MRI-Based High-Accuracy Morphometry in Consideration of Acquisition-Related Morphological Variability

Dissertation zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
am Department Informatik
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
der Universität Hamburg

vorgelegt von Arne Littmann

geboren am 02.04.1976 in Neustadt/Holstein

Hamburg 2008

Gutachter:	Prof. Dr.-Ing. Hans-Siegfried Stiehl Prof. Dr. med. Christian Büchel
Tag der Disputation:	07. November 2007

Acknowledgements

This thesis is the result of a PhD research project carried out in cooperation between the Department of Informatics at the University of Hamburg, and Siemens Medical Solutions, Magnetic Resonance. Both parties supported me excellently in the course of this thesis:

Above all, I would like to thank Prof. Dr.-Ing. H. S. Stiehl, Prof. Dr. C. Buechel and Dr. J. Guehring for their support while supervising this thesis. Their valuable remarks also significantly contributed to this dissertation.

I would also very much like to thank Siemens Medical Solutions, Magnetic Resonance, not only for direct financial support, but also for kindly providing access to its marketing and research scanning facilities in Erlangen.

Additionally, many thanks are going to the colleagues from the Advanced-Neuro- and the Physics-department within Siemens Magnetic Resonance for numerous fruitful discussions about the investigated pulse sequences and about hardware details of the scanners incorporated in the experimental study, especially to Dr. G. Krueger, Dr. T. Feiweier and E. Mueller.

Abstract

Over the past decades the image processing and –analysis community has spent significant effort to develop innovative algorithms and to improve existing methods in terms of accuracy, reproducibility and computational efficiency. In contrast, relatively little research has been done to find out to what extent the validity of results obtained with these methods is limited by inherent imperfections of imaging. This observation is especially true for magnetic resonance imaging (MRI)-based morphometry, which aims at the accurate and reproducible measurement of geometrical properties of anatomical structures despite the fact that MRI images are geometrically distorted.

In this thesis, for the first time the impact of the MRI data acquisition process on the validity of morphometric measurements is investigated profoundly enough to answer the - from a clinical point of view - crucial question to what extent the data acquisition process quantitatively affects the detection limit of state-of-the-art MRI-based morphometry. The investigation includes both theoretical considerations regarding the nature of acquisition-related morphological variability in image space as well as an experimental quantification of the resulting limits of image-based detection of subtle morphological changes in intra-subject MRI data series. In practice, the study reveals that acquisition-related morphological variability in MRI data series is too large to be considered irrelevant for high-accuracy applications like early detection of Alzheimer's disease.

As a consequence of this finding, possible strategies for correcting MRI data for this acquisition-related morphological variability are examined. Based upon this analysis, a novel concept for eliminating acquisition-related morphological variability from intra-subject data series is proposed, which - in contrast to existing approaches - does not require any additional phantom-imaging, but rather spatially normalizes these image data series by nonlinear alignment of markers intrinsically contained in clinical patient data. Tailored to the thesis's medical context, namely early detection of Alzheimer's disease, this novel concept can be described as white matter-based spatial normalization. Its prototypical implementation shows promising results in controlled validation experiments as well as in tests with clinical patient data. Therefore, it might serve as the methodological basis for future developments aiming at fully exploiting the clinical potential of MRI-based morphometry.

Zusammenfassung

In den vergangenen Jahrzehnten haben intensive Forschungsaktivitäten im Bereich der computergestützten Bildverarbeitung und –analyse nicht nur eine Vielzahl innovativer Methoden hervorgebracht, sondern auch zu erheblichen Verbesserungen bekannter Verfahren im Hinblick auf deren Genauigkeit, Reproduzierbarkeit und Recheneffizienz geführt. Allerdings konzentrierte man sich dabei – abgesehen von der unvermeidlichen Berücksichtigung der konkreten Bildkontrastverhältnisse – im Wesentlichen auf die Analyse der Bilddaten losgelöst von ihrer Entstehung, und es wurde nur mit vergleichsweise geringem Aufwand der Frage nachgegangen, inwieweit die Validität von mittels Bildverarbeitung erzeugten Ergebnissen begrenzt ist durch solche Imperfektionen der Bildeingangsdaten, die inhärent der speziellen Charakteristik des Datenakquisitionsprozesses geschuldet sind.

Im Kontext der Magnetresonanztomographie (MRT) wurden zuletzt intensive Diskussionen über die Validität von Traktographie-Verfahren in der Diffusions-Tensor-Bildgebung [Basser00] [Jones05], aber auch über die korrekte Interpretation zeitlicher Signaländerungen in der funktionellen Bildgebung geführt [Logothetis04] [Menon99]. Die Frage nach akquisitionsbedingten Grenzen in der MRT-basierten Morphometrie wurde jedoch bislang nicht mit derselben Aufmerksamkeit bedacht, obwohl weltweit intensiv an Hochpräzisions-Anwendungen wie der bildbasierten Früherkennung von Alzheimer gearbeitet wird.

Diese Dissertation schließt diese Lücke, indem zum ersten Mal grundlegend der Einfluss des MRT-Datenakquisitionprozesses auf die Validität von auf der Basis von MRT-Bildern gewonnenen morphometrischen Werten untersucht wird. Diese Untersuchung beinhaltet sowohl physikalisch motivierte Betrachtungen zum Wesen der akquisitionsbedingten morphologischen Variabilität, als auch eine experimentelle Quantifikation der sich aus ihr ergebenden Grenzen für die MRT-basierte Detektion von morphologischen Veränderungen in intra-individuellen MRT Datenserien. Praktisch ergibt sich aus dieser Studie, dass die akquisitionsbedingte morphologische Variabilität in solchen Datenserien im Regelfall zu groß ist, um für Hochpräzisions-Anwendungen wie die bildbasierte Früherkennung von Alzheimer als irrelevant angesehen zu werden.

Basierend auf einer Analyse möglicher Strategien zum retrospektiven Ausgleich der akquisitionsbedingten morphologischen Variabilität wird im zweiten Teil der Arbeit ein neues Konzept zur Eliminierung dieses Störeinflusses vorgestellt, welches im Gegensatz zu bekannten Verfahren keine zusätzlichen Kalibrierungsaufnahmen von speziellen Struktur-Phantomen erfordert, sondern die Datenserien unter Verwendung intrinsischer Marker räumlich normalisiert. Auf den medizinischen Kontext dieser Arbeit zugeschnitten, das heißt auf die Früherkennung von Alzheimer, lässt sich dieses Verfahren charakterisieren als eine auf der räumlichen Verteilung der weißen Hirnsubstanz basierende nichtlineare räumliche Normalisierung. Die prototypische Implementierung dieses Konzepts zeigt vielversprechende Ergebnisse sowohl in kontrollierten Validierungsexperimenten als auch in Tests mit klinischen Patientendaten. Das vorgestellte Verfahren kann somit die methodologische Grundlage für zukünftige Entwicklungen bilden, das Potential MRT-basierter Morphometrie trotz des untersuchten Phänomens der akquisitionsbedingten morphologischen Variabilität voll ausschöpfen zu können.

1. Introduction

1.1 Contributions of the Dissertation

Over the past decades the image processing and –analysis community has spent significant effort to develop innovative algorithms and to improve on existing methods in terms of accuracy, reproducibility and computational efficiency. In contrast, relatively little research was undertaken to find out to what extent the validity of results obtained with these methods is limited by inherent imperfections of imaging. This observation is especially true for magnetic resonance imaging (MRI)-based morphometry, which aims at the accurate and reproducible measurement of geometrical properties of anatomical structures despite the fact that MRI images are geometrically distorted.

Recently has been discussed the validity of fiber tracking in diffusion tensor imaging [Basser00] [Jones05] as well as the correct interpretation of signal changes observed in functional MRI [Logothetis04] [Menon99] to the effect that researchers and clinicians have become aware of errors, and consequently, of risks in using these methods for clinical diagnosis and neurobiological research. However, the problem of acquisition-related limitations in MRI-based morphometry has not gained similarly strong attention even though high-accuracy applications like early detection of neurodegenerative diseases are intensely worked on.

It can be argued that the investigation of error propagation along task-specific clinical workflows from patient positioning via data acquisition and image processing to diagnosis is not the field of duty for the image processing and –analysis community. But to our understanding it is desirable to not only provide physicians with state-of-the art post-processing tools, but also to reliably equip them with the methods' limits in a clinical setting. The latter inevitably requires detailed knowledge about the characteristics of the image formation process and subsequent processing stages.

This thesis addresses this topic in a twofold way:

- The impact of the MRI data acquisition process on the validity of morphometric measurements is investigated for the first time in a long-term study including effects resulting both from imperfect hardware stability and from imperfections in patient repositioning. The investigation includes both theoretical considerations regarding the nature of acquisition-related morphological variability in image space as well as an experimental quantification of the resulting limits of image-based morphometric applications aiming at the detection of subtle changes in intra-subject MRI data series.
- Secondly, a novel concept for eliminating this acquisition-related morphological variability from intra-subject data series is proposed. In contrast to existing phantom-based approaches, this technique purely relies on the nonlinear alignment of markers intrinsically contained in clinical patient data. The method is specifically tailored for the use in early detection of neurodegenerative diseases like Alzheimer's.

1.2 Structure of the Dissertation

Chapter 2 provides the medical and technical background for this thesis: By briefly introducing Alzheimer's disease (AD) to the non-expert reader, it is clarified which phenomena to concentrate on in MRI-based early detection of AD, and a deep understanding of the high requirements that this task imposes on post-processing applications in terms of accuracy and reproducibility is established. These medical considerations are followed in the chapter's second part by an overview of the most important computational methods known from literature to quantitatively describe morphology and morphological changes based upon the analysis of 3D image data. However, none of these approaches explicitly takes into consideration potential morphological inaccuracies inherently contained in the input images already. Instead, all these techniques implicitly and tacitly assume that the data acquisition process does not introduce any morphological variability into image space.

Chapter 3 examines by way of a system analysis whether this assumption is correct in the context of MRI-based morphometry. This is done by investigating the impact of the MRI data acquisition process on morphometric analyses both qualitatively by an inspection of the underlying theory of MRI physics and quantitatively by means of a specifically designed experimental long-term study. The study reveals that acquisition-related morphological variability cannot be neglected in high-accuracy morphometry: Diseases that are accompanied by gradual morphological changes like AD cannot be reliably detected in MRI data series until the actual anatomical change exceeds the experimentally determined detection limit of about 2% local volume change. Since this is too much to be considered irrelevant for early detection of AD, possible strategies for correcting MRI data for this acquisition-related morphological variability are examined. Based upon these considerations, it is finally developed a novel concept for eliminating acquisition-related morphological variability in image space from intra-subject MRI data series, which does not require any additional phantom-imaging, but rather spatially normalizes these image data series by nonlinear registration of white matter (WM).

