Implementational Aspects

So far, we have proposed a novel, discrete PDF estimator that combines the advantages of anisotropic binning with the efficiency of discrete, isotropic histograms. The presented approach contains several data-driven estimation steps: the computation of a binning layout that is optimal with respect to a minimal quantization error, the automatic selection for the number of bins, and the calculation of a threshold value

that separates the background intensities from the objects within each image. In addition, the coincidence weighting method utilizes these thresholds to decrease the influence of the B-B contributions in the joint PDF, if necessary.

We suggest incorporating the proposed methods at specific points of a registration algorithm. The computational overhead is minimal compared to fixed settings, as some of the data-driven schemes are performed only once in a preprocessing step before the actual multi-level registration is started. Primarily, the preprocessing is necessary in order to achieve a nonlinear requantization for the quasi-adaptive kernel estimation, and to compute the background intensity thresholds for the coincidence weighting. The requantization can be efficiently combined with the creation of a multi-resolution image pyramid used in a multi-level nonlinear optimization afterwards. The main registration loop usually implements an iterative numerical scheme over several resolution levels to optimize the transformation between the images. We assume that the optimal kernel widths for the Parzen-window estimation can be used for an entire nonlinear optimization on a single level. Therefore, we also introduce a preprocessing step for a single level where the data-driven kernel width computation is performed. During the nonlinear optimization, the resulting kernel width values are used for the actual PDF estimation with the efficient, discrete Parzen-windowing. The remaining steps to complete the registration algorithm may vary between different applications and have been omitted for the sake of clarity and generality. The following listing summarizes the required implementation steps along with references to the corresponding sections:

Preprocessing

- compute the background intensity thresholds $\psi_{\mathcal{R}}, \psi_{\mathcal{T}}$ used for the coincidence weighting (Section 4.3.4)
- compute the number of bins $b_{\mathcal{R}}, b_{\mathcal{T}}$, for the reference and the template image as a good trade-off between accuracy and quantization error (Section 4.3.3)
- calculate the adaptive binning layout of the histograms given $b_{\mathcal{R}}, b_{\mathcal{T}}$, which then determines the nonlinear intensity requantization for each image (Section 4.3.3)
- apply the nonlinear requantization of the image intensities and threshold values to apply the proposed Parzen-window estimator (Section 4.3.2)
- create multiple resolutions of the requantized images for a multi-level optimization

Main Registration Loop

1. Multi-Level Preprocessing
 - estimate optimal kernel width values for the current resolution of input images and the given (or also automatically determined) number of samples by optimizing the log-likelihood function (4.4)
2. Nonlinear Optimization until Convergence

- determine the joint histogram by jittered sampling to avoid numerical grid artifacts (Section 4.1)
- apply the proposed coincidence weighting with the thresholds $\psi_{\mathcal{R}}, \psi_{\mathcal{T}}$ to reduce the influence of background noise on the similarity measure (Section 4.3.4)
- estimate the joint PDF using the efficient discretization scheme for the Parzen-window approach with the optimal kernel widths for this level (Section 4.3.2)

4.5 Results

The first step in the practical implementation is comprised of the automatic computation of the background intensity thresholds. These values are used for the coincidence weighting, and their computation has to be robust and invariant to the medical modality. Figure 4.10 shows results of the background threshold detection for the images provided in Figure 4.8a, which shows a CT and PET image. Figure 4.11 provides

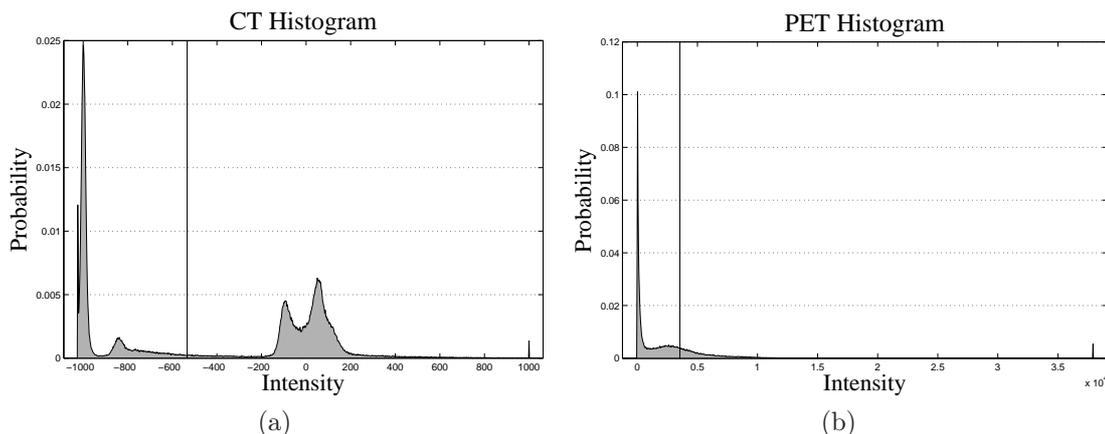


Figure 4.10: Intensity histograms for the images in Figure 4.8a. The (a) histogram for the CT is shown with the computed threshold marked as a vertical line at around -550 HU. The (b) plot for the PET is provided, respectively.

examples for the joint density estimation of the same images at an initial starting position. Although the image pair shows a large amount of background information, the proposed weighting scheme leads to similar joint density estimates, even when the number of samples is reduced. For this example, the percentage of samples taken from the entire number of voxels within the overlap domain ranged between 10% and 0.01%. To compute the results, the number of histogram bins, the adaptive binning, the coincidence weighting, and the kernel width optimization have been applied in the Parzen-window estimation. The data-driven estimation of the parameters accounts for the uncertainty within the data, which is increasing with each reduction of the sampling rate. The kernel width is always chosen optimally with respect to the data and adapts to the sampling rate. The PDF estimates for very low sampling

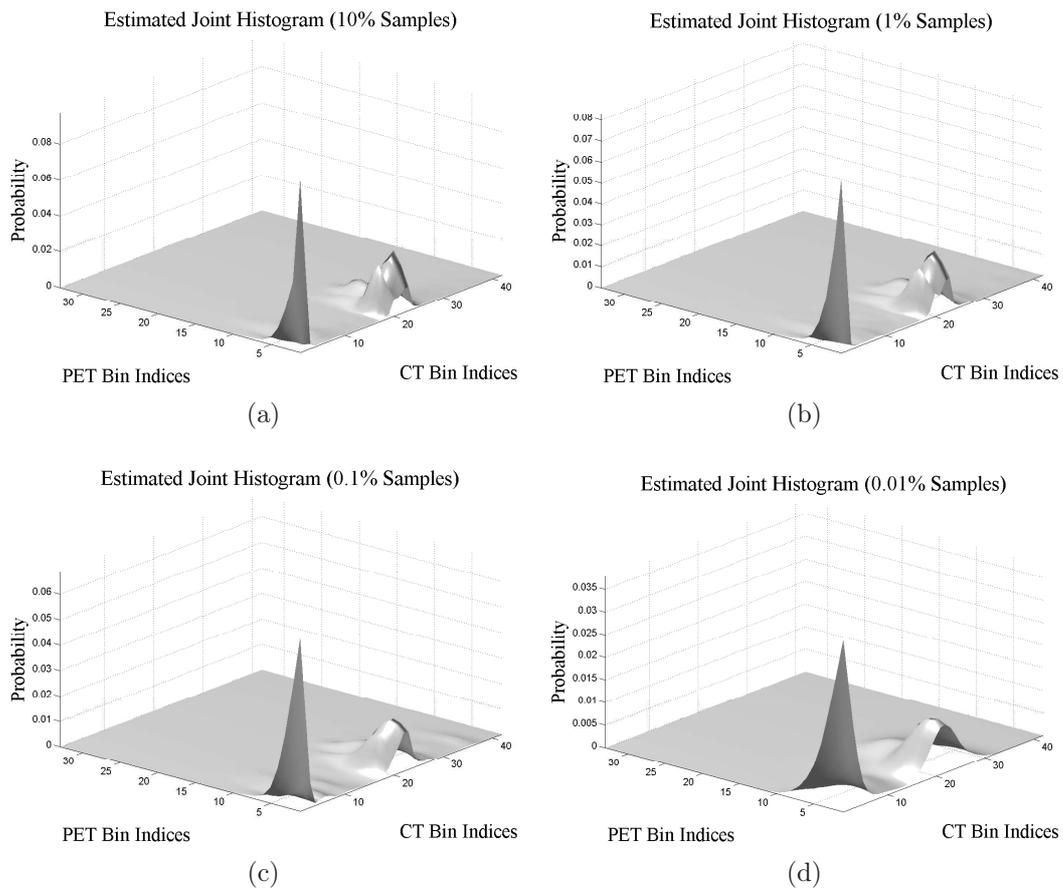


Figure 4.11: Results of the proposed PDF estimation for a PET-CT image pair. The number of samples for the estimation has been (a) 10%, (b) 1%, (c) 0.1%, and (d) 0.01% of the number of voxels within the overlap domain.

rates, therefore, are smoothed representations of the high sampling rate PDF, but its overall structure can still be recognized.

Picking up the registration example at the beginning of this chapter (see Figure 4.1), we have applied the proposed estimators in a new registration attempt. Figure 4.12 shows the positive impact on the registration that is achieved by optimal

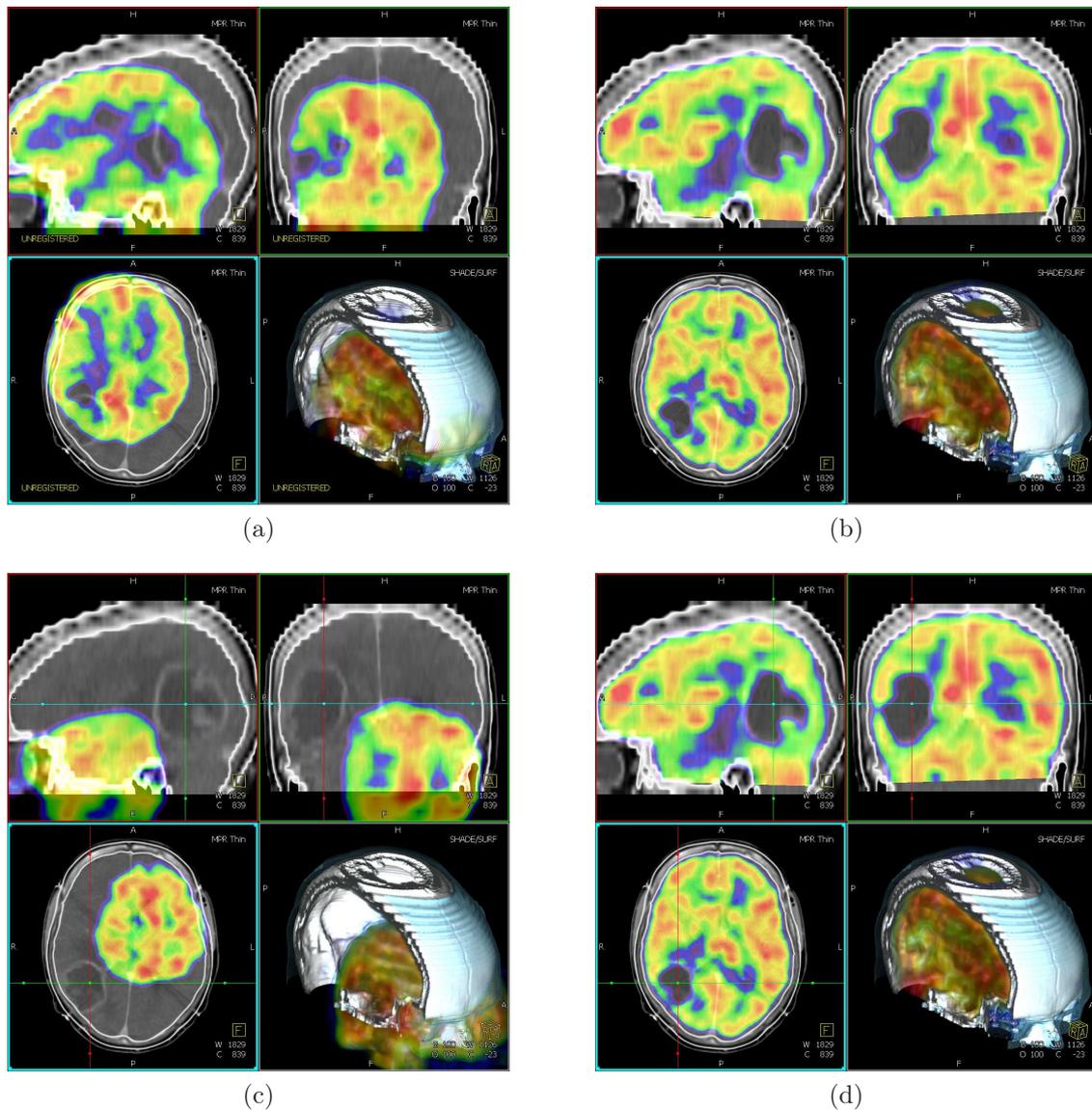


Figure 4.12: Registration of the images provided in Figure 4.1 using the proposed PDF estimation techniques. The final transforms computed for both initial starting positions shown in (a), (c) yield the results in the right column (b), (d). Here, the alignment is consistent and more accurate.

parameters values. The proposed approach clearly indicates an improvement of both the accuracy and the robustness compared to the initial result in Figure 4.1. A more detailed evaluation of the data-driven approaches can be found in Chapter 7, which contains the results for the fully automatic PDF estimation used in a registration of the RIRE database.

Chapter 5

Projection-Based Registration

From a computational point-of-view, the evaluation of the similarity measure is the most time-consuming operation within a registration algorithm. In non-parametric techniques, the optimization problem is ill-conditioned and, therefore, an additional regularization energy is required. The situation is different for parametric techniques. In the case of rigid transforms, the range for the measure computation is very large, compared to the six parameters for the transform, which yields a well-conditioned optimization problem. As this system is overdetermined for rigid transformations, sparse sampling and statistical estimation methods (see Chapter 4) are often used to reduce computational costs. In the following, we present an alternative to previously mentioned approaches to reduce the amount of computations by a division of the parameter space into mutually disjoint subsets of lower dimensions. The related optimization problems can be computed separately. This technique is based on the concept of marginalization to become invariant to specific parameters. In addition, the novel projection scheme for the marginalization is suitable for a parallel implementation, for example on a graphics processing unit (GPU). An example implementation is presented in Chapter A.

There is still the need for fast – here, the term *fast* actually means close to real time – rigid registration algorithms on state-of-the-art desktop workstations. Of course, registration algorithms can be solved in near real time on dedicated, highly parallel workstation clusters or supercomputers (e.g. see the work of Christensen [Chri98], Ino [Ino05], or Salomon et al. [Salo05]). Nonetheless, rigid registrations have to be computed also on general workstations in many clinical scenarios, such as in tumor therapy monitoring, digital subtraction angiography (DSA), or as an initial step for a subsequently applied, deformable registration algorithm. As described in Section 3.3, there exist quite a number of intensity similarity measures. Some are only suitable for mono-modal cases, while others can also deal with different types of multi-modal problems. They all have in common that a straightforward implementation leads to a pixel-by-pixel distance measure evaluation, which is expensive for very large images. Therefore, approaches for solving the nonlinear optimization problems (e.g. Newton, Gradient Descent, Levenberg-Marquardt) tend to become very time-consuming on standard hardware when it comes to real clinical data. Although multi-resolution approaches can reduce the computational load to a certain amount, the highest accuracy can only be achieved by performing at least some optimization

iterations at the highest resolution level of the input data. These last few iterations may take too much time, even on the latest workstations.

In the following sections, we present a novel decomposition scheme that results in parameter subsets that are optimized disjointly on projections of the image data. The decomposition yields significant speed-ups of the computation time required for the optimization. Compared to state-of-the-art projection-based registration algorithms, the proposed approach is able to deal with subvolume matching problems and enables a complete decomposition of the 6-DOF into disjoint 1-DOF problems. A common issue of related algorithms is the need for reprojections of the high-dimensional space. While solving the 1-DOF optimization problems, the projections do not have to be recomputed for the proposed method, because the parameter subset is related linearly to the projection space. With commonly applied orthogonal projections, this holds true only for the translational parameters. Therefore, we introduce a circular and cylindrical projection geometry for 2-D and 3-D images, respectively, to account for rotation parameters as well. Additionally, the proposed approach can be ported onto dedicated graphics hardware.

5.1 Decomposition of the Parameter Search Space

Widely established, intensity-based registration algorithms usually require the computation of an image similarity measure defined on all pixels of the overlap region between a reference and a template image. Although the partial derivatives of this measure with respect to the transform parameters can often be analytically formulated, its evaluation on a pixel-by-pixel basis is very time-consuming. In addition, it is well known from theory that the conditioning of optimization techniques decreases with an increasing number of dimensions of the parameter space.

In 2000 and 2002, Hornegger and Niemann [Horn00; Horn02] proposed a new method to reduce the complexity of the optimization by projections. The approach is closely related to a technique often used in stochastic frameworks, where the dependencies on specific random variables can be eliminated by marginalization. The authors applied the method to determine the position and orientation of 3-D objects in the field of object recognition. Their work describes how the projection of the data onto a lower dimensional subspace yields an optimization problem of reduced complexity. Solutions for the objective function on these projected sets can be found with significantly less computations, compared to the original problem. In case of the rigid registration, an optimal parameter set for the 6-DOF transform is computed iteratively by solving lower dimensional optimization problems until the global solution converges. In the field of object detection and localization, a similar approach was developed by Ratan et al. [Rata00]. They detected and localized faces within 2-D images by utilizing a small, 2-D template that is warped into an image. Using dynamic programming on the correlation between columns of the template and the image, their approach allowed to convert the localization problem to a 1-D sub-problem to search for the best match.

The principle of reduction of dimensionality of the search space was applied to 3-D image registration by Khamene et al. [Kham06] who proposed to iteratively project the 3-D volumes onto 2-D images along the coordinate axes. This orthogonal

projection scheme eliminates three DOF with each projection: two rotational and one translational parameter in the direction of the projection. Although the presented results indicated a huge gain in computational speed, their approach still couples three DOF within the optimization problems for the 2-D projections. This is due to the usage of orthogonal projections, which are only invariant to changes in the rotation around the projection axis and translations within the projected image plane. Their algorithm can, therefore, reduce the complexity only to 3-DOF optimization problems.

An additional problem often arises from projection methods, if the content of the input data differs largely from each other. This is most often the case for subvolume registration tasks, where one image only consists of a subset of the first one. In related work, ideal assumptions are made on the content of the images, and the authors rely on the distance measure to be invariant to those parts of the images that are only present in one of the projections. We, therefore, propose an extension of the projection scheme to incorporate the overlap domain into the projection geometry. This additional constraint ensures that the projections are only computed within the overlap domain, which depends on the transform parameters.

5.2 Orthogonal Projections

This section illustrates the basic ideas of using projections for the computation of image similarities and how they accelerate the registration. In addition, the projection-based registration approach is extended to support subvolume problems by fitting a suitable projection geometry into the overlap domain.

As mentioned before, the intensity-based, rigid registration of two 3-D images aligns the reference image \mathcal{R} and the template image \mathcal{T} with respect to a 6-DOF transform $\Phi_{\mathbf{a}}^{\mathbf{R}}$. For the sake of simplicity, we present a formulation based on the SSD distance measure between the two images. The parameters of the transform are usually estimated by a nonlinear optimization approach. Finding an optimum is equivalent to finding the zero crossing of its derivative, which consists of a sum of partial derivatives of (3.23) with respect to the transform parameters \mathbf{a} :

$$\begin{aligned} \nabla_{\mathbf{a}} \mathcal{D}^{\text{SSD}}[\mathcal{R}, \mathcal{T}, \Phi_{\mathbf{a}}^{\mathbf{R}}] &= \frac{-2}{|\Omega(\mathcal{R}, \mathcal{T}_{\Phi})|} \int_{\Omega(\mathcal{R}, \mathcal{T}_{\Phi})} (\mathcal{R}(\mathbf{x}) - \mathcal{T}_{\Phi}(\mathbf{x})) J_{\Phi_{\mathbf{a}}^{\mathbf{R}}}^T(\mathbf{x}) \nabla_{\mathbf{x}} \mathcal{T}_{\Phi}(\mathbf{x}) \, d\mathbf{x} \\ &\stackrel{!}{=} 0. \end{aligned} \quad (5.1)$$

The partial derivative (5.1) yields a vector determining the ascent direction of the distance measure. This information is used during the optimization in order to find the direction towards the desired optimum. The drawback of this standard approach becomes obvious when considering the integral: both the measure and the derivative have to be evaluated on the entire overlap domain $\Omega(\mathcal{R}, \mathcal{T}_{\Phi})$. In practice, large sets of samples have to be drawn from this domain, which requires interpolation operations (see also Section 4.1). This has to be repeated for each iteration in the nonlinear optimization process.

The general idea of projection approaches is to eliminate some DOF in the parameter space by generating projections, i. e. to perform a marginalization of the

corresponding transform parameters. Although the optimization of the similarity measure between the projections is not the same as for the original images, we assume that the result is identical, i. e. the similarity between the projections yields an optimum for a correct alignment of the 3-D images. Ideally, the projections depend linearly on the remaining parameters. This crucial requirement allows to refine the remaining parameters in a single projection step without having to reproject. The principle is illustrated in the Figure 5.1 taken from [Horn 00]. The figure shows the

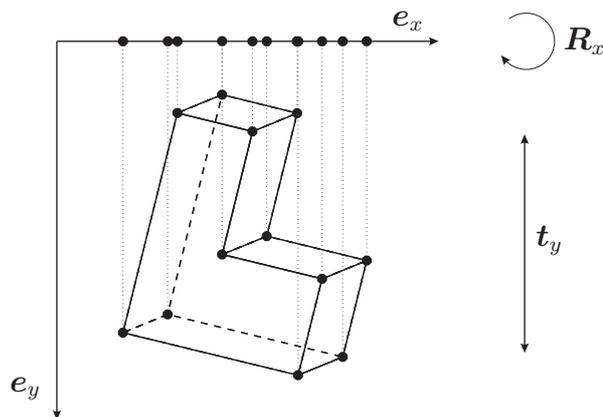


Figure 5.1: Illustration of a linear projection of features onto the x -axis e_x . The projection remains the same, regardless of how the object is translated along the y -axis (t_y), or rotated around the x -axis (R_x). If the object is translated parallel to the x -axis, the projection is translated accordingly. In these cases, it is, therefore, not necessary to reproject. This figure was taken from [Horn 00].

projection of a set of 3-D features onto the x -axis. Both a translation of the features along the projection direction and a rotation around the x -axis do not affect the projection. In the case of medical image registration, the spatial sampling positions may be regarded as features. In general, shifting an image along a direction that is orthogonal to the projection axis translates its projection by the same amount, and vice versa. An optimization within the reduced parameter space after the marginalization does not require reprojections. Of course, if the volume is rotated around a different axis, the previous projection no longer fits to the changes in the parameter set and a reprojection becomes necessary. As far as only changes in the translational part of the parameter vector are considered, orthogonal, axis-aligned projections onto 1-D images are sufficient to cover the entire variation in the parameters. This is illustrated in Figure 5.2a. Using orthogonal projections along the coordinate axes corresponds to a marginalization over the translational parameters.

5.3 Nonlinear Projection Schemes

To account for changes within the parameter subset that belongs to rotations, we propose a nonlinear projection scheme that is described in the following section. Orthogonal projections cannot be used in these cases, as they are valid only for one rotational parameter setting, and a change of the rotation requires a reprojection,

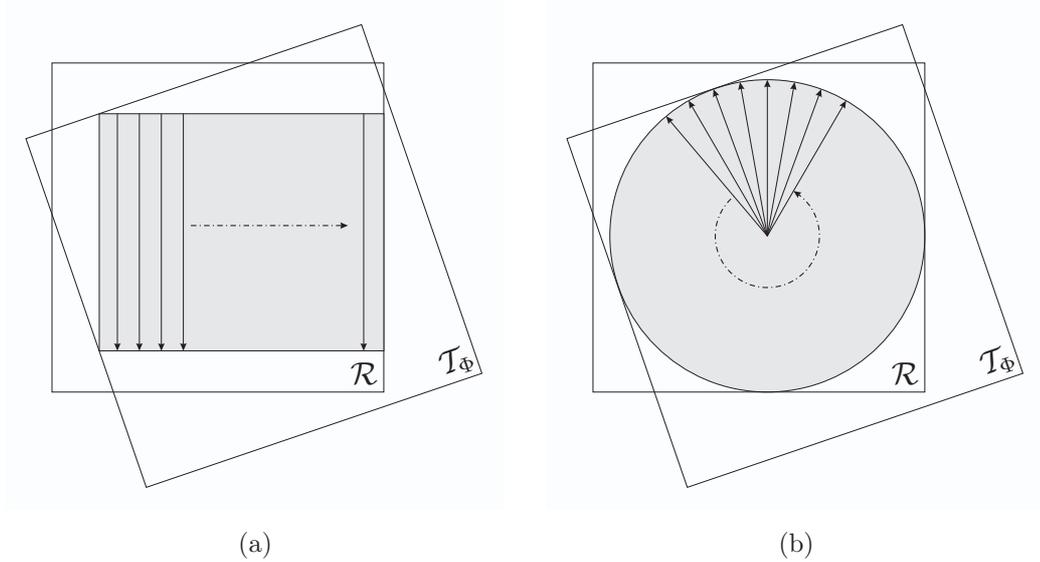


Figure 5.2: Projection schemes for 2-D images to separate (a) translational and (b) rotational parameters.

which is prohibitive. We will extend the approach for 2-D problems first, before considering the more difficult projection scheme for 3-D images.

5.3.1 Circular Projections for 2-D Problems

In order to achieve a complete separation of the registration parameters into disjoint 1-DOF subsets, it is necessary to account for the rotational parameters. A projection geometry has to be selected for the rotations that exhibits the same property as the axial projection in Figure 5.2a for the translational parameters: a linear relation between the projection and the parameter. We propose to cast rays from the center of a circle onto its boundary for the rotational parameter optimization in 2-D registration problems. The principle is shown in Figure 5.2b. A rotation of the template image around the center of the circle is reflected as a translational shift of the circular projection on the boundary. This enables an optimization of the rotation parameter independent of the translation, and also without having to reproject when the rotation angle is changed. If the same circular projection geometry is chosen for both the reference and template image, the parameter for a rotation around the center of the circle can, therefore, be optimized in the 1-D circular projections. Only one circle projection for each image is required for several optimization steps in this reduced 1-DOF parameter space, similar to projections onto the coordinate axes for the translations.

Unfortunately, the optimization results in the several projections are to some degree dependent on each other: a translation along an axis has some influence on the overlap and, thus, on the center for the circular projection. Therefore, the content within the projection images depends on the optimization result of the previous parameter subsets. In the following, we will combine the axis-aligned and the circular projections in an iterative optimization scheme that can be applied to 2-D registra-

tion problems. It is highly efficient with respect to the operations performed in the original 2-D image space, as a single projection of the data has approximately the same computational complexity as one evaluation of the standard distance measure (5.1) on the entire overlap domain.

5.3.2 Cylinder Projections for 3-D Problems

Lifting the 2-D circular projection approach into 3-D space in a straightforward manner results in a spherical geometry. However, the parametric representation of the surface implies a nonlinear relation and, additionally, the discretization at the poles is difficult. Representing the rotational parameters in spherical coordinates does not yield a linear relation to the projection onto the surface of the sphere. Modifying one parameter on the surface, therefore, results in a nonlinear warping within the projection domain, which can only be compensated by means of a reprojection.

We propose to use a cylindrical projection geometry that does not separate between all rotational and translational parameters, but rather combines one rotational with one translational DOF. This, again, yields linear relations between the projections and the parameter values. A rotation around the axis of the cylinder, similar to the rotation around the center of the circle in the 2-D case, relates to a shift of the projection along the radial component of the cylinder surface. Likewise, a translation along the axis of the cylinder corresponds to a translation of the surface projection along its second coordinate axis. The two operations are mutually independent from each other, which enables additional projections to separate the parameters into 1-DOF problems, and, in both cases, without having to reproject during their optimization. Dealing with the surface data on basis of the cylinder coordinates is not necessary. The surface projections may be regarded as common 2-D images that result from unfolding the cylinder into a 2-D image plane. The x -axis then corresponds to the rotation angle, the y -axis to the translation along the cylinder axis. For the x -axis, periodic boundary conditions have to be applied, as the original cylinder surface is connected along the entire boundary. A rotation around the cylinder axis in the 3-D image space is reflected by shifting the projection image along its x -axis. Analogously, a translation along the cylinder axis relates to a translation of the projection along its y -axis. An example for a cylinder projection of a CT image of a human head is given in Figure 5.3. The surface image can then be orthogonally projected along its axes, as described in Section 5.2, which results in a fully decoupled sequence of 1-DOF optimizations. A single cylinder projection corresponds to two parameters: one rotation angle and a translation. In order to represent the entire six DOF of a rigid transform, three cylinder projections are necessary, with the axes being orthogonal to each other. Since the cylinder projections are not independent of each other, it is necessary to perform the optimization iteratively by repeatedly projecting onto each of the cylinder geometries in an alternating order.

Computing the projection along the rays is similar to standard volume rendering techniques [Enge06]. Performing a single projection from 3-D to 2-D is very similar to creating a digitally reconstructed radiograph (DRR), a technique used for 2-D/3-D registrations. The projections along a single axis correspond to line integrals where the intensities are accumulated. The approach can be ported to dedicated graphics

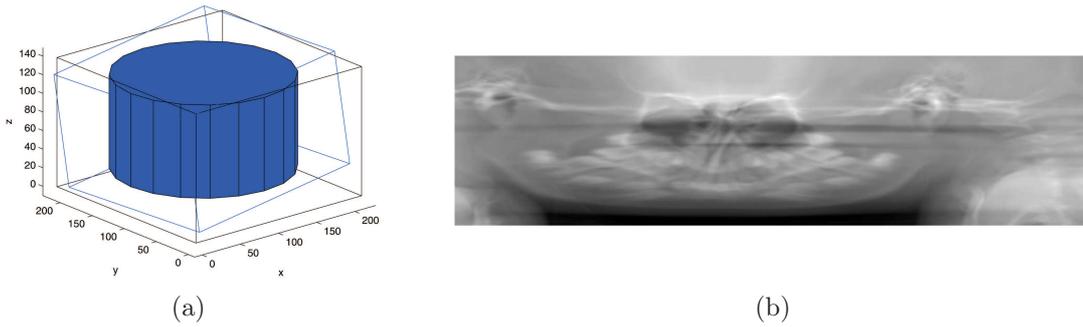


Figure 5.3: The images contain (a) the illustration cylinder geometry fitted into the overlap domain of two CT head datasets, the 3-D cubes mark the boundaries of the image domains, and (b) one corresponding unfolded cylinder projection of the 3-D images. Its x -axis corresponds to a rotation parameter around the cylinder axis, the y -axis to a translation along the axis.

hardware (see the corresponding section in the Appendix A). Additionally, it is possible to incorporate transfer functions or segmentation results (e.g. see Hahn et al. [Hahn 05]).

5.3.3 Fitting the Projection Geometry

Medical image registration often involves subvolume problems, where only a part of the content of one image is shown in the other one. If the projection approach does not account for the overlap domain, the projection is superimposed by the additional content of the larger image, and, thus, an unwanted bias is introduced into the similarity measure. An example for this problem is provided in Figure 5.4, that, for the sake of simplicity, shows the effect of varying image domains for an orthogonal projection along the main axis of a full body CT. As we want to evaluate the similarity in the reduced space, we have to make sure that the projections are computed only within the overlap domain. Otherwise, the superimposed structures lead to a displacement of the correct optimum within the subspace objective function. Given the images in Figure 5.4, the restriction of the projections to the overlap domain removes the body parts below the pelvis seen in Figure 5.4a.

In order to incorporate this requirement into the presented scheme, the projection geometry has to be fitted into the overlap domain between the two images. Depending on the dimensionality of the images, various optimization problems have to be solved. For 2-D geometries, we present solutions for the fitting of a square, a rectangle, and a circle. In 3-D, the corresponding problems are to find the optimally fitting positions of a right cuboid, a rectangular box, and a cylinder. Suitable techniques to tackle the constrained optimization problems are comprised of simplex methods and sequential quadratic programming [Wach 06]. By fitting the geometry, we also maximize the information contained within the projections. Therefore, we retain the ability to deal with subvolume registration problems while keeping a maximal amount of overlap information within the objective functions.

