

Normalization of Magnetic Resonance
Images and its Application to the
Diagnosis of the Scoliotic Spine

Der Technischen Fakultät der
Universität Erlangen–Nürnberg

zur Erlangung des Grades

DOKTOR–INGENIEUR

vorgelegt von

Florian Jäger

Erlangen — 2010

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

Tag der Einreichung:	21.04.2010
Tag der Promotion:	15.07.2010
Dekan:	Prof. Dr.-Ing. R. German
Berichterstatter:	Prof. Dr.-Ing. J. Hornegger Prof. D. Rückert, Ph.D

Normalisierung von
Magnetresonanzaufnahmen und ihre
Anwendung auf die Diagnose der
skoliotischen Wirbelsäule

Der Technischen Fakultät der
Universität Erlangen–Nürnberg

zur Erlangung des Grades

DOKTOR–INGENIEUR

vorgelegt von

Florian Jäger

Erlangen — 2010

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

Tag der Einreichung:	21.04.2010
Tag der Promotion:	15.07.2010
Dekan:	Prof. Dr.-Ing. R. German
Berichterstatter:	Prof. Dr.-Ing. J. Hornegger Prof. D. Rückert, Ph.D

Abstract

Due to its excellent soft tissue contrast and novel innovative acquisition sequences, Magnetic Resonance Imaging has become one of the most popular imaging modalities in health care. However, associated acquisition artifacts can significantly reduce image quality. Consequently, these imperfections can disturb the assessment of the acquired images. In the worst case, they may even lead to false decisions by the physician. Moreover, they can negatively influence an automatic processing of the data, e.g., image segmentation or registration. The most commonly observed artifacts are intensity inhomogeneities and a missing sequence-dependent general intensity scale. In this thesis, several novel techniques for the correction of the intensity variations are introduced. Further on, we demonstrate their advantages in a clinical application.

Many state-of-the-art approaches for correction of inhomogeneities lack either generalizability, efficiency, or accuracy. We present novel methods that overcome these drawbacks by introducing prior knowledge in the objective function and by mapping the optimization process onto a divide-and-conquer like strategy. The experiments show that we can increase the average separability of tissue classes in clinical relevant 3-d angiographies by approximately 18.2% whereas state-of-the-art methods could only achieve 11.6%.

The mapping of the intensities of a newly acquired image to a general intensity scale has to preserve the structural characteristics of the image's histogram. Further, it has to be invertible. Hence, many standardization approaches estimate a rather coarse intensity transformation. We propose several methods for standardization that are closely related to image registration techniques. These methods compute a per-intensity mapping. In addition, the methods presented are the only ones known that do a joint standardization and that can handle images with a very large field-of-view. The experiments show that our method achieves an average intensity overlap of the major tissue classes of T1w images of about 86.2%. The most commonly used state-of-the-art method resulted in only 70.1% overlap.

In order to illustrate the applicability and importance of the proposed normalization techniques, we introduce a system for the computer-aided assessment of anomalies in the scoliotic spine. It is based on the segmentation of the spinal cord using Markov random field theory. All required steps are presented, from the pre-processing to the visualization of the results. In order to evaluate the system, we use the angle between automatically computed planes through the vertebrae and planes estimated by medical experts. This results in a mean angle difference of less than six degrees being accurate enough to be applicable in a clinical environment.

Kurzzusammenfassung

Auf Grund des hervorragenden Weichteilkontrasts und neuen innovativen Aufnahme-sequenzen wurde die Magnetresonanztomographie zu einer der meist verwendeten bildgebenden Modalität im modernen Gesundheitswesen. Artefakte können allerdings in den Aufnahmen eine stark verminderte Bildqualität bewirken. Die Störungen erschweren die Sichtung der Daten und können schlimmstenfalls sogar zu falschen Entscheidungen des Radiologen führen. Darüber hinaus beeinflussen diese eine automatische Weiterverarbeitung, z.B. eine Segmentierung oder eine Bildregistrierung, auf eine negative Art und Weise. Die am häufigsten beobachteten Artefakte sind Intensitätsvariationen innerhalb eines Bildes und zwischen mehreren Aufnahmen. In dieser Arbeit stellen wir neue Techniken zur Korrektur dieser vor und demonstrieren deren Anwendung in einer klinischen Applikation.

Viele aktuelle Ansätze zur Korrektur von Inhomogenitäten mangelt es entweder an Generalisierbarkeit, Effizienz oder Präzision. Die präsentierten Methoden lösen diese Probleme durch die Integration von a-priori Wissen und durch die Abbildung der Optimierung auf eine Teile-und-Herrsche Strategie. Die Experimente zeigen, dass unsere Ansätze eine durchschnittliche Verbesserung der Separierbarkeit von Gewebeklassen in 3-d Angiographien um ca. 18.2% erreichen. Im Vergleich hierzu erlangen Standardtechniken nur 11.6%.

Die Abbildung der Intensitäten eines neu aufgenommenen Bildes auf eine Standardintensitätsskala muss die Struktur des ursprünglichen Histogramms erhalten. Weiterhin muss die Abbildung invertierbar sein. Daher berechnen viele Standardansätze eine grobe Intensitätstransformation. In dieser Arbeit stellen wir mehrere Techniken zur Intensitätsstandardisierung vor, die zu Methoden der Bildregistrierung verwandt sind. Darüber hinaus sind die vorgestellten Methoden die einzigen bekannten Ansätze, die gleichzeitig mehrere Wichtungen sowie einen großen Aufnahmebereich korrigieren können. Die Experimente zeigen, dass unsere Methoden eine durchschnittliche Überlappung der Gewebeklassen von ca. 86.2% erreichen. Der aktuell meistverwendete Ansatz kommt nur auf eine Überlappung von 70.1%.

Um die Anwendbarkeit der vorgestellten Methoden zu demonstrieren stellen wir ein System zur computergestützten Sichtung von Anomalien in der skoliotischen Wirbelsäule vor. Das System basiert auf einer Segmentierung des Spinalkanals durch Markov Zufallsfelder. Es werden alle Schritte, von der Vorverarbeitung bis hin zur Visualisierung der Daten, präsentiert. Zur Evaluierung verwenden wir den Winkel zwischen automatisch berechneten Ebenen zu Ebenen, die von einem medizinischen Experten bestimmt wurden. Die Winkelabweichung beträgt dabei weniger als sechs Grad. Dies ist für den klinischen Einsatz ausreichend.

Acknowledgment

I would like to use this opportunity to thank a couple of people that supported me during the last years. Most likely this work would not exist in this form without them.

Thank you:

Prof. Joachim Hornegger for putting your trust in me, and for convincing me to do my PhD at the Pattern Recognition Lab.

Dr. med. Rolf Janka for his support and for all his good advices.

Prof. Daniel Rückert for reviewing my work.

Prof. Elmar Nöth for always having an open ear for my problems and helping to solve them.

Prof. Paul Finn for giving me the opportunity to work in his lab at UCLA.

Eva Kollorz for enduring all my little idiosyncrasies as office mate, for being a very good friend with an open ear for all my problems, for helping me whenever possible, and for always offering good advice.

Marcus Prümmer for becoming a good friend, and fruitful discussions.

Simone Gaffling, Markus Mayer, Rüdiger Bock, Arne Militzer, Maria Polyanskaya, Hannes Hofmann, and Eva Rothgang for proofreading.

All colleagues for a very good time at the lab in the last years.

Last but not least my friends, my girlfriend Nadine, and my family for their support.

Additionally, I gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German National Science Foundation (DFG) in the framework of the excellence initiative.

Moreover, I gratefully acknowledge funding of the Bavaria California Technology Center (BaCaTeC).

Florian Jäger

Contents

1	Introduction	1
1.1	Imaging in Medicine	1
1.2	Influence of Artifacts on the Diagnosis and the Computerized Analysis	2
1.3	Contribution to the Progress of Research	5
1.4	Structure of this Work	6
2	Magnetic Resonance Imaging Theory	9
2.1	Acquisition	9
2.1.1	Spin and Magnetization	10
2.1.2	Spatial Encoding	11
2.1.3	Pulse Sequences	13
2.2	Reconstruction	17
2.2.1	k-Space	18
2.2.2	Parallel Acquisition Techniques	19
2.3	Artifacts	24
2.3.1	Motion Artifacts	24
2.3.2	Aliasing	25
2.3.3	Chemical Shift Artifacts	25
2.3.4	Truncation Artifacts	25
2.3.5	Magnetic Susceptibility	26
2.3.6	Intensity Non-uniformity	26
2.4	Summary	27
3	Correction of Signal Intensity Inhomogeneities	29
3.1	Problem Definition	29
3.2	State-of-the-Art	32
3.3	A Divide-and-Conquer Approach to Bias Correction	36
3.3.1	Model Definition	36
3.3.2	Divide-Step	38
3.3.3	Conquer-Step	40
3.4	Entropy Optimization using a Histogram-based Regularizer	44
3.4.1	Entropy Minimization	44
3.4.2	Histogram-based Regularization	45
3.5	Experiments & Results	47
3.5.1	Divide-and-Conquer-based Bias Correction	47
3.5.2	Histogram-based Regularization	58

3.6	Application to Other Modalities	64
3.6.1	Microslice Photographs	64
3.6.2	Retina Fundus Images	65
3.7	Summary & Discussion	68
4	Standardization of MRI images	71
4.1	Problem Definition	72
4.2	State-of-the-Art	72
4.3	Matching of 1-D Histograms	74
4.3.1	Affine Alignment of 1-D Histograms	75
4.3.2	Non-Rigid Alignment of 1-D Histograms	76
4.4	Non-Rigid Alignment of n-D Histograms	79
4.5	Standardization of Whole-Body MRI scans	82
4.6	Experiments & Results	85
4.6.1	1-d Histogram Alignment	85
4.6.2	N-d Histogram Alignment	95
4.6.3	Standardization of Whole Body Data Sets	105
4.7	Summary & Discussion	112
5	Computer-aided Assessment of Anomalies in the Scoliotic Spine	115
5.1	Motivation	115
5.2	State-of-the-art	116
5.3	Pre-processing	116
5.4	Segmentation using MRFs	119
5.4.1	Initialization	119
5.4.2	Iteration Step	119
5.5	Labeling of the Vertebrae and Visualization	124
5.6	Experiments & Results	125
5.7	Summary & Discussion	126
6	Outlook	129
7	Summary	131
A	Mathematical Symbols	137
B	Derivative of the Jeffrey Divergence	141
	List of Figures	143
	List of Tables	145
	Bibliography	147
	Index	155

Chapter 1

Introduction

Un bon croquis vaut mieux qu'un long discours.

Napoleon Bonaparte

1.1 Imaging in Medicine

There is little doubt that imaging of the morphology and functionality of the human body has revolutionized health care. Nowadays it is possible to examine cells on a microscopic level, visualize their metabolism and even record a specific “fingerprint” to identify them. On the other hand, physicians are able to create morphologic 3-d images of the human body as a whole and distinguish between the functionality of malign and benign tissues on a macroscopic level.

This has led to various tremendous improvements in health care. New possibilities for acquisition and visualization have enabled physicians and engineers to develop novel techniques for the diagnosis and treatment which would not have been feasible without modern imaging systems. Cancer diagnostics is a very impressive example for this. For instance, radiotherapy would hardly be possible without using the information about the exact location of malign target structures.

Beyond that imaging has not only increased the quality of patient care, but also reduced the costs for health care. The reason for this is that there are much less unnecessary procedures performed. Additionally, many diseases can be diagnosed in a much earlier stage of the progression, since it is not necessary that symptoms show on the outside of the human body, as it used to be. Moreover, if a disease is detected, procedures can be performed much less invasive.

In particular, *Magnetic Resonance Imaging* (MRI) gains more and more importance in diagnostics and treatment. The reason for this is that it offers a very good imaging resolution. This holds especially for structures containing soft tissue materials. Additionally, it has the advantage that there are no side effects known. For instance, in other modalities like computed tomography (CT) ionizing radiation is used that can cause irreparable damage to cells. MRI is based on the discovery of nuclear magnetic resonance (NMR) by Felix Bloch and Edward M. Purcell in 1946. Both received the Nobel price for their work in 1952. The introduction of field gradients by Paul C. Lauterbur (Nobel price together with Peter Mansfield in 2003) in

the early 1970s enabled spatial encoding and thus imaging using MRI. Consequently, more and more acquisition protocols for medical applications were developed since.

Nowadays, MRI is applied in nearly all areas of medical diagnostics and research. Still most prominent are morphological acquisitions. In this field MRI has become indispensable for many kinds of application because of its outstanding soft tissue contrast. It gives an excellent insight in all body regions, from the imaging of the knee to whole-body examinations of tumor patients and the assessment of anomalies in the spine.

On top of this, due to new innovative acquisition protocols more and more functional images are acquired. Functional MRI (fMRI), for instance, bases on the change of local blood flow in the vasculature induced by brain activity. Because of the associated reduction of the paramagnetic deoxyhemoglobin, the change of image contrast can be used to visualize the centers of activity in the brain [Buxt 02].

Diffusion MRI, describing diffusion weighted imaging (DWI) and diffusion tensor imaging (DTI) [Fill 09], relies on the local characteristics of the diffusion of water molecules. E.g., DTI uses the diffusion properties of tissues containing fibrous structures like muscle fibres in the heart or axons of white matter in the brain. The diffusion in the fibers along its main axis is largest. Thus, the diffusion can be measured in several directions and stored in a tensor at each voxel location. The gained information can be used for tractography afterwards, for instance.

Magnetic resonance angiography (MRA), like time-of-flight imaging also called inflow angiography [Pike 92], exploits the unsaturated inflowing blood into the saturated static tissue to create image contrast. Furthermore, angiographies can also be done by phase-contrast MRA (PC-MRA) and contrast-enhanced MRA (CE-MRA) that utilizes MRI contrast agents. Arterial spin labeling (ASL), on the other hand, labels fluids before they enter the slice to be measured. Using this, perfusion maps of the brain can be calculated. Recently, strategies for other regions, like the kidneys [Mart 04, Jank 08, Jank 09, Uder 09], were introduced as well.

1.2 Influence of Artifacts on the Diagnosis and the Computerized Analysis

Despite all the progress in imaging technology, a severe problem for using these devices for diagnostics and treatment planning are associated imaging artifacts. The main issue is that these image imperfections can influence the diagnosis or in the worst case even lead to false decisions by the physician.

If possible, the sources of artifacts have to be understood and removed during the acquisition process. In many cases, however, this cannot be achieved due to physical, financial or time issues. Then, they have to be dealt with using retrospective correction methods. Artifacts have to be handled in all kinds of acquisition procedures, like among others CT, MRI, or optical methods like endoscopy. For many of the associated image distortions satisfactory solutions are already available, e.g., the interpolation of defect pixels in digital radiographies.

As the focus of the thesis lies on MRI, we concentrate on this modality because of the sheer mass of different acquisition techniques and related artifacts. The most

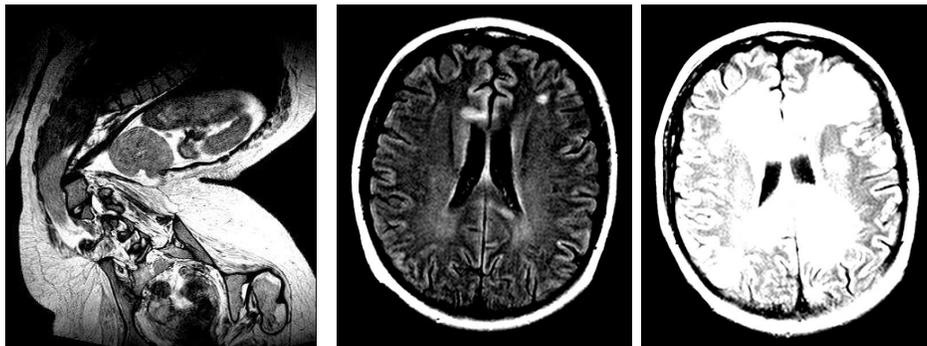


Figure 1.1: Difference between intra- and inter-image intensity variations. Left: Slice from a 3-d SPACE data set showing the spine. It can be observed that the middle part is brighter than the upper and lower area. Middle, Right: Two MRI FLAIR images of the same patient acquired in a distance of 6 months using the same acquisition protocol. Apart from the changed pathologies, large intensity differences can be observed. The transfer function in both images was set to center: 225, width: 150.

prominent acquisition artifacts in MRI are among many others *warp around* artifacts, *chemical shift* artifacts, *truncation* artifacts, and *signal intensity non-uniformities*. The sources of these and other common artifacts visible in MRI images are discussed in chapter 2 in more detail. Some of the artifacts can even simulate pathologies that are not visible in the artifact-free case. For instance, warp around artifacts can result in high signal intensities in some regions and might be mistaken for a pathology. Furthermore, artificial thickening of the optic nerves or pararenal fluid collections are simulated by *chemical shift* artifacts. However, many artifacts do not create false pathologies in the images, but make the diagnosis much more complicated for the radiologist. E.g. truncation artifacts can cause difficulties in the evaluation of the meniscus [Puse 86]. Most mentionable, though, are intensity non-uniformities, also denoted as signal intensity variations.

We differentiate two distinct types of signal intensity variations:

1. intra-image signal intensity variations and
2. inter-image signal intensity variations.

Both kinds of variations are illustrated in Fig. 1.1.

Intra-image signal intensity variations lead to intensity non-uniformities within a single image. Usually, these variations are very smooth. Thus, in many cases a human observer does not even recognize them. Their impact on an automatic processing of the images, on the other hand, is large. In literature, these variations are often denoted as intensity inhomogeneities caused by a gain or bias field. In some areas of application the correction of intensity inhomogeneities is called intensity equalization.

Inter-image signal intensity variations denote intensity changes between two images showing the same content but that are acquired on different scanners and/or at different time-points. The correction of these variations is denoted as intensity standardization. It requires the existence of a general intensity scale.

We denote the process of removing both kinds of variations as *normalization*. This is summarized in the following definition:

Definition 1 (*Normalization*)

The normalization of an image consists of (1) the correction of signal intensity inhomogeneities as well as (2) the standardization of the image intensities to a given general intensity scale.

A special case of intensity variations is induced by some interleaved MRI acquisition sequences. They create intensity differences in adjacent slices which significantly reduce the image quality if the volume is visualized in 3-d. The correction of these variations can be achieved using methods for intensity standardization. Furthermore, M. Schmidt introduced a fast and reliable correction method in [Schm05].

The inter- and intra-image signal intensity variations and thus the lack of a protocol dependent standard intensity scale makes it impossible to predefine standard transfer functions to visualize certain tissue classes. The radiologist has to perform the adjustment by himself in every single case and even for different regions within the images. This process can be very time consuming.

Moreover, it is often a necessity to further process acquired images to be able to provide the physician with an intuitive access to the images' content. The required post-processing can include multi-modal image registration to visualize the data of different modalities within the same coordinate system (image fusion), PET and MRI images for instance. Additionally, in many cases it is beneficial to segment certain defined structures within the acquired images to be able to conduct automatic measurements, like lesion sizes, or to support the physician in finding a diagnosis by automatically classifying anatomical structures.

Especially artifacts like intensity variations that change the appearance of structures within the images have a significant influence on the quality of the results of these image processing methods. For instance, in image registration, the methods are usually guided by intensity based difference measures. In general, a direct comparison of the intensities of two images is done using the sum-of-squared-differences (SSD), or a statistical measure describing the mutual statistical dependance between the images' intensities (mutual information, MI). In the case of the SSD measure, artifacts introducing intensity variations have a obvious influence on the resulting quality of the image registration. However, even statistical measures suffer from artifacts causing intensity variations, as these alter the intensity distributions of the images [Hahn09, Lotj10].

For most segmentation methods, the influence of image artifacts on the result is even more severe than in image registration. Many segmentation approaches assume that objects to be segmented have homogeneous intensity properties. This includes color information as well as information about their texturing. If these are altered due to imaging artifacts, many approaches will perform significantly worse [Zhug09] or even fail.

1.3 Contribution to the Progress of Research

The main contributions of this work can be split into three different categories: methods for bias correction, intensity standardization, and the utilization of MR normalization on a clinically relevant application dealing with computer-aided assessment of anomalies in the scoliotic spine. Although all methods were developed for MRI, they can be utilized in other fields of application, too. In the following, we will briefly introduce the major scientific contributions. For more information, we want to refer to the corresponding chapters.

Correction of signal intensity inhomogeneities

1. Divide-and-Conquer based approach for intensity inhomogeneity correction: A novel fast and robust method for bias correction. The main advantage, compared to other state-of-the-art algorithms, is its good generalization property. Thus, the approach can easily deal with weak and strong bias fields, it is applicable in 2-d as well as for 3-d problems, and it can handle complex bias fields.
2. Histogram-based regularization strategy for bias correction: Technique for a further regularization of existing methods to increase their reliability and their robustness against interfering structures in the images. Moreover, using the regularization technique, prior knowledge that is stored in histograms can be integrated into the optimization process.

Signal intensity standardization

1. Affine alignment of 1-d histograms: An approach to perform a very coarse matching of the intensity domains of two images. As the model consists of only two unknown parameters, it is very fast. Although the results are not very precise because of the model's simplicity, the results are very well suited as an initialization for more complex approaches.
2. Non-rigid alignment of 1-d histograms: Fast and precise method for intensity standardization. Since a non-parametric intensity transformation is estimated, it can handle small local intensity deviations in the histograms. The incorporation of a tunable regularization enables the adaptation of the approach to most problem domains. Moreover, it estimates a real-valued transformation.
3. Non-rigid alignment of n-d histograms: It is the first approach that standardizes all acquired MR images jointly by mapping the standardization to a non-rigid image registration problem. Thus, it can handle more complex intensity deviations than 1-d approaches are capable of. The method is integrated into the Slicer3 framework¹ by the National Alliance for Medical Image Computing.
4. Whole-body MRI intensity standardization: Images with a large *field-of-view* or images that are a composition of several acquisitions have very complex

¹http://www.na-mic.org/Wiki/images/f/fe/Slicer3Training_WhiteMatterLesions_v2.3.pdf

and often even spatially dependent intensity statistics. In order to deal with this problem we developed a novel distance measure that is integrated into the non-rigid alignment of n-d histograms.

Assessment of anomalies in the scoliotic spine

1. The assessment of anomalies in the scoliotic spine using MRI is an essential task during the planning phase of a patient's treatment and surgeries. Due to the pathologic bending of the spine, this is an extremely time consuming process, as an orthogonal view onto every vertebra is required. In this thesis, we present a system for computer-aided assessment of anomalies in 3-d MRI images of the spine relying on the segmentation of the spinal cord. We introduce all necessary steps, from the pre-processing of the data to the visualization component.

1.4 Structure of this Work

In this work, we will follow the data from the very basic physical principles of the acquisition modality, over the correction of artifacts that are an inevitable result of the used imaging techniques, to a sample application that relies on the presented pre-processing steps. The overall focus is put on MRI. However, all methods are not restricted to this modality.

In order to better understand the intensity distortions the thesis deals with, the basic theoretical foundations involved in forming an MR image are introduced in chapter 2. First, the physical principles, like nuclear spin, are presented. Then, simple acquisition protocols are explained. This is followed by basic reconstruction techniques and parallel image acquisition. The chapter is closed by a brief overview about various imaging artifacts that influence image quality.

Chapter 3 explicitly deals with intra-image signal intensity variations. In this field a lot of research was already done. For this reason, we start the chapter with a short overview and classification of state-of-the-art methods. Further on, we will introduce extensions to these methods that make the proposed methods more reliable and robust against pathological structures. Additionally, we will present a novel method that is able to solve some of the major disadvantages of many state-of-the-art approaches. Finally, an evaluation is presented that compares our methods to others that are commonly used.

In chapter 4 the correction of inter-image signal intensity variations is treated. In the beginning, the normalization of single images is presented. The normalization relies on the alignment of image histograms. In the following, this is extended to multi-channel images and to the standardization of whole-body MR images. The algorithms are evaluated and compared to state-of-the-art methods.

In chapter 5, we give an example in which the normalization techniques that were previously introduced are used in a clinically relevant application. We present a framework for the assessment of anomalies in the scoliotic spine. The whole processing chain is introduced, from the pre-processing of the acquired data sets to the final presentation of the results to the physician.

The thesis is concluded by a short outlook on possible next steps in the correction of intensity variations in chapter 6, and a brief summary of the the whole thesis in chapter 7.

Chapter 2

Magnetic Resonance Imaging Theory

In this chapter, we briefly introduce the theory of MRI. First, the physical background is explained including spatial encoding. Additionally, we present three imaging sequences used for experiments and examples throughout the thesis. The next section deals with algorithms for the reconstruction of the acquired MR data. Here, single coil as well as parallel reconstruction methods are treated. Finally, the chapter is concluded by a short summary of artifacts encountered in MRI. Because the intention of this chapter is just to overview the principles and the problems using MRI for *in-vivo* imaging, many vital details of MR theory are skipped. For further information, we refer to state-of-the-art literature (e.g., Brown and Semelka [Brow 03] or Oppelt [Oppe05]). Nevertheless, the chapter gives all the necessary information to understand the sources of intensity variations within the images. This understanding, however, is required to develop mathematical models that describe these variations in MR images. Without this, it is not possible to introduce required restrictions, like information about the smoothness of intensity inhomogeneities, into the optimization stage of the correction methods making a satisfactory solution feasible. Moreover, a good intuition helps to provide and integrate prior knowledge for the normalization of acquired MR images.

2.1 Acquisition

The human body consists of approximately 70% of bound and unbound water molecules. The most abundant isotope of hydrogen is the nucleus 1H . Its spin is $1/2$. For these reasons, usually the nucleus 1H is used for *in-vivo* imaging. However, other nuclei with a nuclear spin are used as well in MR spectroscopy for instance [Brow 03]. Although the theoretical physical basis of MRI relies on quantum mechanics, all necessary effects can be explained using a model of a proton described by a small bar magnet. In the quantum mechanic viewpoint, the spin of all stable atomic nuclei can only take discrete angles with respect to the magnetic field (e.g. for protons with spin $1/2$ two angles). Using the macroscopic viewpoint, the angles can be arbitrary. Due to this fact, spin mechanics can be described using the Bloch equations [Oppe05].

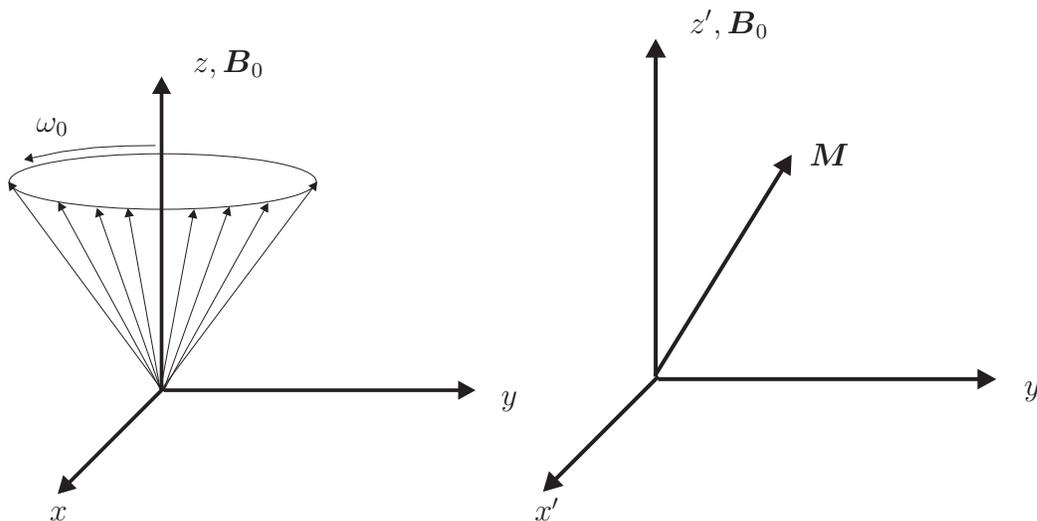


Figure 2.1: Left: Precession motion of the nuclear spins around the main magnetic field \mathbf{B}_0 with angular velocity ω_0 . Right: The magnetization vector \mathbf{M} in the rotating frame of reference.

2.1.1 Spin and Magnetization

If a sample of protons is exposed to an external magnetic field the orientation of their spins align parallel or anti-parallel to the direction of the magnetic field. The proportion of the direction of the spins is defined by the Boltzmann statistics [Brow 03]. Because of the lower energy level, there is a surplus of spins aligning parallel to the magnetic field yielding a magnetization $\|\mathbf{M}\|$ of the sample. $\|\mathbf{M}\|$ is the Euclidean norm of the magnetization vector $\mathbf{M} \in \mathbb{R}^3$. If the magnetization vector of the sample is not completely parallel to the magnetic field \mathbf{B}_0 , a torque is exerted on it. The torque is perpendicular to the direction of the magnetic field and the vector \mathbf{M} . Hence, the magnetization vector describes a precession motion around the main magnetic field \mathbf{B}_0 . The angular velocity of the spins is proportional to the Euclidean length of \mathbf{B}_0 [Brow 03] and can be computed as

$$\omega_0 = \gamma \|\mathbf{B}_0\|. \quad (2.1)$$

The angular velocity ω_0 is also called Larmor or resonance frequency. The proportionality constant connecting the main magnetic field \mathbf{B}_0 and the angular precession velocity ω_0 is called the gyromagnetic constant γ . The gyromagnetic ratio for the nucleus ^1H (proton with spin 1/2) is $\gamma/(2\pi) = 42.6 \text{ MHz/T}$. As the precession velocity is known, it is possible to switch from the static frame of reference with coordinate axis x , y and z to a rotating frame of reference with coordinate axis x' , y' and z' . The rotation axis is the direction of \mathbf{B}_0 that corresponds to the z -axis. As a result of the new rotating coordinate system, the magnetization vector does not rotate anymore around the main magnetic field \mathbf{B}_0 . This has the advantage that the following equation can be written w.r.t. a static coordinate system. The connection between the coordinate systems is illustrated in Fig. 2.1.

The basic idea behind MRI is to flip the sample's magnetization vector away from the parallel alignment to the main magnetic field. Subsequently the vector begins to realign with the main magnetic field. The realignment can be measured from the outside of the sample. The flipping is called excitation as the system is supplied by additional energy from an external source. In order to excite the spins, the energy pulse (radiofrequency pulse, RF pulse) has to match the resonance frequency of the sample defined by the gyromagnetic ratio (see Eq. (2.1)). In general, flip angles of 30° , 90° , or 180° are used. After excitation with a RF pulse, all spins precess in the same phase. Their moments add up to the signal induced in the receiving coil. Because of the influence of molecular magnetic fields caused by macro-molecules in the sample and susceptibility variations in the excited tissue, the spins lose their phase coherence. The transversal magnetization \mathbf{M}_t induces a signal in a receiver coil placed perpendicular to the transverse plane. This signal is called free induction decay (FID). The decay is characterized by the tissue dependent time constants T_2 and T_2^* . The decay caused by the molecular fields only is described by the constant T_2 . It is up to three times longer than T_2^* . In literature, this is often referred to as transverse relaxation time or spin-spin relaxation time.

A second very important tissue dependent time constant is T_1 . It describes the longitudinal relaxation or spin-lattice relaxation. It corresponds to the time that is needed to restore 63% of the z -component of the magnetization vector. The relaxation follows

$$M_z(t) = M_0(1 - e^{-t/T_1}) \quad (2.2)$$

with t being the elapsed time since the RF pulse and M_0 being the magnetization before the excitation. Some typical relaxation times for tissue classes relevant in the thesis are shown in the following table [Oppe 05]:

	T_1	T_2
kidney tissue	$0.59\text{s} \pm 0.16$	$T_2 = 58\text{ms} \pm 24$
gray matter	$T_1 = 0.81\text{s} \pm 0.14$	$T_2 = 101\text{ms} \pm 13$
white matter	$T_1 = 0.68\text{s} \pm 0.12$	$T_2 = 92\text{ms} \pm 22$

Note that T_2 times are much smaller than T_1 times.

2.1.2 Spatial Encoding

So far, only the excitation and the measurement of relaxation times of the whole sample was considered. For imaging purposes, a spatial encoding of the measurements has to be introduced.

Slice Selection

For simplicity we are considering axial slices only. However, arbitrary orientations are possible. The slice selection directly makes use of Eq. (2.1). In order to change the precessing frequency a magnetic gradient field G_z in z -direction is applied to the main magnetic field. Thus the total field strength is z -dependent and can be computed as

$$B_z(z) = \|\mathbf{B}_0\| + G_z z. \quad (2.3)$$

As the precession frequency depends on the field strength B_z , we have to apply a RF impulse including the frequencies $f_1 < f < f_2$ defined by the interval $z_1 < z < z_2$

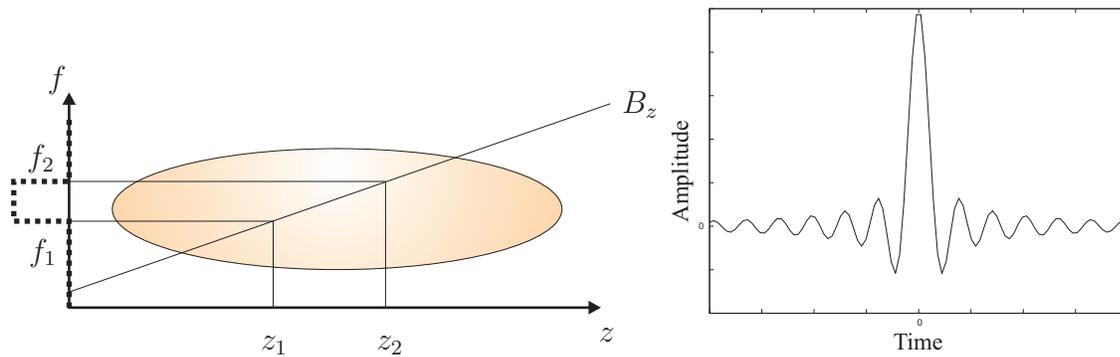


Figure 2.2: Illustration of the slice selection: The z -component B_z of the main magnetic field linearly varies due to the applied gradient. To excite a slice of slice thickness $z_1 < z < z_2$, the corresponding frequency profile has to be computed (dotted line). The profile has to be Fourier transformed. The result is a **sinc** function in time domain (plot on the right) that corresponds to the required RF pulse. In practice a truncated **sinc** is used.

to excite a slice with a slice thickness of $z_2 - z_1$. In the following, we assume that the center of the interval corresponds to $z = 0$. With that the field strength is equal to $\|\mathbf{B}_0\|$ in the center of the interval. The Fourier transform of the rectangular frequency spectrum corresponds to the $\text{sinc}(t) = \sin(t)/t$ function in time domain. The link between gradient field, rectangular spectrum, and the signal envelope of the RF pulse is illustrated in Fig. 2.2. A source for artifacts arises from the fact that in theory the **sinc** function has an infinite duration, in practice, however, it has to be cut off after the first or second lobe due to the protocol timing.

Frequency Encoding

Let us assume that a slice centered around $z = 0$ and parallel to the x, y plane has already been excited. Thus, all spins have been flipped down a defined angle and they are in-phase. Due to inhomogeneities in the magnetic field, the spins begin to dephase and the induced signal in the receiver coils gets reduced. From Eq. (2.1) it is known that the precession frequency is dependent on the local magnetic field strength. If the local magnetic field strength is altered, the spins in this region precess with a different angular velocity corresponding to the new field strength. In order to do spatial frequency encoding a linear gradient is applied to the main magnetic field. This gradient varies the magnetic field in x -direction. The direction of the gradient field is the z -axis. This can be formulated as

$$B_x(x) = \|\mathbf{B}_0\| + G_x x \quad (2.4)$$

with G_x being the gradient. If this gradient field is applied during the readout, it is obvious that the induced signal contains various frequencies, each corresponding to a spatial position along the gradient direction. The difference in angular velocity of the initial to the altered magnetic field strength for location x can be calculated as $\delta\omega = \gamma G_x x$. In the following we call the x -axis also the readout direction or the direction of the frequency encoding.

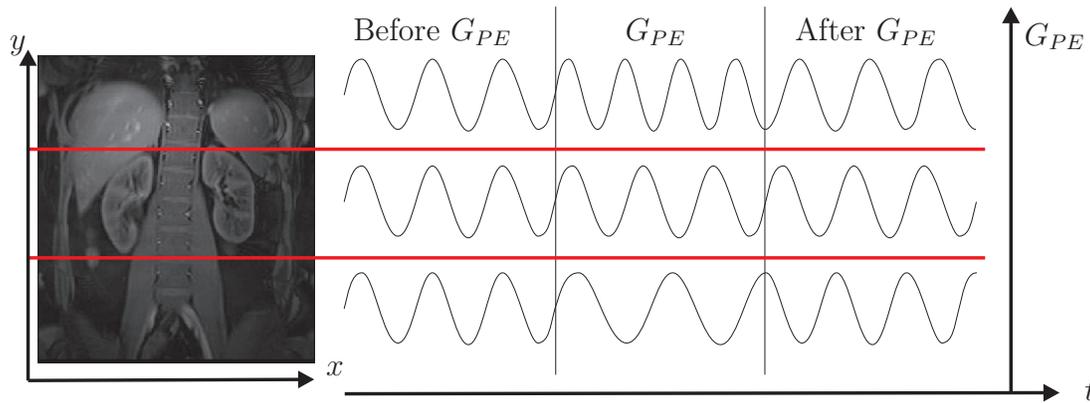


Figure 2.3: Illustration of phase encoding: Before the application of the gradient G_{PE} , all spins in one column precess in phase and with the same precession frequency. During the application, the frequency of the spins is dependent on their y -coordinate. After the gradient has been turned off again, the spins continue to precess with their original frequency but they have a phase difference.

Phase Encoding

Using the slice selection and frequency encoding two of the required three coordinates can be selected. In fact the result of the frequency encoding is the integral over the y -direction at a fixed x - and z -coordinate. In order to add further information to the excited slice an additional gradient G_{PE} is applied for a defined time interval before the readout. The gradient linearly alters the magnetic field along the y -direction. In the following, we will refer to this as the phase encoding direction. The field inhomogeneity induced by the gradient yields that the precession frequency is dependent on the y -coordinate. After the gradient has been turned off, the spins will continue to precess at the original frequency. However, the different precession frequencies yielded a y coordinate dependent phase shift of the spins. The magnitude of the phase shift is defined by the strength of the gradient G_{PE} and the time it is applied. The phase encoding principle is illustrated in Fig. 2.3.

2.1.3 Pulse Sequences

Pulse sequences define the acquisition of the MR images. For instance, they include the information about when to turn on a gradient or when to apply a RF impulse. We will introduce three MRI pulse sequences used throughout the thesis. However, there are a lot more important sequences used in clinical routine. Especially, sequences that attenuate certain tissue classes, like fluids in FLAIR (FLuid Attenuated Inversion Recovery) are widely used.

Spin Echo

One of the first imaging sequences was the Spin Echo sequence. It is directly derived from the position encoding techniques introduced in the previous section. In Fig. 2.4 the sequence timing diagram is shown. First, together with the slice selection

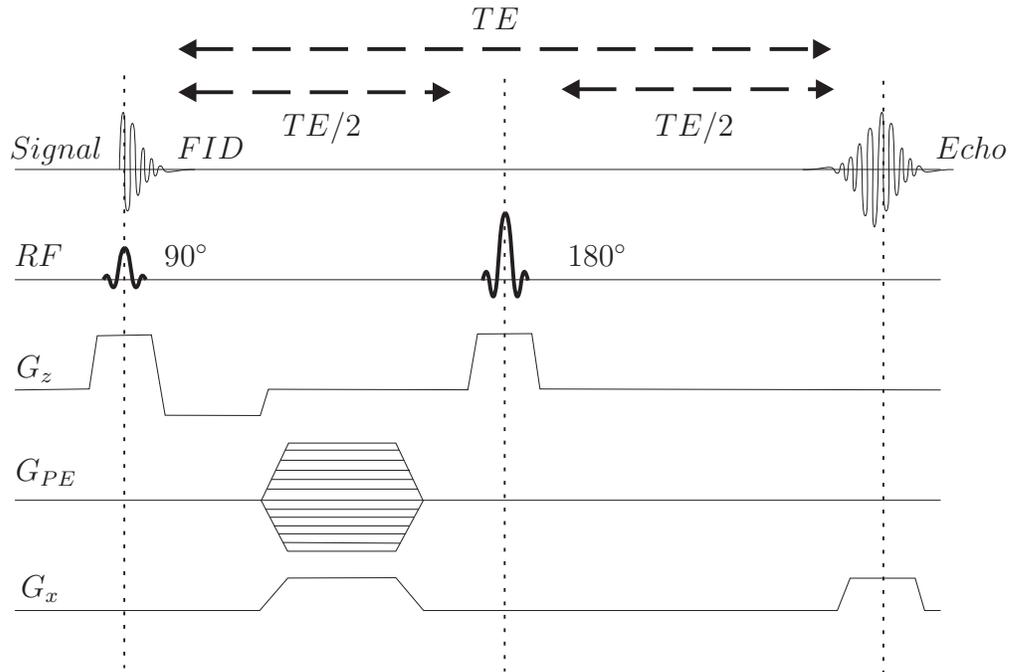


Figure 2.4: Spin echo timing diagram: The first line shows the expected signal induced in the receiving coil. In the second line the RF excitation pulse and the refocusing pulse are shown. Next, the timing of the slice selection gradient G_z is illustrated. This is followed by a phase encoding gradient G_{PE} . For every phase encoding step, schematically illustrated by the parallel lines, the whole diagram has to be repeated. Finally, the frequency encoding is illustrated.

gradient, a 90° RF pulse is applied. This pulse flips all spins and brings them into phase alignment. Next, the phase encoding gradient G_{PE} is applied. This changes the phases of the spins along the y -direction. In order to be able to do more precise measurements with respect to the fast T_2^* relaxation times a 180° refocusing RF pulse is applied after $TE/2$. TE is called echo time. This yields an echo of the signal after TE . The reason for this is that immediately after the 90° RF excitation pulse the spins start to de-phase depending on local magnetic inhomogeneities. If a 180° refocusing RF pulse is applied all spins are flipped by 180° but continue to precess with the same velocity in the same direction. Thus, the spins are in phase again after $TE/2$ has elapsed after the 180° pulse. This is illustrated in Fig. 2.4. During readout, the frequency encoding gradient is turned on and the signal is read/sampled using an analog to digital converter (ADC). Before continuing the acquisition using a differing phase encoding gradient G_{PE} to create data with another phase shift, one has to wait for a certain time period. The time period between two acquisitions is called repetition time (TR). This time interval is necessary as the magnetization has to return towards its equilibrium state before the next measurement can be performed.

The contrast in spin echo sequences is defined by the two timing parameters TE and TR . Between the initial excitation and the received echo the magnetization has decreased by a factor of $\exp(-TE/T_2)$ because of the T_2 decay. If the echo times are chosen to be much smaller than T_2 ($TE \ll T_2$), this has hardly any effect as the factor is close to one. If, on the other hand, TE and T_2 are in close proximity, the factor is dominated by the T_2 times of the underlying tissue. Thus, the resulting image is T_2 -weighted. The other important contrast parameter is the repetition time TR . If it is chosen to be smaller than the longitudinal relaxation time T_1 , the longitudinal magnetization cannot completely relax. This means, that the magnetization vector \mathbf{M} is not in its equilibrium state parallel to the main magnetic field \mathbf{B}_0 . This yields T_1 -weighted images. The recovered magnetization $M_{||}$ can be computed as

$$M_{||} = M_0(1 - e^{-TR/T_1}) \quad (2.5)$$

with M_0 being the magnetization directly before the excitation pulse. A summary of the influence of the contrast parameters is given in the following table:

	short TR	long TR
short TE	T_1 -weighted	PD-weighted
long TE	not used	T_2 -weighted

In Fig. 2.5 typical examples for T_1 and T_2 weighted images of the head are given.

Turbo Spin Echo

Turbo Spin Echo (TSE) sequences, also called Fast Spin Echo (FSE) are closely related to the standard SE sequence introduced before. Instead of just one single 180° refocusing pulse multiple refocusing pulses follow each 90° excitation pulse. These pulses produce multiple echoes of a single excitation. The echoes are called echo-train. To every refocusing pulse a different phase encoding gradient G_{PE} is applied. Thus, every echo can be used to acquire a phase encoded line. The length of the

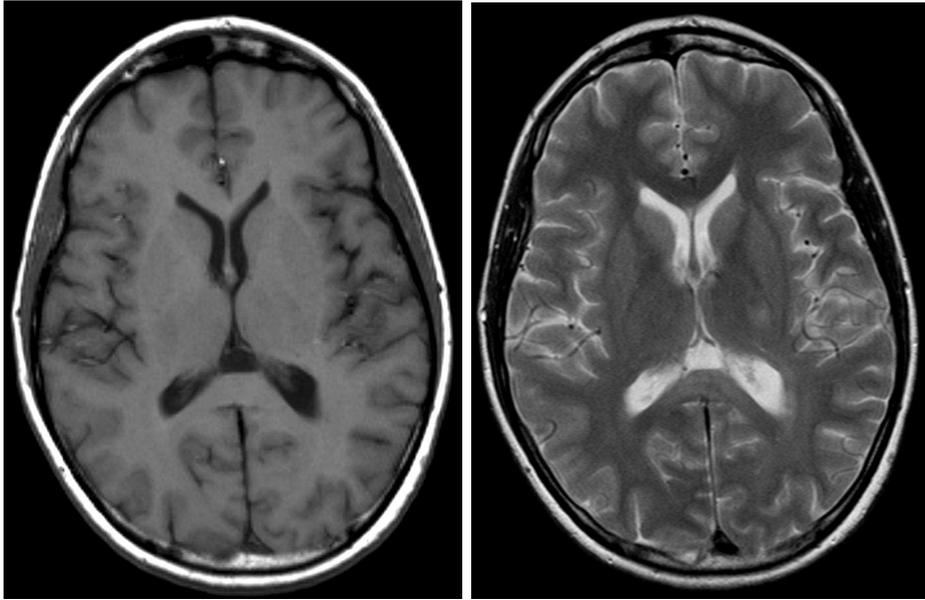


Figure 2.5: Left: T1-weighted image. Right: T2-weighted image. In the images the same patient is acquired with different imaging parameters. For instance, it can be seen that in the T1w image CSF appears dark whereas it appears bright in the T2w image. Both images are acquired at the Radiologic Institute, University of Erlangen.

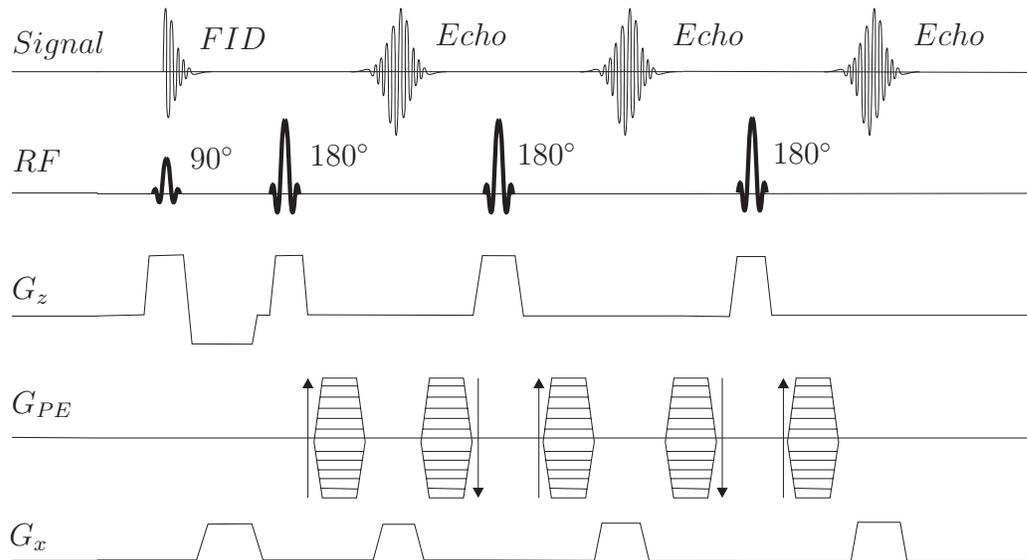


Figure 2.6: Turbo spin echo (TSE) timing diagram: The timing diagram corresponds to a TSE sequence with a echo train length of three. The first line illustrates the expected signal induced in the receiver coil. Then, the excitation RF pulse and the refocusing pulses are shown. In the third line the schematic timing of the slice selection gradient G_z is presented. This is followed by the schematic phase encoding gradient G_{PE} . Here, every gradient is followed by a reverse one indicated by opposite arrows. Finally, the readout gradient G_x is shown.

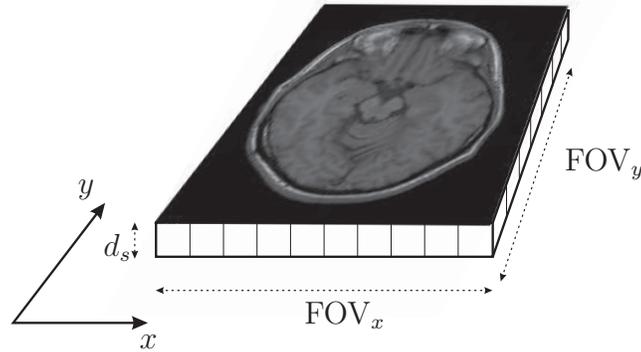


Figure 2.7: Field-of-view of an axial slice through the head.

echo-train (number of echoes) is called turbo factor. The timing diagram is shown in Fig. 2.6. As multiple echo times are used, the echo time that is used for acquiring the line with $G_{PE} = 0$ is called the effective TE. $G_{PE} = 0$ is used as it has the highest contribution to the image contrast.

There are several adapted versions of the TSE sequence. Later, we will deal with images acquired using *sampling perfection with application optimized contrasts using different flip angle evolutions* (SPACE) a 3-d TSE sequence. SPACE has extremely long echo trains. In order to achieve this, the refocussing pulses are no longer 180° but vary in flip angle. Consequently, high resolution 3-d images can be acquired in a very fast manner.

Arterial Spin Labeling

The principle of arterial spin labeling (ASL) is quite straight forward. In a first step, a slice to be imaged is selected. Then, the area below the slice to be imaged is labeled magnetically. This area contains the arteries. The labeling is done by an 180° inversion pulse. After the labeling the arterial blood flows into the selected slice. Due to the inversion of the spins of the blood the magnetization of the vessels is reduced. After some time, called the transit time, the slice is imaged. The result is called tagged image. Because of the reduced magnetization within the vessels, they have a lower signal intensity in the image. After all spins have reached their equilibrium state the slice is acquired again yielding the control image. By subtracting both images the perfusion of the tissue is computed.

Compared to other perfusion imaging methods, ASL has the advantage that no contrast bolus has to be given to the subject. This reduces problems regarding allergic reactions. Further, it is still an open research topic if contrast agents influence the metabolism of organs like the kidneys.

2.2 Reconstruction

The acquired MR image corresponds to a slice through the region of interest of the patient. The imaged part is called the field-of-view (FOV). It is parameterized by

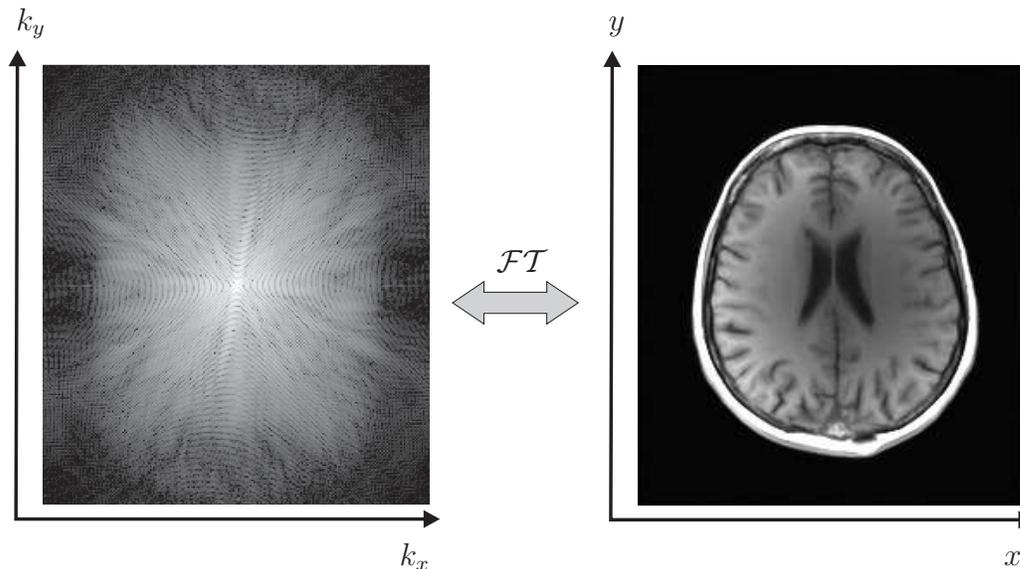


Figure 2.8: k -space / image domain correspondence: Left: the k -space image is shown. For presentation the logarithmical absolute values are shown. Right: The corresponding slice in image domain. The connection between both domains is the 2-d Fourier transform.

the position and orientation within the acquired object, the slice thickness d_s , the dimension in the readout direction, and the dimension in phase encode direction. The size of a voxel in x - and y -direction is defined by the number of samples during the readout N_{RO} , the number of phase encoding steps N_{PE} as well as the size of the FOV. Thus, the physical size s_v of a voxel can be computed as

$$s_v = d_s \cdot \text{FOV}_x / N_{RO} \cdot \text{FOV}_y / N_{PE} \quad (2.6)$$

where FOV_x is the size of the FOV in readout direction and FOV_y is the size in phase encoding direction respectively. The 3-d voxel in the acquired slice corresponds to a 2-d pixel in the image domain. In Fig. 2.7 an axial slice and its FOV is illustrated.

2.2.1 k -Space

The acquired raw frequency and phase data is stored in k -space. In the case of 2-d acquisition protocols like the spin echo sequence, it is a 2-d data matrix. The k_x -axis of the matrix represents the time component (readout direction) of the acquisition, the k_y -axis encodes the phase shift information (phase encode direction). The number of sample points N_{RO} in the readout direction is defined by the sampling in the ADC. In the y -direction the sampling is defined by the number of phase encoding steps N_{PE} . In order to convert the acquired data from k -space to the image domain, a 2-d Fourier transform of the data has to be performed. The correspondence between k -space and image domain is illustrated in Fig. 2.8.

In the center of k -space, the amplitude variations are dominant. Thus, these lines are primarily responsible for the contrast in the image. Signals acquired with

high amplitudes of G_{PE} contain high frequencies. For this reason, these k -space lines represent the edges mainly. They are located away from the center of k -space. This corresponds to basic image processing principles.