Chapter 4 presents this novel concept in terms of a software prototype implemented in order to be able to experimentally test the correctness of the previously presented system analysis. It starts with a focused review of the state-of-the-art in image registration and WM segmentation in order to identify the most promising image processing methods to be used for a prototypical implementation of WM-based spatial normalization. Subsequently, this general specification is refined by a discussion of the technical details of the prototype implemented in this thesis. Finally, experimental results obtained with this prototype are presented, whereby special emphasis is put on controlled validation experiments in order to allow for a profound discussion of both the method's general potential, and its limitations in its prototypical implementation. The latter also serves as the foundation for suggestions for future work.

Chapter 5 evaluates how well the proposed method for elimination of acquisition-related morphological variability is compatible with MRI contrasts other than T1-weighting, as it can be assumed that post-processing applications in the near future are extended to allow for analysis of multiple contrasts in a combined way.

Finally, the main findings of the previous chapters are summarized in chapter 6 in an overall conclusion.

For better readability, each chapter starts with an introduction that briefly outlines the chapter's internal structure as well as its position within the thesis's overall context. Each chapter is completed by a summary of the main results and of the resulting conclusions with regard to high-accuracy morphometry.

1.3 Additional Remarks

Since the problem of acquisition-related morphological variability is of relevance not only for the image processing and –analysis community, but also for medical professionals, the dissertation is written in a style of being of benefit to both communities. Accordingly, in chapters 2 and 3, mathematical and physical principles of the discussed topics are described in such a way that they are comprehensible for both communities. However, chapter 4 inevitably requires the incorporation of mathematical details to allow for an adequate presentation of the methodological principles and implementational details of WM-based elimination of acquisition-related morphological variability.

Regarding the terminology used throughout this dissertation, care was taken to consistently use the termini technici as defined in appendix A. To further avoid misunderstanding, it is stressed here in advance that results determined by the application of an image analysis algorithm are merely estimations of the true magnitude of the anatomical quantity, and that the term acquisition-related morphological variability exclusively refers to the imaged object's morphological representation in image space and does not imply any true variability in the physical world.

Finally, it should be pointed out that in this thesis, MRI pulse sequences are described using "Siemens terminology". A comprehensive overview of the different MRI vendors' terminology is given in [Nitz99]. For example, General Electric uses the acronym SPGR (Fast Spoiled Gradient Refocused Acquisition into steady state), and Philips the acronym T1-FFE (T1-weighted Fast Field Echo) for the imaging technique referred to by Siemens as FLASH (Fast Low Angle SHot). For better readability, all acronyms used in this thesis are listed and explained in appendix B.

2. Medical and Technical Context

2.1 The Anatomy of Alzheimer's Disease

2.1.1 Introduction

In the last years, image-based early detection of Alzheimer's disease (AD) has become one of the most intensely worked on topics in MRI-based morphometry. As the investigation of the relevance of imperfections in the data acquisition process for the validity of results obtained with state-of-the-art methods for early AD detection requires profound knowledge of the way the disorder manifests itself in image space, this chapter briefly introduces AD for the non-expert reader. Besides a coarse outline of its pathological characteristics, special attention is paid to the description of the pattern of progression and the way this differs from other neurodegenerative diseases.

2.1.2 Pathology

Alzheimer is a neurodegenerative disorder the main characteristic of which is preceding atrophy in certain brain structures. In detail, communication between neurons declines, so that these slowly lose synaptic connections to other nerve cells, until they stop functioning and finally die [NIA02] [Scahill02]. Although the reasons for this massive reduction in the number of neurons have not finally been found, two undisputed hallmarks have arised from autopsies of AD patients' brains [Braak91]:

- Amyloid plaques (AP) are an insoluble aggregation of protein fragments along with fragments of dead or dying neurons and other cells. They are formed when the amyloid precursor protein (APP), a component of the cell membrane, is cleaved apart into beta-amyloid by specific enzymes. Several studies revealed a toxic effect of beta-amyloid to neurons: Concretely, it is discussed that beta-amyloid might cause inflammation in the brain, block receptors for neurotransmitters, deplete intracellular stores of these neurotransmitters or generate free radicals that are suspected to play a key role in ageing and degeneration of neurons. However, the question whether AP are the reasons for AD or if they merely are a byproduct of the disease could not finally be answered up to now [NIA00].
- Neurofibrillary tangles (NFT) are pathological deposits of helically twisted filaments inside of neurons, mainly consisting of chemically modified tau-proteins. In healthy neurons, tau-proteins stabilize the cells' microtubules, but the chemically altered variant is no longer fully capable of fulfilling this task. This leads to a pathological disruption of intracellular transport of molecules to the axon and, thus, to a disordered communication between affected nerve cells and may later even cause neurons' death [NIA00].

Amyloid plaques and neurofibrillary tangles are not the only characteristics of AD on a molecular level: It was shown that it is also accompanied by locally increased iron concentration, especially in those areas that are typically affected by atrophy [Bartzokis00] [Ong05]. However, elevated iron accumulation holds true for other neurodegenerative diseases as well, which is why AD cannot be reliably diagnosed based on this symptom alone.

2.1.3 Progression

AD neither equally affects all brain areas nor proceeds linearly in the affected regions. Autopsies as well as (functional) imaging techniques rather proved that the progression of AD follows a characteristic pattern, starting in the medial temporal lobe, moving on to the temporoparietal association cortex, then to frontal, and finally, to primary sensory and visual areas[1] [Thompson03] [Scahill02] [Thacker02].

Studies about atrophy in presymptomatic AD patients revealed significant loss of brain volume in the entorhinal cortex and the hippocampus [Scahill02] [Reiman96] [Lehtovirta00]. This is remarkably consistent with observations of higher concentration of AP and NFT in these regions [Braak91]. Furthermore, these findings very much correspond to early clinical symptoms of AD patients, as the hippocampus plays a key role in memory processes.

Whereas some studies report that in early AD the precuneus and the anterior frontal lobe show significant atrophy [Scahill02], other scientists found reduced metabolism in the posterior cingulated gyrus as well as in parietal, temporal and frontal lobe regions [Minishima97] [Reiman96]. However, with the disease proceeding and the physiological processes becoming more prominent, the differences between existing theories more or less vanish.

In patients mildly affected[2] with AD, atrophy is still superproportionally present in the hippocampus, entorhinal cortex, precuneus and frontal lobe, but additionally also in the inferior and lateral parts of the temporal lobes as well as in the posterior part of the cingulated gyrus [Thompson03] [Scahill02].

Moderate AD is characterized by significant volume loss in the frontal cortices, the posterior cingulated gyrus as well as the precuneus and the inferior and lateral temporal lobe structures [Thompson03] [Scahill02]. Whereas the sensory and motor areas are still comparatively left out, hippocampal atrophy is no longer significantly higher than in healthy patients [Scahill02]. In detail, Thompson et al. report an annual loss of local GM volume of up to 5% in the reported brain regions at this stage of the disease (decline in average MMSE score from 18 to 13). The annual GM loss rate in age-matched healthy controls is reported to be less than 1%.

Whether the atrophic process also reduces the corpus callosum's volume to a higher extent than in normal ageing is not clear, yet [Thompson03] [Janke00].

Generally spoken, studies on the progression of AD in the brain do not always agree in every detail, but the overall pattern of atrophy is quite clear: AD seems to start in the entorhinal cortex and in the hippocampus, and from there to move on to the temporoparietal cortex and to frontal lobe regions, leaving out sensory and motor areas until the very end of the disorder. It is furthermore undisputed that the left hemisphere is more severely affected by brain volume loss than the right one [Thompson03] [Scahill02] [Janke01] [Thompson01b].

---

[1] "Area" by definition is a physical quantity expressing the size of a part of a surface. In the study of brain function, however, the term is also used to describe regions that are known to be responsible for a certain cognitive ability. In accordance to this practice, this thesis names a cortical region an "area" whenever it is defined by its function rather than its anatomical location.

[2] It is difficult to classify AD patients into groups according to the disease stage they are in. Here, following Scahill and collegues, mild AD corresponds to a MMSE score of 20-27, and moderate AD to a score of 8-19 [Scahill02].

This pathological cascade needs years to reach a stadium that corresponds to clinically relevant symptoms like memory deficits or even of higher-order cognitive malfunctions. And even from the first obvious symptoms it lasts several years until the dementia has fully developed [Elias00].

The question why AD spreads in the described manner still cannot be fully answered. According to the disconnection theory [Jobst92], the process can be understood as a chain reaction. When hippocampal and entorhinal neurons die because of AD, the synaptic input of nerve cells in the temporoparietal cortex that were previously connected to them declines. This is accompanied by slower metabolism, a decrease in blood flow, and it might even cause these neurons to develop NFT [Smith02a]. As these nerve cells die, their connections to other cortical regions break down, and the whole process will start again.

Another theory suggests that AD is a consequence of plasticity failure in a way that excessive plasticity-related cellular activity might eventually cause errors like loss of synapses, dendrites and neurons, and that this might also be the underlying reason for the generation of AP and NFT [Mesulam00].

2.1.4 Diagnosis

AD is defined by its pathology, but as neither AP nor NFT can yet be detected in vivo, AD so far is diagnosed by means of specific neuropsychiatric tests, mostly the Mini Mental State Examination (MMSE). These tests are designed to evaluate the patient's performance in tasks stressing mental abilities that are typically affected by AD. Especially the patients' short time memory and their capability in word finding are tested for. However, for a profound confirmation of the diagnosis, a histological analysis of the patient's brain is required.

As a result of the upcoming financial impact of AD on the health-care system, a lot of research activities regarding early AD detection have been started. Various approaches are being followed, ranging from identification of AD-specific markers in CSF [NIA02] to analysing atrophy in the brain on the basis of brain imaging [Thompson03] [Scahill02] [Thacker02].

2.1.5 Therapy

As stated above, the factors initializing AD are not clear yet. Consequently, there is no way to cure AD so far. Topical therapies concentrate on slowing down the process of degeneration, but are unable to stop or even reverse it.

Again, there are several approaches currently being followed. Some researchers are developing drugs that aim at prohibiting enzymes from breaking down neurotransmitters into their constituents, hoping to slow down the disconnection phenomenon in the brain [NIA00], while others work on radical-blocking antioxidant substances or on anti-inflammation drugs. Some studies even showed promising results by simple application of vitamin e or estrogen [NIA02]. Of course, there is also a lot of research taking place to reduce the psychiatric and behavioural problems of patients with advanced AD.

Topical animal experiments on the one hand focus on prohibiting the generation of beta-amyloid and on the other hand aim at generating new fully functional neurons in the brain by making use of stem cells [NIA02].