Fortunately, the overlap domain, if not being empty, is always a convex polytope for medical images, as it is defined by a space partitioning with half spaces

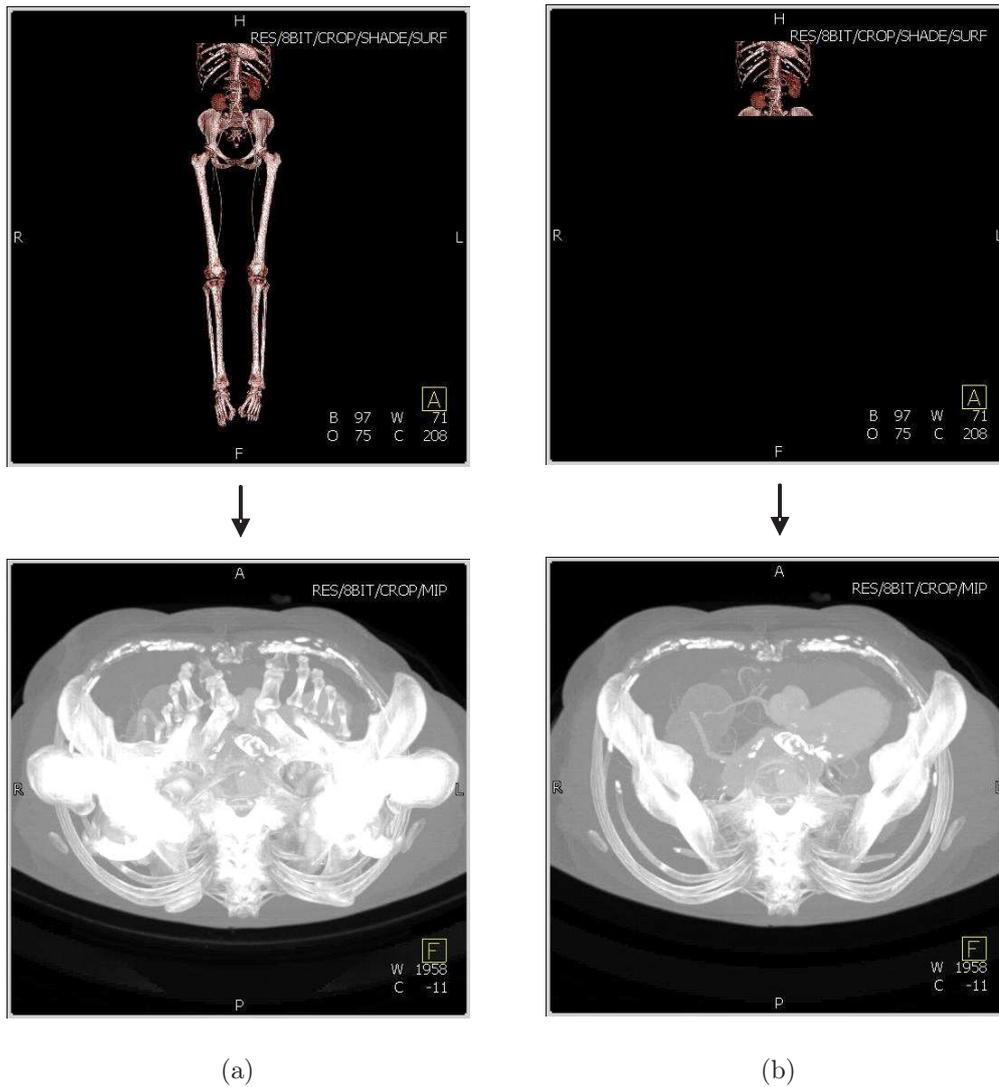


Figure 5.4: The images (a) within the left column show a projection along the longitudinal axis of the full image domain. The reduction of the domain (b) results in a different projection. The structures that are only visible in the left images (the legs or the hip) contribute to the projection image and lead to additional information that impairs a subsequent registration.

representing the boundaries of the image domains. In addition, we demand that the projection geometries have to be axis-aligned, thus, the optimization problems can be solved very efficiently using linear and quadratic programming.

2-D Problems

We will first describe the problems of fitting a square, a rectangle, and a circle into the overlapping region of two intersecting rectangles. An axis-aligned rectangle is

parameterized with an offset corner point $\mathbf{x}_{\text{rect}} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ and two side lengths $l_x, l_y \in \mathbb{R}$:

$$\text{rect}_{\text{AA}} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x + l_x \\ y \end{pmatrix}, \begin{pmatrix} x + l_x \\ y + l_y \end{pmatrix}, \begin{pmatrix} x \\ y + l_y \end{pmatrix} \right\} = \{\mathbf{x}_{\text{rect}}, l_x, l_y\} . \quad (5.2)$$

An arbitrarily oriented rectangle is defined by boundary lines that are pairwise orthogonal to each other. The intersections of these lines specify the corner points of the rectangle:

$$\text{rect} = \{L_1, L_2, L_3, L_4\} = \{\mathbf{n}_1, \dots, \mathbf{n}_4, d_1, \dots, d_4\} , \quad (5.3)$$

where L_i is the i -th line that defines the boundary of the rectangle. Each line partitions the coordinate system into two half spaces:

$$L = \{\mathbf{y} \mid \mathbf{n}^T \mathbf{x} = d\} , \quad (5.4)$$

with the vector \mathbf{n} , $\|\mathbf{n}\| = 1$, denoting the normal to the line, and d its distance to the origin. Thus, a point \mathbf{x} lies within the negative half space, which is the half space opposite to the normal vector direction, if:

$$\mathbf{n}^T \mathbf{x} < d , \quad (5.5)$$

or within the positive half space, otherwise. We define the boundaries of the rectangle with the normal vectors pointing to the outside of the rectangular region. The inner region of the rectangle is, therefore, defined by the intersection of all negative half spaces that are given by the boundary lines. In the following, we represent the image domains in 2-D by two rectangles (5.3). Therefore, the overlap domain between the images is determined by the inner region of the convex polytope that is shaped by their boundary lines (5.4).

Fitting a Square The following parameter vector is used to describe an axis oriented square:

$$\mathbf{a} = (\mathbf{x}_{\text{rect}}, l_x)^T . \quad (5.6)$$

As $l_x = l_y$ is a property of the square, we can simplify the parameters in (5.2). The problem of fitting a square into the overlapping region can then be stated as a linear maximization problem for the side length l_x , but subject to each corner of the square being located within the overlap region. Suppose the boundaries of the 2-D reference image domain are defined by the rectangle $\text{rect}_{\mathcal{R}}$, and for the transformed template by $\text{rect}_{\mathcal{T}_{\Phi}}$, accordingly. Then, the overlap domain is described by a set of eight normal vectors $\mathbf{n}_1, \dots, \mathbf{n}_8$ and the corresponding distance values $\mathbf{d} = (d_1, \dots, d_8)^T$, with the first four elements in each set belonging to $\text{rect}_{\mathcal{R}}$ and the others to $\text{rect}_{\mathcal{T}_{\Phi}}$. The optimization problem for the fitting of a square with maximal size is then:

$$\begin{aligned} \max \quad & l_x \quad , \quad \text{with } l_x \in \mathbb{R} \\ \text{s. t.} \quad & \mathbf{M}\mathbf{a} \leq \mathbf{d} . \end{aligned} \quad (5.7)$$

with $\mathbf{M} \in \mathbb{R}^{32 \times 3}$ being the measurement matrix for the constraints. It consists of the components of each rectangle's line normals, which are arranged in a matrix that defines the linear system of inequalities:

$$\mathbf{M} = \begin{pmatrix} n_{11} & n_{12} & 0 \\ n_{11} & n_{12} & n_{11} \\ n_{11} & n_{12} & n_{11} + n_{12} \\ n_{11} & n_{12} & n_{12} \\ \vdots & \vdots & \vdots \\ n_{81} & n_{82} & n_{82} \end{pmatrix}. \quad (5.8)$$

Here, n_{ij} is the j -th element of the i -th normal vector \mathbf{n}_i . The problem 5.7 can be solved by means of linear programming, for example with a two-phase simplex approach [Padb99].

Fitting a Rectangle In order to fit an axis-oriented 2-D rectangle into the overlap domain, a quadratic optimization problem has to be solved with similar constraints as in (5.7). The parameter set, which describes a general 2-D rectangle, is given by:

$$\mathbf{a} = (\mathbf{x}_{\text{rect}}, l_x, l_y)^T. \quad (5.9)$$

In the case of a square, the maximization problem is linear, as the two parameters l_x and l_y are equal. Now, we have to incorporate the area of the rectangle and, therefore, the problem becomes quadratic:

$$\begin{aligned} \max \quad & l_x l_y, & \text{with } l_x, l_y \in \mathbb{R} \\ \text{s. t.} \quad & \mathbf{M}\mathbf{a} \leq \mathbf{d}. \end{aligned} \quad (5.10)$$

The measurement matrix for (5.10) consists of the following elements:

$$\mathbf{M} = \begin{pmatrix} n_{11} & n_{12} & 0 & 0 \\ n_{11} & n_{12} & n_{11} & 0 \\ n_{11} & n_{12} & n_{11} & n_{12} \\ n_{11} & n_{12} & 0 & n_{12} \\ \vdots & \vdots & \vdots & \vdots \\ n_{81} & n_{82} & 0 & n_{82} \end{pmatrix}. \quad (5.11)$$

A solution for this problem can, for instance, be computed numerically using sequential quadratic programming. We applied an interior point method described by Wächter and Biegler [Wach06] after calculating an initial starting point based on the simplex solution for a square.

Fitting a Circle Using the fitted rectangular geometry, the translational parameters along the x - and y -axis can be separated. In order to achieve a decomposition of the rotational parameters for a 2-D registration problem as well, it is necessary to fit a circle geometry into the overlap polytope. Changes in the angle of a rotation within a 2-D image plane are, then, reflected by a shift on the boundary of the circle that is

located at the center of the rotation. The ideal circle representing the image content within the overlap domain has the maximal area and is located completely within the intersection polytope. Thus, the related fitting problem yields parameters for its position and radius. Similar to the case for the square, the convexity of the polytope can be utilized to reduce the fitting to a linear, constrained optimization problem. The criterion for an optimal circle is a maximal radius with respect to the circle being located completely within the overlap domain. The circle is parameterized by:

$$\mathbf{a} = (\mathbf{x}_c, r)^T, \quad (5.12)$$

with \mathbf{x}_c being the center of the circle, and r its radius. The radius constraint for a circle located at \mathbf{x}_c and limited by a line L is given by:

$$r = d - \mathbf{n}^T \mathbf{x}_c. \quad (5.13)$$

The linear inequality constraints can, therefore, be reformulated to:

$$\mathbf{n}_i^T \mathbf{x} + r \leq d_i, \quad \forall i = 1, \dots, 8. \quad (5.14)$$

These constraints yield the following measurement matrix:

$$\mathbf{M} = \begin{pmatrix} n_{11} & n_{12} & 1 \\ n_{21} & n_{22} & 1 \\ \vdots & \vdots & \vdots \\ n_{81} & n_{82} & 1 \end{pmatrix}. \quad (5.15)$$

The resulting linear problem is similar to (5.7), but now with respect to the radius of the circle:

$$\begin{aligned} \max \quad & r, \quad \text{with } r \in \mathbb{R} \\ \text{s. t.} \quad & \mathbf{M}\mathbf{a} \leq \mathbf{d}. \end{aligned} \quad (5.16)$$

Although the maximization of the area within a circle, in general, is a quadratic problem, the convexity property of the problem domain allows for a linear formulation. As it was the case for (5.7), the problem can be solved by linear programming.

3-D Problems

The 2-D problem of fitting a rectangle into the convex polytope of the overlap domain boundaries can be analogously extended to 3-D. Fitting an axis-aligned, right cuboid leads to a linear problem similar to (5.7), but with the additional component of the z -axis and six half spaces, instead of four. The complexity for a general, rectangular box becomes cubic, as the maximal volume is subject to the optimization in 3-D. However, the formulas may be easily derived from (5.10) described above. Therefore, we will concentrate on the more complex case of fitting a cylinder in the following.

Fitting a Cylinder The complete separation of the six DOF for a rigid registration of 3-D images requires three cylinders. The axes of the cylinders should be mutually orthogonal to achieve the maximal coverage of the rotational parameters. Thus, we formulate the problem for arbitrary cylinder orientations. The objective for the optimization is the maximization of the cylinder volume within the intersection domain. As the intersection polytope is convex, it is sufficient to define constraints for the points that are lying on the boundaries of the top and bottom circular caps of the cylinder. If all of these points are located within the intersection polytope, the entire cylinder is contained within the intersection domain, as well. The problem is, therefore, similar to (5.16) and the formulations are based on the 2-D circle fitting. The minimal distance between the center point of the circular cap to the polytope half spaces defines the radius, which has to be maximized. Within this maximization, we also incorporate the height of the cylinder, as it is defined by the distance between the two center points of the circular caps, and achieve a coupling between both the radius and the volume. The axis of the cylinder determines the spatial orientation and defines the vector between the cap centers.

In 3-D, the overlap domain can be described by a convex polytope, which consists of 12 planes P , each being parameterized by a normal vector and a distance value:

$$\text{poly} = \{P_1, \dots, P_{12}\} = \{\mathbf{n}_1, \dots, \mathbf{n}_{12}, d_1, \dots, d_{12}\} . \quad (5.17)$$

The points \mathbf{x} within the cylinder are defined by the center point of the lower circular cap $\mathbf{x}_{\text{cap}} \in \mathbb{R}^3$, the normalized axis orientation vector $\mathbf{o} \in \mathbb{R}^3, \|\mathbf{o}\| = 1$, the height $h \in \mathbb{R}$, and the radius $r \in \mathbb{R}$:

$$\text{cyl} = \left\{ \mathbf{x} \mid \underbrace{\mathbf{o}^T \mathbf{x}_{\text{cap}} \leq \mathbf{o}^T \mathbf{x} \leq \mathbf{o}^T \mathbf{x}_{\text{cap}} + h}_{\text{location between the cap planes}} \wedge \underbrace{\|\mathbf{x}_{\text{cap}} - (\mathbf{x} - (\mathbf{o}^T \mathbf{x})\mathbf{o})\| \leq r}_{\text{location within the radius}} \right\} . \quad (5.18)$$

An illustration of the cylinder geometry is shown in Figure 5.5. Here, the cylinder is defined by the intersection points \mathbf{x}_i , located on the planes P_i , with $i = 1, 2, 3$.

Similar to the maximal radius constraints for the circle (5.14), the distances of the cylinder cap center points, $\mathbf{x}_{\text{cap}}, \mathbf{x}_{\text{cap}} + h\mathbf{o}$, to the planes $P_i, i = 1, \dots, 12$, are used to constrain the radius of the cylinder. However, the distance is not measured directly between the cap centers and the planes, but within the top and bottom cap planes. The boundary planes of the polytope and the cylinder caps intersect either in a line, a common plane, or not at all, depending on whether the plane's normal vector is parallel to the cylinder axis.

For a normal vector \mathbf{n}_j of the j -th polytope plane, with $\mathbf{n}_j \not\parallel \mathbf{o}$, its normalized projection onto the lower cap plane is denoted by $\hat{\mathbf{n}}$. The distance \hat{d}_j to the intersection between P_j and \mathbf{x}_{cap} can then be expressed in terms of a multiple of the projected normal (see also Figure 5.5b):

$$\bar{\mathbf{n}}_j = \mathbf{n}_j - (\mathbf{o}^T \mathbf{n}_j)\mathbf{o} \quad (5.19)$$

$$\hat{\mathbf{n}}_j = \frac{\bar{\mathbf{n}}_j}{\|\bar{\mathbf{n}}_j\|} \quad (5.20)$$

$$\hat{d}_j = \hat{\mathbf{n}}_j^T (\mathbf{x}_{\text{cap}} + \lambda \hat{\mathbf{n}}_j) . \quad (5.21)$$

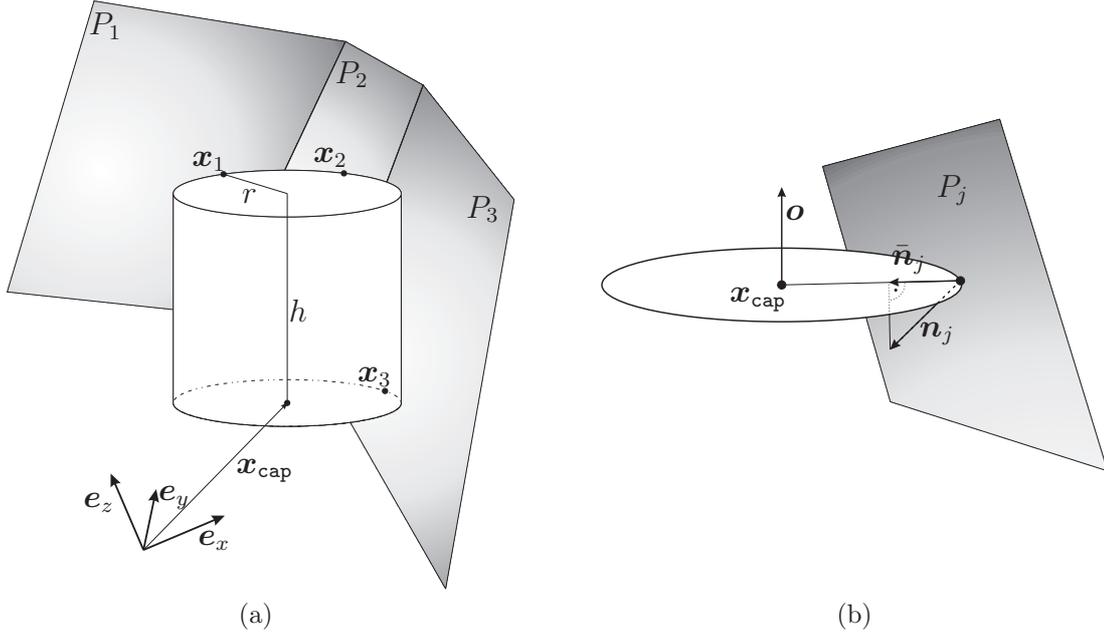


Figure 5.5: Illustrations of (a) the cylinder geometry, and (b) the projection of the j -th plane normal vector onto the cap plane.

Here, $\lambda \in \mathbb{R}$ is the scalar value that determines the distance between the cap center and the intersection. If the plane is tangent to the lower cap circle, the value of the parameter is defined by the distance of the plane to the coordinate origin, and we can formulate an expression for \hat{d}_j :

$$\lambda = \frac{d_j - \mathbf{n}_j^T \mathbf{x}_{\text{cap}}}{\mathbf{n}_j^T \hat{\mathbf{n}}_j} \quad \text{using} \quad \mathbf{n}_j^T (\mathbf{x}_{\text{cap}} + \lambda \hat{\mathbf{n}}_j) = d_j \quad (5.22)$$

$$\hat{d}_j = \frac{d_j}{\mathbf{n}_j^T \hat{\mathbf{n}}_j} + \left(\hat{\mathbf{n}}_j^T - \frac{1}{\mathbf{n}_j^T \hat{\mathbf{n}}_j} \mathbf{n}_j^T \right) \mathbf{x}_{\text{cap}} . \quad (5.23)$$

Maximizing the distance \hat{d}_j between the center points of the caps and the j -th intersection line is equivalent to maximizing the radius of the cylinder, and yields the following constraints for the j -th plane:

$$\hat{\mathbf{n}}_j^T \mathbf{x}_{\text{cap}} + r \leq \hat{d}_j \quad (5.24)$$

$$\hat{\mathbf{n}}_j^T (\mathbf{x}_{\text{cap}} + h\mathbf{o}) + r \leq \hat{d}_j . \quad (5.25)$$

Those polytope planes with normal vectors parallel to the cylinder axis constrain the height, but not the radius. Suppose the normal vector of the k -th polytope plane is parallel to the cylinder axis, i. e. $\mathbf{n}_k \parallel \mathbf{o}$, the length of the projection of \mathbf{n}_k onto the cylinder cap planes has zero length. Therefore, this plane only constrains the cylinder's position and height:

$$\mathbf{n}_k^T \mathbf{x}_{\text{cap}} \leq d_k \quad (5.26)$$

$$\mathbf{n}_k^T (\mathbf{x}_{\text{cap}} + h\mathbf{o}) \leq d_k . \quad (5.27)$$

Using (5.23), the optimization problem for the cylinder fitting is:

$$\max r^2 h \quad , \text{ with } \mathbf{x}_{\text{cap}} \in \mathbb{R}^3, r, h \in \mathbb{R} \quad (5.28)$$

$$\text{s. t. } \mathbf{n}_k^T \mathbf{x}_{\text{cap}} \leq d_k \quad , \forall k \mid (\mathbf{n}_k \parallel \mathbf{o}) \quad (5.29)$$

$$\text{s. t. } \mathbf{n}_k^T \mathbf{x}_{\text{cap}} + h \mathbf{n}_k^T \mathbf{o} \leq d_k \quad , \forall k \mid (\mathbf{n}_k \parallel \mathbf{o}) \quad (5.30)$$

$$\text{s. t. } \mathbf{n}_j^T \mathbf{x}_{\text{cap}} + r \mathbf{n}_j^T \hat{\mathbf{n}}_j \leq d_j \quad , \forall j \mid (\mathbf{n}_j \not\parallel \mathbf{o}) \quad (5.31)$$

$$\text{s. t. } \mathbf{n}_j^T \mathbf{x}_{\text{cap}} + h \mathbf{n}_j^T \mathbf{o} + r \mathbf{n}_j^T \hat{\mathbf{n}}_j \leq d_j \quad , \forall j \mid (\mathbf{n}_j \not\parallel \mathbf{o}) . \quad (5.32)$$

The constraints (5.29)-(5.32) determine the linear system of inequalities, which guarantees that the optimal cylinder is located within the overlap domain. The problem may be solved by sequential quadratic programming [Wach 06].

5.3.4 Uniqueness of the Solution

The optimization with respect to only the distances to the boundaries of the overlap domain does not lead to a unique solution in all cases. For example, if the reference and template image domains are equal except for a translation in one direction, the solution for a cylinder geometry is unique for the radius and the height, but not for the cap center point \mathbf{x}_{cap} . Any cylinder with the optimal height and radius placed on the line segment depicted in Figure 5.6 fulfills the constraints. All cylinders located

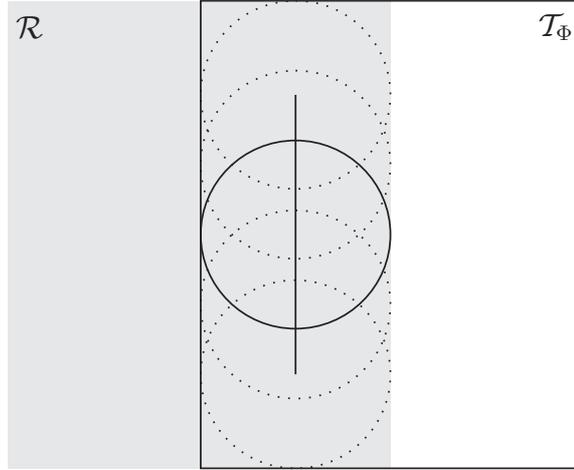


Figure 5.6: Schematic view from top onto two equal but shifted image domains and the solution space for the location of the cylinders, indicated by the line segment.

in this solution space have the same volume, however, from a registration point of view, it is important which position is chosen for the projection. If we assume that the important content within the images is located at the center of the domain, we prefer a solution that accounts for this additional criterion. It is also beneficial for the optimization of the rotation parameters to place the cylinder axis close to the center of the reference volume. In order to achieve a unique solution with this additional requirement, we propose to extend the objective for the maximal cylinder volume by an additional term that punishes the distance to the center point $\mathbf{c}_\mathcal{R} \in \mathbb{R}^3$ of the reference domain. The direct Euclidean distance to the cap centers would bias the

result towards cylinders of smaller height. Instead, we measure the distance $d_{\mathcal{R}}$ from \mathbf{x}_{cap} to the perpendicular of $\mathbf{c}_{\mathcal{R}}$ within the plane of the lower cap center:

$$d_{\mathcal{R}} = \|\mathbf{x}_{\text{cap}} - (\mathbf{c}_{\mathcal{R}} - (\mathbf{o}^T \mathbf{c}_{\mathcal{R}}) \mathbf{o})\| . \quad (5.33)$$

This additional distance is incorporated into the maximization problem in order to resolve ambiguities and locate the cylinder more in the center of the reference image. Applied to the example in Figure 5.6, the solution then consists only of one cylinder located in the middle of the line segment:

$$\max \alpha r^2 h - (1 - \alpha) d_{\mathcal{R}} \quad , \quad \text{with } \mathbf{x}_{\text{cap}} \in \mathbb{R}^3, r, h \in \mathbb{R} , \quad (5.34)$$

where $\alpha \in [0, 1]$ is the weighting factor between the two objectives. Of course, the constraints (5.29)-(5.32) still apply.

This concept to enforce a unique solution can be analogously adapted to the other fitting problems.

5.4 Comparison to Marginal Space Learning

A technique that is similar to the aforementioned projection approach is marginal space learning (MSL), introduced by Zheng et al. [Zhen07]. The idea of MSL is to reduce the parameter search space for the initial placement of a model within an image, for example in image segmentation. The problem can also be regarded as a registration between the model and the image content. Usually, the goodness of a specific model placement is determined using classifiers, which have been trained on specific image features. The projection approach bears a resemblance to MSL, as both techniques try to reduce the parameter space by marginalization. In MSL, the PDF for the image features, with respect to all of the transform parameters, is not known a priori, but it is assumed to contain a cluster around the optimal position and orientation parameters. An exhaustive search for this cluster in the high-dimensional parameter space is prohibitive, as the computational complexity increases exponentially with each dimension. A reduction could be achieved by a marginalization for a subset of the parameters, however, as the PDF of this reduced set is also unknown, the marginalization has to be accomplished by specific types of features that are invariant to the parameters one seeks to eliminate. Determining the PDF of these features is assumed to be equivalent to the marginalization of the high-dimensional PDF. Thus, the classifiers are trained directly on the ‘‘marginalized’’ PDFs. Large portions of the entire parameter space are quickly eliminated by the search for clusters in the marginals. The resulting, restricted regions of the high-dimensional space can then be effectively processed using exhaustive search methods for the optimal parameters.

In the projection approach, one wants to avoid evaluations in the full parameter space, as this is time-consuming compared to the measure evaluation in the subspaces. In MSL, the optimal marginal parameters are lifted into the higher space and refined there, as well. Instead, we propose to iteratively refine the parameters only in the projections, without having to calculate the measure for the entire image domain.

5.5 Results

The proposed algorithm has been implemented for 2-D and 3-D registrations on the central processing unit (CPU). In addition, results have been acquired with a hardware-accelerated version on the GPU. Details about the latter implementation, which is using the OpenGL Shading Language to specify the shaders on the GPU, can be found in the Appendix A. The following comparison of the algorithm with a rigid registration implemented within the *Insight Segmentation and Registration Toolkit* (ITK)[Iban 05] on the CPU was performed on an Intel Pentium M 2.26 GHz with 2 GByte of main memory, and an Nvidia GeForce 7800 Go. The 2-D CPU registration of two 256^2 medical images takes approximately 10 seconds using a common ITK-based rigid registration approach and no more than 0.3 seconds with the proposed approach. There is basically no GPU acceleration necessary for 2-D registrations to achieve registrations in real time, however, we expect an increase of the frame rate of the registration by a factor of 10 when porting the approach onto the graphics card. In 3-D, an ITK registration of two volumes with 256^3 voxels each lasts more than 10 minutes, if the entire overlap domain is used for the computation of the similarity measure and no statistical sampling techniques are applied. The proposed projection approach on the CPU takes approximately 3 minutes. The hardware-accelerated implementation yields a noticeable speed-up. A single cylinder projection of a 256^3 dataset requires a mean time of approximately 40 milliseconds, which results in a total runtime for the entire approach of less than 10 seconds. The runtime for the projection registration is, therefore, comparable to approaches that make use of sparse statistical sampling. A result for a registration of a CT image pair is shown in Figure 5.7. The images have been acquired at different stages of the therapy. Therefore, a rigid registration can be used to visualize the treatment-related differences. There is still

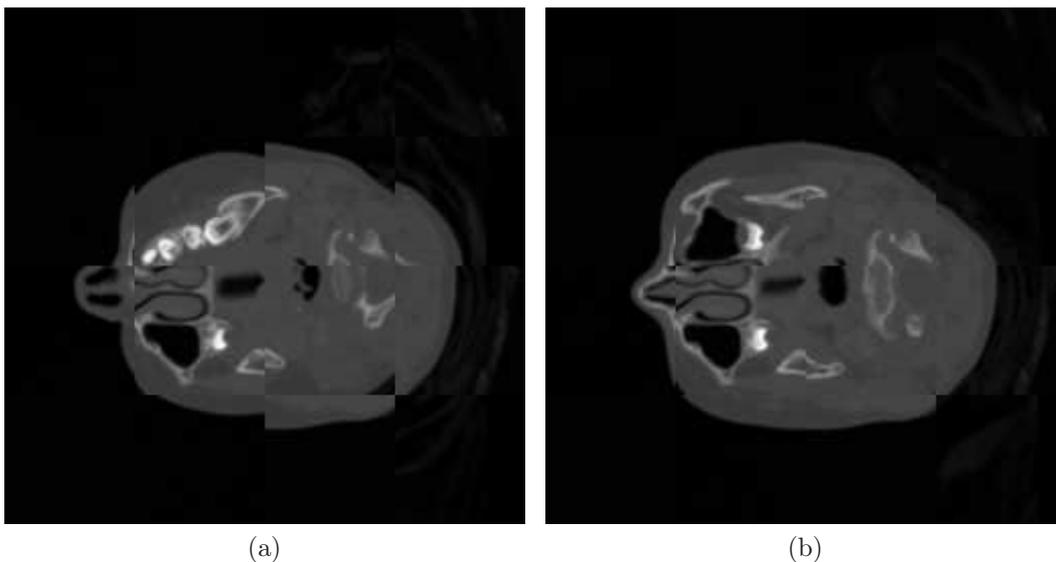


Figure 5.7: Slice taken from the 3-D checkerboard images resulting from a registration between a pre- and post-therapy CT dataset. The images reflect the differences in position and orientation (a) before, and (b) after the projection-based registration.

room for improvement of the GPU implementation. Currently, the bottleneck of the processing chain is the marginalization of the cylinder surface image. These images are still transferred from the GPU into the main memory to perform the summation on the CPU. Newer GPU technologies allow for even faster approaches, because of extended sets of operations implemented in hardware that supersede the costly transfer operations, as the entire computation can be carried out within the programmable graphics pipeline of the GPU.

Chapter 6

An Evaluation Approach for Registration Algorithms

The presented automatic registration approaches rely on an intensity-based similarity measure as the criterion for an optimal alignment. So far, we have discussed techniques to improve stochastic measures by a data-driven parameter estimation, which yields an optimal adaption to the data, and a reduction of complexity by projection schemes. Numerical aspects also play an important role during the optimization and have to be taken into consideration to avoid grid effects. The goal is to design a similarity measure that is both robust and accurate to meet the requirements of clinical applications. Unfortunately, the similarity measure alone is not sufficient to determine the quality of the registration result. We need evaluation schemes to measure the performance of the registration algorithm in terms of accuracy and robustness. Designing an evaluation scheme for image registrations is not an easy task, as a gold standard is often missing. Different placements of the patient for each scan come along with combinations of soft tissue deformations and rigid body movements. It gets even harder for inter-patient registrations, which are needed, for instance, to determine statistical shape variations. The problem is that a ground truth definition of the correspondences between the points is usually unknown.