In order to reconstruct the images from the raw data, it is a necessity for today's that the k -space is completely filled and that the data points are equidistantly sampled. This means that distances Δk_x and Δk_y between two data points in k_x - and k_y -direction are constant. The number of sample points in k -space ($N_{PE} \cdot N_{RO}$), Δk_x and Δk_y define the resolution of the final image. There are several strategies to fill the k -space matrix. Acquisition sequences like the introduced spin echo sequence (see section 2.1.3), fill the k -space matrix linearly. This means that always a complete k -space line is acquired at once and written into the data matrix. However, other k -space trajectories are also used like a spiral trajectory. These acquisition trajectories can be generated by dynamically altering the imaging gradients [Brow 03].

2.2.2 Parallel Acquisition Techniques

A major disadvantage of MRI compared to other tomographic techniques like CT is the acquisition time. The limiting factor is the phase encoding step. After every acquired line one has to wait until the spins relax to their equilibrium state again. For example, a standard T1w spin echo sequence with 384 acquired lines, a TR time of TR = 559 ms and no averaging yields a total acquisition time of approximately

$$t_{ac} = N_{PE} \cdot TR = 384 \cdot 559 \text{ ms} \approx 3 \text{ min } 35 \text{ s} \quad (2.7)$$

for every slice. The goal of parallel acquisition techniques (PAT) is to reduce the number of phase encoding steps to accelerate the overall acquisition. The improvement is measured by the acceleration factor r . The main idea is to use additional receiving coils. Due to the different spatial location of the coils and their sensitivity profile the information can be combined to form the resulting image. Modern coil arrays can have up to 32 single receiving coils¹.

There are two different categories of PAT. The first one combines the information of the receiving coils in the image domain. Methods of the second category merge the data in the frequency domain. Further, there are a few methods that combine image domain and frequency domain. In the following, we briefly introduce one commonly used approach of both main categories: *Sensitivity encoding* (SENSE) and *generalized auto calibrating partially parallel acquisition* (GRAPPA).

Sensitivity Encoding

SENSE is an approach operating in the image domain. In order to accelerate the acquisition, the phase encoding step width is increased by a factor r . Thus, only every r th line has to be acquired yielding an acceleration factor of r . The missing k -space lines lead to overlaps of the acquired object as some frequencies that are required for reconstruction are not available. The overlaps are illustrated on the left of Fig. 2.9.

¹<http://www.medical.siemens.com>

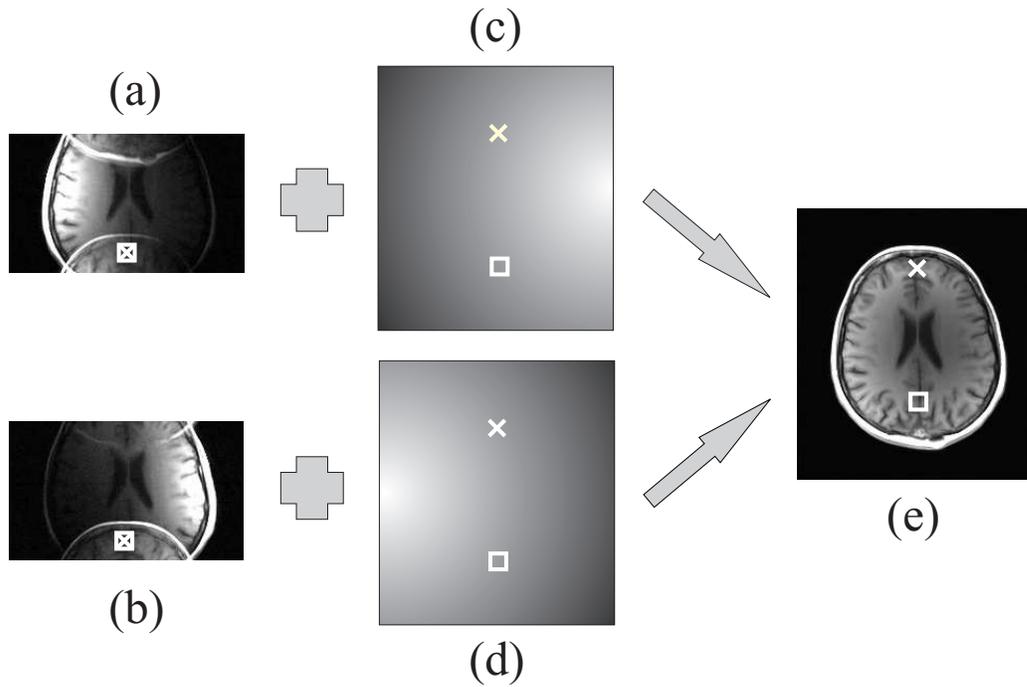


Figure 2.9: Schematic illustration of SENSE: The figure shows the principle of SENSE for a two coil setting ($r = 2$). Figures (a) and (b) correspond to the images S_1 and S_2 acquired using coil 1 and coil 2. The overlaps in the images are a result of the missing k -space data. For an acceleration factor of $r = 2$, data for every second phase encoding step is acquired only. Image (c) and (d) are the coil sensitivity profiles C_1 and C_2 measured beforehand. Using the sensitivities and the acquired images, the original image ρ can be reconstructed (image (e)) using equations 2.8. The cross and the rectangle show the relationship of the pixel contributions and locations in all images.

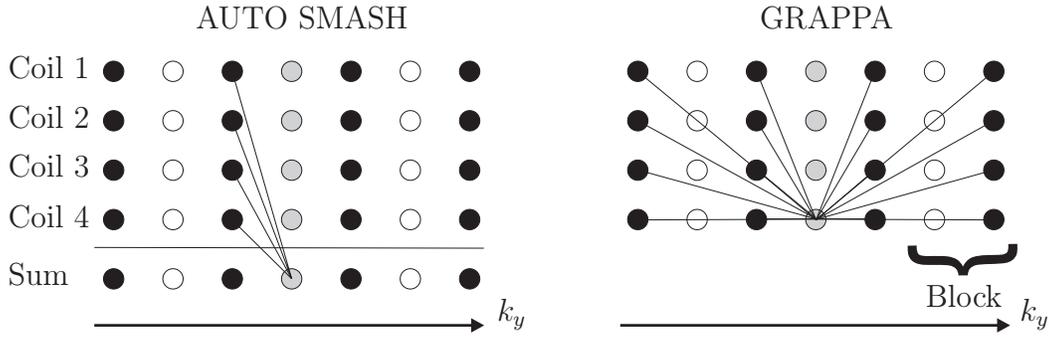


Figure 2.10: Schematic illustration of AUTO SMASH and GRAPPA: Solid circles illustrate measured lines, blank circles missing lines, and gray circles auto-calibration signals S^{ACS} . The figure shows an acquisition with acceleration factor $r = 2$ and one additional S^{ACS} line. Left figure: Illustration of the estimation of the weights using the AUTO SMASH approach. Right figure: Using the GRAPPA algorithm multiple lines are used for the estimation.

Every pixel in the acquired image of every coil represents the weighted sum of the spin density ρ at r different locations within the images. The function of weighting factors C_i of coil i are location dependent and correspond to the sensitivity profile of the coils. An example of the system of equations for two receiving coils and thus an acceleration factor of $r = 2$ is

$$\begin{aligned} S_1(x, y) &= C_1(x, y)\rho(x, y) + C_1\left(x, y + \frac{\text{FOV}_y}{2}\right)\rho\left(x, y + \frac{\text{FOV}_y}{2}\right) \\ S_2(x, y) &= C_2(x, y)\rho(x, y) + C_2\left(x, y + \frac{\text{FOV}_y}{2}\right)\rho\left(x, y + \frac{\text{FOV}_y}{2}\right) \end{aligned} \quad (2.8)$$

where S_i is the signal perceived at coil i . The connection between the two equations is shown in Fig. 2.9. Here, the sensitivity profiles C_i of the coils have to be known. They can be computed by a preceding reference scan. Both equations have two unknowns for every measured position. As there are two unknowns and two equations the system of equations can easily be solved. Thus, the “real” signal intensity can be computed for both locations in the image.

Generalized auto calibrating partially parallel acquisition

In contrast to SENSE, GRAPPA [Gris02] is a method operating in k -space. The underlying theoretical concept of k -space methods is that the phase encoding gradient G_{PE} bears a sinusoidal phase modulation across the whole excited object [Lark07]. As the measured phase is encoded along the y -direction in k -space, it yields a shift of the acquired line in that direction. This is used in the reconstruction to approximate the missing lines.

The k -space entries of the j th coil can be written as

$$S_j(k_x, k_y) = \iint C_j(x, y)\rho(x, y)e^{-ik_x x}e^{-ik_y y}dx dy \quad (2.9)$$

using the explicit formulation of the Fourier transform as integrals over x and y . The simplest k -space based method is *Simultaneous acquisition of spatial harmonics* (SMASH) [Sodi97]. In order to fill the missing k -space lines, it explicitly uses the sensitivity profiles of the coils. It is assumed that the profiles of all coils differ from each other. Because of the arrangement of the coils, this assumption is valid. SMASH uses a linear combination of the coil sensitivities to approximate the m th spatial harmonic:

$$\sum_{l=1}^L n_c(l, m) C_l(x, y) = e^{-im\Delta k_y y} \quad (2.10)$$

where L is the number of available coils and $n_c(l, m)$ defines the unknown weighting of the l th coil to the approximation of the m th spatial harmonic. The weighting factors $n_c(l, m)$ are estimated by a least-squares fit. Inserting Eq. (2.10) into Eq. (2.9) yields

$$\begin{aligned} \sum_{l=1}^L n_c(l, m) S_l(k_x, k_y) &\approx \iint \rho(x, y) e^{-ik_x x} e^{-ik_y y} e^{-im\Delta k_y y} dx dy \\ &= S(k_x, k_y + m\Delta k_y). \end{aligned} \quad (2.11)$$

Using Eq. (2.11) all missing k -space lines can be approximated. Consequently, the complete image can be reconstructed.

As the coil sensitivities are usually not known, an adaption of SMASH referred to as AUTO-SMASH [Jako98] determines the weights by acquiring additional lines in the center of the k -space. These lines are used as reference lines. In order to estimate the necessary weights $n_c(l, m)$, a fit to these lines is done. Theoretically, only a single additional line is needed, if the acceleration factor r is two. In general, at least $(r - 1)$ reference lines are required. The additional acquired k -space lines are called auto-calibration signals S^{ACS} . The problem can be formulated as

$$\sum_{l=1}^L S^{\text{ACS}}(k_x, k_y + m\Delta k_y) = \sum_{l=1}^L n_c(l, m) S_l(k_x, k_y). \quad (2.12)$$

The fitting of the data is illustrated in Fig. 2.10. Here only a single line is used to compute the missing lines in between. The reconstruction itself is equivalent to SMASH.

In contrast to SMASH-like methods, GRAPPA does not only compute the combined image but provides a reconstruction for every receiver coil. Further on, GRAPPA does not use a single line to compute the missing k -space lines but a neighborhood. The neighborhood consists of a number of blocks N_b . Each block consists of one acquired line and $(r - 1)$ missing lines. A block is illustrated in Fig. 2.10. As the missing lines are approximated for all coils, weights for every receiving coil j have to be estimated. Using this the reconstruction formula of GRAPPA can be written as

$$S_j(k_x, k_y + m\Delta k_y) = \sum_{l=1}^L \sum_{b=0}^{N_b-1} n_c(j, b, l, m) S_l(k_x, k_y + br\Delta k_y). \quad (2.13)$$

The corresponding fitting process is illustrated in Fig. 2.10.

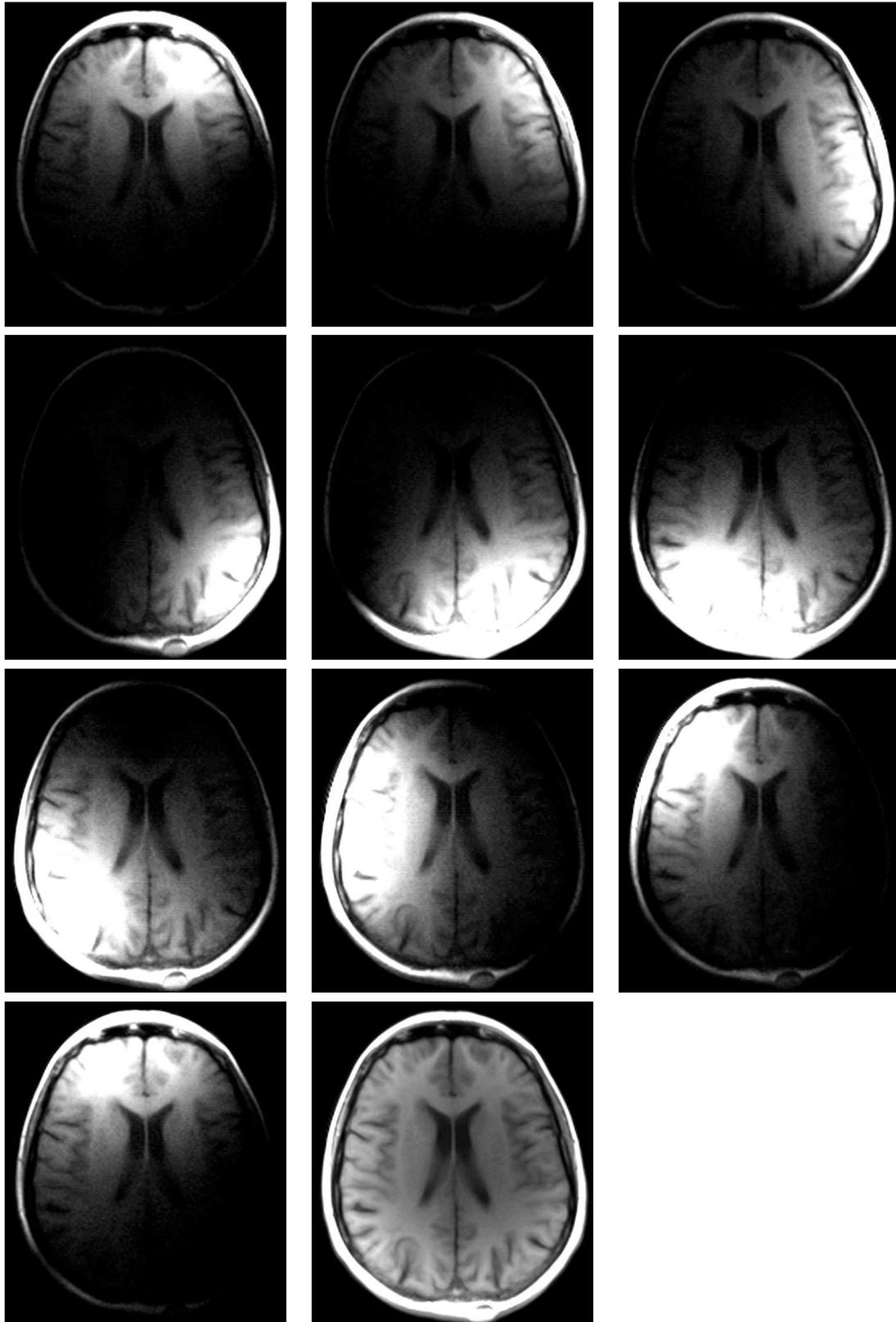


Figure 2.11: Parallel acquisition and sum-of-squares reconstruction: The image in the bottom row and second column shows the sum of squares reconstruction result using a 32 channel head receiver coil array. The other ten images are example images of single coils. The images are acquired at the Radiologic Institute at the University of Erlangen.

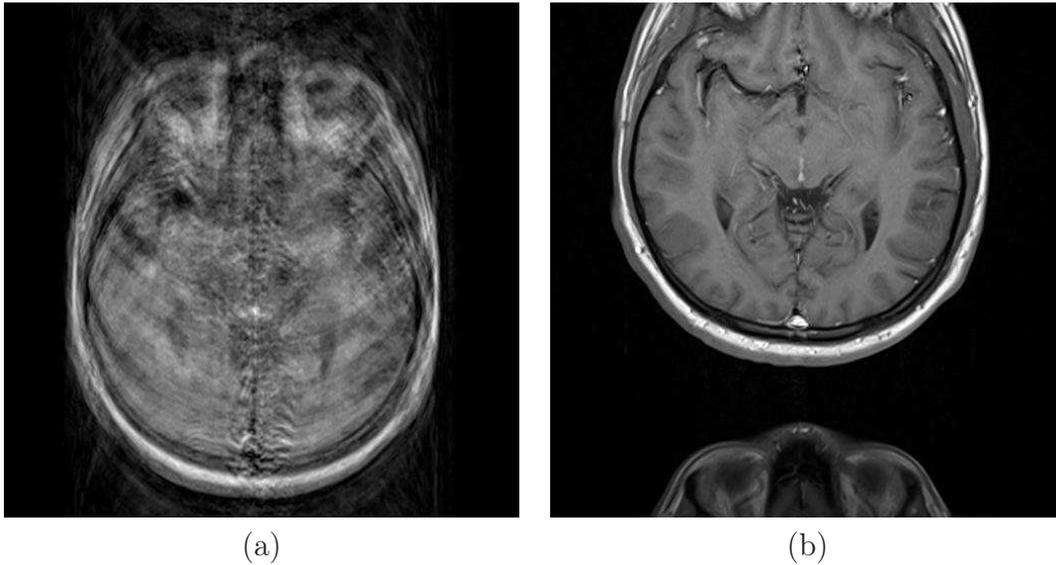


Figure 2.12: Image (a) shows a motion artefact on a T1w head image. In image (b) a wrap around artifact is illustrated. Both images are acquired at Siemens AG, Healthcare Sector, Erlangen, with TE=8.4 ms, TR=550 ms and an flip angle of 90°.

After the computation of the single coil images, they have to be merged to form the result image. In general a sum-of-squares (SOS) [Lars03] reconstruction is used for this task. The SOS reconstruction can be formulated as

$$\tilde{\rho}(x, y) = \sqrt{\sum_{l=1}^L |S_l(x, y)|^2} \quad (2.14)$$

where $\tilde{\rho}(x, y)$ is the estimated object density at position (x, y) . An example for the SOS reconstruction using a 32 channel head coil array is shown in Fig. 2.11.

2.3 Artifacts

In this section, we briefly introduce some of the most commonly encountered imaging artifacts in MRI. However, many details are skipped. Intensity non-uniformities are discussed in more detail in the following chapters about the correction of signal intensity variations. The correction of many of the introduced artifacts gains more and more importance as their impact on image quality of modern scanners with higher field strengths increases.

2.3.1 Motion Artifacts

Motion artifacts are a very common problem in MR imaging. One of the basic assumption of MRI is that there is absolutely no patient movement during the acquisition. The only change that occurs is due to the variation of the phase encoding gradient G_{PE} . Since the time for a complete acquisition can be rather long, motion

is unavoidable. This can either be the result of respiratory motion, heart beat, blood flow, or patient movement. The result is that the tissue that is excited at a certain position within the 3-d world coordinate system produces its signal at a different, translated position. Thus, the acquired information is put in the wrong location within the k -space matrix. Moreover, if after a certain time the patient moves but stays in the new position for the rest of the acquisition, the k -space is filled with two different patient settings.

Motion disturbs phase encoding more than on the frequency encoding as phase encoding is applied before the readout, whereas frequency encoding is done concurrently with signal acquisition. A motion artifact is illustrated in Fig. 2.12 (a).

2.3.2 Aliasing

Aliasing, also called warp-around artifacts, results from objects that are excited but that are outside of the actual FOV. In the images they manifest as overlaps on the opposite side of the FOV showing the object parts that are outside the FOV. This artifact can occur both because of phase encoding and frequency encoding. However, artifacts caused by frequency encoding can be eliminated by oversampling of the data during the readout or by removing the high frequencies using a bandpass filter. In phase encoding direction the aliasing can only be corrected by increasing the FOV in the phase encoding direction. A wrap around artifact is illustrated in Fig. 2.12 (b).

2.3.3 Chemical Shift Artifacts

There are two different artifacts that are summarized as chemical shift artifacts. The assumption of frequency encoding is that all protons within an excited voxel have the same precession frequency. However, at voxels that contain both water and fat molecules this is not the case. Fat molecules have a lower resonant frequency than water. This leads to a frequency mapping of the fat protons to a voxel with a lower frequency. Especially at borders between tissue classes with an equal distribution of water and fat molecules the tissue misalignment is visible. An example for a problematic region is the spine area between the discs and vertebrae.

The second artifact based on the chemical shift is called phase cancelation. This artifact occurs within out-of-phase gradient echo images. The reason for the artifact is that in gradient echo images the phase cycling is not reversed using a 180° pulse. Thus, fat and water protons contribute differently to the received signal. The contribution of different molecules depends on TE. This is illustrated in Fig. 2.13.

2.3.4 Truncation Artifacts

Truncation artifacts, also called Gibbs ringing, are related to the detection and sampling of the generated echoes. During the readout the signal is sampled. After and before the readout period the signal is cut off. If there is still a significant signal after this time period it is ignored. For instance, this can happen when high frequencies from high contrast edges are still present. In the images truncation artifacts can be seen as ring-like oscillations. The correction of this artifact can be done by using

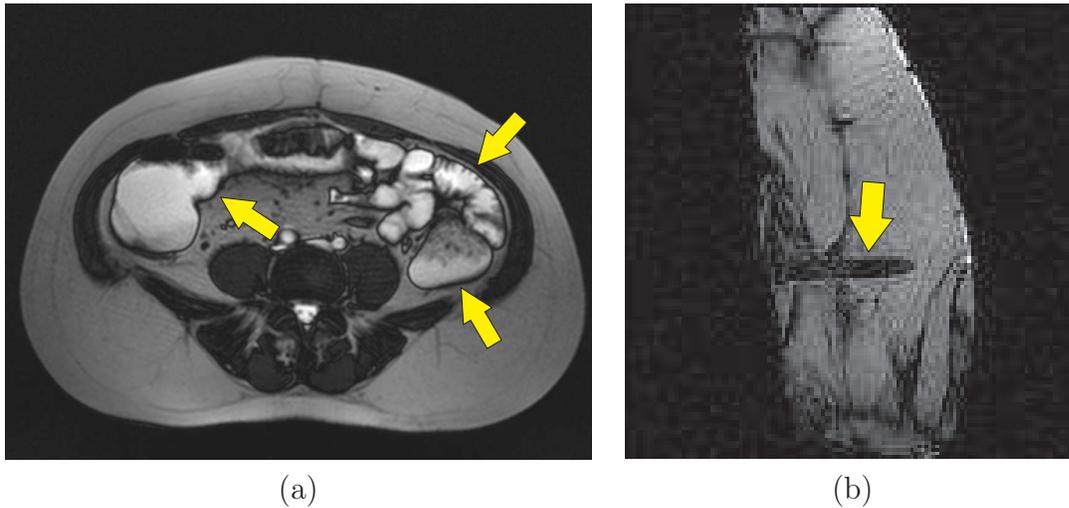


Figure 2.13: Chemical shift artifacts and magnetic susceptibility: Image (a) shows chemical shift artifact. Black borders around the structures are a result of this imperfection. Some of these are marked by the arrows. The image was acquired at the Radiologic Institute at the University of Erlangen. In image (b) a magnetic susceptibility artifact is illustrated. The black region marked by the arrow in the image results from a magnetic susceptibility artifact caused by a biopsy needle. The needle has a diameter of about 1 mm the artifact a diameter of about 6 mm. The image was acquired at Siemens AG, Healthcare Sector, Erlangen.

apodization filters like Hanning or Gaussian filters to reduce the signal intensity at the borders to zero. However, this can lead to blurred images.

2.3.5 Magnetic Susceptibility

If two different tissues like air and soft tissue with different magnetic susceptibility are in close vicinity to each other, local distortions are induced to the magnetic field. This yields a de-phasing of surrounding spins. Additionally, frequency shifts of spins may occur. In the images, this is observable as signal loss or a distortion of the image. Especially metal objects produce large artifacts due to magnetic susceptibility. This is illustrated in Fig. 2.13.

Magnetic susceptibility is used in paramagnetic contrast agents. They cause an observable signal loss in T_2 -weighted images, from which the tissue perfusion can be calculated for instance.

2.3.6 Intensity Non-uniformity

There are a couple of reasons for the occurrence of intensity non-uniformities. Mostly they are dependent on the scanned object. Thus, it is not possible to correct these non-uniformities by a calibration of the MR scanner. Instead, the inhomogeneity correction has to be done retrospectively in general. Methods for correction of these

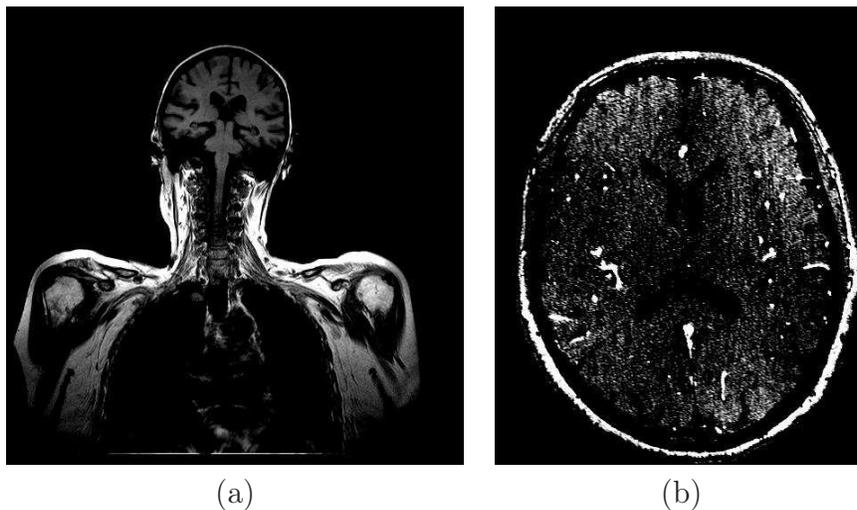


Figure 2.14: Intensity non-uniformities: Both images show examples for intensity non-uniformities. Image (a) was acquired at the Radiologic Institute, Charité, Berlin. Image (b) was acquired at the Radiologic Institute, University of Erlangen.

artifacts are one of the major topics of this thesis and will be discussed in the next chapters in more detail. Two examples are given in Fig. 2.14.

A major cause of signal intensity variations are inhomogeneities in the B_1 field induced by the RF pulse. In literature this is also referred to as *RF inhomogeneities*. It results in varying flip angles within the excited object. This yields a reduced magnetization in areas with smaller flip angles and thus, varying intensities in the acquired images.

Furthermore, the tuning of the coils is of importance. This is especially the fact for parallel acquisitions. For SENSE the coil sensitivities have to be approximated beforehand. For GRAPPA they are estimated implicitly. In general, however, the estimation is not perfect leading to smooth signal intensity inhomogeneities within the reconstructed images.

Interleaved acquisition schemes are usually used to accelerate the acquisition of 3-d volumes by using a multi-slice technique. However, in some cases neighboring slices are partly excited by the applied RF pulse. This can lead to a reduced signal intensity when the wrongly excited slice is excited again for its own acquisition. This is called *slab boundary artefact* in literature. The intensity deviations in neighboring slices are tissue dependent.

2.4 Summary

The basic principle of MRI is to put an object in a strong homogeneous magnetic field. Afterwards, the spins of the nuclei are flipped away from the parallel alignment to the external field using a RF pulse. This is called excitation. When the pulse is turned off, the spins begin to realign with the main magnetic field. The signal that they emit in the meantime can be measured using receiver coils. It is denoted as FID. So far, the whole sample is excited and emits the measured signal. Thus, no

imaging is possible. However, by introducing gradient fields that locally change the magnetic field strength, spatial encoding becomes feasible. The succession of excitation pulses, gradient fields and readout is called pulse sequence. The most popular ones are spine echo sequences. They fill the raw data matrix, the k -space, sequentially row by row. In order to decrease the time needed for acquisition, recently parallel acquisition techniques were introduced. Here, several images with reduced data are acquired in parallel. To compute the missing data, all images are combined during the reconstruction. The most commonly used reconstruction techniques used for this case are SENSE and GRAPPA. However, still many MR images suffer from imaging artifacts. The most observed artifacts are intra- and inter-image signal intensity inhomogeneities.

Mainly, there are three sources that characterize the intensity inhomogeneities in the data sets: the coils, the acquisition sequence, and inhomogeneities in the RF pulse. The induced inhomogeneities by the coils are multiplicative to the signal defined by their spatially dependent sensitivity profile. The profile smooth over the whole image domain. Thus, the inhomogeneities are composed of low frequencies only. Moreover, close to the receiver coils, the signal-to-noise ratio is higher, as the measured signal is stronger from these regions. On the other hand, farer away from the coils the signal gets weaker. Further on, the inhomogeneities are also influenced by the imaged sample, due to local magnetic fields altering the signal. However, also the used pulse sequence has a major influence on the signal intensity variations. E.g., the slab boundary artefact is often observed in MR images. It is a result of an interleaved acquisition strategy. Although often differently assumed, the intensity deviation is tissue dependent.

Chapter 3

Correction of Signal Intensity Inhomogeneities

In this chapter, we deal with the correction of signal intensity variations that occur within a single volume. In literature, this is often called correction of signal intensity inhomogeneities or estimation of a bias or gain field respectively. In some cases, especially related to retinal fundus images, the correction of intra-image signal intensity variations is named intensity equalization following the broadly used histogram equalization. First, in this chapter, we present a mathematical formulation of the correction problem that is based on the physical principles introduced in chapter 2. Then, a brief overview about state-of-the-art methods dealing with the topic is given. Because there has been a lot of work done in the last decades, we present a taxonomy for the approaches. Additionally, the most commonly used methods will be explained shortly. Further on, we introduce a novel approach to inhomogeneity correction and an extension to a given method. The latter one can be seen as an additional regularization term to stabilize the estimation of bias fields. This is followed by the experimental evaluation of the methods. The experiments show that we can increase the average separability of tissue classes in 3-d TOF MRA images by approximately 18.2% whereas state-of-the-art methods could only achieve 11.6%. Additionally, the applicability of the proposed methods for the correction of inhomogeneities occurring in other modalities is demonstrated. Finally, we conclude the section by a brief summary and a discussion of the proposed methods and experiments.

3.1 Problem Definition

In general, the intra-scan signal intensity variations are due to physical properties of the acquisition device and/or the acquisition setup. Usually, the bias field is said to be smooth and slowly varying as a result of the sensitivity profiles of the receiver coils (compare section 2.2.2). Consequently, it is composed of low frequency components only. This stands in contrast to multiplicative and additive noise, which is very rapidly changing. Thus, it relies on high frequency components. Basically, there are two different ways to model the influence of inhomogeneities on an acquired image: the additive and the multiplicative model. For many modalities, however, a mixture of both models is the source of the intensity variations. Thus, the choice of the model

should be based on the subsequent application. The additive model can be written as

$$v(\mathbf{x}) = u(\mathbf{x}) + b(\mathbf{x}) + n(\mathbf{x}) \quad (3.1)$$

where $v : \mathbb{R}^D \mapsto \mathbb{R}$ is the observed image including the bias field effects, $u : \mathbb{R}^D \mapsto \mathbb{R}$ is the ideal image, $b : \mathbb{R}^D \mapsto \mathbb{R}$ is the smooth bias field and $n : \mathbb{R}^D \mapsto \mathbb{R}$ represents the noise component of the model. The vector $\mathbf{x} \in \Omega$ defines an arbitrary position within the image domain $\Omega \in \mathbb{R}^D$ where D is the image dimensionality. Usually, Ω is a subset of \mathbb{R}^2 or \mathbb{R}^3 . The most commonly assumed model is the multiplicative bias field model. The reason for this is that the coil sensitivities as well as the RF pulse inhomogeneities are multiplicative to the MR signal (compare chapter 2). The model can be formulated as

$$v(\mathbf{x}) = u(\mathbf{x}) \cdot b(\mathbf{x}) + n(\mathbf{x}). \quad (3.2)$$

For both types of models the noise part n is usually neglected. In order to avoid problems associated with multiplicative models like its non-linearity, in literature Eq. (3.2) is often used in its logarithmic form:

$$\begin{aligned} \hat{v}(\mathbf{x}) &= \log v(\mathbf{x}) = \log u(\mathbf{x}) + \log b(\mathbf{x}) \\ &= \hat{u}(\mathbf{x}) + \hat{b}(\mathbf{x}) \end{aligned} \quad (3.3)$$

where $\hat{\cdot}$ denotes the logarithmic transformed signals. This, however, induces a different kind of problem. As the intensity quantification is discrete and the logarithmic transformation changes the spacing between the intensities, it gets more difficult to compute good approximations of the logarithmic images' intensity distributions. An example for this is shown in Fig. 3.1. A possible solution to the mentioned problem is an Parzen estimation of the histograms using varying kernel sizes which depend on the position within the histogram.

Moreover, a technique called NP windows [Dows06] can be used. We use an adaption of this. Here, the image is assumed to be a continuous function that is sampled at the grid points corresponding to the positions $\mathbf{x} \in \Omega$. In order to obtain intensities between the grid points, we use a linear interpolation between them. If an intensity $v(\mathbf{x})$ at a given position \mathbf{x} in an image v has to be inserted into the histogram \mathcal{V} the intensity information of neighboring pixel is used as well. For each neighboring intensity $\mathbf{x}' \in \mathcal{N}_{\mathbf{x}}$, all bins between the bins corresponding to the intensities \mathbf{x} and \mathbf{x}' are increased by a contribution factor. This factor is calculated with respect to the intensity distance of \mathbf{x} and \mathbf{x}' . The advantage of this technique is twofold. First, contextual information is integrated into the histograms. Second, less data samples are needed to compute a meaningful histogram.

The advantage of the additive bias field model is that the signal-to-noise ratio (SNR) is not changed due to the intensity transformation. Using the multiplicative case this is not the case. Assuming the standard SNR definition $\text{SNR} = \mu/\sigma$ with μ being the mean value of a single class and σ its standard deviation, the SNR change can be seen in the multiplicative case computed as

$$\text{SNR} = \frac{\int_{\Omega} u(\mathbf{x})b(\mathbf{x})d\Omega}{\sigma_{n(x)} \cdot \#\Omega} \quad (3.4)$$

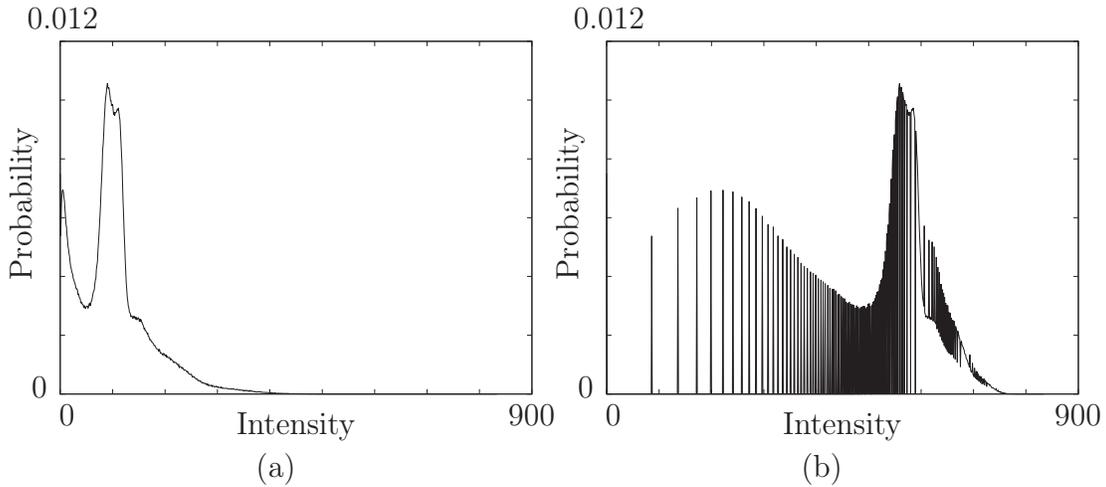


Figure 3.1: Example for log-transformed histograms: In plot (a) the histogram of a T1w head image is shown. The intensity range is zero to 833. Plot (b) corresponds to the histogram of the log-transformed image. The intensities are scaled so that the image has the same intensity range than the input. The plots illustrate that it is much more difficult to approximate a good histogram in the log-transformed case.

where $\sigma_{n(x)}$ is the standard deviation of the noise model. It is obvious that the SNR is directly dependent on the strength of the bias field. If the values of the bias field get smaller then the SNR gets lower. Correcting the image in these areas using a multiplicative bias field model increases the tissue contrast within the dark regions of the images. In many cases this is a very advantageous property, for subsequent segmentation methods for instance. For the additive model the contrast is not increased.

In general, the correction of the intensity inhomogeneities corresponds to a reduction of the information in the sense of information theory within the images. However, the reduced information has to correspond to the bias field only. No tissue information of any kind is acceptable to be lost during the correction process. For this reason, the estimation of the bias field and thus the reduction is restricted by physical properties of the bias field, like the smoothness and the slow variance. Thus, many trivial solutions are impossible, for instance $b(\mathbf{x}) = v(\mathbf{x})$. Furthermore, during the correction of the signal intensity inhomogeneities, it has to be kept in mind that the intensities of the images are invariant against a global affine transformation. Consequently, the scaling of the intensities are arbitrary. Furthermore, the origin of the intensities can be chosen arbitrarily. For this reason the following equation

$$u(\mathbf{x}) \equiv s_{\text{MRI}}u(\mathbf{x}) + t_{\text{MRI}} \quad (3.5)$$

where $s_{\text{MRI}} \in \mathbb{R}^+$ is an intensity scaling and $t_{\text{MRI}} \in \mathbb{R}$ is an intensity shift holds. Usually, in MRI the lowest intensity is zero corresponding to “no received signal”. In general, the MRI images are corrected in a mean-preserving manner. Thus, for instance, in the multiplicative case

$$\int_{\Omega} v(\mathbf{x})d\Omega = \int_{\Omega} \frac{v(\mathbf{x})}{b(\mathbf{x})}d\Omega \quad (3.6)$$

has to hold. Certainly, this can also be formulated for the additive case as well.

3.2 State-of-the-Art

In the last decades, a lot of different approaches for the correction of intensity inhomogeneities in MRI have been developed. The great majority of the methods can be classified to fall into a small number of different categories. First of all, one has to differentiate between prospective and retrospective methods.

Prospective methods usually use information gained from the scanner system to correct the acquired images. The sources of the used information are scans of phantoms [Simm 94, Coll 02], data from multi-coil images (e.g. surface and body coils [Nara 88, Fan 03]) and the utilization of special acquisition sequences like SENSE and GRAPPA (compare section 2.2.2) [Prue 99, Chio 03]. The advantage of prospective methods is that the additional information is closely related to the problem in general. Thus, the correction is very good for most approaches. On the other hand, the methods are often restricted to a small area of application. Furthermore, in many cases additional scans are required that introduce higher overall costs and time consumption. Most important, however, often the image acquisition is done long before the processing of the data. At this time, no additional correction or calibration scans can be performed anymore. For these reasons, we concentrate on retrospective methods in this thesis.

In this context, retrospective means that the correction takes place after the acquisition and the reconstruction of the images. Although, there are multiple ways to define categories to classify **retrospective correction approaches** [Bald 06], we use the classification scheme introduced by Vovk *et al.* in [Vovk 07]. They use the following categorization:

1. filtering methods,
2. surface fitting approaches,
3. segmentation-based methods and
4. histogram-based approaches.

Filtering: Methods based on the filtering [Axel 87, Lewi 04, Cohe 00] of the acquired images rely on the assumption that the inhomogeneities do not contain any high frequency information at all. The image content, on the other hand, is assumed to be composed of high frequency signals. Thus, the image content and the inhomogeneities can be separated using low pass filtering. This implies that there is no frequency overlap between both classes. In general, this assumption is not valid. Consequently, new artifacts are introduced due to the filtering operation, e.g. edge effects (see Experimental Section 3.5.1). The most prominent used method is called *Homomorphic Unsharp Masking* (HUM). We describe this method later on in this section in more detail.

Surface Fitting: Methods based on surface fitting match a parametric model to previously selected features [Dawa 93, Vemu 05, Voku 99]. The features used are often the image intensities of a single dominant tissue class. There are many disadvantages

for methods using this kind of features. First of all, pixel corresponding to the tissue have to be selected within the images. Further on, the correction is based on a single tissue class, that has to be very large and distributed over the whole image domain. Finally, the bias field has to be extrapolated to other regions what is very error-prone. As this class is hardly used in practice, we do not introduce an example approach in more detail.

Segmentation: The segmentation of images is very closely related to the correction of the intensity inhomogeneities. If a complete segmentation of the image is available, the computation of the bias field is rather trivial. On the other hand, it is very advantageous for the segmentation process if there are no intensity variations within the images that are induced by an imperfect acquisition. This close relationship is utilized by many authors to jointly do a segmentation and an inhomogeneity correction. Especially, statistical frameworks are suitable for the incorporation of the estimation of the bias field additionally to the determination of segmentation model parameters (e.g. segmentation methods based on maximum-likelihood (ML) or the maximum *a posteriori* (MAP) criterion [Well96, Van99a, Zhan01]). A very commonly used approach for the joint estimation of the bias field and the segmentation of images is the *Modified Fuzzy-C-Means* (MFCM) segmentation approach introduced by Ahmed *et al.* [Ahme02]. We discuss this method later in this section.

Histogram: The most generalizable class of approaches can be labeled as histogram-based methods. All information required for the correction is directly computed from the images' histograms. Usually these methods base on the assumption that the application of a bias field to undistorted images corresponds to the convolution of the image histograms with a smoothing kernel. Due to this smoothing the information content within the images is increased. Consequently, the entropy is raised as well. Histogram-based methods use these properties to correct the intensity inhomogeneities [Salv06]. Due to the little assumptions made, these approaches are in general more robust in many fields of application and to structural changes in the images because of evolving pathologies, for instance. The most commonly used algorithms of this class are the *Nonparametric Nonuniformity Normalization* (N3) introduced by Sled *et al.* [Sled98] and the *Correction using Information Minimization* by Likar *et al.* [Lika01]. Both methods are introduced in more detail later in this section.

In the following we briefly discuss the most commonly used methods for signal intensity inhomogeneity correction. For more details about the mentioned approaches, we want to refer to the corresponding literature.

Homomorphic Unsharp Masking

HUM, a filtering method, is the most commonly used bias field correction method. It was first introduced by Axel *et al.* in 1987 [Axel87]. HUM assumes a multiplicative inhomogeneity model. The correction is done by a simple division of the observed image $v(\mathbf{x})$ by its low passed version $\text{LPF}\{v(\mathbf{x})\}$. In order to maintain the image mean or sometimes the median, the resulting image is multiplied by a constant corresponding to the image mean/median μ . Thus, the correction can be written as

$$u(\mathbf{x}) = v(\mathbf{x})/b(\mathbf{x}) = \mu \cdot v(\mathbf{x})/\text{LPF}\{v(\mathbf{x})\}. \quad (3.7)$$

Besides of the choice of using the mean or the median as multiplicative constant, the only parameter that has to be chosen is the size of the smoothing kernel. Brinkmann *et al.* showed in [Brin98] that a kernel width of approximately 65 is optimal for the correction of MRI brain images. A further advantage of HUM is that it can be easily implemented very efficiently.

Modified Fuzzy C-Means

MFCM proposed by Ahmed *et al.* in [Ahme02] is a typical example for a segmentation based bias field correction approach. The method is designed to segment the brain into the three classes *background*, *white matter*, and *gray matter*. In order to improve the segmentation of a standard Fuzzy C-Means (FCM) classifier, they added two additional components: a regularizer that incorporates the spatial coherence of the tissue classes and an additional parameter \hat{b} that represents the bias field. They use the logarithmic version of the multiplicative model. Although, the method can also be applied to multi-spectral data, for simplicity we present the single-spectral case only. The overall objective function for the joint segmentation and bias field estimation can be written as

$$J = \sum_{i=1}^C \sum_{\mathbf{x} \in \Omega} a_{i\mathbf{x}}^p |\hat{v}(\mathbf{x}) - \hat{b}(\mathbf{x}) - c_i|^2 + \frac{\alpha}{\#\mathcal{N}_{\mathbf{x}}} \sum_{i=1}^C \sum_{\mathbf{x} \in \Omega} a_{i\mathbf{x}}^p \left(\sum_{\mathbf{r} \in \mathcal{N}_{\mathbf{x}}} |\hat{v}(\mathbf{r}) - \hat{b}(\mathbf{r}) - c_i|^2 \right) \quad (3.8)$$

with C being the number of classes, α being the factor steering the influence of the neighborhood, p being a weighting exponent controlling the fuzziness, $\#\mathcal{N}_{\mathbf{x}}$ being the cardinality of $\mathcal{N}_{\mathbf{x}}$ and $a_{i\mathbf{x}}$ being the fuzzy class membership of voxel \mathbf{x} to class i . The parameters to be estimated during the minimization of J are the class centers $\{c_i\}_{i=1}^C$, the bias field estimates $\{b(\mathbf{x})\}_{\mathbf{x} \in \Omega}$, and the partition matrix $[a_{i\mathbf{x}}] = A$ with $a_{i\mathbf{x}} \in [0, 1]$, $\sum_{i=1}^C a_{i\mathbf{x}} = 1 \forall \mathbf{x}$, and $0 < \sum_{\mathbf{x} \in \Omega} a_{i\mathbf{x}} < N \forall i$.

Nonparametric Nonuniformity Normalization (N3)

Besides HUM, N3, a histogram based method, is one of the most commonly used algorithms for inhomogeneity correction. Many authors compare their results to this approach. The reason is that it is fast, relatively robust and can be applied to any body region as well as acquisition sequence. Sled *et al.* assume in [Sled98] a multiplicative bias field model. In order to be able to use an additive model, they employ the logarithmic case. The principle idea of their correction strategy is based on the assumption that the application of a bias field b to an image u corresponds to the convolution of the histogram \mathcal{U} of the image with the histogram \mathcal{B} of the bias field. As the bias field is very smooth and slowly varying, the correction problem can be reduced to finding a multiplicative field that maximizes the high frequencies of the histogram $\tilde{\mathcal{U}}$ of the approximation of the ideal image \tilde{u} . However, the distributions of u and b are not known. Sled *et al.* propose that the distribution \mathcal{U} can be approximated by sharpening the distribution \mathcal{V} of the observed image v . Then a smooth field is estimated that produces a distribution $\tilde{\mathcal{U}}$ that is close to the sharpened one. Furthermore, the distribution of \hat{b} can be assumed to be Gaussian with zero mean and given variance. Thus, the only parameter to be estimated is

the width of the distribution. \mathcal{B} is assumed to be a Gaussian distribution. As any Gaussian can be decomposed into a convolution of narrower Gaussian distributions, the method is iteratively executed starting with a very narrow distribution [Sled 98].

Given the estimation $\tilde{\mathcal{U}}$, the Gaussian shaped distribution \mathcal{B} can be computed. Using this and the convolution assumption the expected value of \hat{u} given \hat{v} can be written as

$$\begin{aligned} E[\hat{u}|\hat{v}] &= \int_{-\infty}^{\infty} \hat{u}p(\hat{u}|\hat{v})d\hat{u} \\ &= \frac{\int_{-\infty}^{\infty} \hat{u}\mathcal{B}(\hat{v} - \hat{u})\tilde{\mathcal{U}}(\hat{u})d\hat{u}}{\int_{-\infty}^{\infty} \mathcal{B}(\hat{v} - \hat{u})\tilde{\mathcal{U}}(\hat{u})d\hat{u}}. \end{aligned} \quad (3.9)$$

This holds since the location \mathbf{x} is arbitrary and it is assumed that \hat{u} and \hat{v} are independent and uncorrelated random variables. Consequently, using Eq. (3.9), the bias field \hat{b} can be approximated by

$$\hat{b}_s(\hat{v}) = S \{ \hat{v} - E[\hat{u}|\hat{v}] \} \quad (3.10)$$

where S is a smoothing operator reducing the effect of noise within the images.

Correction using Information Minimization

In [Lika 01] a histogram based method is introduced that directly uses the assumption that the application of a bias field implies an increase of the information content within the image. They model the degeneration of the bias field by an additive and a multiplicative part. However, they show that in most of the cases the additive part can be neglected. In order to describe the bias field they use a linear combination of K smoothly varying basis functions $s_i(\mathbf{x})$, $i \in \{1, \dots, K\}$, with a_i being the weight of the i -th basis function. As basis function Likar *et al.* use Legendre polynomials $q_i(\mathbf{x})$. The degree of the polynomials was up to fourth-order. Further on, they keep the mean value constant (mean-preserving criterion) and ensure that an equal change of any parameter a_i results in a similar magnitude of model variation. This is realized by

$$\frac{1}{\#\Omega} \int_{\Omega} |v(\mathbf{x})s_i(\mathbf{x})|d\mathbf{x} = 1 \quad \forall i. \quad (3.11)$$

with $\#\Omega$ being the cardinality of Ω . Using this the model of the signal intensity inhomogeneities can be written as

$$\begin{aligned} b^{-1}(\mathbf{x}) &= 1 + \sum_{i=1}^K a_i s_i(\mathbf{x}) \\ &= 1 + \sum_{i=1}^K a_i \frac{q_i(\mathbf{x}) - c_i}{d_i} \end{aligned} \quad (3.12)$$

where the c_i are the neutralization constants needed to fulfill the mean-preserving condition and d_i being the normalization parameters corresponding to the condition shown in Eq. (3.11). Thus, the optimization problem can be written as

$$\{a_i^*\} = \arg \min_{\{a_i\}} H(v(\mathbf{x})b^{-1}(\mathbf{x})) \quad (3.13)$$

with H being the Shannon entropy and $\{a_i^*\}$ being the set of optimal parameters. The solution to this problem can be found by Powell’s multidimensional directional set method and Brent’s one-dimensional optimization algorithm [Lika 01].

3.3 A Divide-and-Conquer Approach to Bias Correction

In this section, we present a Divide-and-Conquer-like method for the correction of intensity inhomogeneities in images. In the following the method will be referred to as DaC approach. The main idea is that the approximation of a complex global parametric model is reduced to the estimation of many much simpler local models (compare Fig. 3.2). This is a valid assumption as the bias field is smooth and slowly varying (compare sections 2.2.2 and 3.1). The number of required parameters to be estimated can differ severely between the local and the global model. Typical numbers are two or three unknowns for each local model and approximately 100 for the complex one. Subsequently, the local models are put together in the “conquer” phase to form the global one. Finally, the global model is used for the actual correction of the whole image. The process is illustrated in Fig. 3.2.

The biggest advantage of the separation into two phases lies in the complexity of the optimization. For the low-order local models the optimization bases on the entropy of the image. Thus, it is non-linear. However, due to the small number of unknowns feasible and efficiently solvable. The composition of the global model in the “conquer” phase can be mapped to a linear operation. Consequently, it can be computed very fast as well.

3.3.1 Model Definition

Throughout literature a lot of different parametric models are used to approximate the bias field within images. Particular suitable are linear combinations of basis functions. Here, especially radial basis functions as well as polynomials are used.

Although other models are also possible, we use polynomials

$$b_r(\mathbf{x}) = n_e + \sum_{i=1}^D a_i^r x_i = n_e + \mathbf{a}^{rT} \mathbf{x} \quad (3.14)$$

of degree one as local model. D corresponds to the dimension of the image, e.g., $D = 2$ or $D = 3$. The vector $\mathbf{x} = (x_1, \dots, x_D)^T$ defines the location of a voxel within the image. The scalars $a_i^r \in \mathbb{R}$ are the coefficients of the local polynomial model r . The set \mathbf{a}^r is defined as $\mathbf{a}^r = \{a_1^r, \dots, a_D^r\}$. The coefficients a_i^r are the only unknowns in the system of equations. Let the scalar n_e be the neutral element of the bias correction operation. For a multiplicative model, the neutral element is set to $n_e = 1$ and for an additive model to $n_e = 0$. This means that if there is no correction necessary, all coefficients a_i^r are set to $a_i^r = 0$.

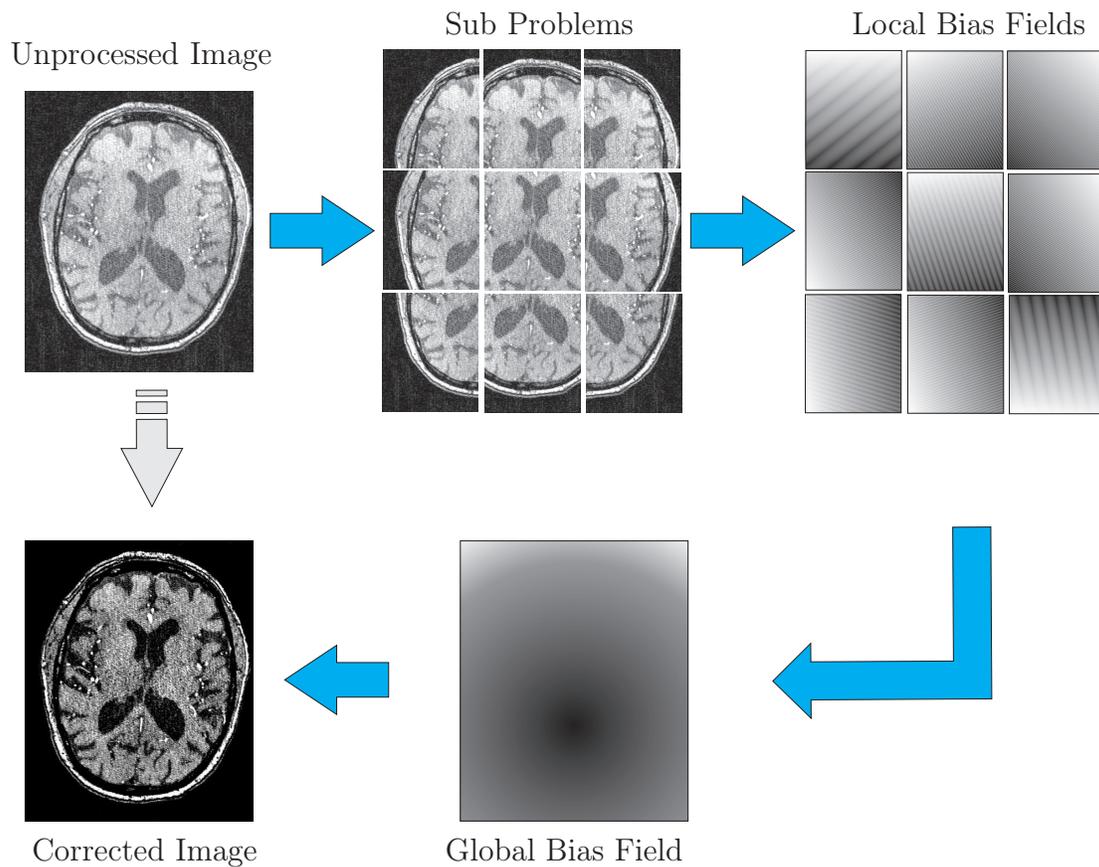


Figure 3.2: Workflow of the divide-and-conquer based bias correction: The upper row corresponds to the “divide”-phase of the bias correction; the lower row to the “conquer”-phase respectively.