2.1.6 Other dementias

There are several forms of dementia besides AD, among them frontotemporal dementia (FTD) and vascular dementia (VAD), both of which show similar patterns of atrophy. Furthermore, these patterns are close to that of normal ageing, which makes it difficult to discriminate between these disorders.

The question whether the pattern of atrophy in AD is unique in principle can be answered positively. Recent cross-sectional studies revealed that each dementia is accompanied by a slightly different atrophic pattern, but nonetheless it was impossible to reliably assign an individual patient to the correct group based on the detected atrophy alone, since the distributions of the measures, which atrophy was quantified with, substantially overlapped [Thacker02]. Stated differences were that FTD-affected brains tend to show more atrophy in larger areas of the cortex and that atrophy in the medial temporal lobe [Forstl96] [Kitagaki98], though being highly accelerated, is not as severe as in AD patients [Frisoni99].

Discriminating AD from normal healthy ageing based upon the characterization of atrophy proved to yield higher sensitivity and specificity, as the rate of atrophy in the medial temporal lobe exceeds the normal rate more clearly [Jack97]. However, depression and alcohol-related cognitive impairment, being phenomena ageing often is accompanied by, significantly reduced the sensitivity [O'Brien97]. In a clinical setting, it obviously is important to take these confounders into consideration as well.

2.1.7 Conclusion

Profound knowledge of AD-induced morphological changes in the brain is a substantial foundation for the evaluation of existing image-based approaches for early detection of AD as well as the basis for the investigation of the relevance of imperfections in the data acquisition process for this task.

Summarizing the preceding paragraphs, the following characteristics of atrophy in AD can be safely stated that are relevant for image-based early diagnosis:

- Atrophy starts in the medial temporal lobe: Structures of primary interest are the hippocampus and the entorhinal cortex, but further regions might be affected by atrophy before clinical symptoms of AD become apparent. As reported, candidate structures are the precuneus and the anterior frontal lobe. Since evidence on the locations of atrophy in AD is not complete yet, it might not be sufficient to solely analyze these structures.
- In early AD there is no increased atrophy in sensory and motor areas.
- Atrophy is more severe in the left hemisphere than in the right; due to this asymmetry, brain volume of the left hemisphere has to be smaller than the right one
- As a result of atrophy in AD, the ventricular volume significantly increases. Restricting the morphometric analysis to the ventricles, however, is insufficient due to the indirect character of this measure and due to the similarity in the pattern of atrophy of other types of dementia: A small difference in volume loss in brain regions that are spatially close might lead to undistinguishable changes in ventricular volume and shape.
- Early AD detection is most promisingly done in longitudinal studies rather than comparing a patient's brain to a typical AD-brain, as individual differences in brain structure exceed the effects resulting from atrophy: Not

only the total brain volume, but even the volume of the hippocampus alone varies depending on the person's everyday life [Maguire00].

Even though the literature so far does not make any statement on how accurately these AD-induced characteristics have to be quantified in order to be able to separate AD from other dementias, it is obvious that image-based early detection of AD requires the reliable detection of very subtle changes throughout the whole brain in order to be clinically useful. For the remainder of this dissertation, the annual local GM loss rate of up to 5% as reported by Thompson et al for moderate AD patients can serve as a helpful orientation [Thompson03].

2.2 Image-Based Morphometry – Measures and Approaches

2.2.1 Introduction

In this section the most important computational methods known from literature to quantitatively describe anatomy based upon the analysis of 3D medical image data arte introduced. Further insight into these approaches is given in the presentation of widespread morphometric measures as most of them are connected to a specific method. The primary goal of this overview is to establish common understanding about how the image processing community so far has (or has not) incorporated knowledge about the image formation process into the development of morphometric image analysis tools. Consequently, this overview is focused on the methods' underlying principal concepts, rather than exhaustively presenting every technical variant of the general ideas. However, it also comprises a brief discussion of the various methods' suitability for early detection of AD.

2.2.2 Voxel-Based Morphometry

2.2.2.1 Principle

Voxel-based morphometry (VBM) involves a voxel-wise comparison of the local value of the chosen morphometric measure [Ashburner00]. To ensure that anatomically corresponding structures are compared, this approach requires spatial normalization of the datasets. While this is relatively straightforward in cases of intra-subject studies where structural change over time is investigated, this spatial normalization is rather difficult when morphological change between different populations is to be quantified: on the one hand, brain structure varies significantly between subjects especially in the cortex such that the underlying transformation must be allowed to substantially deform the datasets, but on the other hand only global changes are to be evened out and not inter-individual differences in the size and shape of the structures of interest. Practically, this spatial normalization is mostly done by an affine registration to a common template followed by a low-frequency nonlinear registration. On of the most common criticisms of VBM is that the decision on how many degrees of freedom are incorporated in the spatial normalization step is made rather arbitrary, as it is never quite clear which morphological differences are to be considered global and which are rather local (see figure 2.1).

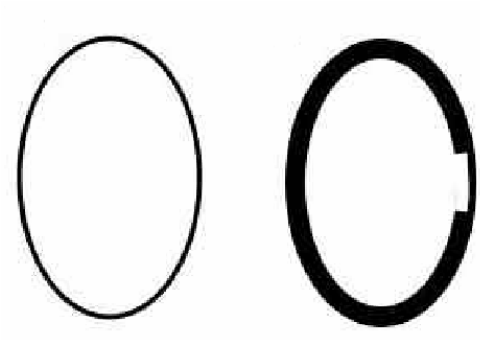


Figure 2.1: *Sketch of the difficulty encountered in cross-sectional studies which morphological differences are to be considered global and which are rather local. The difference in thickness can be interpreted as a global difference between these "subjects" that needs to be eliminated in the spatial normalization or as a local growth of the "tissue" that must be preserved throughout the normalization procedure.*

Following spatial registration tissue density maps are calculated for each voxel of the normalized images. Some implementations of VBM additionally incorporate a so-called "modulation" which uses the Jacobian (see 2.2.3.2.2) of the nonlinear deformation field to modulate the intensity in the tissue density maps (and thus, to correct for volumetric changes due to nonlinear spatial normalization). In cross-sectional studies voxel-wise parametric statistical tests between the populations are carried out to locate regions that show significant differences with respect to the tissue density of interest.

2.2.2.2 Measures – Tissue Density

Tissue densities reflect the proportion of certain tissues in a voxel. Their computation requires the datasets to have been segmented into the candidate tissue types. Most recent segmentation algorithms yield a soft segmentation, i.e. they generate maps that contain for each voxel the probability of class membership. This is typically done by modeling the intensity distribution in the brain as a mixture of three (GM, WM, CSF) independent Gaussian distributed probability density functions and then iteratively searching for that parameter set which best explains the given histogram. Tissue density maps are then computed by smoothing each segmentation result with an isotropic Gaussian kernel. Practically, especially GM density is used for detecting disease-specific anatomical changes between different populations [Ashburner00] [Davatzikos01].

Apparently, any change in position or volume of a cortical structure inevitably leads to a change in GM density, which makes it a powerful measure for morphometry including early AD detection. The crucial step for tissue density based morphometric applications is an accurate and reproducible automatic segmentation that pays special attention to partial volume effects (for an illustration of partial volume effects see figure 2.4).

2.2.3 Deformation-Based Morphometry

2.2.3.1 Principle

Another popular strategy in morphometry is to evaluate structural changes by analyzing deformation fields [Gaser01] [Chung01]. To do so the datasets are first spatially normalized to compensate for global differences in brain size etc. via affine registration. In deformation-based morphometry (DBM), this normalization is followed by an elastic registration step to bring the datasets into best possible anatomical alignment, meaning that in contrast to registration in VBM, inter-

individual differences in size of anatomical structures have to be compensated for to allow for valid statistical results.

After elastic registration of the images, ideally every change in shape and volume is contained in the deformation fields. Statistical tests are then performed on these deformation fields or on measures derived from them.

2.2.3.2 Measures

2.2.3.2.1 Displacement Field

The most straightforward way to statistically evaluate the deformation fields is to directly compare the displacements for each voxel. When applied to cross-sectional studies, however, this technique does not directly localize regions with different structures, but rather identifies brain structures that have translated to different positions, which often is not of primary interest [Chung01].

2.2.3.2.2 Jacobian Determinant

Due to these difficulties, in most cases the deformation fields (here: $\vec{u}(x,y,z)$, see equation 2.1 and 2.2) are further analyzed to detect regions of significant local compression or contraction. One possibility to do this is the evaluation of the determinant of the Jacobian J of these deformation fields for every voxel [Gaser01]. Whereas $J>1$ indicates local expansion, $J<1$ corresponds to regions of compression.

$$\vec{x} = (x, y, z) \rightarrow (x + u_x(x, y, z), y + u_y(x, y, z), z + u_z(x, y, z)) \qquad (2.1)$$

$$J(\vec{x}) = \det(\nabla \vec{u}\,|_{\vec{x}}) = \begin{vmatrix} \frac{\partial u_x(\vec{x})}{\partial x} & \frac{\partial u_y(\vec{x})}{\partial x} & \frac{\partial u_z(\vec{x})}{\partial x} \\ \frac{\partial u_x(\vec{x})}{\partial y} & \frac{\partial u_y(\vec{x})}{\partial y} & \frac{\partial u_z(\vec{x})}{\partial y} \\ \frac{\partial u_x(\vec{x})}{\partial z} & \frac{\partial u_y(\vec{x})}{\partial z} & \frac{\partial u_z(\vec{x})}{\partial z} \end{vmatrix} \qquad (2.2)$$

If an atrophic pattern is known well enough, a comparison of atrophy rates in the regions of interest is sufficient for diagnosis. Nevertheless, an identification of these regions within the individual brain has to be done either manually or by registration to an atlas[3].

In contrast to VBM, the crucial step in characterizing the atrophic pattern by the analysis of nonlinear deformation fields no longer is an accurate segmentation, but a similarly accurate brain registration. If this can be achieved, any disease being accompanied with characteristic morphologic changes can, in principle, be detected with the Jacobian determinant, among them AD.

[3] It should be mentioned that such a registration requires a special atlas despite the small amount of AD induced atrophy in cases of early Alzheimer's disease, as AD patients are significantly older than the normal population. Consequently, it would be necessary to build an atlas based on image data of an age-matched population to ensure that the atlas adequately represents normal ageing effects.

2.2.4 Surface-Based Morphometry

2.2.4.1 Principle

In contrast to the preceding strategies, surface-based morphometry does not statistically evaluate the complete image volume, but rather concentrates on characteristics of surfaces [Chung03]. Obviously, this requires accurate segmentation, most commonly of tissue classes, but it could be equally applied to anatomical structures. Based upon these segmentation results, approximations of the structures' surfaces are reconstructed which subsequently are statistically analyzed, regarding both local and global features.