Fitzpatrick et al. [Fitz98] and West et al. [West97] presented an evaluation for rigid registrations based on ground truth transformations determined from implanted fiducial markers. Although this is a very accurate technique, its field of application is currently limited to rigid, intra-patient registrations of the head, as the markers are screwed into the skull. For other parts of the human body, a rigid transform is insufficient and the placement of the markers raises problems, especially for tracking soft tissue movements. This complexity is one of the main reasons why existing approaches to evaluate non-rigid registration algorithms are currently limited to brain images, because reliable segmentation databases are available for this purpose and the morphology of the brain is not as complex as, for instance, the anatomy of the abdomen. Hellier et al. [Hell03] proposed an evaluation framework based on segmentations of cortical regions. They applied a combination of a global measure, for the overlap between the segmented regions after the registration, with local distances to the cortical sulci. Christensen et al. [Chri06] presented a similar evaluation approach that makes use of richly annotated 3-D MR images of the brain and a series of metrics

to measure the alignment quality. Although there is a website related to the project [NIRE09], it still seems to be in the development phase and uploading of results is currently not supported. Urschler et al. [Urs07] have recently proposed an open-source and open-data framework for the comparison and evaluation of intra-subject image registrations. They proposed to use synthetic deformations and quantitative measures on the deformation fields to assess the validity of deformable registration transforms.

In this chapter, we will present an evaluation framework that is applicable to both parametric and non-parametric registration algorithms. The proposed system requires a database of annotations by medical experts. In contrast to other approaches, it incorporates the observer reliability, i. e. the intra- and inter-observer variation, within a statistical model of the segmented data. The models are evaluated for each registration result using a Kullback-Leibler (KL) divergence between the spatial landmark distributions of the medical experts as measure for the registration quality with respect to the transformation.

6.1 Statistical Evaluation Framework

In general, the knowledge of medical experts is a reliable source of information to determine correspondences between the images. Here, we use it to design statistical models based on manually selected landmarks, labeled contours, and segmented image regions. Commonly, one cannot expect that an automatic registration algorithm delivers “correct” results for parts of the images where even experts are unsure about the true alignment. On the other hand, the registration has to be very accurate in regions where the observer reliability is high. Thus, it is necessary to collect multiple segmentations from each medical expert to incorporate the intra- and inter-observer variability into the statistical model. The segmentations may consist of combinations between point, surface, and region labels. The similarity between the labels $\mathbf{L}_{\mathcal{R}}$ and $\mathbf{L}_{\mathcal{T}}$ with respect to a registration transform $\Phi(\mathbf{x})$ is measured with the symmetric KL divergence between the spatial distributions $p(\mathbf{x}|\mathbf{L}_{\mathcal{R}}^j)$ for the reference and $p(\mathbf{x}|\mathbf{L}_{\mathcal{T}}^j \circ \Phi)$ for the transformed template labels:

$$\mathcal{D}^{\text{SymKL}}[\mathbf{L}_{\mathcal{R}}, \mathbf{L}_{\mathcal{T}}, \Phi] = \frac{1}{L} \sum_{j=1}^L \frac{\mathcal{D}^{\text{KL}}[\mathbf{L}_{\mathcal{R}}^j, \mathbf{L}_{\mathcal{T}}^j \circ \Phi] + \mathcal{D}^{\text{KL}}[\mathbf{L}_{\mathcal{T}}^j \circ \Phi, \mathbf{L}_{\mathcal{R}}^j]}{2} \quad (6.1)$$

$$\mathcal{D}^{\text{KL}}[\mathbf{L}_{\mathcal{R}}^j, \mathbf{L}_{\mathcal{T}}^j \circ \Phi] = \int_{\Omega(\mathcal{R}, \mathcal{T}_{\Phi})} p(\mathbf{x}|\mathbf{L}_{\mathcal{R}}^j) \log \frac{p(\mathbf{x}|\mathbf{L}_{\mathcal{R}}^j)}{p(\mathbf{x}|\mathbf{L}_{\mathcal{T}}^j \circ \Phi)} d\mathbf{x} , \quad (6.2)$$

with L being the total number of landmark sets and the operator “ \circ ” specifying the application of the transformation. If we assume that the selection of each landmark is a random process, which follows a specific distribution, the labels defined by the medical experts can be regarded as random measures. The shape of the distribution provides information about the uncertainty within the selection process. With an increasing sharpness of the distribution the certainty also increases, and vice versa.

6.2 Spatial Density Estimation

Depending on the modalities of the reference and template image, some anatomical shapes are well suited for being point landmarks, for example centers of small lesions that are visible in both images, or structures that can be located precisely. Others may be more suitable for a representation as surfaces or regions. A list of labels is specified for the entire body and the evaluation algorithm dynamically selects the joint subset of labels contained within the overlap domain. Without loss of generality, we assume in the following that the label description consists of a set of d -dimensional points, which are located either at a specific anatomical landmark, on a surface, or within a segmented region. Let $\mathbf{L}_{\mathcal{R}}^j = \{\mathbf{x}_{\mathcal{R},j}^i\}_{i=1,\dots,M}$ be the set of spatial samples for the j -th label in the reference and $\mathbf{L}_{\mathcal{T}}^j = \{\mathbf{x}_{\mathcal{T},j}^i\}_{i=1,\dots,N}$ the points for the corresponding label within the domain of the template image. The samples may originate, for instance, from multiple raters who pick specific locations within the image. The true distributions of the selected labels are not known, therefore, the PDFs required to compute (6.2) have to be estimated. The related problems are very similar to the estimation of intensity distributions in the context of statistical image similarity measures, with the difference that the random measures are now the spatial position samples for the labels.

6.2.1 Parametric Models

We first examine the estimation of the distribution of the labeled samples with a parametric density function. A Gaussian model, for instance, requires only the calculation of the mean and variance, which directly provides information about the uncertainty within the process of the label selection. The distribution for the j -th reference landmark label is given by:

$$p(\mathbf{x}|\mathbf{L}_{\mathcal{R}}^j) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_{\mathcal{R},j}, \boldsymbol{\Sigma}_{\mathcal{R},j}) \quad (6.3)$$

$$\boldsymbol{\mu}_{\mathcal{R},j} = \frac{1}{M} \sum_{i=1}^M \mathbf{x}_{\mathcal{R},j}^i \quad (6.4)$$

$$\boldsymbol{\Sigma}_{\mathcal{R},j} = \frac{1}{M} \sum_{i=1}^M (\mathbf{x}_{\mathcal{R},j}^i - \boldsymbol{\mu}_{\mathcal{R},j}) (\mathbf{x}_{\mathcal{R},j}^i - \boldsymbol{\mu}_{\mathcal{R},j})^T, \quad (6.5)$$

where $\boldsymbol{\mu}_{\mathcal{R},j}$ is the mean and $\boldsymbol{\Sigma}_{\mathcal{R},j}$ the covariance matrix of the normal distribution for the reference landmark samples. The density estimate for the template label is analogue to (6.3), but incorporates the spatial registration transform Φ in addition:

$$p(\mathbf{x}|\mathbf{L}_{\mathcal{T}}^j \circ \Phi) \approx \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_{\mathcal{T}_{\Phi},j}, \boldsymbol{\Sigma}_{\mathcal{T}_{\Phi},j}) \quad (6.6)$$

$$\boldsymbol{\mu}_{\mathcal{T}_{\Phi},j} \approx \frac{1}{N} \sum_{i=1}^N \Phi(\mathbf{x}_{\mathcal{T},j}^i) \quad (6.7)$$

$$\boldsymbol{\Sigma}_{\mathcal{T}_{\Phi},j} \approx \frac{1}{N} \sum_{i=1}^N (\Phi(\mathbf{x}_{\mathcal{T},j}^i) - \boldsymbol{\mu}_{\mathcal{T}_{\Phi},j}) (\Phi(\mathbf{x}_{\mathcal{T},j}^i) - \boldsymbol{\mu}_{\mathcal{T}_{\Phi},j})^T. \quad (6.8)$$

Depending on the type of spatial transform, the equality in the equations above is no longer guaranteed and the assumptions may be violated. In case of a rigid transform,

however, the density of the transformed label set can be specified immediately from the original, non-transformed mean and covariance of the template label [Horn 96]:

$$p(\mathbf{x}|\mathbf{L}_T^j \circ \Phi_a^R) = \mathcal{N}(\mathbf{x}; \mathbf{R}\boldsymbol{\mu}_{T,j} + \mathbf{t}, \mathbf{R}\boldsymbol{\Sigma}_{T,j}\mathbf{R}^T) . \quad (6.9)$$

There exists a closed form, analytical formulation for the result of the KL divergence between the two normal densities [Yosh 99] that can be applied in this case to measure the accuracy of a rigid registration with respect to the segmented landmarks:

$$\begin{aligned} \mathcal{D}^{\text{KL}}[\mathbf{L}_{\mathcal{R}}^j, \mathbf{L}_T^j \circ \Phi_a^R] &= \frac{1}{2} \log \left(\frac{\det \boldsymbol{\Sigma}_{\mathcal{T}_{\Phi,j}}}{\det \boldsymbol{\Sigma}_{\mathcal{R},j}} \right) \\ &+ \frac{1}{2} \text{Tr} \left(\boldsymbol{\Sigma}_{\mathcal{R},j} \boldsymbol{\Sigma}_{\mathcal{T}_{\Phi,j}}^{-1} + \boldsymbol{\Sigma}_{\mathcal{T}_{\Phi,j}}^{-1} (\boldsymbol{\mu}_{\mathcal{R},j} - \boldsymbol{\mu}_{\mathcal{T}_{\Phi,j}}) (\boldsymbol{\mu}_{\mathcal{R},j} - \boldsymbol{\mu}_{\mathcal{T}_{\Phi,j}})^T \right) - \frac{d}{2} . \end{aligned} \quad (6.10)$$

From an implementation point of view, the parametric densities in (6.3) and (6.6) are very appealing. The analytical formulation (6.10) provides an immediate solution for the distance, but it is certainly questionable whether the assumption of normally distributed label samples is justified. Even if the samples are indeed normally distributed, the application of a local, non-parametric transformation yields another kind of distribution, which likely follows no longer a normal distribution. In addition, the mono-modal PDF model for landmark points has to be extended to multi-modal mixture models to represent the spatial distributions of surfaces or regions. The number of the mixture components is not known beforehand and the computational complexity for the estimation of the parameters for the mixture components increases with a growing number of components. Additionally, the calculation of the KL divergence on mixtures is not a trivial task.

6.2.2 Non-Parametric Models

In order to achieve the highest flexibility for the estimation of the spatial densities, we propose to apply a non-parametric density estimation using Parzen-windowing, similar to the methods described in Section 4.2. The spatial PDF estimator for the j -th reference label is, again, discretized using a histogram $h_{\mathbf{L}_{\mathcal{R}}}^j$:

$$\hat{p}(\mathbf{x}|\mathbf{L}_{\mathcal{R}}^j) = (h_{\mathbf{L}_{\mathcal{R}}}^j \star K_{\boldsymbol{\lambda}}^B)(\mathbf{x}) , \quad (6.11)$$

with $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_d)^T$ being the d -dimensional vector that contains the kernel width values for the multi-variate Parzen-windowing. Again, $K_{\boldsymbol{\lambda}}^B$ denotes a cubic B-spline kernel function. The number of bins for the histogram discretization is determined by the spatial resolution of the images and may have a higher resolution than the image in order to achieve subpixel-accurate evaluation results. Figure 6.1a shows an example set of selected points for a landmark label in a 2-D image. The uncertainty within the data is estimated by a maximum likelihood optimization of the optimal kernel widths, as formulated in (4.4). The corresponding log-likelihood function for the example problem is shown in Figure 6.1b. The estimated spatial distributions for a point label using different kernel widths are shown in Figure 6.2. Applying the optimal kernel width with respect to the data-driven cross-validation

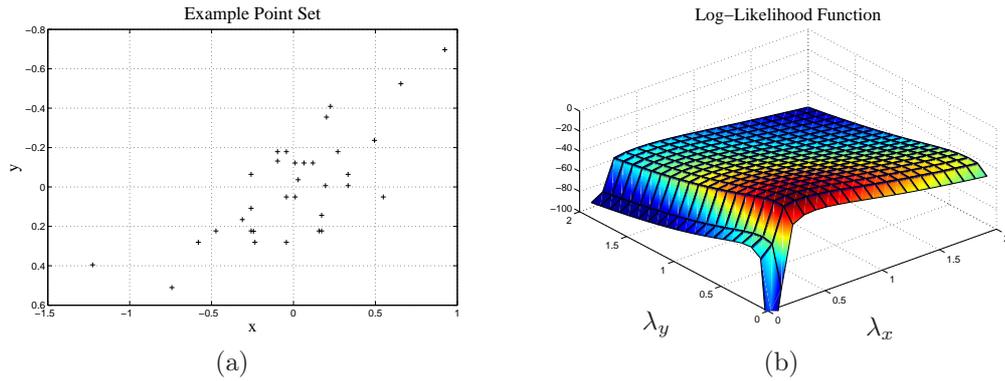


Figure 6.1: (a) Shows an example 2-D point set for a label within an image, (b) is the corresponding log-likelihood objective function that has been sampled at various B-spline kernel widths.

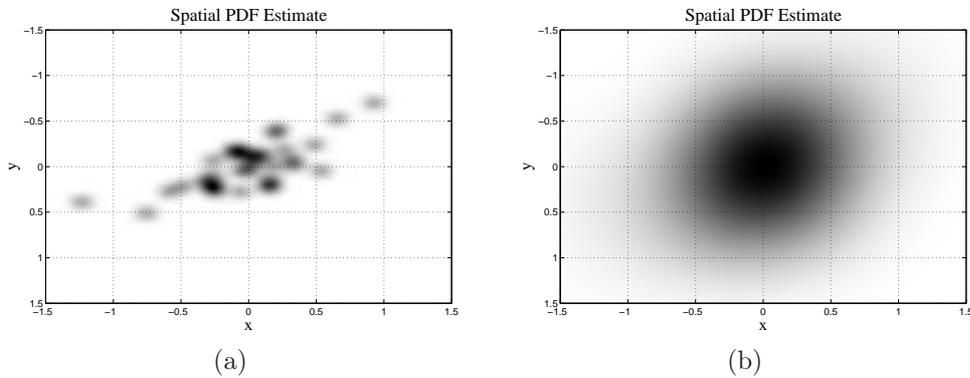


Figure 6.2: The images show (a) the estimated spatial distribution for the data in Figure 6.1 using a B-spline kernel width of $\lambda = (0.01, 0.01)^T$, compared to (b) the estimate using an optimal kernel width of $\lambda = (0.63, 0.99)^T$. The kernel width values are multiples of the pixel resolution in each dimension.

yields the estimate in Figure 6.2b. Similar to the PDF estimation for stochastic similarity measures, the kernel widths need to be optimal to achieve accurate evaluation results. The regions for the plotted distributions are basically defined on the entire image domain, however, for the sake of clarity, the zero-valued parts outside the support range of the kernel function have been omitted. The KL divergence between such spatial distributions can simply be evaluated pixel-wise in a discrete manner using only the non-zero pixels.

Compared to the parametric formulation, the non-parametric density estimation can be easily extended to spatial distributions of surfaces or segmented regions. In Figure 6.3, it is demonstrated how the variation within the label set affects the distributions. The distribution plots 6.3b, 6.3c, 6.3e, and 6.3f show the results of a discrete, B-spline Parzen-window estimator for manual segmentations within an MR slice through the lateral ventricle from a human brain. The label data in this example consists of ten samples that have been created by interactive, manual segmentation with B-spline contours. Darker values within the distribution plots indicate a higher probability for a point belonging to the label at the corresponding pixel position.

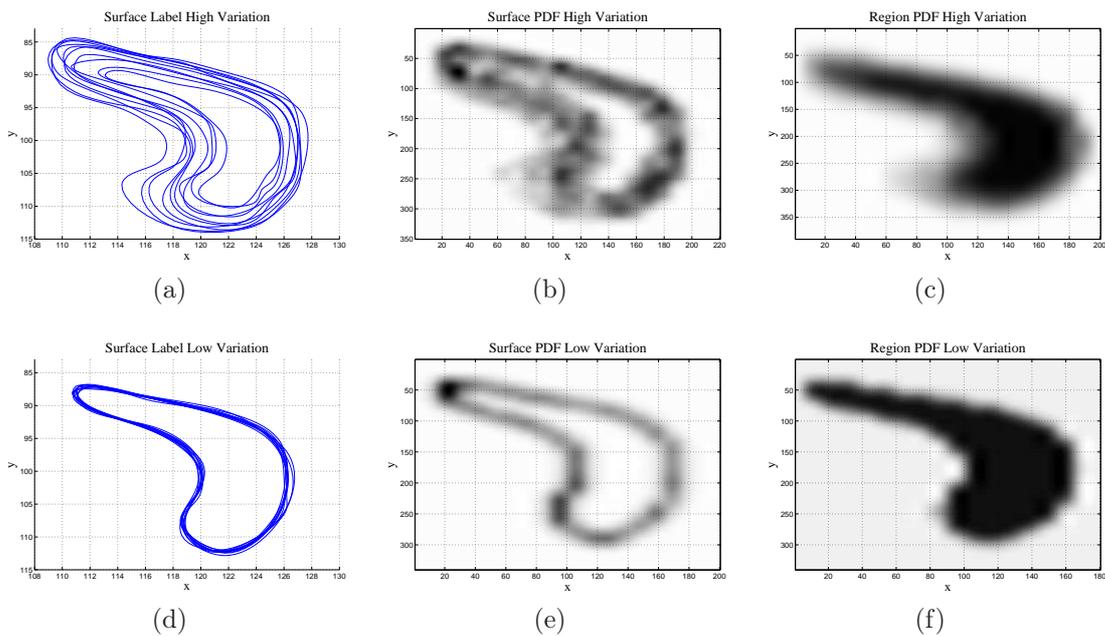


Figure 6.3: The images in the first row show (a) the plot of the segmentations for a label with a high variation, (b) the estimated spatial surface distribution, and (b) the region distribution. In the second row, the plots are arranged in the same ordering, but with a lower variance within the label data.

The intra-label variation is reflected by the spread (the kernel width) of the PDF estimates: it is smaller for the distributions of the second label compared to those in the first row.

6.3 Experiments

In the following, we examine the feasibility of the evaluation framework using six label positions that have been chosen for the evaluation: the sections of the horns of the lateral and temporal ventricles, the cerebellum, the center of the fornix and the lower corner of the insula. Figure 6.4 shows the image together with a delineation of the labels. For the segmentation of the object contours, a manual approach has been implemented that allows the user to interactively adjust control points, which are then interpolated by cubic B-spline curves. The samples for the point labels have been acquired by tracking the positions clicked by the user. Each label consists of 20 samples, the variation within the sampled data is illustrated in Figure 6.5. In the following experiment, an artificial 2-D B-spline transform (3.6) is applied to both images in Figure 6.4 and the segmented labels. The influence of the local mismatch on the evaluation criterion (6.1) is then measured for varying strengths of the deformation field. Figure 6.6 illustrates the manipulation of the control grid, the resulting B-spline transformation within the image domain, the original input image, and the transformed image. Let \mathcal{C}_0 be the control grid that defines an identity B-spline transform and \mathbf{u} the deformation of \mathcal{C}_0 into the grid shown in Figure 6.6a. The

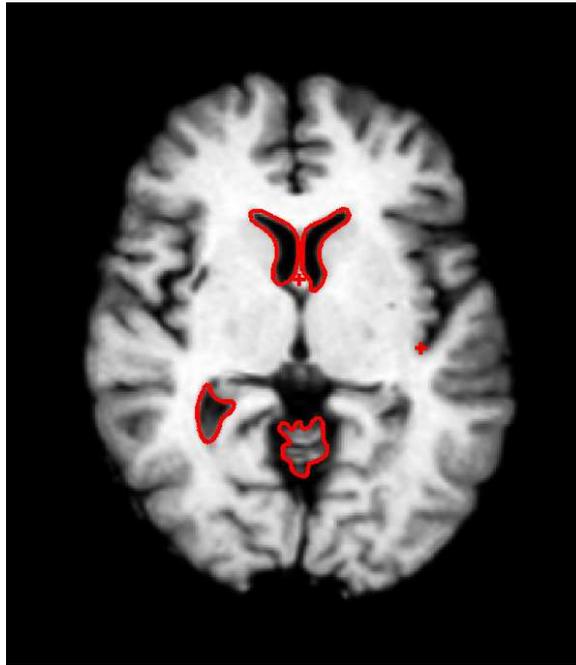


Figure 6.4: Label sets for a slice taken from an MR image of a human brain. The set consists of six labels: three regions for the ventricles, the surface of the cerebellum, the center point of the fornix, and the corner of the insula.

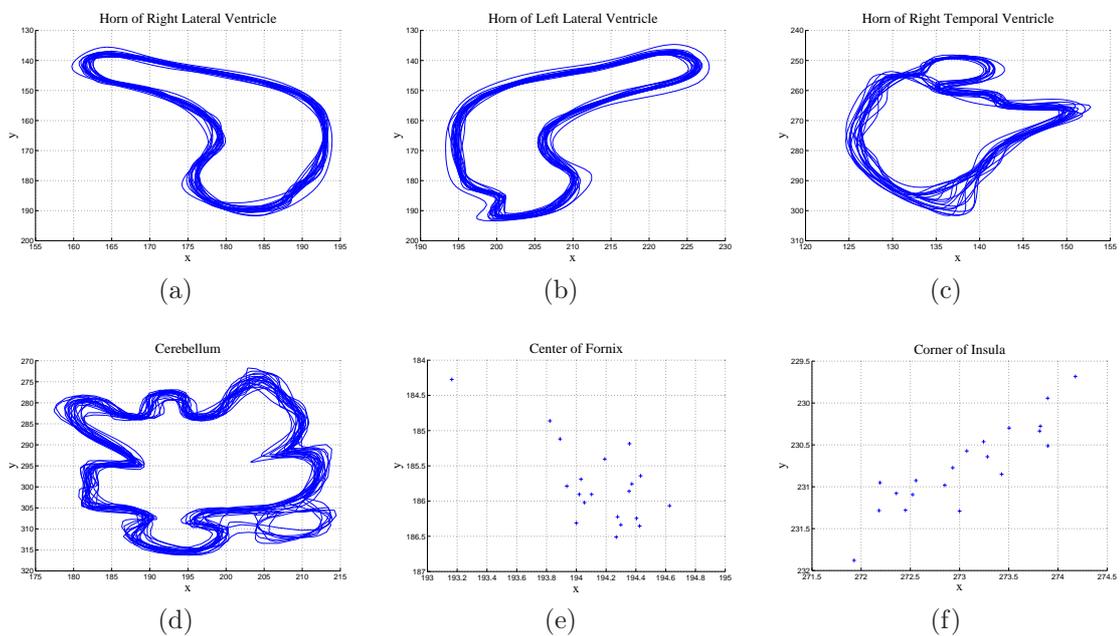


Figure 6.5: Variations in the segmented label data for the MRI brain slice in Figure 6.4. The plots show the manual segmentation results for the horns of (a) the right lateral, (b) the left lateral, and (c) the temporal ventricle, (d) the outline of the cerebellum, (e) the center of the fornix, and (f) the corner of the insula within the slice. The units along the axes correspond to multiples of image pixels.

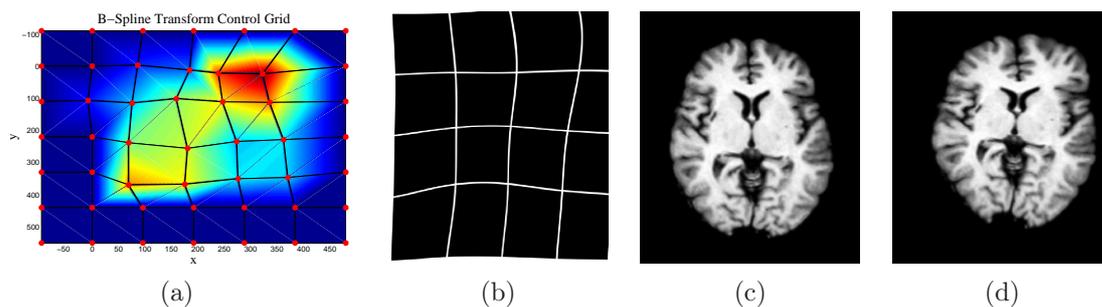


Figure 6.6: Artificial deformation defined by modification of (a) the control grid for a cubic B-spline transform in 2-D. The strength of the control grid deformation is color coded from blue (no deformation) to red (largest deformation). (b) Shows the resulting transformation within the image grid, (c) is the input image, and (d) the transformed image.

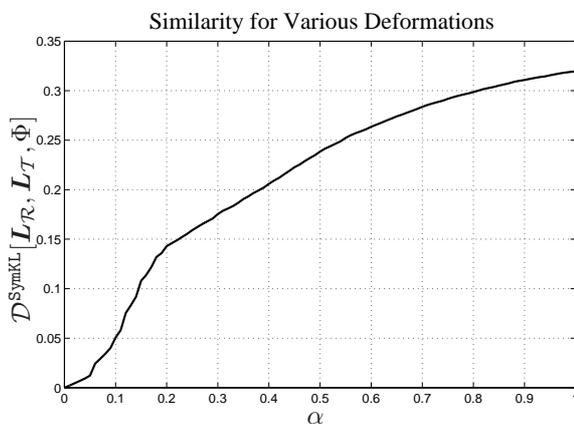


Figure 6.7: The proposed evaluation criterion (6.1) applied to the labels in example 6.4 with varying values for α .

strength of the deformation is then determined by a linear weighting factor $\alpha \in [0, 1]$, which results in the corresponding 2-D B-spline transform:

$$\Phi_{\alpha}^{\text{BS}}(\mathbf{x}) = \mathbf{x} + \sum_{l=0}^3 \sum_{m=0}^3 B_l(u)B_m(v) (C_0(i+l, j+m) + \alpha \mathbf{u}(i+l, j+m)) . \quad (6.12)$$

The values for the distance (6.1) with respect to various strengths of the deformation are plotted in Figure 6.7. The results indicate that the measure (6.1) is suitable to assess the registration accuracy, however, it is necessary to collect a large number of labeled data. In order to keep the segmentation effort for the physicians low, we suggest to use semi-automatic segmentation approaches that can be controlled by user inputs, for example the *random walk image segmentation* [Grad06]. Still, a lot of effort is necessary to establish an evaluation database for various modalities and a number of image pairs that allow statistically relevant comparisons between registration algorithms. The creation of this database is outside the scope of this thesis, therefore, we will leave it to future work.

Part III
Applications

Chapter 7

3-D/3-D Multi-Modal Rigid Registration of the RIRE Database

The statistical methods described in Chapter 4 have been integrated into the NMI distance measure of a state-of-the-art rigid registration application to perform an evaluation, which is based on the retrospective image registration evaluation project (RIRE version 2.0) database of brain images. This database consists of CT, PET, and MR images. The MR data is comprised of T1, T2, and MP-Rage sequences. For some of the MR images, the database also contains corrected versions with respect to scanner-dependent geometry distortions. These images are denoted, in the following, by the term “rectified”. The entire evaluation consists of CT-MR, PET-MR, as well as MR-MR image pairs, which yields 114 registrations in total. West et al. [West 97] and Fitzpatrick et al. [Fitz 98] proposed a gold standard registration based on the detection and measurement of implanted fiducial markers and evaluated the target registration error (TRE) for the transformations. The markers have been erased before distributing the data to conduct a blind study.

Using the RIRE database, we compare the data-driven methods for the automatic parameter selection to standard settings found in literature. The settings for the standard registration are: discretization of the PDFs with equidistantly-spaced histograms, standard deviations for the multi-resolution approach as proposed in Wells et al. [Well 96], Halton quasi-random sampling of 10% of the overlap domain with jittering, and no special handling of structured background noise. In addition to the fully automatic algorithm, the registrations were also manually determined by a medical expert. Only results from automatic approaches were hitherto known for this database. Table 7.1 shows the abbreviations used for the compared methods. Combinations between the techniques are also evaluated, for instance, the proposed NMI approach with all parameters being optimally adapted to the input data is achieved by the combination KCR. The effect of the number of bins for the histograms with respect to the registration accuracy is analyzed as well. For the combinations without requantization (no R), the values 16, 64 and 256 denote the fixed setting for the histogram sizes. When R is enabled, the values are used as initializations in the minimization of the quantization error trade-off, which yields the number of bins. Most of the medical images result in an estimated number of bins in the range of 30 to 60. A minimum requirement of 64 or 256 bins, therefore, already fulfills the

Abb.	Description
S	standard parameter settings
K	automatic Parzen-window kernel width selection
C	coincidence weighting
R	adaptive requantization with number of histogram bins selection
M	manual registration by medical expert

Table 7.1: Abbreviations used for the comparisons between the parameter selection methods described in Chapter 4. K,C, and R are also combined with each other.

trade-off criterion between quantization error and computational efficiency in most of the cases. The resulting values for the number of bins are then 64 or 256, respectively. For a comparison of the number of bins in the R combinations, we suggest to use results from the entries for 16 and 256 bins presented in the following tables and plots. Quasi-random sampling is performed using Halton sequences with a length of 10% of the voxels contained within the overlap domain at the current iteration. A minimum number of 10.000 samples is set for lower resolutions. The jittering and partial volume interpolation are performed for all automatic methods, including S. Those algorithm combinations not containing a specific identifier use the standard settings instead, e.g. the CR method applies the default kernel widths instead of data-driven estimates. The registration algorithm incorporates multi-resolution stages down to a size of 32 voxels along a single image direction. On an Intel Core 2 Duo, 2.6 GHz CPU with 3 GByte of main memory, the average registration time is within a range of 10 to 20 seconds for a single image pair, compared to several minutes needed with larger numbers of samples. It took the medical expert an average (standard deviation) time of 3.5 (± 1.8) minutes for one registration and 6 hours and 37 minutes in total for all image pairs (these times do not include the loading of data or breaks during the registration).

7.1 Significance Tests

A two-tailed, paired t -test at a 5% level of significance was applied to analyze statistical differences between the approaches based on the median TRE values. Tables 7.2 and 7.3 present the TRE values for the conducted CT-MR and PET-MR registrations of the RIRE data.

The first tests comprise a statistical comparison between the proposed, fully automatic NMI approach, i. e. the KCR method with an initialization of 16 bins, and an implementation using standard parameters from literature with 16, 64 and 256 histogram bins. KCR yields a significantly higher accuracy compared to S for 16, 64 and 256 bins. A comparison with the manual registration also results in a significantly higher accuracy in favor of KCR. The level of significance gets slightly worse for tests between KCR with 256 bins and S (6.1%), and also KCR 256 with M (5.5%).

We also examine whether a specific number of histogram bins performs better than another. There is indeed a significant difference within the results achieved by S 16 compared to both S 64 and S 256. Between S 64 and S 256, however, a comparison

MR Modality		Technique									N
		S	K	C	R	KC	KR	CR	KCR	M	
T1	16	1.9	1.7	1.5	1.7	1.6	1.6	1.6	1.6	1.8	16
	64	1.7	1.6	1.6	1.8	1.6	1.6	1.6	1.6	1.8	
	256	1.6	1.5	1.6	1.6	1.6	1.5	1.6	1.6	1.8	
PD	16	2.2	2.1	2.1	2.2	2.3	2.1	2.3	2.4	2.4	11
	64	2.4	2.3	2.4	2.3	2.5	2.4	2.4	2.5	2.4	
	256	2.2	2.2	2.5	2.5	2.5	2.5	2.5	2.6	2.4	
T2	16	2.4	2.2	2.1	2.2	1.9	2.1	2.0	1.9	1.9	15
	64	2.2	2.2	1.9	2.2	1.9	2.2	1.9	1.9	1.9	
	256	2.2	2.3	2.0	2.2	1.9	2.3	2.0	2.0	1.9	
T1 rect.	16	1.8	1.3	0.8	1.4	0.8	1.0	0.8	0.7	1.5	6
	64	1.3	1.2	0.7	1.3	0.8	1.2	0.8	0.7	1.5	
	256	1.1	2.0	0.6	1.1	0.6	1.5	0.6	0.8	1.5	
PD rect.	16	1.2	1.1	0.9	1.1	0.8	1.1	0.8	0.8	1.5	7
	64	1.1	1.1	0.9	1.1	0.8	1.1	0.8	0.9	1.5	
	256	1.1	1.6	0.8	1.2	0.8	1.2	0.8	0.9	1.5	
T2 rect.	16	1.7	1.3	1.1	1.8	1.5	1.2	1.3	1.2	1.5	7
	64	1.6	1.3	1.3	1.5	1.2	1.3	1.2	1.2	1.5	
	256	1.4	1.9	1.2	1.3	1.2	1.9	1.1	1.1	1.5	

Table 7.2: Median TRE values for the evaluation of CT-MR registrations in millimeters. N denotes the number of patients available for each modality combination.