As global bias model, we use a polynomial model as well. However, the degree d of this polynomial is usually much higher than for the local one. In the 2-d case the global polynomial model can be written as

$$b(\mathbf{x}) = \sum_{i=0}^d \sum_{j=0}^{d-i} a_{i,j} x_1^i x_2^j \quad (3.15)$$

with $a_{i,j}$ being the unknown coefficients and $a_{0,0} = n_e$. Consequently, the number of unknowns N_u can be computed from the image dimension D and the polynomial degree d as

$$N_u = \frac{(d+D)!}{d!D!} - 1. \quad (3.16)$$

As introduced in section 3.1, from an information theoretic viewpoint the bias field correction reduces the information content of the images. For this reason, we use the Shannon entropy

$$H(v) = - \int_{\Gamma} p_v(\chi) \log p_v(\chi) d\chi \quad (3.17)$$

as criterion for optimality with Γ corresponding to the intensity domain. Consequently, the optimal bias field minimizes the entropy of the image with respect to the constraints introduced by the bias field model. The densities that are required for the computation of the entropy are approximated using either the NP windows based estimation technique [Dows06] introduced in section 3.1 or a Parzen estimation with a Gaussian window function G_σ . The scalar σ is the standard deviation of the window function. The bin sizes of the histograms are chosen to be smaller than the quantization step size of the intensities within the images. Using Parzen window estimation the approximated probability density function can be written as

$$p_\sigma(\chi) = \frac{1}{N} \sum_{j=1}^N G_\sigma(\chi - \chi_j) \quad (3.18)$$

with N being the number of intensity samples, χ being the intensity of interest and χ_j being the intensity associated with bin j . In order to increase the computational efficiency this can be approximated by

$$\tilde{p}_\sigma(\chi) = \mathcal{H} \star G_\sigma \approx p_\sigma(\chi) \quad (3.19)$$

where \mathcal{H} corresponds to the histogram of the image. For more details, we want to refer to Hahn *et al.* [Hahn09]. The advantage of using Parzen estimation (without optimal kernel width estimation) compared to the NP windows based technique is runtime. However, with Parzen estimation the histograms are more sensitive to the number of samples, the size of the bins as well as noise within the images.

3.3.2 Divide-Step

In a first step, the global correction problem is divided into smaller sub-problems. This is done by defining R regions \mathcal{R}_i that cover the whole region of interest: $\bigcup_{i=1}^R \mathcal{R}_i =$

Ω . Although it is not mandatory, the regions are equally sized and chosen to be disjoint. The splitting of the problem domain is illustrated in Fig. 3.2. For all regions, it is assumed that its bias field can be approximated by the local bias field models defined in Eq. (3.14). The search for the optimal parameter set \mathbf{a}^{r*} of the local bias field can be formulated as the non-linear minimization problem

$$\mathbf{a}^{r*} = \arg \min_{\mathbf{a}^r} H(r; \mathbf{a}^r) \quad (3.20)$$

where $H(r; \mathbf{a}^r)$ is the entropy of sub-region r corrected by the local bias field with parameter set \mathbf{a}^r . As all sub-problems are independent of neighboring regions they all have their own coordinate system. The point of origin is chosen to be the center of each sub-region. Thus, the signal intensity of this center point is never altered independent of the type of the local bias field model and its strength (compare Eq. (3.14)). This has the advantage that local models with an extreme slope are no longer preferred.

The entropy minimization itself is done by a gradient descent approach. Due to the low degree of the local bias model, in the 2-d case only two and in the 3-d case only three parameters have to be estimated for each sub-region. In order to emphasize the spatial coherence of neighboring sub-regions an additional regularizer, suppressing large variations of neighboring parameter sets, can be introduced. Such a regularization term can be formulated as

$$S(r; \mathbf{a}^r) = \sum_{i=1}^D |a_i^r - \mu_i^r| \quad (3.21)$$

where a_i^r are the coefficients of the local polynomial model and

$$\mu_i^r = \frac{1}{\#\mathcal{N}_r} \sum_{l \in \mathcal{N}_r} a_i^l \quad (3.22)$$

are the mean values of the coefficients of the neighboring regions. \mathcal{N}_r is the neighborhood of sub-region r and $\#\mathcal{N}_r$ is its cardinality. Given the neighboring coefficients the resulting regularized optimization problem can be written as

$$\mathbf{a}^{r*} = \arg \min_{\mathbf{a}^r} H(r; \mathbf{a}^r) + \alpha S(r; \mathbf{a}^r) \quad (3.23)$$

where α steers the influence of the regularization term.

In order to find the optimal local model for the inhomogeneities including the neighborhood information (see Eq. (3.23)), we use an iterative minimization strategy. As initialization all coefficients of the sub-regions are set to $a_i^r = 0 \forall r \in \{1, \dots, R\}$. Then, in each iteration, the coefficients of the neighboring regions of the previous step $k-1$ are used to update the coefficients of the current sub-region in iteration k . This is done for all R sub-regions and is repeated until the solution for all sub-problems does not change anymore.

The result of the first optimization step are R sets of coefficients that locally solve the bias correction problem. Due to the locality of the coefficients, the parameters cannot directly be used to do a global intensity inhomogeneity correction. However, as the union of all sub-regions covers the complete region of interest, we can estimate a global parametric model from the local approximations.

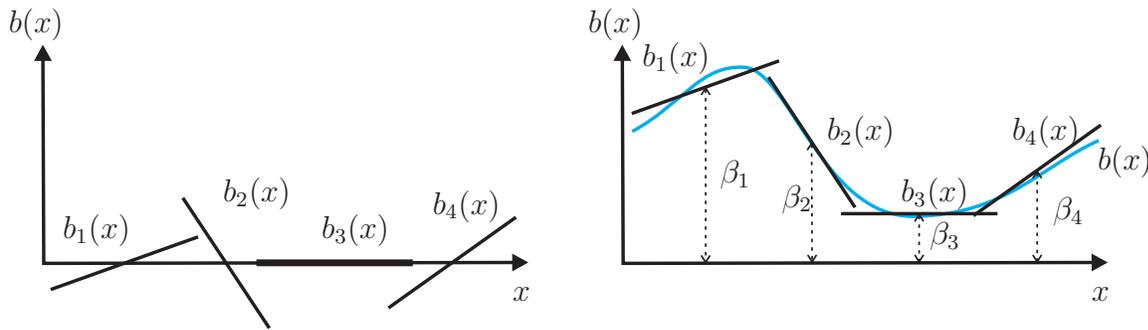


Figure 3.3: Approximation of the additive inhomogeneity model: The left plot illustrates the estimated local bias fields. In order to approximate the global field additive scalars β_r have to be estimated for all local models.

3.3.3 Conquer-Step

In the conquer phase of the proposed method, the solution for all independent sub-problems are combined to form the global correction model. However, due to the invariance of the image data against intensity shifts and scaling the solution of the additive bias correction model is invariant with respect to the summation with a scalar constant (compare Eq. (3.5)):

$$\begin{aligned}
 u(\mathbf{x}) &\equiv s_{\text{MRI}}u(\mathbf{x}) + t_{\text{MRI}} \\
 &= s_{\text{MRI}}(v(\mathbf{x}) - b(\mathbf{x})) + t_{\text{MRI}} = s_{\text{MRI}}(v(\mathbf{x}) - (\check{b}(\mathbf{x}) + \beta)) + t_{\text{MRI}} \\
 &= s_{\text{MRI}}(v(\mathbf{x}) - \check{b}(\mathbf{x})) + \check{t}_{\text{MRI}}.
 \end{aligned} \tag{3.24}$$

The multiplicative model, on the other hand, is invariant to the multiplication with a scalar factor:

$$\begin{aligned}
 u(\mathbf{x}) &\equiv s_{\text{MRI}}u(\mathbf{x}) + t_{\text{MRI}} \\
 &= s_{\text{MRI}}\frac{v(\mathbf{x})}{b(\mathbf{x})} + t_{\text{MRI}} = s_{\text{MRI}}\frac{v(\mathbf{x})}{\check{\beta}b(\mathbf{x})} + t_{\text{MRI}} \\
 &= \check{s}_{\text{MRI}}\frac{v(\mathbf{x})}{\check{b}(\mathbf{x})} + t_{\text{MRI}}.
 \end{aligned} \tag{3.25}$$

For this reason, different strategies for the combination of the local models have to be applied. For both types, the additive as well as the multiplicative model, the local bias fields are evaluated at a number n_r of points uniformly distributed over their region's domain. From this, we receive R sets of n_r bias field values $\mathcal{B}_r = \{\nu_1^r, \dots, \nu_{n_r}^r\}$. Additionally, we get a set $\mathcal{X}_r = \{\mathbf{x}_1^r, \dots, \mathbf{x}_{n_r}^r\}$ of associated global coordinates for all regions.

Additive Inhomogeneity Model

In the *additive case*, as mentioned before, the bias field correction is invariant to the summation of the bias field model with a constant scalar. Because of this and because

of the choice of the origin of the coordinate system of the single sub-regions, neighboring local inhomogeneity models differ by an additive scalar. Thus, the relation between two local bias fields b_l and b_j of two neighboring sub-regions \mathcal{R}_l and \mathcal{R}_j can be described by

$$b_l(\mathbf{x} - \mathbf{x}_c^l) + \beta_l = b_j(\mathbf{x} - \mathbf{x}_c^j) + \beta_j \quad (3.26)$$

where the β_i s are corresponding additive scalars of region \mathcal{R}_i with $i \in \{l, j\}$. Note that the correspondence of the β_i s is related to an arbitrary global shift of the bias field by a scalar constant. The vector \mathbf{x}_c^i defines the origin of the local coordinate system of region \mathcal{R}_i in global coordinates. The relation between the global and the local models can be approximated as

$$b(\mathbf{x}) = b_r(\mathbf{x} - \mathbf{x}_c^r) + \beta_r \quad (3.27)$$

with b being the global bias field as described in the equations (3.1) and (3.15) and b_r being the local bias field of sub-region \mathcal{R}_r with $s \in \{1, \dots, R\}$. The relation between the local bias fields, the summands β_r , and the global correction model is illustrated in Fig. 3.3. In order to be able to transfer the local bias field values \mathcal{B}_r in the global context, for all regions the corresponding β_r has to be estimated. If these are known, the parameter set \mathbf{a} of the global model can be computed in a straight forward manner. In the additive case, the vector of summands $\boldsymbol{\beta} = \{\beta_1, \dots, \beta_R\}$ as well as the parameter set \mathbf{a} can be estimated jointly. Using the sets \mathcal{B}_r and \mathcal{X}_r , the overall optimization objective for the joint estimation can be written as

$$(\mathbf{a}^*, \boldsymbol{\beta}^*) = \arg \min_{\mathbf{a}, \boldsymbol{\beta}} \sum_{r=1}^R \sum_{l=1}^{n_r} (b(\mathbf{x}_l^r; \mathbf{a}) - \nu_l^r - \beta_r)^2. \quad (3.28)$$

where \mathbf{a}^* is the optimal global parameter set of Eq. (3.15), $\boldsymbol{\beta}^*$ is the corresponding vector of summands, and $b(\mathbf{x}; \mathbf{a})$ is the bias value at location \mathbf{x} given the parameter set \mathbf{a} . With $\mathbf{a} = (a_1, \dots, a_{N_u})$ this yields in the 2-d case the following linear problem

$$\begin{pmatrix} (x_{1,1}^1) & (x_{1,1}^1)^2 & \dots & (x_{2,1}^1)^d & 1 & 0 & \dots & 0 \\ (x_{1,2}^1) & (x_{1,2}^1)^2 & \dots & (x_{2,2}^1)^d & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (x_{1,n_r}^1) & (x_{1,n_r}^1)^2 & \dots & (x_{2,n_r}^1)^d & 1 & 0 & \dots & 0 \\ (x_{1,1}^2) & (x_{1,1}^2)^2 & \dots & (x_{2,1}^2)^d & 0 & 1 & \dots & 0 \\ (x_{1,2}^2) & (x_{1,2}^2)^2 & \dots & (x_{2,2}^2)^d & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (x_{1,n_r}^2) & (x_{1,n_r}^2)^2 & \dots & (x_{2,n_r}^2)^d & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (x_{1,1}^R) & (x_{1,1}^R)^2 & \dots & (x_{2,1}^R)^d & 0 & 0 & \dots & 1 \\ (x_{1,2}^R) & (x_{1,2}^R)^2 & \dots & (x_{2,2}^R)^d & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (x_{1,n_r}^R) & (x_{1,n_r}^R)^2 & \dots & (x_{2,n_r}^R)^d & 0 & 0 & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_{N_u} \\ \beta_1 \\ \vdots \\ \beta_R \end{pmatrix} = \begin{pmatrix} \nu_1^1 \\ \vdots \\ \nu_{n_r}^1 \\ \nu_2^1 \\ \vdots \\ \nu_{n_r}^2 \\ \vdots \\ \nu_R^1 \\ \vdots \\ \nu_{n_r}^R \end{pmatrix} \quad (3.29)$$

The system matrix of the linear system is very sparse, as for every region, and thus in every matrix row, there is exactly one β_r unequal to zero. The system can be solved using QR decomposition of the system matrix for instance.

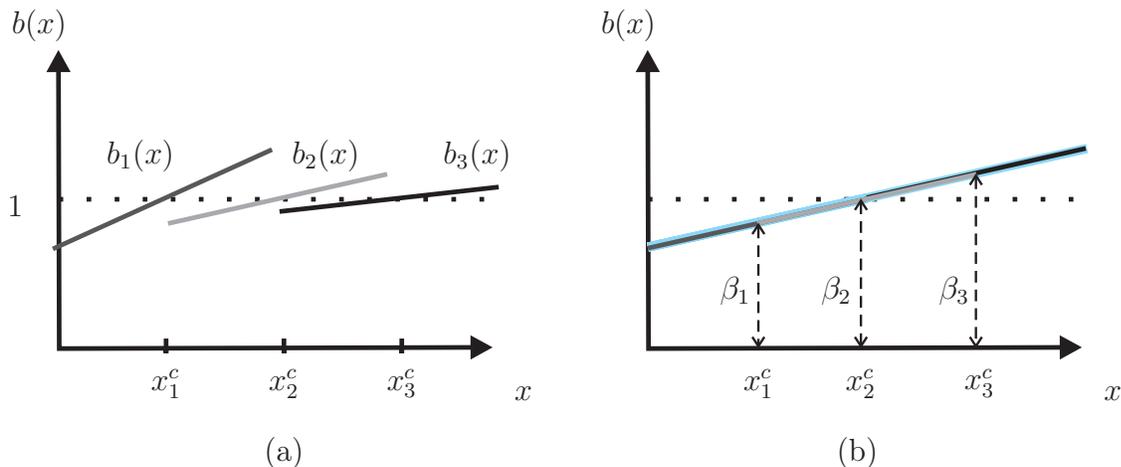


Figure 3.4: Approximation of the multiplicative inhomogeneity model: The left plot illustrates the estimated local bias fields. It can be seen that the slopes of all fields are different. In the right plot the estimated β_r values are multiplied. Now the overall model is a straight line corresponding to the global model.

Multiplicative Inhomogeneity Model

In contrast to the additive case, the bias field in the multiplicative inhomogeneity model is invariant to the multiplication with a scalar factor β . This is a direct result from the missing general intensity scale of MRI images and the multiplicative model defined in Eq. (3.2). Using this and because the local bias field value at the center of every local bias field model equals $b_l(0) = n_e = 1$, the relation of the global to the local bias field can be written as

$$b(\mathbf{x}) = \beta_l b_l(\mathbf{x} - \mathbf{x}_c^l). \quad (3.30)$$

This corresponds to Eq. (3.27) of the additive model. Thus, the relationship between two local polynomial models of neighboring sub-regions can be formulated as

$$\beta_l b_l(\mathbf{x} - \mathbf{x}_c^l) = \beta_j b_j(\mathbf{x} - \mathbf{x}_c^j). \quad (3.31)$$

Note that the correspondence of the β_i s is related to an arbitrary global multiplication of the bias field by a scalar constant. However, due to the multiplicative factors β_r with $r \in [1, \dots, R]$, the optimization process to find the parameter sets \mathbf{a} and $\boldsymbol{\beta}$ is no longer linear. This can easily be explained by the fact that the parameters \mathbf{a}^r are altered by the factors β_r . Consequently, the β_r values are more than a simple “shift”. This is illustrated in Fig. 3.4. In the figure, it is assumed that the global bias field follows an affine model. Because the bias field value at centers of all local bias fields corresponds to $b_l(0) = 1$, the slopes of the estimated fields differ from each other (Fig. 3.4 (a)). After the estimation of the factors β_i the bias fields form the global model (Fig. 3.4 (b)). In this example, the parameter set \mathbf{a} does not have to be estimated, as the degree of the local and global models match.

Instead of using a non-linear optimization scheme for parameter estimation, we divide the parameter estimation in two parts:

1. the estimation of the factors β_r and
2. afterwards, by keeping β fixed, the approximation of the parameters \mathbf{a} of the global bias model.

Using this decoupling of the optimization process, both objectives can be estimated linearly. The major advantage of this is the much smaller computational complexity. For the computation of the factors β_r , we use the logarithmic version of Eq. (3.31). Rearranging the equation yields

$$\log \beta_l - \log \beta_j = \log b_j(\mathbf{x} - \mathbf{x}_c^j) - \log b_l(\mathbf{x} - \mathbf{x}_c^l). \quad (3.32)$$

Here, all unknown β_r are on the left and the known bias values are on the righthand side. From Eq. (3.32) a simple least squares problem can be generated. The error induced by the equation is rather small if the vector \mathbf{x} is in the area of overlap of the two sub-regions l and j . We draw n_o samples from this overlapping area. If this is done for all overlapping areas, we can create a large over-determined linear system of equations. The solution to this system are the $\hat{\beta}_r = \log \beta_r$ values. This requires that all β_r are $\beta_r > 0$. This is a consequence from all local bias fields being $b_r(\mathbf{x}) > 0$ in $\mathbf{x} \in \mathcal{R}_r$. The problem itself is very sparse as there are just two non-zero elements in every line of the system matrix. Using $\hat{b}_i(\mathbf{x} - \mathbf{x}_c^i) = \log b_i(\mathbf{x} - \mathbf{x}_c^i)$ gives the system

$$\begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 1 & -1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & -1 & 0 & \dots & 0 & 0 \\ 1 & 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & -1 & \dots & 0 & 0 \\ \dots & & & & & \\ 0 & 0 & 0 & \dots & 1 & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_R \end{pmatrix} = \begin{pmatrix} \hat{b}_2(\mathbf{x}_1^{1,2} - \mathbf{x}_c^2) - \hat{b}_1(\mathbf{x}_1^{1,2} - \mathbf{x}_c^1) \\ \hat{b}_2(\mathbf{x}_2^{1,2} - \mathbf{x}_c^2) - \hat{b}_1(\mathbf{x}_2^{1,2} - \mathbf{x}_c^1) \\ \vdots \\ \hat{b}_2(\mathbf{x}_{n_o}^{1,2} - \mathbf{x}_c^2) - \hat{b}_1(\mathbf{x}_{n_o}^{1,2} - \mathbf{x}_c^1) \\ \hat{b}_3(\mathbf{x}_1^{1,3} - \mathbf{x}_c^3) - \hat{b}_1(\mathbf{x}_1^{1,3} - \mathbf{x}_c^1) \\ \vdots \\ \hat{b}_3(\mathbf{x}_{n_o}^{1,3} - \mathbf{x}_c^3) - \hat{b}_1(\mathbf{x}_{n_o}^{1,3} - \mathbf{x}_c^1) \\ \vdots \\ \hat{b}_R(\mathbf{x}_1^{R-1,R} - \mathbf{x}_c^R) - \hat{b}_{R-1}(\mathbf{x}_1^{R-1,R} - \mathbf{x}_c^{R-1}) \\ \vdots \\ \hat{b}_R(\mathbf{x}_{n_o}^{R-1,R} - \mathbf{x}_c^R) - \hat{b}_{R-1}(\mathbf{x}_{n_o}^{R-1,R} - \mathbf{x}_c^{R-1}) \end{pmatrix}. \quad (3.33)$$

In order to avoid the trivial solution $\beta = (\mathbf{0}, \dots, \mathbf{0})^T$ an arbitrary β_i can be set to one.

If all β_r are known, the coefficients of the global inhomogeneity model can be estimated by solving

$$\mathbf{a}^* = \arg \min_{\mathbf{a}} \sum_{r=1}^R \sum_{l=1}^{n_r} (b(\mathbf{x}_l^r; \mathbf{a}) - \beta_r \nu_l^r)^2. \quad (3.34)$$

This can be easily written in matrix notation as well. The system matrix looks very similar to the matrix in Eq. (3.29). However, now the β_r are already known. Again, arbitrary matrix solvers can be used to estimate the optimal parameter set \mathbf{a}^* .

3.4 Entropy Optimization using a Histogram-based Regularizer

A problem of many methods for inhomogeneity correction is the lack of the integration of problem specific knowledge into the process of finding the corresponding parameter set. Without this kind of information, it can happen that the correction fails as the optimizer gets stuck in a wrong local or even global minimum. In this section, we present a method that includes prior knowledge into the normalization approach. We model the problem specific information by a histogram representing the image statistics of the reference image. This histogram is introduced as an additional regularizer into the objective function. In the following, we will briefly present Salvado's method [Salv06]. Afterwards, an extension to this approach is described that incorporates the histogram information. Initial correction results using this extension were first described in [Jage08] and [Bald06].

3.4.1 Entropy Minimization

Salvado's method [Salv06] is based on the assumption that the observed images are composed of an ideal image corrupted by multiplicative signal intensity inhomogeneities. Thus, the assumed model corresponds to Eq. (3.2). The bias field itself is modeled by bi-cubic splines. Like introduced in section 3.1 the observed image contains information of the bias field as well as the image content. Thus, the entropy of the corrected image has to be smaller than that of the observed image. Salvado's correction approach iteratively estimates the parameters of a bi-cubic spline model that minimize the entropy in the corrected image. Although theoretically applicable to any dimension, in practice, it is applied to two-dimensional images only due to runtime issues. In literature the method is often referred to as *Local Entropy Minimization with bi-cubic Spline model* (LEMS).

In order to initialize the spline model a polynomial

$$b_0(\mathbf{x}) = \sum_{n=0}^d \sum_{m=0}^{d-n} a_{m,n} x_1^m x_2^n \quad (3.35)$$

is fit to the image data. The degree of the polynomial is denoted by d . Usually, $d = 4$ is used. The unknown coefficients of the polynomial are called $a_{m,n}$. In order to estimate the unknown coefficients a least-squares approximation is used, utilizing SVD for instance.

After the unknown coefficients $a_{m,n}$ are estimated, an initial bi-cubic spline is created. The spacing between the nodes is denoted by s_n . Salvado *et al.* come to the conclusion that in most cases a spacing of $s_n = 21$ mm is a good choice. $\Lambda = \{\kappa_1, \dots, \kappa_{K_n}\}$ is the set of all K_n nodes $\kappa_i \in \mathbb{R}$, with $i \in \{1, \dots, K_n\}$, of the bi-cubic spline model. The pre-sequently computed polynomial bias field is used to initialize all nodes κ_i .

We use a different optimization strategy than Salvado *et al.*. Instead of a golden section search and parabolic interpolation in each iteration, here, a simple gradient descent approach is applied. Furthermore, in [Salv06] a bin size of half a quantization

step size of the images is used. In our experiments it turned out, that a better choice is to compute the histograms with a bin size of double the quantization step size and to employ a partial volume approach to compute the required histograms. Overall, for our data sets, these changes made the approach much more reliable.

The goal of the optimization is to find the bi-cubic spline model that minimizes the objective function O_E . In LEMS O_E corresponds to the Shannon entropy

$$H_{\Gamma}(v; \Lambda) = - \sum_{l=g_{\min}}^{g_{\max}} \mathcal{V}_{\Lambda}[l] \log(\mathcal{V}_{\Lambda}[l]), \quad (3.36)$$

with g_{\min} and g_{\max} being the minimal/maximal bin of the histograms. \mathcal{V}_{Λ} is the histogram of the image v corrected using a spline model with node set Λ . The minimization is done iteratively. In one iteration, all K nodes $\kappa_i \in \Lambda$ are optimized sequentially using a gradient descent approach. The nodes are optimized in a descending order starting with the node that corresponds to the highest intensity within the image. The reason for this is that regions with high intensity in the images usually also have a higher SNR. The nodes are scaled returning a bias field with mean one after each optimization step. After all nodes have been optimized, the next iteration starts. The optimization is stopped, if the entropy changes drops below a threshold ϵ or the maximal number of iterations is reached.

The overall bias field b is obtained by evaluating the estimated bi-cubic spline at each image position.

3.4.2 Histogram-based Regularization

In order to increase the robustness of LEMS, we introduce a new regularizer into the objective function O_E . The regularization term is based on a measure that computes the difference of the histogram of the current optimization stage to a reference histogram $\tilde{\mathcal{U}}$. The reference histogram approximates the histogram of the ideal image u .

There is a close relationship between the histogram difference measure and the entropy measure. The entropy measure computes the inverse distance to a uniform distribution. For the measure, it is assumed that the ideal image contains a single tissue class. Thus, its histogram \mathcal{U} consists of a single peak only. A generalization to this is the proposed histogram-based measure that is founded on the distance to a reference histogram. In the ideal case, the reference histogram is known as prior knowledge.

Basically, there are two sources for the reference histogram $\tilde{\mathcal{U}}$:

1. from reference data and
2. approximated from the histogram \mathcal{V} of the observed image v .

In general, reference data is not available for a specific problem environment. This is mainly a result of the broad range of different acquisition protocols in MRI. A possible source for reference data are atlases or images for which gold standard segmentations are available.

However, the presence of intensity inhomogeneities in the images yields blurred image histograms compared to the ideal histogram (compare section 3.1). This property can be used to approximate the reference histogram $\tilde{\mathcal{U}}$ using sharpening methods applied to the observed histogram \mathcal{V} [Sled 98]. Consequently, the approximated reference histogram can be computed as

$$\tilde{\mathcal{U}} = \mathcal{V} - \alpha_{HP} \mathcal{FT}^{-1} \{ \mathcal{K}_\sigma \cdot \mathcal{FT} \{ \mathcal{V} \} \} \quad (3.37)$$

where \cdot is an elementwise multiplication and α_{HP} is a weighting factor [Gonz 02]. We use a Gaussian highpass filter kernel

$$\mathcal{K}_\sigma = 1 - e^{-\frac{\zeta^2}{2\sigma^2}} \quad (3.38)$$

where σ is the standard deviation of the kernel and ζ is a frequency.

In order to compute the difference between the reference histogram $\tilde{\mathcal{U}}$ and the histogram of the current optimization stage \mathcal{V}_Λ we employ the discrete version of the Jeffrey divergence (JD)

$$D_{\text{JD}}(\tilde{\mathcal{U}}, \mathcal{V}_\Lambda) = \sum_{l=g_{\min}}^{g_{\max}} \tilde{\mathcal{U}}[l] \log \frac{\tilde{\mathcal{U}}[l]}{\mu(l)} + \mathcal{V}_\Lambda[l] \log \frac{\mathcal{V}_\Lambda[l]}{\mu(l)} \quad (3.39)$$

with

$$\mu(l) = \frac{\tilde{\mathcal{U}}[l] + \mathcal{V}_\Lambda[l]}{2} \quad (3.40)$$

as a distance measure between histograms.

The combination of the entropy and the JD-based similarity measure yields the new objective function

$$O_{\text{EH}} = (1 - \alpha) H_\Gamma + \alpha \eta D_{\text{JD}}, \quad (3.41)$$

where α steers the influence of the single components and η adjusts the ranges of both terms. If α is chosen to be $\alpha = 0$, the resulting measure corresponds to the measure proposed by Salvado *et al.* On the other hand, if α is set to $\alpha = 1$ the correction is independent of the image entropy and is based on the histogram distance only.

An estimation for the scaling factor η can be computed beforehand. First, α is set to zero. Thus, the optimization depends on the entropy only. Let ΔH be the resulting entropy difference of the initial to the optimized histogram. Second, α is set to one. Now, only the histogram measure is used. Let ΔD be the resulting difference. The scaling factor η is best set to

$$\eta = \frac{\Delta H}{\Delta D}. \quad (3.42)$$

Consequently, if α is set to $\alpha = 0.5$ yields an equal influence of both terms on the optimization process.

As before, the optimization is done using a gradient descent approach. However, note that these two measures still have to be computed on their own histograms: The entropy is calculated on the local histogram of the region of influence $\Omega_i \subseteq \Omega$ of the currently treated node $\kappa_i \in \Lambda$ whereas D_{JD} is always based on the global histogram.

In the following, we call the proposed method *Local Entropy minimization with Histogram Regularization* (LEHR).

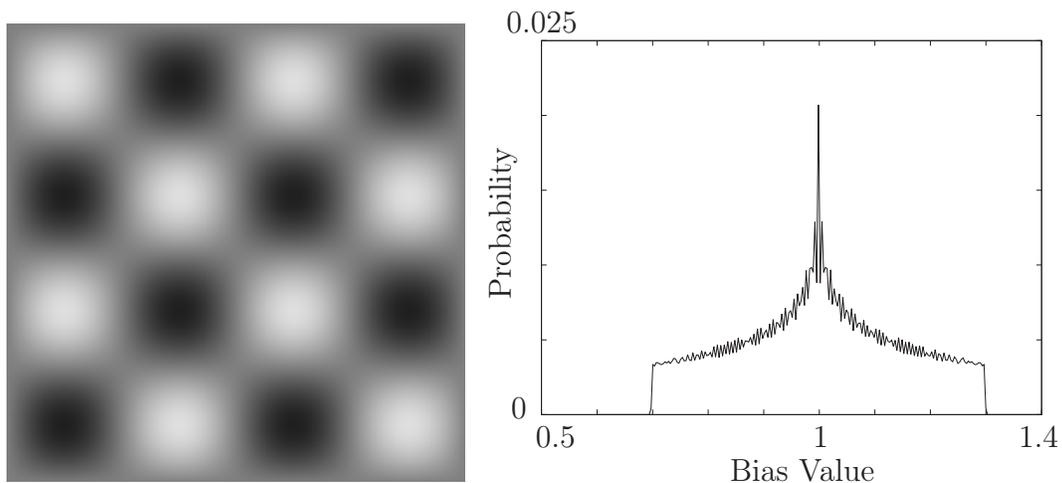


Figure 3.5: Synthetic inhomogeneities: The left image shows the sinusoidal shaped inhomogeneities applied to the checkerboard images. The values range between 0.7 and 1.3. In the right plot the histogram corresponding to the bias field is shown.

3.5 Experiments & Results

In this section, several experiments for the evaluation of the proposed bias correction methods are presented. Each experiment is splitted in a section describing the used data sets, a section introducing the evaluation strategy, and finally a results section.

3.5.1 Divide-and-Conquer-based Bias Correction

In order to evaluate the proposed DaC inhomogeneity correction method, we chose five different scenarios. First, we use synthetic images to illustrate the major pros and cons of our algorithm. This is done for both the 2-d and 3-d case. Then, we apply the proposed correction approach to synthetic 3-d MRI images. Finally, the method is used for the inhomogeneity correction of clinically relevant 3-d T1w MRI data sets as well as 3-d TOF MRA images. The experiments were performed on a 2.00 GHz Intel Core2 CPU with 2 GB RAM. DaC is implemented in C++ and integrated in the ITK framework¹. The bias correction of an image with a size 256×256 , a sub-region size of 40, and an overlap of two requires approximately one second.

Synthetic Images

Checkerboard image:

Data Sets: We use two different settings for the experiments using artificial data sets. The images in the first experiment are composed by two intensity classes that are arranged in a checkerboard-like fashion. The intensity classes simulate two different tissue types. The lower intensity has a value of $i_{\text{low}} = 100$ and the higher intensity is $i_{\text{high}} = 150$. The intensities are distorted by a multiplicative sinusoidal inhomogeneity field. Its frequency is slightly lower than the frequency of the checkerboard. The

¹<http://www.itk.org>

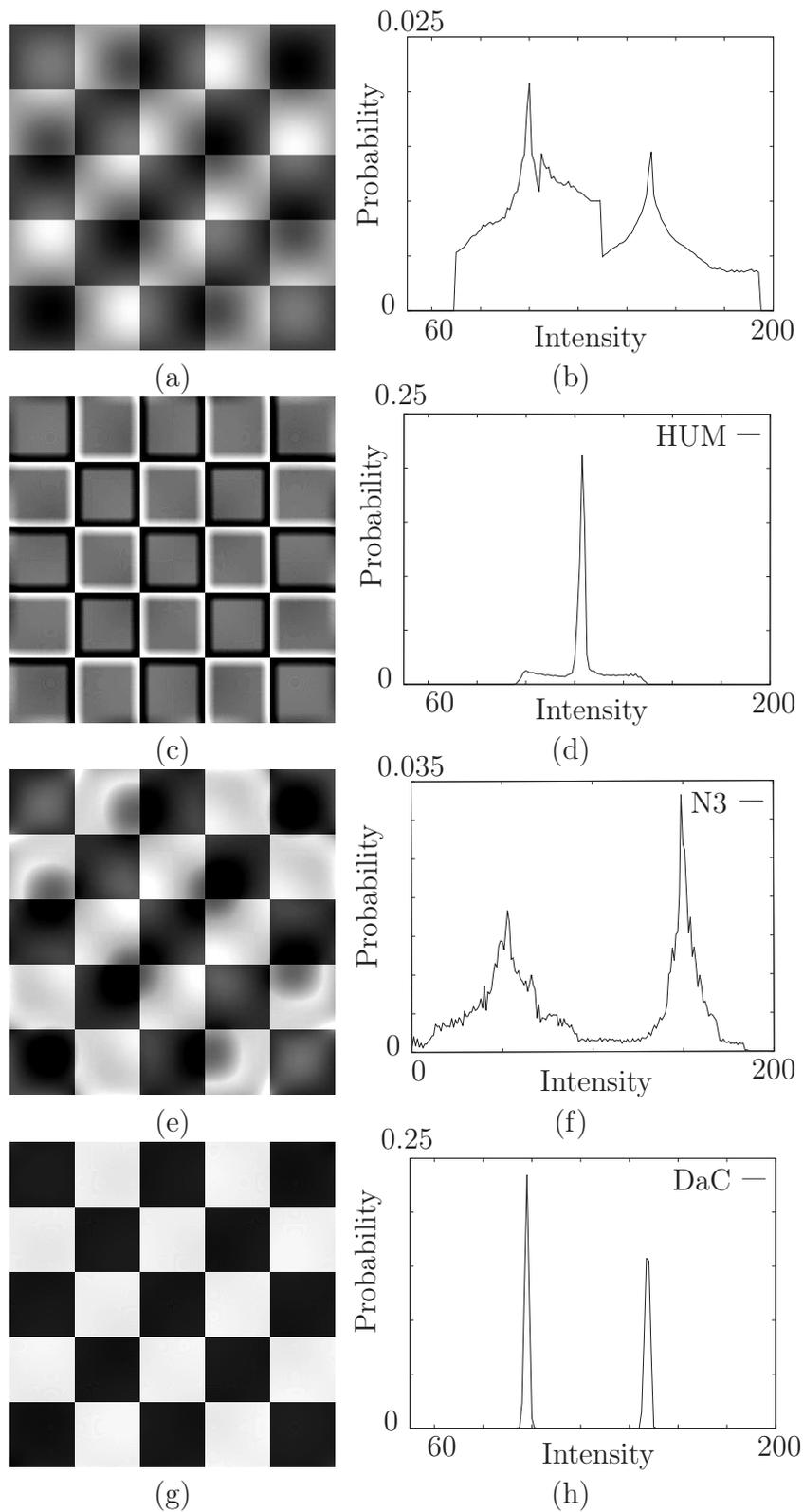


Figure 3.6: Correction of synthetic checkerboard images: Left: The resulting checkerboard images are shown. Right: The histograms of the images to the left. In the first row the checkerboard affected from the artificial bias field is shown. In the second row, the results of the correction using HUM are illustrated. Next, the results using N3 and finally the result using the proposed DaC approach.

strength of the inhomogeneity field is 60%. This means that the intensities of the bias field are bound between the values 0.7 and 1.3. The overall design is chosen in such a way that the correction is difficult for homomorphic approaches [Well96]. Segmentation-based as well as histogram-based methods, on the other hand, can more easily solve this problem. The used bias field and its histogram are shown in Fig. 3.5. The distorted image is shown in the upper left and its histogram in the upper right of Fig. 3.6.

Evaluation Method: Here, a qualitative evaluation is done only. In order to compare the results of the different inhomogeneity correction techniques, the resulting images as well as their histograms are shown.

Results: As expected the correction using HUM yielded very unconvincing results. The initial inhomogeneity field is removed, but due to the structure of the image, the image intensities within regions of a single intensity are altered incorrectly. It can be observed that the intensities are shifted towards the mean of the image. This effect becomes more and more obvious if the filter size is decreased. In literature, this effect is called edge effect [Vovk07]. In the histogram of the HUM-corrected image the shift of the intensities towards the mean $\mu_m = 124$ can be seen. The result is shown in Fig. 3.6 (c) and the corresponding histogram in Fig. 3.6 (d).

N3 succeeds in separating most of the pixels of both classes. However, the inhomogeneity correction is not satisfactory. The resulting image as well as its histogram are illustrated in Fig. 3.6 (e) and (f). The reason for this behavior is that N3 sequentially estimates the bias field by a sharpening of the original histogram. Though, the artificial bias field is so strong that a good estimation of the underlying ideal histogram cannot be computed that way. Consequently, the image cannot be restored completely.

The correction results using the proposed DaC approach is very convincing. It can be seen that the tissue classes are completely separated. Only small intensity variations are still visible within the image histogram. These variations have the following reasons: In multiple steps of the proposed method a discretization of the data is done inducing a certain discretization error. More important, the inhomogeneity field has a sinusoidal shape, however, the correction is done using a polynomial model. This cannot compensate all the intensity variations induced by the sinusoidal model. The correction result of DaC is illustrated in the last row of Fig. 3.6.

Random Intensities:

Data Sets: In the second experiment using synthetic data, the intensities of a 3-d volume are randomly drawn from a pre-defined intensity distribution. The distribution chosen corresponds to the histogram of a real 3-d MRI head data set. The data set containing the random numbers is then altered using a sinusoidal inhomogeneity field. A slice through the image data is shown in Fig. 3.7 (a). In Fig. 3.7 (b) the probability density of the undistorted volume is plotted. The random structure of the images is very advantageous for homomorphic approaches because they rely on uniformly distributed intensities throughout the whole volume domain. Moreover, as the histograms are independent of the underlying structure within the image, histogram-based methods are not influenced by the randomness, too. On the other hand, approaches that base on a segmentation of the data sets cannot be applied in this scenario. Again, we compare the proposed DaC approach, HUM and finally N3.

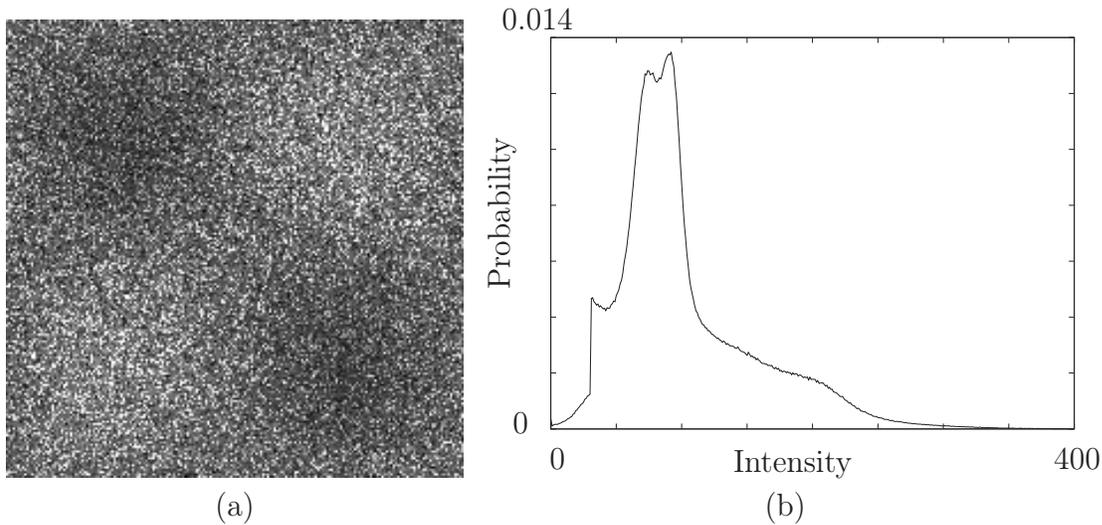


Figure 3.7: Slice through the synthetic volume: The image on the left side shows a slice through a synthetic volume. The intensities are randomly drawn from the density shown on the right side. Moreover, the volume was altered by a 3-d sinusoidal bias field. The density shown corresponds to a real MRI head image.

Evaluation Method: As a measure for the quality of the inhomogeneity correction, we use the standardized deviation from the known ground truth inhomogeneity field $\tilde{b}(\mathbf{x})$ to the estimated bias fields $b(\mathbf{x})$. The measure corresponds to the quality criteria described in Sled *et al.* [Sled98]. In a first step the ratio $r_{\mathbf{x}} = b(\mathbf{x})/\tilde{b}(\mathbf{x})$ at every pixel location is computed. Then, the mean value μ_r and the standard deviation σ_r for all $r_{\mathbf{x}}$ are calculated. In order to measure the quality q of a correction result, we compute

$$q = \frac{\sigma_r}{\mu_r}. \quad (3.43)$$

Thus, q -values close to zero indicate a good correction, larger values a bad one.

Results: The results show that there is no significant performance difference between the proposed method (DaC) and HUM. The q values for the methods are for DaC $q_{\text{DaC}} = 0.030$ and for HUM $q_{\text{HUM}} = 0.039$. A slice through the estimated 3-d inhomogeneity fields can be seen in Fig. 3.8. The resulting estimation of the inhomogeneity field using the N3 method is slightly worse. Here, the correction yielded a q -value of $q_{\text{N3}} = 0.093$. Like in the previous experiment, the reason for the correction problems of N3 are due to the sequential estimation of the bias field. Nevertheless, the shape of the inhomogeneities for all evaluated methods corresponded to the shape of the ground truth field. The estimated bias fields of the three methods as well as the reference field are shown in Fig. 3.8.

Synthetic MRI brain images

Data Sets: The synthetic MRI images used are taken from the *BrainWeb: Simulated Brain Database*² of the McConnell Brain Imaging Center of the Montreal Neurological

²<http://www.bic.mni.mcgill.ca/brainweb/>

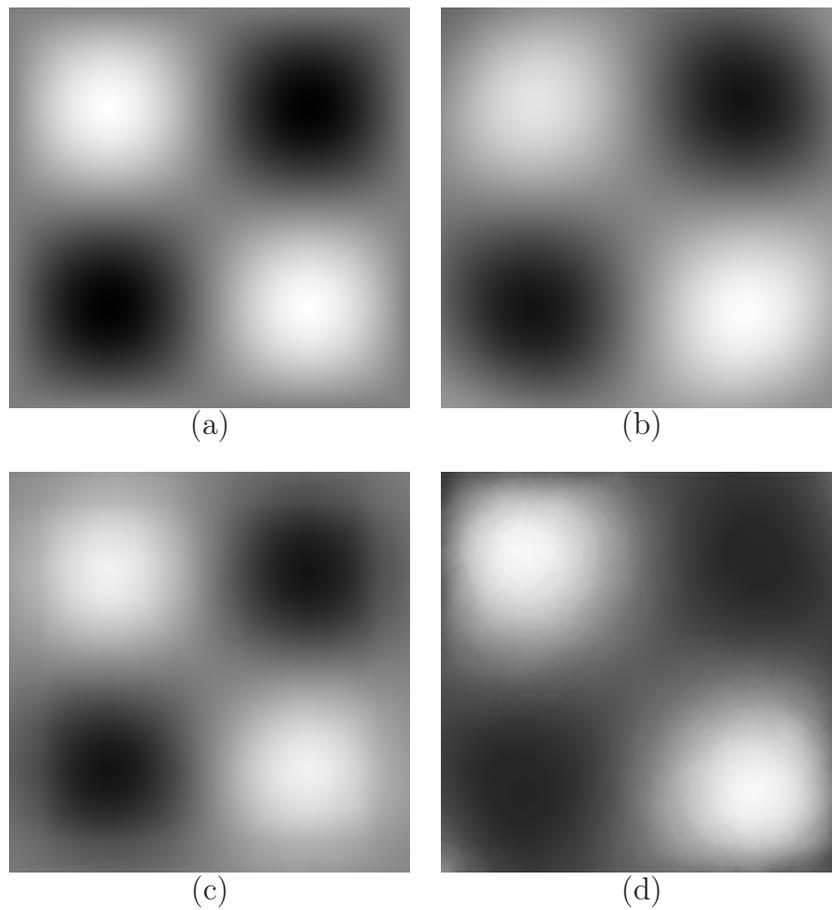


Figure 3.8: Estimated bias fields of synthetic 3-d data: (a): Slice through ground truth bias field. (b): Estimated field using DaC. (c): Estimated field using HUM. (d): Estimated field using N3. It can be seen that all algorithms recover the structure of the ground truth field quite well.

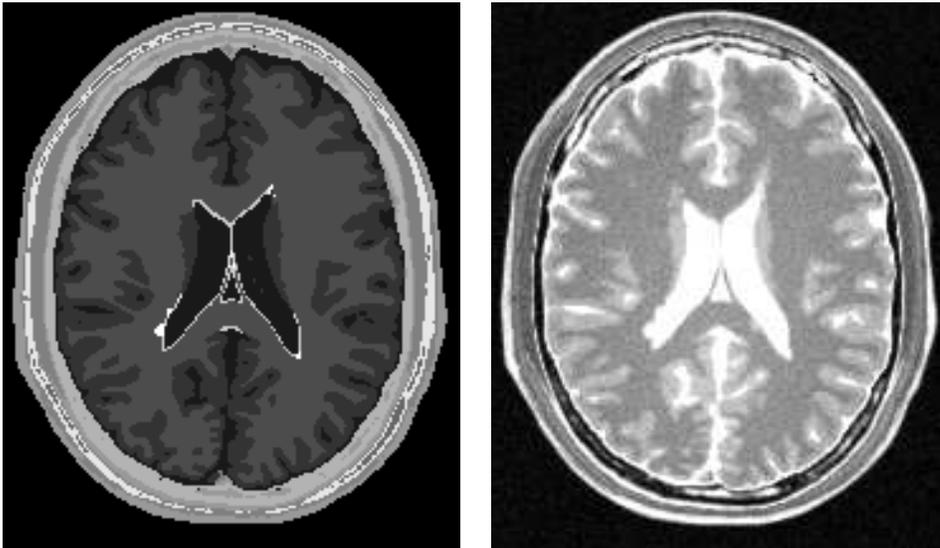


Figure 3.9: Ground truth segmentation (left) and simulated T2w MRI brain data (right). The images are taken from the BrainWeb database [Coll98].

Institute, McGill University [Coll98]. In total we use 12 different data sets. Six data sets are simulations of normal brains and the rest simulates a brain with Multiple Sclerosis. All data sets have an isotropic resolution of 1 mm and a size of $181 \times 217 \times 181$. For each scan T1-weighted, T2-weighted and Proton Density (PD) images were available. The strength of the simulated bias field is 0% and 40% for all scans. A level of 40% means that the values of the applied multiplicative bias field range between 0.8 and 1.2. An example T2w slice and its segmentation is shown in Fig. 3.9.

Evaluation Method: In order to evaluate the inhomogeneity correction, we use two quality measures. The first one is the coefficient of variations $cv(C)$ of a tissue class C . It describes the normalized standard deviation within a single class. The normalization of $\sigma(C)$ is done by its mean $\mu(C)$. Thus, the criteria can be written as

$$cv(C) = \frac{\sigma(C)}{\mu(C)}. \quad (3.44)$$

The coefficient is invariant to uniform linear intensity transformations. To be able to use this criterion, a segmentation of the tissue classes is needed. For the synthetic MR images ground truth data is available. We apply this criteria to the white and gray matter tissue class. However, there are cases where the intra-class intensity deviation is reduced by the inhomogeneity correction but on the other hand the inter-class distance is worsened. For segmentation purposes this “class overlap” is extremely important as it describes the separability of the tissue classes. In order to deal with the class overlap, we use a second quality measure, the so-called coefficient of joint variations (see [Lika01]):

$$c_jv(C_1, C_2) = \frac{\sigma(C_1) + \sigma(C_2)}{|\mu(C_1) - \mu(C_2)|}, \quad (3.45)$$

	Normal Brain						Brain with MS Lesions					
	T1w		T2w		PD		T1w		T2w		PD	
	0%	40%	0%	40%	0%	40%	0%	40%	0%	40%	0%	40%
<i>cv(GM)</i>												
Orig.	9.9	13.5	17.7	20.3	4.5	9.7	10.1	13.7	9.7	13.7	3.8	9.4
DaC	9.9	9.9	17.8	17.9	4.5	4.5	10.1	10.1	9.8	9.8	3.9	3.8
N3	9.9	10.0	17.7	20.0	4.6	4.6	10.1	10.2	9.9	10.1	4.0	4.0
HUM	14.0	13.7	20.2	20.2	5.4	5.1	14.2	14.0	11.3	11.1	4.6	4.4
M4*	9.9	10.0	17.8	17.8	4.5	4.5	10.1	10.2	9.9	9.9	3.9	3.8
FMI*	11.5	13.7	22.3	20.3	6.3	6.9	13.2	11.4	15.0	11.3	6.3	6.8
<i>cv(WM)</i>												
Orig.	5.2	9.2	12.0	13.3	4.9	7.5	5.4	9.3	8.3	10.6	4.5	7.4
DaC	5.0	5.0	12.1	11.6	4.9	4.6	5.2	5.1	8.4	8.1	4.6	4.3
N3	5.2	5.2	12.0	13.0	4.9	4.6	5.3	5.3	8.4	8.2	4.5	4.3
HUM	8.4	8.1	14.4	14.2	5.3	4.8	8.5	8.2	9.3	9.0	4.9	4.4
M4*	5.1	5.0	12.2	11.8	4.9	4.7	5.3	5.3	8.5	8.4	4.5	4.3
FMI*	7.3	9.2	17.6	14.8	6.6	6.1	8.5	6.9	12.4	9.6	6.9	6.3
<i>cjv</i>												
Orig.	51.6	69.3	83.2	106.4	64.9	163.0	50.9	68.0	74.9	123.8	66.9	195.6
DaC	52.0	51.5	82.7	83.8	64.4	64.2	51.4	50.9	74.4	75.6	66.2	66.8
N3	51.9	51.8	83.1	103.9	66.4	66.7	51.3	51.1	74.9	77.8	68.7	68.7
HUM	95.0	90.6	106.0	107.9	75.4	73.5	92.5	88.2	90.7	90.8	76.6	74.5
M4*	52.1	51.7	82.5	82.4	65.0	64.1	51.5	51.1	74.3	74.2	67.0	66.1
FMI*	72.5	96.4	125.5	116.8	93.5	107.9	88.0	65.3	150.3	98.3	117.3	125.9

Table 3.1: DaC-inhomogeneity correction results (simulated MR brain images). All values given are in %. The results for methods marked with * are taken from [Lika 01].

with C_1 and C_2 being the two regarded tissue classes. It again is invariant against linear intensity transformations.

For the evaluation DaC is compared to the following state-of-the-art inhomogeneity correction methods: *HUM* [Brin 98], N3 [Sled 98], *Intensity Inhomogeneity Correction by Information Minimization* (M4) [Lika 01], and finally a method called *Fast Model Independent* (FMI) [Voku 99].

Results: The results are shown in Table 3.1. The results labeled by a star are taken from an article of Likar *et al.* [Lika 01]. The results show that DaC, M4 and N3 yielded very similar correction results. After the inhomogeneity correction all *cv*s and *cjv*s are very close to the coefficients of the undistorted images. For instance, the average deviation of the *cjv* of the corrected images to the *cjv* of the ground truth images are $cjv_{\text{DaC}} = 0.7\%$ for DaC, $cjv_{\text{M4}} = 0.8\%$ for M4 and $cjv_{\text{N3}} = 3.5\%$ for N3. On the other hand, the correction results of HUM and FMI are not convincing. Here, the average deviations of the *cjv* to the undistorted coefficients are $cjv_{\text{HUM}} = 39.1\%$ for HUM and $cjv_{\text{FMI}} = 60.5\%$ for FMI. For HUM this is a logical consequence of the results of the pure synthetic data sets. It can be observed, that the and FMI even corrupt images containing no intensity inhomogeneities. All other methods kept these images nearly constant.

Real MRI brain images

In this evaluation section, we use two types of real clinical relevant MRI images. First, we evaluate the inhomogeneity correction algorithms using two data sets taken

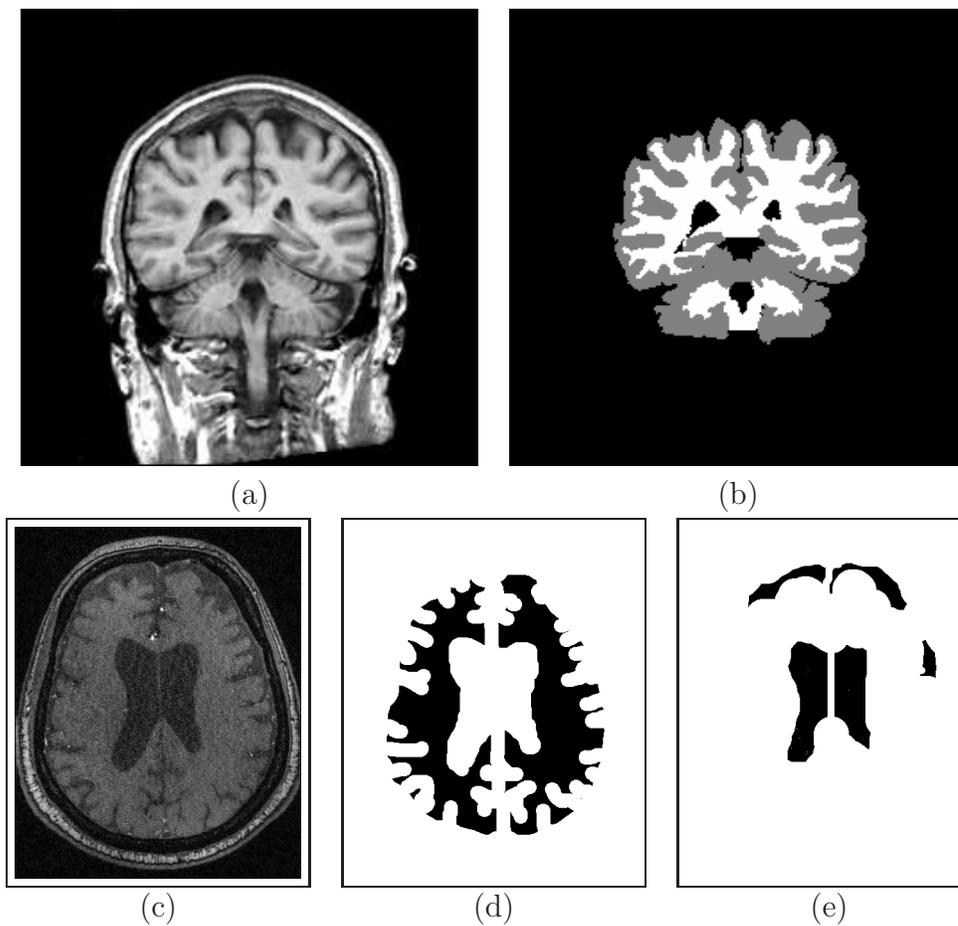


Figure 3.10: Real MRI data sets with gold standard segmentation: Upper row: T1w slice through clinical relevant human head scan (left) and its corresponding ground truth segmentation (right). The images are taken from the IBSR database. Lower row: Slice through a 3-d TOF MRA image (left). Segmentation of WM/GM brain tissue (middle) and rough segmentation of CSF (right).

from the *Internet Brain Segmentation Repository* (IBSR)³ of the Massachusetts General Hospital, Center for Morphometric Analysis. For both data sets a gold standard segmentation including the classes “white matter”, “gray matter” and “other” is available. In Fig. 3.10 (a) a slice through one of the data sets and in (b) through its segmentation is shown. Second, we use 12 TOF MRA data sets that were acquired at the Department of Radiologic Sciences at the University of California, Los Angeles (UCLA). For these data sets no complete gold standard segmentation is available. However, for all data sets 3-4 slices that are segmented into the classes CSF and brain tissue (gray and white matter) are available. An example slice and its segmentations can be seen in the Figures 3.10 (c), (d) and (e).