One of the major challenges in local surface-based morphometry is the reliable identification of corresponding points lying on identical cortical surfaces, but in different datasets.

One possible way to do this is to establish a uniform surface-based coordinate system and subsequently perform elastic registrations and statistical comparisons in this new coordinate system. The most suitable surface-model for the cortex is a sphere respectively an ellipsoid, as this allows for preservation of the original surface's topological structure, especially the local connectivity. There are several approaches for transforming the highly convoluted cortical surface into a sphere [Thompson01a] [Fischl99b]. However, it is not undisputed, whether surface flattening does not substantially distort the inherent geometry of the cortex for further analyses.

A reasonable alternative is to constrain the process of surface-reconstruction to guarantee a one-to-one correspondence of vertices, and thus, triangles. This is done by segmenting the inner and outer cortical surface by inflating an ellipsoid in such a way that no self-intersections occur and that each surface has exactly the same number of vertices. As every vertex can be traced back to a vertex of the original ellipsoid there is an automatic linkage between vertices [MacDonald00]. Nevertheless, the constraints imposed on surface-reconstruction almost certainly confound the results as they limit the surface's spatial resolution.

2.2.4.2 Measures

2.2.4.2.1 Cortical Thickness

Cortical thickness naturally corresponds to the cerebral cortex's organization. Concretely, the cortex consists of six layers, in which the functional units extend approximately vertically. Since their true orientation cannot be identified in MRI due to limited spatial resolution, the most common approach for estimating cortical thickness is the bi-directional straight-line approach, in which the length of the shortest line from the cortical surface to the GM and WM boundary is computed and vice versa (see figure 2.2).

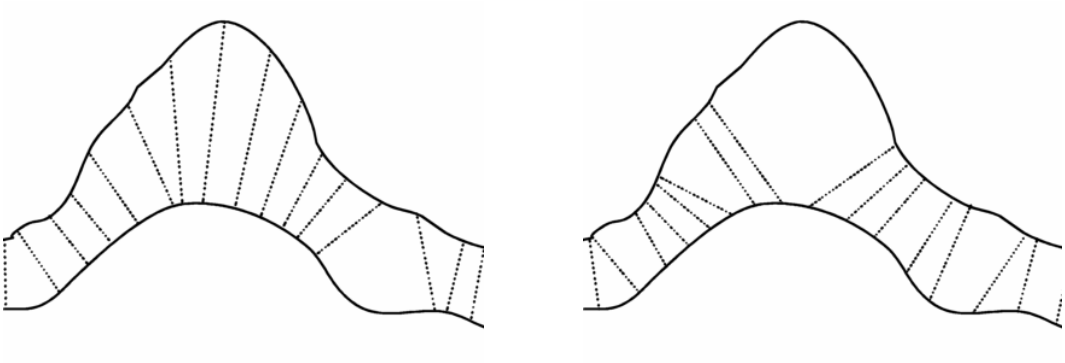


Figure 2.2: *Sketch of the bi-directional straight-line approach: To avoid ambiguities, cortical thickness is estimated as the average of the shortest line from the cortical surface to the WM/GM-interface (left) and the shortest line from the WM/GM-interface to the cortical surface (right).*

Other implementations make use of their inherent linkage between vertices and simply compute the Euclidean distance between two vertices that originated from identical vertices on the ellipsoid [Chung03].

2.2.4.2.2 Surface Area

When forcing a dynamic surface to always consist of the same number of vertices and faces, morphological change automatically manifests itself in a changed surface area of triangles that are situated at the border of a degenerating or expanding structure. The value of surface area of every triangle can, thus, be analyzed to detect morphological change [Chung03][4]. Obviously, surface triangles have to be linked in the above-mentioned way to be analyzed like this.

2.2.4.2.3 Curvature

Deformation of an arbitrary structure automatically leads to changes in the structure's surface curvature. Local morphological change can, thus, also be detected by comparing the curvature at corresponding vertices [Chung03]. Significantly increased local curvature can be interpreted as folding of the underlying brain regions, whereas decreasing curvature indicates flattening of structures.

2.2.5 Direct Volumetric Morphometry

2.2.5.1 Principle

Probably the most straightforward way to quantitatively describe anatomy is to first delineate the structures of interest in the datasets, and then compute their approximate volumes as the resulting voxel sets' volumes. In modern applications this is done automatically by specially designed segmentation algorithms, since manual delineation is very time-consuming and less reproducible. In contrast to the preceding approaches this method needs a priori knowledge about which structures to segment.

2.2.5.2 Measures – Volume and Volume Ratios

Apart from directly comparing the estimated volume of (a) brain structures like the hippocampus [Starck02] [Shen02], (b) tissue classes like GM [Sastre05] [Karas04] or (c) the ventricles [Wu02] [Hahn01], it is common to compute volume ratios to gain specificity. An increase in the volume of the lateral ventricles, for instance, may have a number of different causes. Thus, without assessment of other regions, a particular

[4] In fact, Chung defines surface area change on vertices rather than triangles, but as this requires a specific surface parameterization (2D Riemannian manifold), it seemed to be more appropriate to present it in a more general way [Chung03].

loss or increase may lack specificity. One example of such a global measure is the ventricle-brain ratio (VBR) that is often used in the study of schizophrenia [Gaser04].

2.2.6 Dedicated Approaches for Longitudinal Morphometry

As pointed out earlier, image-based early detection of Alzheimer's disease is most promisingly done in longitudinal studies rather than comparing a patient's brain to a typical Alzheimer's brain, as inter-individual differences in brain structure exceed the effects resulting from atrophy. None of the morphometric approaches outlined so far is specifically dedicated to this longitudinal imaging scenario. However, they can be used for longitudinal studies as well by analyzing each dataset of the longitudinal data series individually as accurately as possible.

The only measure developed explicitly for quantification of morphological changes in longitudinal data series is the so-called boundary shift integral (BSI), proposed by Freeborough and Fox [Freeborough97]. Its basic idea is to integrate intensity differences along the brain-CSF border in registered MR images within an appropriate intensity interval (see figure 2.3).

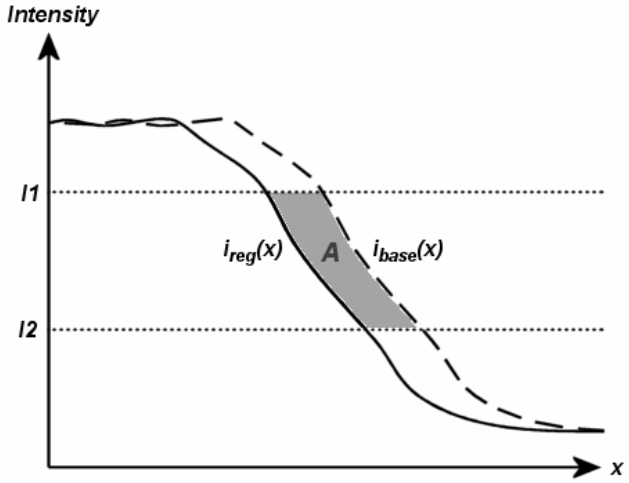


Figure 2.3: *A one-dimensional representation of the boundary shift between a baseline scan ibase(x) and a registered follow-up scan ireg(x). An estimate of the shift Δx is computed as the area A divided by I1-I2 [Freeborough97].*

When applied to whole brain morphometry, this method clearly can only be used to estimate the overall brain volume change rather than localize it, and thus, it cannot differentiate between different pathologies being accompanied by atrophy. Some research groups therefore combine the BSI method with a semi-automated identification of regions of interest in order to be able to estimate individual atrophy rates for different anatomical regions [Barnes04] [Paviour06]. The validity of results obtained from these variants, however, strongly relies on the analyzed subvolumes to be anatomically identical not only across subjects, but even in an intra-subject evaluation.

2.2.7 Discussion

Morphometric measures can be divided into three groups based upon the central postprocessing step required for the computation of theirs:

- Tissue density, surface metrics, and absolute or relative volumetric measures rely on an accurate segmentation of either tissue classes or anatomical structures.
- The estimation of specific atrophy patterns and derived measures is based upon a similarly accurate nonlinear registration of datasets.
- BSI relies purely on a reproducible segmentation of longitudinal datasets into brain and non-brain voxels, followed by an accurate linear registration of the brain regions.

Out of segmentation-based measures, volume ratios or absolute volumes have to be considered least suitable for early AD detection, as it is not clear yet, which structures are necessary to segment in order to be able to reliably differentiate between AD and dementias showing similar atrophy patterns. Focusing on an indirect measure like ventricular volume that comprises all volumetric changes in the brain, however, does not truly solve this problem, since a small difference in volume loss in brain regions that are spatially close might lead to undistinguishable changes in ventricular volume and shape. To overcome the outlined limitations of volume ratios and absolute volumes, a complete segmentation of all anatomical structures within the brain would be necessary. Given the high demands for accuracy and reproducibility that early AD detection imposes, such a segmentation is a very difficult task, as some brain structures are small in volume, but have a comparatively large surface, which makes valid volumetry of theirs prone to errors caused by partial volume effects.

Moreover, neighbouring inner brain structures do not necessarily show any noticeable contrast in image intensity at resolutions currently achievable in structural MRI (see figure 2.4). The border between the hippocampus and the amygdala, for example, can only be defined statistically based upon prior knowledge. Early AD detection based upon automated whole-brain segmentation therefore is promising only, if the segmentation procedure incorporates a high-quality probabilistic atlas that also contains morphological changes normal ageing is accompanied by.