MR Modality		Technique									N
		S	K	C	R	KC	KR	CR	KCR	M	
T1	16	45.5	3.1	3.2	6.0	2.5	3.1	2.9	3.0	3.1	7
	64	3.8	2.9	3.0	3.5	3.2	3.0	3.4	3.3	3.1	
	256	3.9	18.8	3.2	3.8	3.4	7.8	3.7	3.7	3.1	
PD	16	35.6	3.2	2.6	6.1	2.8	2.9	2.8	2.7	3.9	7
	64	3.4	3.2	2.7	3.4	2.7	3.2	2.8	2.6	3.9	
	256	3.5	20.1	2.7	3.6	2.9	5.2	2.9	2.7	3.9	
T2	16	33.4	2.8	3.0	4.3	2.9	2.5	3.0	2.4	2.6	7
	64	3.6	2.8	2.8	3.4	2.7	2.8	2.8	2.8	2.6	
	256	4.1	21.6	2.7	4.5	2.7	6.2	2.5	2.6	2.6	
T1 rect.	16	47.5	2.1	2.7	2.3	2.1	1.9	2.1	2.0	2.9	4
	64	2.0	1.9	1.8	1.9	1.9	2.1	2.0	1.9	2.9	
	256	2.0	3.9	2.2	2.0	2.1	3.7	2.1	2.1	2.9	
PD rect.	16	24.8	2.8	2.5	3.0	2.7	2.5	2.6	2.7	3.3	5
	64	2.5	2.8	2.5	2.8	2.7	2.9	2.7	2.7	3.3	
	256	3.2	19.9	2.5	2.8	2.7	12.5	2.4	2.9	3.3	
T2 rect.	16	35.1	2.7	2.9	3.1	2.7	2.1	2.9	2.7	2.6	5
	64	2.6	2.0	2.6	2.6	2.9	2.1	2.6	2.6	2.6	
	256	2.9	3.9	2.6	2.4	2.7	4.1	2.7	2.6	2.6	

Table 7.3: Median TRE values for the evaluation of PET-MR registrations in millimeters. See notes in Table 7.2.

yields no significant differences. The proposed KCR approach shows no significant influence of the number of bins on the results for a minimum required number of bins of 16, 64, or 256.

7.2 Target Registration Errors

Figures 7.1, 7.2 and 7.3 show the mean TRE values of the various techniques along with their standard deviations for the evaluated modality pairs. For the MP-Rage results, all presented techniques achieve similar accuracies, and a single, outstanding approach cannot be identified. The automatic CT-MP-Rage registrations show a slightly increased accuracy compared to the medical expert, which is contrary to the MP-Rage-T2 pairs. In the plots of Figure 7.2, which correspond to CT-MR

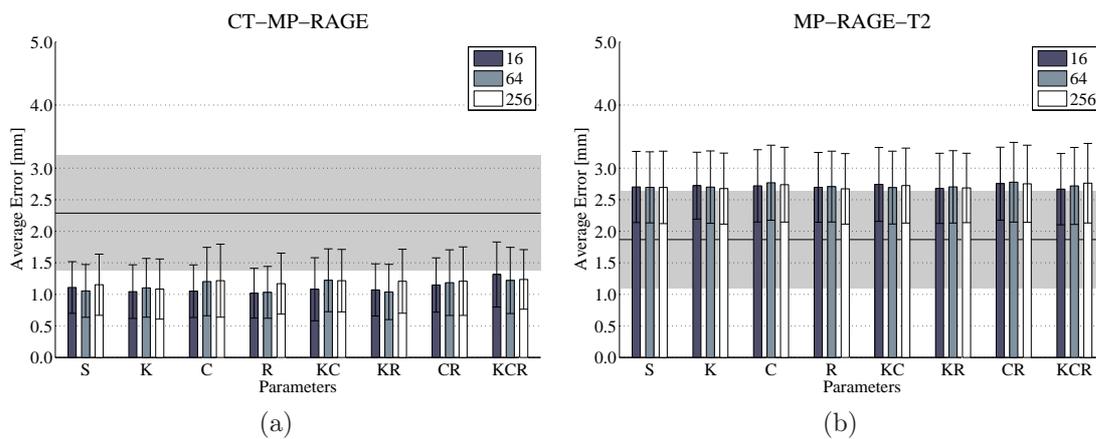


Figure 7.1: Mean and standard deviation of the TRE values for MP-Rage MR sequence image pairs: (a) CT-MP-Rage and (b) MP-Rage-T2 combinations. The vertical bars depict the mean TRE for a specific parameter estimation technique together with the standard deviations (vertical lines). The solid horizontal line indicates the mean TRE of the medical expert, the vertical bar marks the corresponding standard deviation. Results for 16, 64, and 256 number of bins for each parameter estimation technique are presented in order to examine the influence of the histogram size on the registration accuracy.

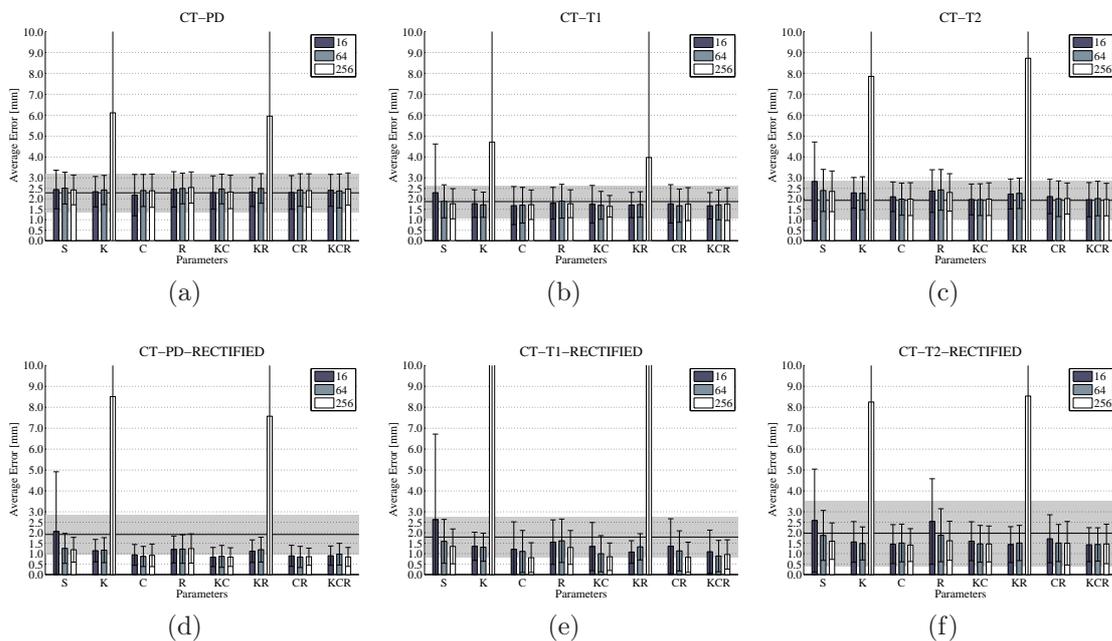


Figure 7.2: Mean and standard deviation of the TRE for CT-MR registrations: (a) CT-PD, (b) CT-T1, (c) CT-T2 and between the distortion corrected MR sequences (d) CT-PD rect., (e) CT-T1 rect. and (f) CT-T2 rect. See also Figure 7.1.

registrations, a distinct improvement is achieved for the rectified MR sequences. In these cases, the automatic registration techniques yield enhancements of the accuracy of up to one millimeter. The NMI implementation with all data-driven parameter

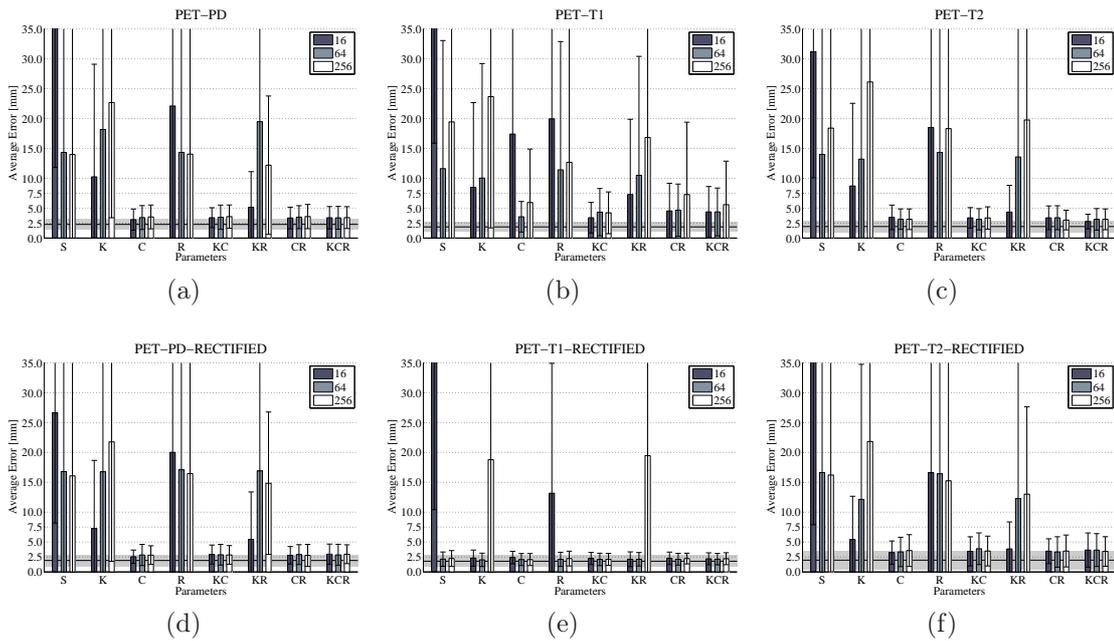


Figure 7.3: Mean and standard deviation of the TRE for PET-MR registrations in the same ordering as in Figure 7.2.

selection methods enabled (KCR) shows an overall good performance for all image pairs. Similar statements can be made for the results of the PET-MR image pairs in Figure 7.3, where the increased TRE values are due to the coarse resolution of the PET data. The geometry correction of the MR images seems to play a more important role for combinations with PET. Again, the proposed KCR technique is performing very well for all PET-MR image pairs. The overall mean TRE measured for all image pairs (CT-MR, PET-MR and MR-MR) is 2.34mm for KCR with 16 bins, compared to 2.54mm for the manual registration, and 6.48mm for the standard implementation with 64 bins.

7.3 Acceptance Rates

Besides the median and mean TRE analyses, we also investigated the overall landmark acceptance rates of the fully automatic KCR approach for the NMI with at least 16 bins, compared to the standard approach with 64 bins, and the manual registration. The acceptance rate for a specific error threshold is computed as the ratio between the number of landmarks with a TRE smaller than the threshold and the total number of landmarks. The values for the acceptance rate are, therefore, in the range of $[0, 1]$. The plots in Figure 7.4 show the increase of the acceptance rates with respect to a decrease of the error threshold. The visual appearance of the curves allows to directly compare the performance of the techniques for various modality pairs. The acceptance rate curves can be regarded as a special form of receiver operating characteristics (ROC), which are often used for the evaluation of pattern recognition systems [Egan 75]. Apart from small landmark error levels in the

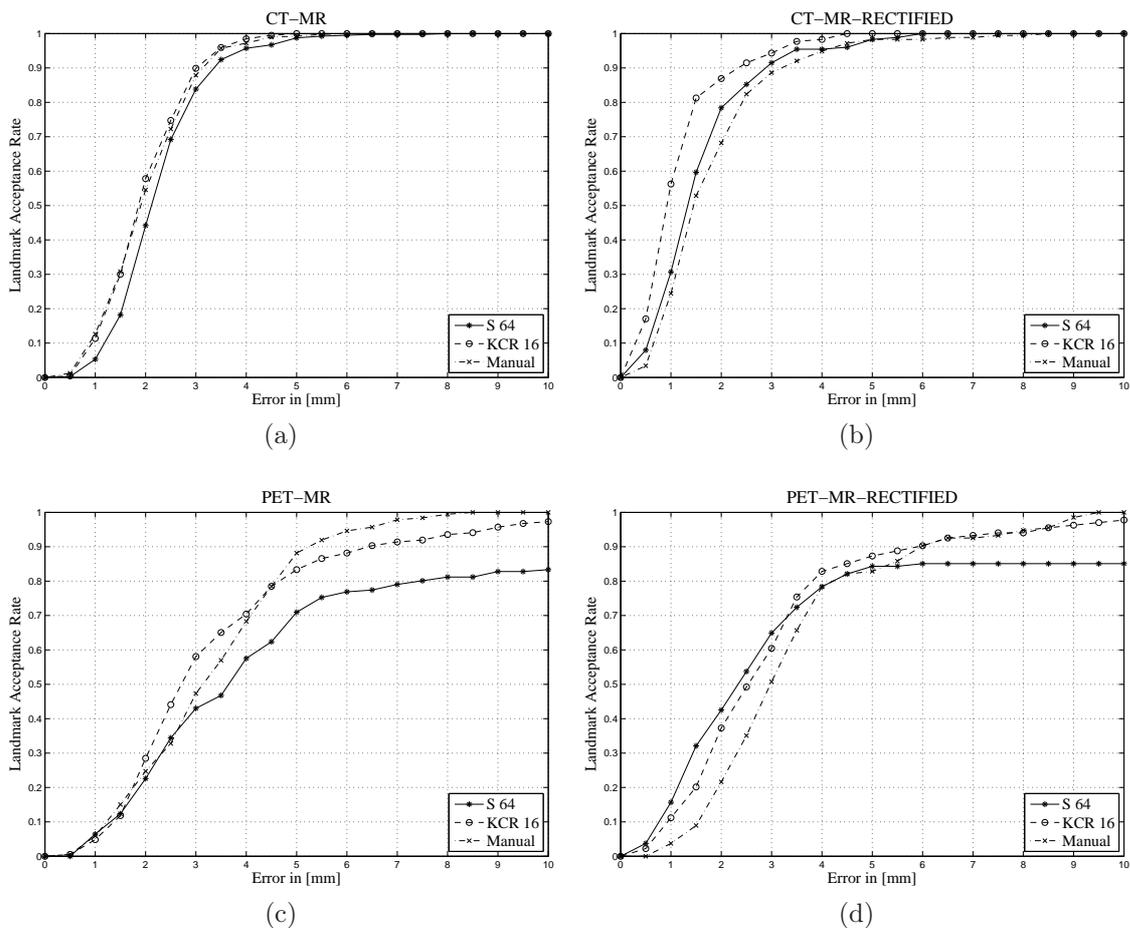


Figure 7.4: Landmark acceptance rates for the modality combinations (a) CT-MR (PD, T1, T2), (b) CT and rectified MR, (c) PET-MR, and (d) PET with rectified MR. The curves are plotted for the registration using the completely data-driven parameter selection for the NMI, a standard parameter set and the medical expert.

PET-MR rectified images, the KCR approach performs better than the standard. It yields acceptance rates of 90% for a TRE of 2.5mm for distortion corrected CT-MR and 6mm for PET-MR combinations. The manual registration by the medical expert achieves higher acceptance rates for the PET combinations with the non-corrected MR images, but for other cases, the automatic registration using the KCR approach performs better.

7.4 Influence of MR Distortion Correction

West et al. [West 97] have statistically analyzed whether a geometry correction, as proposed by Chang and Fitzpatrick [Chan 92], of the PD, T1 and T2 MR sequences yields a better registration accuracy. The geometry correction technique is called “rectification” and is based on the acquisition of two distorted images using spin-echo pulse sequences that are identical except for a simple change in the gradients. Chang and Fitzpatrick showed that a perturbation of the magnetic field can, to some

extent, be manipulated by adjusting the gradient between the two acquisitions. The undistorted intensities of the rectified image can then be calculated from the distorted positions and the ratio between the applied gradients. West et al. found significant differences between the registrations of MR images with and without corrections only for CT-T2 pairs in one out of eleven registration approaches. Other registrations showed minor significances for CT-T2 and CT-T1 pairs at a level of 10%. In contrast, statistical tests on our results confirm that the distortion correction yields a significant improvement. A comparison between the median TRE values with and without correction leads to 95% confidence intervals of [0.76, 0.95]mm for CT-MR images and [0.43, 1.80]mm for PET-MR pairs.

Chapter 8

SPECT Subtraction Imaging for Epilepsy Diagnoses

One of the main goals in epilepsy surgery planning is to localize the region of seizure onset. For this purpose, intra- and inter-ictal SPECT images are acquired between epileptic seizures and closely after a seizure. The comparison of these images is non-trivial due to the low spatial resolution, varying image intensity ranges (uptake values), and different acquisition times. Therefore, conventional side-by-side, visual assessments are extremely difficult. We propose an alternative workflow that has a reduced complexity for the physician compared to commonly applied techniques. Our method introduces digital subtraction techniques, which are already successfully applied in digital subtraction angiography (DSA), into the context of SPECT epilepsy imaging. The presented clinical workflow consists of a series of automatic steps that do not require any user interactions. First, we apply the techniques presented in Chapter 4 to perform a rigid registration based on the intensity statistics. This is necessary since, in the beginning, the intensities may largely differ. Second, the estimated joint PDF after the application of the registration transform is used to establish an intensity mapping between the two images, which we refer to as *normalization*. In the following, we will illustrate that local differences in the blood flow between the two acquisitions lead to a global bias of the rigid registration. The third step, therefore, consists of a non-parametric registration to correct these errors and also possible geometric distortions between the two images. In the context of brain imaging without surgery, using non-parametric registration techniques seems questionable, however, in this case, it is the variation within the cerebral blood flow (CBF) that leads to different focal regions, which directly affects the global optimization in a negative way. Making use of an additional, but heavily regularized, non-parametric registration helps to decrease this bias. The final step of the workflow consists of a subtraction between the processed SPECT images. The resulting difference image directly depicts changes in the CBF between the two acquisitions. A previously acquired MRI is integrated into the workflow and used to spatially localize the differences. We show in a multiple observer study that the additional application of a non-parametric registration leads to an increased accuracy, a higher confidence of the physician and a better quality of the difference image. The resulting information

about the changes within the CBF can, for instance, be used for further diagnostics or surgery planning.

8.1 Related Work

The CBF is known to increase during epileptic seizures at the areas of the seizure onset, which leads to differences between the two images. Subtraction methods were introduced into SPECT imaging by Zubal et al. [Zuba 95] and Spanaki et al. [Span 99]. They applied it after a rigid registration and normalization to analyze the location of seizure onset. O'Brien et al. [OBri 98; OBri 99] additionally incorporated MR images through an image fusion with the difference image to visualize the spatial location of the focal spots. In order to identify statistically relevant image differences, both Chang et al. [Chan 02] and McNally et al. [McNa 05] used techniques of statistical parametric mapping for a comparison of the subtraction values with a collection of normal subjects in order to identify those variations that are due to the epileptic disease. Koo et al. [Koo 03] have shown in a preliminary study that the retrospective subtraction of inter- from intra-ictal SPECT images at a visualization window of 75% to 100% shows good concordance with the seizure foci determined by other, well established techniques.

Varying tracer concentrations between the image acquisitions lead to different uptake values that have to be normalized before subtraction. Otherwise, there is a systematic error in the intensities of the difference image. Several approaches for this normalization are described in literature. Chang et al. [Chan 02], for instance, fit a linear mapping model into the joint histogram of both SPECT images after the registration such that the entropy of the difference image is minimized. Other methods include normalizations to the maximal image values, to the mean uptake value in the entire brain, or to selected regions. The normalization in the present work approximates a linear model as proposed also by Liao et al. [Liao 03].

Regarding the registrations, we apply the techniques described within the Sections 3.1, 3.2, and Chapter 4. As the SPECT images usually contain a lot of low-intensity, structured noise within the background region, we utilize the KCR parameter estimation for NMI, which has been shown to work for the RIRE data in Chapter 7. The approach of comparing inter with intra-ictal SPECT images is very similar to techniques used in DSA. There, an additional non-parametric registration is incorporated in order to get rid of artifacts within the difference image that arise from patient movements. An overview of related techniques to account for such motion artifacts can be found, for instance, in Meijering et al. [Meij 99]. For the presented approach, the non-parametric registration is applied to compensate for the bias in the rigid registration. In the following, we make use of a non-parametric registration with a curvature regularization energy term constraining the deformation.

8.2 Clinical Workflow

The following section describes the steps of the workflow in a chronologic order. Both the rigid and the non-parametric registrations are important steps – we refer to the

previous Chapters 3 and 4 for the details. The intensity normalization is necessary for the comparison between the intensity values of the two SPECT images. If an MRI of the patient exists, it is incorporated by a registration with the reference SPECT image. A registration of the difference image directly with the MRI is an ill-posed problem, because there is insufficient anatomical information left within the differences that could be adequately aligned. Instead, the registration transform determined for the SPECT-MR pair is applied to the subtraction image, which is located within the same coordinate system of the reference SPECT. For this task, a rigid registration is used.

8.2.1 Rigid Registration

The computation of the normalization between the inter- and intra-ictal SPECT data requires that the images are suitably aligned. To correct for different patient positions between the acquisitions, a rigid, intensity-based registration between the SPECT images is performed. We choose as objective function for this problem the NMI distance measure between the two images with respect to a rigid transform (i. e. only rotations and translations). In Figure 8.1, the rigid position and orientation

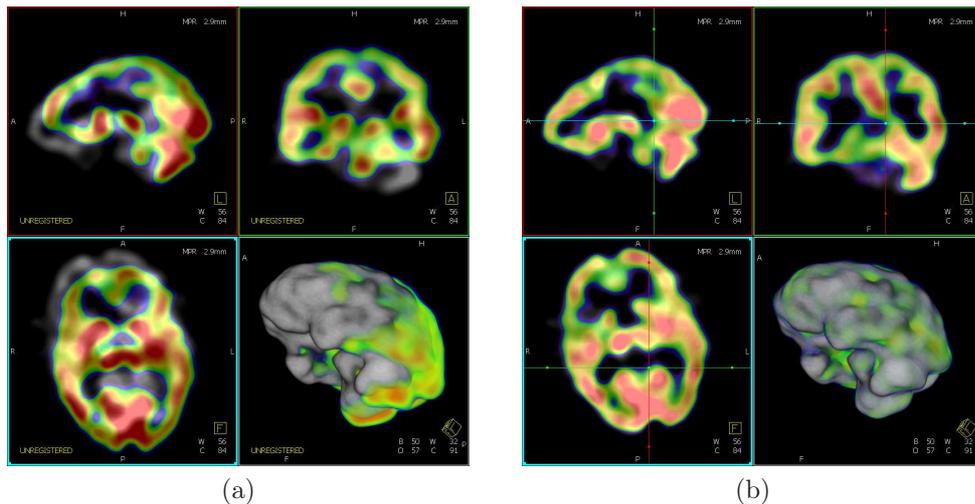


Figure 8.1: Input SPECT image dataset (a) before, and (b) after the rigid registration.

alignments are computed for an example image pair. Here, the initial misalignment, which is compensated by the rigid registration, is mainly due to a different placement of the head between the two acquisitions.

8.2.2 Intensity Normalization

As mentioned above, for a correct interpretation of the differences between the SPECT images, the intensities have to be normalized to a common intensity range. This is necessary due to different acquisition times and changes in the overall tracer uptake within the human body. We model the mapping by an affine intensity transform similar to the proposals of Liao et al. [Liao03]. In order to be invariant to

the intensities of the background, and also the structured noise contained within, we restrict the affine mapping to the region of the joint PDF above the background thresholds, i. e. the probabilities for joint intensities that both belong to brain tissue. The thresholds are determined using the binary quantization approach described in Section 4.3.4. The components of the affine model are then computed by linear regression within the joint PDF without the background intensities. This yields an affine intensity mapping that transforms the principal axis in the joint PDF region onto the diagonal, as the entries of a discrete PDF are located on the diagonal for perfectly aligned and normalized images. An example for this normalization can be seen in Figure 8.2. After the application of the affine mapping, the entries within the

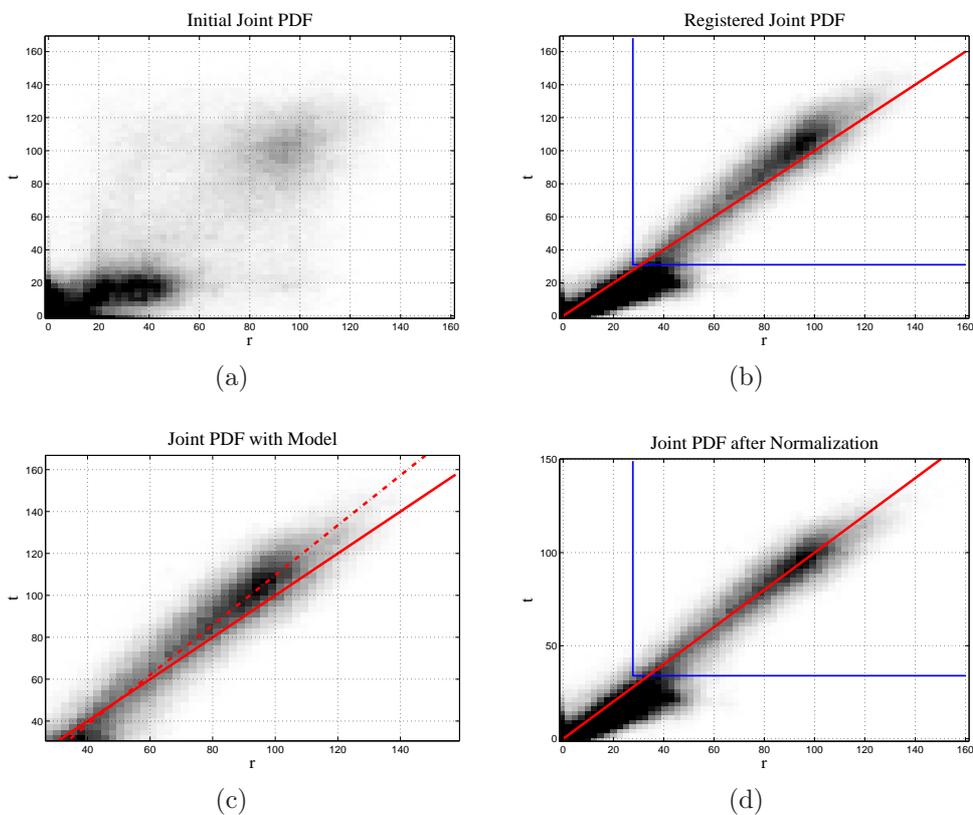


Figure 8.2: The figures show the joint PDF of two input SPECT images (a) before, (b) after the registration, (c) with the model fitted into the joint PDF without background content, and (d) the joint PDF after applying the intensity normalization. The figures also show the diagonal through the histogram space (solid red line), the fitted affine model (dashed red line), and the threshold region for the background (solid blue lines).

joint PDF are located closer to the diagonal, which indicates a better correspondence between the intensities. Figure 8.3 shows an example image pair before and after the normalization. The same transfer function (window level settings) has been used for the visualization in this example.

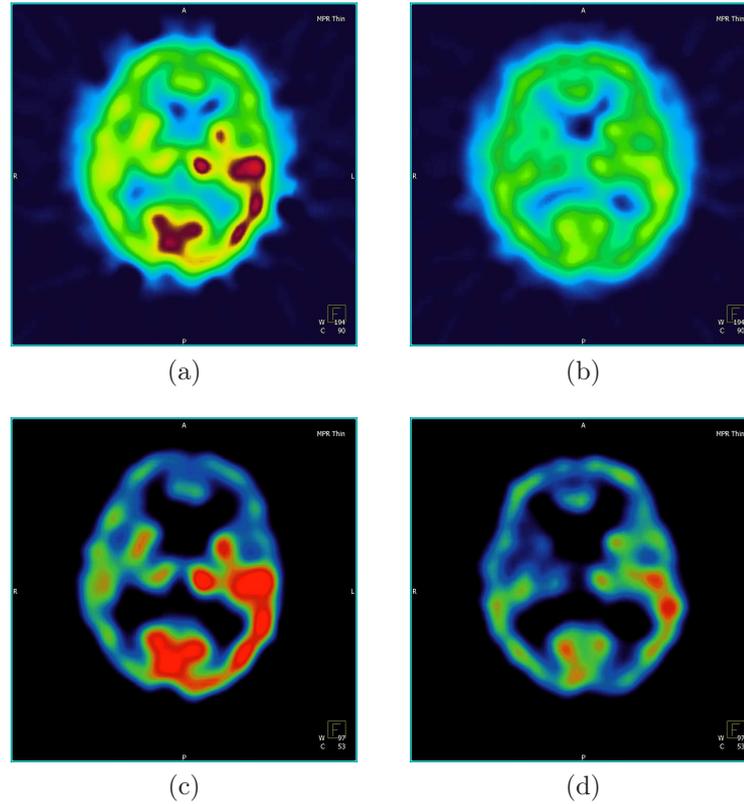


Figure 8.3: An example result of the proposed SPECT normalization approach. The images show (a) the intra-, (b) the inter-ictal SPECT images before the normalization, and (c), (d) the corresponding images afterwards.

8.2.3 Non-Parametric Registration

The rigid registration results in a transform that can only be optimal in a global sense, similar to a least squares solution. We found that if there is a large amount of variation in the CBF across the data sets at different locations, the intensity similarity measure is biased. This tends to impair the registration accuracy at lower contrast image regions, for instance at the boundaries of the brain. If such a biased registration result is used as input for the subtraction stage, the difference image falsely contains information at the misaligned regions, which may lead to wrong diagnoses. Figure 8.4 shows a 2-D registration example that demonstrates the effect of bias on the difference image. Both the reference and the template image show the same slice of an arbitrarily chosen SPECT image. In order to demonstrate the impact of a local intensity variation on the registration result, an artificial focal spot is added to each image. The spots have the same extent, but differ in their location. After a rigid 2-D registration of the images and the examination of the resulting difference image shown in Figure 8.4c, the artifacts at the boundaries of the brain outline are clearly visible. The errors at these regions are due to the bias, which is introduced by the local motion of the artificial focal spot and affected the result of the rigid registration. The non-parametric registration is less impaired by the local variation, as can be seen in Figure 8.4d, although the deformation is restricted by a high weighting factor for the regularization energy term. There is still a local

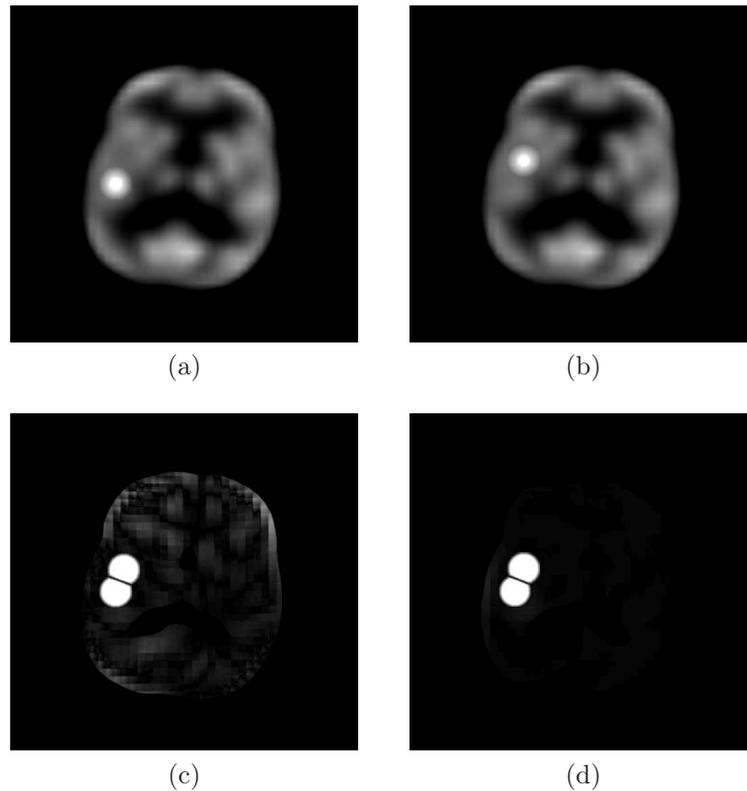


Figure 8.4: Illustration of the bias of a rigid registration introduced by local variations in the image content. The top row consists of (a) the reference and (b) the template image, which only differ by the location of an artificially introduced focal spot. The bottom row shows (c) the difference image based on the results of a rigid registration, and (d) of a non-parametric registration. The same visualization settings have been applied for both subtraction images.

difference at the artificial spot, which is exactly what we want to retain, but the artifacts at the remaining image regions are substantially smaller.