IBSR:

Data Sets: The first IBSR data set, labeled 788, used shows the brain of a male subject. The scanner used is a 1.5 Tesla General Electric Signa machine. The acquisition parameters for the T1-weighted coronal scan are TR=40 ms and TE=5 ms. The reconstructed in-plane resolution is 1.0 mm². The slice thickness of the image is 3.0 mm. The size of the reconstructed volume is 256 × 256 × 60. Further, the methods are evaluated using a T1-weighted IBSR data set, labeled 1320, of a brain of a 5 year old male subject. Again the Signa scanner is used for acquisition. The parameters are TR=40 ms and TE=5 ms. The in-plane resolution is 0.9375 mm × 0.9375 mm and the image has a slice thickness of 1.5 mm. The data set has a resolution of 256 × 256 × 128.

Evaluation Method: In order to evaluate the correction of the data sets in the IBSR database, we use the coefficient of variation (see Eq. (3.44)) as well as the coefficient of joint variation (see Eq. (3.45)) introduced in the previous section as quality criteria for the correction results. In order to compare DaC, we also compute the correction results for N3 and HUM.

Results: The correction results are illustrated in the first two lines of Table 3.2. For data set 788 it can be observed that all methods have a similar correction performance. The unprocessed data set itself is affected by a very weak bias field. Thus, the overall change induced by the methods is rather small. The results for image 1320 show that the coefficient of variation is reduced equally by N3 and DaC. However, the *cjv* is lower using DaC compared to N3. In both cases, HUM has a negative effect on the separability of gray and white matter.

3-d TOF MRA:

Data Sets: The 3-d TOF MRA data sets were acquired in clinical routine at UCLA on a 3T TrioTim Siemens scanner. All data sets show the head region. The scanning parameters were TR=24 ms and TE=3.6 ms. The in-plane resolution of the data sets was around 0.41 mm × 0.9375 mm with a slice thickness of about 0.64 mm. The size of the data sets was $\approx 416 \times 512 \times 144$.

Evaluation Method: The results of the correction experiment using 3-d TOF MRA images are illustrated in Table 3.2. In order to do a quantitative evaluation of the proposed method, for all N_t data set the coefficient of variation, compare Eq. (3.44), for CSF and brain tissue (GM/WM) is computed. Additionally, we calculate the coefficient of joint variation *cjv* for these tissue classes to measure the influence of the bias correction onto subsequent processing stages. In order to summarize all three

³<http://www.cma.mgh.harvard.edu/ibsr/>

	Orig.	DaC	HUM	N3	Orig.	DaC	HUM	N3	Orig.	DaC	HUM	N3
IBSR	<i>cv(WM)</i>				<i>cv(GM)</i>				<i>cjv(GM,WM)</i>			
788	22.0	21.5	21.7	22.3	10.5	9.5	10.1	10.6	94.8	93.5	97.3	93.9
1320	19.5	15.5	16.0	15.4	12.9	9.0	9.2	9.2	85.0	80.9	85.9	84.8
UCLA	<i>cv(Brain)</i>				<i>cv(CSF)</i>				<i>cjv(Brain,CSF)</i>			
P1	21.0	18.2	18.6	18.4	31.7	31.4	53.6	31.5	87.5	77.9	98.0	78.1
P2	19.6	15.5	16.4	15.4	33.7	33.0	60.6	32.8	77.4	65.4	89.5	66.1
P3	21.5	18.0	18.7	18.4	31.3	31.3	56.0	31.4	86.3	77.4	111.1	80.4
P4	29.0	23.8	27.4	24.8	39.3	38.5	71.4	38.2	128.5	94.8	123.0	106.2
P5	21.6	17.1	17.5	17.4	33.9	35.3	62.2	34.7	78.8	65.2	92.8	71.8
P6	23.2	17.9	18.7	29.4	33.9	34.9	60.3	37.6	81.0	73.2	95.4	312.9
P7	26.6	21.4	25.9	22.9	33.8	34.9	76.1	33.0	107.2	85.1	116.1	97.4
P8	20.6	17.8	18.6	18.5	33.7	30.7	58.2	31.7	109.3	81.5	119.3	95.0
P9	21.8	17.6	18.3	18.3	31.1	29.5	45.2	26.2	89.5	72.9	113.0	77.3
P10	18.3	15.0	15.1	14.9	31.0	26.6	42.2	26.9	95.9	72.2	97.3	77.8
P11	20.0	17.6	17.6	17.8	37.1	34.1	53.7	33.6	91.9	76.4	114.2	80.9
P23	19.5	16.1	16.5	16.3	27.9	27.9	48.5	27.3	80.5	69.5	114.0	78.3

Table 3.2: DaC-inhomogeneity correction results (clinical relevant MR brain images). All values given are in %.

coefficients in a single measure for each class, we compute the average ratio of the coefficient of (joint) variation after and before correction:

$$\mu_{cv} = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{cv_{\text{corr}}(i)}{cv_{\text{orig}}(i)} \quad (3.46)$$

where $cv_{\text{corr}}(i)$ is the cv/cvj of the corrected image i and $cv_{\text{orig}}(i)$ is the cv/cvj of the original image i . Moreover, we do a k-means clustering of the images to further illustrate the influence of the bias correction on segmentation and classification methods.

Results: The cjv numbers show that there is a significant improvement of the separability of the two considered tissue classes using DaC as well as N3. Although the correction succeeded for both methods, the results indicate that DaC performs better than N3. Especially, the cjv value of subject P6 shows that N3 can fail in some cases whereas DaC shows more robustness. The average ratio is $\mu_{cgv,N3} = 113.2\%$ (excluding P6: $\mu_{cgv,N3} = 88.4\%$) for N3 and $\mu_{cgv,DaC} = 81.8\%$ for DaC. As expected the separability of the tissue classes are worsened using HUM. Here, the average ratio is $\mu_{cgv,HUM} = 116.6\%$.

Similar results can be seen regarding the composite class containing GM and WM. Both algorithms DaC and N3 performed well, however, it can be clearly seen from the resulting cv values that DaC does in most of the cases better and overall a more reliable correction. The average ratios in percentage for these algorithms are $\mu_{brain,N3} = 88.2\%$ for N3 and $\mu_{brain,DaC} = 82.8\%$ for DaC. The variation coefficient for the CSF class does not show a significant improvement for both algorithms. In this case, the results of N3 are slightly better than the ones of DaC. The percentages for the two tested methods are $\mu_{csf,N3} = 96.6\%$ for N3 and $\mu_{csf,DaC} = 96.9\%$ for DaC.

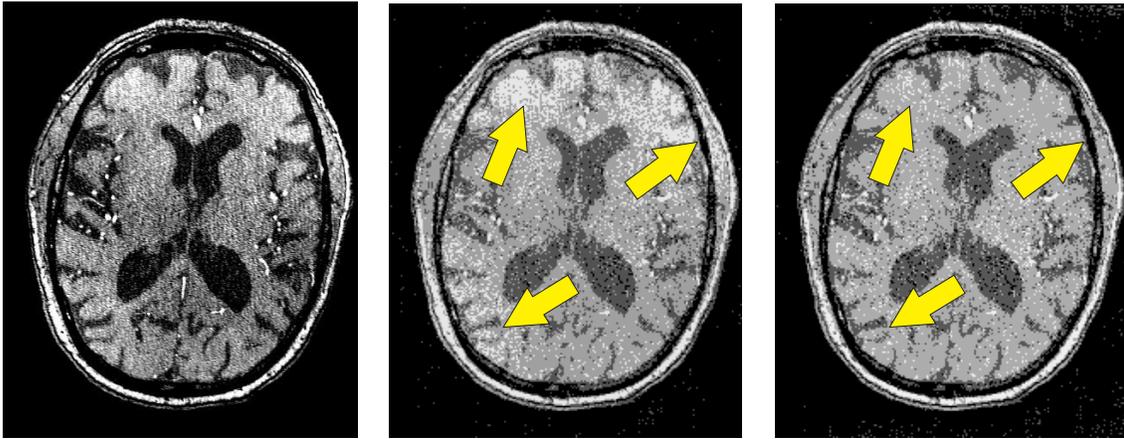


Figure 3.11: K-means clustering of TOF images: The left image shows a slice through a MRA TOF volume. The clustering result in the middle image was achieved using no pre-processing of the volume. The right image shows the k-means clustering result on bias corrected image utilizing DaC.

There are two simple reasons for this. First, CSF appears very dark in TOF images. Thus, the multiplicative bias field has not an equally strong impact on it as it has on other “brighter” tissue classes. Second and more important, the segmentation of CSF is mainly composed of the ventricles that have a relatively small size compared to the whole volume and that are basically located within a few slices only. Due to the slow varying bias fields, the effect on the ventricles is rather small. Looking at the correction results of HUM shows, that the method works well for the dominant brain tissue class but significantly worsens the variation coefficient for the *CSF* class. The average ratios are $\mu_{brain,HUM} = 86.8\%$ and $\mu_{csf,HUM} = 182.0\%$. This results from edge effects that dominate the correction of the small CSF class. The larger class, on the other side, is hardly affected by this.

In order to further illustrate the effect of the bias correction on a subsequential processing of the TOF data sets, we do a k-means clustering of the images. The clustering can be mathematically formulated as the minimization problem

$$J = \sum_{i=1}^R \sum_{\mathbf{x} \in \Omega} \|v(\mathbf{x}) - c_i\|^2 \quad (3.47)$$

where R is the number of classes and c_i is the intensity prototype of class i . We use four classes: background, CSF, brain tissue (WM,GM) and vessels. An illustrative result is shown in Fig. 3.11. The left image shows the original uncorrected slice. In the middle image the clustering result without any bias correction is shown. In the last image the clustering result after bias correction using DaC is shown. Looking at the front region of the brain (arrows in Fig. 3.11) indicates that the clustering result significantly improves using bias correction as a pre-processing step.

3.5.2 Histogram-based Regularization

Especially in combination with the histogram-based regularization, Salvado's method has a rather high complexity. Thus, 3-d correction is not feasible. For this reason, we present results of 2-d experiments only. First, we evaluate the method on synthetic data sets. Here, we utilize the same data sets as for the evaluation of DaC. This is followed by experiments on real clinically relevant images. Quantitative results are presented for slices of 3-d TOF images of head. Further on, we qualitatively evaluate the method on slices of 3-d SPACE images of the spine. The experiments were performed on a 2.00 GHz Intel Core2 CPU with 2 GB RAM. The method is implemented in C++ and integrated in the ITK framework⁴. The bias correction of an image with a size 256×256 , and a node spacing of 50 requires approximately three minutes.

Synthetic Data Sets

Data Sets: The synthetic data sets correspond to the artificial images used in the experiments of section 3.5.1. First, we perform the evaluation on checkerboard images corrupted by a sinusoidal bias field. The frequency of the sine wave is slightly lower than the frequency of the checker board. The lower intensity is $i_{\text{low}} = 100$ and the higher intensity is $i_{\text{high}} = 150$. In the second experiment, we use random intensity values. The distribution corresponds to a real clinical head data set (compare section 3.5.1). However, in this section we do no 3-d evaluations. Thus, just a single slice was computed and altered by a sinusoidal bias field. The pdf as well as an example image showing a random distribution and a sinusoidal distortion is shown in Fig. 3.7.

Evaluation Method: In the experiment using the checkerboard images, only qualitative results are shown. As reference we use both, the ground truth histogram as well as the artificially sharpened image histograms. Like in section 3.5.1, we utilize the normalized relative deviation of the gold standard bias field to the estimated one in order to evaluate the correction quality of the data sets containing random numbers.

Results: In the checkerboard experiment, it can be observed, that LEMS ($\alpha = 0$) yielded very good results. Both classes are successfully separated. Similar to the DaC approach smaller deviations from the correct values can be observed (Fig. 3.12 (a)). The corresponding image histogram is plotted in Fig. 3.12 (d). LEHR achieved similar results utilizing a gold standard reference and an α -value of $\alpha = 1$ (Fig. 3.12 (b)). Only in the bottom of the image slight intensity deviations can be seen. This, however, is a result of the precision of the optimization. In Fig. 3.12 (e) the gold standard histogram as well as the histogram of the corrected image are plotted. If the required reference histogram is approximated by a sharpening of the observed histogram, the results become significantly worse (Fig. 3.12 (c)). The corresponding reference histogram as well as the histogram of the corrected image are shown in Fig. 3.12 (f). The reason for the worse result is that the ideal histogram cannot be approximated in a satisfactory manner. Thus, there are more intensities relevant for the correction method than barely the two gold standard intensity peaks. This result

⁴<http://www.itk.org>

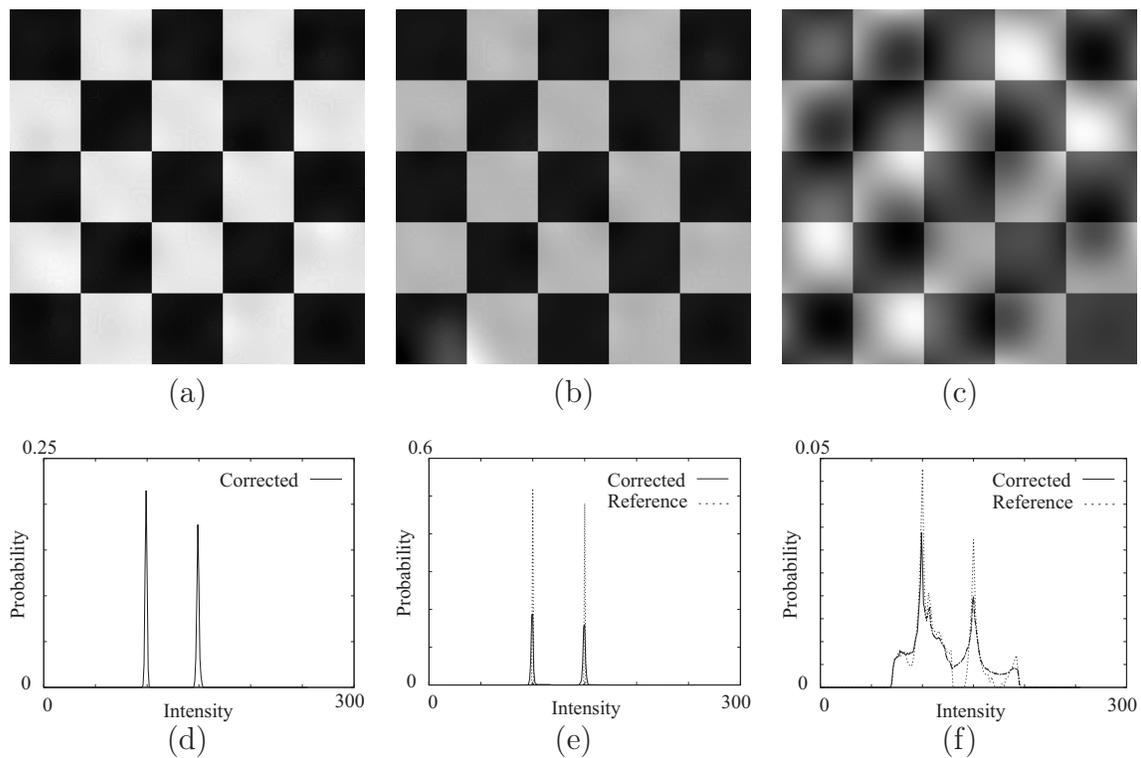


Figure 3.12: Checkerboard results for entropy and histogram optimization: In the upper row the corrected images using (a) LEMS, (b) LEHR ($\alpha = 1$) using a ground truth histogram and (c) LEHR ($\alpha = 1$) using an approximated ground truth histogram are shown. The plots (d), (e) and (f) correspond to the histograms of the respective images. Furthermore, in (e) and (f) the used reference histograms are shown.

is closely related to the correction results of N3 in section 3.5.1. N3 has the same problems correcting this synthetic data set as it also relies on a sharpening of the input histograms.

In the experiments using randomly generated images, the results for LEMS and LEHR are differing very much. The quality criterion Eq. (3.43) yielded for LEMS $q_E = 0.170$. As we are correcting 2-d images only, the measured q -numbers cannot be compared in a one-to-one fashion to the numbers of DaC, N3 and HUM in section 3.5.1. However, q_E is significantly worse compared indicating problems in the bias field estimation of LEMS. A reason for this is the missing regularization during a single optimization step, even though all nodes of the bi cubic spline are rescaled after every iteration to satisfy the mean preservation criterion.

LEHR, on the other hand, achieved very good correction results. Using an approximated reference histogram a q -value of $q_{AH} = 0.036$ can be achieved for $\alpha = 1$. Surprisingly, the result using ground truth histogram is with $q_{GTH} = 0.044$ slightly worse. An explanation for this behavior is that in the case of using a ground truth reference, in the close proximity of the correct solution the distance values between the reference and the corrected histograms are very small. Thus, the optimization slows down and finally it stops. In the case of an approximated reference histogram the standard deviation of a tissue class is usually smaller than in the ground truth histogram. Thus, the optimization stops at a later stage.

The best correction result is obtained for LEHR using the ground truth histogram as reference and a weighting of $\alpha = 0.6$. Here, a q -value of $q_{GTH,0.6} = 0.024$ is achieved. The best configuration for the approximated histogram is $\alpha = 0.5$. This results in a q -value of $q_{AH,0.5} = 0.031$. These values are slightly better than the results for DaC, N3 and HUM. However, as mentioned before, they are not directly comparable. They are just an indication that the performance of the histogram regularized optimization is close to the performance of the before mentioned algorithms.

Clinical Relevant Data Sets

Slices from 3-d TOF images

Data Sets: In correspondence to the experiments evaluating DaC in section 3.5.1, we use images of the head acquired with a 3-d TOF sequence at the Magnetic Resonance Research Center at UCLA for evaluation. Although, the images are acquired using a 3-d sequence only single slices are used for the experiments as the complexity of the evaluated correction algorithms is too large for bigger data sets. For all selected slices reference segmentations including the classes brain containing WM and GM as well as the class CSF are available.

Methods: As only single slices are used, the CSF class is small and restricted to a small area. Thus, its coefficient and the coefficient of joint variation representing the separability of the classes are rather unreliable. For this reason, only the coefficient of variation (compare Eq. (3.44)) of the large brain class is used for measuring the correction quality. However, the other coefficients are given as well for completeness.

Results: The correction results LEMS and LEHR are shown in Fig. 3.13. In order to give a better impression of the performed bias correction, a threshold is applied to the images. The arrows in the image point to regions where the differences of the correction using LEHR, $\alpha = 1$, and LEMS ($\alpha = 0$) can be seen best. The first column

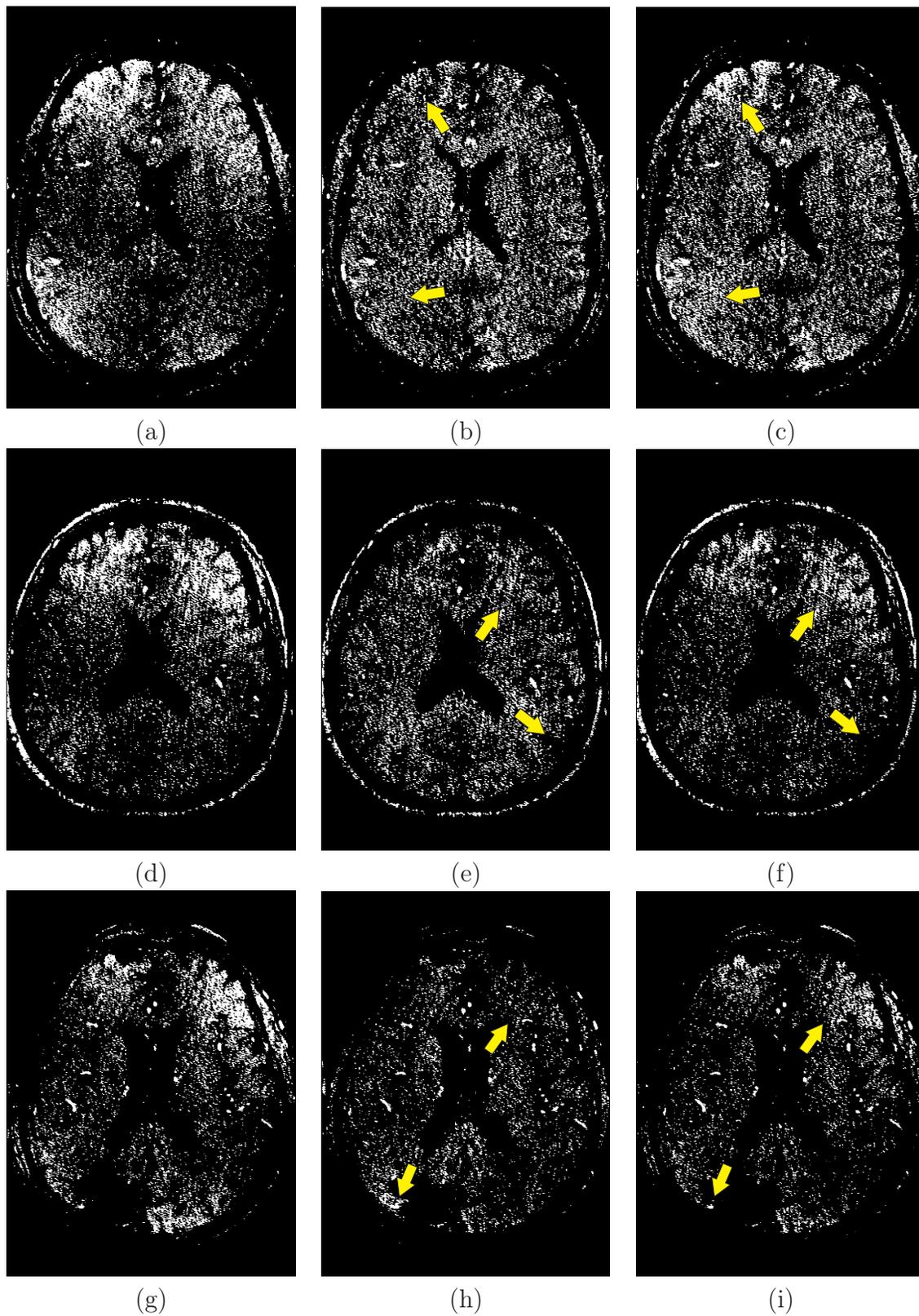


Figure 3.13: 3-d TOF results for entropy and histogram optimization: The images (a), (d) and (g) show the original uncorrected slices. In the images (b), (e) and (h) the correction results using the histogram optimization are illustrated. Finally, in the images (c), (f) and (i) the results using entropy optimization are shown. The arrows indicate regions of interest where the differences between the corrections can be seen best.

shows the original uncorrected slice. In the middle column, images (b),(e) and (h) show the result using LEHR and the images (c),(f) and (i) illustrate the correction using LEMS. By comparing the results, it can be seen that LEHR performed better than LEMS. This result is supported by the coefficient of variation of the brain tissue class. To compute the overall performance, the mean of all coefficients of variation is computed. The mean coefficient using LEHR is $\mu_h(\text{Brain}) = 0.156$. On the other hand, the mean coefficient using LEMS resulted in $\mu_e(\text{Brain}) = 0.164$. The best value yields $\alpha = 0.75$ with $\mu_{75}(\text{Brain}) = 0.155$. The coefficients of variation for the small CSF class resulted in $\mu_h(\text{CSF}) = 0.308$ for $\alpha = 1$ and $\mu_e(\text{CSF}) = 0.313$ for $\alpha = 0$. The best value was achieved for $\alpha = 0.75$ again with $\mu_{75}(\text{Brain}) = 0.307$. As mentioned before, the results for the CSF class have to be handled with care due to its size. The coefficient of joint variation resulted in $\mu_h(\text{CSF}, \text{Brain}) = 0.717$ using the histogram measure and $\mu_e(\text{CSF}, \text{Brain}) = 0.692$. The best result yielded $\alpha = 0.25$ with $\mu_{25}(\text{CSF}, \text{Brain}) = 0.676$.

Slices from 3-d SPACE images

Data Sets: In this experiment, slices from 3-d SPACE images acquired in clinical routine are used. Compared to the bias fields visible in the TOF images from the previous experiment, here, the intensity inhomogeneities present in the images are rather large. The slices used are either coronal or sagittal. They have a size of 384×384 with an in-plane resolution of approximately $1.04 \text{ mm} \times 1.04 \text{ mm}$. The imaging parameters were TE=130 ms and TR=1000 ms with a flip angle of 150° . Example images can be seen in the Figures 3.14 (a) and (d). All images were acquired at the Radiologic Institute at the University of Erlangen.

Evaluation Method: As there is no gold standard segmentation or bias field available for these images, a qualitative evaluation is done only. Before correction all image intensities below a certain threshold are set to zero in the images to concentrate the computations on foreground pixel. For visual evaluation LEHR is compared to LEMS.

Results: The results of the inhomogeneity correction using LEMS as well as LEHR ($\alpha = 1$) are shown in Fig. 3.14. In the first column, Figures 3.14 (a) and (d), the original images are displayed. In the images Fig. 3.14 (b) and (e) the resulting images using LEHR are illustrated. Finally, in the last column, Figures 3.14 (c) and (f), the corrected images using LEMS are given. In this experiment, it can be seen clearly that the results achieved using LEMS are much more reasonable than the corrections using LEHR. The reason for the big correction difference is that the approximated reference histograms do not match the ideal ones very well. Because of the extreme intensity inhomogeneities, the ideal histogram cannot be resolved by a simple sharpening of the input histograms. Thus, the correction does not cover the large global intensity variations induced by the bias field in the images. The entropy optimization on the other hand does not suffer from these drawbacks as it does not rely on a previously approximated reference. Regions of interest that illustrate the correction differences best are marked by arrows within Fig. 3.14.

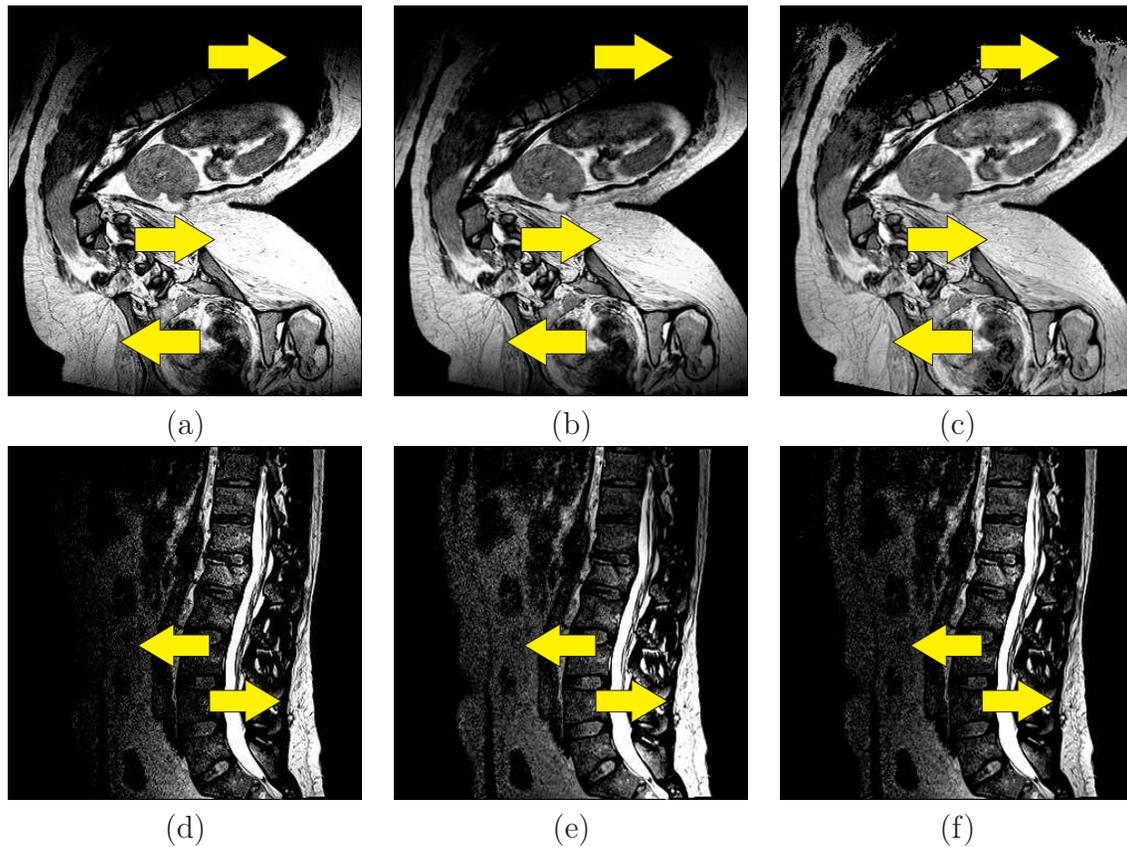


Figure 3.14: 3-d SPACE results for LEMS and LEHR: The images (a) and (d) show the original uncorrected slices. In the images (b) and (e) the correction results using LEHR are illustrated. Finally, in the images (c) and (f) the results using LEMS are shown. The arrows indicate regions of interest where the differences between the corrections can be seen best.

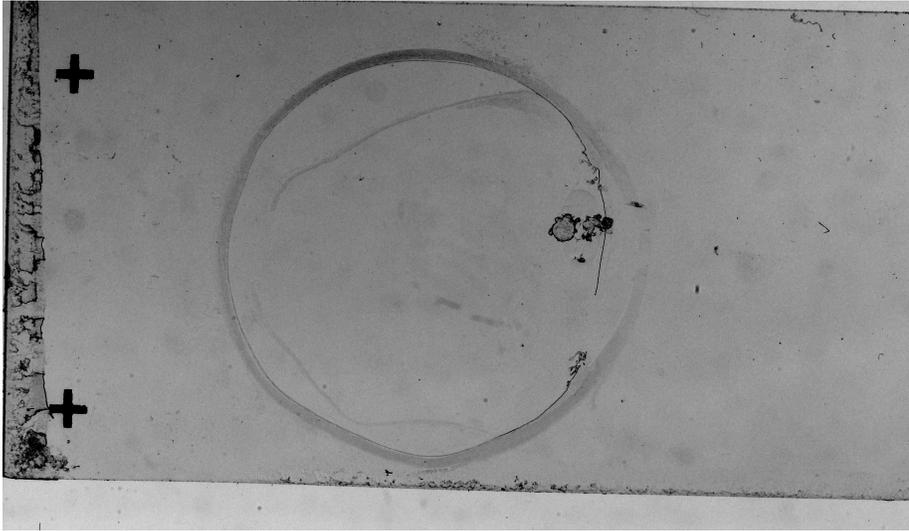


Figure 3.15: Microslice photograph of a slice through the human eyeball.

3.6 Application to Other Modalities

In this section, we show the applicability of the DaC approach to other modalities that suffer from similar signal intensity inhomogeneity problems. First, we present experiments for the correction of intensity inhomogeneities of microphotographs of cuts through the eye. Here, the intensity variations are mainly due to soiling after the slicing. Thus, the inhomogeneities are assumed to be additive. The second experiment deals with retina fundus images. The retina is illuminated by a single flash. The reflected light is acquired using a CCD camera. The main distortions are due to the optical system of the human eye. Here, the additive as well as the multiplicative model are applied in literature [Yous08, Hoov03]. We assume the multiplicative model as otherwise there is no contrast enhancement in regions that were initially dark. However, one has to keep in mind the worsening of the SNR that can negatively affect subsequent pre-processing steps. Both intensity distortions have in common that they are rather complex. For this reason bias correction methods, that rely on very low polynomial approximations cannot be used in these cases, M4 proposed by Likar *et al.* [Lika01] for instance.

3.6.1 Microslice Photographs

Data Sets: In Fig. 3.15 a typical microslice photograph of a human eye is shown. The image has a size of 3801×2201 pixel. It is obvious that there is a principal tendency that the upper left part of the image is darker than the lower right area. The images were acquired by S. Gaffling, Erlangen Gradient School in Advanced Optical Technologies (SAOT), University of Erlangen and G. Michelson, Interdisciplinary Center of Ophthalmic Preventive Medicine and Imaging, University of Erlangen.

Evaluation Method: As there is no gold standard available, we cannot do a quantitative analysis of the methods. Here, however, the results show significant differences.

Consequently, the pros and cons of each method can be illustrated in a good manner on this clinically relevant data set.

We evaluate HUM and the proposed DaC approach for the correction of additive intensity inhomogeneities within the images. In the version described in section 3.2, HUM corrects multiplicative intensity inhomogeneities. For the current experiment, we change the correction of an observed intensity χ_o with a local mean μ_l and a global mean μ_g in the images to $\chi^* = \chi_o + \mu_g - \mu_l$ so that HUM can handle additive inhomogeneities.

Results: In Fig. 3.16 the correction results of both methods are illustrated. In order to visualize the results in a more convenient way, we applied a threshold to the images. For each image it was chosen so that the background shows approximately the same level of masked pixel in all images. Fig. 3.16 (a) shows the thresholded unprocessed image. Here, the non-uniformities can clearly be seen. Image (b) corresponds to the correction result of the proposed DaC approach. In Fig. 3.16 (c) the HUM corrected image is presented. Both methods removed the global ramp very well. However, locally the proposed method outperforms HUM. In the areas marked by the arrows 1, it can be seen that HUM emphasized the border by increasing the background intensity next to darker structures. This is known as edge effects. The reason for this is that due to the lower signal intensities of the structure itself the local mean is decreased. Thus, after the correction these parts are brightened. The formally bright homogeneous area (arrows 2) is falsely darkened to adapt it to the global mean value. The main problem is that therewith also anatomical structures are altered. This can be seen in the area of the lens (arrow 3). The correction result of the proposed method does not show these drawbacks.

3.6.2 Retina Fundus Images

Data Sets: For the second 2-d experiment, we use retina fundus images provided by the Shiley Eye Center of the University of California, San Diego in the scope of the STARE (Structured Analysis of the Retina) Project⁵. Example images are shown in the left column of Fig. 3.17.

Evaluation Method: As in the previous experiment no gold standard was available for the correction. For this reason only a qualitative evaluation is done. In order to evaluate the proposed DaC approach, we compare it to HUM.

Results: The correction results of HUM and DaC can be seen in Fig. 3.17. The middle column corresponds to DaC and the right column to HUM respectively. In order to visualize the differences between the correction results, we applied a threshold to the images. The threshold for each image is chosen so that the difference between the methods can be observed best. Like in the previous experiment it can be observed that the results of the DaC approach are more reasonable from an anatomical viewpoint. In many HUM corrected images the edge effects dominate the crossing between two types of tissue classes. This can be seen especially in images including pathologies or next to large vessels. Moreover, wrong intensity shifts can be observed, if two large regions with different dominating tissue types are visible within

⁵<http://www.ces.clemson.edu/~ahoover/stare/>

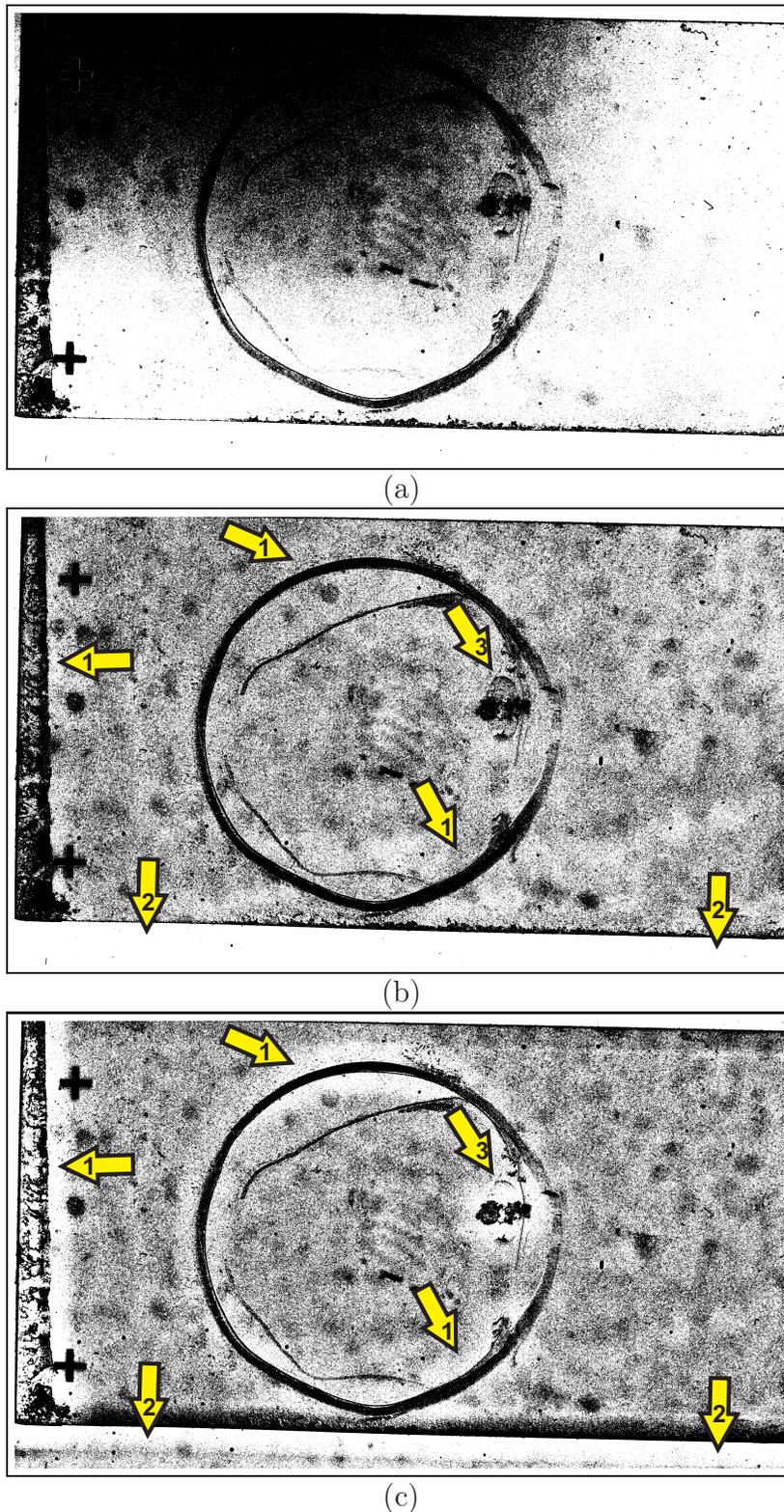


Figure 3.16: Bias field correction results for microslice photographs: The image (a) shows the thresholded unprocessed image. Image (b) shows the DaC corrected image. Image (c) is the correction result of HUM. The arrows indicate the main differences between the two correction methods: 1: edge effects; 2: two large classes with different intensities; 3: changes in the image content.

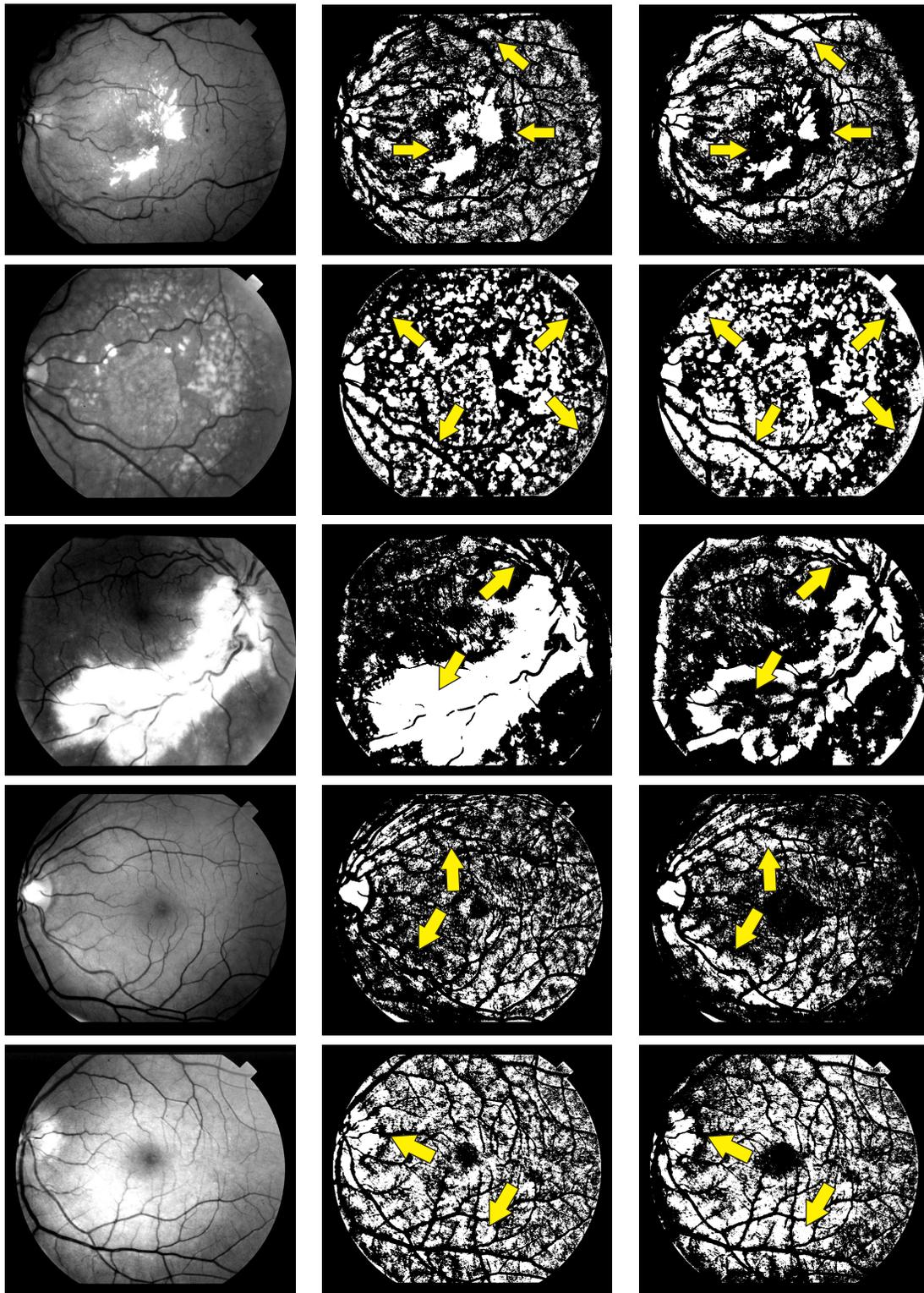


Figure 3.17: Bias correction of retina fundus images: Left column: Green channel of retina fundus images. Middle column: Correction result using the DaC approach. Right column: Correction result using HUM. All images are taken from the STARE database. The arrows emphasize regions where the differences between the methods can be clearly seen.

the images. The proposed approach does not have these drawbacks. In the images of Fig. 3.17 this can be best seen at the tips of the red arrows.

3.7 Summary & Discussion

Many MRI images suffer from severe intra-volume intensity inhomogeneities. In order to improve the performance of post-processing algorithms like segmentation or image registration, these artifacts have to be removed. In the last two decades several methods have been developed to deal with bias fields. They can be classified in the categories retrospective and prospective. Many state-of-the-art methods are either developed for a few specific applications and/or they are not precise enough. We presented in this section two novel approaches for bias correction. DaC is a fast and very flexible method that relies on the divide-and-conquer principle. LEHR, on the other hand, introduces prior knowledge into the estimation of the inhomogeneities. Consequently, the correction includes a signal intensity standardization component. For more information about standardization see section 4. However, the computational complexity of LEHR is rather large. Thus, its practical relevance is low.

The most noticeable property of DaC is its generalizability. The experiments show that it is for all kind of data one of the best methods for the estimation of intensity inhomogeneities. It can easily deal with large as well as weak bias fields. Further on, the method is, compared to many other entropy based approaches, rather fast due to its divide-and-conquer design. On the other hand, it has all the advantages of entropy based algorithms, like the independence of large homogeneous tissue regions, in contrast to HUM, for instance. The experiments show that DaC can increase the average separability of tissue classes in 3-d TOF MRA images by approximately 18.2% whereas N3, for instance, could only achieve 11.6%.

LEHR suffers of similar issues as N3. Both methods are extremely dependent on the approximation of the reference histogram. A good approximation of the reference is possible for weak bias fields only. If the inhomogeneities are too large, the single tissue classes cannot be correctly separated in the histograms leading to bad correction results. If, on the other hand, the inhomogeneities are rather small then the estimation of the reference histograms is quite good and reliable. Thus, due to the additional knowledge, the correction is very good. Available ground truth data is always advantageous if integrated as an additional regularizer in the estimation of the intensity inhomogeneities. However, for almost all applications there is no ground truth available.

The choice of the right correction method is highly dependent on the subsequent application and the statistics of the images. In general, methods basing on the images' entropy are more precise and robust against (pathological) changes in the data. Moreover, they have better generalization properties than other methods. Their complexity, however, is usually much higher than that of other methods. For instance, HUM, basing on the smoothness assumption, can be implemented very efficiently. Its correction results are not very convincing in an anatomical sense. But, it is sufficient for several post-processing methods. N3 and LEHR rely on the estimation of reference histograms. If these can be provided in a reliable manner, by using ground truth data, for instance, the results are very good. This, however, is not possible in all cases.

The performance of N3 is usually rather good, whereas LEHR is too slow for clinical relevant applications. Although, only slightly slower than N3, DaC generalizes best of all tested inhomogeneity correction methods. The correction results are in the great majority of cases convincing in a medical sense. The reason for the good performance of DaC is that the entropy optimization is solved using very simple models. The composition of the global model is done after all local models are estimated.

Chapter 4

Standardization of MRI images

Following definition 1 describing the normalization of MR images in section 1.2, besides the intra-scan intensity correction, the inter-scan intensity correction is a very important pre-processing step in MR imaging. The major source of these variations is the calibration of the MR system and intensity non-uniformities. A similar problem that is solved by intensity standardization methods, is the correction of the slab boundary artefact (compare section 2.3.6). The methods introduced in this chapter deal with these problems. Although a lot of research was done in the field of intensity inhomogeneity correction, compare chapter 3 for instance, the standardization of MR images did not get the same attention so far. In general, a general intensity scale has no direct impact on medical diagnostics by experts; however, volume renderers cannot use standard presets (transfer functions) to visualize certain organs or tissue classes. The physician has to adjust the settings for every single scan. Furthermore, more sophisticated automatic segmentation and quantification methods are needed, as they have to adapt their parameters to the observed image intensities. Additionally, currently a new class of hybrid imaging systems combining MR and Positron Emission Tomography (PET) is being developed. In order to increase the PET image quality, an standardized attenuation correction utilizing the MR data has to be performed [Zaid07]. For this purpose the MR intensities have to be mapped to attenuation coefficients which correlate to tissue classes. Furthermore, Madabhushi *et al.* conclude in [Mada05] that a subsequent application of both approaches, inhomogeneity correction and intensity standardization, has major advantages for most processing methods, like segmentation or registration. In this chapter, we first define the standardization problem. This is followed by a short overview of state-of-the-art algorithms. Next, a 1-d histogram matching approach is presented. One advantage of MR is the possibility to obtain images of a body region using different acquisition protocols emphasizing different types of tissue. We present a method for the joint standardization of all acquired spatially coherent images. Furthermore, an approach for the inter-scan correction of whole body MR images is introduced. Afterwards the methods are evaluated and the results are presented. The experiments show that our method achieves an average intensity overlap of the major tissue classes of T1w images taken from the public available IBSR database of about 86.2%. The most commonly used state-of-the-art method resulted in only 70.1% overlap. The section is concluded by a brief summary and conclusions.

4.1 Problem Definition

Inter-scan intensity variations are a well known artifact in image processing. In general, they describe intensity differences of pixel belonging to a single homogeneous class in two or more different images. In MR imaging the classes correspond to tissues, like white matter, fat, etc., present in the images.

Let $\{(\chi_1, \xi_1), (\chi_2, \xi_2), \dots, (\chi_k, \xi_k)\}$ be a set of k pairs of corresponding intensities in two images v and u with χ being the intensities in v and ξ being the intensities in u respectively. In the following the image v is called target image and u is denoted as reference or source image. The goal of intensity standardization is to find a mapping $\varphi : \Gamma \mapsto \Gamma$ of the intensities of the target image to the intensities of the source image such that

$$\varphi(\chi_i) = \xi_i \quad \forall i \in \{1, \dots, k\}. \quad (4.1)$$

$\Gamma \subset \mathbb{R}$ is the intensity domain of the images. For medical applications it is a necessity that φ is a bijective mapping. The reason for this constraint is that otherwise information included in the images can get lost due to the standardization process, for instance, if intensities of two different tissues are mapped to the same intensity in the output image. Furthermore, let $p_v : \Gamma \mapsto [0, 1]$ be the pdf of the target image and $p_u : \Gamma \mapsto [0, 1]$ the pdf of the reference image respectively.

In practice, in many images pathologies are present. These pathologies can have a major impact on the shape of the images' histograms. It is very important that these structural characteristics are preserved during the standardization process. This is a very challenging task as the content of the images and thus their histogram characteristics are not known beforehand. Moreover, anatomical structures can be visible in the reference histogram and not in the target histogram or vice versa.

In contrast to the correction of intensity inhomogeneities, it is assumed that the standardization mapping performs a global correction. This means, that the intensity mapping is independent of the spatial position \mathbf{x} within the images v and u . Thus, it is assumed that there are no spatial intensity inhomogeneities present within the images.

4.2 State-of-the-Art

The most commonly used intensity standardization method is histogram equalization [Gonz 02]. As reference it assumes a uniform distribution of the image intensities in the final image. The mapping can be written as

$$\xi_i = \varphi_{\text{HE}}(\chi_i) = \int_0^{\chi_i} p_v(w) dw \quad (4.2)$$

with p_v being the pdf of v and w being a dummy variable for integration. In the continuous case the inverse of the mapping can be computed. If p_v is larger than zero in the covered intensity range, the mapping is even strictly monotonically increasing. Consequently, the inversion is a single valued function. In the discrete case, however, usually the mapping cannot be inverted. Furthermore, due to the uniform mapping onto the interval $[0, 1]$ structural information, about pathologies for instance, is lost.

For these reasons, histogram equalization is not suitable for the usage in medical applications if the information about tissue intensities in the data sets is explicitly used in subsequent processing stages.

A generalization of histogram equalization is histogram specification, also called histogram matching. In contrast to histogram equalization an arbitrary pdf can be assumed as reference. Equivalent to Eq. (4.2) the pdf of the reference image can be transformed to a uniform distribution. Let the corresponding mapping be called ϑ . Using ϑ , the histogram specification can be written as

$$\xi_i = \varphi_{\text{HS}}(\chi_i) = \vartheta^{-1}[\varphi_{\text{HE}}(\chi_i)]. \quad (4.3)$$

First, the distribution of v is mapped to a uniform distribution. Afterwards, the uniform distribution is mapped by ϑ^{-1} to the distribution of the reference. Like before, the mapping φ_{HS} is exact in the continuous case, for discrete images, however, this cannot be guaranteed. Nevertheless, it is only invertible if ϑ is strictly monotonic. For further details we want to refer to [Gonz 02]. Although a reference histogram can be specified, the structural information is not preserved if it is not contained in the reference histogram. Moreover, as histogram specification maps the histograms very precisely, differences within the images have a huge impact on the standardization results.

In the following, we briefly describe state-of-the-art methods that were explicitly designed for application in MRI. They try to overcome the problems associated with invertibility and the loss of structural information by introducing prior knowledge about the problem domain into the standardization process.

A histogram based method is proposed by Pierre Hellier [Hell 03] who, given an image, estimates a mixture of Gaussians that approximates the histogram of the image. The fitting is done utilizing the approach presented in [Van 99b]. He then computes a polynomial correction function that aligns the mean intensities of the different tissue classes. Unfortunately, this approach is applicable to the head region only, as this is the only region where the tissue classes can be estimated by a few distinct categories. Furthermore, it is highly dependent on the quality of the result of the fitting of the Gaussians.

Weisenfeld and Warfield propose a combined intensity standardization and inhomogeneity correction method in [Weis 04]. They estimate a multiplicative correction field that adapts the intensity statistics of an acquired MR volume to a previously created model. As a model they employ the histogram of a previously acquired image with appropriate image properties. The correction is achieved by minimizing the Kullback-Leibler divergence between the model and the template intensity distribution. The standardization method has been evaluated on brain scans but can potentially be applied to other regions of the body, too. However, the computational complexity of the method is rather high. Thus, it is not suitable for the daily clinical routine.

A signal intensity standardization method using spatial tissue correlations between a reference and a template image is presented in [Schm 05] by M. Schmidt. In order to match the images, a non-linear registration algorithm is used. Once the images are aligned, a scalar multiplicative correction weight is computed. Though potentially applicable to all body regions, the standardization is highly dependent on

the result of the registration step. Thus, the results will be worse in regions of the body with significant anatomical differences and deformations, like the thorax or the abdominal area or regions including anomalies as well as pathologies.