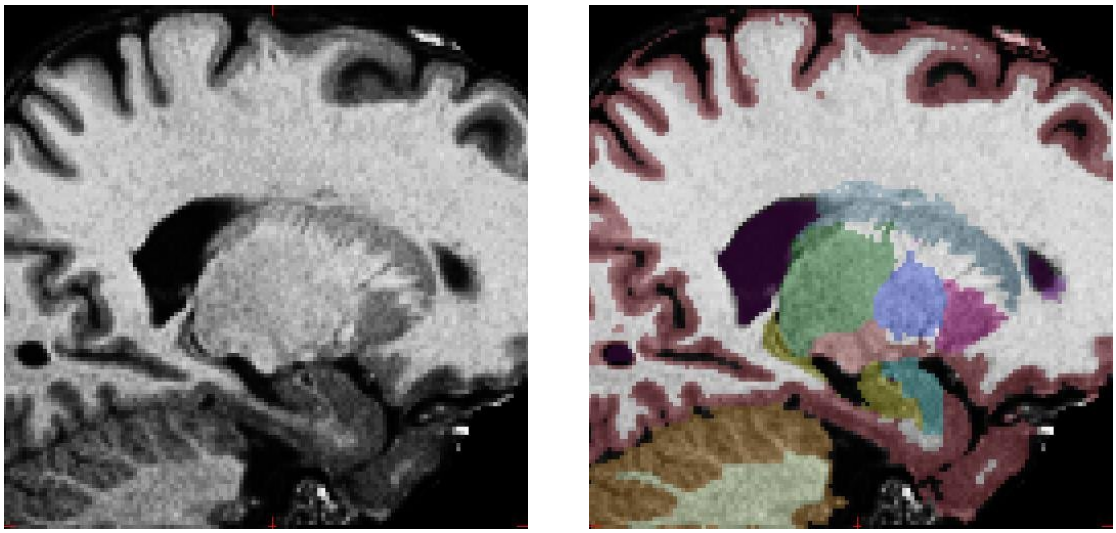


Figure 2.4: *Visualization of two problems faced in (subcortical) segmentation: (a) some subcortical structures do not show any intensity difference from their neighbours, so that their identification has to be purely based on model knowledge; (b) segmentation of small thin but stretched structures like the hippocampus are very likely to be severely affected by partial volume effects and – in a longitudinal study – by slight changes in patient positioning; (segmentation was done by the author with the tools developed at MGH by Fischl et al. [Fischl02])*

Cortical thickness techniques as well as other surface-based approaches require a very consistent parametric representation of the highly convoluted cortical surface such that changes in area, curvature or distance are not merely resulting from slight differences in triangulation e. g. due to differences in image discretization during the imaging process caused by a slightly altered position of the patient in the scanner. Additionally, surface-based methods imply that a soft segmentation result has to be converted into binary images for each tissue class in order to reconstruct surfaces from these binary voxel sets. As illustrated in figure 2.5, this conversion is not a trivial task, since different interpretations of voxels containing more than one tissue class can seriously affect the shape of the subsequently reconstructed surface.

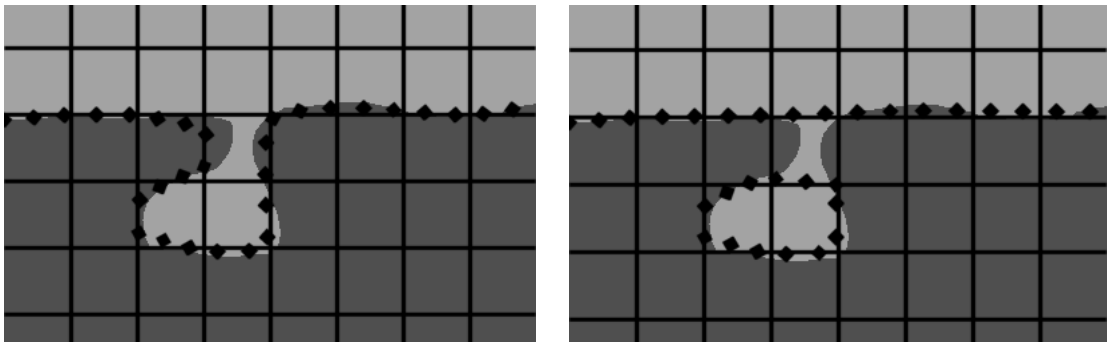


Figure 2.5: *Different interpretations of voxels containing more than one tissue class (partial volume effect) can seriously affect the shape of a reconstructed surface (here: dotted line).*

Furthermore, it is all but trivial to compare complex surfaces, as it is difficult to identify homologous points on them. Methods that tackle this correspondence-problem by guaranteeing linkage between vertices during surface-reconstruction do not seem to be appropriate for early AD detection: Though being effective in cross-sectional studies, this restriction implicitly reduces the surfaces' spatial resolution and, thus, automatically limits the accuracy achievable with these techniques which is a major drawback for early AD detection. Other approaches try to solve the correspondence-problem by establishing a new surface-based spherical coordinate system. However, the mapping of highly convoluted surfaces like the outer cortex onto a sphere bears the risk of confounding further analyses, as the geometrical properties of the original surface, e. g. point-to-point distances, may be too drastically distorted.

In total, the described difficulties in surface-reconstruction from discrete segmentation images in the presence of partial volume effects as well as the outlined challenges regarding a consistent parametric representation of the cortical surface without reducing its spatial resolution (due to too restrictive smoothness constraints) have to be considered potential error sources that are specific to surface-based morphometry. These methods therefore appear to be more error-prone than GM density approaches or deformation-based techniques that simply do not require these critical processing steps.

BSI is a robust and computationally rather simple approach for estimating the whole-brain volume change from longitudinal MRI data. However, it is not capable of localizing this change as it integrates the intensity differences over the complete brain-CSF border. BSI computation alone therefore does not allow for a reliable differentiation between AD and other diseases that are accompanied by brain atrophy. At first glance, regional BSI analyses seem to solve this problem, but practically they merely shift it to the necessary step of segmenting the ROIs with high accuracy and

reproducibility. The problems involved in automated delineation of inner brain structures already were described in the critical inspection of whole-brain segmentation approaches which is why they are not repeated here.

Tissue density approaches and deformation-based techniques prima facie seem to be comparable both in significance and in computational complexity. However, a concluding evaluation of these methods' suitability for early detection of neurodegenerative diseases has to be based upon a profound analysis of the image formation process.

2.2.8 Conclusion

A great variety of different computational methods to quantitatively describe anatomy based upon the analysis of 3D medical image data have been proposed by the image processing and –analysis community so far. The assessment of these methods' suitability for high-accuracy applications like early detection of AD revealed that surface-based approaches appear more error-prone than tissue density approaches or deformation-based techniques because of additional technical challenges in accurate and reproducible surface-reconstruction from discrete segmentation images in the presence of partial volume effects. Moreover, it became evident that direct volumetry is least suitable for early AD detection, as it is not clear yet, which structures are necessary to segment in order to be able to reliably differentiate between AD and dementias showing similar atrophy patterns.

Most importantly, however, it can be stated that none of the various approaches explicitly takes into consideration potential morphological inaccuracies inherently contained in the input images. Instead, all these methods implicitly and tacitly assume that the data acquisition process does not introduce any morphological variability into image space. The question whether this assumption is correct in the case of MRI or whether the MRI data acquisition process imposes an intrinsic detection limit for subsequent image-based morphometric analyses is elaborately investigated in the next chapter.

3 The Impact of the MRI Data Acquisition Process on High-Accuracy Morphometry

3.1 Introduction

Having shown that MRI-based early detection of Alzheimer's disease requires the accurate and reproducible quantification of very subtle changes in a patient's brain, this chapter qualitatively and quantitatively investigates the impact of the MRI data acquisition process on such high-accuracy morphometric applications. As this investigation also takes into consideration effects resulting from imperfect hardware stability and from imperfections in patient repositioning, it answers for the first time the question to what extent the MRI data acquisition process affects the detection limit of existing morphometric applications.

Based upon this analysis, it is theoretically developed a novel concept for eliminating acquisition-related morphological variability in image space from intra-subject MRI data series, which does not require any additional phantom-imaging. This concept's general superiority to existing approaches is shown by an evaluation with respect to the approaches' capabilities to spatially correct intra-subject MRI head data series such that radiologists can be provided with – ideally – perfectly reproducible high-resolution image series of a patient's brain. In the ideal case, this would technically reduce image-based detection of subtle morphological changes to the computation of difference images[5].

As a result of the topic's complexity, the chapter is divided into three parts:

- Firstly, it is introduced into those characteristics of the MRI data acquisition process that are relevant for understanding its impact on high-accuracy morphometric analyses. It is thereby put special emphasis on a profound, i.e. physics-based, but still intuitively understandable explanation of these characteristics.
- Based upon these considerations, an experimental study is designed to quantify the impact of the MRI data acquisition process on the validity of morphometric measurements. The results of this study for the first time answer the question to what extent the MRI data acquisition process affects the detection limit of existing morphometric applications.
- Finally, possible strategies for correcting MRI data series for acquisition-related morphological imperfections are investigated, and they are evaluated with respect to their suitability for the use in longitudinal imaging scenarios. As outlined above, this includes the presentation of a novel concept for eliminating acquisition-related morphological variability in image space which does not require any additional phantom-imaging.

For better readability, each of these three subchapters is self-contained insofar as they start with an introduction that briefly outlines their internal structure as well as their

[5] Apparently, subtle changes can only be detected if the images' spatial resolution and contrast-to-noise-ratio (CNR) is sufficiently high. Moreover, the anatomical interpretation of detected changes, of course, still would have to be done either by a medical expert or by an appropriate segmentation procedure.

position within the chapter's overall context. Moreover, each of these subchapters is completed by a summary of the main results and of the resulting conclusions with regard to high-accuracy morphometry.

It is explicitly pointed out again this chapter is written in a style of being of benefit not only to the image processing and –analysis community, but also for medical professionals, since the problem of acquisition-related morphological variability is of relevance for both communities. Accordingly, mathematical and physical principles of the discussed topics are described in such a way that they are comprehensible for both communities.

3.2 Understanding Acquisition-Related Morphological Variability in MRI

3.2.1 Introduction

As shown in the previous chapter, significant effort has been spent over the past decade in order to develop innovative image processing and -analysis algorithms and to improve existing methods in terms of accuracy, reproducibility and computational efficiency. Surprisingly, however, relatively little research was undertaken to find out to what extent the validity of results obtained with these methods is limited by inherent imperfections of the input images. This observation is especially true for MRI-based high-accuracy morphometry like early detection of AD, which was shown to require the reliable detection of very subtle changes throughout the whole brain despite the fact that MRI images are geometrically distorted.

Since a quantitative investigation of the impact that imperfections in the MRI data acquisition process might have on the validity of image-based morphometric results requires a profound knowledge of the nature of these imperfections, this subchapter introduces the different sources of geometrical distortions in MRI and the way they affect an imaged object's morphology in image space. Additionally, special emphasis is put on the characterization of these effects' influences on longitudinal image analyses taking into consideration imperfections in patient repositioning. Finally, it is pointed at the phenomenon of natural morphological variability, e.g. caused by normal variations in a subject's water balance.

3.2.2 Nature of Geometrical Distortions in MRI[6]

When talking about geometrical distortions in MRI, it is advisable to distinguish between object-dependent and –independent sources: Whereas the former can be – in a physically inexact attempt to simplify the understanding of their effects – considered stationary in the object's coordinate system, the latter would have to be thought of as stationary in scanner-based coordinates.