Consequently, the rigid registration can only result in a global fit of the images, and variations in the CBF lead to local SPECT signal changes that impair the accuracy of result. This can be compensated by a subsequent non-parametric registration, where local changes have only a limited effect on the global registration transform. Without the non-parametric registration step, slight rigid mis-registrations, due to the aforementioned bias, show up in the difference image and might be falsely regarded as lesions. Of course, the proposed method is not the only way to handle this problem, however, we found that it works very well in practice. The visual quality of the subtraction images, as a result from the described workflow, is an important factor for the confidence of the physician to localize the focal spots.

8.2.4 Visualization

The spatial transformation, which is composed of the rigid transform and the deformation field, is then incorporated into the subtraction of the normalized images. The intensity differences can be directly interpreted as changes of the blood flow activity

between the acquisitions, however, the spatial information is lost. The missing spatial information can, for instance, be contributed by an MRI of the patient. The direct registration of the difference image with the MRI is an ill-posed problem. Instead, we make use of the fact that the difference image is contained within the same coordinate system as the two registered SPECT images. The fusion of the MRI with the difference image can, hence, be realized indirectly by a registration of the MR with the intra-ictal SPECT image. This is performed, again, by calculating a rigid transform that maximizes the NMI between the MRI and the SPECT. The resulting transform can finally be used to fuse the difference image with the MRI, which yields additional information for the physician.

8.3 Results

The proposed method has been applied to a collective of 26 epilepsy patients and assessed by physicians. Each patient went through the standard diagnostic procedures. The locations and the number of the focal spots are known from the patient charts. The images have been anonymized within the clinics and handed to us afterwards. The subtraction images were generated with two versions of the proposed workflow: the first version consisted of the workflow described above. In the second version, the non-parametric registration was disabled. The subtraction images were randomly ordered into a set of evaluation protocols with a maximal distance between the two workflow results from a single patient. The physicians then had to specify the location and the number of the focal spots based on a fusion of the subtraction image with the corresponding MRI in a blind study. Figure 8.5 shows the implemented user interface for the evaluation. The participants are distinguished by name and password in order to avoid a mixture between inputs from different participants. During the evaluation, the intermediate results are stored on the hard disk. The participants can specify the number and locations of the focal spots for each evaluation sample, together with an indication of the image quality and their certainty. The results have been collected from two physicians from the clinics of nuclear medicine of the University Erlangen.

Objective for Comparison	Rigid	Non-Parametric
Correct Number of Foci	41	41
Correct Location	65	64
Certainty	38	50
Correct Number of Foci (<i>certain</i> only)	58	54
Correct Location (<i>certain</i> only)	80	81
Observer Reliability	84	92
Quality	1.12	1.33

Table 8.1: Comparison between the evaluation results of the two workflow versions.

The results of the blind evaluation are presented in Table 8.1. The first column specifies the objective for the comparison, the second column shows the results of the workflow that contains only the rigid registration. In the third column, the values correspond to the complete workflow with both the rigid and the non-parametric reg-

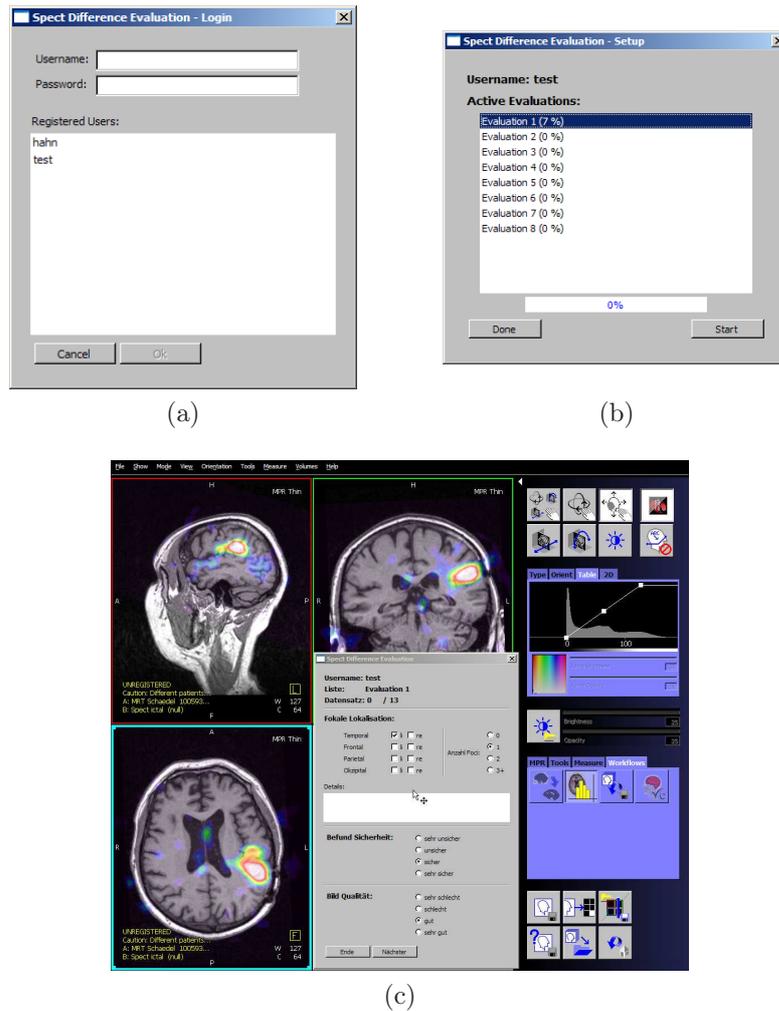


Figure 8.5: Illustration of the several steps during the evaluation of the SPECT difference images: (a) the user login screen in order to distinguish between the participants, (b) the progress screen for each participant, and (c) the actual evaluation screen with the fusion between the anonymized MRI and the SPECT difference image.

istration. The values relate to the means of the evaluations and, except for the quality statement, are scaled to percentages. On an ordinal scale, the certainty had to be classified as *very uncertain*, *uncertain*, *certain*, or *very certain*. The third comparison objective measures the percentage of samples that have been classified as *certain* or higher. The effects of the two versions of the workflow on the resulting difference images are shown for an example dataset in Figure 8.6. From the visualization of the difference images at the same significance level, it can be seen that the result using the additional non-parametric registration step contains an overall lower noise level and the contours of the lesion are more focused. In Figure 8.7, the fusion visualization for two datasets within the study collective are presented. The image 8.7a shows a lesion within the left parietal lobe that can be located in both the difference images and the MRI. For the image 8.7b, the MRI did not indicate any lesions, however, the focal spot in the right temporal lobe, which can be well recognized within the difference image, correlates also with the EEG measurements and visual side-by-side

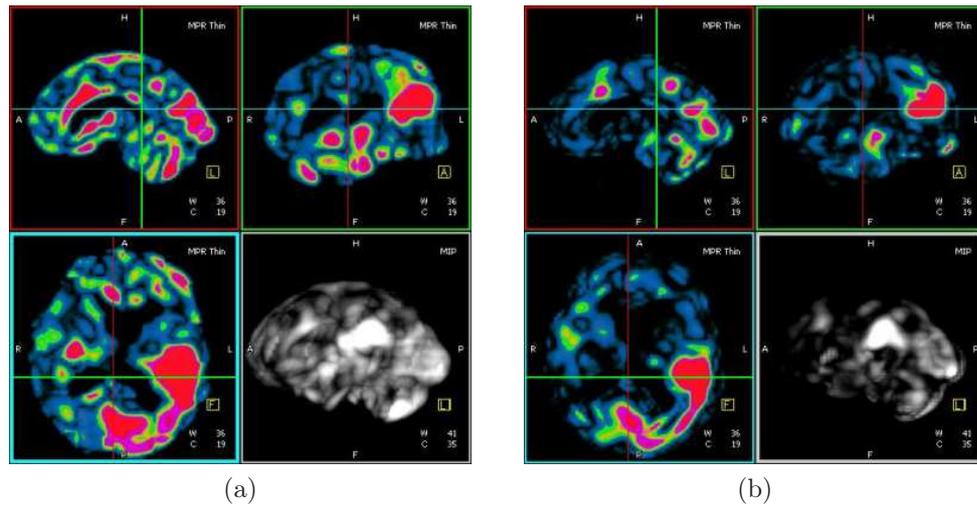


Figure 8.6: A comparison between the results of the two workflow versions for an example dataset. The renderings show (a) the difference image using only a rigid registration, and (b) with an additional non-parametric registration using the same transfer function settings.

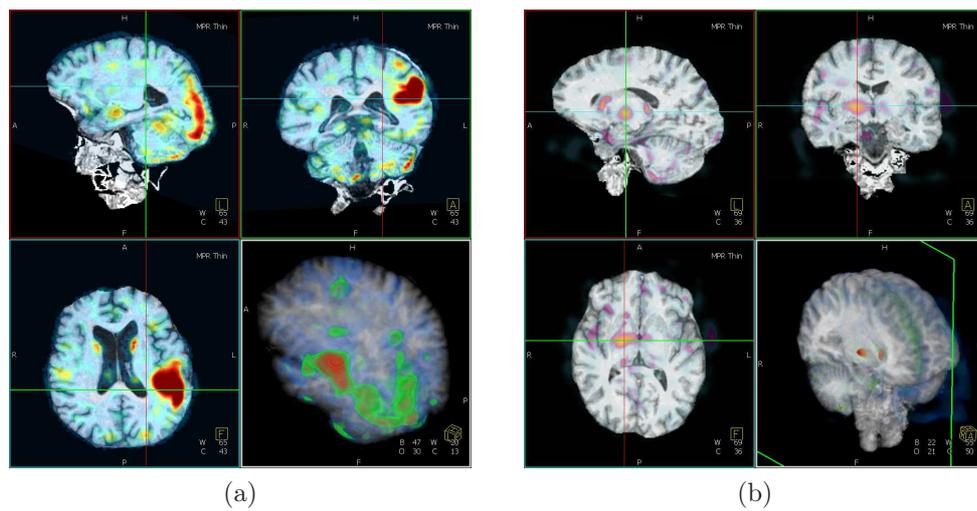


Figure 8.7: Fusion visualization between an MRI and the subtraction result of the proposed workflow for two epilepsy patients.

assessment of the SPECT data. Two final examples for the fusion of the subtraction workflow results with the corresponding MR images are presented in Figure 8.8.

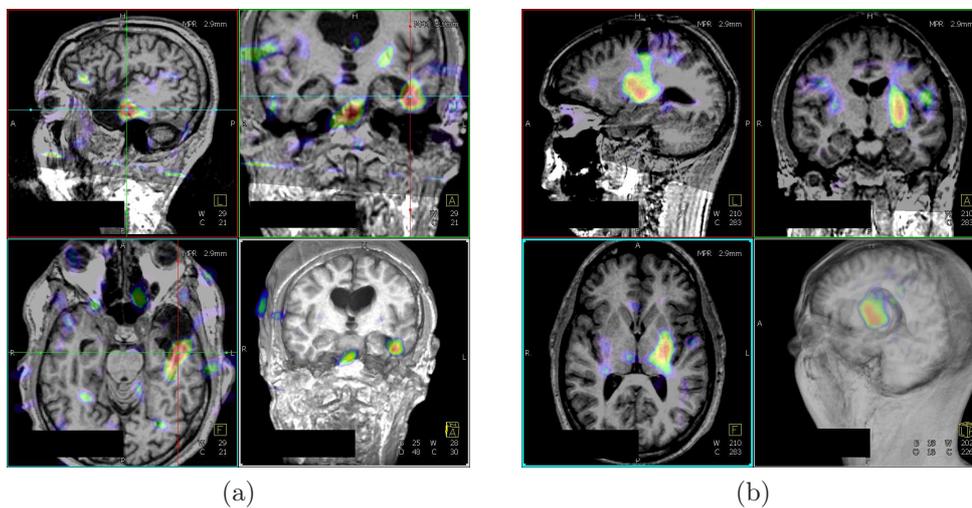


Figure 8.8: Two example patients taken from the collective. The images show the SPECT subtraction image fused with the corresponding MRI.

Chapter 9

Registration Applied to Shape-Based Segmentation

The segmentation of medical images is an important preprocessing step for further medical analyses or diagnoses. Many different methods are used for the classification of anatomical structures of interest, for example region growing, level set, or active shape model (ASM) segmentations. A major problem for most of the segmentation algorithms are structures that are not clearly delineated from the surrounding tissue. An unconstrained segmentation usually leaks into other structures within these regions. Introducing prior knowledge into the segmentation approach helps to increase the robustness of a segmentation and to avoid leakage in cases of noisy or missing data. The ASM technique is an example for a segmentation algorithm that makes use of statistical information about shape variations, which are extracted from a set of training data. The crucial aspect of an ASM is the correct determination of the correspondences between the training data, as errors during the model generation phase lead to wrong statistics about the shape and, in consequence, to wrong results.

A state-of-the-art solution consists of the optimization of a minimum description length (MDL) measure between a set of points (landmarks) placed on a parametric surface onto which all training shapes are mapped. The training shapes are represented as surface meshes with potentially different numbers of vertices. Optimizing the MDL is highly complex and may be very time consuming, even on newest hardware. In addition, updating the model with new training data requires rerunning the procedure all over again. Finding a suitable parametric space for the mapping may also pose a problem for some applications, for example if the transformation is complex, or if the object largely deviate from the shape of the target space.

From another perspective, the correspondence problem for the model generation may be regarded as a registration task: depending on the DOF in the spatial transform, the alignment between the structures within the training data can be determined by a registration algorithm. Using a combination of parametric and non-parametric registration algorithms provides, therefore, an alternative to the MDL technique in order to address the correspondence problem. In the following chapter, we concentrate on solving the correspondence problem with a non-parametric registration, propose a suitable distance measure for this specific registration, and compare the novel approach with an already established MDL method. The results

are presented for the clinical application of kidney segmentation from CT images to gain information about their extent or volume, which is of interest for nephrologists.

9.1 Model Generation

In 1995, Cootes [Coot 95] introduced the ASMs to regularize a segmentation with prior knowledge of the statistical variation of the target shape. This additional information allows to achieve robust results even if the signal to noise ratio is low or the delineation of the organ boundaries is distorted. For 3-D applications, a common approach to represent the training shapes is to use surface meshes composed of triangles. The necessary steps to generate an ASM model are shown in Figure 9.1. Each ASM is

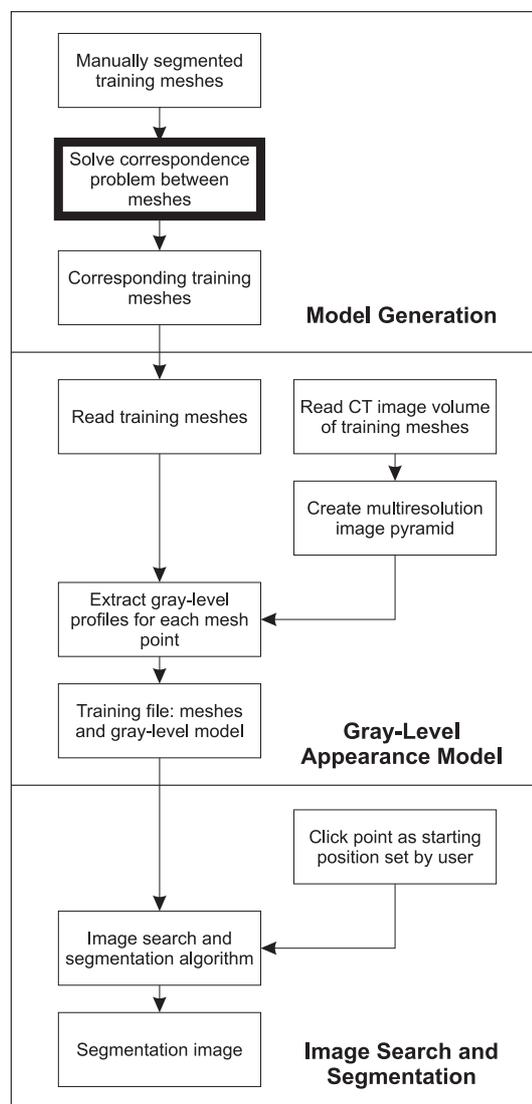


Figure 9.1: The components of a state-of-the-art ASM segmentation system. The computation of correct correspondences between the training shapes is crucial for the entire approach.

based on a point distribution model that is created from a set of N training shapes, each composed of n sampled surface points. For a segmentation problem in 3-D, let $\mathbf{x}_i^j \in \mathbb{R}^3$, with $i = 1, \dots, n$ be the vertex coordinates on the j -th surface. The shape $\mathbf{s}_j \in \mathbb{R}^{3n}$ is composed of these vertices by rearranging the coordinate components (x, y, z) :

$$\mathbf{s}_j = \underbrace{(x_1^j, \dots, x_n^j)}_{x\text{-components}}, \underbrace{(y_1^j, \dots, y_n^j)}_{y\text{-components}}, \underbrace{(z_1^j, \dots, z_n^j)}_{z\text{-components}})^T. \quad (9.1)$$

The N training samples are used to extract the mean shape $\bar{\mathbf{s}}$ and the principal modes of variation by a principal component analysis (PCA):

$$\begin{aligned} \bar{\mathbf{s}} &= \frac{1}{N} \sum_{j=1}^N \mathbf{s}_j \\ \mathbf{S} &= ((\mathbf{s}_1 - \bar{\mathbf{s}}), \dots, (\mathbf{s}_N - \bar{\mathbf{s}})) \\ \mathbf{S}\mathbf{S}^T \mathbf{v}_i &= \varphi_i \mathbf{v}_i \quad , \text{ with } i = 1, \dots, T \\ \text{s. t. } \|\mathbf{v}_i\| &= 1 \quad , \end{aligned} \quad (9.2)$$

where $\mathbf{S}\mathbf{S}^T$ denotes the covariance matrix of the training shapes that is decomposed into the eigenvectors \mathbf{v}_i belonging to the T -largest eigenvalues φ_i . The decomposition is performed as described in Murase and Lindenbaum [Mura 95].

The mutually orthogonal eigenvectors are sorted in descending order of their respective eigenvalues and describe the principal modes of variation from the mean shape. Thus, a linear combination of the eigenvectors with the mean shape spans the subset of shapes that are composed of the principal modes of variation. Each shape within this space can be expressed by:

$$\mathbf{s} = \bar{\mathbf{s}} + \sum_{i=1}^T \alpha_i \mathbf{v}_i \quad , \quad (9.3)$$

with α_i being the weighting factor for the corresponding i -th variation.

The approach heavily depends on the correspondences between the training shapes that are incorporated in (9.2). Difficulties arise especially for complex shapes in higher dimensions. Unfortunately, the surface representations using triangular meshes, in general, is not uniquely determined for all training shapes. Therefore, the sample meshes may differ in the cell structure and the number of vertices. This makes it difficult to select surface points that occur only in the joint subset of the training shapes. If the information about the location of these points is not given by construction, a correspondence problem has to be solved prior to the application of the PCA. Davies et al. [Davi02a] described an automatic method for the construction of optimal 3-D shape models based on statistics. The authors proposed to estimate an optimal parameterization of each surface mesh with respect to an MDL criterion [Davi02b]. An extension for dense correspondences, which are optimally distributed across the set of training shapes, was proposed by Heimann et al. [Heim 05; Heim 06]. They included the distribution of the vertex points within the mapped parametric space into the objective function.

In the following, we propose an alternative to MDL optimization in order to solve the correspondence problem by utilizing a non-parametric image registration algorithm, which is described in Section 3.2.

9.2 Registered Point Correspondences

The ASM model generation with respect to triangle surface meshes requires the computation of correspondences between the points on the surfaces. In contrast, we propose to shift the problem into the discrete image domain using segmentation images of the training shapes, which enables intensity-based image registration techniques. The registration has to align the shape images from different patients, which requires a spatial transform with sufficient DOF. Therefore, we propose to incorporate a non-parametric image registration into the ASM generation to compute dense deformation fields between the training shape images. The deformation is constrained by a curvature regularization energy. In addition, the representation of the segmented shapes within the discrete image domain enables the use of intensity-based similarity measures. The techniques introduced in Sections 3.2 and 3.3 can be directly applied to this application. A common reference shape is selected from the training set and registered with the remaining samples. The surface mesh of this reference shape can then be mapped through the resulting spatial deformations. This guarantees a one-to-one match between all points on the surface meshes.

As the basis for the registration, we use an implicit representation of the i -th training shape by a bi-valued segmentation image Γ_i . It is defined within the discrete image domain Ω_i , which is specified by the spatial sampling properties of the segmented image:

$$\Gamma_i(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \Omega_i \text{ inside } \mathbf{s}_i \\ 0, & \text{otherwise .} \end{cases} \quad (9.4)$$

Thus, Γ_i represents the corresponding surface mesh \mathbf{s}_i within the discrete image domain. The point correspondence problem between two shapes $\mathbf{s}_i, \mathbf{s}_j, i \neq j$, can now be formulated as a registration between Γ_i and Γ_j . It optimizes the spatial transform that maps corresponding structures within the segmentation images onto each other. The resulting dense deformation field $\mathbf{u}_{i,j}$ is the spatial transform between Ω_i and Ω_j , and provides a solution for the point correspondence problem:

$$\mathbf{x}_k - \mathbf{u}_{i,j}(\mathbf{x}_k) = \mathbf{x}_l \quad \text{with } \mathbf{x}_k \in \mathbf{s}_i, \mathbf{x}_l \in \mathbf{s}_j . \quad (9.5)$$

It is worth mentioning that this formulation of the problem within the image domain does not require an explicit representation of the surfaces. It can be applied to both implicit segmentation images and explicit surface meshes, because both representations can be converted into each other. In comparison, the MDL-based methods rely on an explicit surface representation, which is a problem for shapes that are composed of several elements (e.g. as a result of an implicit level set segmentation). The training data is acquired by a supervised segmentation of the images and stored either as binary images or surface meshes. Attention has to be paid to retain substantial structural information in the discretization. Usually, the resolution of the input images is sufficient to be used for the shape images alike.

As described above, the optimization problem (3.14) is composed of a curvature regularization (3.17) on the deformation field and an appropriate similarity measure. Regarding the point correspondence problem, the measure has to operate on binary shape images. Points on the surface of the shapes have to be registered correctly and intrinsic properties, such as the curvature, have to be retained between corresponding

surface regions as closely as possible. As all shape images share the same intensities, it is reasonable to use the SSD (3.23) as a suitable distance measure for this task. A drawback of this measure in the context of finding shape correspondences, however, is that it does not account for surface properties. Therefore, we propose an extension of this similarity that incorporates the curvature κ_i of the i -th segmentation, which is also used in the context of level set segmentations, for instance, in Sethian [Seth99]. The curvature is based on the signed Euclidean distance transform $d_{\text{SE}}(\Gamma_i, \mathbf{x})$ of the shape image Γ_i at position \mathbf{x} :

$$\kappa_i(\mathbf{x}) = -\nabla \cdot \frac{\nabla d_{\text{SE}}(\Gamma_i, \mathbf{x})}{|\nabla d_{\text{SE}}(\Gamma_i, \mathbf{x})|}, \quad \mathbf{x} \in \Omega_i. \quad (9.6)$$

A detailed description of d_{SE} can be found in Danielsson [Dani80] or Ye [Ye88]. The distance-transformed shape image contains the signed Euclidean distance to the closest point on the shape boundary. The curvature-extended sum of squared differences (CSSD) similarity measure $\mathcal{D}^{\text{CSSD}}$ between the shape images Γ_i and Γ_j is then formulated as a convex combination of the SSD and a squared distance \mathcal{D}^κ between the surface curvature values:

$$\mathcal{D}^\kappa[\Gamma_i, \Gamma_j, \mathbf{u}_{i,j}] = \frac{1}{|\Omega(\Gamma_i, \Gamma_j)|} \int_{\Omega(\Gamma_i, \Gamma_j)} (\kappa_i(\mathbf{x}) - \kappa_j(\mathbf{x} - \mathbf{u}_{i,j}(\mathbf{x})))^2 d\mathbf{x} \quad (9.7)$$

$$\mathcal{D}^{\text{CSSD}}[\Gamma_i, \Gamma_j, \mathbf{u}_{i,j}] = (1 - \beta)\mathcal{D}^{\text{SSD}}[\Gamma_i, \Gamma_j, \mathbf{u}_{i,j}] + \beta\mathcal{D}^\kappa[\Gamma_i, \Gamma_j, \mathbf{u}_{i,j}],$$

where $\beta \in [0, 1]$ is a weighting factor between the two distances. A value of $\beta = 0.5$ is used for the experiments in the following. Figure 9.2 illustrates the difference between the resulting deformation for \mathcal{D}^{SSD} and \mathcal{D}^κ for an example registration between two simple shapes. The curvature distance results in deformation fields that align regions

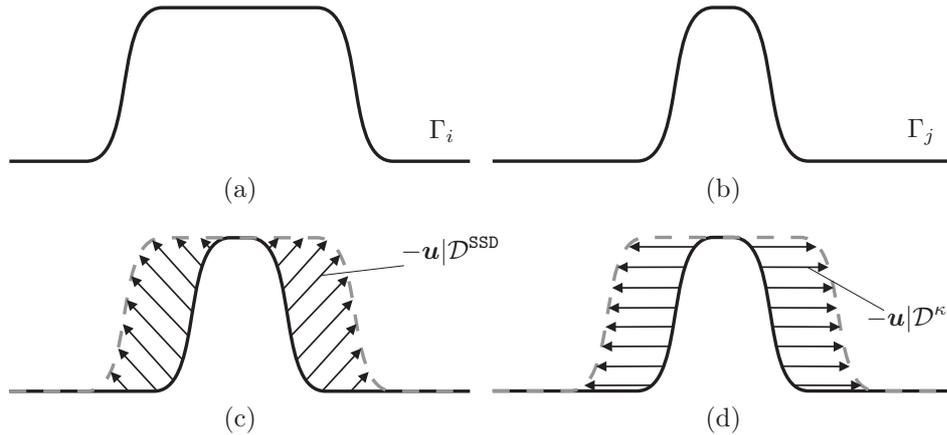


Figure 9.2: Differences between deformation fields resulting from similarity measures \mathcal{D}^{SSD} and \mathcal{D}^κ . The illustrations show examples for (a) a reference, (b) a template shape, (c) the deformation with respect to an optimization of \mathcal{D}^{SSD} , and (d) the deformation for \mathcal{D}^κ .

with similar curvature. In contrast, optimizing only the SSD distance yields a least squares deformation field with respect to the regularization energy. In consequence,

the optimal transform based on the standard SSD approach might just smooth out a bulge in the surface, while the extended distance (9.7) tries to match it with a corresponding bulge in its local neighborhood, first.

In the implementation, the curvature (9.6) is calculated on the original, undeformed shape images and interpolated during the iterative refinement of the registration transform. As it contains second order derivatives, the curvature (9.6) is usually very sensitive to noise in the images. In this case, however, it is computed on the distance transforms of the binary shape images. Noise in the original images has, therefore, only effects on the boundary of the segmentation results. Nonetheless, we propose to use derivatives of low pass filtering kernels, e.g. a Gaussian, with kernel widths chosen with respect to the spatial resolution of the shape images.

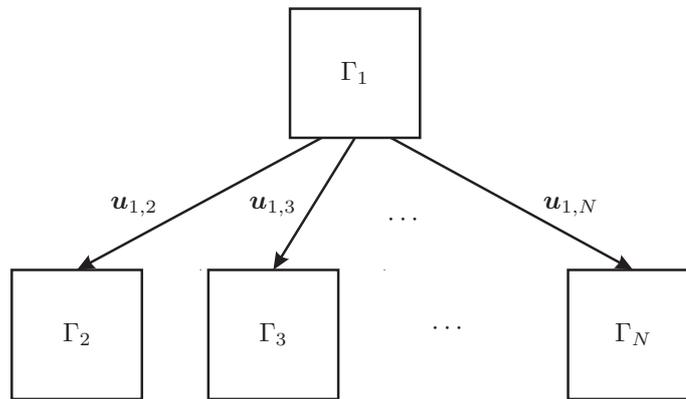


Figure 9.3: Directions of the registrations for the N shape images used for the training of the ASM. The arrows point from the reference to the template image. The resulting deformation fields are used to map between the coordinate system of Γ_1 into the coordinate systems of all other training shapes.

The registration scheme for the registration-based approach to solve the correspondence problem between the training shapes is shown in Figure 9.3. The reference shape image Γ_1 is registered with the remaining training samples $\Gamma_{[2,\dots,N]}$ such that the resulting deformation fields $\mathbf{u}_{1,[2,\dots,N]}$ can be used to transform a spatial coordinate $\mathbf{x} \in \Omega_1$ into all other coordinate systems. In the aforementioned equation (9.2) for the determination of the shape variance by a PCA, it is required that the shapes composing \mathbf{S} contain corresponding spatial points in the correct ordering. The registration results can now be incorporated into the shape vectors to fulfill this requirement:

$$\begin{aligned} \mathbf{s}_j &= \mathbf{s}_1 \circ \mathbf{u}_{1,j} \\ \mathbf{x}_i^j &= \mathbf{x}_i^1 - \mathbf{u}_{1,j}(\mathbf{x}_i^1) . \end{aligned} \quad (9.8)$$

In (9.8), the surface mesh \mathbf{s}_1 , which is defined by the points $\mathbf{x}_i^1, i = 1, \dots, n$, is mapped through the deformation field $\mathbf{u}_{1,j}$ to compute the corresponding points \mathbf{x}_i^j of the j -th training shape. This mapping is indicated by the operator “ \circ ”. It implies an interpolation within the deformation field, because the vertices of \mathbf{s}_1 are not necessarily located at the knot positions of the discrete grid of the shape image Γ_1 . If the segmentation results only consist of the shape images, the surface mesh may

be extracted using a suitable surface extraction technique, for example the marching cubes algorithm that was introduced by William and Harvey [Will 87]. The proposed registration approach yields N surface meshes with points that are corresponding by construction, as the cell structure of the mesh is not modified by the spatial transform. Finally, the covariance matrix of the registered shape surfaces can be decomposed into the principal variations via PCA.

9.3 Results

Three different algorithms for solving the point correspondence problem are compared to each other based on an evaluation using 3-D CT images of kidneys. We compare an established MDL approach [Heim 05] to the proposed non-parametric registration of shape images using a normal SSD and the novel CSSD measure. Each algorithm has been embedded into the ASM framework shown in Figure 9.1. The medical data for the evaluation consists of 3-D abdominal CT images from 41 different patients of mixed gender and age. The images have been acquired using two different Siemens CT scanners (Sensation 10 and Sensation 16) with spatial resolutions ranging between 0.6/0.6/5 to 0.75/0.75/5 [mm] in x/y/z and provided in DICOM format. The volume sizes for the experiments range between 512x512x120 to 512x512x300 voxels. In order to evaluate differences in the point correspondence algorithms, all 41 kidney pairs have been manually segmented. The results of the segmentations have been approved by a nephrologist and used as the gold standard in the following evaluations. The entire set of labeled segmentation data is divided into two disjoint parts for the evaluation: one set of varying sizes between 7 and 20 is used for the training of the ASM, the remaining shapes for testing. In the following, we will refer to the resulting ASM models as MDL, SSD and CSSD. All ASMs have been created on the same training data and tested with equal initialization parameters. Table 9.1 provides a brief description of the properties for the MDL approach for a single kidney. For the other two methods, the surface mesh \mathbf{s}_1 has been extracted from Γ_1 with 2000 vertices by a marching cubes algorithm. The SSD and CSSD models are based on surface meshes of 2000 vertices, as well.

Mean radius in voxels	22
Number of samples	41
Sample complexity for the MDL (# vertices)	2000-3000
Model complexity for the MDL (# landmarks)	2562

Table 9.1: Characteristics of the clinical datasets for one side of the kidney pairs used for the MDL approach.

In order to make the approach robust to the selection of the reference shape Γ_1 , a state-of-the-art rigid registration is applied to all shape images first, see Section 3.1. The components of an affine transform are not constrained by the curvature regularization term in the non-parametric case. For the conducted experiments, choosing the shape with the minimal mean curvature as reference worked fine for the registration of the entire training set.