The intensity standardization methods described in Nyúl *et al.* [Nyl00] and Ge *et al.* [Ge00] are based on a 1-d histogram matching. Currently, Nyúl's method is the most frequently used standardization approach in MR imaging. First, they detect landmarks (e.g., percentiles, modes) on a training set of histograms. Then all detected landmarks are averaged, in order to generate a standard landmark set. When a new image is acquired, the detected landmarks of its histogram are matched to the previously computed standard positions. Finally, in order to create a continuous intensity mapping, the histogram positions between the landmarks are linearly interpolated. This standardization method has been evaluated on brain scans but can potentially be applied to other regions of the body, too. Nevertheless, for every new body region and protocol, the corresponding histograms have to be analyzed and appropriate landmarks have to be chosen. The reliable detection of the landmarks is a very challenging and essential task, as the quality of the standardization heavily depends on it. This can be seen as the major drawback of this method. Moreover, the intensities between the landmarks are linearly interpolated, in practice, however, these intensity deformations are non-linear.

We presented a comparison of several methods for intensity standardization at the conference “Bildverarbeitung für die Medizin 2008” [Berg08]. In the remainder of this chapter we introduce several approaches for intensity standardization that do not rely on any assumptions about the shape of the images' histograms. Thus, our methods are completely independent of the application, region of interest and acquisition sequence, as long as there are reference histograms available for the given task. These can be easily computed by choosing representative images for the new body regions and/or acquisition protocols. Because only histograms are employed, no non-rigid spatial alignment of the data sets has to be done. For this reason, deformations or pathologies hardly influence the normalization results. To the best of our knowledge, all state-of-the-art algorithms standardize the observed intensities using a single image at a time and ignore spatially adjacent images. For many applications this is sufficient, because in many regions of the body a gray value in one image acquired with a specific pulse sequence is associated with exactly one intensity in another pulse sequence. This, however, is not always the case. We introduce a novel method that can jointly standardize all acquired images. Moreover, the proposed method is the only approach that can deal with whole body MRI data sets, which have much more complex statistical properties than spatially constrained images, like the head.

4.3 Matching of 1-D Histograms

In this section, an approach for the standardization of single spectral data sets is introduced. The method itself is split into two parts; first an affine alignment is done to overcome larger intensity deformations. After this coarse matching, a non-parametric non-rigid alignment of the target with the reference histogram is performed. The separation of the estimation of the matching into the described parts has the advantage that the rough alignment can be done in a very fast manner. Moreover, in order to

make the non-parametric alignment more stable it can be regularized stronger if only small changes have to be considered.

Throughout the section the overall intensity mapping φ_{1D} is defined as

$$\varphi_{1D}(\chi) = \chi - \phi_{1D}(\chi) \quad (4.4)$$

with $\chi \in \Gamma$ being an arbitrary intensity within v . The function $\phi_{1D} : \Gamma \mapsto \Gamma$ describes the deformation of the intensities also denoted as intensity shift. As mentioned before, the deformation is split into the affine ϕ_A and the non-rigid ϕ_D part:

$$\phi_{1D}(\chi) = \phi_A \circ \phi_D(\chi). \quad (4.5)$$

4.3.1 Affine Alignment of 1-D Histograms

The affine transformation model for the alignment of the reference with the target histogram can be written as

$$\tilde{\xi} = \tilde{s}_A \cdot \chi + \tilde{t}_A \quad (4.6)$$

where $\tilde{\xi}$ is the transformed intensity. The only unknown parameters of the model are the scaling $\tilde{s}_A \in \mathbb{R}$ and the translation $\tilde{t}_A \in \mathbb{R}$ of the intensities. Given \tilde{s}_A and \tilde{t}_A the unknown parameters s_A and t_A of the deformation field

$$\phi_A(\chi) = \chi - s_A \cdot \chi + t_A \quad (4.7)$$

can be computed. For MR images the translation part can be neglected in most cases. The resulting optimization problem is easily solvable. Even brute force methods are applicable here because the problem dimension is low.

In order to measure the difference between the two histograms, we either use the Sum of Squared Differences (SSD) or the Jeffrey Divergence (JD). SSD is defined as

$$\mathcal{D}_{\text{SSD}} = \int_{\Gamma} D_{\text{SSD}}(\chi, p_u, p_v, \phi) d\chi = \frac{1}{2} \int_{\Gamma} (p_v \circ \varphi(\chi) - p_u(\chi))^2 d\chi \quad (4.8)$$

where φ is the intensity mapping as defined in Eq. (4.1). JD is defined as

$$\mathcal{D}_{\text{JD}} = \int_{\Gamma} D_{\text{JD}}(\chi, p_u, p_v, \phi) d\chi = \int_{\Gamma} p_u(\chi) \log \frac{p_u(\chi)}{\mu(\chi)} + p_v \circ \varphi(\chi) \log \frac{p_v \circ \varphi(\chi)}{\mu(\chi)} d\chi \quad (4.9)$$

with

$$\mu(\chi) = \frac{p_u(\chi) + p_v \circ \varphi(\chi)}{2}. \quad (4.10)$$

Theoretically there are many measures applicable and known from literature, most notably the Kullback Leibler Divergence (KL-Divergence). In practice, however, these two measures perform best. Although JD is very similar to the symmetric KL Divergence, in our experiments JD showed to be more robust against the number of histogram bins, bin size, initial alignment of the histograms etc.

Consequently, the optimization problem can be written as

$$\varphi_{1D}^* = \arg \min_{\varphi} \mathcal{D}_{\{\text{SSD}, \text{JD}\}} \quad (4.11)$$

where $\phi_D(\chi)$ is set to the identity.

4.3.2 Non-Rigid Alignment of 1-D Histograms

After the affine alignment a non-rigid matching using a variational registration approach is utilized to approximate the non-parametric deformation ϕ_D . The goal of the proposed intensity normalization is to maximize the similarity of the transformed target pdf $p_v \circ \varphi_{1D}$ to the reference with respect to the intensity shift ϕ_D and a given similarity measure \mathcal{D} . Like in the affine alignment stage, we utilize either SSD, compare Eq. (4.8), or JD, compare Eq. (4.9), to measure the similarity between the transformed target pdf and the reference. The affine intensity shift ϕ_A is assumed to be known.

Relying on the similarity measures only yields an ill-posed optimization problem. For this reason, we introduce an additional term controlling the smoothness of the solution. As we want neighboring intensities to be mapped as similar as possible we chose a regularizer that is based on the first derivative of the mapping ϕ_D . In literature this is called diffusion type regularization and can be formulated as

$$\mathcal{S} = \frac{\alpha}{2} (\phi'_D(\chi))^2 \quad (4.12)$$

where α controls the influence of the regularization term \mathcal{S} on the solution of the optimization problem and $\phi'_D(\chi)$ is the derivative of the intensity dependent mapping. However, other regularization approaches, like curvature or elastic regularization [Mode04], can be applied as well.

Besides stabilizing the optimization process, the regularizer \mathcal{S} introduces the possibility to integrate prior knowledge about the problem domain into the solution of the intensity standardization. This is particularly advantageous if the target and source images are affected by different kinds of image artifacts such as noise or defect pixel. Further on, in medical applications, like the analysis of brain tumors in MRI, the source images do not show the same pathologic structures than the target image in general. Moreover, the structures can have different sizes in the data sets. These facts affect the image's pdf and with that the normalization of the target image. Here the regularization prevents a mismatch of the intensities of these structures.

Using the presented similarity measures as well as the regularization term, the standardization can be formulated as the variational problem of first order

$$\mathcal{I} = \int_{\Gamma} D(\chi, p_u, p_v, \phi_D) + \frac{\alpha}{2} (\phi'_D(\chi))^2 d\chi \quad (4.13)$$

with \mathcal{D} representing either SSD or the JD. The goal of the optimization is to find the intensity deformation ϕ_D . This corresponds to the functional

$$\mathcal{I} = \int_{\Gamma} \mathcal{L}(\chi, \phi_D, \phi'_D) d\chi \quad (4.14)$$

with $\mathcal{L}(\chi, \phi_D, \phi'_D) = D(\chi, p_u, p_v, \phi) + \frac{\alpha}{2} (\phi'_D(\chi))^2$. As we want to find its minimum, the Euler differential equation

$$\frac{\partial}{\partial \phi_D} \mathcal{L}(\chi, \phi_D, \phi'_D) - \frac{d}{d\chi} \frac{\partial}{\partial \phi'_D} \mathcal{L}(\chi, \phi_D, \phi'_D) = 0 \quad (4.15)$$

has to be fulfilled. This is equivalent to setting the directional derivative of \mathcal{I} to zero. In practice, this yields the minimization problem

$$\phi^* = \arg \min_{\phi} \left(\frac{\partial}{\partial \phi_D} \mathcal{L}(\chi, \phi_D, \phi'_D) - \frac{d}{d\chi} \frac{\partial}{\partial \phi'_D} \mathcal{L}(\chi, \phi_D, \phi'_D) \right). \quad (4.16)$$

For all used similarity measures \mathcal{D} the second part $\frac{d}{d\chi} \frac{\partial}{\partial \phi'_D} \mathcal{L}(\chi, \phi, \phi')$ is zero. For the regularization the first part $\frac{\partial}{\partial \phi_D} \mathcal{L}(\chi, \phi_D, \phi'_D)$ becomes zero. Thus, using SSD Eq. 4.15 yields

$$\left(p_v \circ \varphi(\chi) - p_u(\chi) \right) (p_v \circ \varphi)'(\chi) - \alpha \phi''_D(\chi) = 0, \quad (4.17)$$

and employing JD

$$(p_v \circ \varphi)'(\chi) \left(\frac{1}{2} + \log \frac{\mu(\chi)}{p_v \circ \varphi(\chi)} \right) - \alpha \phi''(\chi) = 0. \quad (4.18)$$

The differentiation of JD is described in detail in appendix B.

In order to find a solution to the presented differential equations, we introduce an additional artificial time component τ . The partial derivative regarding the time component τ is formulated as

$$\frac{\partial}{\partial \tau} \phi_D(\chi, \tau) = \frac{\partial}{\partial \phi_D} \mathcal{L}(\chi, \phi_D, \phi'_D) - \frac{d}{d\chi} \frac{\partial}{\partial \phi'_D} \mathcal{L}(\chi, \phi_D, \phi'_D). \quad (4.19)$$

The solution for the non-rigid matching corresponds now to the steady state solution $\frac{\partial}{\partial \tau} \phi_D(\chi, \tau) = 0$ of the time-dependent PDE [Mode04, Weic98]. Eq. (4.19) can be reformulated as

$$\frac{\partial}{\partial \tau} \phi_D(\chi, \tau) = \mathcal{A}[\phi_D](\chi) - f(\chi, \phi_D(\chi, \tau)) \quad (4.20)$$

with $\mathcal{A}[\phi_D(\chi, \tau)](\chi) = \frac{\partial}{\partial \phi_D} \mathcal{L}(\chi, \phi_D, \phi'_D)$ being a partial differential operator corresponding to the regularization term and $f(\chi, \phi_D(\chi, \tau)) = \frac{d}{d\chi} \frac{\partial}{\partial \phi'_D} \mathcal{L}(\chi, \phi_D, \phi'_D)$ being a force term corresponding to the distance measure. By setting $\phi_D^{(i)}(\chi) = \phi_D(\chi, i \cdot \Delta\tau)$ and using $\Delta\tau$ as discrete time step a semi-implicit scheme

$$\frac{\phi_D^{(i+1)}(\chi) - \phi_D^{(i)}(\chi)}{\Delta\tau} = \mathcal{A}[\phi_D^{(i+1)}](\chi) - f(\chi, \phi_D^{(i)}(\chi)) \quad (4.21)$$

for the solution of the non-rigid alignment problem can be derived from Eq. (4.20) [Mode04].

To solve the described optimization problem, Eq. (4.21) has to be discretized. The discretization is done by sampling the pdfs on an equidistant grid. The spacing between two grid points is denoted by $\Delta\chi$. The discrete version of p_u is denoted by \mathcal{U} and the discretization of p_v is called \mathcal{V} . Moreover, for the approximation of the continuous pdfs of the images by discrete histograms \mathcal{U} and \mathcal{V} a partial volume based interpolation scheme is used. This means, that during the computation of the histograms, observed intensities are added to the two neighboring histogram bins weighted by their relative distance to the borders of the bins. Furthermore

a coarse resolution only the remaining small deformations have to be treated at the computational more expensive higher resolution. Let the number of levels be called M . Usually, $M = 2$ is enough. Let further the downsampled variables be marked by a subscript denoting the resolution of the discretization. For instance $\Phi_{D,m}^{(i)}$ corresponds to the deformation field in level m .

The proposed intensity standardization is steered by the parameters $\Delta\tau$, α , the maximal number of iterations i_{\max} and the number of used levels M . Without the inversion of the matrix $(\mathbf{I} - \Delta\tau\mathbf{A})$ that has to be performed only once, the complexity of the method corresponds to the complexity of a matrix vector multiplication. Consequently, it is $\mathcal{O}(n_h^2)$ dependent on the number of histogram bins n_h . The overall approach is summarized in Algorithm 1.

Algorithm 1 Algorithm for the alignment of 1-d histograms

Require: v , u , α , $\Delta\tau$

Ensure: Standardized image intensities of image v

- 1: compute the histograms \mathcal{V} and \mathcal{U}
 - 2: compute the affine mapping ϕ_A by minimizing Eq. (4.11)
 - 3: apply ϕ_A to \mathcal{V}
 - 4: initialize $\Phi_{D,0}^{(i)} = 0 \forall \chi \in \Gamma; i = 1$
 - 5: compute the downsampled histograms \mathcal{U}_0 and $\mathcal{V}_0^{(0)}$
 - 6: **for** $m = 0$ to M **do**
 - 7: **repeat**
 - 8: compute $\Phi_{D,m}^{(i)}$ using Eq. (4.24)
 - 9: compute $\mathcal{V}_m^{(i)}$ using $\Phi_{D,m}^{(i)}$
 - 10: $i = i + 1$
 - 11: **until** $\|\Delta\Phi\| < \varepsilon \wedge i < i_{\max}$
 - 12: upsample the transformation $\Phi_{D,m}^{(i)}$
 - 13: **end for**
 - 14: **for** all voxels **do**
 - 15: transform the intensities of v to obtain v_ϕ using the estimated deformation Φ_D
 - 16: **end for**
 - 17: **return** v_ϕ
-

In a way the proposed method can be seen as a continuous version of dynamic programming (DP). In both methods an ideal mapping is computed. However, this mapping is discrete in the DP case whereas for the 1-D histogram matching a real valued mapping is computed. Furthermore, in the proposed algorithm deformation constraints are implicitly introduced into the computation using a regularization term. In the case of DP these constraints can be explicitly modeled.

4.4 Non-Rigid Alignment of n-D Histograms

The goal of the non-rigid alignment of n-D histograms is to find a mapping between the intensities of a set of images $\mathbf{v} = (v_1, v_2, \dots, v_n)$, where n is the number of images acquired with different modalities and a reference set of images $\mathbf{u} = (u_1, u_2, \dots, u_n)$

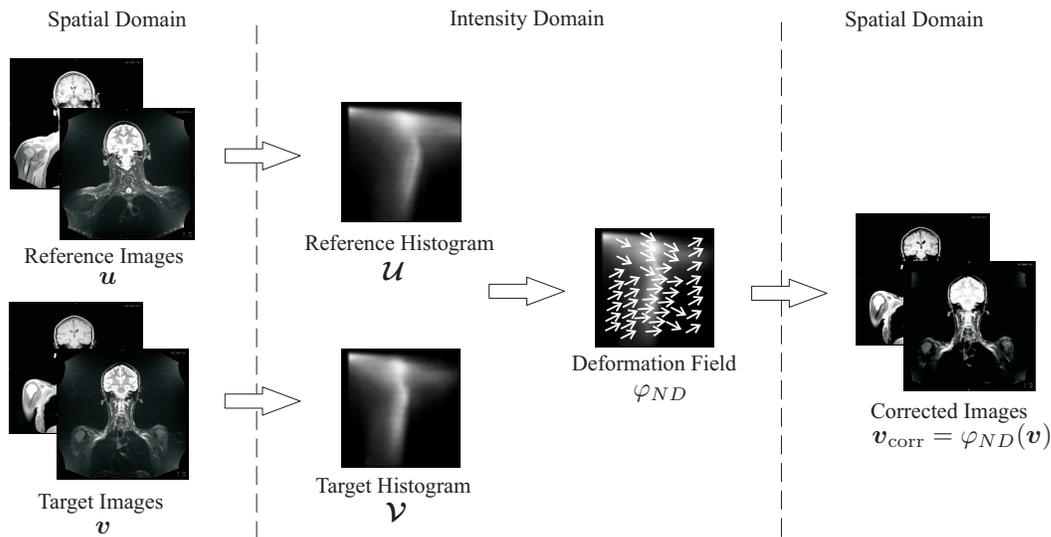


Figure 4.1: Schematic illustration of the intensity standardization. First, from the reference images a reference joint histogram is created. This is the training component of the approach. Then from the current MRI images a joint histogram is generated. In the next step these histograms are non-rigidly registered. Using the gained transformation function, the current images are standardized. For visualization purposes, the proposed method is shown for $n = 2$, but it is applicable to any n .

so that an arbitrary intensity vector $\chi \in \Gamma^n$ describes the same tissue class in both sets with $\Gamma^n \subseteq \mathbb{R}^n$ being the intensity space of the image sets. The method, we present in this section, was first introduced at the Conference “Vision, Modeling and Visualization 2006” [Jage06a] in the context of the segmentation of Multiple Sclerosis in MRI head images. Later on, it was refined and published in [Jage09a]. The main idea of our method is that this mapping can be approximated by the minimization of the distance between the joint histograms of the two sets of images. The required joint histograms are of dimensionality n , corresponding to the number of images. The domain of the histograms is Γ^n . Note that, at least for real data sets, no plausible transformation of the relative joint histograms can be found, such that, the difference is zero, because the volume of tissue classes in the image sets \mathbf{v} and \mathbf{u} differ for inter- as well as intra-patient measurements (e.g., anatomical differences, partial volume averaging effects, positioning of the patient). Thus, the search for a mapping between the intensity spaces is equivalent to finding the transformation φ_{ND} between the corresponding joint histograms which minimizes a given distance measure \mathcal{D} :

$$\tilde{\varphi}_{ND} = \arg \min_{\varphi_{ND}} \mathcal{D}(\mathbf{u}, \mathbf{v}; \varphi_{ND}), \quad (4.25)$$

with \mathbf{u} and \mathbf{v} being the joint histograms of the image sets \mathbf{u} and \mathbf{v} . If the joint histograms are treated as images, this task can be viewed as a non-rigid image registration problem. Note that, although in theory histograms of arbitrary dimensionality can be used, in practice n should be small enough to be able to reliably compute the required densities. Otherwise, the registration results may be no longer satisfactory.

Techniques like Parzen estimation can solve the problem of insufficient samples; however, this leads to high computational costs.

For image registration a variety of algorithms are available. A survey about image registration is given in Maintz et al. [Main 98] and Hill et al. [Hill 01]. We employed the variational non-rigid registration approach which was first introduced by Modersitzki et al. [Mode04]; however, other deformable registrations schemes are applicable like the spline based non-rigid registration presented in [Ruec99]. The result of this method’s optimization is the transformation $\varphi_{ND} : \mathbb{R}^n \mapsto \mathbb{R}^n$. In the context of the registration of multi-dimensional joint histograms, it describes how to transform the gray values of one set of images \mathbf{v} such that its intensity distribution best matches the reference distribution, with respect to the used distance measure and regularizer. The objective functional \mathcal{J} of the non-rigid registration can be written as

$$\mathcal{J}[\mathbf{u}, \mathbf{v}; \phi_{ND}] = \mathcal{D}[\mathbf{u}, \mathbf{v}; \phi_{ND}] + \alpha \mathcal{S}[\phi_{ND}], \quad (4.26)$$

with \mathcal{D} being the distance measure (e.g., Sum of Squared Differences or Mutual Information), \mathcal{S} being a regularizer (e.g., Curvature or Elastic regularization), α defining the influence of the regularizer on the optimization and ϕ_{ND} representing the deformation between the joint histograms. The applied non-rigid image registration using a variational framework is very similar to the non-rigid alignment of 1-d histograms in section 4.3.2. For further in depth information about variational registration theory, we want to refer to the literature [Mode04, Jage06b]. The intensity standardization can be done by

$$\boldsymbol{\chi}_{corr} = \boldsymbol{\chi}_{orig} + \phi_{ND}(\boldsymbol{\chi}_{orig}) = \tilde{\varphi}_{ND}(\boldsymbol{\chi}_{orig}), \quad (4.27)$$

where $\boldsymbol{\chi}_{orig} \in \Gamma^n$ describes the intensity vector in the original target image set \mathbf{v} and $\boldsymbol{\chi}_{corr} \in \Gamma^n$ is the intensity vector in the corrected images respectively. A schematic overview of the standardization process is given in Figure 4.1. Here the relationship between the spatial and the intensity domain is illustrated.

If the computed transformation is applied to the joint pdf p of the target image set, it is not guaranteed that the resulting function is still a pdf, as the constraint $\sum_{\boldsymbol{\chi}} (p \circ \tilde{\varphi}_{ND})(\boldsymbol{\chi}) = 1$ might be invalid. However, as the derived mapping is applied to images, the resulting pdfs of the corrected images will fulfill the constraint again. Nevertheless, volume preserving non-rigid registration approaches can be used as well [Habe04].

In a pre-processing step, the joint histograms were equalized in order to increase the performance of the registration in regions with small tissue support. This is very important for data sets acquired with protocols that highlight small structures (e.g. blood vessels or kidneys in TIRM images). Without the equalization step, areas in the joint histograms representing such structures are not treated satisfactorily in the registration process, as small histogram values hardly influence the distance measure. Thus, the registration concentrates on structures in the histograms with high tissue support.

For data sets being “statistically simple”, e.g. consisting of only a few dominating classes like the head region, the proposed method returns satisfactory results (see results section). However, the following problems may arise in more complex data sets: a) tissue classes with a small number of voxels do not have enough support to be transformed in a reliable manner; b) if a previous bias field correction step has

failed, the histograms are blurred and the statistical information of a tissue class is spread to a broad range of gray values. Consequently, it is no longer possible to find a plausible global transformation of the intensity vectors. One straightforward solution to this is to split the data sets into smaller sub-volumes. These sub-volumes can then be intensity-standardized separately. However, this can still lead to problems if the statistical content of a sub-volume is not sufficient for reliable registration. In order to have sufficient statistical content, a partition should have the same dominating tissue classes as the corresponding partition in the reference image. Furthermore, the histogram has to have a similar morphology as the histograms of the neighboring partitions. If the sufficient content assumption does not hold, some tissue classes might be transformed in a false way. Furthermore, due to the independent standardization of the sub-volumes, intensity discontinuities can occur at the junctions between the partitions.

4.5 Standardization of Whole-Body MRI scans

In order to overcome these drawbacks we proposed parts of the following sub-volume based technique at the Conference “Bildverarbeitung für die Medizin 2007” [Jage07] and refined it in an article in IEEE Transactions on Medical Imaging [Jage09a]. The core of our method is the new distance measure for the registration of the joint histograms which utilizes the statistical information of neighboring partitions in the joint histograms in regularizing the computed transformation of the intensities. The modified method can be summarized as follows: The data is split into K partitions. For each partition $k \in \{1, \dots, K\}$ a joint histogram for the target and one for the reference volume is created. The partitioning of the input data sets is arbitrary. We use a linear partitioning along the z -axis; however, the sub-volumes can also be identified using other strategies (e.g., segmentation). Each partition is independently registered, but all the remaining $K - 1$ partitions are utilized for regularization of the registration. Consequently, the new distance measure can be described as a weighted linear combination of the simple distance measures of all K histograms. Thus, the deformation computed for partition j depends on all histogram partitions, as it utilizes their properties as well. This leads to the following distance measure:

$$\mathcal{D}^j[\mathbf{u}, \mathbf{v}; \phi_{ND}^j] = \sum_{k=1}^K a_{j,k} \mathcal{D}[\mathbf{u}^k, \mathbf{v}^k; \phi_{ND}^j] \quad (4.28)$$

with

$$\sum_{k=1}^K a_{j,k} = 1, \quad (4.29)$$

where \mathbf{u}^k and \mathbf{v}^k are the reference histogram of sub-volume k and the target histogram of sub-volume k respectively, ϕ_{ND}^j is the deformation field of partition j corresponding to the transformation φ_{ND}^j and \mathcal{D} is a simple distance measure as mentioned above. Furthermore, j is the current sub-volume and $a_{j,k}$ is the influence of the force of partition k in the context of the standardization of sub-volume j .

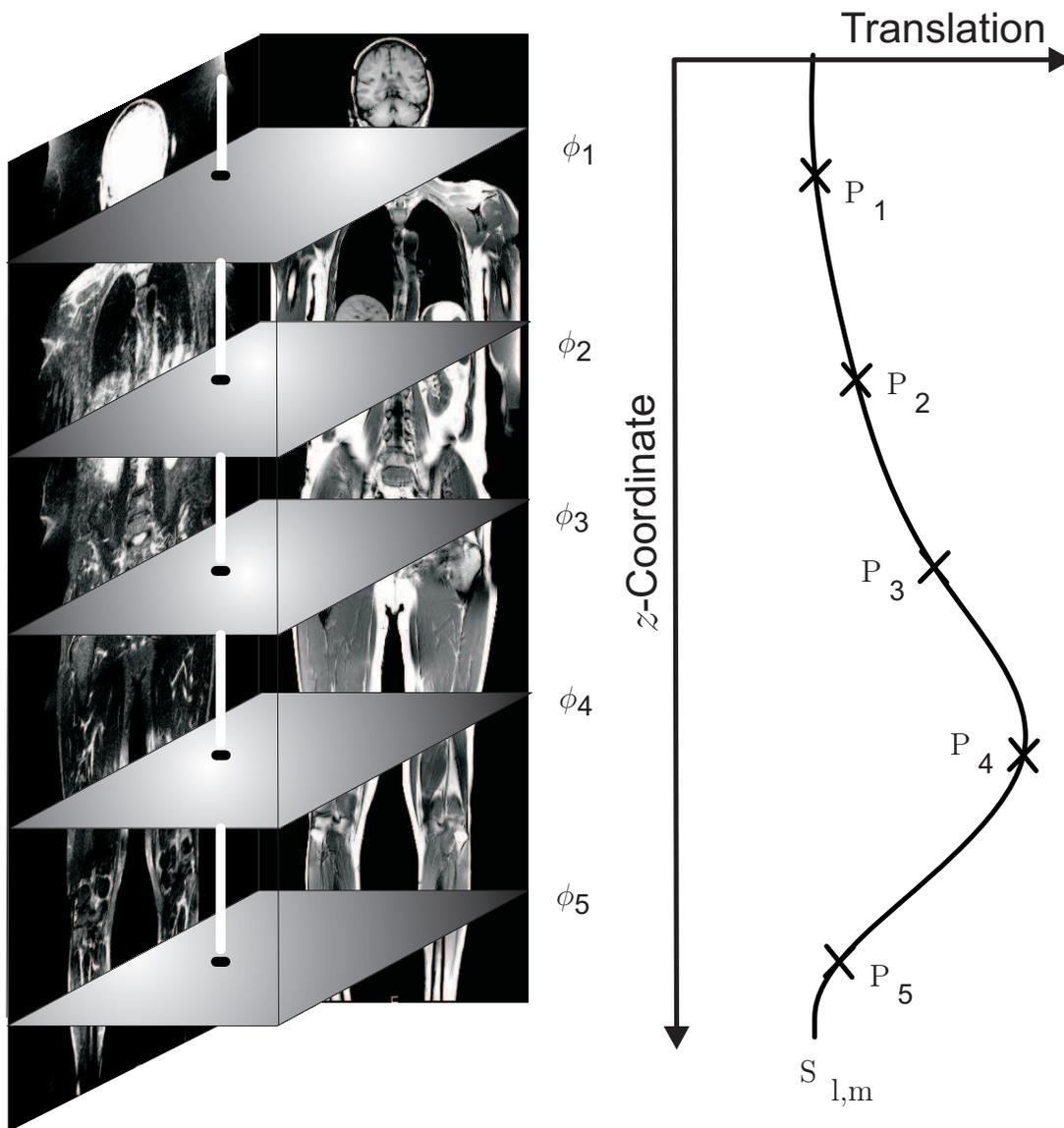


Figure 4.2: Schematic illustration of the B-spline interpolation of the transformation vectors. The example uses $K = 5$ partitions. The planes ϕ_1 - ϕ_5 represent the magnitude of the deformation in a specific sub-volume. The plot shows the relationship between the control points P_1 - P_5 of the spline $S_{l,m}$, the z-coordinate within the volumes and the magnitude of the translation.

If all $a_{j,k}$ are set to $a_{j,k} = \frac{1}{K}$ the resulting deformation field of all partitions is the same. Setting $a_{j,j} = 1$ for all j and $a_{j,k} = 0, \forall j \neq k$ results in a unconstrained non-rigid registration of the joint histograms of all K blocks. The proposed approach can still lead to discontinuities at the junctions of the partitions. However, these can be significantly reduced, depending on the chosen weighting factors $a_{j,k}$. Furthermore, these discontinuities can be further reduced by interpolating the computed transformations φ_{ND}^j along the spatial positions of the corresponding intensities. Depending on the partitioning of the volume into blocks, different interpolation schemes have to be applied. As we decompose our volume along the z -axis only, we apply a cubic B-spline interpolation as follows. The sampling of the intensity spaces of the n acquired images is denoted by l_1, \dots, l_n ; thus, the joint histograms have $L = l_1 \times \dots \times l_n$ bins. Consequently, the transformation φ_{ND}^j corresponding to the deformation field ϕ_{ND}^j , is sampled at L positions. Thus, $n \cdot L$ B-splines have to be computed; one for each bin and for each dimension of the deformation field. Correcting an intensity χ_m with z -coordinate z in an image v_m given an intensity vector $\boldsymbol{\chi} = (\chi_1, \dots, \chi_m, \dots, \chi_n)^T$ results in

$$\chi_{m,\text{corr}} = \chi_m + S_{l,m}(z), \quad (4.30)$$

where $m \in \{1, \dots, n\}$ and $l \in \{1, \dots, L\}$ corresponds to the intensity vector $\boldsymbol{\chi}$. The relationship between the control points of the splines $S_{l,m}$, the z -coordinate of the volumes and the signal intensity transformation is illustrated in Figure 4.2. The approach can be summarized as shown in Algorithm 2.

Algorithm 2 Algorithm for the alignment of n -d histograms

Require: $\mathcal{U}^k, a_{j,k} \quad \forall k, j$

Ensure: Standard intensity scale

for $k = 1$ to K **do**

 compute current histogram \mathcal{V}^k

 equalize \mathcal{V}^k

end for

for $j = 1$ to K **do**

 compute ϕ^j ; the previously computed histograms $\mathcal{U}^k, \mathcal{V}^k$ and $a_{j,k}$ are required

end for

for $l = 1$ to L **do**

for $m = 1$ to n **do**

 compute B-spline $S_{l,m}$ using ϕ^j

end for

end for

for all voxels **do**

 transform the intensities using the interpolated transformations; $S_{l,m}$ and the z -coordinate are required.

end for

A further advantage of the proposed intensity standardization method arises from the use of joint histograms. As only pdfs are utilized, the standardization results do

not become worse if the region of interest in the two volumes differs slightly. For this reason just a coarse rigid registration of the reference with the current image data sets is required in the inter-patient case to generate the K sub-volumes for the proposed approach. If the whole statistical information is used at once ($K = 1$) then no alignment of the volume sets \mathbf{u} and \mathbf{v} has to be done. As a result of the flexibility concerning the region of interest, the method can deal with anatomical differences in inter-patient scans or anatomical changes due to evolving lesions, for instance. However, if the source and target images are significantly different anatomically, or if there is a big variance in the localization of the regions of interest, the registration may fail. However, a possible solution for increasing the quality of the standardization results is to introduce additional reference histograms that cover the possible anatomical differences. Hence, before the signal intensity standardization, the proper set of reference histograms has to be chosen. This can be achieved, for instance, by selecting the histograms with a minimal distance to the current histograms. Furthermore, we assume that the images are roughly aligned. If this cannot be guaranteed a rigid pre-registration has to be performed beforehand. If this is not done then the joint histograms will not fit and hence the quality of the standardization will be worse.

4.6 Experiments & Results

In this section, several experiments for the evaluation of the proposed histogram alignment methods are presented. Each experiment is split into a section describing the used data sets, a section introducing the evaluation strategy, and finally a results section.

4.6.1 1-d Histogram Alignment

In order to evaluate the proposed 1-d histogram matching intensity standardization method, various experiments using synthetic images as well as clinically relevant data sets are performed. First, we present results of experiments using artificially altered images to demonstrate the basic properties of the developed approach. Then, clinically relevant image data of different body regions, pulse sequences as well as scanner brands are used for the evaluation. Whenever possible, gold standard information is utilized to measure the standardization performance. In some cases, however, only a qualitative evaluation is performed. The experiments were done on a 2.00 GHz Intel Core2 CPU with 2 GB RAM. The intensity standardization methods are implemented in C++ and integrated in the ITK framework¹. As a global transformation is computed, the signal intensity standardization is independent of the dimensionality of the reference and target images. Using 128 bins, two levels, and 100000 iterations, it took approximately one second to perform the standardization including the affine pre-alignment.

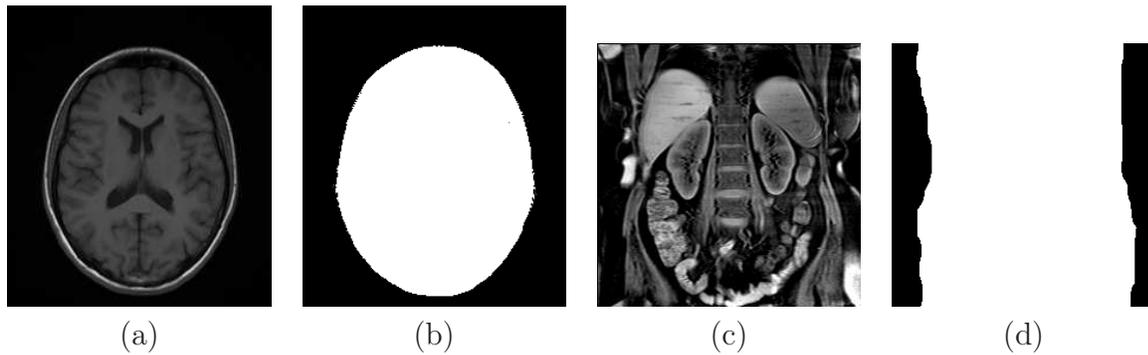


Figure 4.3: Images used for the synthetic 1-d standardization experiments: (a) T1w slice through a 3-d volume of the head. (b) The corresponding mask to exclude the background from the calculations. (c) T2w weighted slice showing the kidneys. (d) Mask image corresponding to image (c).

Artificial Perturbation of Real Images

Data sets: In this experiment, we alter images by introducing an artificial perturbation of its intensities. Throughout the experiment we are using 2-d images, however, all methods are independent of the data sets' dimensionality. All results are generated using two different kinds of images: a slice through a T1w head data set and a T1w image of the kidneys. The image showing the kidneys was bias corrected using DaC in a subsequent step. For the computation of the intensity mappings a mask was used to restrict the calculations to foreground pixels. The masks were generated using a semi-automatic segmentation scheme. The images as well as their masks are shown in Fig. 4.3.

Evaluation Method: The considered methods are evaluated in three different settings using synthetic perturbations of the images' intensities.

First, we apply *Pepper Noise* to the images. The percentage of corrupted pixels ranges between $p_c = 0$ and $p_c = 0.5$. As the histograms are computed using a NP window based technique, the noise cannot be simply added to the histogram directly but has to be introduced into the images. The corrupted images \tilde{u}_p are computed as

$$\tilde{u}_p(\mathbf{x}) = \begin{cases} 0 & \text{if } \text{rand}() < p_c \\ u(\mathbf{x}) & \text{otherwise.} \end{cases} \quad (4.31)$$

where $\text{rand}()$ is a random numbers generator that produces random numbers with a uniform distribution in the range of $[0, 1]$. Thus, the noise depends on the position within the image what is reflected in the histograms.

Second, we alter the images using additive *Gaussian Noise*. Due to the spatial dependence of the NP window histogram approximation, in this case the intensities are distorted within the images, too. The images \tilde{u}_g including the Gaussian noise can be computed as

$$\tilde{u}_g(\mathbf{x}) = u(\mathbf{x}) + \text{g}_{\text{rand}}(\sigma) \quad (4.32)$$

¹<http://www.itk.org>

with $g_{\text{rand}}(\sigma)$ being a random number generator that produces random numbers with a Gaussian distribution with standard deviation σ and zero mean. We use a random number generator that is based on the Polar method. The standard deviation σ varies between zero and 50 intensity units.

Finally, we evaluate the considered methods by introducing an *additional artificial tissue class* to the images. The intensity distribution of the artificial class is assumed to be Gaussian with a standard deviation of $\sigma = 10$ gray values. As the tissue class is locally bounded within the images, it is sufficient to alter the histogram of the original images. The new histogram can be written as

$$\tilde{\mathcal{U}}_a(i) = \frac{1}{1 + \alpha_g}(\mathcal{U}_a(i) + \alpha_g G_\sigma(i - i_c)) \quad (4.33)$$

with G_σ being a Gaussian density with standard deviation σ . The variable i_c defines the mean intensity of the artificial tissue class and the scalar factor α_g its size. In the experiments we use three different center points i_c at the intensities 20, 120 and 220. The histograms include the intensity range $\Gamma = \{0, \dots, 255\}$. The size factor α_g varies between 0 and 0.5; here, 0.5 means that the size of the additional class is 50% of the size of the initial foreground pixel. Thus, a class size between 0% and 50% of the initial number of pixel within the images is used.

In all three scenarios considering an artificial perturbation of the images' intensities, we evaluate histogram specification (HS), a simple affine alignment (AR) of the data sets, nonrigid 1-d histogram matching only (1D), and the proposed combined 1-d histogram matching (AR1D). The latter one includes the affine histogram alignment as a pre-processing step. In order to test the proposed method in a relevant scenario, we utilize the same parameters as the ones used for the standardization of 3-d clinical data sets (Tumorbases, see next experiment section). The intensity transformation is computed using the distorted image \tilde{u} as reference and the undistorted image u as target. The transformed image u^* is then compared to the undistorted image. As quality measure, we utilize the mean square error (MSE) as well as the root mean square error (RMSE) between both images. If the methods are robust against the perturbations, the resulting measure has to be equal to zero.

Results: In Figures 4.4 (a) and 4.4 (b) the results of the standardization of the T1w head image as well as the image of the kidneys perturbed with *pepper noise* are shown. Due to the one-to-one mapping of the histogram specification the MSE and with it the RMSE induced increases, as expected, linearly dependent on the probability of the pepper defect. With a pepper probability of 50% the MSE using histogram specification is $\text{MSE}_{\text{HS}} = 21.39$ for the head data set and $\text{MSE}_{\text{HS}} = 40.54$ for the kidneys respectively. This means that there is a mean intensity deviation of about $\text{RMSE}_{\text{HS}} = 4.63$ in the head case and $\text{RMSE}_{\text{HS}} = 6.37$ for the kidney image. Using an affine registration of the histograms only, it can be seen in the plots that the standardization error increases with the probability of a "defect" pixel. The reason for this is the way of doing the histogram computation. Due to the NP windows technique the histograms are blurred because of the spatial distribution of the pepper noise. This fact confuses the search for the optimal scaling parameter. However, due to the reduced flexibility the error is still less than the error introduced by histogram specification. For the head image the MSE is $\text{MSE}_{\text{AR}} = 3.53$ and $\text{MSE}_{\text{AR}} =$

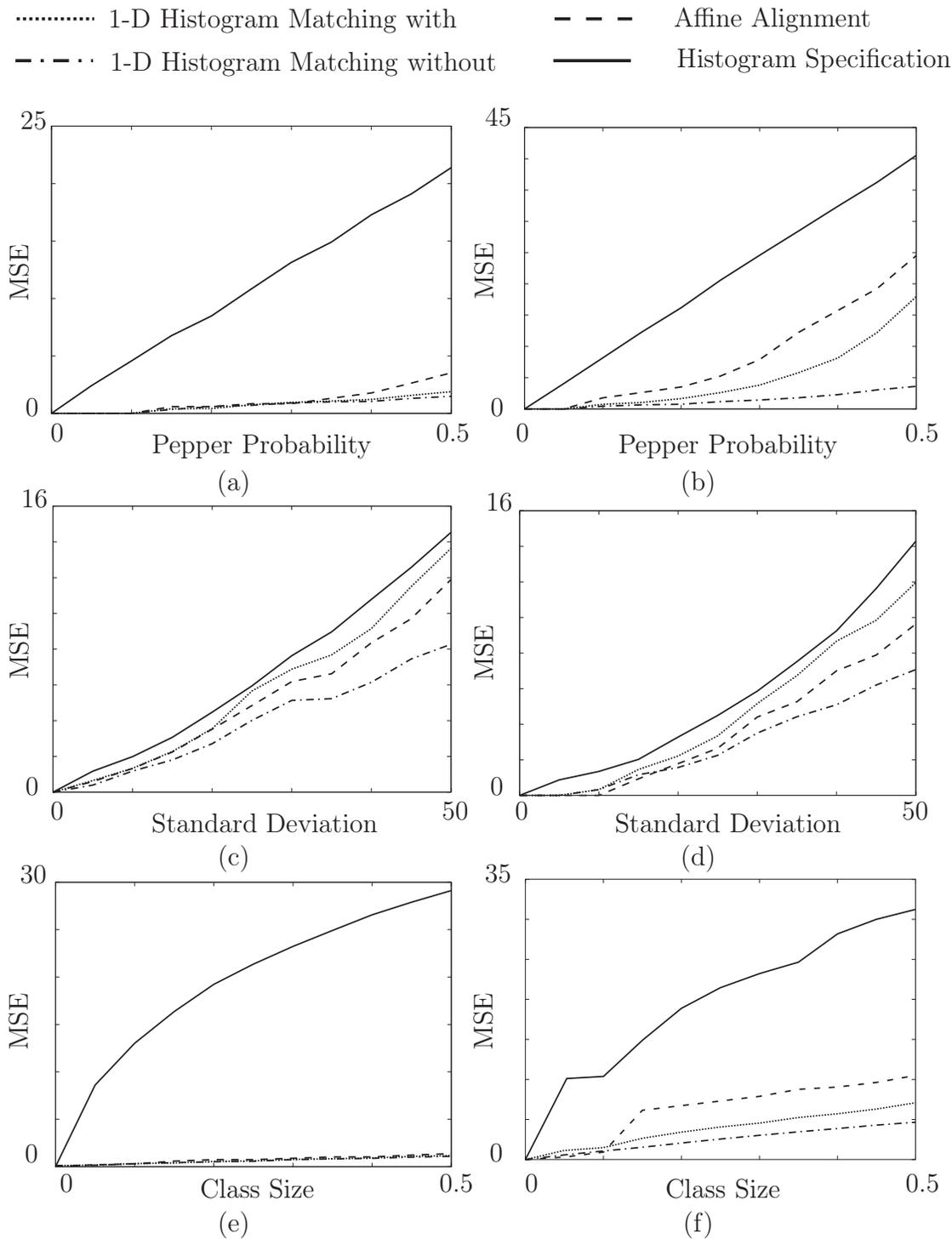


Figure 4.4: Results of the synthetic 1-D standardization experiments: In all experiments, we compare histogram specification, an affine alignment, 1-d histogram matching *without* and *with* a previous affine alignment step. The left column ((a), (c), and (e)) shows the result using the head images, the right column ((b), (d), and (f)) the results using the kidney image respectively. The plots in (a) and (b) correspond to the pepper noise setting, (c) and (d) correspond to the Gaussian noise experiment and (e) and (f) illustrate the results of the additional tissue class experiment. The shown MSE corresponds to the MSE from the ground truth data.

24.50 in the kidney case. This corresponds to an induced intensity deviation of about $\text{RMSE}_{\text{AR}} = 1.88$ in the head case and $\text{RMSE}_{\text{AR}} = 4.95$ for the kidney image. Applying the 1-D histogram matching without subsequent affine registration shows that there is hardly any influence of the pepper noise onto the standardization result. The reason for this is that the structural differences of the histograms are too small. Here, the MSE is $\text{MSE}_{\text{1D}} = 1.49$ ($\text{RMSE}_{\text{1D}} = 1.22$) for the head image and $\text{MSE}_{\text{1D}} = 3.61$ ($\text{RMSE}_{\text{1D}} = 1.90$) in the kidney case. However, if the affine registration is done before the 1-D non-rigid registration of the histograms as an initialization step the error gets higher compared to a single 1-D non-rigid registration. The MSE becomes $\text{MSE}_{\text{AR1D}} = 1.89$ for the head images and $\text{MSE}_{\text{AR1D}} = 17.96$ for the kidneys that corresponds to a mean intensity deviation of $\text{RMSE}_{\text{AR1D}} = 1.37$ and $\text{RMSE}_{\text{AR1D}} = 4.24$ at a probability of 50% for pepper noise. The reason for the increasing error is that after the affine registration the histograms do not match with sufficient accuracy in many areas. The non-rigid registration tries to revert the affine alignment but the optimization gets stuck in a local minima. Nevertheless, the error is still smaller than the MSE induced by histogram specification and a pure affine alignment.

The results of the experiments testing the influence of *additive Gaussian noise* are shown in the Figures 4.4 (c) and 4.4 (d). With an increasing standard deviation of the noise the MSE introduced by histogram specification increases as well. Given a standard deviation of $\sigma = 50$ intensity units, the introduced MSE is $\text{MSE}_{\text{HS}} = 14.54$ ($\text{RMSE}_{\text{HS}} = 3.81$) for the head images and $\text{MSE}_{\text{HS}} = 14.29$ ($\text{RMSE}_{\text{HS}} = 3.78$) in the case of the kidney data set. The reason for this behavior is that the Gaussian noise blurs the image histogram. Thus, in order to fit the reference (in this case the histogram including the artificial noise) the target histogram has to be spread in some areas and has to be clinched in others yielding an increased MSE. This effect also affects the standardization results of the affine registration, the single 1-D histogram matching and the combined approach, too. Compared to the pepper noise scenario, however, even the 1-D histogram matching without a prior affine registration yields a significant MSE. The proposed combined registration approach also performs much worse than in the pepper noise experiment. The reason for this is exactly the same as for histogram specification. The locality of the correction tries to minimize the distance to the blurred version of the histogram. Thus, the standardization induces an error dependent on the standard deviation of the noise.

To evaluate the experiment that introduces an *additional artificial tissue class* into the data, we compute the MSE dependent on the relative size of the artificial tissue class. Moreover, the MSE is computed for three different mean intensities of the artificial tissue class. Then, the three resulting curves are averaged. This is done for all four methods for the head as well as the kidney data set. The resulting plots are shown in the Figures 4.4 (e) and 4.4 (f). As expected for the standardization using histogram specification the MSE increases with the relative size of the artificial tissue class. The error results from the mapping of multiple pixel intensities to the additional class. This also means that if the value of the class is further away from the mean value of the foreground pixel the error becomes larger. At a level of 50% the MSE is in average $\text{MSE}_{\text{HS}} = 29.13$ ($\text{RMSE}_{\text{HS}} = 5.40$) for the head image. In the case of the kidney data set the average MSE is $\text{MSE}_{\text{HS}} = 31.24$ ($\text{RMSE}_{\text{HS}} = 5.59$). The affine registration is hardly affected by the additional tissue class in the case of the head

Experiment	Data Set	HS	AR	1D	AR1D
Pepper Noise	Head	21.39	3.53	1.49	1.89
Pepper Noise	Kidney	40.54	24.50	1.90	17.96
Gaussian Noise	Head	14.54	11.90	8.27	13.64
Gaussian Noise	Kidney	14.29	9.63	7.07	11.96
Add. Class	Head	29.13	1.37	1.08	1.17
Add. Class	Kidney	31.24	10.50	4.66	7.09

Table 4.1: 1-d non-rigid standardization results using artificial perturbations. All numbers correspond to the MSE of the standardization result. For the pepper noise experiment a pepper probability of 50% is assumed. A standard deviation of $\sigma = 50$ is used in the Gaussian noise experiment. Finally, in the experiment with an additional artificial tissue class, the class has a size of 50% of the initial image pixel.

image. Here, the MSE is in average $MSE_{AR} = 1.37$ ($RMSE_{AR} = 1.17$). However, for the kidneys the MSE has a saltus at a relative class size of 15%. This results from the experiment with the mean intensity of the class being 220. The center intensity of the new class is slightly larger than the main intensity range of the kidney image. With the increasing class size, the affine registration includes this region into the main part of the histogram. This results in an average MSE of $MSE_{AR} = 10.50$ ($RMSE_{AR} = 3.16$) at a relative class size of 50%. Applying the 1-D histogram matching only induces a rather small MSE to the standardization result. At a relative class size of 50% this resulted in an average MSE of $MSE_{1D} = 1.08$ ($RMSE_{1D} = 1.04$) for the head image and a MSE of $MSE_{1D} = 4.66$ ($RMSE_{1D} = 2.16$) for the image of the kidneys. For the 1-D non-rigid histogram matching including a subsequent affine registration, the same behavior can be observed. However, the saltus in the affine registration case of the kidney is hardly visible in the result. Looking at the average MSE at a relative class size of 50%, this results in $MSE_{AR1D} = 1.17$ ($RMSE_{AR1D} = 1.08$) in the case of the head image and an average MSE of $MSE_{AR1D} = 7.09$ ($RMSE_{AR1D} = 2.66$) for the kidney image.

In table 4.1, the results of all three experiments are briefly summarized. The numbers correspond to the plots in Fig. 4.4.

MRI Head Data Sets

In the experiments using real clinically relevant head images, we use two public available data bases. First, 17 T1w data sets provided by the *Internet Brain Segmentation Repository* (IBSR)² of the Massachusetts General Hospital, Center for Morphometric Analysis are used. Second, we do qualitative experiments using ten public available SPL braun tumor resection image data sets (Tumorbase) provided by the Surgical Planning Laboratory (SPL), Brigham & Women’s Hospital, Boston Massachusetts, USA³.

²<http://www.cma.mgh.harvard.edu/ibsr/>

³<http://www.spl.harvard.edu/>

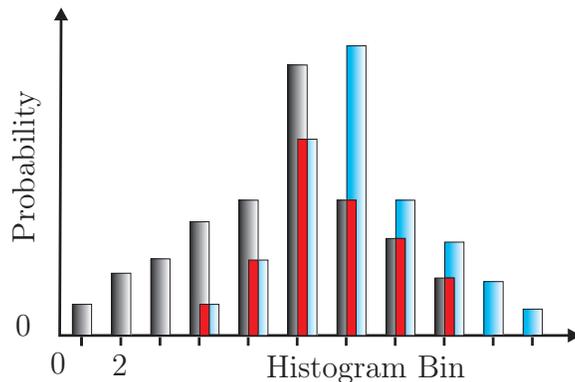


Figure 4.5: Histogram overlap distance measure: The distance between two relative histograms (gray and blue bars) corresponds to the sum of their overlap (red bars).

IBSR:

Data Sets: Originally, 18 data sets are included in the IBSR database, however, for one data set the gold standard segmentation is missing. For all other data sets a gold standard segmentation into the classes white matter (WM), gray matter (GM), CSF and background is available. All data sets are provided in a bias corrected form. The images have a slice thickness of 1.5 mm. The in-plane resolution is about $0.94 \text{ mm} \times 0.94 \text{ mm}$. The size of the data sets is $256 \times 256 \times 128$.

Evaluation Method: In this experimental setup the evaluation is based on the gold standard segmentations. To measure the quality of the intensity standardization the overlap between the histogram of a tissue class before and after the standardization to the histogram of the same tissue class in the reference image is calculated. Here, the overlap of two relative histograms is calculated as

$$d_O(\mathcal{H}_1, \mathcal{H}_2) = \sum_{i \in \Gamma} \min\{\mathcal{H}_1(i), \mathcal{H}_2(i)\} \quad (4.34)$$

where Γ is the intensity range of the histograms and $\mathcal{H}_1, \mathcal{H}_2$ are the histograms to be compared. This means that in the case of a bad standardization the overlap of the relative histograms is $d_O(\mathcal{H}_1, \mathcal{H}_2) = 0$ and for a perfect standardization the measure gives $d_O(\mathcal{H}_1, \mathcal{H}_2) = 1$. The similarity measure is illustrated in Fig. 4.5. As the gold standard segmentation contains three different tissue classes, we compute the measure for all three and average the result for all tissue classes. For evaluation one data set is used as reference. Then, all other images are standardized to the reference. Thus, in total $N_s = 17 \cdot 16/2 = 136$ standardizations are performed. From all results the mean as well as the standard deviation is computed.