3.2.2.1 Object-Independent Sources

The major object-independent sources of geometrical distortions are static field inhomogeneities, gradient nonlinearities, eddy current induced dynamic field

[6] Since state-of-the-art MRI-based morphometry is done using 3D image acquisition techniques only, the delineation of sources of geometrical distortions is restricted to those relevant for these acquisition schemes. 2D sequences are additionally affected because field inhomogeneities (no matter whether static or dynamic) corrupt the position, shape and thickness of the excited slice.

distortions, and imperfections in the gradient amplifiers' calibration. In contrast, resonance frequency drifts during scan, fluctuations in power supply, and concomitant Maxwell field contributions do not noticeably affect an object's morphology when acquiring images at a spatial resolution lower than 0.25 mm^3. As this resolution is not expected to be increased within the next years, these effects are not further described here; this section is rather focused on those influences that have a practical impact on clinical MRI today.

– **Static field inhomogeneities** are practically unavoidable in the construction of whole-body magnets. As they are also influenced by the scanner's local environment, manufacturers adjust the homogeneity on-site (a procedure typically referred to as "passive shimming") but residual imperfections remain. Static field inhomogeneities cause distortions in frequency-encoding direction[7], as the spins' precession-frequency *f* depends linearly on the local magnetic field strength *B* (see equation 3.1; the gyromagnetic ratio γ is a characteristic of the object's material). Practically, local differences in *B* entail dislocations *d* which scale inversely with the imaging protocol's readout pixel bandwidth *BW*, as illustrated in equation 3.2 (distortions are given in pixels). Obviously, distortions arising from static field inhomogeneities are indistinguishable from those caused by local differences in magnetic susceptibility (see below), as these also manifest themselves as local changes of the magnetic field. Since - from an image processing point of view - it is sufficient to know about the remaining field inhomogeneities after "active shimming" (see below), the order of magnitude of static field inhomogeneities before and after "passive shimming" is not given here.

$$2 \cdot \pi \cdot \Delta f(\vec{x}) = \Delta \omega(\vec{x}) = \gamma \cdot \Delta B(\vec{x}) \qquad (3.1)$$

$$\Delta d(\vec{x}) = \frac{\gamma \cdot \Delta B(\vec{x})}{2 \cdot \pi \cdot BW} \qquad (3.2)$$

– **Eddy currents** are induced in conductive material within the scanner every time a gradient amplitude changes in the course of the scanning process. These eddy currents cause dynamic magnetic fields that – known from Lenz's law – always oppose their origin. As shown in figure 3.1, an important characteristic in the analysis of eddy currents is their temporal decay: Even high-amplitude gradient pulses with a minimum rise time do not necessarily have to be highly affected by eddy currents if the time constant of the corresponding eddy currents is short compared to the gradient pulse's duration. Quite obviously, distortions caused by eddy currents are dependent on physical scanner properties as well as on the actual imaging sequence and its parameterization, e.g. they scale linearly with the change of the gradient amplitude. In typical clinical three-dimensional structural datasets acquired with state-of-the art scanners, eddy current induced geometrical distortions maximally are in the medium sub-millimeter range (see section 5.2).

[7] For those not familiar with the basics of MRI data acquisition: Spatial encoding of the MRI signal is achieved by application of orthogonal magnetic field gradients during data acquisition. In 3D acquisition schemes, one gradient is used to set the spins' frequency ("frequency encoding" or "readout" axis), and the other two to set their phase ("phase encoding" and "partition encoding" axis, respectively). The images are then reconstructed by 3D Fourier transformation.

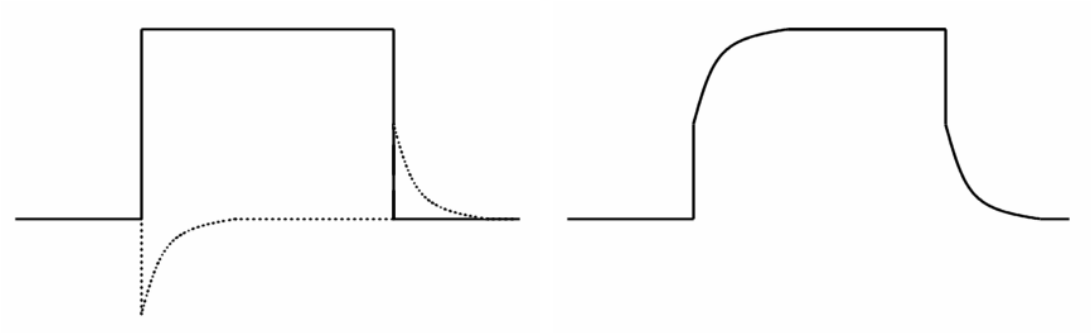


Figure 3.1: *Illustration of the effect of field distortions caused by eddy currents on the shape of an ideal rectangular gradient pulse. Left: dynamic field changes resulting from eddy currents (dotted lines) oppose the gradient pulse changes they are caused by (solid lines); Right: the effective gradient pulse is the superposition of the originally planned pulse and the field distortions caused by eddy currents*

– **Gradient coil nonlinearities** cause the signal to be spatially encoded in a curved rather than in Euclidean space. Since gradient coils are designed to be practically perfectly linear around the magnet's isocenter, nonlinearities are stronger in outer regions, and such are the corresponding geometrical distortions. Apparently, geometrical errors introduced by gradient nonlinearities are an intrinsic characteristic of the actual gradient coils' design and, thus, are in principle known to the manufacturers. As they are also independent of the imaged object, the magnetic field strength and the imaging protocol, they can be "universally" compensated for. In a subvolume relevant for whole-brain morphometry, distortions caused by gradient coil nonlinearities are in the range of millimeters, but they can vary significantly between different scanners (see section 3.3).

– **Gradient amplifier calibration** errors lead to scaling errors in each gradient direction. Typically, the calibration consists of an adjustment of the gradient drivers such that a dedicated phantom's geometry measured in image space matches its geometrical specification known from phantom construction. The underlying algorithm works very accurately such that the calibration procedure's accuracy is rather limited by the tolerance in phantom fabrication. As illustrated in figure 3.2, imperfections in gradient amplifier calibration lead to a systematic scaling error separate for each gradient direction. With typical state-of-the art scanners, these errors are – in a subvolume relevant for whole-brain morphometry - in a low sub-millimeter range.

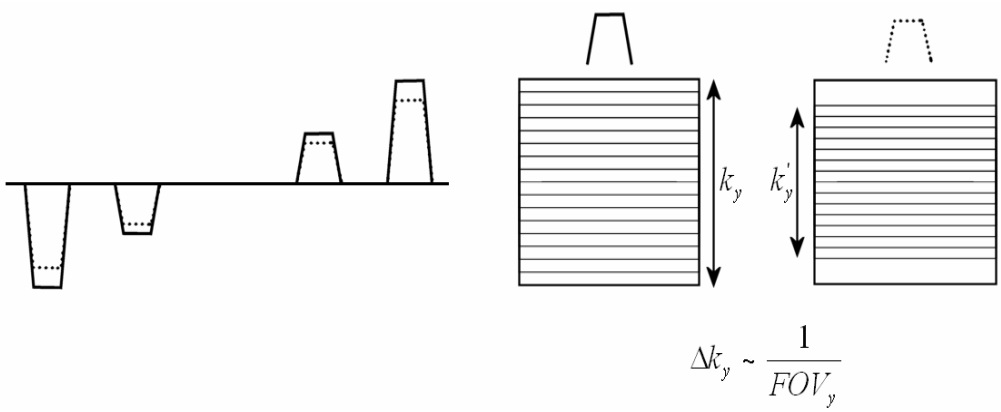


Figure 3.2: *Illustration of the effect of incorrect gradient calibration on an imaged object's geometry in image space (the planned gradient pulses are drawn in solid lines, the effective gradient pulses are drawn in dotted lines): as a result of the smaller k-space sampling interval in y-direction, the field of view in this direction is larger, and the object appears compressed in this direction*

3.2.2.2 Object-Dependent Sources

The most important object-dependent sources of geometrical distortions in MRI are magnetic susceptiblity artefacts, chemical shift, and – apart from phantom imaging – patient movements during scanning.

– **Magnetic susceptibility** differences between spatially neighbouring materials cause local magnetic field inhomogeneities which are especially prominent at the object's interface to air. Like static field inhomogenities, these lead to geometrical distortions in frequency-encoding direction the magnitude of which in image space is linearly dependent on the imaging protocol's bandwidth[8]. Probably even more important is their linear dependence on the magnetic field strength (see equation 3.3; χ_n being the magnetic susceptibility of tissue n) which means that susceptibility artefacts are twice as large for 3T as for 1.5T scanners (provided that the same imaging protocol is used). Note also that not necessarily just a patient's outer shape, but also his respectively her inner structure can be affected by susceptibility artefacts due to air filled cavities like the paranasal sinus etc.. Practically, susceptibility induced field distortions are reduced by a subject specific shim adjustment procedure ("active shimming"). This consists of an automatic determination of the field homogeneity (with the imaged object in the scanner) immediately before the scanning process starts. This magnetic field map is then used to automatically adjust the current setting of the dedicated shimming coils in order to minimize inhomogeneities. After "active shimming", the off-resonance frequency Δf at 1.5T can be up to 80 Hz at the frontal sinus, and thus (applying equation 4.2) lead to a dislocation $\Delta d = \Delta f/BW$ in a noticeable subpixel range when using structural imaging protocols with a typical readout bandwidth BW of 100-300 Hz/pixel.

$$\Delta B(\vec{x}) = \left(\chi_1(\vec{x}) - \chi_2(\vec{x})\right) \cdot B(\vec{x}) \qquad (3.3)$$

[8] Reducing distortions resulting from local susceptibility gradients by increasing the readout bandwidth, however, compromises the signal-to-noise ratio (SNR). High bandwidth multi-echo acquisition schemes, therefore, seem to be a promising approach for morphometry, since they reduce susceptibility artefacts, while maintaining high SNR by combining the individual images [Fischl04]. It should be kept in mind, however, that a higher bandwidth also means stronger eddy currents.

– **Patient movements during scanning** also corrupt the geometrical correctness of the patient's morphology. Apparently, movements can result from a patient's inability to keep perfectly still, but also from rather unavoidable physiological processes like CSF pulsation or patient breathing. Note that the different types of movements manifest themselves differently in image space depending on their magnitude and time course. Whereas periodic small-scale movements like CSF pulsation lead to a blurring of tissue borders, aperiodic low frequency movements like the typical "slow nodding" of the head in brain imaging introduce morphological errors in image space.

– **Chemical shift** between different tissue types arises because of different electron configurations of hydrogen in these substances. Electrons surrounding the nucleus also interact with the magnetic field, which causes a slight change in the proton spins' precession frequency. For materials with different electron configurations, this results in relative dislocations in frequency encoding direction, the magnitude of which again scales linearly with the magnetic field strength. In brain imaging, the chemical shift is especially prominent between subcutaneous fatty tissue and its surroundings. At 1.5 Tesla the fat-water shift is about 220 Hz, which leads to a relative dislocation in the range of a pixel for typical structural imaging protocols.