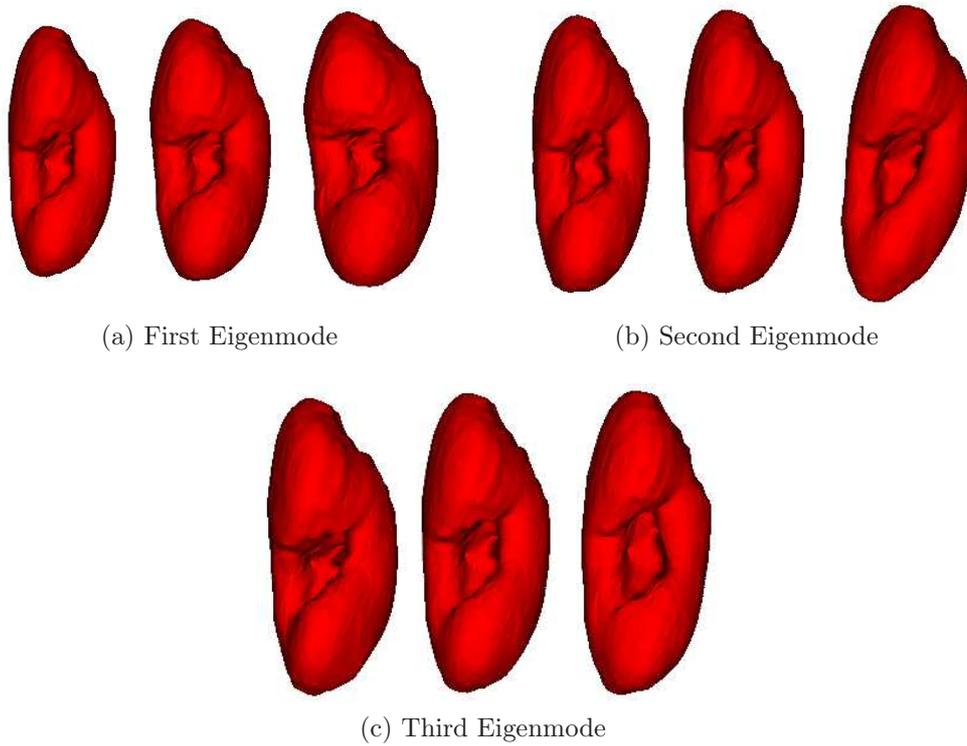


Figure 9.4: The three major principal modes of variation of a kidney shape model trained with 20 samples and created with the CSSD registration approach.

In order to get an impression of the statistical information contained within the training samples, Figure 9.4 illustrates the variations along the three major principal modes. The model has been created from 20 training samples using the CSSD registration approach to compute the correspondences between the training shapes.

9.3.1 Evaluation Measures

The evaluation of the results for the segmentation of the kidneys with the three ASMs is based on the comparison with the gold standard. We focused on measuring the generalization of the models, i.e. the ability to adapt the shape to a kidney image that is not contained within the training data, and the segmentation error with respect to the gold standard. The ability to adapt the models to new input data depends on the number of incorporated principal components and the variation within the training set. Therefore, we constructed models with all three approaches using different numbers of training samples.

The segmentation error is measured by the MSE and the sensitivity (SE). Given the i -th gold standard shape image Γ_i , which is not contained within the training data, and the segmentation result $\hat{\Gamma}_i$ from an ASM, the MSE is given by:

$$\text{MSE}[\Gamma_i, \hat{\Gamma}_i] = \frac{1}{|\Omega_i|} \int_{\Omega_i} \left(\Gamma_i(\mathbf{x}) - \hat{\Gamma}_i(\mathbf{x}) \right)^2 d\mathbf{x} . \quad (9.9)$$

It is a measure for the mean squared distance between the segmentations and depends on the resolution used in the discretization. Therefore, the same discretization

parameters have been applied to $\hat{\Gamma}_i$ as used for the shape image Γ_i . The SE, as the second evaluation criterion, is defined as:

$$\text{SE} = \frac{\text{TP}}{\text{TP} + \text{FN}} , \quad (9.10)$$

where TP is the number of true positives and FN the number of false negatives within $\hat{\Gamma}_i$. In our case, TP is given by the number of voxels that are segmented consistently as kidney tissue in both shape images $\hat{\Gamma}_i$ and Γ_i . FN is the number of voxels that have been falsely classified as background. The SE is used to measure the conditional probability for a correct segmentation of kidney structure according to the gold standard. Therefore, this measurement states how well the shape model generalizes to new kidneys that are not contained within the training set.

In literature, the specificity is often used as an additional criterion for the quantification of segmentation results. In practice, however, there is a problem with this particular measure due to a missing background normalization: an increase of the background region in the segmented images reduces the influence of the actual segmented region on the specificity value. Background voxels are usually not recognized as members of the segmented structure anyway. Hence, it is difficult to use the specificity value as an evaluation measure.

9.3.2 Experimental Results

This section presents the evaluation results for three ASM generation methods. All experiments have been performed on a Pentium 4, 2.8 GHz single core CPU with 2 GBytes of main memory. On this hardware, a single registration of two shape images from the training set takes approximately 8 minutes for both the SSD and the CSSD approach, the registration of the largest model with 20 training shapes approximately 170 minutes in total. In comparison, the computation of the correspondences with the MDL approach lasts up to 20 hours. All timings refer to implementations of the algorithms that have not been optimized. As several tasks may be performed in parallel, further improvements could be achieved by utilizing multi-core processor architectures.

Since the presented segmentation system has to be initialized by a seed point, the following experiments are divided into two parts. First, we present results for varying locations of the seed point placements in the model initialization. Second, the center of gravities of the test segmentations are used as initialization in order to analyze the performance of each model with respect to a larger test set and different numbers of training shapes. As mentioned before, the parameters for the ASM segmentations were identical for all three models throughout the corresponding experiments. The segmentation itself, i. e. the adaption of the gray level appearance model to the current image, was implemented using a multi-resolution technique to increase the attraction range of the optimization algorithm and to accelerate the computation. The numerical convergence criterion for each resolution level is based on the variation of the MSE of the segmentation between two subsequent iterations. In general, the segmentation algorithm converges in less than 30 seconds, where the maximum number of iterations for each level is set to 70. Analyzing the results, 30 iterations for one level are generally sufficient to achieve numerical convergence.

Sensitivity to Seed Point Variations

A series of tests has been carried out to evaluate the robustness with respect to the placement of the initial seed points. The proposed ASMs are evaluated for various seed positions with respect to the MSE (9.9) and the SE (9.10). The incorporated modes of variation cover 99.9% of the training set. 13 different positions for the initial seed point placement have been chosen to reflect typical user inputs. From the test set, samples for one left and right kidney were used for this experiment. Results on the left kidney for the models generated with 20 training samples are provided in Table 9.2. Table 9.3 contains the corresponding results for the right kidney and, additionally, shows a comparison with the models based on 10 training samples.

Initialization Sensitivity - Left Kidney

$N = 20$	MDL	SSD	CSSD
SE	0.67 ± 0.12	0.94 ± 0.02	0.95 ± 0.01
MSE	100490 ± 33599	22991 ± 3545	22835 ± 3046

Table 9.2: Results for the ASM segmentation of a left kidney for 13 different starting positions. The values correspond to the mean and the standard deviation of the evaluation measures. The standard deviation within the initial placements was (3.58, 1.98, 4.14) [mm] in (x, y, z) coordinates. The models were trained with 20 sample shapes.

Initialization Sensitivity - Right Kidney

$N = 10$	MDL	SSD	CSSD
SE	0.69 ± 0.18	0.74 ± 0.13	0.74 ± 0.12
MSE	79330 ± 44930	61064 ± 33263	60084 ± 33043

$N = 20$	MDL	SSD	CSSD
SE	0.77 ± 0.15	0.91 ± 0.01	0.91 ± 0.01
MSE	54680 ± 37344	13866 ± 2159	13501 ± 1098

Table 9.3: Results for the ASM segmentation of a right kidney for 13 different starting positions. The values correspond to the mean and the standard deviation of the evaluation measures. The seed point variation was (3.04, 1.98, 4.14) [mm] in (x, y, z) . In addition to the results for models generated from 20 samples, the evaluation measures for a training set of 10 samples are presented, as well.

Both the SSD and the CSSD ASMs achieved better scores compared to the MDL method. The curvature-extended similarity measure yields slight improvements over the standard SSD approach. According to these experiments, the models generated by the registration approach are less sensitive to variations of the initial seed point placement. Figure 9.5 shows a graphical comparison of the values in Table 9.3. The segmentation was improved for all three models trained with a larger set of samples. The results from the registration approaches indicate a higher relative gain in performance for more samples than for the MDL model. Figure 9.6 shows an

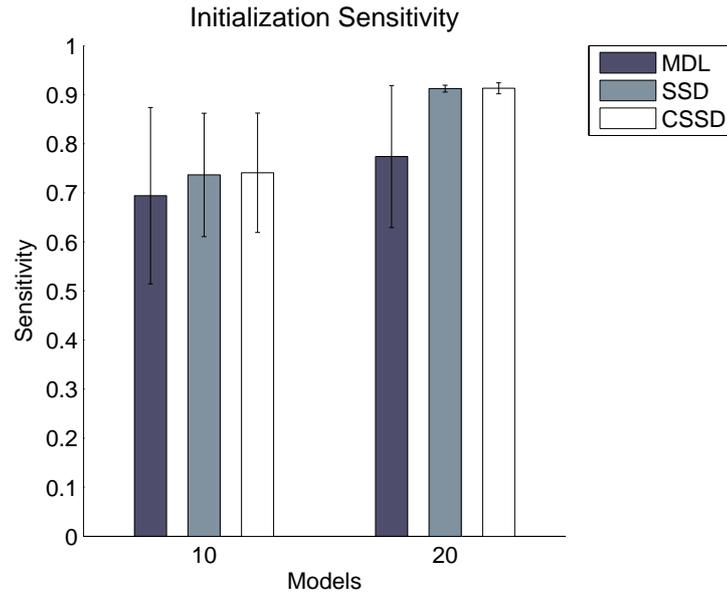


Figure 9.5: Comparison of the three different ASMs based on 10 and 20 training samples for varying starting positions in a segmentation of a right kidney. The chart reflects the results of Table 9.3. The registration approaches show an increased benefit from a larger training set compared to the MDL model.

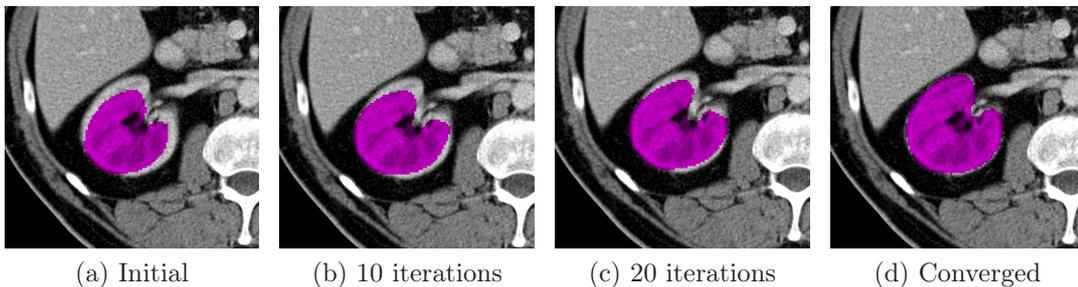


Figure 9.6: The images show the iterative refinement a right kidney segmentation example using the CSSD model trained with 20 samples. The sequence illustrates the progress from (a) the initial placement of the mean shape, (b) after 10 iterations, (c) after 20 iterations, and (d) through to numerical convergence.

example for the segmentation using the CSSD model trained with 20 samples. The progress of the iterative refinement of the model position is illustrated for a single 2-D slice taken from the 3-D test volume. The images show the initial placement of the mean shape, intermediate results after 10 and 20 iterations, and the final, converged segmentation. The 3-D visualization of the final result after convergence is provided in Figure 9.7.

Cross-Validation Segmentation Results

In the second part of the experiments, we present the evaluation results for the cross-validation regarding the overall segmentation errors for fixed, initial seed point

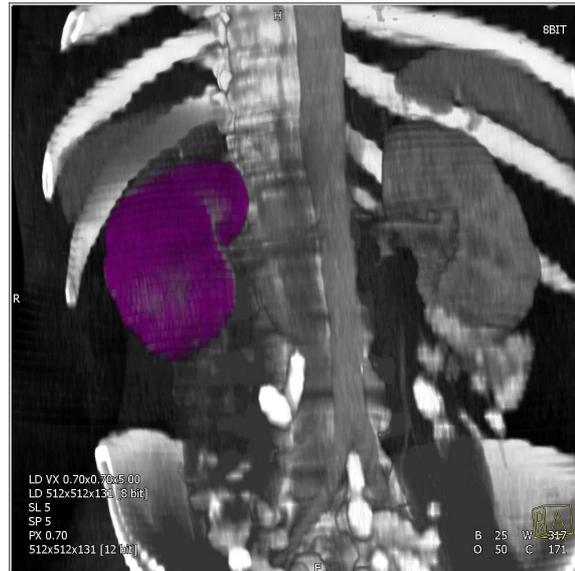


Figure 9.7: 3-D visualization of the segmentation result for the example shown in Figure 9.6. The content colored in purple marks the segmentation result after convergence for the CSSD model trained with 20 samples.

placements. The labeled data has been divided into disjoint sets for the training and the testing. The training was performed with 7, 10, 15 and 20 different training samples, the remaining samples have been added to the test set. Considering the MDL, SSD, and CSSD approaches, this yields 12 models in total. Compared to the previous tests, where the initial placement varied, the seed point has been automatically placed at the center of gravity of the gold standard segmentations. This initializes the segmentation phase (see also Figure 9.1) equally for all compared models.

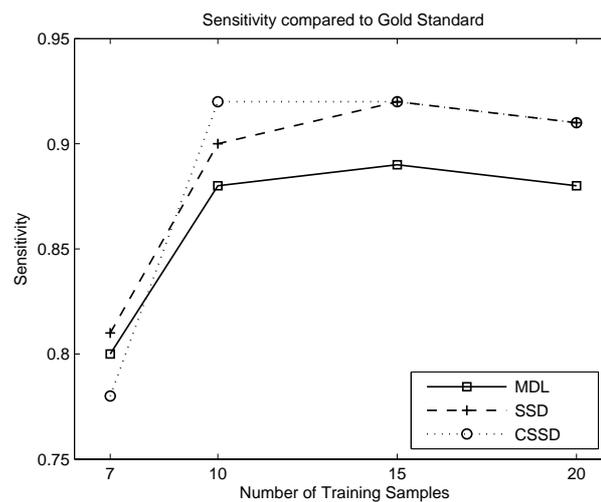


Figure 9.8: A comparison of the sensitivity for the MDL, SSD and CSSD models generated by a different number of training samples. For 10 or more training samples, both the SSD and CSSD approaches yield better results than the model created with MDL point correspondences.

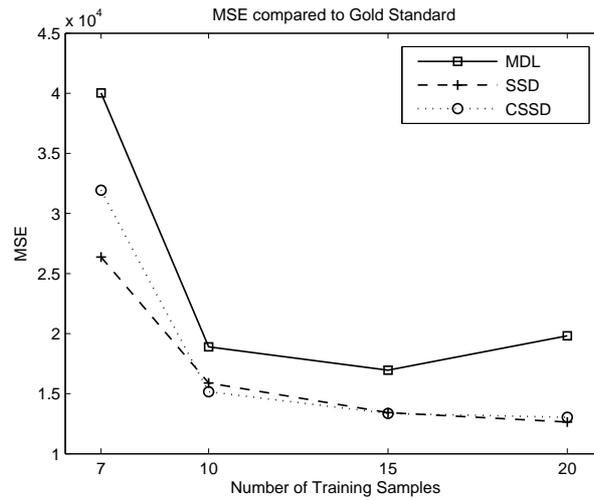


Figure 9.9: The MSE for the three compared model with respect to the number of training samples. Again, the registration approaches outperform the MDL technique in terms of registration accuracy.

Figure 9.8 shows the chart for the comparison of the three models with respect to the SE for varying numbers of training samples. The chart for the MSE is provided in Figure 9.9, respectively. The corresponding numerical values for each measure are given in Table 9.4. The results acquired from all models clearly indicate that

SE				MSE			
#	MDL	SSD	CSSD	#	MDL	SSD	CSSD
7	0.8	0.81	0.78	7	40037	26366	31930
10	0.88	0.9	0.92	10	18892	15883	15162
15	0.89	0.92	0.92	15	16955	13425	13357
20	0.88	0.91	0.91	20	19820	12635	13043

Table 9.4: Mean values for the SE and MSE of the models with respect to varying numbers of training samples for the ASM generation.

the ability of an ASM to adapt to a new shape depends on the number of training samples used for its creation. In our experiments, the SSD and CSSD approaches deliver better results for models trained with 10 or more samples. In the case of 7 training shapes, the SSD model performs best, however, for larger training sets, the CSSD model yields better results. In terms of the MSE, the MDL model delivered the worst results. Considering the ASMs trained with 10 or more training samples, the MSE values decrease almost linearly for SSD and CSSD, whereas for MDL, it seems that it is not capable of further improvements with more training data. The suspicion substantiates from a comparison between the mean shapes of the CSSD and the MDL models created from 20 samples, see Figure 9.10. The CSSD model provides more morphological information on concave parts of the surface, while the MDL mean surface contains considerably less anatomical detail. Although the mean surface of the MDL model is composed of a similar number of triangles, the distribution of points on the surface of the CSSD model leads to a better representation of the

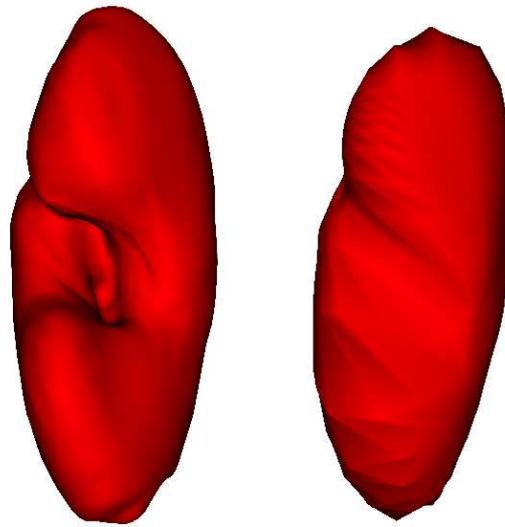


Figure 9.10: A comparison of the mean shapes between the CSSD and the MDL models created from 20 training samples. Left: CSSD registration approach model. Right: MDL model.

surface properties of the kidneys. The more detailed surface representation in the

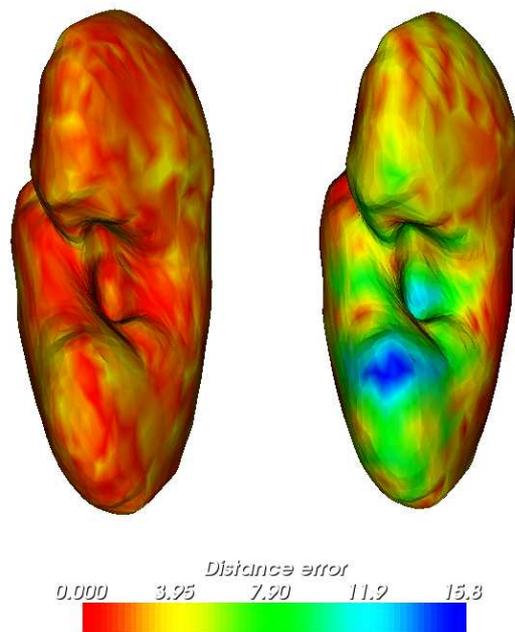


Figure 9.11: Distances between the surfaces of a test shape to the resulting segmentations with the CSSD (left) and the MDL model (right), both models were trained with 20 samples. The colors mapped onto the surface mesh of the test sample indicate the distances in mm between the closest points to the segmentation results.

CSSD model yields a higher accuracy in concave segmentation areas. The distances to the surface of the ground truth shape are smaller than those achieved with the

MDL mode, as shown in Figure 9.11. The numerical values are visualized by a color mapping onto the surface of the test segmentation. The CSSD result has a better accuracy at the area around the renal hilus, where the errors for MDL are large.

Chapter 10

Outlook

The statistical methods presented in this work are used to estimate intensity distributions in multi-modal image registration algorithms. The parameters for the implementation are optimally adapted to the input images, and the results show that the resulting robustness of the solutions cannot be achieved with fixed standard settings found in literature or determined empirically. The methods have been applied to a rigid registration with NMI used as a similarity measure.

An open issue is whether sparse sampling techniques are suitable for being applied in a non-parametric registration framework. For the rigid transform, the derivative of the measure can easily be computed using either finite differences or an analytical formulation. From a numerical point of view, the derivatives of the similarity measure with respect to a parametric transform are better conditioned than in the non-parametric case. There, the value of the derivative is required at each element of the deformation field and is not summarized over all elements within the overlap domain. It is certainly of interest whether sparse sampling for the estimation of the densities is providing enough information for a numerically stable optimization in the latter case. Additional experiments have to be conducted with the focus on the validity and numerical stability of the derivative values.

Besides statistical sampling, a projection approach was presented to divide the search space for the registration transform parameters into disjoint, lower-dimensional subspaces in order to reduce the computational complexity. The method has been applied to mono-modal images in 2-D/2-D and 3-D/3-D registrations. The projection theory with disjoint parameter sets for the optimization might be applicable also to 2-D/3-D registration problems, where 2-D digitally reconstructed radiographs (DRR) are usually created from a 3-D volume by ray casting algorithms. The similarity measure with respect to the spatial transformation of the volume is then computed between the DRRs and the 2-D X-ray images. A change in the position and orientation parameters of the 3-D image, hence, requires a re-computation of the DRRs. For the perspective projections used in the DRR generation, a projection geometry may be found that allows for the refinement of a subset of the transform parameters with a single projection. Additional complexity is introduced by the projection geometry, but less projections of the 3-D volume are required in the end. This might help to reduce the computation times for the 2-D/3-D registrations. Additional challenges for projection-based approaches arise from multi-modal input data: a projection com-

puted from a CT volume has a completely different visual appearance than those computed from an MR or PET volume. The reduction of the dimension from 3-D to 2-D, or even to 1-D projections, reduces the statistical information. A direct application of a multi-modal similarity measure, such as the NMI might yield a substantial amount of local optima in the objective functions for the subspace optimizations. It has to be investigated whether the correct optimum can still be located in the projected space and the approach is applicable to multi-modal problems.

New technologies in computer graphics hardware can be utilized to further reduce the runtime requirements of the registration algorithms. An implementation of the cylinder projection scheme in OpenGL Shading Language is proposed in Chapter A. For example, the Compute Unified Device Architecture (CUDA), which was recently introduced by the Nvidia company, provides the necessary programming interfaces to port also statistical similarity measures onto the GPU. Compared to OpenGL Shading Language, CUDA is more suited for general purpose programming tasks. Dedicated high-end workstations with just a few of those GPU cards yield an enormous amount of processing power. With increasing graphics memory that is available on new graphics hardware, the computation of the similarity measure can be entirely performed on the GPU, and transfers of large amounts of data between the CPU and the GPU memory can be avoided.

The results of the registration algorithms need to be computed both in a short amount of time and with a high accuracy. Especially the results of non-parametric algorithms with a large number of DOF have to be assessed thoroughly to determine whether the deformation makes sense. There is currently no database available that provides gold standard deformations similar to the RIRE database for the rigid registration. In this thesis, we proposed a statistical evaluation framework that incorporates the reliability of medical experts who provide the manual segmentations of the images. In order to establish this system in practice, a multi-modal database of images has to be segmented by various physicians, which is not a trivial task. The physicians have to be supported by reliable, semi-automatic segmentation algorithms to reduce the workload. Similar to the RIRE project, a blind study of the registration results requires an internet site that provides support for distributing the data without the information about the segmented labels, uploading the spatial transformations, and an automatic evaluation of the achieved accuracy.

Regarding the presented application, extended studies of the SPECT subtraction results using image registration should be conducted to further examine the influence of the deformable registration. One of the interesting aspects is how much regularization on the deformation field is sufficient to compensate the bias in the rigid registration. Additional models for the intensity transform between the SPECT images may be investigated as well.

Registration techniques were used to solve the correspondence problem for ASMs based on explicit representations of shape surfaces by triangle meshes. Similarly, the methods can be integrated into shape-based level set formulations. In an implicit formulation, the shape variations may be described by the eigenvectors of the deformed shape images directly, or by an analysis of the deformations between the training shapes. The latter approach might be incorporated directly into the iterative calcu-

lation of the spatial position update of each point located on the moving zero level set instead of a regularization of the already updated curve.

In the case of shape-based segmentation algorithms, the prior knowledge about the statistical variations of an object helps to increase the robustness of the solutions. Similar benefits can be expected when incorporating prior knowledge into registration algorithms. Gütter et al. [Guet05] and Zöllei [Zoll06] have proposed to add prior knowledge to statistical similarity measures in terms of intensity distributions learned from ground truth data. At least for CT images, the intensities relate to the density of the tissue. A modeling of the physical deformation properties of various tissue types might be incorporated into the regularization of the deformation fields as well. In a further step, one can think of identifying anatomical structures within the images first and adding this information to the registration functional afterwards, for instance, as a regularization energy that models the physical capabilities of the deformations of the human body in a larger scale. Of course, the latter approach requires a robust segmentation of the organs within the images, however, the reward is a deformation field that is reasonable in a physical sense.

Chapter 11

Summary

Statistical methods are successfully employed in a wide range of image processing algorithms. In medical registration applications, intensity distributions provide necessary information about the similarities between various types of images. Such registration techniques are required, for example, in radiation therapy, longitudinal patient studies, and image reconstruction. A small selection of applications are described in Chapter 2. A fusion visualization of the registered images provides valuable, additional information to the physician, but requires registration algorithms that have to support many modalities and are both robust and accurate, no matter what images are used as input.

Theory

The mathematical notation used throughout this thesis is introduced in Chapter 3. Each registration approach is based on an objective function that measures the degree of alignment between some input images with respect to a spatial transformation. Based on the type of this transform, we divide the registration algorithms into parametric and non-parametric approaches. In the former class of algorithms, the problem is well-conditioned, because the number of parameters is small compared to the spatial information within the images. The presented parametric mappings include rigid, affine, and cubic B-spline transforms. The non-parametric registration algorithms require a regularization of the spatial transform, which consists of a dense deformation field, in addition to the image similarity measure. We briefly summarize the steps towards a solution for the partial differential equation utilizing the calculus of variation. The curvature regularization energy is specified and used throughout this thesis for the non-parametric registration tasks. The similarity measure is the main part of the objective function. Besides a coarse classification of existing measures, Section 3.3 describes commonly applied feature-based and intensity-based similarity measures. The latter class is divided into direct, indirect non-statistical, and indirect statistical measures.

Especially in clinical applications, the input images vary in size, field of view, modality, content, and signal to noise ratio. Statistical similarity measures are currently state-of-the-art in the registration of medical images. Some engineering parameters have to be specified for an implementation in an automatic, intensity-based image registration algorithm. Standard settings for these values can be found in

literature, however, these parameters are actually dependent on the input data and mutually influence each other. Therefore, a single, empirically determined set of parameter values that works for all input data cannot be found. In Chapter 4, we focus on crucial numerical aspects of the estimation of the intensity distributions and data-driven schemes to optimally adapt the implementation variables to the input data. A jittering technique helps to reduce sampling artifacts in the objective function. The artifacts mainly occur due to regular sampling patterns and grid-aligning effects. The sampled image intensities are regarded as random measures from an unknown PDF. We apply Parzen-windowing to estimate this PDF using a kernel width determined by optimizing a log-likelihood function in a leave-one-out cross-validation. A discretization scheme based on histogram binning is presented, which greatly reduces the computational costs for both the density estimation and the data-driven computation of the kernel width. The optimization of the log-likelihood function can be achieved by numerical methods that require derivatives of the kernel functions with respect to their widths. The derivatives are formulated for a Gaussian and a cubic B-spline kernel. The discretization of the density estimator involves a convolution of the kernel with the histogram. A quasi-adaptive Parzen-window estimator is proposed in order to account for variations in the true PDF and to keep up the efficient convolution scheme. In addition, we introduce a trade-off for the discretization error to automatically determine a number of bins for the histogram. As medical images usually contain structured noise within the background region, we propose a coincidence weighting scheme that is based on automatically determined threshold values. The weighting within the joint PDF reduces the influence of the background and the noise artifacts contained within. Implementation guidelines are provided before showing results of the data-driven density estimation. In a multi-modal registration example with different initial positions and orientations, the optimal parameter values yield much better alignment results than those achieved with standard settings.

The similarity measure computation has the highest computational complexity of the entire registration algorithm. In a straightforward implementation, it has to be recomputed for a large number of samples whenever the transform parameters change during the nonlinear optimization. An alternative approach to reduce the computational complexity is described in Chapter 5, where the transform parameters are regarded as random variables. It is known from statistics that marginalization can be used to eliminate the influence of random variables within the measurements. The presented projection scheme in image registration is analogue to marginalization, as the projections are invariant to some of the transform parameters. The key aspect of the approach for rigid registrations is that a reduced subset of the parameters can be optimized from the projection images without having to reproject. In the case of 2-D registrations, a complete separation of the parameters into 1-DOF subspaces is achieved by projecting along the coordinate axes and onto a circle. In 3-D, oriented cylinder geometries enable the separation into three disjoint parameter subsets for a rigid registration, each containing a rotational and a translational parameter. The 3-D volume is projected onto the 2-D image of the cylinder surface, which can, again, be reduced to 1-D projections by marginalization along the axes. This scheme yields a separate 1-DOF optimization problem for each transform parameter, which is computationally very efficient. As a necessary requirement to achieve valid re-

sults, the projection geometry has to be fully enclosed within the overlap domain of the input images. We propose solutions for the fitting problems together with an additional constraint to enforce uniqueness. The basic concepts of the projection approach are similar to those incorporated within the recently developed marginal space learning. Therefore, we compare the two approaches and delineate the differences. Additional information on the implementation of the projection scheme on a GPU, using OpenGL Shading Language, is provided in the Appendix A.

In the final Chapter 6 of the theoretical part of this thesis, we introduce a statistical framework for the evaluation of registration algorithms based on segmentations supervised by medical experts. There is currently no database with an established gold standard available for deformable registration techniques, such as the RIRE project for rigid transformations. Establishing a gold standard for spatial deformations is a highly non-trivial task due to several reasons: evaluations based on artificial deformations are always biased towards the technique that is used for the creation of the ground truth; fiducial markers can hardly be attached to deformable tissue within the human body; and the intensity-based similarity measure alone is not sufficient to deduce whether a deformation was physically correct. One possibility to acquire information about the correspondences between the images is to employ the knowledge of medical experts who have to label the data either manually or semi-automatically. Of course, this process leads to results that may vary with respect to the reliability of the observers. We incorporate this uncertainty into the evaluation approach by a representation of the segmented labels with spatial distributions. Each labeled region, surface, and anatomical landmark is regarded as a random measure, which, again, allows for non-parametric density estimation with Parzen-windowing. The distance measure between two segmented images with respect to the spatial transform is based on a symmetric KL divergence. We briefly describe the techniques for the spatial density estimation and show an example for the change in the quality criterion of a labeled MR image that is transformed by a B-spline deformation.

Applications

Chapter 7 consists of an evaluation of the proposed statistical techniques on the RIRE database, which contains CT, MR, and PET images together with gold standard transformations determined from implanted fiducial markers. A single evaluation on this data requires 114 registrations. We examine influences of various parameter settings for the implementation of the NMI measure and compare the results also to a manual registration. Along with three values for the minimal number of bins in the histogram layout, this yields 25 evaluations in total. The significance tests on the registration errors are performed with a two-tailed, paired t -test at a significance level of 5% for the median errors. The parameters selected by data-driven estimation schemes outperform the standard settings with a high significance. If the variables are automatically adapted, there are no indications for significant changes in the results with respect to the number of histogram bins. The landmark acceptance rates feature a better performance of the data-driven approach compared to both the manual registration and the standard settings. We also conclude from the results that a geometry correction of the MR images yields more accurate transforms.