Results: In Fig. 4.6 a few standardized images of the IBSR database that were corrected using the proposed 1-D histogram matching are shown. It can be seen that all images' histograms are matched to the source histogram in a very good manner. The visual impression is supported by the quantitative numbers. The proposed standardization approach reached a mean value of $\mu_{AR1D} = 0.862$. This means that all classes had an average overlap of 86.2%. The standard deviation is with $\sigma_{AR1D} = 0.058$

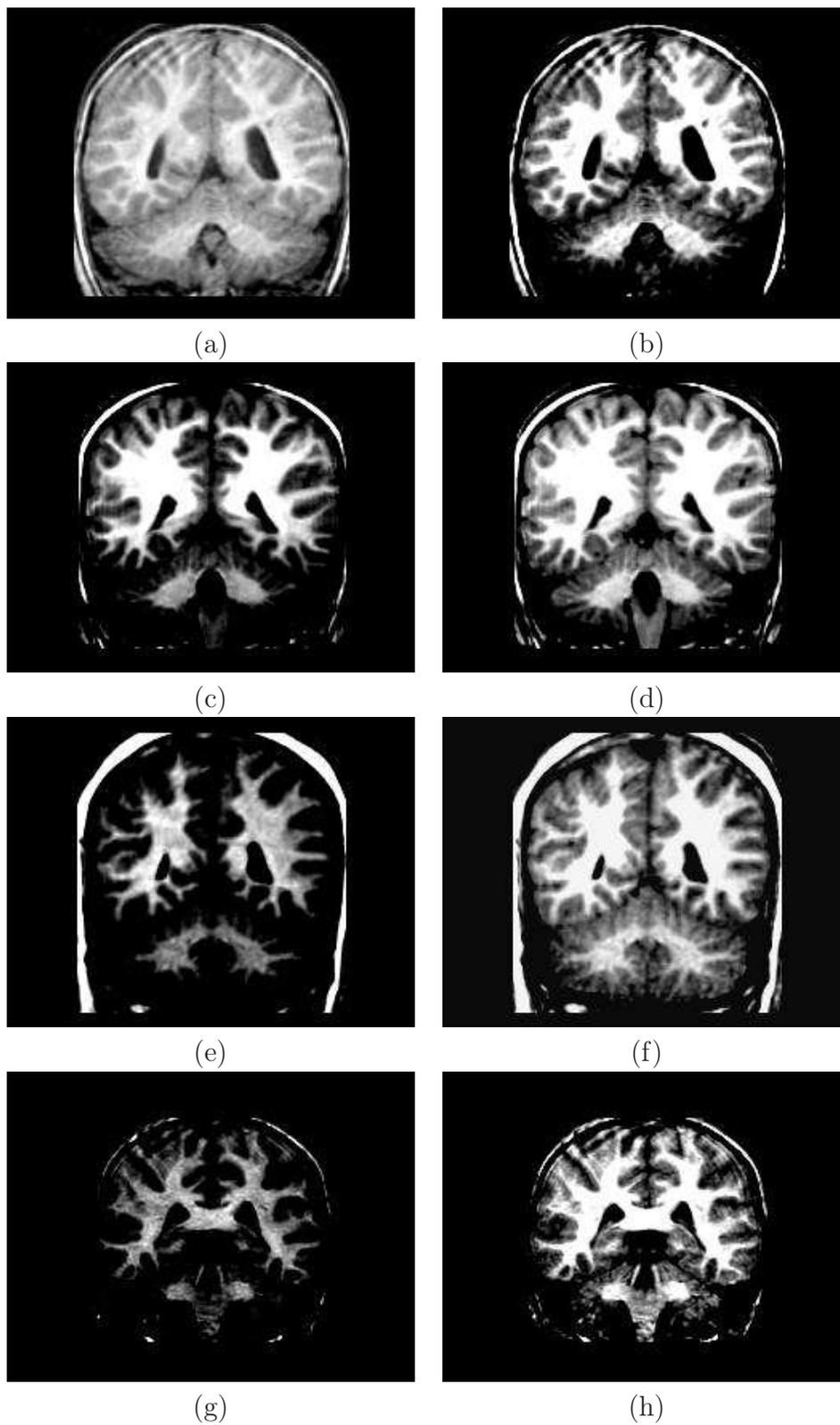


Figure 4.6: 1-d standardization results on IBSR T1w data sets: (a) Reference image (transfer function: center 65, width 130); (b) Reference image (transfer function: center 90, width 40); (c),(e),(g): target images (transfer function: center 90, width 40); (d),(f),(h): standardized images (transfer function: center 90, width 40)

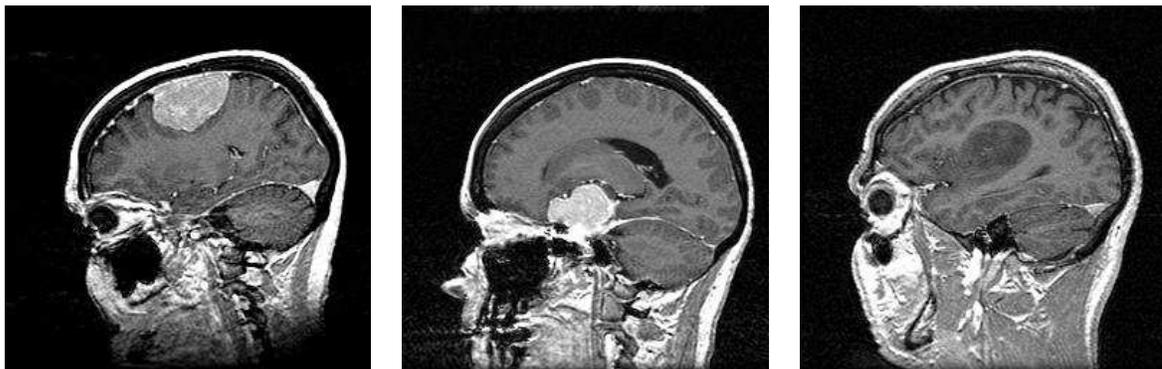


Figure 4.7: Three example slices from three different T1w images taken from Tumorbase.

rather small. The mean overlap before the standardization was $\mu_{bef} = 0.395$ with a standard deviation of $\sigma_{bef} = 0.310$. Consequently, the overlap was improved by approximately 218%.

A similar good result was achieved using histogram specification. Here, the mean value of the overlapping criteria was $\mu_{HS} = 0.842$. The standard deviation was with $\sigma_{HS} = 0.076$ also slightly higher than the deviation of AR1D. Due to the large number of evaluations, it is obvious that AR1D performed slightly better than histogram specification. The overlap is improved by 213% in average. However, it has to be kept in mind that the specification approach tries to find a one-to-one mapping between the histograms. Thus, pathologic structures that are not included in this quality measure are mapped to other tissue classes as well.

Surprisingly, Nyúl’s method has considerable problems dealing with the T1w data sets. Using the landmark configuration that relies on the second mode of the histograms yields a mean overlap of $\mu_N = 0.643$ with a quite large standard deviation of $\sigma_N = 0.188$. The reason for the bad result is that in some cases the second mode is not detected very well. That leads to a very bad overlap measure for these data sets. Consequently, the standard deviation is also higher than for the other methods. Using a different landmark configuration yields slightly better results. If the 50% percentile is used as second landmark, the mean overlap is $\mu_{N,50} = 0.701$ with a standard deviation of $\sigma_{N,50} = 0.123$. For these more heuristic configurations the results can never be as excellent because the data sets are from different patients. Consequently, all tissue classes have different sizes what alters the percentiles.

Tumorbase:

Data Sets: All ten public available T1w data sets provided by the SPL have a size of $256 \times 256 \times 124$. Their in-plane pixel resolution is $0.94 \text{ mm} \times 0.94 \text{ mm}$ with a slice thickness of 1.5 mm. All images show pathologies like meningioma, low grade glioma and astrocytoma. The acquisition order was LR. In Fig. 4.7 some example slices taken from the volumes are given.

Evaluation Method: For the Tumorbase data sets no gold standard segmentation is available. For this reason only a qualitative evaluation is done. In order to restrict

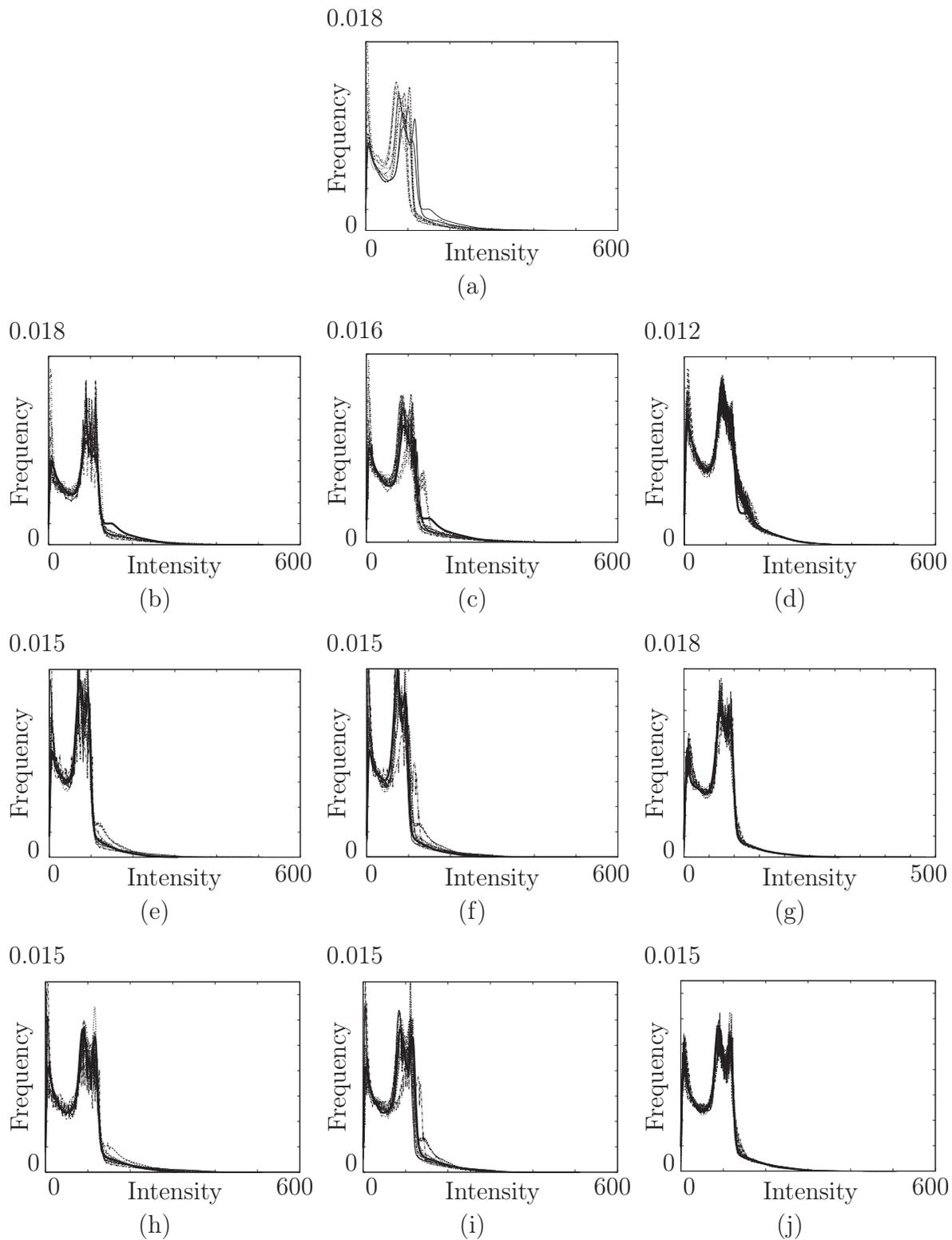


Figure 4.8: Standardization results of the Tumorbase images: In plot (a) the histograms of the unprocessed images are shown. The plots (b), (e) and (h) show the results using the proposed 1-D non-rigid histogram registration method using three different reference histograms. The results using Nyúl's method are illustrated in (c), (f) and (i). Finally, the output of histogram specification is shown in the plots (d), (g) and (j).

the calculations on relevant tissue classes, we use a spherical mask. It was centered in the middle of the volume and had a radius of half the size of each direction.

Results: In Fig. 4.8 the results of the standardization using the proposed AR1D method as well as the results of histogram specification and Nyúl’s method are shown. In order to illustrate the performance of the methods, we arbitrarily choose three data sets from the database. These data sets are used as reference for the intensity standardization. All other data sets are standardized to these references. Fig. 4.8 (a) shows the histograms of the masked data sets without any standardization. All histograms have very similar shapes. However, the distribution of the intensities do not fit very well. The Figures 4.8 (b), (e) and (h) show the result of AR1D. In all three cases it can be seen, that the approach aligns the histograms very well. Moreover, the structure of all histograms is preserved.

The standardization results for Nyúl’s method are shown in the Figures 4.8 (c), (f) and (i). As the second mode of the histograms did not match the same tissue class in all histograms, we could not use it as landmark. Thus, we choose the median of the intensities as landmark. Nyúl *et al.* propose this approach in [Ny100]. The plots show that the method improves the similarity between the histograms while keeping the structure of the single histograms. However, for one data set the median did not fit very well to the general median landmark. Thus, this histogram is corrected in a very bad manner. Overall the result is very good, but the intensity deviation of the standardized data sets to the reference one is obviously larger compared to the results of AR1D.

Finally, in the Figures 4.8 (d), (g) and (j) the histograms of the standardized images using histogram specification are shown. As expected the overall matching of the histograms is very good. But important shape characteristics of the histograms are not preserved at all. This can be observed in all three examples. Thus, the result of this approach is not usable in the medical context as this corresponds to a reduction of intensity information within the images. Most likely the information that was taken away corresponds to pathologies not visible in all images.

4.6.2 N-d Histogram Alignment

For evaluation of the proposed non-rigid alignment of n -D histograms (NAND) for intensity standardization various body regions, pulse sequences and different machines of the same brand were used. If possible the approach was evaluated using time-to-time studies (follow-up studies). However, in the case of whole body MRI, no voxel-wise evaluation was possible, as the deformation of the anatomical structures was too large even for follow-up studies, due to different positioning of the patients, non-rigid transformations of body regions, partial volume averaging effects and pathologies. As for the 1-d histogram matching, the method is independent from the size of the volumes. However, the performance is highly dependent on the number of bins. In a two image problem with a joint histogram size of 128×128 the standardization took about seven seconds.

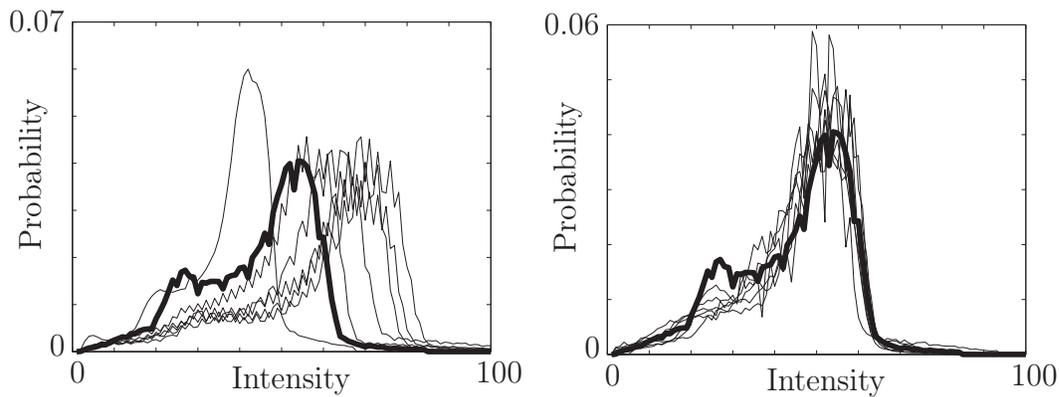


Figure 4.9: Left: the marginals of the T1-weighted head images before standardization; Right: the marginals of the same images after signal intensity standardization. The thick line in both plots represents the reference histogram.

Standardization of MRI Head Images

Data Sets: For the evaluation of the proposed intensity standardization method on MRI head data sets T1- and T2/FLAIR images were used. All data sets were acquired on a Siemens Symphony 1.5 T scanner. The T2-weighted FLAIR data sets had a resolution of $408 \times 512 \times 19$, an isotropic in-plane resolution of $0.43 \text{ mm} \times 0.43 \text{ mm}$, 7.2 mm slice thickness and $TE = 143$ and $TR = 9000$ ms. The T1-weighted images had a resolution of $208 \times 256 \times 19$ with an isotropic in-plane resolution of $0.86 \text{ mm} \times 0.86 \text{ mm}$, 7.2 mm slice thickness and $TE = 14$ and $TR = 510$ ms. In total 25 volumes from eleven different patients were used. All the images were of real patient data including evolving lesions.

Evaluation Method: For the experiments, a single T1-weighted volume and its corresponding FLAIR data set was chosen as reference for each patient. Then the follow ups were standardized to the intensities of the reference volumes. As the treated body region is small and the pdfs of brain images are not very complex, the volumes were not partitioned ($K = 1$). By subtracting the two image sets \mathbf{u} and \mathbf{v} intuitive quality measures can be computed.

In order to evaluate the intensity standardization the mean distance and the standard deviation between each patient's reference and target volumes were chosen as quality measures. The systematic error between the reference and the standardized images is expected to be smaller than that between the reference and the unprocessed images due to the signal intensity standardization. However, patient data with evolving structures from clinical routine was used and thus the anatomy of the brain changed. Consequently the evaluation method has the drawback that the difference between the volumes will never vanish. The lesions were removed by a segmentation step beforehand (just for evaluation, not for the signal intensity standardization step). Only those voxels that were classified as healthy brain tissue in both volumes were considered in computing the quality of the standardization. In order to be able to compute the difference between the volumes, a rigid registration using normalized mutual information [Hahn05] as distance measure had to be utilized, so that the reference and template volumes match each other.

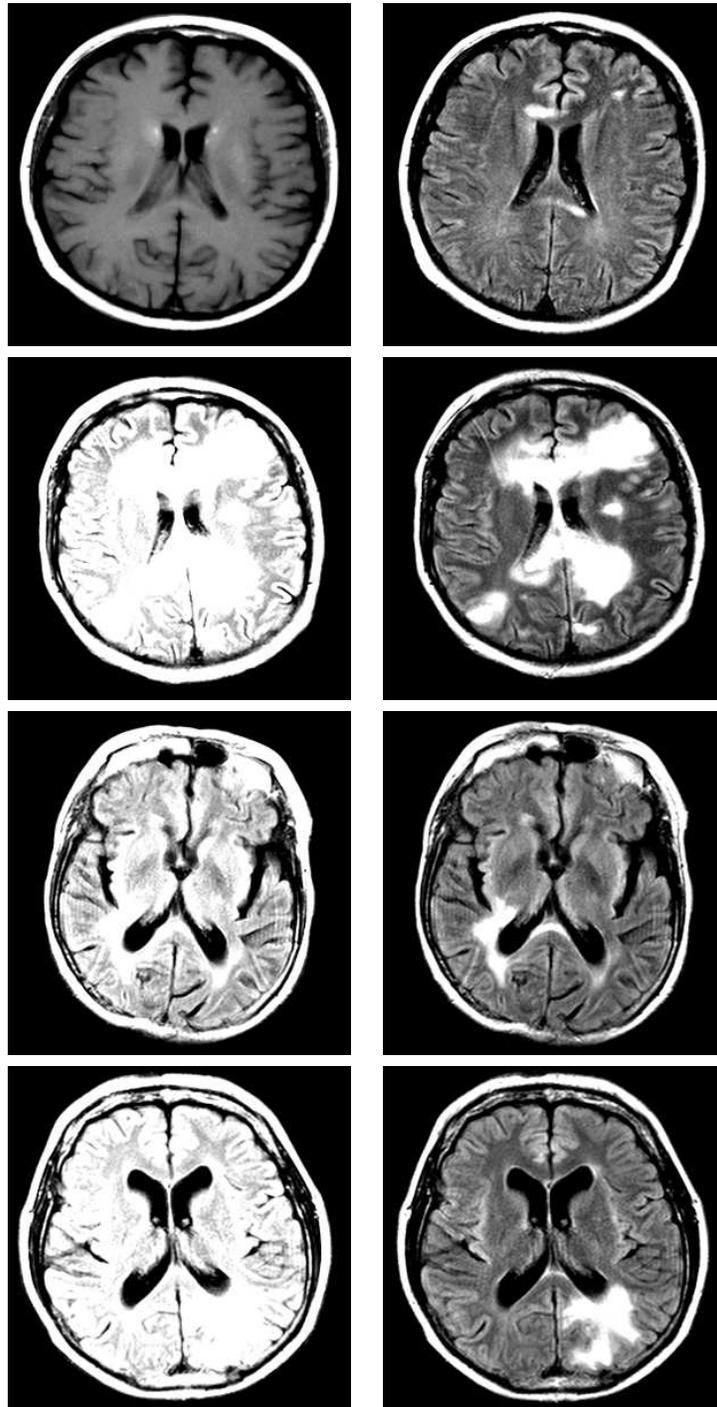


Figure 4.10: First row: On the left side a slice of the reference T1-weighted volume is shown, on the right the FLAIR slice respectively. Beginning with line 2: Left column: the original FLAIR slices. Right column: the corrected FLAIR slices. All images (except the T1w slice) are displayed with the same transfer function: center 200 and width 200.

Volume	FLAIR				T1			
	original		standardized		original		standardized	
	μ	σ	μ	σ	μ	σ	μ	σ
1a	19.94	62.41	27.51	62.27	90.98	57.82	4.65	55.01
1b	42.25	66.70	6.30	70.75	75.19	62.82	3.17	61.23
2	89.76	100.88	3.64	72.99	45.37	61.92	9.36	67.46
3	7.09	58.17	1.02	59.00	19.24	69.83	1.64	65.01
4a	80.11	79.58	8.67	58.95	53.39	74.52	0.39	77.93
4b	79.36	50.44	2.14	57.64	17.79	81.75	10.18	86.32
5	22.82	46.84	0.32	49.50	25.49	53.70	9.74	56.45
6	22.82	34.79	13.39	36.37	7.51	68.70	1.81	61.00
7	106.70	73.07	7.70	77.73	28.09	61.78	4.50	58.29
8	38.24	87.34	30.33	88.17	87.53	75.91	12.27	63.13
9	107.80	110.41	12.48	110.74	2.23	96.49	0.90	84.50
10	25.05	39.58	1.84	41.44	19.49	57.27	5.56	60.58
11a	2.48	56.81	3.02	55.87	15.49	52.84	16.53	51.74
11b	11.28	47.09	2.72	46.54	8.47	47.57	15.03	47.39
Means	46.84	36.41	8.65	9.17	35.45	29.02	6.84	5.12

Table 4.2: Standardization results of the MRI head images. Absolute mean and standard deviation of the difference images with and without standardization.

For the signal intensity standardization, we used the following parametrization for the registration: $\Delta\tau = 0.01$, $\alpha = 0.001$ and 128 bins. As interpolation method, a bi-linear interpolation was utilized. The histograms were created by a partial volume technique. This means that the bins were filled relative to the distance to the intensity vector. This configuration was suitable for most of the experiments. In some cases the parameter α had to be slightly adapted.

Results: The detailed results of the standardization are shown in Table 4.2. For the FLAIR images the absolute mean $\mu = (1/N_e) \sum_{j=1}^{N_e} |\mu_j|$ of all N_e experiments using the images without standardization was $\mu_{\text{FLAIR}} = 46.84$ and the corresponding standard deviation was $\sigma_{\text{FLAIR}} = 36.41$. Using the proposed intensity standardization approach the absolute mean was $\mu'_{\text{FLAIR}} = 8.65$ with a standard deviation of $\sigma'_{\text{FLAIR}} = 9.17$. For the original T1-weighted images the absolute mean was $\mu_{\text{T1}} = 35.45$ with a standard deviation of $\sigma_{\text{T1}} = 29.02$. Using the intensity standardization the mean was $\mu'_{\text{T1}} = 6.84$ with a standard deviation of $\sigma'_{\text{T1}} = 5.12$. Although, in general, the method significantly decreases the difference, there are a few cases that this difference may increase. There are two reasons for this effect:

First, the optimization of the registration might be stuck in a local minimum of the objective function.

Second, if the structure of the histograms does not fit precisely, the resulting intensity mapping might be incorrect. However, this just happens if the distance between the histograms was small from the beginning; thus, the resulting differences are still very small.

If the method proposed by Nyúl et al. [Nyl00] is employed the results are: $\mu_{\text{FLAIR}}^{\text{Nyul}} = 10.73$, $\sigma_{\text{FLAIR}}^{\text{Nyul}} = 8.30$, $\mu_{\text{T1}}^{\text{Nyul}} = 11.27$ and $\sigma_{\text{T1}}^{\text{Nyul}} = 7.57$. As the used images were bimodal, we utilized the second mode as histogram landmark. As mentioned before, a slight difference between the volume is expected as the anatomy of the brain changed [Jage06a]. Furthermore, we applied no bias correction method to the head images, because the intensity inhomogeneities of the acquired images were rather small. Figure 4.9 shows the marginals of the joint histograms of the T1-weighted head images before and after standardization. The marginals of the joint histograms correspond to the histograms of the single volumes. Here six volumes were randomly selected and standardized to a single reference. The thick line represents the reference histogram in both plots. In Figure 4.10 the effect of the signal standardization on FLAIR images is shown. The first row shows corresponding slices from the reference volume set illustrating all dominant tissue classes. On the left side the original image slices are shown. On the right side the processed images are shown. All images are displayed with the same transfer function; thus, comparable tissue classes should have similar gray values in the corrected images. Furthermore, all images are from different acquisitions, whereas the images in the second and third row show a follow-up study of the same patient. It can be clearly seen that the intensities are adapted to the intensities of the reference image.

Parametric Perturbation of Intensities

Data Sets: In this experiment parametric perturbations were applied on a pair of T1- and T2/FLAIR weighted images. They were acquired on a Siemens Symphony 1.5 T scanner. The T2-weighted FLAIR data set had a resolution of $408 \times 512 \times 19$, an isotropic in-plane resolution of $0.43 \text{ mm} \times 0.43 \text{ mm}$, 7.2 mm slice thickness and TE = 143 and TR = 9000 ms. The T1-weighted volume had a resolution of $208 \times 256 \times 19$ with an isotropic in-plane resolution of $0.86 \text{ mm} \times 0.86 \text{ mm}$, 7.2 mm slice thickness and TE = 14 and TR = 510 ms.

Evaluation Method: In all experiments in this section the used image pair was utilized as reference. Then the T1- and the T2/FLAIR- weighted images were perturbed by a parametric intensity mapping; thus, both data sets are transformed independently. With that setting, the joint standardization of all volumes has no advantage over the separate standardization. In this case the difference of NAND to Nyúl's approach [Nyl00] is reduced to the non-linearity of our approach and that it is not necessary for us to locate any landmarks in the histograms. In order to evaluate the standardization methods, the perturbed images were mapped back to the previously chosen reference data sets. As a measure for the quality of the standardization, the mean distance between the reference and the intensity standardized images was used.

The first intensity mapping we evaluated was

$$\chi_{\text{new}} = \chi_{\text{ref}} \cdot \left(\frac{(v - 1)}{\chi_{99.8}} \cdot \chi_{\text{ref}} + 1 \right), \quad (4.35)$$

with χ_{new} being the perturbed intensity, χ_{ref} the intensity in the reference image, $\chi_{99.8}$ the 99.8-percentile of all intensities in the reference image and v a factor to

control the strength of the perturbation. This means that if $v = 1$, the intensities do not change at all. If v is smaller than one then $\chi_{\text{new}} < \chi_{\text{ref}}$; if $v > 1$ then $\chi_{\text{new}} > \chi_{\text{ref}}$. Furthermore, the higher the value of the initial intensity the larger is the perturbation. We evaluated the signal intensity standardization for $v \in [0.3; 2.0]$ for Nyúl’s method [Nyl00] and the proposed approach. As the histograms were bimodal, in Nyúl’s method the second mode was utilized as landmark as recommended in [Nyl00].

Secondly, we evaluated

$$\chi_{\text{new}} = \chi_{\text{ref}} \cdot \left(1 + c \sin \left(\zeta \frac{\chi_{\text{ref}}}{\chi_{99.8}} \right) \right) \quad (4.36)$$

as an intensity mapping. Here the parameter c describes the amplitude of the distortion and ζ its frequency. We evaluated both algorithms with a frequency of $\zeta = 1$ and $\zeta = 4$. For $\zeta = 1$ we evaluated the amplitude $a \in [0; 0.5]$; for $\zeta = 4$ we chose $c \in [0; 0.35]$. An amplitude of $c = 0$ yields no intensity distortion.

Results: The results are illustrated in the upper three plots in Figure 4.11. It can be seen that both approaches decrease the mean distance to the reference images. However, in general, the NAND yields better results than Nyúl’s method [Nyl00]. Due to the piecewise linear intensity mapping between the chosen landmarks, the results of Nyúl’s method get worse if the perturbations become more non-linear. As the proposed signal intensity standardization method does not rely on the position of any landmarks, it is much more flexible and can easily adapt the correction to the non-linearity. If there are small intensity distortions only, Nyúl’s method slightly outperforms the proposed method. The reason for this is, that if there are hardly any intensity distortions then it is easy to exactly determine the same landmark positions in the histograms. As the proposed approach does not use any assumptions about the shape of the histograms and we are using a global regularization in the non-rigid registration, small changes in the perturbed histograms yield global changes in the intensity mapping.

Artificial Perturbation of the Intensities of a Single Tissue Class in Synthetic Images

Data Sets: For this experiment we used synthetic data sets from the McConnell Brain Imaging Centre (BIC) of the Montreal Neurological Institute, McGill University [Coll98]. The simulated data sets had a resolution of $181 \times 217 \times 181$ and a slice thickness of 1mm. We chose the T1w and T2w-images with a noise level of 3% and no signal intensity inhomogeneities. In order to evaluate the proposed algorithm, the intensities were artificially perturbed by a warping of the joint histograms using thin plate splines. Slices from the synthetic images are illustrated in Figure 4.12. Figure 4.13(a) shows the joint histogram of the original T2w and T1w image. In Figure 4.13(b) the joint histogram with the maximal intensity deformation is shown. For the experiments the intensity deformation is varied between no and maximal distortion.

Evaluation Method: For the evaluation, the mean absolute distance of the reference image to the standardized intensity transformed images was chosen. In order to

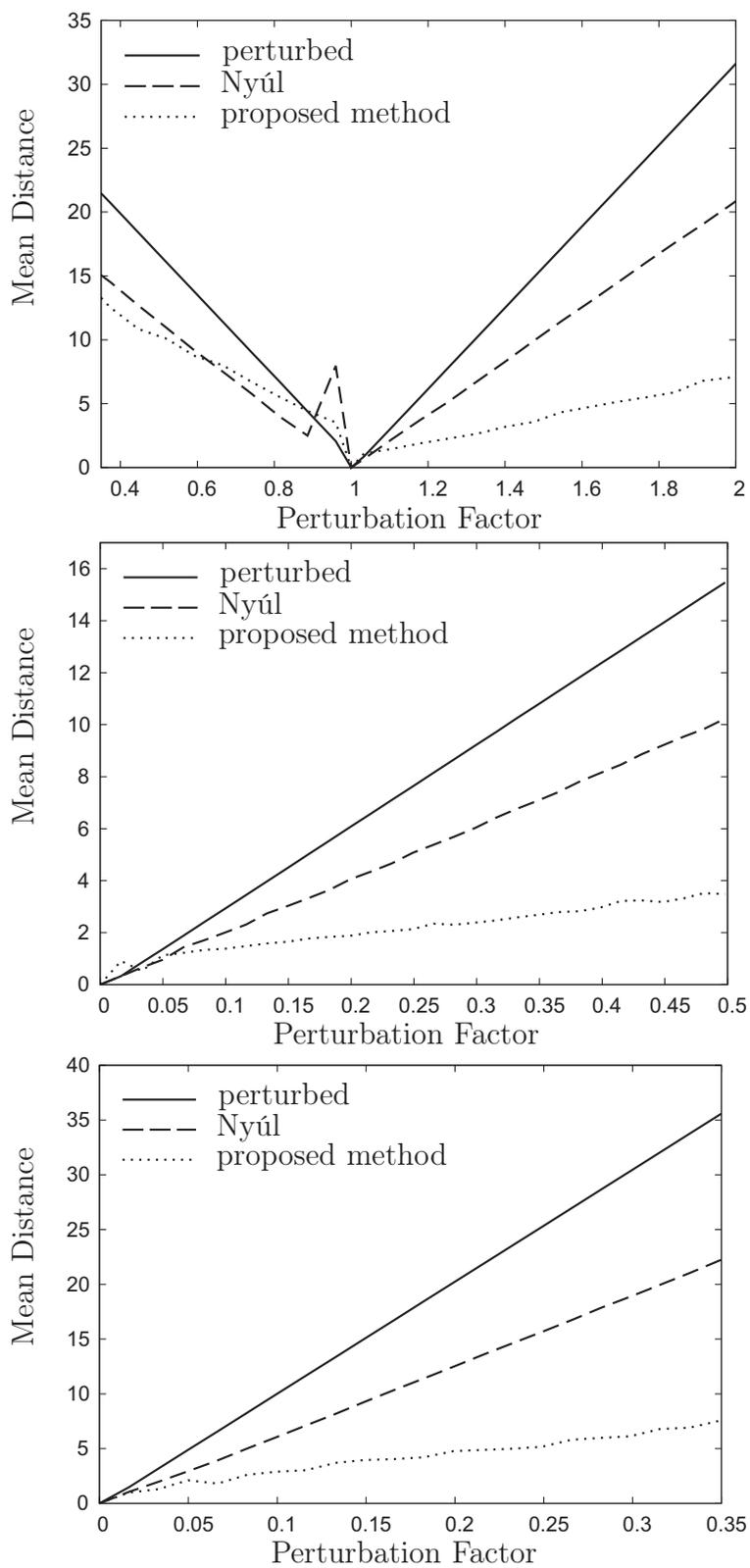


Figure 4.11: N-d non-rigid standardization results using an artificial perturbation: The first plot shows the results of the standardization for the linear intensity perturbation of the images. In the second and third plot the results for the sinusoidal perturbation are shown ($\zeta = 1$ and $\zeta = 4$).

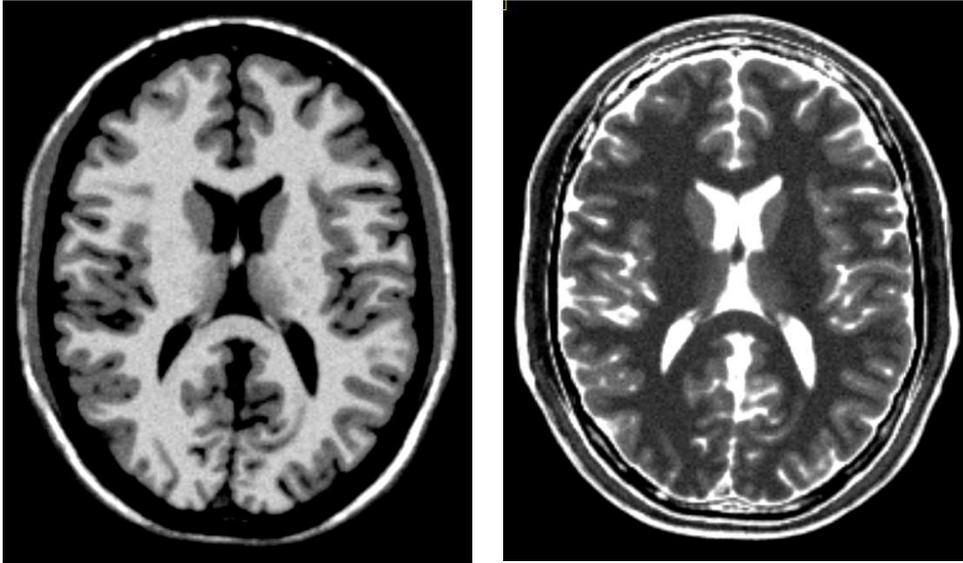


Figure 4.12: The left image shows the used synthetic T1w image. On the right side a corresponding T2w slice is shown. The images are taken from the BrainWeb database [Coll98].

create a set of transformed images the intensity deformation was scaled between no deformation and maximal deformation. The measures of NAND were compared to Nyúl’s method [Nyl00]. Here, we tried different configurations and found, that using the percentile set $\mathbf{p} = \{10, 20, \dots, 90\}$, where each element $p_i \in \mathbf{p}$ represents a p_i th percentile, in the histogram yields the best results.

Results: This experiment demonstrates the main advantages of the signal intensity standardization using all acquired images jointly compared to the normalization of the images separately. It can be seen that if just the intensities of a certain tissue class are deformed, these changes cannot be corrected by aligning the 1-d histograms to a given reference. The reason for this is that the translation of an intensity in the projection results in the translation of a complete row in the joint histogram. Thus, not only the distorted tissue class is modified. The result of the experiment regarding the T1w image is shown in the last plot of Figure 4.14. As illustrated in Figure 4.13, the plot shows that the proposed method outperforms Nyúl’s method [Nyl00] given the artificial intensity distortion. Figure 4.13(d) shows the standardization result of the images with maximal intensity deformation. The intensities of the transformed intensity class were not corrected properly. Consequently, the mean distances after standardization are worse than before. Furthermore, streak artifacts can be seen at the borders of the percentile positions. Using a continuous transformation instead of the piece-wise linear transformations could reduce these histogram artifacts. In Figure 4.13(c) the normalization result of the proposed method is shown. The intensity deformation was reduced significantly. However, due to the regularization and the loss of some intensities during the distortion step, no perfect standardization, in the sense of zero distance, was reached.

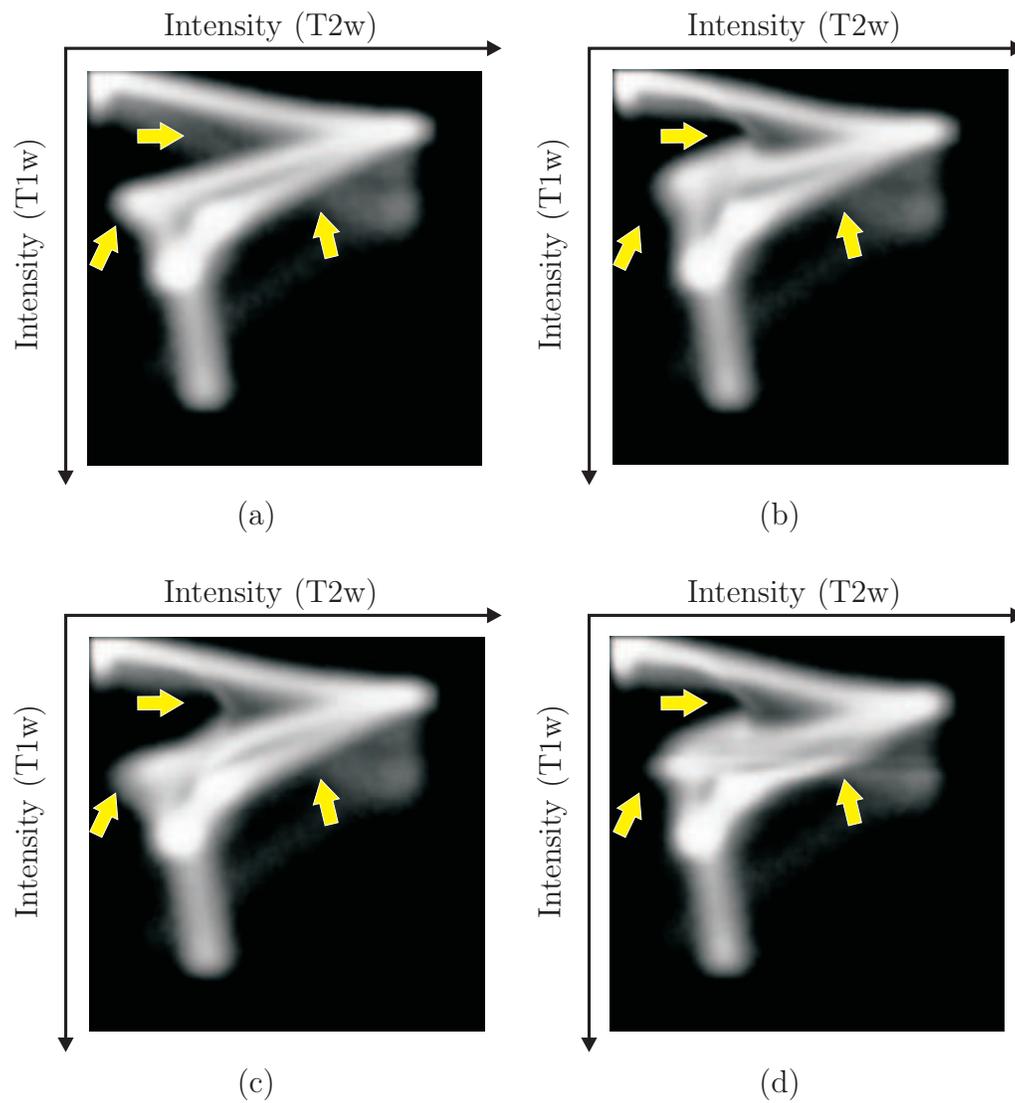


Figure 4.13: Image (a) shows the original joint histogram of the synthetic T2w and T1w image. The histogram is equalized for presentation. The x -direction represents the T2w image and the y -direction the T1w image. In figure (b) the joint histogram of the intensity transformed images is shown. (c) shows the correction result using the proposed method. (d) illustrates the correction result of Nyúl's method [Nyl00]. The arrows indicate regions where the differences between the standardization techniques can be seen best.

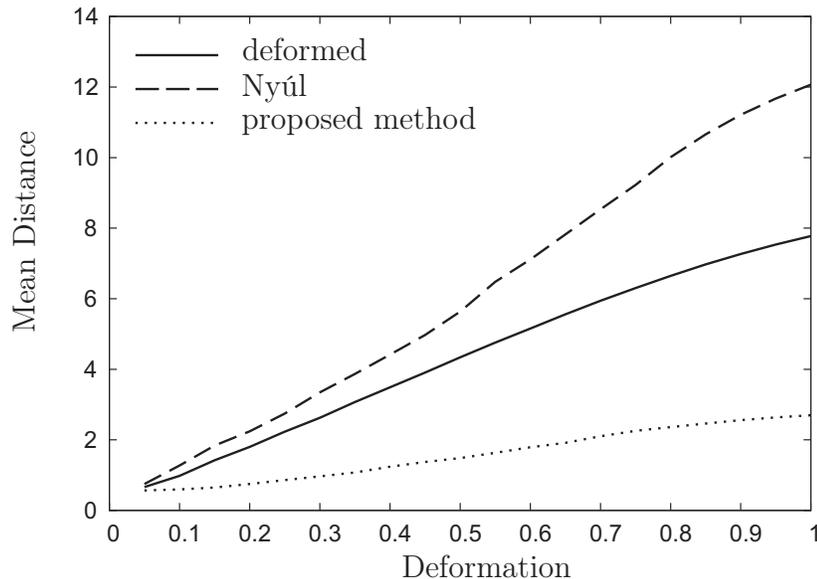


Figure 4.14: N-d standardization results using an artificial intensity transformation of a single tissue class: The plot shows the mean distance of the intensity perturbed T1w image, the corrected version using Nyúl’s method [Nyl00] and the standardized method using the proposed method with respect to the degree of the intensity distortion.

Standardization of Head Data Sets Including Artificially Evolving Pathologies

Data Sets: For the evaluation of the influence of pathologies on the proposed signal intensity standardization T2/FLAIR weighted images and T1 weighted images with contrast bolus were used. All data sets were acquired on a Siemens Symphony 1.5 T scanner. The T2-weighted FLAIR data sets had a resolution of $408 \times 512 \times 19$, an isotropic in-plane resolution of $0.43 \text{ mm} \times 0.43 \text{ mm}$, 7.2 mm slice thickness and $TE = 143$ and $TR = 9000$ ms. The T1-weighted images had a resolution of $208 \times 256 \times 19$ with an isotropic in-plane resolution of $0.86 \text{ mm} \times 0.86 \text{ mm}$, 7.2 mm slice thickness, $TE = 14$ and $TR = 510$ ms and 14 ml of *Magnevist* contrast bolus.

Evaluation Method: Two evolving pathologies were simulated in this experiment. First, the ventricles were increased (see Fig. 4.15) and second, a lesion within the data set was artificially enlarged (see Fig. 4.15). In order to simulate the growth of the pathologies, a landmark based non-rigid registration approach was employed. For this purpose a set of landmarks $\mathbf{L} = \{\mathbf{l}_1, \dots, \mathbf{l}_n\}$, where $\mathbf{l}_i \in \mathbb{R}^3$, $i \in \{1, \dots, n\}$, on the border of the structure (ventricles / lesion) were selected in the treated volume pair \mathbf{v} . These n landmarks were manually assigned to new positions $\mathbf{L}' = \{\mathbf{l}'_1, \dots, \mathbf{l}'_n\}$. Additionally, for each landmark \mathbf{l}_i a second pair of landmarks at a distance δ outside the structure was introduced to keep the transformation as local as possible. Finally, the corners of the volumes were used as static landmarks. Thus, in total $n * 2 + 8$ landmarks were utilized. The registration approach used based on thin plate splines [Rohr01]. After the alignment the landmarks are matched exactly. The values between the landmarks are interpolated smoothly. The result of the registra-

tion is a mapping $\Upsilon : \mathbb{R}^3 \mapsto \mathbb{R}^3$ that describes the transformation of a voxel. By multiplying Υ by a constant factor $g_a \in [0, 1]$, the resulting mapping Υ_a generates a new volume pair \mathbf{v}_a , where the size of the treated structure lies between the size in the original volumes and the size in the volumes mapped by Υ . This is illustrated in the Figures 4.15 and 4.16.

Let \mathbf{u} be a reference volume set. The computed mapping of the intensity space of a target volume set \mathbf{v}_a to the reference set \mathbf{u} is denoted by $\gamma_a : \mathbf{v}_a \mapsto \mathbf{u}$. In order to be able to compare the result of the signal intensity standardization, the resulting intensity transformation γ_a is applied to the original volume pair \mathbf{v} yielding $\mathbf{v}_{\gamma_a} = \gamma_a(\mathbf{v})$. If pathologies have no influence on the signal standardization, all corresponding volume pairs \mathbf{v}_{γ_a} should have the same intensities for all voxels. Let $\bar{\mathbf{v}}$ be the mean of all signal standardized volumes \mathbf{v}_{γ_a} . Thus, the pathology dependency can be measured by the mean distance μ_p and its standard deviation of an observed volume pair \mathbf{v}_{γ_a} to the mean volume pair $\bar{\mathbf{v}}$.

Results: In the experiment using the artificially enlarged lesion, the volume of the lesion varied between approximately $v_0 = 38\text{cm}^3$ for $g_0 = 0$ to about $v_1 = 136\text{cm}^3$ for $g_1 = 1$. This corresponded to a lesion to head ratio of about $r_0 = 0.02$ to $r_1 = 0.08$. The results are illustrated in the plots of Figure 4.17. The mean absolute difference of all transformed FLAIR images yielded $\mu_l^{\text{FLAIR}} = 0.26$ with a standard deviation of $\sigma_l^{\text{FLAIR}} = 0.13$; the mean absolute difference of the T1w images was $\mu_l^{\text{T1}} = 0.54$ with a standard deviation of $\sigma_l^{\text{T1}} = 0.37$.

In the experiment using the artificially increased ventricles, the volume of the ventricles varied between approximately $v_0 = 85\text{cm}^3$ for $g_0 = 0$ to about $v_1 = 316\text{cm}^3$ for $g_1 = 1$. This corresponded to a ventricle to head ratio of about $r_0 = 0.04$ to $r_1 = 0.14$. The results are illustrated in the plots of Figure 4.18. The mean absolute difference of all transformed FLAIR images yielded $\mu_v^{\text{FLAIR}} = 0.40$ with a standard deviation of $\sigma_v^{\text{FLAIR}} = 0.23$; the mean absolute difference of the T1w images was $\mu_v^{\text{T1}} = 0.64$ with a standard deviation of $\sigma_v^{\text{T1}} = 0.49$.

The input images had gray values in the range between zero and approximately 1200. Thus, the maximal measured intensity deviations show that the proposed algorithm is robust against severe pathological changes within the volumes.

4.6.3 Standardization of Whole Body Data Sets

Data Sets: All data sets were acquired on a Siemens Avanto 1.5 T whole body MRI scanner. The TIRM images had a resolution of $512 \times 512 \times 30$ (each block) with an isotropic in-plane resolution of $0.98 \text{ mm} \times 0.98 \text{ mm}$ and 5.5 mm slice thickness and $\text{TE} = 83$, $\text{TR} = 1660 \text{ ms}$ and the FL2D images had a resolution of $512 \times 410 \times 30$ (each block) with an isotropic in-plane resolution of $0.98 \text{ mm} \times 0.98 \text{ mm}$ and 5.5 mm slice thickness and $\text{TE} = 4.7$ and $\text{TR} = 291 \text{ ms}$. The size of the composed whole body images was $542 \times 1746 \times 20$ for both protocols. Only the composed volumes were used for the experiments. All images were acquired in clinical routine, thus, including pathologies. In total nine whole body MRI data sets were used for evaluation. Two of these were from the same patient, acquired with a time delay of six months. In order to reduce the effects of intensity inhomogeneities homomorphic unsharp masking was employed.

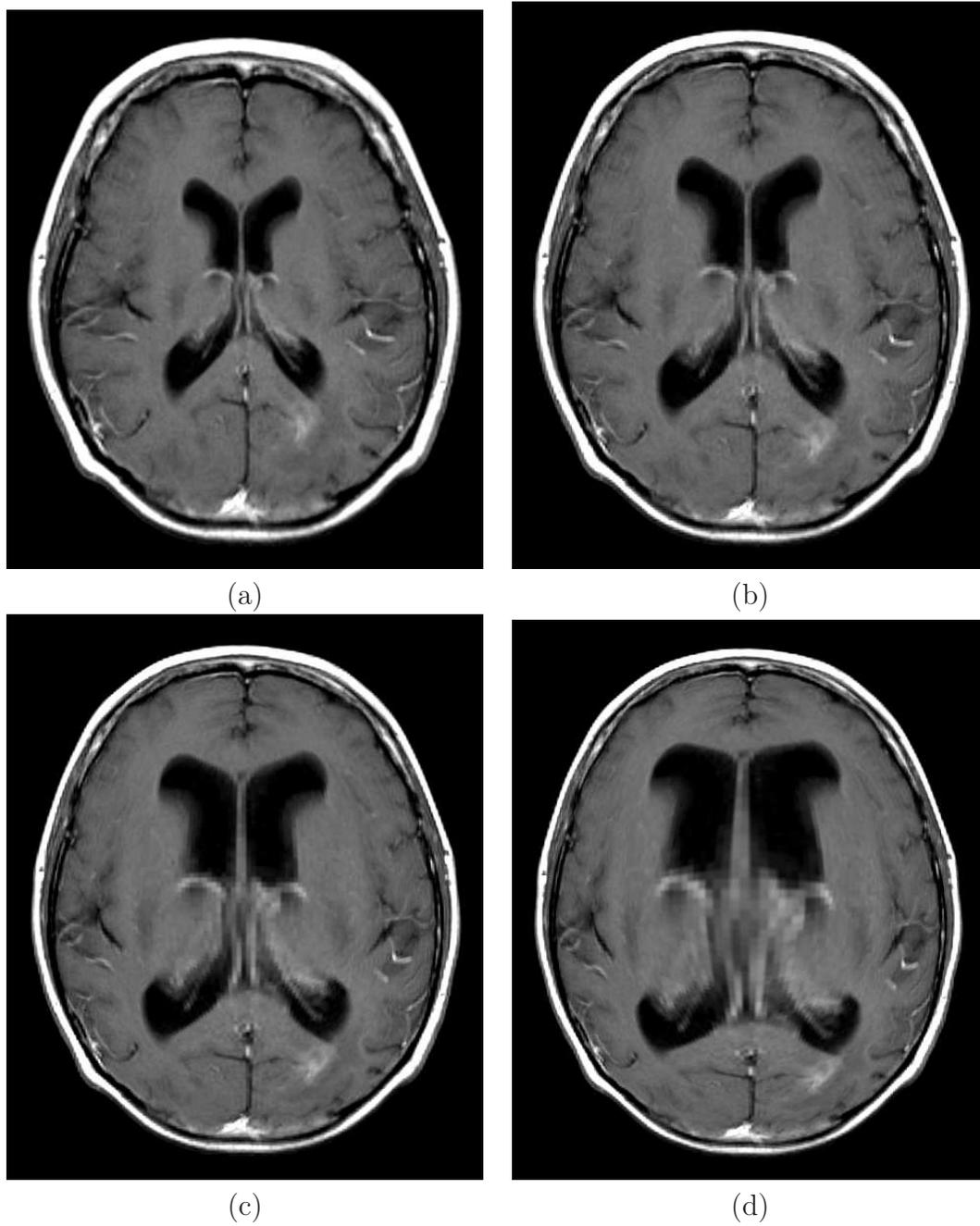


Figure 4.15: The images (a)-(d) show the artificial growth of the first two ventricles: (a) original slice, (b) 33%, (c) 66% and (d) maximal deformation.

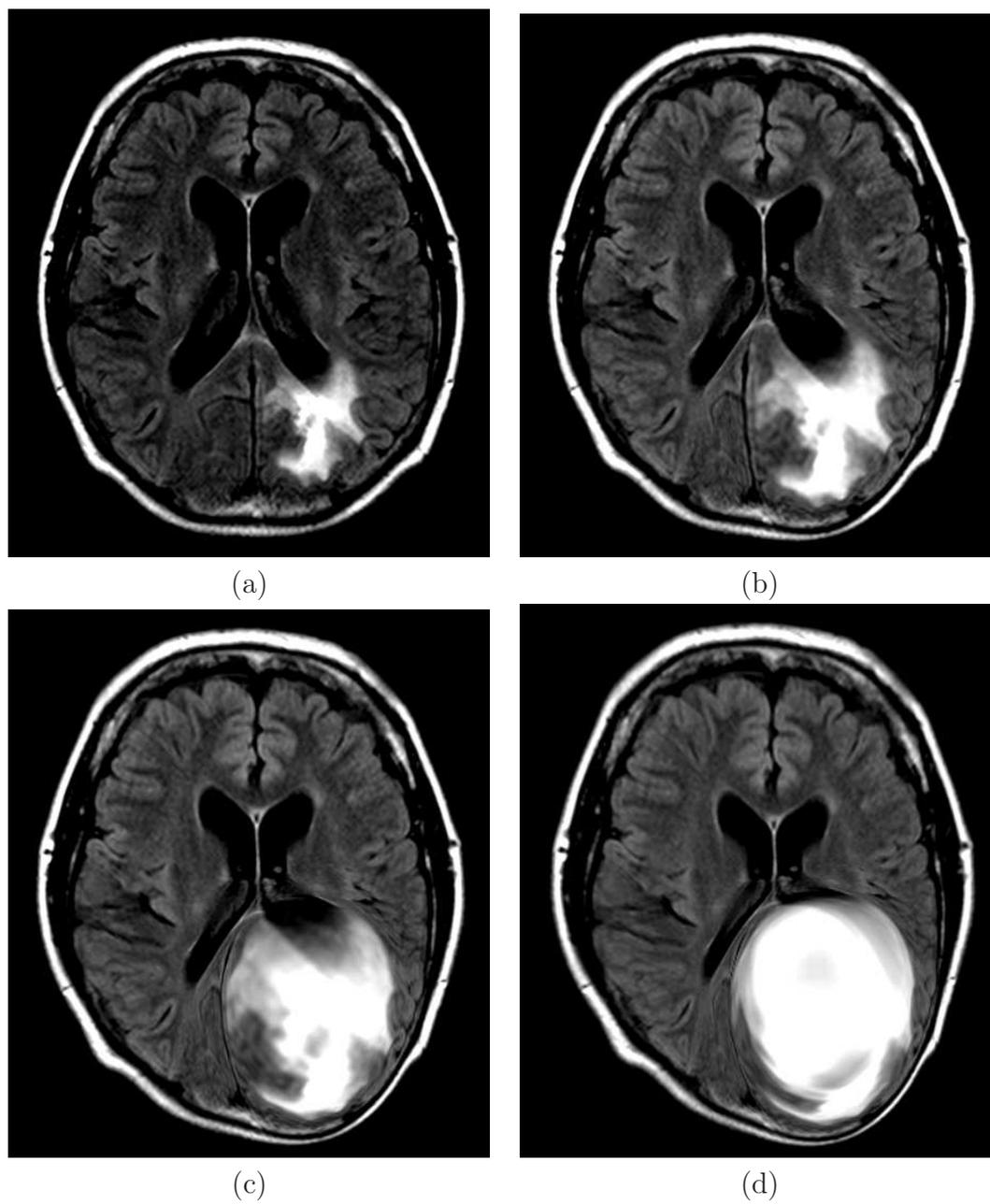


Figure 4.16: In the images (a)-(d) the artificial growth of an lesion is shown. Image (a) shows the original FLAIR slice; (d) the maximal deformation; (b) 33% of the maximal deformation and (c) 66% respectively.