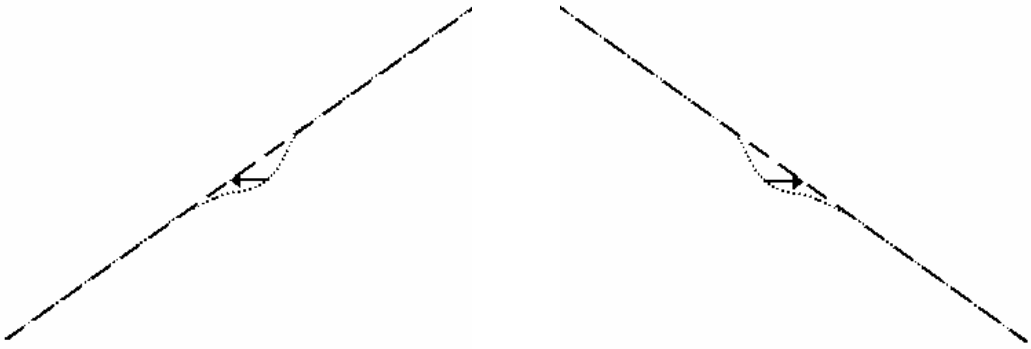


Figure 3.3: *Geometric distortions due to differences in magnetic susceptibility in dependence on the readout gradient; the readout gradient is represented by the dashed line, the effective gradient (readout gradient plus local susceptibility influences) is visualized as the dotted line; the arrow indicates a shift in image space (in frequency encoding direction) due to local differences in magnetic susceptibility*

3.2.3 Implications for Longitudinal Imaging

Whereas exact image-based measuring of geometrical properties of anatomical structures requires a correction for all sources of acquisition-related geometrical distortions, no matter their physical characteristics, perfect longitudinal morphometry in theory would be possible also if these distortions were identical in the acquired data. Unfortunately, however, this is merely a fictitious advantage even for perfectly stable scanners, as this identity of distortions would have to be given in patient-based coordinates. Thinking in patient-based coordinates, however, imperfections in patient repositioning automatically imply a mismatch between the acquired datasets in the scanner-based coordinate systems. Consequently, object-independent distortions would be different in these images. This is the reason why it is important in longitudinal imaging to establish a patient positioning protocol that aims at optimizing the reproducibility of patient positioning.

Neglecting the rather unsystematic patient movements, one can – for longitudinal imaging - classify the different sources of geometrical distortions in three different

groups depending on their effects on the object's morphology in image space: gradient nonlinearities and eddy currents result in relative nonlinear deformations, gradient amplifier calibration imperfections cause a relative shearing transformation between longitudinal datasets, and differences in magnetic susceptibility and static field inhomogeneities lead to relative geometrical differences like those illustrated in figure 3.4.

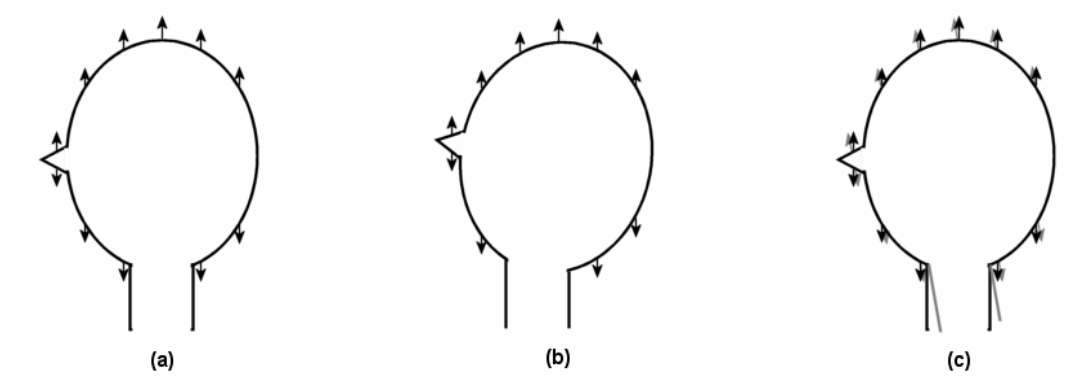


Figure 3.4: *Illustration of the effect of geometrical distortions (indicated by arrows) resulting from local susceptibility gradients on morphological consistency in longitudinal MRI: (a) distortions in the baseline image, (b) distortions in a follow-up image, (c) relative morphological changes (positional differences between black and grey arrows) after a linear registration of (b) to (a). Due to imperfections in patient repositioning, the susceptibility artifacts no longer point in the same direction in patient-based coordinates (but only in the scanner's coordinate system). Practically, the distortions will not even be perfectly identical in magnitude due to imperfections in the shimming procedure. Note that these effects are not eliminated by preprocessing steps like skull-stripping because there are also local differences in susceptibility in the interior parts of the head.*

Additional to these purely repositioning-related effects, imperfect active shimming reproducibility, regular scanner maintenance activities, and imperfect long-term hardware stability also introduce relative morphological changes into longitudinal MRI data of a physically unchanged object. Transferred to a clinical setting, the sum of all these potential error sources represent the overall acquisition-related morphological variability superimposed onto the pathological process's manifestation in image space. The question to what extent these influences affect the detection limit of morphometric applications aiming at detecting subtle changes between longitudinal MRI data is further gone into in section 3.3.

3.2.4 Natural Morphological Variability

All sources of morphological variability presented so far are characteristics of the data acquisition process in the presence of imperfections in object repositioning. It should be kept in mind, however, that every living subject's morphology underlies a certain amount of natural variability that is not necessarily a symptom of pathology. Practically, there are physiological processes beyond ageing or loss, respectively gain of weight that can lead to noticeable changes in a patient's morphology. Probably the most important one of these "confounders" in brain imaging is the variability in the patients' water balance (see figure 3.5).

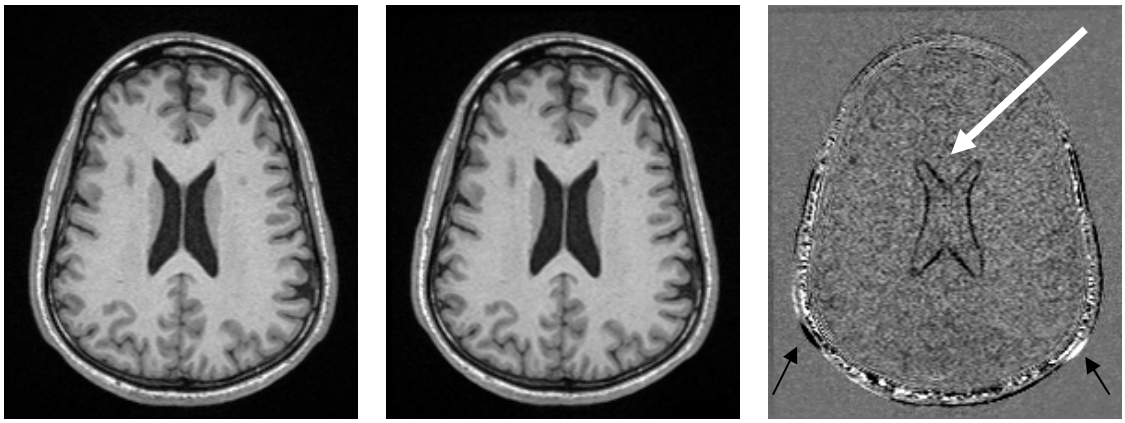


Figure 3.5: *Morphological variability due to differences in a person's water balance: A human volunteer was scanned on two consecutive days on the same scanner with identical protocols (left: baseline scan, middle: linearly registered follow-up scan, right: difference image). Between the two scans dehydration was induced by light workout activity without any uptake of liquid. The black arrows indicate changes due to different positions of the head relative to the dorsal pads used to increase patient comfort. More importantly, one can observe a significant volumetric change of the ventricles pointed at by the white arrow (note that for registration only brain voxels were used in order not to be confounded by differences in the positions of the eyes or the pads).*

3.2.5 Discussion

Even though a lot of effort has been spent over the past decades to develop innovative image processing and –analysis algorithms and to improve existing methods in terms of accuracy, reproducibility and computational efficiency, relatively few research was undertaken to find out to what extent the validity of results obtained with these methods is limited by inherent imperfections of the input images. In fact, validation studies in image processing and -analysis largely consist of comparisons of the performance of an algorithm to results produced by medical experts in an interactive manner. In this subchapter, it was shown qualitatively that in MRI the assessment of an application's clinical validity needs to go beyond this standard image analysis validation-strategy and should additionally incorporate the data acquisition process as well as patient-positioning issues.

In practice, this was shown by separately describing each source of geometrical distortions in MRI relevant for image-based morphometry, including a detailed characterization of the way this source affects an imaged object's morphology in image space. Accurate image-based measurements of geometrical properties of anatomical structures obviously require correcting for all the outlined sources of geometrical distortions.

Special emphasis was put on the exploration of these phenomena's consequences on longitudinal imaging: As a result of imperfect reproducibility in patient positioning, acquisition-determined geometrical distortions are not perfectly reproduced in the patient-based coordinate system, which inevitably leads to morphological changes in image space even for a physically unchanged object that is imaged with a perfectly stable scanner. Along with variability caused by imperfect hardware stability and by differences in the patient's motion during scanning, this morphological variability in image space was named acquisition-related morphological variability to indicate that it is related, but not determined by the data acquisition process. More importantly, it does not reflect an actual change in the patient's morphology.

Regarding the complexity of this variability, it can be stated that it corresponds to a nonlinear spatial transformation of an object's morphology in image space, whereby

the underlying deformation field is dependent on the scanner, the imaging sequence, its parameterization and the imaged object itself.

3.3 Quantification of Acquisition-Related Morphological Variability

3.3.1 Introduction

The previous section has shown that the MRI data acquisition process has a confounding effect on an imaged object's morphological representation in image space. Moreover, the considerations regarding the impact of these acquisition-determined geometrical distortions on morphological accuracy and reproducibility in longitudinal imaging scenarios have revealed that - in the presence of imperfections in object repositioning - they introduce morphological variability into image space even for an otherwise physically unchanged object. Along with variability caused by imperfect hardware stability and by differences in the patient's motion during scanning, this morphological variability in image space was named acquisition-related morphological variability.

Since this variability does not reflect an actual change in the patient's anatomy, but rather can be considered a superposition onto it in image space, acquisition-related morphological variability makes it more difficult for physicians to correctly interpret subtle changes detected in longitudinal MRI data by means of image processing and –analysis algorithms.