Statistical registration techniques can also be applied to SPECT subtraction imaging for epilepsy diagnosis. The main goal is to localize the region of seizure onset. Two SPECT images are acquired at different activation stages in order to visualize the changes in the CBF. The comparison is impaired by different positions of the patient between the acquisitions and a missing intensity normalization. Therefore, statistical image registration techniques are utilized to align the two SPECT images in a first step. Afterwards, a model for an affine intensity transformation is fitted into the estimate of the joint PDF after the registration to enable a direct comparison of the intensity values. The automatic background thresholds, which are required for the coincidence weighting during the registration, are used to exclude the background from the intensity mapping. Otherwise, it may have a detrimental effect on the model placement and a large background region introduces a bias in the intensity transform. We demonstrate that local differences in the CBF have a global effect on the rigid transform, which can be improved by a strongly regularized, non-parametric registration. The sparse information of the final subtraction images can be spatially localized through a registration with an MRI in the final step. The subtraction imaging is applied to a collective of 26 patients and assessed in a blind evaluation study to compare a workflow with only a rigid registration to a version with an additional non-parametric alignment. Both approaches show similar percentages for the localization of the lesions, as well as the number of foci. In the cases of an additional non-parametric registration, however, the physicians report a higher certainty, which is also reflected by a larger correlation coefficient for the observer reliability.

Chapter 9 dealt with the incorporation of image registration methods in the generation of a shape model for image segmentation. Compared to unconstrained algorithms, the prior knowledge of shape variations used in an ASM segmentation helps to increase the robustness of the results and to avoid leakage into adjacent tissue. The shape variations are determined from a set of training samples by solving a point correspondence problem. If errors are made within this stage, the shape statistics are corrupted by noise and a decomposition into the principal modes of variation yields wrong statistics. State-of-the-art solutions for the problem are based on a reparameterization of the surface mesh and the optimization of an MDL criterion. Unfortunately, the entire approach has to be recomputed, if an additional shape is added to the training set. The mapping into a suitable parametric space, where the MDL function is optimized, may also be hard to achieve, especially for concave objects. We propose to tackle the correspondence problem by means of image registration. The training shapes are represented as discrete images using a signed distance transform. Establishing a spatial relation between the structures within the shape images by registration is then a solution for the correspondence problem. Therefore, a combination of parametric and non-parametric registration provides an alternative to an MDL formulation of the problem. We briefly summarize the basics of an ASM before proposing a novel scheme for the registration of the shape images. The SSD distance measure is extended by a curvature term to incorporate surface properties into the objective function. The deformation fields, which result from the registration between a selected reference shape and the remaining training samples, are used to transform the surface mesh of the reference shape into the spatial domains of the remaining samples. The elements within the measurement matrix for the decomposi-

tion into the principal modes of variation are, hence, corresponding by construction. Results are presented for the 3-D segmentation of kidneys from a collective of 41 CT images. In summary, the proposed registration approach yields a higher sensitivity and lower MSE values than corresponding models generated with an established MDL algorithm. A cross-validation is performed between training sets of 7, 10, 15, and 20 samples. In all experiments, the models created from registered training shapes yield better segmentation results. A comparison between the mean shapes of both approaches for a model with 20 samples reveals that the mean shape determined by registration contains more detail on the shape surface and smaller distances between the gold standard and the ASM segmentation results can be achieved.

Finally, we comment in Chapter 10 on future work to improve and further investigate the techniques presented in this thesis. The incorporation of prior knowledge into the registration algorithms may lead to major improvements, especially for establishing non-parametric deformations that relate to physical properties of the human body.

Appendix A

Cylinder Projections on the GPU

In chapter 5, we introduced a novel registration approach that reduces the search space by nonlinear projections. In order to rigidly register 3-D images, a cylinder projection scheme allows to separate the optimizations within the parameter space into disjoint 1-DOF subproblems. For the projection-based registration approach, it is necessary to project a 3-D input image onto the surface of a cylinder, which is discretized and unfolded into a 2-D projection image. The image axes correspond to a radial and an axial parameter. This projection image can be used in the registration approach to optimize parameters for the rotation angle around the cylinder axis, and the translation along of it, respectively. The algorithm for the computation of the cylinder projection is suitable for an implementation on dedicated graphics hardware. Modern GPUs contain fast texturing units that can be utilized for the interpolation along the samples of the projection rays.

The problem of projecting image content onto a cylinder surface is very similar to a standard ray casting algorithm [Fole97]. The required coordinate transformation is formulated in homogeneous coordinates and split into a nonlinear and a linear part. An additional rigid transformation, e. g. the transform that is currently refined in a registration algorithm, can be easily combined with the linear components of the coordinate transform without loss of performance. In addition, we present a multi-level extension of the cylinder projection scheme that leads to an increase in computational efficiency, as the number of computed rays is reduced.

A.1 Discretization of the Projection Image

A sampling scheme for the cylinder surface is chosen for the discrete representation of the projection such that each projection ray hits every voxel of the 3-D image along its path at least once. This requires either some sort of rasterization of the rays or an interpolation scheme. Rasterization techniques are prone to aliasing in cases where the resolution of the input image largely differs between the image axes whereas a sampling of the rays at equidistant positions implies an interpolation within the image grid. The latter technique, therefore, is less dependent on the resolution of the input image. In the following, we focus on a ray casting technique with equidistant sampling positions along the ray.

Calculating the line integral of a ray that is cast onto the surface with an orthogonal direction to the cylinder axis and requires a discretization into N_R sampling points along its way. The directions of the rays are determined by the intersections with the projection image \mathcal{P} , which is discretized into N_ϕ elements in the radial direction and N_t positions along the cylinder axis, i. e. $\mathcal{P} \in \mathbb{R}^{N_\phi \times N_t}$. Each pixel within the projection image \mathcal{P} of the cylinder surface is assigned a radial component ϕ and an axial coordinate position t for the position along the cylinder orientation axis \mathbf{o} . The tuple (ϕ, t) acts as a discrete, zero-based index into the projection image pixel space, with $\phi = 0, 1, \dots, N_\phi - 1$, and $t = 0, 1, \dots, N_t - 1$. The right-hand coordinate system of the unit vectors \mathbf{e}_x , \mathbf{e}_y , and \mathbf{o} , which point into the direction of the coordinate axes, describes a cylinder parallel to the z -coordinate axis. The vectors \mathbf{e}_x and \mathbf{e}_y may be simply exchanged to form a similar coordinate system for other orientations. The value of the numerical line integral along the ray through an image \mathcal{I} is:

$$\mathcal{P}(\phi, t) = \frac{1}{N_R} \sum_{i=1}^{N_R} \mathcal{I} \left(\mathbf{x}_{\text{cap}} + \underbrace{\frac{th}{N_t - 1} \mathbf{o}}_{\text{axial-}} + \underbrace{\frac{ir}{N_R - 1} \left(\cos \left(\frac{\phi 2\pi}{N_\phi} \right) \mathbf{e}_y - \sin \left(\frac{\phi 2\pi}{N_\phi} \right) \mathbf{e}_x \right)}_{\text{radial-distance}} \right). \quad (\text{A.1})$$

The cylinder radius is denoted by r , its height by h , and the spatial position of the lower cap center is provided in \mathbf{x}_{cap} (see also section 5.3.2).

A.2 Multi-Level Cylinder Projection

In order to keep the discretization error of the projection \mathcal{P} low, the values N_ϕ, N_t for the projection size in radial and axial direction have to be large enough to retain the information of the input image within the projection. A disadvantage of the straightforward application of the direct ray casting strategy is that the distance between the rays is decreasing towards the axis of the cylinder, which is due to the fact that we describe a rotation around the cylinder axis. In a discretization of the approach, the distances between the rays are usually chosen small enough to hit each voxel near the cylinder surface at least once. Accordingly, the voxels located closer towards the cylinder axis are intersected by far more rays, which results in an increasing amount of oversampling. As each row of \mathcal{P} corresponds to a slice through the cylinder, which basically is a circle projection within the slice, Figure 5.2b can be consulted to illustrate this problem: the projection onto a cylinder or a circle results in varying sampling distances between the projection rays. In terms of computation efficiency, this oversampling leads to more interpolation operations performed in the straightforward approach than necessary.

Therefore, we propose a multi-level computation scheme for the cylinder projection to reduce the oversampling and save computation time. The scheme is illustrated in Figure A.1, where the usage of three levels for the reduction of oversampling is shown. The number of required rays is reduced within the inner segments. As the projection images of the inner segments feature a lower resolution along the radial dimension, the values for the line integrals have to be interpolated at the starting positions of the rays for the next segment. Each segment border specifies a sub-cylinder

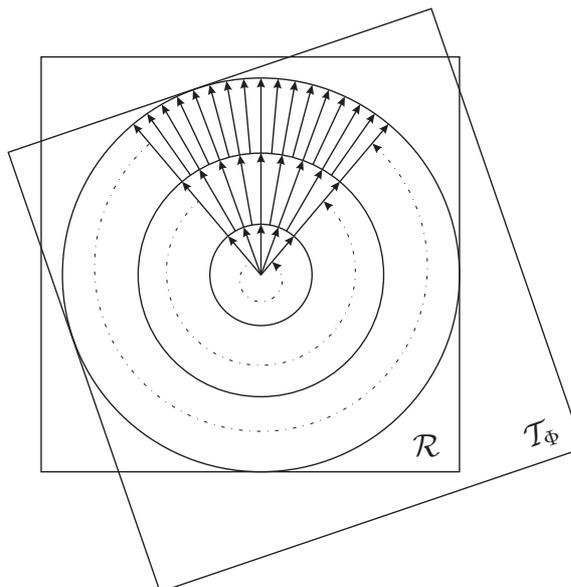


Figure A.1: Multi-level cylinder projection using three intermediate projection levels. The radial sampling distance is decreased in the projections of the inner segments. The results of the intermediate projections are lifted into the next segment by means of interpolation.

surface that is discretized such that each voxel on the surface is at least hit once by a projection ray. The first level consists of rays emanating from the projection center and ending at the border surface of the first level. The projections for the first level are calculated as given in (A.1). The next levels consist of decreasing sampling distances between the rays that start at the surface of the previous level and end at the next level or the cylinder surface, respectively. As the entire projection is a summation of sampling points from the center along the ray, the projection portions of the lower levels have to be lifted into the next level by interpolation on the low-level surfaces. The entire multi-level projection scheme can be computed efficiently on the GPU.

A.3 Coordinate Transformation

The coordinate transformation for the sampling positions along the ray to compute the integral (A.1) is split into two parts using homogeneous coordinates. The linear transformation is precomputed outside the ray casting routines and contains the necessary transformations to map a cylinder coordinate into the coordinate system of the image for the interpolation of the intensity value. This may also include a spatial transform on the volume, e.g. the registration transform that changes in each iteration of the registration algorithm. The position in cylinder coordinates is computed for each angle and distance of the ray with respect to the cylinder axis within the nonlinear part.

The sample positions along the rays are discretized using the 3-tuple (ϕ, t, i) , which describes the location of the i -th sample along the ray within the cylinder. The

transformation $\Phi_{\text{cyl}}(\phi, t, i)$ yields the spatial coordinate \mathbf{x}_c^i without the components for orientation and offset:

$$\mathbf{x}_c^i = \begin{pmatrix} x_c^i \\ y_c^i \\ z_c^i \end{pmatrix} = \Phi_{\text{cyl}}(\phi, t, i) = \begin{pmatrix} -\frac{ir}{N_R-1} \sin\left(\frac{\phi 2\pi}{N_\phi}\right) \\ \frac{ir}{N_R-1} \cos\left(\frac{\phi 2\pi}{N_\phi}\right) \\ \frac{th}{N_t-1} \end{pmatrix}. \quad (\text{A.2})$$

In the linear part, the transformation has to incorporate the offset and orientation of the cylinder, a spatial rotation \mathbf{R} , and a translation \mathbf{t} of the image. In addition, we include the necessary transformations into the voxel coordinate system of the image, which makes use of the orthonormal vectors $\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z \in \mathbb{R}^3$ for the orientation of the image coordinate axes, the origin of the image \mathbf{x}_o , and the sizes $s_x, s_y, s_z \in \mathbb{R}$ for a voxel along each coordinate axis. The voxel coordinate \mathbf{v}^i for the i -th position is then:

$$\mathbf{v}^i = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{V}^{-1} \mathbf{D}^T \mathbf{T} \begin{pmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{o} & \mathbf{x}_{\text{cap}} \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\text{linear}} \underbrace{\begin{pmatrix} \mathbf{x}_c^i \\ 1 \end{pmatrix}}_{\text{nonlinear}} \quad (\text{A.3})$$

$$\mathbf{V} = \begin{pmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{A.4})$$

$$\mathbf{D} = \begin{pmatrix} \mathbf{d}_x & \mathbf{d}_y & \mathbf{d}_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{A.5})$$

$$\mathbf{T} = \begin{pmatrix} \mathbf{R} & \mathbf{t} - \mathbf{x}_o \\ 0 & 1 \end{pmatrix}. \quad (\text{A.6})$$

The linear part of the coordinate transformation is constant for all rays of the current projection.

A.4 GPU Implementation

Current state-of-the-art GPUs provide hardware support for texture interpolation together with programmable vertex and fragment shading pipelines with at least 32bit floating point accuracy. In the case of the aforementioned projection algorithm, this yields an advantage over an implementation on a state-of-the-art CPU, which does not provide hardware-accelerated interpolation of the image intensities and has less processing cores. In the following, we present an implementation scheme for the GPU that implements the multi-level cylinder projection by ray casting. Both the nonlinear transformation and the interpolation of the image intensities are incorporated within a fragment shader. The sum in (A.1) is realized by alpha blending using the framebuffer and an additional accumulation texture. Finally, we present code for a fragment shader for the cylinder projection written in OpenGL Shading Language.

In the direct projection scheme, a ray is cast from the projection center towards the cylinder surface, image intensities are collected along the line of the ray, and the value for the integral is written to the output. This scheme cannot be ported directly to the GPU, as loops and conditional statements are computationally expensive in a parallel processing pipeline. Instead, we make use of a technique applied in direct volume rendering, where the intensities along the rays are summarized using alpha blending [Enge06]. An implementation on the graphics hardware for the back-to-front rendering of view-aligned planes through the volume, as it is required for volume rendering, can utilize appropriate alpha blending techniques for the summation. In the cylinder projection, the analogue to the view-aligned planes are the projections at each sampling step i for the distance to the cylinder axis. The back-to-front rendering is, therefore, the collection of intensity values along the rays from the inside to the outside of the cylinder towards its surface. The summation is achieved by alpha blending between an accumulation texture and the framebuffer. For a GPU implementation of the cylinder projection, the cylinder surfaces are treated as unfolded 2-D images that are rendered directly into the framebuffer of the graphics card. The nonlinear part of the transformation is computed for each fragment and combined with the linear transform.

Using the texturing unit for the interpolation requires a representation of the images as textures: a 3-D texture for the input image and 2-D textures for the framebuffer and the accumulation texture. The mapping into the voxel space (A.3), thus, requires an additional transformation into the corresponding texture spaces. Commonly, the discretization on the CPU is node-based, whereas the OpenGL texture interpolation is performed in a cell-based notation. In a cell-based texture, each texture element represents the center of a spatial region, whereas in a node-based approach, the region is defined by the neighboring knots of the image grid. The dif-

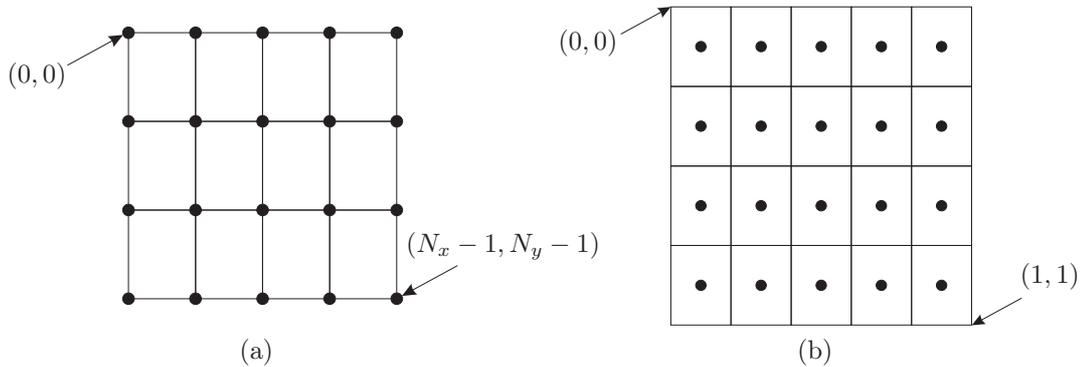


Figure A.2: The (a) node-based image grid is commonly used for image interpolation on the CPU, whereas in OpenGL, the (b) cell-based notation is applied for the representation of textures.

ference between the two representations is shown in Figure A.2. They are related to each other by a conversion between the index of a voxel v and a cell c :

$$c = \frac{v}{N} + \frac{1}{2N} \quad (\text{A.7})$$

$$v = cN - \frac{1}{2}. \quad (\text{A.8})$$

Here, $c \in [0, 1]$ denotes the cell-based texture coordinate and $v \in [0, N - 1]$ the node-based coordinate for the element of an image of size N . Within the fragment shader, we account for the change into the cell-based representation using an adapted nonlinear transform to map the texture coordinates in the corresponding framebuffer pixel to the cylinder coordinates in a node-based representation:

$$\mathbf{x}_c^i = \begin{pmatrix} x_c^i \\ y_c^i \\ z_c^i \end{pmatrix} = \Phi_{\text{cyl}}^{\text{cell}}(\phi, t, v) = \begin{pmatrix} -\frac{ir}{N_R-1} \sin\left(\left(\phi N_\phi - \frac{1}{2}\right) \frac{2\pi}{N_\phi}\right) \\ \frac{ir}{N_R-1} \cos\left(\left(\phi N_\phi - \frac{1}{2}\right) \frac{2\pi}{N_\phi}\right) \\ \left(t N_t - \frac{1}{2}\right) \frac{h}{N_t-1} \end{pmatrix}. \quad (\text{A.9})$$

The 3-D image is stored as a texture in the graphics card memory as well. Therefore, the coordinate transform (A.3) has to include the conversion from node- to cell-based representation, too. The coordinate \mathbf{c}^i within the GPU texture space is converted from the voxel index \mathbf{v}^i using:

$$\mathbf{c}^i = \begin{pmatrix} N_x & 0 & 0 \\ 0 & N_y & 0 \\ 0 & 0 & N_z \end{pmatrix}^{-1} \mathbf{v}^i + \begin{pmatrix} \frac{1}{2N_x} \\ \frac{1}{2N_y} \\ \frac{1}{2N_z} \end{pmatrix}, \quad (\text{A.10})$$

where (N_x, N_y, N_z) denotes the number of voxels of the image along the coordinate axes of the image space.

The loop over all samples along the rays is implemented by rendering the projection value at a specific distance from the cylinder axis into an accumulation texture. The accumulation texture is then added to the framebuffer using alpha blending, and the result is, again, written into the accumulation texture. The blending in the framebuffer can also be utilized to interpolate the projections of a previous segment of the multi-level projection scheme, as described in section A.2. The main program, which is evaluated on the CPU, controls the graphics pipeline to render the cylinder planes into the framebuffer from inside of the cylinder to the outside through all multi-level segments and distances to the axis. Finally, the pixel values of the framebuffer after the last run of the fragment shader contain the corresponding line integrals (A.1). In Figure A.3, an example implementation of the fragment shader, which is used to compute the i -th summation term in (A.1), is presented in OpenGL Shading Language. The variables used within the program are explained in Table A.1.

```
uniform sampler3D imageTex;
uniform sampler2D accumTex;
uniform float radialSpacing;
uniform float axialSpacing;
uniform float distance;
uniform mat4 linearTransform;
uniform vec2 framebufferSize;
uniform vec2 accumTexSize;
uniform float rayNormalization;

void main(void)
{
    float phi = (gl_TexCoord[0].x * framebufferSize.x - 0.5) * radialSpacing;
    vec4 cylCoord;
    cylCoord.x = -distance * sin(phi);
    cylCoord.y = distance * cos(phi);
    cylCoord.z = (gl_TexCoord[0].y * framebufferSize.y - 0.5) * axialSpacing;
    cylCoord.w = 1.0;
    vec4 imageTexCoord;
    imageTexCoord = linearTransform * cylCoord;
    imageTexCoord = clamp(imageTexCoord, vec4(0.0), vec4(1.0));
    vec2 accumTexCoord;
    accumTexCoord = accumTexSize * gl_TexCoord[0].xy;
    accumTexCoord = clamp(accumTexCoord, vec2(0.0), accumTexSize);
    float colAccum = texture2D(accumTex, accumTexCoord.xy).x;
    float colVol = texture3D(imageTex, imageTexCoord.xyz).x;
    gl_FragColor.x = colVol * rayNormalization + colAccum;
}
```

Figure A.3: Example implementation for a GPU cylinder projection fragment program written in OpenGL Shading Language. See Table A.1 for an explanation of the variables.

Program Variable	Explanation
imageTex	texture identifier for the input image \mathcal{I}
accumTex	texture identifier for the accumulation texture
radialSpacing	$\frac{2\pi}{N_\phi}$
axialSpacing	$\frac{h}{N_t-1}$
distance	$\frac{ir}{N_R-1}$
linearTransform	linear transformation from the cylinder into the texture space, as in (A.3), (A.10)
frameBufferSize	size of the framebuffer (N_ϕ, N_t)
accumTexSize	size of the previous projection, which may be smaller than the framebuffer when entering the next segment
rayNormalization	$\frac{1}{N_R}$

Table A.1: Explanation of the variables used within the fragment program in Figure A.3.

Appendix B

Notation

\mathbb{R}	set of real numbers
\mathbb{Z}	set of integers
Φ	general transformation
\mathbf{a}	parameter vector for parametric transformations
$\Phi_{\mathbf{a}}$	parametric transformation
\mathcal{R}	reference image
\mathcal{T}	template image
\mathcal{T}_{Φ}	transformed (and interpolated) template image
$\Omega(\mathcal{R})$	spatial domain of the reference image
$\Omega(\mathcal{T}_{\Phi})$	spatial domain of the transformed template image
$\Omega(\mathcal{R}, \mathcal{T}_{\Phi})$	spatial overlap domain between the reference and the template image
\mathcal{D}	similarity measure
\mathbf{R}	rotation matrix
\mathbf{t}	translation vector
B	cubic B-spline basis function
C^d	d-continuity
\mathcal{C}	3-D-control grid
\mathbf{u}	displacement vector field
\mathcal{U}	set of all displacement fields
\mathcal{J}	registration functional in the variational formulation
\mathcal{S}	regularization energy
∇	gradient operator
Δ	Laplace operator (∇^2)
$\nabla \cdot$	divergence operator
Γ	segmented surface and shape image
\mathbb{I}	identity matrix
$J_{\Phi}(\mathbf{x})$	Jacobian of the transform Φ at the spatial position \mathbf{x}
$\ \mathbf{x}\ $	Euclidean norm $\sqrt{\mathbf{x}^T \mathbf{x}}$ of vector \mathbf{x}
$\ \mathbf{A}\ $	Frobenius norm $\sqrt{\sum_i \sum_j a_{ij} ^2}$ of matrix \mathbf{A}

r	random sample associated with a reference image intensity
t	random sample associated with a transformed template image intensity
$p_{\mathcal{R}}$	PDF of the reference image intensities
$p_{\mathcal{T}_{\Phi}}$	PDF of the transformed template image intensities
$\mathbf{p}_{\mathcal{R}, \mathcal{T}_{\Phi}}$	joint PDF of the reference and the template image intensities
$\mu_{\mathcal{R}}$	mean value of the reference image intensities within the overlap domain
$\mu_{\mathcal{T}_{\Phi}}$	mean value of the transformed template image intensities within the overlap domain
$\text{VAR}[X]$	variance of the random variable X
$\text{E}[X]$	expected value of the random variable X
$\mathcal{H}(X)$	entropy of the random variable X
\mathcal{O}	Landau notation for limiting behavior of a function
K_{λ}	kernel PDF of width λ
$p_{\lambda, n}$	Parzen-window PDF estimator
$\hat{p}_{\lambda, n}$	discrete Parzen-window PDF estimator
\star	convolution operator
\mathcal{L}	log-likelihood function
h	discrete histogram containing the fractions of samples that fall into each bin
b	number of histogram bins
w	bin width of an equidistantly-spaced histogram
c_i	center value of i -th histogram bin
$[l_{i-1}, l_i]$	interval of the i -th histogram bin
g	1-D Gaussian kernel PDF
G	2-D Gaussian kernel PDF
\hat{p}_{0-0}	estimated probability for O-O intensity combinations
\hat{p}_{B-B}	estimated probability for B-B intensity combinations
$\mathbf{h}_{\mathcal{R}, \mathcal{T}_{\Phi}}$	discrete joint histogram with fractions of sample intensity combinations between the reference and the template image
$\psi_{\mathcal{R}}$	automatically determined threshold value that separates the background region from the object within the reference image
$\psi_{\mathcal{T}}$	automatically determined threshold value that separates the background region from the object within the template image
$b_{\mathcal{R}}$	number of bins determined automatically for the reference image
$b_{\mathcal{T}}$	number of bins determined automatically for the template image
\mathbf{e}_x	unit vector along the x -coordinate axis
\mathbf{x}_{rect}	corner point of an axis-oriented rectangle
l_x	side length of an axis-oriented rectangle in x -direction
\mathbf{x}_{cap}	spatial location of the lower cap center point of a cylinder
\mathbf{o}	cylinder direction axis

h	cylinder height
r	cylinder radius
\mathbf{n}	normal vector
$\mathbf{c}_{\mathcal{R}}$	spatial center of the reference image
$\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}; \boldsymbol{\Sigma})$	normal distribution of \mathbf{x} with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$\text{Tr}(\mathbf{M})$	trace of matrix \mathbf{M}
\mathbf{s}_j	j -th training shape representation (surface mesh) for an ASM
\mathbf{S}	matrix of all training shapes
φ_i	i -th eigenvalue of the covariance matrix of all training shapes
\mathbf{v}_i	i -th eigenvector of the covariance matrix of all training shapes
$d_{\text{SE}}(\Gamma_i, \mathbf{x})$	signed Euclidean distance transform of the shape image Γ_i at position \mathbf{x}
κ	curvature
$\mathcal{P}(\phi, t)$	cylinder projection image at radial position ϕ and the position t along the cylinder axis

Appendix C

Abbreviations

1-D	one-dimensional
2-D	two-dimensional
3-D	three-dimensional
ASM	active shape model
AX	C-arm computed tomography
B-B	background-background intensity combinations in the joint PDF
CBF	cerebral blood flow
CC	cross correlation
CPU	central processing unit
CSSD	curvature-extended sum of squared differences
CR	correlation ratio
CT	computed tomography
CUDA	compute unified device architecture
DICOM	digital imaging and communications in medicine
DOF	degrees of freedom
DRR	digitally reconstructed radiograph
DSA	digital subtraction angiography
EEG	electroencephalography
FN	number of false negative classifications
GM	Gaussian maps
GPU	graphics processing unit
HU	Hounsfield units
ICP	iterative closest point
i.i.d.	independent and identically distributed
ITK	insight segmentation and registration toolkit
KL	Kullback-Leibler
LM	landmark
MDL	minimum description length
MI	mutual information
MRI	magnetic resonance imaging

MSE	mean squared error
MSL	marginal space learning
NCC	normalized cross correlation
NN	nearest neighbor
NMI	normalized mutual information
O-O	object-object intensity combinations in the joint PDF
PCA	principal component analysis
PDE	partial differential equation
PDF	probability density function
PET	positron emission tomography
PIU	partition image uniformity
PVI	partial volume interpolation
TP	number of true positive classifications
TPS	thin-plate splines
TRE	target registration error
RIRE	retrospective image registration evaluation project
RIU	ratio image uniformity
SAD	sum of absolute differences
SE	sensitivity
SPECT	single photon emission computed tomography
SSD	sum of squared differences
s. t.	such that
SVD	singular value decomposition
US	ultrasound

Bibliography

- [Arun 87] K. S. Arun, T. S. Huang, and S. D. Blostein. “Least-Squares Fitting of Two 3-D Point Sets”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-9, No. 5, pp. 698–700, 1987.
- [Ashb 07] J. Ashburner. “A fast diffeomorphic image registration algorithm”. *NeuroImage*, Vol. 38, No. 1, pp. 95–113, 2007.
- [Ashb 99] J. Ashburner and K. J. Friston. “Nonlinear Spatial Normalization Using Basis Functions”. *Human Brain Mapping*, Vol. 7, No. 4, pp. 254–266, 1999.
- [Bai 09] W. Bai and M. Brady. “Regularized B-spline deformable registration for respiratory motion correction in PET images”. *Physics in Medicine and Biology*, Vol. 54, No. 9, pp. 2719–2736, 2009.
- [Beg 05] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes. “Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms”. *International Journal of Computer Vision*, Vol. 61, No. 2, pp. 139–157, 2005.
- [Besl 92] P. J. Besl and N. D. McKay. “A Method for Registration of 3-D Shapes”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, pp. 239–256, 1992.
- [Bish 97] C. M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford, 1997.
- [Brem 87] P. Brémaud. *An Introduction to Probabilistic Modeling*. Springer, New York, 1987. p. 130.
- [Brow 92] L. G. Brown. “A survey of image registration techniques”. *ACM Computing Surveys*, Vol. 24, No. 4, pp. 325–376, 1992.
- [Brox 07] T. Brox, B. Rosenhahn, D. Cremers, and H.-P. Seidel. “Nonparametric Density Estimation with Adaptive, Anisotropic Kernels for Human Motion Tracking”. In: A. M. Elgammal, B. Rosenhahn, and R. Klette, Eds., *Workshop on Human Motion*, pp. 152–165, Springer, Rio de Janeiro, Brazil, October 2007.
- [Bucc 05] M. K. Bucci, A. Bevan, and M. Roach. “Advances in Radiation Therapy: Conventional to 3D, to IMRT, to 4D, and Beyond”. *A Cancer Journal for Clinicians*, Vol. 55, No. 2, pp. 117–134, 2005.
- [Chan 02] D. J. Chang, I. G. Zubal, C. Gottschalk, A. Necochea, R. Stokking, C. Studholme, M. Corsi, J. Slawski, S. S. Spencer, and H. Blumenfeld. “Comparison of Statistical Parametric Mapping and SPECT Difference Imaging in Patients with Temporal Lobe Epilepsy”. *Epilepsia*, Vol. 43, pp. 68–74, 2002.

- [Chan 92] H. Chan and J. M. Fitzpatrick. “A Technique for Accurate Magnetic Resonance Imaging in the Presence of Field Inhomogeneities”. *IEEE Transactions on Medical Imaging*, Vol. 11, No. 3, pp. 319–329, 1992.
- [Chow 83] Y.-S. Chow, S. Geman, and L.-D. Wu. “Consistent Cross-Validated Density Estimation”. *The Annals of Statistics*, Vol. 11, No. 1, pp. 25–38, 1983.
- [Chri 01] G. E. Christensen and H. J. Johnson. “Consistent Image Registration”. *IEEE Transactions on Medical Imaging*, Vol. 20, No. 7, pp. 568–582, 2001.
- [Chri 06] G. E. Christensen, X. Geng, J. G. Kuhl, J. Bruss, T. J. Grabowski, I. A. Pirwani, M. W. Vannier, J. S. Allen, and H. Damasio. “Introduction to the Non-rigid Image Registration Evaluation Project NIREP”. In: J. P. W. Pluim, B. Likar, and F. A. Gerritsen, Eds., *Biomedical Image Registration, Third International Workshop, WBIR 2006, Proceedings*, pp. 128–135, Springer, Utrecht, The Netherlands, July 9-11 2006.
- [Chri 96] G. E. Christensen, R. D. Rabbit, and M. I. Miller. “Deformable Templates Using Large Deformation Kinematics”. *IEEE Transactions on Image Processing*, Vol. 5, No. 10, pp. 1435–1447, 1996.
- [Chri 97] G. E. Christensen, S. C. Joshi, and M. I. Miller. “Volumetric Transformation of Brain Anatomy”. *IEEE Transactions on Medical Imaging*, Vol. 16, No. 6, pp. 864–877, 1997.
- [Chri 98] G. E. Christensen. “MIMD vs. SIMD parallel processing: A case study in 3D medical image registration”. *Parallel Computing*, Vol. 24, No. 9-10, pp. 1369–1383, 1998.
- [Clar 06] U. Clarenz, M. Droske, S. Henn, M. Rumpf, and K. Witsch. *Computational methods for nonlinear image registration*, Chap. 1, pp. 81–101. Vol. 10 of *Mathematics in Industry*, Springer, Berlin Heidelberg, 2006.
- [Coot 95] T. Cootes, C. Taylor, D. Cooper, and J. Graham. “Active Shape Models - Their Training and Application”. *Computer Vision and Image Understanding*, Vol. 61, No. 1, pp. 38–59, 1995.
- [Csap 07] I. Csapo, C. Holland, and C. Guttman. “Image registration framework for large-scale longitudinal MRI data sets: strategy and validation”. *Magnetic Resonance Imaging*, Vol. 25, No. 6, pp. 889–893, 2007.
- [Dani 80] P. E. Danielsson. “Euclidean Distance Mapping”. *Computer Graphics and Image Processing*, Vol. 14, No. 3, pp. 227–248, 1980.
- [Davi 02a] R. Davies, C. Twining, T. Cootes, J. Waterton, and C. Taylor. “3D Statistical Shape Models Using Direct Optimisation of Description Length”. In: A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, Eds., *Computer Vision - ECCV 2002, 7th European Conference on Computer Vision, Proceedings, Part III*, pp. 3–20, Springer, Copenhagen, Denmark, May 28-31 2002.
- [Davi 02b] R. Davies, C. Twining, T. Cootes, J. Waterton, and C. Taylor. “A Minimum Description Length Approach to Statistical Shape Modelling”. *IEEE Transactions on Medical Imaging*, Vol. 21, No. 5, pp. 525–537, May 2002.