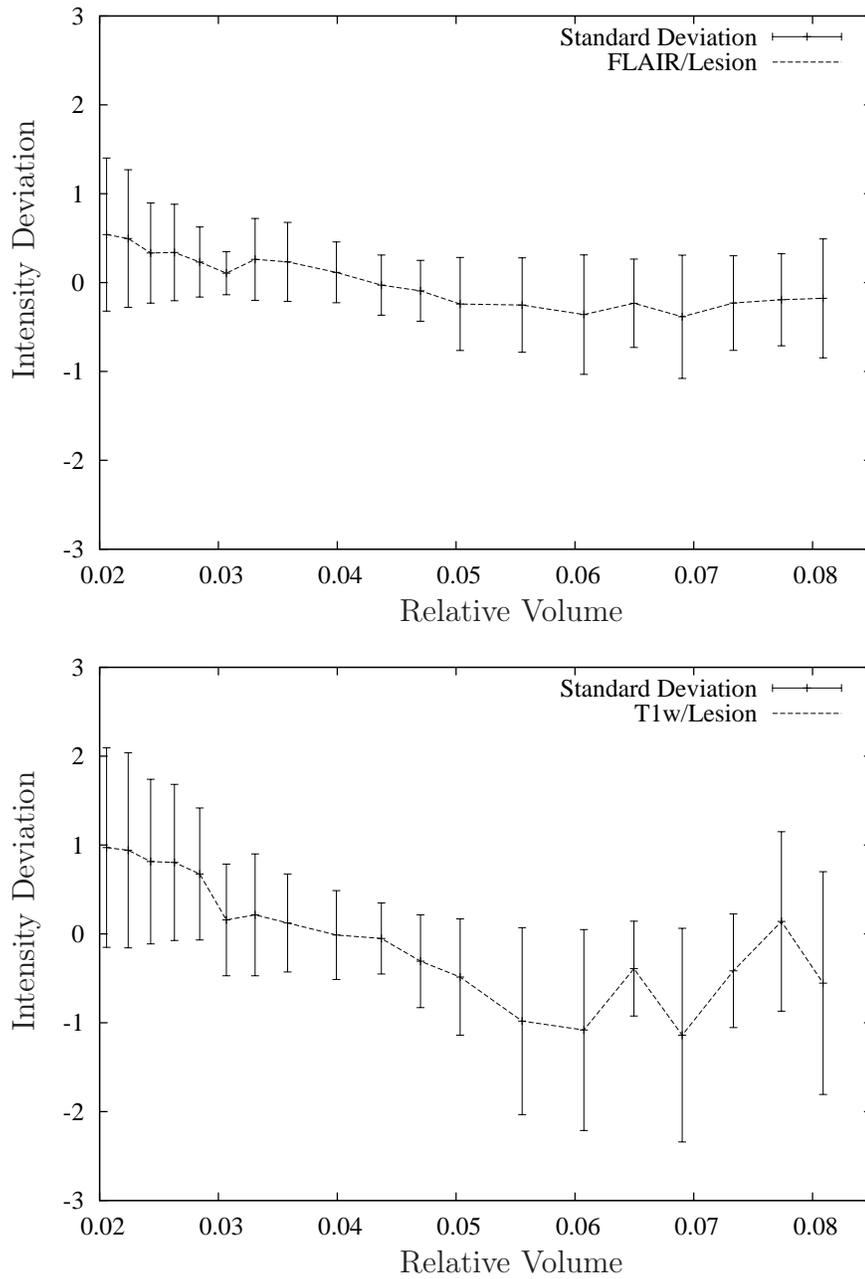


Figure 4.17: Results for n -D non-rigid histogram standardization of artificially growing ventricles: The plots show the difference of the transformed volumes with different sizes of the lesion to the "mean" transformed volume. The intensity range of the images is approximately $[0, 1200]$.

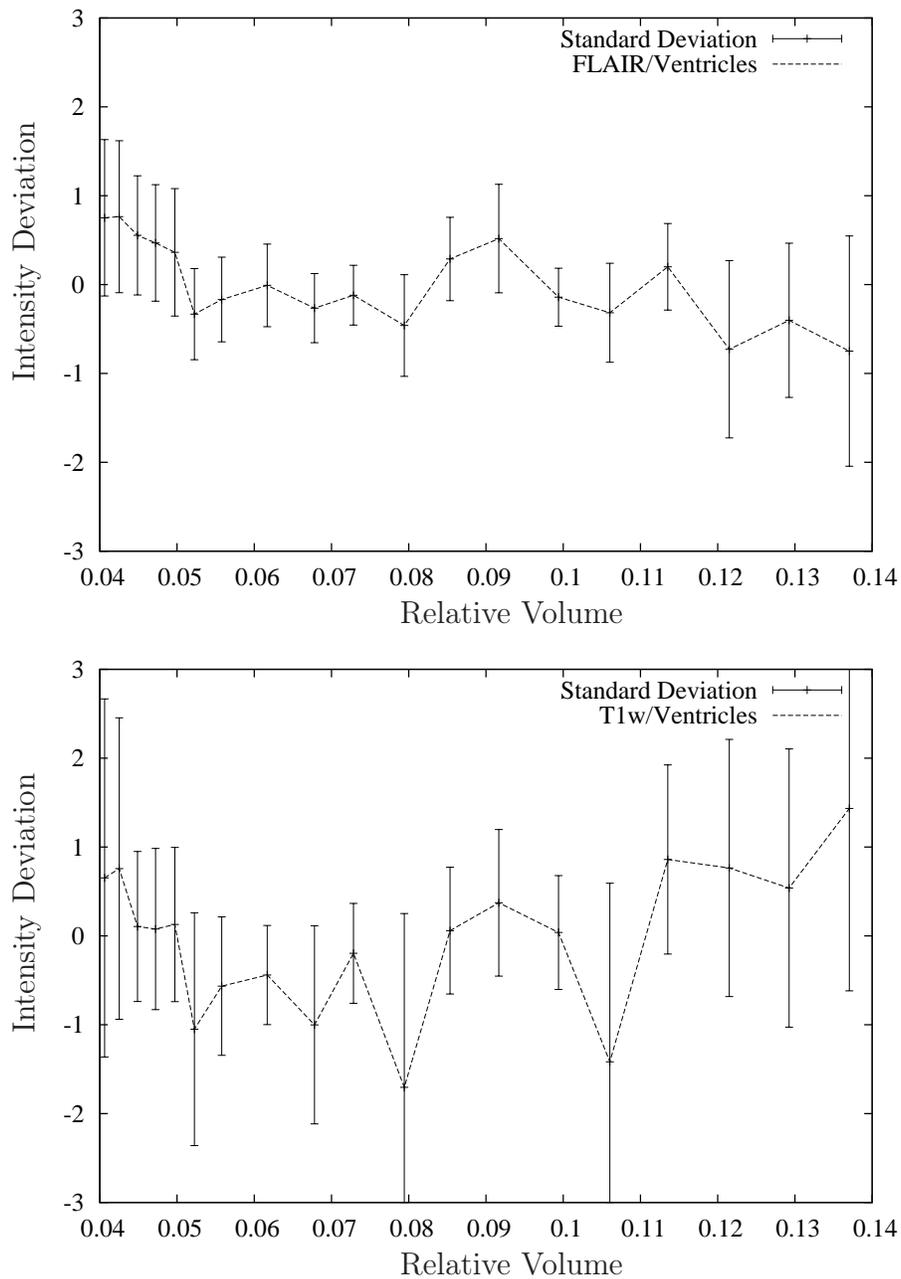


Figure 4.18: Results for n -D non-rigid histogram standardization of artificially growing ventricles: The plots show the difference of the transformed volumes with different sizes of the ventricles to the "mean" transformed volume. The intensity range of the images is approximately $[0, 1200]$.

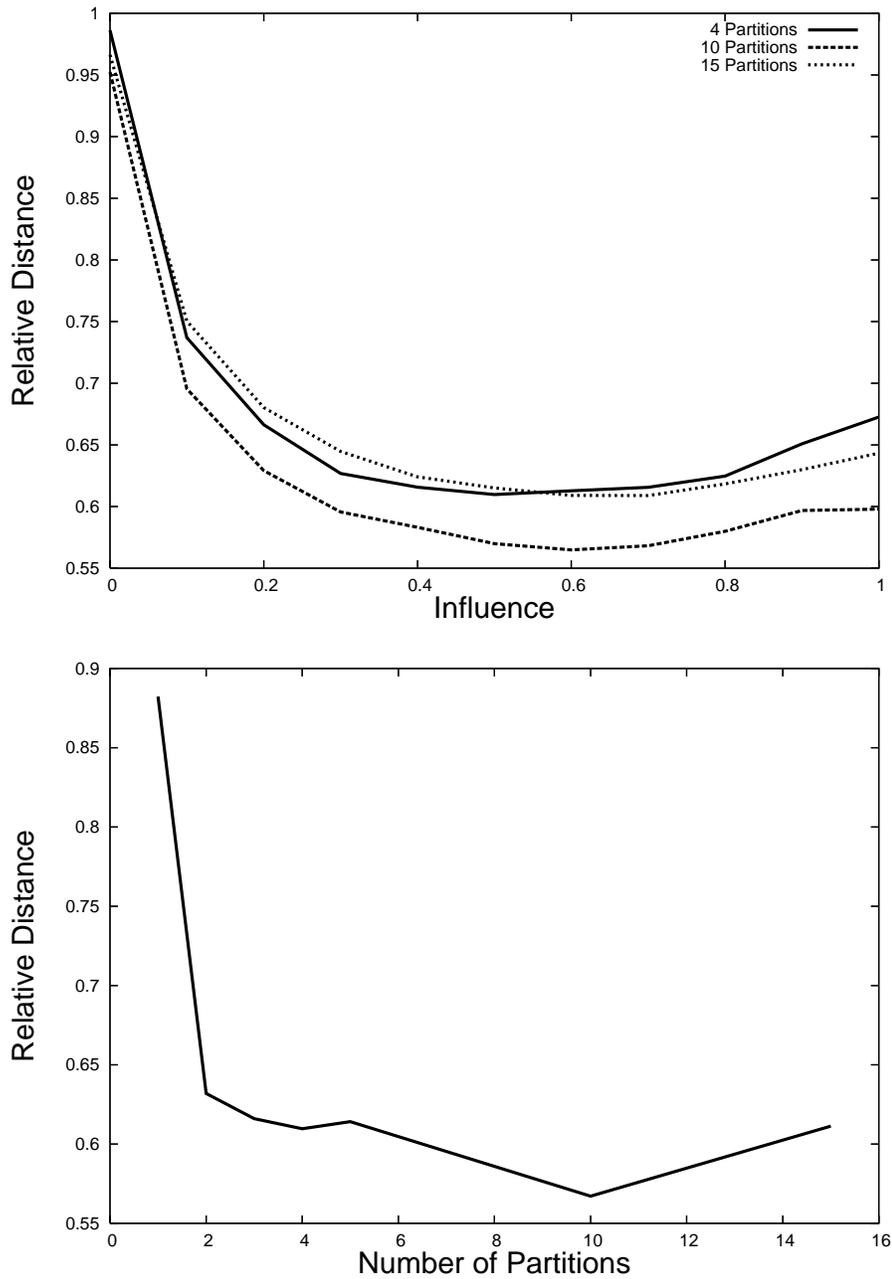


Figure 4.19: The top plot shows the results of the intensity standardization with respect to the influence parameter ψ . The bottom plot illustrates the effect of different numbers of partitions on the standardization.

Evaluation Method: Due to the anatomical differences and the large deformations within follow-up studies, no voxel-wise evaluation method was used. As quality measure the relative distance

$$r_{\text{JH}} = \frac{d_{\text{JD,After}}}{d_{\text{JD,Before}}}. \quad (4.37)$$

between the reference and current joint histogram before and after the standardization was chosen. Consequently, if $r_{\text{JH}} < 1$ the distance between the images was reduced and if $r_{\text{JH}} \geq 1$ the standardization failed. As distance measure d_{JD} between the joint histograms the discrete version of JD, introduced in Eq. (4.9), is used. Similar to the KL-Divergence, JD measures how efficient, on average, it would be to code one histogram using the other as code book [Rubn00]. Our assumption is that if the histograms of two images are as close as possible, a tissue class in both images covers the same intensity range in both volumes. Hence the distance between the joint histograms is a good measure for the quality of the results of the signal intensity standardization.

In the first experiment we evaluated the standardization quality related to the influence of neighboring histograms on the regularization. We used a uniform partitioning along the z-axis. Only direct neighbors to the current histogram are allowed to affect the registration. For this reason we introduced an influence parameter $\psi \in [0; 1]$ with

$$a_{j,k} = \begin{cases} \psi, & \text{if } j = k, \\ (1 - \psi)/2, & \text{if } k = j - 1 \vee k = j + 1, \\ 0, & \text{else.} \end{cases} \quad (4.38)$$

where j is the current sub-volume to be registered. Consequently, $\psi = 1.0$ does not use statistical information of neighboring histograms and $\psi = 0.0$ uses neighboring pdfs only for the standardization. The results showed that the parameter ψ has to be chosen between $\psi \in [0.5; 0.7]$, depending on the number of partitions. If fewer partitions are used, then the influence has to be reduced. Note that, because the first and the last sub-volume are mirrored, their influence on the registration result is not just ψ but $\psi + (1 - \psi)/2$. The standardization results using different numbers of partitions and a varying parameter ψ are illustrated in Figure 4.19 (upper plot).

The second experiment deals with the number of partitions to be used. In order to evaluate this parameter, the best signal intensity standardization result for each number of partitions (the parameter ψ may vary) is selected. This shows that approximately ten sub-volumes have to be chosen. If more partitions are used the pdfs cannot be estimated in a reliable manner anymore. The reason for this is that the data sets are rather noisy and consequently there is not enough intensity information included in smaller partitions. If fewer partitions are used, the standardization is affected by intensity inhomogeneities and small structures are neglected in the registration. The standardization results are illustrated in Figure 4.19 (lower plot).

Results: The results of both experiments show that it is best to use approximately $K = 10$ partitions. They also show that the statistical information of neighboring partitions has a significant impact on the quality of the standardization of a sub volume. For the given number of partitions a parameter $\psi \approx 0.6$ provided the best results. Using the Jeffrey divergence it was possible to achieve a relative distance of $r_{\text{JH}} = 0.567$. This means that the distance between the joint histograms has been

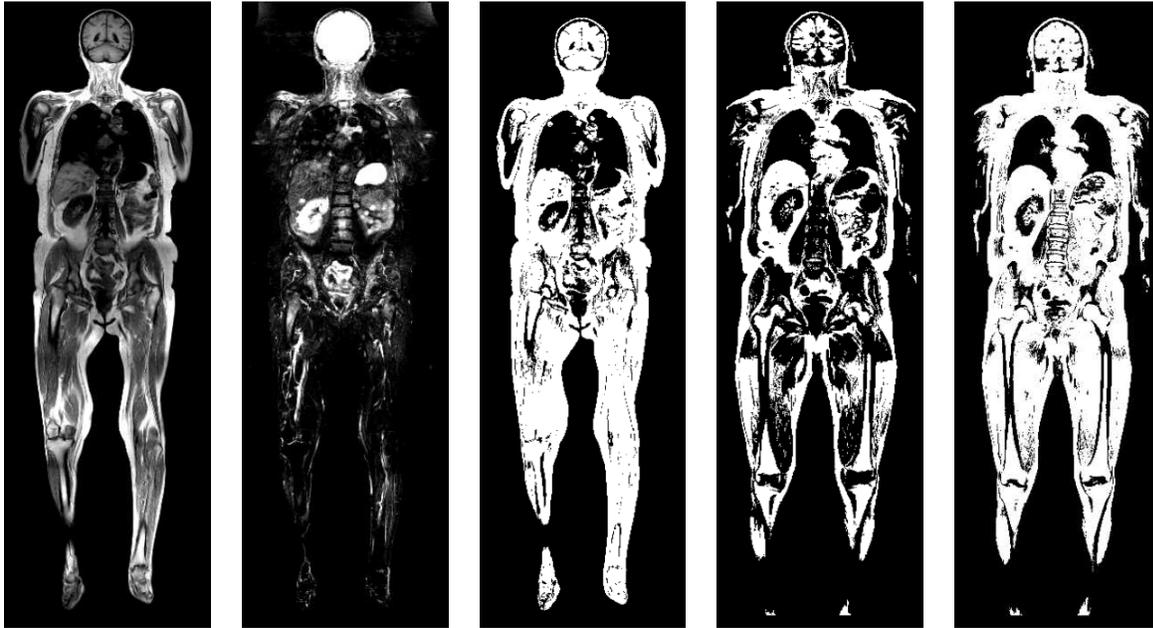


Figure 4.20: From left to right: slice from the reference T1w volume; slice from the reference TIRM weighted data set. A threshold of 400 was applied to the third, fourth and fifth image to better visualize the differences. The third image shows a reference T1w slice, the fourth the original T1w slice of a different patient and the fifth image the standardized T1w slice.

reduced by about 45%. However, it is not possible to achieve significantly better results as the anatomical differences between the data sets are too large. Consequently, even if the signal intensities of all tissue classes are transformed in an ideal manner, there still have to be differences between the histograms. The results of the whole body signal intensity standardization are illustrated in Figure 4.20.

4.7 Summary & Discussion

A major disadvantage of MRI compared to other modalities is that its intensities do not follow a pulse sequence dependent general intensity scale like Hounsfield units in CT. In the last decade, several methods for this standardization problem were developed. In the medical context, however, it is a necessity that the estimated intensity transformation is invertible and that it preserves the structural characteristics of the images' histograms. Due to these restrictions most approaches are application dependent, unprecise, and or have a high computational complexity. All methods for intensity standardization have in common that a global intensity transformation is computed. This is the major difference to the correction of intensity inhomogeneities where a spatially dependent mapping is estimated (compare chapter 3).

All novel methods for MRI signal intensity standardization, we presented in this chapter, are closely related to image registration. In contrast to most of the previously published methods, the proposed approaches are independent from any prior

knowledge about the structure of the data sets and rely on the image statistics only. Hence, no prior registration or segmentation of the data sets are necessary. However, a separation of foreground and background pixel is highly recommendable. Otherwise, the optimization process concentrates on the dominant background peak within the histograms. Furthermore, the methods are independent of the application, body region and pulse sequence used for acquisition, if a reference standard intensity scale is available.

The most basic method for intensity standardization, introduced in this chapter, relies on the estimation of an affine mapping between two 1-d histograms. Due to the very limited number of unknown parameters, it is fast and can be implemented in a straight forward manner. However, because of the lack of flexibility the results are not very accurate. This is the reason, why it is usually used for pre-alignment only to cover large intensity deviations.

In order to refine the correction of local deviations in the histograms, we presented a 1-d non-rigid matching approach. Here, a non-parametric deformation of the intensities is estimated. To overcome unrealistic transformations and to integrate knowledge about the smoothness of the deformation, an additional regularization term is introduced into the objective function. The optimization problem is variational problem.

So far, the information of a single scan is used only. In most cases, however, radiologist acquire more than one weighting for a body region at a time to gain further contrast information about anatomical structures. We presented the first approach that uses the properties of all acquired images jointly (e.g., T1- and T2-weighted images). The image properties are stored in multi-dimensional joint histograms. In order to normalize the intensities of a newly acquired data set, a non-rigid registration is performed between a reference and the target joint histogram. From this matching a non-parametric transformation is obtained, which describes a mapping between the corresponding intensity spaces and subsequently adapts the image properties of the newly acquired series to a given standard. Furthermore, we introduced an adaption to this method that is able to deal with images having a large FOV, whole body scans for instance. To our knowledge, this is the only approach that can handle this kind of data.

The presented standardization methods are a reliable way to adjust pdfs of single as well as multiple series of MRI scans to a pulse sequence dependent standard signal intensity scale. The experiments show that the proposed approaches outperform all existing state-of-the-art methods in terms of generalizability, robustness against pathologies and structural changes, as well as accuracy in the standardization results. For instance, the proposed non-rigid 1-d histogram matching achieves an average intensity overlap of the major tissue classes of T1w images taken from the public available IBSR database of about 86.2%. Nyúl's method, on the other hand, resulted in only 70.1% overlap. However, if the image properties are too distinct between the source and the target, the obtained results may still not be satisfying.

Chapter 5

Computer-aided Assessment of Anomalies in the Scoliotic Spine

In this section, we present a real clinical application of the normalization methods introduced in section 3 and 4. Using the recently introduced 3-d SPACE sequence, the fast acquisition of the complete spine is possible. However, the data sets suffer from large intra- as well as inter-scan intensity variations. Without the correction of these artifacts in the data sets, no reliable automatic processing of the acquired images is achievable.

The majority of the proposed framework for computer-aided assessment (CAA) of the scoliotic spine that is introduced in this chapter has already been presented at the conference “Medical Image Computing and Computer-Assisted Intervention - MICCAI 2009” [Jage09b]. We show that it is possible to statistically model the spinal channel and cord. Using this model we perform a segmentation of the spinal structures. Upon this we build an application that enables the physician to assess the scoliotic spine nearly as fast and precisely as a non-scoliotic spine which should improve the pre-operative work-up of this young patient group.

5.1 Motivation

MRI is being used increasingly to investigate children with scoliosis. Although there may be a hereditary component to true idiopathic scoliosis, the condition has no known cause and is not associated with dysraphism. However, in the infantile and juvenile age group the incidence of spinal cord anomalies like tethered cord, syringomyelia, Chiari malformations, diastematomyelia and meningocele / myelomeningocele ranges from 17.6 to 26% [Lewo92, Gupt98]. Furthermore there can be structural changes of the vertebral bodies like wedge vertebra or hemivertebra. As MRI can visualize all these abnormalities it can be extremely important in the pre-operative planning of scoliosis. Failure to detect abnormalities of the neuraxis prior to treatment of scoliosis, particularly with instrumentation that lengthens the spine, can have serious neurological consequences.

With the introduction of 3-d spin echo sequences (3-d SPACE, Siemens, Erlangen, Germany), compare also section 2.1.3, MRI of the scoliotic spine can be acquired with only two sequences (upper spine and lower spine). However, due to the extreme

bending of the vertebral column in all three axes the manual assessment of the spine is a very time consuming process. In some cases it is even impossible for the radiologist to analyze pathological changes within the spine manually. The reason for this is that an orthogonal view with respect to the spinal column is required on every vertebra. Furthermore, it can be very difficult to specify the anatomic localization of the viewed vertebra. A pin-pointed surgery planing, however, is absolutely essential due to the severity of the procedure. In Figure 5.1 pre- and post-operative images of scoliosis patients are shown. In addition two examples of orthogonal slices are given in the figure.

5.2 State-of-the-art

Most state-of-the-art methods for the localization of the spine in tomographic images do a segmentation of the vertebrae (e.g. [Peng 06]). However, these approaches usually use assumptions about the spinal appearance that are not fulfilled in data sets of scoliosis patients. Particularly the shape of the scoliotic spine is altered considerably. Thus, all assumptions concerning the typical "s" shape and with that the orientation of the vertebrae are no longer valid. Additionally, the shape of the vertebrae can vary in a wider range than in the non-scoliotic case. Two typical examples of spinal images of scoliosis patients are shown in Figure 5.1. Recently, Stern *et al.* presented a promising method to automatically determine the centers of the vertebrae and the centerline [Ster 10]. Their approach relies on geometrical properties of spinal structures as well as on their characteristics in CT and MR images. Additionally, there are a few methods for segmentation of the spinal cord. McIntosh and Hamarneh use 3-d "deformable organisms" that complement classical deformable models with high-level mechanisms [McIn 06]. Furthermore, Coulon *et al.* propose a method that is based on the optimization of a B-spline active surface [Coul 02]. Theoretically, most of the approaches that are used for the segmentation of tubular structures can be altered for the segmentation of the spinal cord or at least for parts of it [Jage 05, Feti 02].

The computation of the centerline presented here does not depend on any prior information about the shape of the spine nor on the shape of the vertebra. Further on, it is not restricted to the used MRI protocol but can be adapted to other modalities in a straight forward manner. The only requirement is that the spinal channel or the spinal cord are visible within the images.

The system for the CAA of spine anomalies can be separated into four parts: pre-processing of the data sets, the segmentation of the spinal channel/cord, the labeling of the vertebra and finally the visualization of the data.

5.3 Pre-processing

In order to cover a wide range of cases and to make the segmentation stage more robust against the influences of inhomogeneities and noise within the data we use the following state-of-the-art pre-processing methods. In order to be able to directly use signal intensities additional to structural components within the images, we apply the

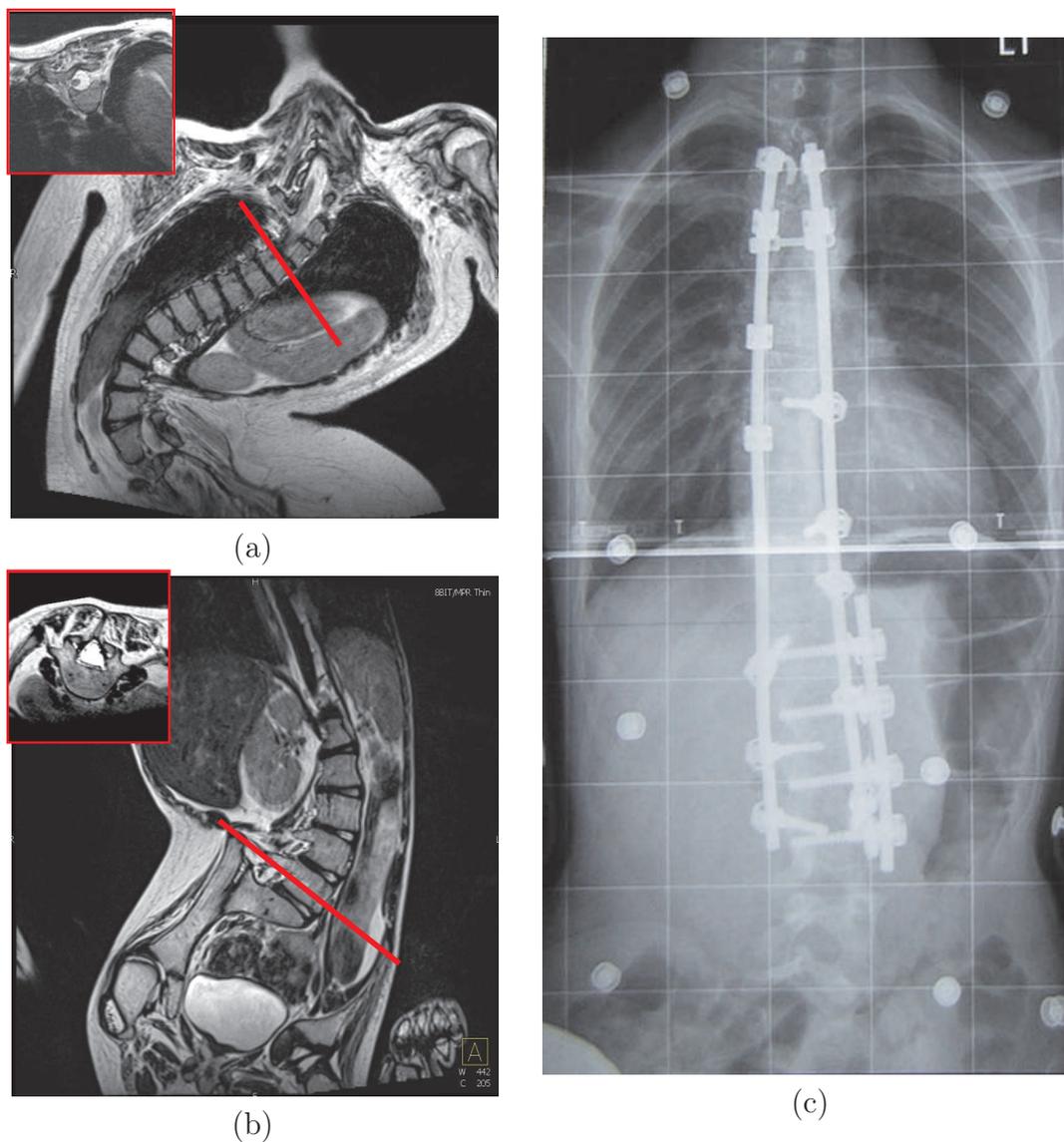
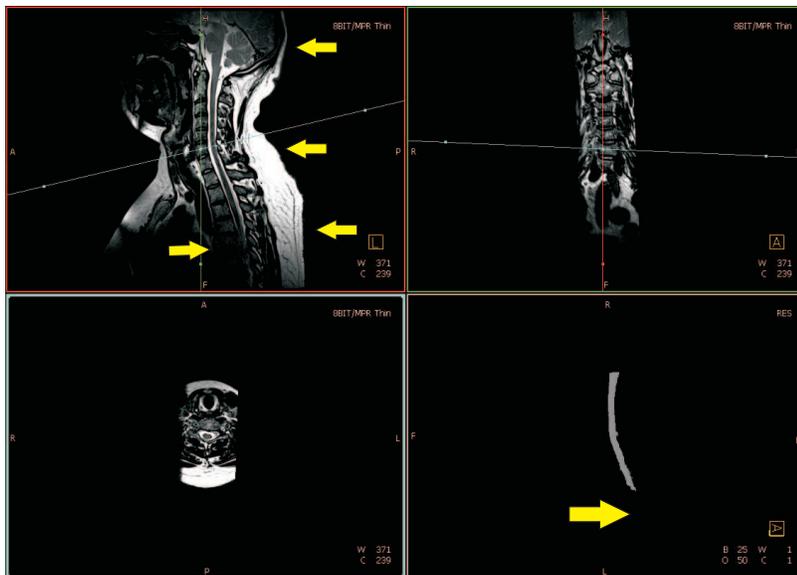
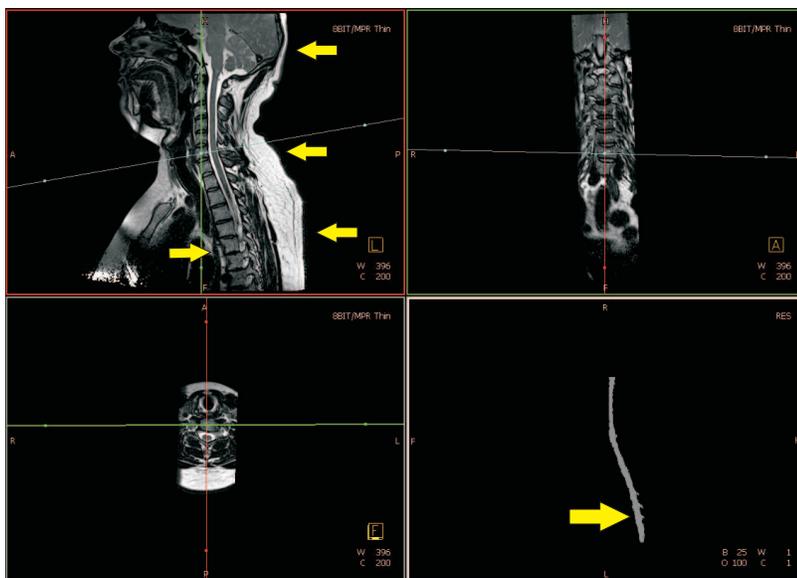


Figure 5.1: Pre- and post-operative images from scoliosis patients: Images (a) and (b) show two coronar slices from pre-operative 3-d SPACE MRI images of two different patients acquired at the Radiologic Institute of the University of Erlangen. In the upper left of both images a slice orthogonal to the spinal column is shown. The slices correspond to the red lines. Image (c) is an anterior-posterior X-ray of a case of adolescent idiopathic scoliosis post-fusion (source: Wikipedia, http://en.wikipedia.org/wiki/File:Wiki_post-op.jpg). There was originally a thoracic curve of 30° and a lumbar curve of 53° .



(a)



(b)

Figure 5.2: Effect of the normalization on the segmentation: Fig. (a) shows a segmentation result without normalization. In (b) the proposed pre-processing was used. It can be seen that using the normalization steps has a significant influence on the quality of the segmentation result. The arrows in both images mark regions of interest where the differences between the input data and between the segmentations can be seen best.

proposed divide-and-conquer based inhomogeneity correction method. The degree of the used polynomial is six. This compensates the influence of coil inhomogeneities during the acquisition of the 3-d MRI images. Subsequently, we use a signal intensity standardization approach to correct inter-scan intensity variations within the data sets. The method used is based on the 1-d non-rigid alignment of image histograms like introduced in section 4.3.2. The size of the histograms was 128 bins. Finally, a median filter is applied to the data sets to reduce noise present in the images. Even with a kernel size of 3×3 , the noise reduction is sufficient for further processing stages. Furthermore, the edge preserving property of the median filter is a necessity for the sequential segmentation. All these methods can be applied during the segmentation step on a per voxel basis. Thus, only voxels that are required for the segmentation of the spinal channel/cord are processed. This yields a reduced computational cost. The difference of the segmentation performance with and without proper pre-processing is illustrated in Fig. 5.2.

5.4 Segmentation using MRFs

The proposed method for the segmentation of the spinal channel and cord is an iterative process. The basic idea is that in each iteration step the segmentation is propagated with respect to the minimization of an energy function. This function is based on local signal intensities as well as local structural information like image gradients and the distance to the current approximation of the spinal centerline. The segmentation is formulated in the context of Markov random field (MRF) theory. Thereby the spatial coherence can be modeled statistically by posterior probabilities and easily be formulated as Gibbs distributions. Moreover, because of the normalization of the data sets, all required intensity distributions can be estimated beforehand. Furthermore, due to the initialization step of the method, a gradient descent strategy can be derived for optimization that can be implemented efficiently.

5.4.1 Initialization

For the initialization of the presented method the radiologist has to set a single seed point within the spinal channel. From this seed point an adaptive region growing is started. The upper and lower intensity boundaries of the segmentation are increased/decreased by one in each growing step until a predefined number of voxels N_0 is contained within the segmented region. Experiments show that a value of $N_0 = 300$ is sufficient. Afterwards we apply morphologic closing to the initial segmentation \mathbf{S}_0 to fill potential holes due to signal intensity variations within the images. All segmentations \mathbf{S}_i with $i > 0$ are binary images with a value of one for voxels within the segmented region and zero as background value.

5.4.2 Iteration Step

In every iteration i the centerline c_i is approximated first using the segmented region \mathbf{S}_{i-1} from the previous iteration. The estimation of c_i is performed by thinning the segmentation \mathbf{S}_{i-1} using the method presented by Lee in [Lee94] and a sequential

polynomial least-squares approximation of the x -, y - and z - components of the skeleton voxels. In general a polynomial degree of four to six is sufficient.

The second phase in every iteration i is the minimization of the energy function U given the previous segmentation \mathcal{S}_{i-1} and the parametric centerline c_i . The objective function is derived from MRF theory [Li 01]. In the following, we will briefly introduce all necessary definitions that are required for the usage of MRFs for segmentation issues.

Markov Random Field Theory

Let $\Psi = \{1, 2, \dots, N_\Psi\}$ be a discrete set of N_Ψ sites. Here, the sites correspond to pixel locations. A *label* or *state* is an event that may happen to a site and may be both continuous or discrete. In our case, two discrete label sets are defined: one for the intensity values denoted as \mathcal{Y} and one for the class memberships denoted as \mathcal{Z} . Note that a label in \mathcal{Y} assumes its value in the range $[0, 2^{d_p} - 1]$ and a state in \mathcal{Z} assumes its value in the range $\{0, \dots, C\}$, with d_p being the pixel depth and $C + 1$ the number of classes. As the spinal cord/channel are treated as a single class $C = 1$.

Next the neighborhood system and cliques are introduced, which build the basis for the contextual constraints. In analogy to most authors, we define a neighborhood for Ψ as

$$\mathcal{N} = \{\mathcal{N}_i, i \in \Psi\},$$

where \mathcal{N}_i is the set of site i 's neighbors (see e.g. [Li 01]). The neighborhood system has the following properties:

1. a site is not a neighbor of itself: $i \notin \mathcal{N}_i$
2. the neighboring relationship is mutual; thus, it is symmetric:
 $i \in \mathcal{N}_j \Leftrightarrow j \in \mathcal{N}_i$

A clique for (Ψ, \mathcal{N}) is defined as a subset of sites in Ψ , in which all the pairs of distinct sites are neighbors, except for single-site cliques. Here, pairwise interaction between pixels is considered only. Every clique therefore consists of one or two sites.

Furthermore, let $\mathbf{Z} = \{Z_1, Z_2, \dots, Z_N\}$ be a set of random variables defined on the set Ψ , in which each random variable Z_i takes a value $z_i \in \mathcal{Z}$. Ξ is the set of all possible configurations of \mathbf{Z} . \mathbf{Z} is said to be a Markov random field (MRF) on Ψ with respect to a neighborhood system \mathcal{N} if and only if the following two conditions are satisfied [Li 01]:

1. positivity: $P(\mathbf{Z}) > 0, \forall \mathbf{Z} \in \Xi$
2. Markovianity: $P(z_i | \mathbf{Z}_{\Psi - \{i\}}) = P(z_i | \mathbf{Z}_{\mathcal{N}_i})$ where $\mathbf{Z}_{\Psi - \{i\}}$ denotes the set of labels at the sites $\Psi - \{i\}$ and $\mathbf{Z}_{\mathcal{N}_i}$ stands for the set of labels of i 's neighbors.

Markovianity depicts the local characteristics of the random field. This yields that a label at a site i depends on its neighboring pixels only. In other words, a site i has direct interaction with its neighbors only. It is possible to always satisfy Markovianity because a sufficiently large neighborhood \mathcal{N}_i can be selected, where the largest neighborhood consists of all other sites.

According to the Hammersley-Clifford theorem [Hamm 71] a MRF is equivalent to a Gibbs random field (GRF). The theorem states that X is called MRF on Ψ with respect to \mathcal{N} if and only if \mathbf{Z} is a GRF on Ψ with respect to \mathcal{N} . The practical value of the theorem is that it provides a simple way to specify the joint probability $P(\mathbf{Z})$, $\mathbf{Z} \in \Xi$ by specifying the clique potential functions $V_c(\mathbf{Z})$ [Li 01].

Here, a configuration \mathbf{Z} corresponds to a segmentation \mathbf{S}_j at iteration j and z_i corresponds to the state $s_{\mathbf{x}}$ of a voxel \mathbf{x} . Thus, the optimal solution of the segmentation problem is defined by the maximum of the Gibbs distribution

$$P(\mathbf{S}_i) = Q^{-1} \exp(-U(\mathbf{S}_i)) \quad (5.1)$$

with Q being a normalization constant, $U(\mathbf{S}_i) = \sum_{\mathbf{x}} V(s_{\mathbf{x}}|\mathbf{S}_i)$ being the objective energy function and $s_{\mathbf{x}}$ being the state of the voxel \mathbf{x} .

In the case of the segmentation of the spinal channel/cord, we are using a 26-connected 3-d neighborhood. Moreover, there are two different states: occupied ($s_{\mathbf{x}} = 1$) if the voxel is part of the spinal channel/cord and free ($s_{\mathbf{x}} = 0$) otherwise. Using \mathbf{S}_{i-1} as initialization, we assume that we are within the area of attraction of the correct minimum [Feti 02]. For this reason a local gradient descent strategy can be used for optimization: \mathbf{S}_i is set to \mathbf{S}_{i-1} initially. All voxel that neighbor the segmentation are put into an open voxel set \mathcal{O} . Then for each voxel $\mathbf{x} \in \mathcal{O}$ the energy for the occupied state e_1 and the energy for the free state e_0 is computed. If $e_1 < e_0$ then \mathbf{x} is included into the segmented region \mathbf{S}_i . Its free neighbors are put into \mathcal{O} . The site \mathbf{x} itself is removed from \mathcal{O} . This is repeated until no more voxels change from state free to occupied. The method is summarized in Algorithm 3. First this is done for the segmentation of the spinal channel. This segmentation is then used as initialization for the segmentation of the spinal cord.

Clique Potentials

The potential $V(s_{\mathbf{x}}|\mathbf{S}_i)$ is composed by the following four parts. The first part of the potential is called smoothness prior as it controls the homogeneity of the segmentation result. It can be formulated as

$$V_s(s_{\mathbf{x}}|\mathbf{S}_i) = 1.0 - \frac{1}{\#\mathcal{N}_{\mathbf{x}}} \sum_{\{\mathbf{x}' \in \mathcal{N}_{\mathbf{x}} | s_{\mathbf{x}'} = s_{\mathbf{x}}\}} 1 \quad (5.2)$$

with $\mathcal{N}_{\mathbf{x}}$ being the neighborhood of \mathbf{x} and $\#\cdot$ the cardinality. This means that $V_s(s_{\mathbf{x}}|\mathbf{S}_i)$ is zero if all voxels in the neighborhood of \mathbf{x} have the state $s_{\mathbf{x}}$. If all neighboring voxels have a different state than $s_{\mathbf{x}}$ the potential is one.

The second component of the potential uses knowledge about the intensity range of the spinal channel and cord. The intensities are modeled by normal densities $N(\mu_c, \sigma_c^2)$ and $N(\mu_o, \sigma_o^2)$ with μ_c and σ_c being the parameters of the spinal channel and μ_o , σ_o of the cord, respectively. Thus, the resulting intensity potential V_v of a voxel \mathbf{x} can be formulated as

$$V_v(s_{\mathbf{x}}|\mathbf{S}_i) = (-1)^{s_{\mathbf{x}}+1} (|u(\mathbf{x}) - \mu_{\{c,o\}}| - 2\sigma_{\{c,o\}}) / (2\sigma_{\{c,o\}}) \quad (5.3)$$

where u is the MRI volume of the spine and $\{c, o\}$ means “c” or “o”. For $s_{\mathbf{x}} = 1$ the potential has the minimum -1 if $u(\mathbf{x}) = \mu_{\{c,o\}}$, it is zero if $u(\mathbf{x}) = \mu_{\{c,o\}} \pm 2\sigma_{\{c,o\}}$



Figure 5.3: Spinal segmentation example: In the image a slice from the original volume is overlaid with the segmentation result in red. In order to close holes the segmentation was dilated.

and positive if the signal intensity differs more than $2\sigma_{\{c,o\}}$ from $\mu_{\{c,o\}}$. If the voxel's state is $s_{\mathbf{x}} = 0$, the maximum is 1 and falls down linearly to $-\infty$. Consequently, voxels having an intensity between $\mu_{\{c,o\}} \pm 2\sigma_{\{c,o\}}$ are preferred to be within the segmentation; voxels with intensities that differ more than $2\sigma_{\{c,o\}}$ from the mean $\mu_{\{c,o\}}$ tend to belong to the background.

The third part of the potential utilizes the relative position of the voxel to the current centerline estimation c_i in iteration i . With $d_c(\mathbf{x})$ being the Euclidean distance from the voxel \mathbf{x} to the centerline c_i the potential V_d can be written as

$$V_d(s_{\mathbf{x}}|\mathbf{S}_i) = (-1)^{s_{\mathbf{x}}+1}(d_c(\mathbf{x}) - r_c)/r_c \quad (5.4)$$

where $r_c/2$ is the average radius of the spinal channel/cord. As a result $V_d(s_{\mathbf{x}} = 1|\mathbf{S}_i) < V_d(s_{\mathbf{x}} = 0|\mathbf{S}_i)$ for voxels that are closer than $r_c/2$ to the approximated center line c_i and $V_d(s_{\mathbf{x}} = 1|\mathbf{S}_i) > V_d(s_{\mathbf{x}} = 0|\mathbf{S}_i)$ if the distance is larger. The minimal value of V_d for the occupied state is -1; the maximal value of the potential for $s_{\mathbf{x}} = 0$ is 1.

Finally, the last part of the potential uses the scalar product between the propagation direction of the segmentation and the gradient of the image intensities. As propagation direction we use the gradient of the segmentation image \mathbf{S}_i at the voxel \mathbf{x} . It can be computed by finite differences, for instance. Thus, the potential can be defined as

$$V_g(s_{\mathbf{x}}|\mathbf{S}_i) = (-1)^{s_{\mathbf{x}}}(1 - |(\nabla u(\mathbf{x}))^T \nabla \mathbf{S}_i(\mathbf{x})|/g_{\max}) \quad (5.5)$$

where g_{\max} is the maximal tolerable magnitude of the gradient. If the gradients are aligned parallel or anti-parallel or if the image gradient is zero, the potential for $s_{\mathbf{x}} = 1$ has its minimal value -1 and its maximal value 1 for $s_{\mathbf{x}} = 0$.

Using the aforementioned clique potentials, the objective function can be written as the minimization problem

$$\mathbf{S}^* = \arg \min_{\mathbf{S}} \sum_{\mathbf{x} \in \Psi} V_s(s_{\mathbf{x}}|\mathbf{S}) + V_v(s_{\mathbf{x}}|\mathbf{S}) + V_d(s_{\mathbf{x}}|\mathbf{S}) + V_g(s_{\mathbf{x}}|\mathbf{S}) \quad (5.6)$$

Algorithm 3 Summary of the spinal cord segmentation

Require: u , seed point \mathbf{p}_s
Ensure: Segmentation of the spinal cord \mathbf{S}

- 1: do intensity inhomogeneity correction (see section 3)
 - 2: do intensity standardization (see section 4)
 - 3: do median filtering of u (kernel size 3×3 or 5×5)
 - 4: compute initial segmentation \mathbf{S}_0
 - 5: $i = 0$
 - 6: **repeat**
 - 7: $i = i + 1$
 - 8: $\mathbf{S}_i = \mathbf{S}_{i-1}$
 - 9: compute centerline c_i from \mathbf{S}_i
 - 10: compute open voxel set \mathcal{O} from \mathbf{S}_i
 - 11: **repeat**
 - 12: take voxel \mathbf{x} from \mathcal{O}
 - 13: remove voxel \mathbf{x} from \mathcal{O}
 - 14: $e_0 = V_s(s_{\mathbf{x}} = 0 | \mathbf{S}_i) + V_v(s_{\mathbf{x}} = 0 | \mathbf{S}_i) + V_d(s_{\mathbf{x}} = 0 | \mathbf{S}_i) + V_g(s_{\mathbf{x}} = 0 | \mathbf{S}_i)$
 - 15: $e_1 = V_s(s_{\mathbf{x}} = 1 | \mathbf{S}_i) + V_v(s_{\mathbf{x}} = 1 | \mathbf{S}_i) + V_d(s_{\mathbf{x}} = 1 | \mathbf{S}_i) + V_g(s_{\mathbf{x}} = 1 | \mathbf{S}_i)$
 - 16: **if** $e_0 > e_1$ **then**
 - 17: set $s_{\mathbf{x}} = 1$ in \mathbf{S}_i
 - 18: insert $\mathbf{x}' \in \mathcal{N}_{\mathbf{x}}$ with $s_{\mathbf{x}'} = 0$ into \mathcal{O}
 - 19: **end if**
 - 20: **until** $\mathcal{O} == \emptyset$
 - 21: **until** $\mathbf{S}_{i-1} == \mathbf{S}_i$
 - 22: **return** \mathbf{S}_i
-

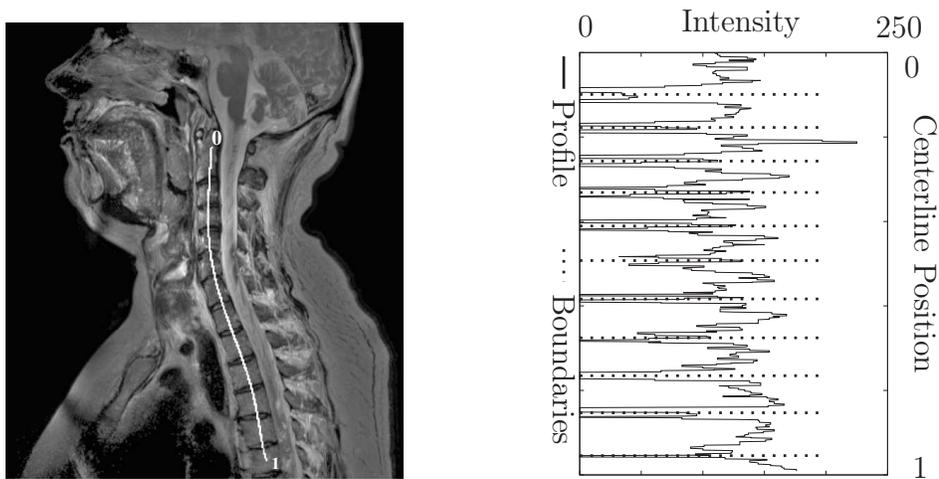


Figure 5.4: Centerline through vertebra bodies: The left image shows a 12 mm thick MIP in the spinal region including the computed centerline through the vertebra (white line). The data set shown is pre-processed. This line is used to compute the boundary positions of the vertebra. In the righthand plot the signal intensities along the centerline are shown. Additionally the estimated boundaries of the vertebrae are illustrated (dotted lines). The computed threshold was $\theta = 70.3$.

where \mathbf{S}^* is the optimal segmentation of the image with domain Ψ . The proposed gradient descent approach for segmenting the spinal channel/cord is summarized in Algorithm 3.

The result of a segmentation is shown in Fig. 5.3. The segmented pixel are marked in red. The segmentation is a little bit larger than the spinal cord. The reason for this is that after the segmentation the result is dilated to close possible holes. This is done because we are interested in the centerline only.

5.5 Labeling of the Vertebrae and Visualization

In order to label the vertebrae within the images, we compute an intensity profile p on the ventral side of the estimated centerline. Then we apply a threshold θ to the computed profile p . From this an initial guess about the positions of the vertebrae is computed. Finally, this guess is refined using the average distances between the vertebrae. The profile as well as the estimated boundaries of the vertebrae are illustrated in Figure 5.4. As there is no slice where the whole line through the vertebrae can be seen we use a 12 mm thick Maximum Intensity Projection (MIP) to be able to cover the whole spine in the illustration. The original slice thickness was 1 mm.

For the visualization, the computed centerline is approximated by splines. Using the parametric approximation we can compute MPRs that are orthogonal to the backbone for every position of the spinal channel/cord. An illustration of the presentation of the MPRs is shown in Figure 5.6.

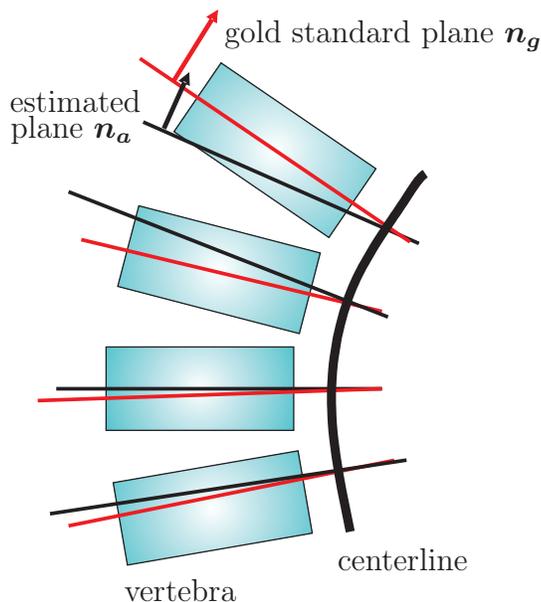


Figure 5.5: Evaluation principle of the automatic centerline segmentation: The illustration shows a schematic view onto the spinal column. The colored rectangles represent the vertebra and the curved thick black line the estimated centerline. The black lines through the vertebrae correspond to the estimated planes. The red one to the gold standard. The arrows represent the normals to the planes.

5.6 Experiments & Results

Data Sets: All data sets were acquired during clinical routine. In total we used 20 3-d SPACE data sets from the spine including ten volumes showing the upper spine and ten data sets covering the lower spine. All images were acquired with a repetition time of $TR = 1000\text{ms}$ and an echo time of $TE = 130\text{ms}$. The volumes had an isotropic in-plane resolution between $0.8\text{ mm} \times 0.8\text{ mm}$ and $1.3\text{ mm} \times 1.3\text{ mm}$ and a slice thickness of 1 mm . The image matrix had a size of 384×384 . Every scan consists of 60 up to 160 slices. All used data sets are publicly available at our homepage¹.

Evaluation Method: The whole processing chain was implemented in C++ and integrated into the ITK Framework (<http://www.itk.org>). For a better presentation of the results and to increase the usability for radiologists everything was integrated into the medical visualization platform InSpace3D. The experiments were performed on a 2.00 GHz Intel Core2 CPU with 2 GB RAM. The whole processing chain took about 5-20s depending on the size and the bending of the backbone.

The focus of this work is an easy-to-use framework for CAA of anomalies in the scoliotic spine. For this reason it is important that radiologists have an orthogonal view onto every vertebra. Thus, we use the following quality measure for evaluation. First, for every vertebra v within the images, a medical expert defines a ground truth plane with normal \mathbf{n}_g^v . Then, the corresponding planes with normal \mathbf{n}_a^v are

¹<http://www5.informatik.uni-erlangen.de/~spine/>

computed using the proposed segmentation method. In order to measure the distance between the corresponding planes we use the angle

$$d_{n,v} = \arccos |(\mathbf{n}_g^v)^T \cdot \mathbf{n}_a^v| \quad (5.7)$$

between the normal vectors. The range of $d_{n,v}$ is $[0^\circ, \dots, 90^\circ]$. If both planes are aligned perfectly parallel or anti-parallel the angle between the normal vectors is $d_{n,v} = 0$ degree. If, on the other hand, the corresponding planes are orthogonal, $d_{n,v} = 90$. The quality q_n of the proposed segmentation method is computed by

$$q_n = \frac{1}{V} \sum_{i=1}^V d_{n,i} \quad (5.8)$$

with V being the number of ground truth planes. It reflects the mean angular deviation of the ground truth to the automatically computed planes. The evaluation principle is illustrated in Fig. 5.5.

Evaluation Results: In total planes through $V = 181$ different vertebra were defined by the radiologist. Using the proposed quality measure for our segmentation method yielded a mean distance of $q_n = 5.65^\circ$ with a standard deviation of $\sigma_{q_n} = 6.09^\circ$. The minimal deviation was 0.50° and the maximal deviation was 25.73° degrees (lower spine: $\mu_l = 7.80^\circ$, $\sigma_l = 8.40$, $\min_l = 0.69^\circ$, $\max_l = 25.73^\circ$, upper spine: $\mu_u = 4.57^\circ$, $\sigma_u = 4.34$, $\min_u = 0.50^\circ$, $\max_u = 13.21^\circ$). If it is assumed that an average vertebra has a size of about $30 \times 30 \times 20 \text{mm}^3$ this means that there is a distance between the two planes of less than 2mm at the border of an average vertebra.

Additionally, a second radiologist defined $V = 61$ planes through vertebrae that had already by labeled by the first radiologist. These planes were used to compute the inter-observer variability of both radiologists. This resulted in a mean angular deviation of $q_o = 2.94^\circ$ ($\sigma_{q_o} = 1.99^\circ$, $\min = 0.48^\circ$, $\max = 9.38^\circ$).