- [Deri 90] R. Deriche. “Fast Algorithms for Low-Level Vision”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, pp. 78–87, 1990.
- [Dros 03] M. Droske, M. Rumpf, and C. Schaller. “Non-rigid morphological image registration and its practical issues”. In: *IEEE International Conference on Image Processing, ICIP 2003, Proceedings*, pp. II: 699–702, Barcelona, Spain, September 2003.
- [Dros 04] M. Droske and M. Rumpf. “A Variational Approach to Non-Rigid Morphological Registration”. *SIAM Journal on Applied Mathematics*, Vol. 64, No. 2, pp. 668–687, 2004.
- [Dros 05] M. Droske. *On Variational Problems and Gradient Flows in Image Processing*. PhD thesis, Universität Duisburg-Essen, Fachbereich Mathematik, 2005.
- [Duda 01] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley & Sons, New York, 2001.
- [Egan 75] J. P. Egan. *Signal Detection Theory and ROC Analysis. Series in Cognition and Perception*, Academic Press, New York, 1975.
- [Enge 06] K. Engel, M. Hadwiger, J. M. Kniss, C. Rezk-Salama, and D. Weiskopf. *Real-Time Volume Graphics*. A K Peters, Ltd., Wellesley, MA, 2006.
- [Fisc 03a] B. Fischer and J. Modersitzki. “Combining landmark and intensity driven registrations”. *Proceedings in Applied Mathematics and Mechanics*, Vol. 3, No. 1, pp. 32–35, 2003.
- [Fisc 03b] B. Fischer and J. Modersitzki. “Curvature based image registration”. *Journal of Mathematical Imaging and Vision*, Vol. 18, No. 1, pp. 81–85, 2003.
- [Fisc 04] B. Fischer and J. Modersitzki. “A unified approach to fast image registration and a new curvature based registration technique”. *Linear Algebra and its Applications*, Vol. 380, pp. 107–124, March 2004.
- [Fitz 98] J. M. Fitzpatrick, J. B. West, and C. R. Maurer, Jr. “Predicting Error in Rigid-Body Point-Based Registration”. *IEEE Transactions on Medical Imaging*, Vol. 17, No. 5, pp. 694–702, 1998.
- [Fole 97] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes. *Computer Graphics: Principles and Practice. The Systems Programming Series*, Addison-Wesley, Boston, 1997.
- [Golu 96] G. H. Golub and C. F. van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, 1996.
- [Grad 06] L. Grady. “Random Walks for Image Segmentation”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 11, pp. 1768–1783, 2006.
- [Guet 05] C. Guetter, C. Xu, F. Sauer, and J. Hornegger. “Learning Based Non-rigid Multi-modal Image Registration Using Kullback-Leibler Divergence”. In: J. S. Duncan and G. Gerig, Eds., *Medical Image Computing and Computer-Assisted Intervention, MICCAI 2005, 8th International Conference, Proceedings, Part II*, pp. 255–262, Springer, Palm Springs CA, USA, October 2005.

- [Habe 06] E. Haber and J. Modersitzki. “Intensity Gradient Based Registration and Fusion of Multi-modal Images”. In: L. Rasmus, N. Mads, and J. Sporning, Eds., *Medical Image Computing and Computer-Assisted Intervention, MICCAI 2005, 8th International Conference, Proceedings, Part II*, pp. 726–733, Springer, Copenhagen, Denmark, October 2006.
- [Hahn 05] D. A. Hahn, J. Hornegger, W. Bautz, T. Kuwert, and W. Römer. “Unbiased rigid registration using transfer functions”. In: J. M. Fitzpatrick and J. M. Reinhardt, Eds., *Proceedings of SPIE on Medical Imaging*, pp. 151–162, San Diego CA, USA, April 2005.
- [Hahn 06] D. A. Hahn, G. Wolz, Y. Sun, J. Hornegger, F. Sauer, T. Kuwert, and X. Xu. “A Practical Salient Region Feature Based 3D Multi-Modality Registration Method for Medical Images”. In: J. M. Reinhardt and J. P. Pluim, Eds., *Proceedings of SPIE on Medical Imaging*, pp. 870–879, San Diego CA, USA, March 2006.
- [Hajn 01] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes. *Medical Image Registration*. CRC Press, Boca Raton FL, USA, 2001. The Biomedical Engineering Series.
- [Han 05] J. Han, J. Hornegger, T. Kuwert, W. Bautz, and W. Römer. “Feature Constrained Non-rigid Image Registration”. In: F. Hülsemann, M. Kowarschik, and U. Rüde, Eds., *Frontiers in Simulation*, pp. 638–643, Erlangen, 2005.
- [Hast 01] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer, New York, 2001.
- [Heim 05] T. Heimann, I. Wolf, T. G. Williams, and H. P. Meinzer. “3D Active Shape Models Using Gradient Descent Optimization of Description Length.”. In: G. E. Christensen and M. Sonka, Eds., *Information Processing in Medical Imaging, IPMI 2005, 19th International Conference, Proceedings*, pp. 566–577, Springer, Glenwood Springs, CO, USA, July 10-15 2005.
- [Heim 06] T. Heimann, I. Wolf, T. G. Williams, and H. P. Meinzer. “Optimal Landmark Distributions for Statistical Shape Model Construction”. In: J. M. Reinhardt and J. P. Pluim, Eds., *Proceedings of SPIE on Medical Imaging*, pp. 518–528, San Diego CA, USA, March 2006.
- [Held 04] S. Heldmann, O. Mahnke, D. Potts, J. Modersitzki, and B. Fischer. “Fast computation of Mutual Information in a variational image registration approach”. In: T. Tolxdorff, J. Braun, H. Handels, A. Horsch, and H.-P. Meinzer, Eds., *Proceedings of the Bildverarbeitung für die Medizin*, Springer, Berlin, 2004.
- [Hell 03] P. Hellier, C. Barillot, I. Corouge, B. Gibaud, G. Le Goualher, D. L. Collins, A. Evans, G. Malandain, N. Ayache, G. E. Christensen, and H. J. Johnson. “Retrospective evaluation of inter-subject brain registration”. *IEEE Transactions on Medical Imaging*, Vol. 22, No. 9, pp. 1120–1130, 2003.
- [Henn 02] S. Henn and K. Witsch. “Iterative Multigrid Regularization Techniques for Image Matching”. *SIAM Journal on Scientific Computing*, Vol. 23, No. 4, pp. 1077–1093, 2002.

- [Herm 02a] G. Hermosillo. *Variational Methods for Multimodal Image Matching*. PhD thesis, University of Nice Sophia Antipolis, France, 2002.
- [Herm 02b] G. Hermosillo, C. Chef d'Hôtel, and O. Faugeras. "Variational Methods for Multimodal Image Matching". *International Journal of Computer Vision*, Vol. 50, No. 3, pp. 329–343, 2002.
- [Hill 01] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes. "Medical image registration". *Physics in Medicine and Biology*, Vol. 46, No. 3, pp. R1–R45, 2001.
- [Horn 00] J. Hornegger and H. Niemann. "Probabilistic Modeling and Recognition of 3-D Objects". *International Journal of Computer Vision*, Vol. 39, No. 3, pp. 229–251, 2000.
- [Horn 02] J. Hornegger, V. Welker, and H. Niemann. "Localization and Classification Based on Projections". *Pattern Recognition*, Vol. 6, No. 35, pp. 1225–1235, 2002.
- [Horn 87] B. K. P. Horn. "Closed-form solution of absolute orientation using unit quaternions". *Journal of the Optical Society of America A*, Vol. 4, No. 4, pp. 629–642, 1987.
- [Horn 96] J. Hornegger. *Statistische Modellierung, Klassifikation und Lokalisation von Objekten*. PhD thesis, Universität Erlangen - Nürnberg, Lehrstuhl für Mustererkennung, 1996.
- [Huan 04] X. Huang, Y. Sun, D. Metaxas, F. Sauer, and C. Xu. "Hybrid Image Registration based on Configurational Matching of Scale-Invariant Salient Region Features". In: *IEEE Workshop on Image and Video Registration*, Washington DC, USA, July 2004.
- [Iban 05] L. Ibáñez, W. Schroeder, L. Ng, and J. Cates. *The ITK Software Guide*. Kitware Inc., second Ed., 2005.
- [Ino 05] F. Ino, K. Ooyama, and K. Hagihara. "A data distributed parallel algorithm for nonrigid image registration". *Parallel Computing*, Vol. 31, No. 1, pp. 19–43, 2005.
- [Josh 00] S. C. Joshi and M. I. Miller. "Landmark Matching via Large Deformation Diffeomorphisms". *IEEE Transactions on Image Processing*, Vol. 9, No. 8, pp. 1357–1370, 2000.
- [Katk 02] V. Katkovnik and I. Shumulevich. "Kernel density estimation with adaptive varying window size". *Pattern Recognition Letters*, Vol. 23, No. 14, pp. 1641–1648, 2002.
- [Kham 06] A. Khamene, R. Chisu, W. Wein, N. Navab, and F. Sauer. "A Novel Projection Based Approach for Medical Image Registration." In: J. P. W. Pluim, B. Likar, and F. A. Gerritsen, Eds., *Biomedical Image Registration, Third International Workshop, WBIR 2006, Proceedings*, pp. 247–256, Springer, Utrecht, The Netherlands, July 9-11 2006.
- [Knop 06] Z. F. Knops, J. B. A. Maintz, M. A. Viergever, and J. P. W. Pluim. "Normalized mutual information based registration using k -means clustering and shading correction". *Medical Image Analysis*, Vol. 10, No. 3, pp. 432–439, 2006.

- [Koo 03] C. W. Koo, O. Devinsky, K. Hari, J. Balasny, M. E. Noz, and E. L. Kramer. "Stratifying Differences on Ictal/Interictal Subtraction SPECT Images". *Epilepsia*, Vol. 44, No. 3, pp. 379–386, 2003.
- [Kybi 03] J. Kybic and M. Unser. "Fast Parametric Elastic Image Registration". *IEEE Transactions on Medical Imaging*, Vol. 12, No. 11, pp. 1427–1442, 2003.
- [Leon 89] M. J. de Leon, A. E. George, B. Reisberg, S. H. Ferris, A. Kluger, L. A. Stylopoulos, J. D. Miller, M. E. La Regina, C. Chen, and J. Cohen. "Alzheimer's disease: longitudinal CT studies of ventricular change". *American Journal of Roentgenology*, Vol. 152, No. 6, pp. 1257–1262, 1989.
- [Liao 03] Y. L. Liao, N. T. Chiu, C. M. Weng, and Y. N. Sun. "Registration and Normalization Techniques for Assessing Brain Functional Images". *Biomedical Engineering: Applications, Basis and Communications*, Vol. 15, pp. 87–94, June 2003.
- [Lloy 82] S. P. Lloyd. "Least Squares Quantization in PCM". *IEEE Transactions on Information Theory*, Vol. 28, No. 2, pp. 129–137, 1982.
- [Maes 97] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. "Multimodality image registration by maximization of mutual information". *IEEE Transactions on Medical Imaging*, Vol. 16, No. 2, pp. 187–198, April 1997.
- [Maes 98] F. Maes. *Segmentation and registration of multimodal medical images: From theory, implementation and validation to a useful tool in clinical practice*. PhD thesis, Catholic University of Leuven, 1998.
- [Main 98] J. B. A. Maintz and M. A. Viergever. "A survey of medical image registration". *Medical Image Analysis*, Vol. 2, No. 1, pp. 1–36, 1998.
- [Max 60] J. Max. "Quantizing for Minimal Distortion". *IRE Transactions on Information Theory*, Vol. 6, No. 1, pp. 7–12, 1960.
- [McNa 05] K. A. McNally, A. L. Paige, G. Varghese, H. Zhang, E. J. Novotny, S. S. Spencer, I. G. Zubal, and H. Blumenfeld. "Seizure localization by ictal and postictal SPECT". *Epilepsia*, Vol. 46, No. 9, pp. 1–15, 2005.
- [Meij 99] E. H. W. Meijering, W. J. Niessen, and M. A. Viergever. "Retrospective Motion Correction in Digital Subtraction Angiography: A Review". *IEEE Transactions on Medical Imaging*, Vol. 18, No. 1, pp. 2–21, 1999.
- [Mitc 87] D. P. Mitchell. "Generating Antialiased Images at Low Sampling Densities". In: M. C. Stone, Ed., *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1987*, pp. 65–72, ACM, New York, NY, USA, 1987.
- [Mode 03] J. Modersitzki and B. Fischer. "Optimal Image Registration with a Guaranteed One-to-one Point Match". In: T. Wittenberg, P. Hastreiter, U. Hoppe, H. Handels, A. Horsch, and H.-P. Meinzer, Eds., *Bildverarbeitung für die Medizin*, pp. 1–5, Springer Verlag, Erlangen, Germany, 2003.
- [Mode 04] J. Modersitzki. *Numerical Methods for Image Registration*. Oxford University Press, Oxford, 2004.

- [Muac 06] A. Muacevic, M. Staehler, C. Drexler, B. Wowra, M. Reiser, and J.-C. Tonn. “Technical description, phantom accuracy, and clinical feasibility for fiducial-free frameless real-time image-guided spinal radiosurgery”. *Journal of Neurosurgery: Spine*, Vol. 5, No. 4, pp. 303–312, 2006.
- [Muac 07] A. Muacevic, C. Drexler, B. Wowra, A. Schweikard, A. Schlaefer, R. T. Hoffmann, R. Wilkowski, H. Winter, and M. Reiser. “Technical Description, Phantom Accuracy, and Clinical Feasibility for Single-session Lung Radiosurgery Using Robotic Image-guided Real-time Respiratory Tumor Tracking”. *Technology in Cancer Research and Treatment*, Vol. 6, No. 4, pp. 321–328, 2007.
- [Mura 95] W. Murase and M. Lindenbaum. “Partial eigenvalue decomposition of large images using spatial temporal adaptive method”. *IEEE Transactions on Image Processing*, Vol. 4, No. 5, pp. 620–629, 1995.
- [NEMA 09] NEMA. “<http://medical.nema.org/>”. 2009.
- [Niem 83] H. Niemann. *Klassifikation von Mustern*. Springer, Heidelberg, 1983.
- [NIRE 09] NIREP. “<http://www.nirep.org/>”. 2009.
- [OBri 98] T. O’Brien, E. So, B. Mullan, M. Hauser, B. Brinkmann, N. Bohnen, D. Hanson, G. Cascino, C. J. Jr, and F. Sharbrough. “Subtraction ictal SPECT co-registered to MRI improves clinical usefulness of SPECT in localizing the surgical seizure focus”. *Neurology*, Vol. 50, No. 2, pp. 445–454, 1998.
- [OBri 99] T. J. O’Brien, E. L. So, B. P. Mullan, M. F. Hauser, B. H. Brinkmann, J. C. R. Jack, G. D. Cascino, F. B. Meyer, and F. W. Sharbrough. “Subtraction SPECT co-registered to MRI improves postictal SPECT localization of seizure foci”. *Neurology*, Vol. 52, No. 1, pp. 137–146, 1999.
- [Padb 99] M. Padberg. *Linear Optimization and Extensions*. Springer, Berlin, 1999.
- [Parz 62] E. Parzen. “On the estimation of probability density function and mode”. *Annals of Mathematical Statistics*, Vol. 33, No. 3, pp. 1065–1076, 1962.
- [Plui 00] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. “Interpolation artefacts in mutual information-based image registration”. *Computer Vision and Image Understanding*, Vol. 77, No. 9, pp. 211–232, 2000.
- [Plui 06] J. P. W. Pluim, B. Likar, and F. A. Gerritsen, Eds. *Biomedical Image Registration, Third International Workshop, WBIR 2006, Proceedings*, Springer, Utrecht, The Netherlands, July 9-11 2006.
- [Prum 06] M. Prümmer, L. Wigstroem, J. Hornegger, J. Boese, G. Lauritsch, N. Strobel, and R. Fahrig. “Cardiac C-arm CT: Efficient Motion Correction for 4D-FBP”. In: G. C. Smith, Ed., *Nuclear Science Symposium, Medical Imaging*, pp. 2620–2628, 2006.
- [Rata 00] A. L. Ratan, W. E. L. Grimson, and W. M. Wells. “Object Detection and Localization by Dynamic Template Warping”. *International Journal of Computer Vision*, Vol. 36, No. 2, pp. 131–147, 2000.
- [Rein 06] J. M. Reinhardt and J. P. Pluim, Eds. *Proceedings of SPIE on Medical Imaging*, San Diego CA, USA, March 2006.

- [Roch 98] A. Roche, G. Malandain, X. Pennec, and N. Ayache. “The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration”. In: *Medical Image Computing and Computer-Assisted Intervention, MICCAI 1998, 1st International Conference, Proceedings*, pp. 1115–1124, Springer Verlag, Cambridge, USA, October 1998.
- [Rohl 00] T. Rohlfing. *Multimodale Datenfusion für die bildgesteuerte Neurochirurgie und Strahlentherapie*. PhD thesis, Technical University Berlin, 2000.
- [Rohl 99] T. Rohlfing and J. Beier. “Improving Reliability and Performance of Voxel-Based Registration by Coincidence Thresholding and Volume Clipping”. In: D. J. Hawkes, D. L. G. Hill, and R. Gaston, Eds., *Proceedings of Medical Image Analysis and Understanding*, pp. 165–168, King’s College, 1999.
- [Rohr 01] K. Rohr. *Landmark based image analysis*. Kluwer Academic, Dordrecht, 2001.
- [Ruec 99] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. “Non-rigid registration using free-form deformations: Application to breast MR images”. *IEEE Transactions on Medical Imaging*, Vol. 18, No. 8, pp. 712–721, 1999.
- [Salo 05] M. Salomon, F. Heitz, G.-R. Perrin, and J.-P. Armspach. “A massively parallel approach to deformable matching of 3D medical images via stochastic differential equations”. *Parallel Computing*, Vol. 31, No. 1, pp. 45–71, 2005.
- [Schm 08] D. Schmidt, A. Szikszai, R. Linke, and T. Kuwert. “Impact of I-131-SPECT/spiral-CT on nodal staging of differentiated thyroid carcinoma at first radioablation”. *Journal of Nuclear Medicine*, Vol. 50, No. 1, pp. 18–23, 2008.
- [Scho 66] P. H. Schönemann. “A Generalized Solution of the Orthogonal Procrustes Problem”. *Psychometrika*, Vol. 31, No. 1, pp. 1–10, 1966.
- [Seth 99] J. A. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge University Press, Cambridge, 1999.
- [Shan 48] C. E. Shannon. “A Mathematical Theory of Communication (parts 1 and 2)”. *The Bell System Technical Journal*, Vol. 27, pp. 379–423, 623–656, 1948.
- [Span 99] M. V. Spanaki, S. S. Spencer, M. Corsi, J. MacMullan, J. Seibyl, and I. G. Zubal. “Sensitivity and Specificity of Quantitative Difference SPECT Analysis in Seizure Localization”. *Journal of Nuclear Medicine*, Vol. 40, No. 5, pp. 730–736, 1999.
- [Stud 99] C. Studholme, D. L. G. Hill, and D. J. Hawkes. “An overlap invariant entropy measure of 3D medical image alignment”. *Pattern Recognition*, Vol. 32, No. 1, pp. 71–86, 1999.
- [Thev 00] P. Thévenaz and M. Unser. “Optimization of Mutual Information for Multiresolution Image Registration”. *IEEE Transactions on Image Processing*, Vol. 9, No. 12, pp. 2083–2099, 2000.

- [Thev 08] P. Thévenaz, M. Bierlaire, and M. Unser. “Halton Sampling for Image Registration Based on Mutual Information”. *Sampling Theory in Signal and Image Processing*, Vol. 7, No. 2, pp. 141–171, 2008.
- [Thev 98] P. Thévenaz, U. E. Ruttiman, and M. Unser. “A Pyramid Approach to Subpixel Registration Based on Intensity”. *IEEE Transactions on Image Processing*, Vol. 7, No. 1, pp. 27–41, 1998.
- [Thir 98] J.-P. Thirion. “Image matching as a diffusion process: an analogy with Maxwell’s demons”. *Medical Image Analysis*, Vol. 2, No. 3, pp. 243–260, 1998.
- [Tikh 77] A. N. Tikhonov and V. Y. Arsenin. *Solutions of ill posed problems*. Wiley, New York, 1977.
- [Tsao 03] J. Tsao. “Interpolation Artifacts in Multimodality Image Registration Based on Maximization of Mutual Information”. *IEEE Transactions on Medical Imaging*, Vol. 22, No. 7, pp. 854–864, 2003.
- [Tsui 00] B. M. W. Tsui, W. P. Segars, and D. S. Lalush. “Effects of Upward Creep and Respiratory Motion in Myocardial SPECT”. *IEEE Transactions on Nuclear Science*, Vol. 47, No. 3, pp. 1192–1195, 2000.
- [Unse 93a] M. Unser, A. Aldroubi, and M. Eden. “B-Spline Signal Processing: Part I - Theory”. *IEEE Transactions on Signal Processing*, Vol. 41, No. 2, pp. 821–833, 1993.
- [Unse 93b] M. Unser, A. Aldroubi, and M. Eden. “B-Spline Signal Processing: Part II - Efficient Design and Applications”. *IEEE Transactions on Signal Processing*, Vol. 41, No. 2, pp. 834–848, 1993.
- [Ursc 07] M. Urschler, S. Kluckner, and H. Bischof. “A Framework for Comparison and Evaluation of Nonlinear Intra-Subject Image Registration Algorithms”. *The Insight Journal - 2007 MICCAI Open Science Workshop*, July-December 2007.
- [van 93] P. A. van den Elsen, J. B. A. Maintz, E. Pol, and M. A. Viergever. “Medical image matching – a review with classification”. *IEEE Engineering in Medicine and Biology*, Vol. 12, No. 4, pp. 26–39, 1993.
- [Verc 07] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. “Non-parametric Diffeomorphic Image Registration with the Demons Algorithm”. In: N. Ayache, S. Ourselin, and A. J. Maeder, Eds., *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2007, 10th International Conference*, pp. 319–326, Springer, Brisbane, Australia, October 29 - November 2 2007.
- [Viol 95] P. Viola. *Alignment by Maximization of Mutual Information*. PhD thesis, Massachusetts Institute of Technology (MIT), Cambridge, USA, 1995.
- [Wach 06] A. Wächter and L. T. Biegler. “On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming”. *Mathematical Programming*, Vol. 106, No. 1, pp. 25–57, 2006.
- [Well 86] W. M. Wells III. “Efficient Synthesis of Gaussian Filters by Cascaded Uniform Filters”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 2, pp. 234–239, 1986.

- [Well96] W. M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. "Multi-modal volume registration by maximization of mutual information". *Medical Image Analysis*, Vol. 1, No. 1, pp. 35–51, 1996.
- [West97] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, Jr., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. G. Hill, D. J. Hawkes, C. Studholme, J. B. A. Maintz, M. A. Viergever, G. Malandain, X. Pennec, M. E. Noz, G. Q. Maguire, Jr., M. Pollack, C. A. Pelizzari, R. A. Robb, D. Hanson, and R. P. Woods. "Comparison and Evaluation of Retrospective Intermodality Brain Image Registration Techniques". *Journal of Computer Assisted Tomography*, Vol. 21, No. 4, pp. 554–566, 1997.
- [Will87] E. L. William and E. C. Harvey. "Marching Cubes: A high resolution 3D surface construction algorithm.". *Computer Graphics*, Vol. 21, No. 4, pp. 163–169, 1987.
- [Wood92] R. P. Woods, S. R. Cherry, and J. C. Mazziotta. "Rapid automated algorithm for aligning and reslicing PET images". *Journal of Computer Assisted Tomography*, Vol. 16, No. 4, pp. 620–633, 1992.
- [Wood93] R. P. Woods, J. C. Mazziotta, and S. R. Cherry. "MRI-PET registration with automated algorithm". *Journal of Computer Assisted Tomography*, Vol. 17, No. 4, pp. 536–546, 1993.
- [Xu08] R. Xu, Y. W. Chen, S. Y. Tang, S. Morikawa, and Y. Kurumi. "Parzen-Window Based Normalized Mutual Information for Medical Image Registration". *IEICE Transactions on Information and Systems*, Vol. E91-D, No. 1, pp. 132–144, 2008.
- [Ye88] Q.-Z. Ye. "The signed Euclidean distance transform and its applications". In: *9th International Conference on Pattern Recognition*, pp. 495–499, IEEE Computer Society Press, Rome, Italy, November 1988.
- [Yosh99] S. Yoshizawa and K. Tanabe. "Dual differential geometry associated with the Kullback-Leibler information on the Gaussian distributions and its 2-parameter deformations". *SUT Journal of Mathematics*, Vol. 35, No. 1, pp. 113–137, 1999.
- [Zhen07] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu. "Fast Automatic Heart Chamber Segmentation from 3D CT Data Using Marginal Space Learning and Steerable Features". In: *IEEE 11th International Conference on Computer Vision, ICCV 2007, Proceedings*, pp. 1–8, IEEE, Rio de Janeiro, Brazil, October 14-20 2007.
- [Zoll06] L. Zöllei. *A Unified Information Theoretic Framework for Pair- and Group-wise Registration of Medical Images*. PhD thesis, Massachusetts Institute of Technology, January 2006.
- [Zuba95] I. G. Zubal, S. S. Spencer, K. Imam, J. Seibyl, E. O. Smith, G. Wisniewski, and P. B. Hoffer. "Difference Images Calculated from Ictal and Interictal Technetium-99m-HMPAO SPECT Scans of Epilepsy". *Journal of Nuclear Medicine*, Vol. 36, No. 4, pp. 684–689, 1995.

Index

- k*-NN, 32
- t*-test, 82

- acceptance rate, 86
- accumulation texture, 128, 130
- active shape model, 99
- alpha blending, 128, 130
- Alzheimer's disease, 8

- B-spline, 12, 33, 75
 - cubic kernel, 36
 - curves, 76
- background noise, 45, 81
- background threshold, 47, 49

- calculus of variations, 14
 - Gâteaux derivative, 14
- cell-based, 129
- cerebral blood flow, 89
- coincidence
 - masking, 46
 - thresholding, 46
 - weighting, 46, 47
- control grid, 12
- correspondence problem, 102, 104
 - registration, 102
- CPU, 68, 82, 107, 129
- cross-validation, 34, 109
- CUDA, 116
- curvature, 103
- curvature regularization, 15, 90, 105

- degrees of freedom, 12

- DICOM, 11
- diffeomorphic deformation model, 15
- digital subtraction angiography, 89
- digital subtraction angiography, 53
- digitally reconstructed radiograph, 115
- displacement field, 13
- distance transform, 18, 103, 104
- distortion correction, 87
- domain, 11, 102
 - overlap, 11, 55
- entropy, 22
 - joint, 23
- epilepsy, 89
- Euler-Lagrange equations, 14
- evaluation, 71, 76, 81, 85, 105, 106, 109
 - protocol, 95

- fitting
 - axis-aligned rectangle, 60
 - axis-aligned square, 61
 - circle, 62
 - cuboid, 63
 - cylinder, 64
 - polytope, 64
 - rectangle, 62
 - rectangular box, 63
 - uniqueness, 67
- fusion, 7

- Gaussian, 36, 73
- GPU, 53, 68, 125

- histogram

- adaptive layout, 42
 - binning, 28
 - discretization error, 43
 - number of bins, 43, 82
 - quasi-adaptive, 28
 - spatial, 74
 - trade-off, 43
- Hounsfield units, 19
- hybrid scanner, 7
- image pyramid, 34
- interpolation artifacts, *see* sampling
- iterative closest point, 17
- ITK, 68
- Jacobian, 20, 40
- jittering, *see* sampling
- kernel width, 27, 33, 34, 40, 48, 49
 - adaptive, 41
 - anisotropic, 34, 42
 - optimal, 35
 - quasi-adaptive, 42
 - spatial, 76
- kidney, 105
- Kullback-Leibler divergence, 72
 - analytical formulation, 74
- level set, 18, 20, 99, 102, 116
- log-likelihood, *see* objective function
- longitudinal studies, 8
- marching cubes algorithm, 105
- marching cubes algorithm, 105
- marginal space learning, 67
- marginalization, *see* projection
- mean squared error, 41, 106
- minimum description length, 99
- node-based, 129
- objective function
 - log-likelihood, 34, 74
 - log-likelihood derivative, 37
 - non-parametric registration, 13
 - parametric registration, 12
 - registration, 11
- OpenGL Shading Language, 68, 128
 - fragment shader, 128, 130
- overlap domain, 28
- partial volume interpolation, 31, 82
- partition of unity, 37
- Parzen-window, 27, 48, 74
 - adaptive, 41
 - B-spline, 33, 36, 75
 - derivative, 38
 - convolution, 33
 - data-driven, 28
 - discrete estimator, 35, 37
 - estimator, 32
 - Gaussian, 36
 - derivative, 37
 - kernel function, 27
 - kernel size, 28
 - spatial estimator, 74
- PDF, 19–22
 - estimation, 28
 - joint, 22
- principal component analysis, 101
- Procrustes problem, 16
- projection, 54, 125
 - circle, 57
 - cylinder, 58, 125
 - multi-level, 126
 - geometry, 55
 - marginalization, 54
 - orthogonal, 55, 56
- quantization, 42

- error trade-off, 42, 81
- Max-Lloyd, 42, 46
- noise, 42
- noise power, 42
- radiation therapy, 8
- random measure, 28
 - affine transform, 40
 - spatial sample, 73
- rasterization, 125
- ray casting, 57, 115, 125
- receiver operating characteristics, 85
- reconstruction, 9
- reference image, 11
- registration
 - bias, 90, 93
 - functional, 14
 - manual, 81, 86
 - non-parametric, 13, 90, 93, 115
 - parametric, 12
 - projection, 54
 - rigid, 90
 - subvolume, 55
- regularization energy, 14
- RIRE, 51, 81, 90, 116
- sampling
 - Halton, 32, 81
 - jittering, 28, 30, 31, 81, 82
 - quasi-random, 32
 - regular, 30
- seed point, 108
- segmentation, 72
 - database, 71
 - image, 102
 - manual, 75
 - random walk, 78
- sensitivity, 106
- similarity measure, 11
 - CC, 19
 - CR, 22
 - CSSD, 103, 106
 - feature-based, 15
 - landmarks, 16
 - MI, 22
 - NCC, 19
 - NMI, 23, 81, 90, 115
 - PIU, 20
 - RIU, 20
 - SAD, 18
 - SSD, 18, 103
 - derivative, 55
 - statistical, 27
 - surfaces, 18
- smoother, *see* regularization energy
- spatial distribution, 72
- spatial transform, 11
- structured noise, 45
- subtraction imaging, 89
- surface mesh, 100, 102, 105
- SVD, 17
- target registration error, 81, 83
- template image, 11
- texture coordinate, 130
- thin plate splines, 17
- transform
 - affine, 12, 105
 - B-spline, 76
 - cubic B-spline, 12
 - rigid, 12, 115
- variance, 21
- variation, 101
- variational framework, 13
- volume rendering, 58, 129

voxel coordinate, 129