The results show that the proposed algorithm works very reliably for the upper spine. Especially in the lower lumbar area and the pelvic region the results get slightly worse. The reason for this is that the medulla ends in this region and separate nerve cords are left. Thus, there is a higher probability that the segmentation follows these cords away from the backbone. In clinical routine this is not a big problem, as these regions are irrelevant for diagnostics in general.

5.7 Summary & Discussion

We presented a novel approach for the segmentation of the spinal cord based on MRF theory. The segmentation is used to compute planes orthogonal to the vertebra column for CAA of anomalies in the scoliotic spine. The advantage of our method is that we do not use any segmentation of the vertebrae itself or information about their relative positioning. Thus, even an extreme bending of the spine or pathologic changes of the vertebra structure can be compensated easily. Further on, the method presented works on 3-d volumes and is not restricted to a good visible coverage of the spine in a single slice. Additionally, no training step is required. Thus, the method can easily adapt if the acquisition protocol changes or other modalities like CT are used.



Figure 5.6: Visualization of the spinal cord: The figures show the presentation of the computed MPRs. The upper two images in each figure show the sagittal plane and the coronal plane respectively. The second row in each figure shows the plane orthogonal to the spinal cord and the orthogonal plane in a 3-d view.

The proposed framework enables the radiologist to intuitively assess anomalies in the scoliotic spine. The errors in orientation observed are small enough for clinical usage. Furthermore, the majority of false centerline estimations occurs in the pelvic region of the spine that is only of little diagnostic interest. The observed errors can further be reduced by post-processing like a rough segmentation of the vertebra using the approximated centerline to improve their pose estimation.

Chapter 6

Outlook

A convenient way of improving the quality of both, the correction of signal intensity inhomogeneities, and the intensity standardization, is to include prior knowledge into the objective function. Usually this information is generated from large amounts of reference or training data what can cause severe problems. In the following, we will present two sketches of possible approaches, one for inhomogeneity correction and one for intensity standardization that make use of prior knowledge.

Active Bias Field Model

Segmentation methods often use statistical shape models of the data to increase the accuracy of the segmentation results. One of the most commonly used methods for representing shape in segmentation are Active Shape Models (ASM) first introduced by Cootes *et al.* [Coot95]. The knowledge about the shape is stored in the model by computing the mean shape and the principal components of shape deviation. This can be done by a Principle Component Analysis (PCA). Thus, the principal shape components can be selected concerning the degree of variation they represent in the training set.

The training is the most challenging part using the proposed technique. As mentioned in section 3.4.2, it is hard to get ground truth data of intensity inhomogeneities. A possibility to generate a gold standard is to manually segment the data sets. Due to the close connection between the estimation of inhomogeneities and segmentation, afterwards, gold standard data can be computed. For the generation of the model, all training data sets have to be in the same coordinate system. However, the alignment does not depend on the anatomy of the imaged object but on the geometry of the surrounding coil system. Following the alignment of the data, the creation of the model itself corresponds to the model generation known from segmentation literature.

If a statistical bias field model is available, it can be used for the inhomogeneity correction. There are two ways how to use this kind of prior information. First, the model can be utilize directly for correction. In this case, starting with the mean bias model, the inhomogeneities are estimated by iteratively minimizing the image entropy using the principal components of the statistical model as decent directions.

Second, the model can be used as an additional regularization term in the optimization process. This strategy allows a more flexible and adjustable handling of the influence of the model in the optimization. The integration into the approaches

described in section 3 can be done easily. For LEMS or LEHR, the nodes can be adapted regarding the bias field model after every iteration. In the case of DaC, the model can be applied after the proper estimation. The influence can be weighted to allow subject dependent differences.

Multiple Mean References Standardization

There are two different possibilities to integrate additional information into the standardization techniques introduced in section 4: Multiple references and the creation of a mean reference. The methods we presented before make use of a single reference only independent of its characteristics. The term “Multiple references” means that there is a reference available for every principal shape of the histogram; a reference for histograms with two modes and a reference for histograms containing three modes for instance. A mean reference, on the other hand, is inverting the principle of multiple references by creating a mean model over all training samples. This has the advantage that the average distance between an observed histogram and the reference is minimized. Both principles can be combined by computing several mean histograms covering a certain shape range. A potential multiple mean references standardization algorithm is composed of a subsequent training step and an intensity standardization step.

From our point of view, there are two possible methods to identify a set of distinct references. First, the user predefines a set of characteristics that separate the histograms in several disjunct classes. As mentioned before, possible criteria for a class are the number of modes, the relative sizes of the modes, the position of the modes relative to histogram percentiles, etc. Second, automatically computed features on the histograms or between histograms like their distance can be used. When the features are computed on the training data, a clustering algorithm has to be applied on the feature sets. The resulting clusters define the histogram classes. After identifying the sets of histograms, for every set a reference has to be computed. This can be done similarly to shape models by computing the deformation field between every histogram within one class. The reference is then created using these deformation fields.

The standardization process is straight forward. In a first step, the reference is selected. This is done by computing either the characteristics of the histogram or by selecting the reference with the closest feature set. Once the reference histogram is selected the intensity standardization is performed using a method proposed in section 4.

Chapter 7

Summary

Without any doubt medical imaging has revolutionized modern health care by offering means to display the morphology and the metabolism of the human body. Especially MRI gained more and more importance in the last decade due to its excellent soft tissue contrast and new innovative acquisition sequences that allow functional imaging (fMRI, ASL, etc.) besides the classic morphological imaging. However, these techniques suffer from artifacts that have a significant impact on quality and image interpretation. Among others signal intensity variations are the most common artifacts in MRI. They can be differentiated in two distinct categories: First, signal intensity inhomogeneities that describe intensity variations resulting from a gain or bias field within a single data set. The methods for correcting these are named bias correction. Second, inter-image signal intensity variations where intensities vary in between different scans. This artifact is corrected using intensity standardization approaches. Because the intensity variations affect directly the appearance of anatomical structures in the images and thus their statistics, they have a large impact on the performance of post-processing methods like image registration and segmentation.

The basis of MRI is the discovery of nuclear magnetic resonance by F. Bloch and E. M. Purcell in 1946. It describes the property that nuclei have in a magnetic field. If they are put into a strong magnetic field, the spin of nuclei start to precess around the axis of the external field. The precession or Larmor frequency ω_0 depends on the field strength and the nuclei. For 1H ω_0 is 42.6 MHz at one Tesla field strength. By applying a RF pulse, the sample's magnetic vector is flipped away from the parallel alignment. After the pulse is turned off, the magnetization begins to realign with the direction of the outer magnetic field. This induces a current in the receiver coils. The decay of the signal is characterized by tissue dependent time constants T_1 and T_2 . The introduction of field gradients by P. C. Lauterbur and P. Mansfield made spatial encoding feasible. By using three types of gradients for slice selection, frequency encoding, and phase encoding 3-d imaging is possible. The specification of the succession of RF pulses and the application of field gradients is denoted as pulse sequence. The most prominent pulse sequence is called spin echo. The acquired raw data is stored in k -space. It is a frequency domain that can be mapped to the image domain by a Fourier transform. As it is often rather time consuming to fill the entire k -space, parallel acquisition techniques were developed. Theoretically, the speed up is equivalent to the number of parallel acquisitions. In order to create the image, two

different composition strategies are used: Methods working in spatial domain like SENSE and methods working in frequency domain, GRAPPA for instance.

The signal intensity inhomogeneities observed in MR images are usually smooth and slowly varying. Thus, they consist of low frequencies only. In literature, these variations are described by two different models, an additive and a multiplicative one, in general. The selection of an appropriate model is very application dependent. As bias fields have a huge impact on a subsequent processing of the data, a lot of research has been done in the last two decades. State-of-the-art methods can be split into prospective and retrospective methods. In clinical environment, however, retrospective methods are much more feasible. This class of methods can be categorized into filtering approaches, like HUM, surface fitting based methods, segmentation-based approaches, e.g. MFCM, and histogram-based approaches, N3 for instance. Most of these lack either generalizability, precision, and/or runtime efficiency.

In order to decrease the computational costs, we introduced a divide-and-conquer based method (DaC). First, the problem domain is split into small sub-regions. It is assumed that the bias fields in these small regions can be approximated by a very simple model. The criteria for optimality is the Shannon entropy. Due to the simplicity of the local models, the non-linear optimization can be done very efficiently using a gradient descent strategy. Afterwards, a global model is computed from the local ones using a least-squares fit in the conquer phase of the approach. This model is used to correct the distorted image.

Salvado *et al.* propose a method that estimates a bi-cubic spline model by iteratively minimizing the image entropy. The method was designed for the correction of MR images for atherosclerosis characterization. To further increase the generalizability and accuracy of the approach, we presented an extension that regularizes the optimization process (LEHR). The regularization is based on the distance of the image histogram to a reference one. This reference can either be approximated from the original histogram or a gold standard histogram can be used. Thus, prior knowledge about the shape of the histograms can be included into the objective function.

The evaluation of both methods is done on synthetic images first to show the principal pros and cons. DaC is used to correct 2-d as well as 3-d images whereas LEHR is applied to 2-d data only due to its computational costs. The experiments show that DaC is superior to all tested methods because of its ability to estimate even very complicated bias fields. Regarding LEHR, it turned out that the choice of the reference is a crucial point in the estimation. If the reference histogram cannot be approximated very well, the results are rather bad. On the other hand, if a reliable approximation is possible, the results are very good. Further on, DaC is applied to simulated 3-d head images. The images are affected by tunable inhomogeneities and noise level. The results are compared to several other methods. The only approach that yielded comparable good results is M4 introduced by Likar *et al.* If possible the experiments using real clinical relevant data are evaluated using a gold standard segmentation. If that kind of data is not available then a qualitative evaluation is performed. The results of these experiments expressively underline the received results from synthetic data sets. It turned out that DaC corrected all images very reliable and that it is robust against pathological changes. For instance, the experiments show that DaC increases the separability of tissue classes in 3-d TOF MRA images

by approximately 18.2% whereas state-of-the-art methods could only achieve 11.6%. Furthermore, if the bias fields are weak LEHR provides satisfactory results. On the other hand, if the inhomogeneities are rather strong, and thus the required reference cannot be approximated very well, the correction quality is low.

Compared to the correction of intensity inhomogeneities, signal intensity standardization did not receive the same amount of attention. However, especially segmentation approaches suffer from a missing standard intensity scale. In the context of clinical applications, the intensity mapping that is computed by the standardization methods has to be invertible. Moreover, it has to keep the structure of the histograms to preserve the medical information. For this reason many state-of-the-art image processing approaches like histogram equalization or histogram specification are not suitable. The most commonly used approach was presented by Nyúl *et al.* Their method relies on the accurate determination of pre-defined landmarks on the reference as well as the target histogram. These are then mapped onto each other. The histogram positions in between are mapped linearly. The exact determination of the landmarks is the biggest drawback of the method as this fails in many cases.

To overcome the drawbacks of the state-of-the-art methods, we presented a couple of methods that are closely related to image registration approaches. In order to increase the performance of a sequent non-rigid alignment of the histograms, a 1-d histogram matching consisting of an affine intensity mapping is presented. Due to two unknowns only, the optimization can be done very efficiently. The used distance measure is JD. The standardization results using the affine mapping are rather weak, but it is very well suited as initialization for more complex methods as large intensity transformations are removed. In order to take local intensity deviations into account, a non-rigid 1-d histogram matching is introduced. Its transformation is non-parametric and it is computed on a per-intensity basis. Because of the affine pre-alignment, the estimation of the transformation can be done in a very fast manner.

So far, all known methods compute the intensity standardization for a single image at a time. However, usually in MRI more than one sequence is used for acquisition, a T1w and a T2w for instance. We presented the first method that uses the complete information jointly for standardization. First, the intensity information of all sequences is stored in a joint histogram. Thus, the dimensionality of the histogram corresponds to the number of input images. Next, the joint histograms are matched. If they are considered to be images, the intensity standardization problem can be mapped to a non-rigid image registration. The gained non-parametric transformation is used to alter the intensities in the target images.

A problem of all methods for intensity standardization is a large FOV. The reason for this is that the intensity statistics are very complex for these data sets and the data sets are often corrupted by severe bias fields. Thus, local intensity characteristics are ignored during the standardization process. To improve the correction of these data sets, first we split the data sets into small sub-regions. Then, these sub-regions are corrected individually. Consequently, there are intensity jumps at the border between regions. In order to reduce these, a novel distance measure is introduced that regularizes the estimation of the transformations by considering adjacent regions.

Similar to the evaluation of bias correction methods, the experiments are split into synthetic ones showing the basic properties of the methods and experiments on clinical relevant data. Whenever possible Nyúl's method is used as a reference. In the synthetic experiments the intensities of real data sets are artificially perturbed. Then, the perturbed image is standardized to the original data set. We could show that the proposed non-rigid methods outperform Nyúl's method in all the experiments. Only if the artificial perturbation is small then Nyúl's method yields better results in some cases. Moreover, we could illustrate that if there are tissue dependent (not intensity dependent!) intensity variations, the joint intensity standardization results are much more reasonable than for the independent case. For the experiments using real clinical data, either images from public available databases or images acquired in daily clinical routine are used. These experiments confirm the results gained from the synthetic data. In most cases the non-rigid alignment outperforms all other methods. Especially, using the T1w images of the public available IBSR database, Nyúl's method has considerable problems whereas the non-rigid alignment does a reliable and robust standardization. Here, Nyúl's method achieved an average intensity overlap of only 70.1% compared to an overlap of about 86.2% using the proposed standardization technique.

With the introduction of the 3-d spin echo sequence SPACE, a fast acquisition of the complete spine is possible. The assessment of anomalies in the spine is an essential task during the planning phase, due to the severeness of the correction procedure of the scoliotic spine. Because of the pathologic bending of the spine, the assessment is an extremely time consuming process as an orthogonal view w.r.t. the spinal column onto every vertebra is required. However, SPACE images showing spinal structures suffer from severe intra- as well as inter-scan intensity variations. Without a proper correction of these artifacts no fast reliable automatic processing of the images is achievable. We introduced a system for CAA of anomalies in the spine relying on CPRs. All necessary steps from the pre-processing of the data including intensity normalization to the visualization component were presented. As the core part of the framework is based on a segmentation of the spinal cord we focused on this. The proposed segmentation method is an iterative process. In every iteration the segmentation is updated by an energy based scheme derived from MRF theory. As some of the used potentials are based on image intensities and gradients, the pre-processing has to include a normalization of the data sets. For this, the previously introduced methods are used. Further, a noise reduction is done that is based on median filtering. From the segmentation of the spinal cord the centerline of the backbone can be approximated. The visualization of the data is done by computing planes that are orthogonal to the subsequently approximated centerline. These planes are shown to the user.

We evaluated the segmentation results on public available clinical relevant 3-d MRI data sets of scoliosis patients. In order to assess the quality of the segmentation we used the angle between automatically computed planes through the vertebra and planes estimated by medical experts. This results in a mean angular difference of less than six degrees. Moreover, we did a qualitative evaluation of the method. The results look very reasonable. The experiments show that the angular deviations of

the estimated planes through the vertebrae are small enough to be applicable in the clinical environment.

Although most of the normalization results including intensity inhomogeneity correction and intensity standardization using the proposed methods are very accurate, in some cases the normalization fails. Often, the reason for this is the changing image content and thus the appearance of the histogram, additional modes for instance. In order to deal with this more prior knowledge has to be included into the optimization process. A possible way of doing so is the introduction of a statistical model of the intensity inhomogeneities observed for a predefined body region and receiver coil setup. This could be used to regularize the optimization process. Moreover, the introduction of more than one reference into the standardization process would positively influence the results.

In conclusion, we created several methods for a reliable normalization of MRI data sets that outperform most state-of-the-art methods. As requested the approaches are applicable in a real clinical setting due to their generalizability, computational complexity, robustness against pathological changes, and their precision. These properties were successfully illustrated in a clinically relevant application that improves the daily workflow.

Appendix A

Mathematical Symbols

List of all mathematical symbols in order of appearance in the text.

M	magnetization of a sample
B_0	main magnetic field
ω_0	angular velocity corresponding to B_0
γ	gyromagnetic constant
$x, y, \text{ and } z$	coordinates
$x', y', \text{ and } z'$	coordinates in the rotating frame
M_t	transversal magnetization
T_1	longitudinal relaxation time
T_2^*, T_2	spin-spin relaxation time
t	time
M_0	magnetization before excitation
$M_z(t)$	magnetization at t
G_z, G_x	gradient field
$B_z(z)$	z -component at position z
$B_x(x)$	x -component at position x
sinc	sinc-fucntion
$\delta\omega$	difference in angular velocity
G_{PE}	phase encoding gradient
TE	echo time
TR	repetition time
M_{\parallel}	parallel magnetization
FOV _x , FOV _y	field of view
d_s	slice thickness
N_{RO}	number of samples during readout
N_{PE}	number of phase encoding steps
s_v	physical size of a voxel
k_x, k_y	coordinates in k -space
$\Delta k_x, \Delta k_y$	sampling step in k -space
\mathcal{FT}	Fourier transform
t_{ac}	acquisition time
r	acceleration factor
ρ	spin density

C_i	sensitivity of coil i
S_i	signal perceived at coil i
L	number of coils
$n_c(\cdot)$	weighting coefficients of the coils
N_b	number of blocks (GRAPPA)
$\tilde{\rho}$	estimated object density
\mathbf{x}	voxel or pixel in image ($x \in \Omega$)
v	observed image
u	ideal image
b	bias field
$n(\cdot)$	noise component
Ω	image domain
D	image dimension
$\hat{\cdot}$	logarithmic version of \cdot
$\mathcal{N}_{\mathbf{x}}$	neighborhood of pixel \mathbf{x}
σ_n	standard deviation of noise component n
μ	mean/median value of data samples
$\text{LPF}\{\cdot\}$	low pass filter
c_i	class centers
$a_{i\mathbf{x}}$	fuzzy membership of voxel \mathbf{x} to class i
C	number of classes
α	weighting factor
p	weighting factor steering the fuzziness
$\mathcal{U}, \mathcal{V}, \mathcal{B}, \mathcal{H}$	histogram of image u, v, b , and h
$E[\cdot]$	expected value
$\#\cdot$	cardinality of \cdot
$\mathbf{a} = (a_0, \dots, a_n)$	weights/coefficients
q_i	Legendre polynomials
s_i	basis function
n_e	neutral element
d	degree of a polynomial
$H(\cdot)$	Shannon entropy
$p(\cdot)$	probability density function
G_σ	Gaussian kernel with standard deviation σ
R	number of regions
\mathcal{R}_i	sub-region
$S(\cdot)$	regularization term
b_i	local bias field
β_i	weighting of a local bias field
\mathcal{B}_r	set of local bias field values of region r
ν_i^r	bias value of region r
\mathcal{X}_r	set of global coordinates of region r
Λ	set of nodes
κ_i	single node
K_n	number of nodes
s_n	spacing between the nodes

O_E	objective function of LEMS
ϵ	threshold
α_{HP}	weight of highpass result in the thinning approach
\mathcal{K}_σ	Gaussian highpass filter kernel with standard deviation σ
ζ	frequency
O_{EH}	objective function of LEHR
η	scaling factor for the histogram difference measure
ΔH	difference in entropy
ΔD	difference of histogram measures
$r_{\mathbf{x}}$	ratio between estimated and ideal bias field
q	quality of a correction result
χ, ξ	intensity
cv	coefficient of variation
cjv	coefficient of joint variation
Γ	intensity domain
φ, ϑ	intensity mapping
ϕ	intensity shift or intensity deformation
s_A, t_A	scaling and translation of the affine model
\mathcal{D}, D	similarity measure
\mathcal{S}	smoother
\mathcal{I}	variational problem of 1-d histogram matching
\mathcal{L}	function depending on the intensities and the intensity shift
τ	artificial time component
$\Delta\tau$	discrete time step
$\Delta\chi$	spacing between intensities
\mathcal{A}	partial differential operator
\mathbf{A}	discrete version of \mathcal{A}
Φ	discrete version of ϕ
n_h	number of histogram bins
f	force term
\mathbf{f}	vector describing the discrete version of f
\mathbf{I}	identity matrix
M	number of used levels
m	current level
n	number of images for the joint standardization
\mathbf{u}, \mathbf{v}	sets containing the n reference and target images
\mathcal{U}, \mathcal{V}	joint histograms of the sets \mathbf{u} and \mathbf{v}
\mathcal{J}	variational problem of n-d histogram matching
K	number of partitions
k	current partition
$a_{j,k}$	weight of partition k while standardizing partition j
$S_{l,m}$	spline for the correction of an image v_m at l
$\text{rand}()$	random number generator with uniform distribution
$\text{g}_{\text{rand}}(\sigma)$	random number generator with Gaussian distribution
d_O	overlap of two relative histograms
c	amplitude

p_i	percentile
\mathbf{p}	set of percentiles p_i
\mathbf{L}	set of landmarks
l_i	landmark
Υ	mapping of a voxel
g_a	factor controlling the magnitude of a mapping
γ_a	intensity mapping
ψ	influence of neighboring partitions on the standardization
r_{JH}	ratio between distances of joint histograms after and before standardization
d_{JD}	Jeffrey divergence of joint histograms
N_0	minimal number of voxels in a segmented region
\mathbf{S}_i	segmentation in iteration i
c_i	centerline in iteration i
U	energy function
Ψ	set of sites
N_{Ψ}	number of sites
\mathcal{Y}, \mathcal{Z}	discrete label sets
d_p	pixel depth
Z_i	random variable
\mathbf{Z}	set of random variables
z_i	label or state
V_c	clique potential
Q	normalization constant of a Gibbs distribution
$s_{\mathbf{x}}$	state of voxel \mathbf{x}
e_0, e_1	energy of state 0 and state 1
$N(\mu, \sigma)$	normal distribution with mean μ and standard deviation σ
d_c	Euclidean distance from centerline
r_c	average radius of spinal cord/channel
g_{max}	maximal gradient magnitude
θ	threshold
$\mathbf{n}_g^v, \mathbf{n}_a^v$	normal to a gold standard and to the estimated plane through vertebra v
d_n	angle between normal vectors
q_n	mean angular deviation

Appendix B

Derivative of the Jeffrey Divergence

In the following, we deduce the derivative of the JD

$$\mathcal{D}_{\text{JD}}(\chi) = \int_{\Gamma} p(\chi) \log \frac{2p(\chi)}{p(\chi) + q(\chi - \phi(\chi))} + q(\chi - \phi(\chi)) \log \frac{2q(\chi - \phi(\chi))}{p(\chi) + q(\chi - \phi(\chi))} d\chi.$$

As we use the derivative to compute the Euler differential equation, only the inner part has to be differentiated with respect to ϕ .

$$\frac{\partial}{\partial \phi} \mathcal{D}_{\text{JD}}(\chi) = \frac{\partial}{\partial \phi} \left[\underbrace{p(\chi) \log \frac{2p(\chi)}{p(\chi) + q(\chi - \phi(\chi))}}_{D_p(\chi)} + \underbrace{q(\chi - \phi(\chi)) \log \frac{2q(\chi - \phi(\chi))}{p(\chi) + q(\chi - \phi(\chi))}}_{D_q(\chi)} \right] \quad (\text{B.1})$$

The derivative of the first part of Eq. (B.1):

$$\begin{aligned} \frac{\partial}{\partial \phi} D_p(\chi) &= \frac{\partial}{\partial \phi} \left[p(\chi) \log \left(2p(\chi) \right) - p(\chi) \log \left(p(\chi) + q(\chi - \phi(\chi)) \right) \right] \\ &= -p(\chi) \cdot \frac{1}{p(\chi) + q(\chi - \phi(\chi))} \cdot q'(\chi - \phi(\chi)) \cdot (-1) \\ &= p(\chi) \frac{q'(\chi - \phi(\chi))}{p(\chi) + q(\chi - \phi(\chi))} \end{aligned} \quad (\text{B.2})$$

The derivative of the second part of Eq. (B.1):

$$\begin{aligned} \frac{\partial}{\partial \phi} D_q(\chi) &= \frac{\partial}{\partial \phi} \left[\underbrace{q(\chi - \phi(\chi)) \log \left(2q(\chi - \phi(\chi)) \right)}_{D_{q1}(\chi)} \right] \\ &\quad - \frac{\partial}{\partial \phi} \left[\underbrace{q(\chi - \phi(\chi)) \log \left(p(\chi) + q(\chi - \phi(\chi)) \right)}_{D_{q2}(\chi)} \right] \end{aligned} \quad (\text{B.3})$$

The derivative of the first part of Eq. (B.3):

$$\begin{aligned}
\frac{\partial}{\partial \phi} D_{q_1}(\chi) &= -q'(\chi - \phi(\chi)) \log \left(2q(\chi - \phi(\chi)) \right) \\
&+ q(\chi - \phi(\chi)) \cdot \frac{1}{2q(\chi - \phi(\chi))} \cdot q'(\chi - \phi(\chi)) \cdot (-1) \\
&= -q'(\chi - \phi(\chi)) \left(\frac{1}{2} + \log \left(2q(\chi - \phi(\chi)) \right) \right)
\end{aligned} \tag{B.4}$$

The derivative of the second part of Eq. (B.3):

$$\begin{aligned}
\frac{\partial}{\partial \phi} D_{q_2}(\chi) &= -q'(\chi - \phi(\chi)) \log \left(p(\chi) + q(\chi - \phi(\chi)) \right) \\
&+ q(\chi - \phi(\chi)) \cdot \frac{1}{p(\chi) + q(\chi - \phi(\chi))} \cdot q'(\chi - \phi(\chi)) \cdot (-1) \\
&= -q'(\chi - \phi(\chi)) \left[\log \left(p(\chi) + q(\chi - \phi(\chi)) \right) + \frac{q(\chi - \phi(\chi))}{p(\chi) + q(\chi - \phi(\chi))} \right]
\end{aligned} \tag{B.5}$$

Plugging together Eq. (B.2), Eq. (B.4) and Eq. (B.5) yields

$$\begin{aligned}
\frac{\partial}{\partial \phi} D_{\text{JD}}(\chi) &= \frac{\partial}{\partial \phi} \left[D_p(\chi) + D_{q_1}(\chi) - D_{q_2}(\chi) \right] \\
&= q'(\chi - \phi(\chi)) \frac{p(\chi)}{p(\chi) + q(\chi - \phi(\chi))} \\
&- q'(\chi - \phi(\chi)) \left[\frac{1}{2} + \log \left(2q(\chi - \phi(\chi)) \right) \right] \\
&+ q'(\chi - \phi(\chi)) \left[\log \left(p(\chi) + q(\chi - \phi(\chi)) \right) + \frac{q(\chi - \phi(\chi))}{p(\chi) + q(\chi - \phi(\chi))} \right] \\
&= q'(\chi - \phi(\chi)) \left(\frac{1}{2} + \log \frac{p(\chi) + q(\chi - \phi(\chi))}{2q(\chi - \phi(\chi))} \right)
\end{aligned} \tag{B.6}$$

List of Figures

1.1	Difference between intra- and inter-image intensity variations	3
2.1	Precession of the spins	10
2.2	Illustration of the slice selection	12
2.3	Illustration of phase encoding	13
2.4	Spin echo timing diagram	14
2.5	T1w and T2w image examples	16
2.6	Turbo spin echo timing diagram	16
2.7	Field-of-view	17
2.8	k -space / image domain correspondence	18
2.9	Schematic illustration of SENSE	20
2.10	Schematic illustration of AUTO SMASH and GRAPPA	21
2.11	Parallel acquisition and sum-of-squares reconstruction	23
2.12	Motion and wrap around artefact	24
2.13	Chemical shift artifacts and magnetic susceptibility	26
2.14	Intensity non-uniformities	27
3.1	Example for log-transformed histograms	31
3.2	Workflow of the divide-and-conquer based bias correction	37
3.3	Approximation of the additive inhomogeneity model	40
3.4	Approximation of the multiplicative inhomogeneity model	42
3.5	Synthetic sinusoidal shaped inhomogeneities	47
3.6	Correction of synthetic checkerboard images	48
3.7	Slice through the synthetic volume	50
3.8	Estimated bias fields of synthetic 3-d data	51
3.9	Ground truth segmentation and simulated T2w brain data	52
3.10	Real MRI data sets with gold standard segmentation	54
3.11	K-means clustering of TOF images	57
3.12	Checkerboard results for LEMS and LEHR	59
3.13	3-d TOF results for entropy and histogram optimization	61
3.14	3-d SPACE results for LEMS and LEHR	63
3.15	Photograph of a slice through the human eyeball	64
3.16	Bias field correction results for microslice photographs	66
3.17	Bias correction of retina fundus images	67
4.1	Schematic illustration of the intensity standardization	80
4.2	Illustration of the interpolation of the transformation vectors	83

4.3	Images used for the synthetic 1-d standardization experiments	86
4.4	Results of the synthetic 1-D standardization experiments	88
4.5	Histogram overlap distance measure	91
4.6	1-d standardization results on IBSR T1w data sets	92
4.7	Example slices from images taken from Tumorbase	93
4.8	Standardization results of the Tumorbase images	94
4.9	Histograms of T1w images before and after standardization	96
4.10	N-d non-rigid alignment results using head images	97
4.11	N-d standardization results using an artificial perturbation	101
4.12	Example slices used for experiments using artificial head images . . .	102
4.13	Resulting joint histograms correction an artificial intensity perturbation	103
4.14	N-d standardization results using an artificial intensity transformation	104
4.15	Artificial growth of ventricles	106
4.16	Artificial growth of a lesion	107
4.17	Results for n -D non-rigid histogram standardization of an artificially growing lesion	108
4.18	Results for n -D non-rigid histogram standardization of artificially grow- ing ventricles	109
4.19	Whole-body MRI intensity standardization plots	110
4.20	Whole-body MRI intensity standardization results	112
5.1	Pre- and post-operative images from scoliosis patients	117
5.2	Effect of the normalization on the segmentation	118
5.3	Spinal segmentation example	122
5.4	Centerline through vertebra bodies	124
5.5	Evaluation principle of the automatic centerline segmentation	125
5.6	Visualization of the spinal cord	127

List of Tables

3.1	DaC-inhomogeneity correction results (simulated MR brain images) .	53
3.2	DaC-inhomogeneity correction results (clinical relevant MR brain images)	56
4.1	1-d non-rigid standardization results using artificial perturbations . .	90
4.2	N-d non-rigid standardization results of MRI head images	98

Bibliography

- [Ahme 02] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty. “A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data”. *IEEE Transactions on Medical Imaging*, Vol. 21, No. 3, pp. 193–199, 2002.
- [Axel 87] L. Axel, J. Constantini, and J. Listerud. “Intensity correction in surface coil MR imaging”. *American Journal of Roentgenology*, Vol. 148, No. 2, pp. 418–420, 1987.
- [Bald 06] M. Balda. *Evaluation of Bias Field Correction Methods for MRI and US*. Master’s thesis, Friedrich-Alexander Universität Erlangen Nürnberg, June 2006.
- [Berg 08] J.-P. Bergeest and F. Jäger. “A Comparison of Five Methods for Signal Intensity Standardization in MRI”. In: T. Tolxdorff, J. Braun, T. M. Deserno, H. Handels, A. Horsch, and H.-P. Meinzer, Eds., *Bildverarbeitung für die Medizin*, pp. 36–40, Berlin, 2008.
- [Brin 98] B. H. Brinkmann, A. Manduca, and R. A. Robb. “Optimized Homomorphic Unsharp Masking for MR Grayscale Inhomogeneity Correction”. *IEEE Transactions on Medical Imaging*, Vol. 17, No. 2, pp. 161–171, April 1998.
- [Brow 03] M. A. Brown and R. C. Semelka. *MRI: Basic Principles and Applications*. Wiley-Liss, New Jersey, 3rd Ed., 2003.
- [Buxt 02] R. B. Buxton. *An Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques*. Cambridge Univ Press, Cambridge, 2002.
- [Chio 03] J.-Y. Chiou, C. B. Ahn, L. T. Muftuler, and O. Nalcioglu. “A Simple Simultaneous Geometric and Intensity Correction Method for Echo-Planar Imaging by EPI-Based Phase Modulation”. *IEEE Transactions on Medical Imaging*, Vol. 22, No. 2, pp. 200–205, February 2003.
- [Cohe 00] M. S. Cohen, R. M. DuBois, and M. M. Zeineh. “Rapid and Effective Correction of RF Inhomogeneity for High Field Magnetic Resonance Imaging”. *Human Brain Mapping*, Vol. 10, No. 4, pp. 204–211, 2000.
- [Coll 02] G. Collewet, A. Davenel, C. Toussaint, and A. S. “Correction of Intensity Nonuniformity in Spin-Echo T1-weighted Images”. *Magnetic Resonance Imaging*, Vol. 20, No. 4, pp. 365–373, May 2002.
- [Coll 98] D. Collins, A. Zijdenbos, V. Kollokian, J. Sled, N. Kabani, C. Holmes, and A. Evans. “Design and Construction of a Realistic Digital Brain Phantom”. *IEEE Transactions on Medical Imaging*, Vol. 17, No. 3, pp. 463–468, June 1998.

- [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, January 1995.
- [Coul 02] O. Coulon, S. J. Hickman, G. J. Barker, D. H. Miller, and S. R. Arridge. "Quantification of Spinal Cord Atrophy From Magnetic Resonance Images Via a B-Spline Active Surface Model". *Magnetic Resonance in Medicine*, Vol. 47, No. 6, pp. 1176–1185, 2002.
- [Dawa 93] B. M. Dawant, A. P. Zijdenbos, and R. A. Margolin. "Correction of Intensity Variations in MR Images for Computer-aided Tissue Classification". *IEEE Transactions on Medical Imaging*, Vol. 12, No. 4, pp. 770–781, December 1993.
- [Dows 06] N. Dowson, R. Bowden, and T. Kadir. "Image template matching using Mutual Information and NP-Windows". In: *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*, pp. 1186–1191, IEEE Computer Society, Washington, DC, USA, 2006.
- [Fan 03] A. Fan, M. Wells, M. Fisher, M. Cetin, S. Haker, R. Mulkern, C. Tempany, and A. S. Willsky. "A Unified Variational Approach to Denoising and Bias Correction in MR". *Information Processing in Medical Imaging*, Vol. 18, pp. 148–159, 2003.
- [Feti 02] C. L. Fetita and F. J. Prêteux. "Quantitative 3D CT bronchography.". In: *IEEE International Symposium on Biomedical Imaging*, pp. 221–224, IEEE, Washington, DC, USA, 2002.
- [Fill 09] A. Filler. "MR Neurography and Diffusion Tensor Imaging: Origins, History & Clinical Impact". *Neurosurgery*, Vol. 65, No. 4 Suppl, pp. 29–43, 2009.
- [Ge 00] Y. Ge, J. K. Udupa, L. G. Nyúl, L. Wei, and R. I. Grossman. "Numerical Tissue Characterization in MS via Standardization of the MR Image Intensity Scale". *Journal of Magnetic Resonance Imaging*, Vol. 12, No. 5, pp. 715–721, 2000.
- [Gonz 02] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice-Hall, Inc., New Jersey, USA, 2nd Ed., 2002.
- [Gris 02] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase. "Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA)". *Magnetic Resonance in Medicine*, Vol. 47, No. 6, pp. 1202–1210, June 2002.
- [Gupt 98] P. Gupta, L. G. Lenke, and K. H. Bridwell. "Incidence of neural axis abnormalities in infantile and juvenile patients with spinal deformity. Is a magnetic resonance screening necessary?". *Spine*, Vol. 23, No. 2, pp. 206–210, 1998.
- [Habe 04] E. Haber and J. Modersitzki. "Numerical methods for volume preserving image registration". *Inverse Problems*, Vol. 20, No. 5, pp. 1621–1638, October 2004.
- [Hahn 05] D. A. Hahn, J. Hornegger, W. Bautz, T. Kuwert, and W. Römer. "Unbiased rigid registration using transfer functions". In: *Proc. of SPIE on Medical Imaging*, pp. 151–162, San Diego, 2005.

- [Hahn 09] D. A. Hahn, V. Daum, and J. Hornegger. “Automatic Parameter Selection for Multi-Modal Image Registration”. *IEEE Transactions on Medical Imaging*, p. accepted for publication, 2009.
- [Hamm 71] J. Hammersley and P. Clifford. “Markov Fields on Finite Graphs and Lattices”. 1971. unpublished.
- [Hell 03] P. Hellier. “Consistent Intensity Correction of MR images”. In: *International Conference on Image Processing (ICIP 2003)*, pp. 1109–12, Barcelona, September 2003.
- [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, March 2001.
- [Hoov 03] A. Hoover and M. Goldbaum. “Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels”. *IEEE Transactions on Medical Imaging*, Vol. 22, No. 8, pp. 951–958, August 2003.
- [Jage 05] F. Jäger, J. Hornegger, and E. G. Hahn. “Formbasierte Segmentierung des Bronchialbaumes”. In: H.-P. Meinzer, H. Handels, A. Horsch, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2005*, pp. 108–112, Springer Berlin Heidelberg New York, Heidelberg, 2005.
- [Jage 06a] F. Jäger, Y. Deuerling-Zheng, B. Frericks, F. Wacker, and J. Hornegger. “A new Method for MRI Intensity Standardization with Application to Lesion Detection in the Brain”. In: L. Kobbelt, T. Kuhlen, T. Aach, and R. Westermann, Eds., *Vision Modeling and Visualization 2006*, pp. 296–276, Aka GmbH, Berlin, 2006.
- [Jage 06b] F. Jäger, J. Han, J. Hornegger, and T. Kuwert. “A Variational Approach to Spatially Dependent Non-Rigid Registration”. In: J. M. Reinhardt and J. P. W. Pluim, Eds., *Proc. of SPIE on Medical Imaging*, pp. 860–869, San Diego, February 2006.
- [Jage 07] F. Jäger, L. G. Nyúl, B. Frericks, F. Wacker, and J. Hornegger. “Whole Body MRI Intensity Standardization”. In: A. Horsch, T. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2007*, pp. 459–463, Springer, Munich, 2007.
- [Jage 08] F. Jäger, M. Balda, and J. Hornegger. “Correction of Intensity Inhomogeneities Utilizing Histogram-based Regularization”. In: R. Bauernschmitt, Y. Chaplygin, H. Feušner, Y. Gulyaev, J. Hornegger, E. Mayr, N. Navab, S. Schookin, S. Selishchev, and S. Umnyashkin, Eds., *Proceedings of the 4th Russian-Bavarian Conference on Biomedical Engineering at Moscow Institute of Electronic Technology*, pp. 23–27, Moscow, 2008.
- [Jage 09a] F. Jäger and J. Hornegger. “Nonrigid registration of joint histograms for intensity standardization in magnetic resonance imaging”. *IEEE Transactions on Medical Imaging*, Vol. 28, No. 1, pp. 137–150, 2009.
- [Jage 09b] F. Jäger, J. Hornegger, S. Schwab, and R. Janka. “Computer-Aided Assessment of Anomalies in the Scoliotic Spine in 3-D MRI Images”. In: G.-Z. Yang, D. J. Hawkes, D. Rueckert, J. A. Noble, and C. J. Taylor, Eds., *MICCAI*, pp. 819–826, Springer, London, 2009.

- [Jako 98] P. M. Jakob, M. A. Grisowld, R. R. Edelman, and D. K. Sodickson. "AUTO-SMASH: A self-calibrating technique for SMASH imaging". *Magnetic Resonance Materials in Physics, Biology and Medicine*, Vol. 7, No. 1, pp. 42–54, November 1998.
- [Jank 08] R. Janka, M. Uder, M. Ritt, F. Jäger, P. Martirosian, R. E. Schmieder, and W. Bautz. "Perfusionsmessung der Niere, gemessen mit arteriellem Spin Labeling (ASL) im Vergleich zur traditionellen Clearance-Bestimmung - Effekte einer 2 wöchigen Therapie mit einem AT1-Rezeptorinhibitor". In: Deutsche Röntgengesellschaft e.V., Ed., *RöFo - Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*, Online, 2008.
- [Jank 09] R. Janka, M. Ritt, F. Jäger, P. Martirosian, R. Schmieder, and M. Uder. "Non-invasive Measurement of Kidney Perfusion in Comparison to Traditional Clearance Technique: Effects of a 2-week Therapy with an AT1-Receptor Inhibitor (Telmisartan)". In: *Radiological Society of North America (RSNA) 2009*, pp. SSA09–01, Radiological Society of North America (RSNA), Chicago, November 2009.
- [Lark 07] D. J. Larkman and R. G. Nunes. "Parallel magnetic resonance imaging". *Physics in Medicine and Biology*, Vol. 52, No. 7, pp. R15–R55, April 2007.
- [Lars 03] E. G. Larsson, E. D., Y. R., P. J. C., and F. J. R. "SNR-optimality of sum-of-squares reconstruction for phased-array magnetic resonance imaging". *Journal of Magnetic Resonance*, Vol. 163, No. 1, pp. 121–123, July 2003.
- [Lee 94] T.-C. Lee, R. L. Kashyap, and C.-N. Chu. "Building skeleton models via 3-D medial surface/axis thinning algorithms". *Computer Vision, Graphics, and Image Processing*, Vol. 56, No. 6, pp. 462–478, 1994.
- [Lewi 04] E. B. Lewis and N. C. Fox. "Correction of differential intensity inhomogeneity in longitudinal MR images". *Neuroimage*, Vol. 23, No. 1, pp. 75–83, 2004.
- [Lewo 92] K. Lewonowski, J. D. King, and M. D. Nelson. "Routine use of magnetic resonance imaging in idiopathic scoliosis patients less than eleven years of age". *Spine*, Vol. 17, No. Suppl. 6, pp. 109–116, 1992.
- [Li 01] S. Z. Li. *Markov Random Field Modeling in Image Analysis*. *Computer Science Workbench*, Springer, Tokyo, 2001.
- [Lika 01] B. Likar, M. A. Viergever, and F. Petruš. "Retrospective correction of MR intensity inhomogeneity by information minimization". *IEEE Transactions on Medical Imaging*, Vol. 20, No. 12, pp. 1398–1410, December 2001.
- [Lotj 10] J. Lötjönen, R. Wolz, J. Koikkalainen, L. Thurfjell, G. Waldemar, H. Soininen, and D. Rueckert. "Fast and robust multi-atlas segmentation of brain magnetic resonance images". *NeuroImage*, Vol. 49, No. 3, pp. 2352–2365, February 2010.
- [Mada 05] A. Madabhushi and J. K. Udupa. "Interplay Between Intensity Standardization and Inhomogeneity Correction in MR Image Processing". *IEEE Transactions on Medical Imaging*, Vol. 24, No. 5, pp. 561–576, May 2005.
- [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, March 1998.

- [Mart 04] P. Martirosian, U. Klose, I. Mader, and F. Schick. “FAIR true-FISP perfusion imaging of the kidneys”. *Magnetic Resonance in Medicine*, Vol. 51, No. 2, pp. 353–361, 2004.
- [McIn 06] C. McIntosh and G. Hamarneh. “Spinal Crawlers: Deformable Organisms for Spinal Cord Segmentation and Analysis”. In: *MICCAI*, pp. 808–815, Springer, Copenhagen, Denmark, 2006.
- [Mode 04] J. Modersitzki. *Numerical Methods for Image Registration*. Oxford University Press, Oxford New York, 2004.
- [Nara 88] P. A. Narayana, W. W. Brey, M. V. Kulkarni, and C. L. Sievenpiper. “Compensation for Surface Coil Sensitivity Variation in Magnetic Resonance Imaging”. *Magnetic Resonance Imaging*, Vol. 6, No. 3, pp. 271–274, 1988.
- [Ny100] L. G. Nyúl, J. K. Udupa, and X. Zhang. “New Variants of a Method of MRI Scale Standardization”. *IEEE Transactions on Medical Imaging*, Vol. 19, No. 2, pp. 143–150, February 2000.
- [Oppe 05] A. Oppelt. *Imaging systems for medical diagnostics: fundamentals, technical solutions and applications for systems applying ionizing radiation, nuclear magnetic resonance and ultrasound*. Publicis Corporate Publishing, Erlangen, 2005.
- [Peng 06] Z. Peng, J. Zhong, W. Wee, and J. Lee. “Automated Vertebra Detection and Segmentation from the Whole Spine MR Images”. In: *IEEE Engineering in Medicine and Biology*, pp. 2527–2530, IEEE, New York City, USA, 2006.
- [Pike 92] G. B. Pike, B. S. Hu, G. H. Glover, and D. R. Enzmann. “Magnetization transfer time-of-flight magnetic resonance angiography”. *Magnetic Resonance in Medicine*, Vol. 25, No. 2, pp. 372–379, 1992.
- [Prue 99] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. “SENSE: Sensitivity encoding for fast MRI”. *Magnetic Resonance in Medicine*, Vol. 42, No. 5, pp. 952–962, 1999.
- [Puse 86] E. Pusey, R. B. Lufkin, R. K. Brown, M. A. Solomon, D. D. Stark, R. W. Tarr, and W. N. Hanafee. “Magnetic resonance imaging artifacts: mechanism and clinical significance.”. *Radiographics*, Vol. 6, No. 5, pp. 891–911, Sep 1986.
- [Rohr 01] K. Rohr. *Landmark-Based Image Analysis Using Geometric and Intensity Models*. *Computational Imaging and Vision*, Kluwer Academic Publishers, Dordrecht, 2001.
- [Rubn 00] Y. Rubner, C. Tomasi, and L. J. Guibas. “The Earth Mover’s Distance as a Metric for Image Retrieval”. *International Journal of Computer Vision*, Vol. 40, No. 2, pp. 99–121, November 2000.
- [Ruec 99] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. “Nonrigid registration using free-form deformations: Application to breast MR images”. *IEEE Transactions on Medical Imaging*, Vol. 18, No. 8, pp. 712–721, August 1999.

- [Salv 06] O. Salvado, C. Hillenbrand, S. Zhang, and D. L. Wilson. “Method to Correct Intensity Inhomogeneity in MR Images for Atherosclerosis Characterization”. *IEEE Transactions on Medical Imaging*, Vol. 25, No. 5, pp. 539–552, May 2006.
- [Schm 05] M. Schmidt. “A Method for Standardizing MR Intensities between Slices and Volumes”. Tech. Rep. TR05-14, University of Alberta, 2005.
- [Simm 94] A. Simmons, P. S. Tofts, G. J. Barker, and S. R. Arridge. “Sources of Intensity Nonuniformity in Spin Echo Images at 1.5 T”. *Magnetic Resonance in Medicine*, Vol. 32, No. 1, pp. 121–128, 1994.
- [Sled 98] J. G. Sled, A. P. Zijdenbos, and A. C. Evans. “A Nonparametric Method for Automatic Correction of Intensity Nonuniformity in MRI Data”. *IEEE Transactions on Medical Imaging*, Vol. 17, No. 1, pp. 87–97, February 1998.
- [Sodi 97] D. K. Sodickson and W. J. Manning. “Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays”. *Magnetic Resonance in Medicine*, Vol. 38, No. 4, pp. 591–603, 1997.
- [Ster 10] D. Stern, B. Likar, F. Pernus, and T. Vrtovec. “Automated detection of spinal centrelines, vertebral bodies and intervertebral discs in CT and MR images of lumbar spine”. *Physics in Medicine and Biology*, Vol. 55, No. 1, pp. 247–264, January 2010.
- [Uder 09] M. Uder, M. Heinrich, F. Jäger, J. Hornegger, R. Schmieder, and R. Janka. “Einfluss neuer Techniken in der Bildgebung der Niere”. *Der Nephrologe*, Vol. 4, No. 1, pp. 26–32, 2009.
- [Van 99a] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. “Automated model-based bias field correction of MR images of the brain”. *IEEE Transactions on Medical Imaging*, Vol. 18, No. 10, pp. 885–896, October 1999.
- [Van 99b] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. “Automated model-based tissue classification of MR images of the brain”. *IEEE Transactions on Medical Imaging*, Vol. 18, No. 10, pp. 897–908, October 1999.
- [Vemu 05] P. Vemuri, E. G. Kholmovski, D. L. Parker, and B. E. Chapman. “Coil Sensitivity Estimation for Optimal SNR Reconstruction and Intensity Inhomogeneity Correction in Phased Array MR Imaging”. *Lecture Notes on Computer Science*, Vol. 3565, pp. 603–614, 2005.
- [Voku 99] E. A. Vokurka, N. A. Thacker, and A. Jackson. “A fast model independent method for automatic correction of intensity nonuniformity in MRI data”. *Journal on Magnetic Resonance Imaging*, Vol. 10, No. 4, pp. 550–562, October 1999.
- [Vovk 07] U. Vovk, F. Pernuš, and B. Likar. “A Review of Methods for Correction of intensity Inhomogeneity in MRI”. *IEEE Transactions on Medical Imaging*, Vol. 26, No. 3, pp. 405–421, March 2007.
- [Weic 98] J. Weickert, B. M. ter Haar Romeny, and M. A. Viergever. “Efficient and Reliable Schemes for Nonlinear Diffusion Filtering”. *IEEE Transactions on Image Processing*, Vol. 7, No. 3, pp. 398–410, March 1998.

- [Weis 04] N. Weisenfeld and S. Warfield. “Normalization of Joint Image-Intensity Statistics in MRI using the Kullback-Leibler Divergence”. In: *IEEE International Symposium on Biomedical Imaging*, pp. 101–104, Arlington, 2004.
- [Well 96] W. M. Wells III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz. “Adaptive Segmentation of MRI data”. *IEEE Transactions on Medical Imaging*, Vol. 15, No. 4, pp. 429–442, August 1996.
- [Yous 08] A.-H. Youssif, A. Ghalwash, and A. Ghoneim. “Optic Disc Detection From Normalized Digital Fundus Images by Means of a Vessels’ Direction Matched Filter”. *IEEE Transactions on Medical Imaging*, Vol. 27, No. 1, pp. 11–18, January 2008.
- [Zaid 07] H. Zaidi. “Is MR-guided Attenuation Correction a Viable Option for Dual-Modality PET/MR Imaging?”. *Radiology*, Vol. 244, No. 3, pp. 639–642, September 2007.
- [Zhan 01] Y. Zhang, M. Brady, and S. Smith. “Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm”. *IEEE Transactions on Medical Imaging*, Vol. 20, No. 1, pp. 45–57, January 2001.
- [Zhug 09] Y. Zhuge and J. K. Udupa. “Intensity Standardization simplifies brain MR image segmentation”. *Computer Vision and Image Understanding*, Vol. 113, No. 10, pp. 1095–1103, October 2009.

Index

- T_2^* decay, 11
- T_1 decay, 11
- T_1 -weighted, 15
- T_2 decay, 11
- T_2 -weighted, 15
- 1-d histogram matching, 71, 74, 81, 85

- Adaptive region growing, 119
- Additive bias model, 40
- Additive model, 29
- Aliasing, 25
- Angular deviation, 126
- Arterial Spin Labeling (ASL), 17
- Auto-calibration signals, 22

- Bias field, 29, 44
- Bloch equation, 9
- Boltzmann statistics, 10
- BrainWeb, 50

- Centerline potential, 122
- Chemical shift artifact, 25
- Clique potentials, 121
- Cliques, 120
- Coefficient of joint variations, 52
- Coefficient of variations, 52
- Coil sensitivities, 27

- Diffusion MRI, 2
- Diffusion tensor imaging, 2
- Diffusion weighted imaging, 2
- Directional potential, 122
- Divide-and-Conquer-based correction, 36

- Dynamic programming (DP), 79

- Echo time, 15
- Echo-train, 17
- Edge effects, 65
- Euler differential equation, 76
- Excitation, 11

- Field-of-view (FOV), 17
- Flip angle, 11
- Force term, 77
- Free induction decay, 11
- Frequency encoding, 12
- Functional MRI, 2

- Gain field, 29
- Gaussian highpass filter, 46
- General intensity scale, 4, 71
- Gibbs distribution, 119, 121
- Gibbs random field, 121
- Global model, 36, 38–41
- Gradient field, 11–13
- GRAPPA, 19, 21
- Gyromagnetic constant, 10

- Histogram equalization, 72
- Histogram specification, 73
- Histogram-based correction, 33
- Histogram-based regularization, 45
- Homomorphic unsharp masking, 32, 33

- Inhomogeneity correction, 26, 29
- Intensity equalization, 29
- Intensity inhomogeneities, 29

- Intensity potential, 121
- Intensity shift, 75, 76
- Intensity standardization, 71, 72, 81, 84
- Inter-image signal intensity variations, 3
- Intra-image signal intensity variations, 3, 29
- Inversion pulse, 17
- Jeffrey divergence, 46, 75
- Joint histogram, 80–82
- k-space, 18
- Kullback Leibler divergence, 75
- Label, 120
- Larmor frequency, 10
- LEHR, 46
- LEMS, 44
- Local model, 36, 39
- Logarithmic model, 30
- Magnetic susceptibility, 26
- Magnetization, 10
- Main magnetic field, 10
- Markov random field (MRF), 119
- Markov random fields, 120
- Markovianity, 120
- Mean-preservation, 31
- Motion artifact, 24
- MR angiography, 2
- Multi level, 78
- Multiplicative bias model, 40, 42
- Multiplicative model, 29
- N-d histogram matching, 71, 79
- N-d histogram matching (NAND), 95
- Neighborhood system, 120
- Neumann boundary condition, 78
- Neutral element, 36
- Noise, 29
- Non-rigid registration, 81
- Nonparametric nonuniformity normalization (N3), 33, 34
- Normalization, 4
- NP windows, 30, 38
- Nyúl's method, 74
- Parallel acquisition, 19
- Partitioning, 82
- Parzen estimation, 30, 38, 81
- Phase cancelation, 25
- Phase encoding, 13
- Prior knowledge, 44, 76
- Prospective correction, 32
- Pulse sequence, 13
- Quantification, 30
- Readout direction, 12
- Reference histogram, 45, 74
- Reference image, 72
- Refocussing RF pulse, 15
- Regularizer, 76
- Relaxation, 11
- Repetition time, 15
- Retrospective correction, 32
- RF inhomogeneities, 27
- RF pulse, 11
- Scoliotic spine, 115
- SENSE, 19
- Shannon entropy, 38, 45
- Sharpening, 46
- Signal-to-noise ratio, 30
- Site, 120
- Slab boundary artefact, 27, 71
- Slice selection, 11
- SMASH, 22
- Smoothness prior, 121

Source image, 72
SPACE, 17, 115
Spatial coherence, 39
Spin echo, 13
Spinal channel, 115
Spinal cord, 115
State, 120
Sum of squared differences, 75
Sum-of-squares reconstruction, 24

Target histogram, 74
Target image, 72
Time-of-flight imaging, 2
Transfer function, 4
Truncation artifact, 25
Turbo spin echo, 15

Vertebral column, 116

Warp-around artifact, 25
Whole body MRI, 71, 105