So far, however, the question to what extent acquisition-related morphological variability affects the detection limit of such morphometric applications has not been addressed yet; the outlined phenomena and underlying theory of MRI physics have been explored purely qualitatively. This chapter therefore aims at answering this question quantitatively. As a consequence of the overall geometrical distortions' dependency on the scanner, the pulse sequence, its parameterization, and even the imaged object itself, a profound quantification of acquisition-related morphological variability requires a detailed characterization of the way it is determined. Therefore, it is put much emphasis on the exact description of the methods used, including data acquisition, phantom characterization, and an elaborate delineation of the data analysis procedure.

3.3.2 Prior Research

Only little information concerning the confounding effect of the MRI data acquisition process on an imaged object's morphology in image space has been published: Jovicich et al. characterize distortions due to gradient coil nonlinearities by the change of a phantom's total diameter [Jovicich04a], and demonstrate the influence of 3D gradient nonlinearity correction offered by scanner manufacturers (see section 3.3.9) on volumetry of deep brain structures [Jovicich04b]. Wang et al. characterize acquisition-determined distortions by absolute positional deviations of a phantom's geometry in image space from its formal specification [Wang04a], and compare a

phantom-based method for 3D distortion correction to the performance of 2D gradient nonlinearity correction algorithms offered by MRI vendors[9] [Wang04b].

All these studies reveal that gradient coil nonlinearities can seriously affect an imaged object's geometry in image space: Averaging across the scanners investigated in [Jovicich04a], the mean error in the phantom's total diameter was 12.5mm (5% of 250mm). 3D gradient coil nonlinearity correction reduced this mean error to 1.0mm (0.4% of 250mm). Regarding volumetry of deep brain structures, 3D gradient nonlinearity correction improved the determined volumes on average by 3% [Jovicich04b]. The studies conducted by Wang et al. - not surprisingly - affirmed that geometrical distortions due to gradient coil nonlinearities differ strongly between different scanner types and showed that 2D gradient nonlinearity correction is not sufficient for high-accuracy morphometry.

Though being helpful to get an idea of the magnitude of distortions caused by gradient coil nonlinearities, these investigations are not sufficient to fully answer the question to what extent acquisition-related morphological variability affects the detection limit of MRI-based morphometric applications.

The latter inevitably requires to accurately determine the confounding effect of acquisition-geometrical distortions on morphological variability in image space in a long-term study, since such an analysis also has to account for potential imperfect hardware stability and for maintenance activities like scanner calibration or software updates. Otherwise the magnitude of the investigated effects could be underestimated, and the results could no longer be transferred to clinical practice.

Moreover, acquisition-related morphological variability has to be evaluated locally, but in the complete field of view relevant for whole brain morphometry, i.e. the inner volume of the head coil. Finally, it has to be assessed with respect to variability in position as well as in volume in order to be useful for clinicians[10]:

- The demand for local evaluation emerges from the fact that all relevant morphometric measures are determined locally as well (see chapter 2.2). A phantom's total diameter – the measure used in [Jovicich04b] - cannot adequately capture a phenomenon as complex as 3D nonlinear deformation field.
- The second requirement – analyzing the complete inner volume of the head coil – is motivated by the fact that acquisition-related morphological variability naturally affects the complete dataset. From a clinical point of view, it is therefore important to know about the magnitude of this variability in the complete field of view, and not only in a few deep brain regions like those analyzed in [Jovicich04b][11].

In total, these considerations reveal that it is necessary to go beyond the outlined prior research in order to be able to fully answer the question to what extent acquisition-related morphological variability affects the detection limit of MRI-based

[9] Software for 3D gradient nonlinearity correction seems to not have been available to the conductors of this study. However, it is product software by now and commercially offered by MRI vendors, which is why the comparison reported in [Wang04b] is not very meaningful any more.

[10] The studies conducted by Wang et al. characterize acquisition-determined distortions only in terms of absolute positional errors.

[11] As mentioned above, these deep brain structures are analyzed in [Jovicich04b] with regard to the effect of gradient nonlinearity correction only, anyway.

morphometric applications. Therefore, a new phantom-based method for characterization of site-specific geometrical distortions has been developed and used for quantitative analyses of acquisition-related morphological variability in a long-term study carried out on state-of-the art MRI scanners. This evaluation method is capable of determining positional and volumetric distortions on a dense 3D grid both in an absolute and in a relative way.

3.3.3 General Remarks on the Experimental Design

The considerations in the preceding section already have identified important requirements that a study on acquisition-related morphological variability has to meet in order to cover all aspects of this phenomenon, e.g. the necessity for it being a long-term investigation and the need for a local evaluation of acquisition-related morphological variability in the complete inner volume of the head coil. Even though all further requirements are rather obvious, they are listed in the remainder of this section for the sake of completeness.

In practice, it is necessary to analyze data from both phantoms as well as human volunteers. The major advantage of phantom experiments is that a phantom's true geometry can be determined at high accuracy, which allows for almost exact comparisons between the scanned object in image space and its true geometry. On the other hand, phantom experiments provide no further insight into patient-specific morphological variability, e.g. arising from susceptibility artifacts or from movements during scanning. Due to a missing ground truth, however, the analysis of human volunteer data can only be used to coarsely verify the results obtained by analyzing phantom data.

Moreover, an experimental study dedicated to this topic must not exclusively rely on one specific scanner type. Even within one product generation, different scanner series have different gradient coils, and thus, differ in gradient nonlinearities as well as in their behaviour in terms of eddy currents. When looking at different scanner generations even the magnets are different[12].

In order to meet all these requirements, a phantom and a healthy human volunteer (28 years old, male) have been scanned on a weekly basis for 6 months, alternating between a 1.5T Siemens Magnetom Avanto and a 3T Siemens Magnetom Trio.

Finally, it is essential for this study's acceptance to the image processing as well as to the medical research community to be able to guarantee that its results are not confounded by the data evaluation procedures. For the phantom experiments this implies the necessity for a thorough validation of the method used for analyzing the phantom data. For the evaluation of the head image series, this goal is most promisingly achieved by making use of a software package that already is accepted by the neuroimaging community as a valid tool.

[12] Ideally, a study on acquisition-related morphological variability includes MRI scanners from all vendors, but as this thesis was strongly supported by Siemens Medical Solutions, and practically carried out at Siemens MRI in Erlangen, it was impossible to include scanners from other vendors in this study.

3.3.4 Data Acquisition

The imaging sequences and protocols have been selected by first analyzing key publications in the field of morphometry, among others [Thompson03][Fischl99a][Davatzikos01], and subsequently discussing their usefulness with the neuroimaging sequence-developers at Siemens MR in Erlangen: Finally, 3DFLASH (3D Fast Low-Angle SHot) and MPRAGE (3D Magnetization-Prepared RApid Gradient Echo) were identified as the most commonly used imaging sequences for structural brain imaging[13]. The protocols are listed in detail in table 3.1 and 3.2.

As a consequence of the previously outlined considerations regarding the active shimming procedure's reproducibility in longitudinal imaging scenarios (see figure 3.4 in section 3.2), special attention has been paid to shimming in the acquisition of phantom data. In order to make sure that this reproducibility is not significantly worse for the phantom compared to the volunteer's head (due to the formers different internal structure), active shimming has been iteratively repeated in phantom imaging[14]. Practically, this iterative shimming has revealed that the phantom used in this thesis (described in the next chapter) is comparable to a human head in terms of reproducibility of active shimming[15].

[13] In this thesis, the "Siemens-terminology" is used to describe the pulse sequences. A brief introduction into the underlying physical principles as well as an overview of the different MRI vendors' terminology is given in [Nitz99]. For example, General Electric uses the acronym SPGR (Fast Spoiled Gradient Refocused Acquisition into steady state), and Philips the acronym T1-FFE (T1-weighted Fast Field Echo) for the imaging technique referred to by Siemens as FLASH (Fast Low Angle SHot).

[14] This so called "interactive shimming" is a product feature implemented at Siemens scanners to allow for a simple quality assurance of the fully automated active shimming procedure's results. In detail, the user is reported both the current setting of the coils used for active shimming as well as the previous setting, such that it is very easy to check whether the shimming procedure converges or shows a rather unstable behaviour. Moreover, this feature allows starting the active shimming from scratch such that it can be quickly assessed whether the shimming always converges to the same setting or not.

[15] Of course, the shim-coils' settings are completely different for the phantom compared to the human head, but their convergence-behaviour as outlined in footnote 14 is very similar.

	3DFLASH	MPRAGE
Acquisition Time	6:04	7:32
TR (ms)	11	2300
TI (ms)	-	1100
TE (ms)	4.94	3.93
Flip Angle (deg)	15	12
Slices	192	192
FoVr (mm)	256	256
FoVp (%)	87.5	87.5
Slice Thickness (mm)	1.0	1.0
Slice Resolution (%)	100	100
Slice Partial Fourier	7/8	7/8
Slice Oversampling (%)	0	0
Phase Resolution (%)	100	100
Phase Partial Fourier	7/8	7/8
Matrix Size	224 x 256	224 x 256
Bandwidth (Hz/Pixel)	140	130
Orientation	Sagittal	Sagittal
Flow Compensation	Off	Off
Asymmetric Echoes	Off	Allowed
iPAT	Off	Off

Table 3.1: *Protocols for data acquisition on the 1.5T Siemens Magnetom Avanto*

	3DFLASH	MPRAGE
Acquisition Time	6:53	8:43
TR (ms)	11	2300
TI (ms)	-	1100
TE (ms)	6.42	3.93
Flip Angle (deg)	15	12
Slices	192	192
FoVr (mm)	256	256
FoVp (%)	100	87.5
Slice Thickness (mm)	1.0	1.0
Slice Resolution (%)	100	100
Slice Partial Fourier	7/8	Off
Slice Oversampling (%)	0	0
Phase Resolution (%)	100	100
Phase Partial Fourier	7/8	7/8
Matrix Size	256 x 256	224 x 256
Bandwidth (Hz/Pixel)	170	130
Orientation	Sagittal	Sagittal
Flow Compensation	Read	Off
Asymmetric Echoes	Off	Allowed
iPAT	Off	Off

Table 3.2: *Protocols for measurements on the 3T Siemens Magnetom Trio*

3.3.5 Phantom Description and Analysis

The phantom used in this long-term study is cylindrically shaped (diameter 250 mm) with an integrated 3D grid of small spherical holes (approximately 4 mm radius) each having a distance of 20 mm to its nearest neighbour in x- and y-, and 19 mm in z-direction (979 spheres in total). These holes are connected in z-direction by cylindrical bars (approximately 2 mm radius), and the complete system of cavities is filled with MR visible fluid.

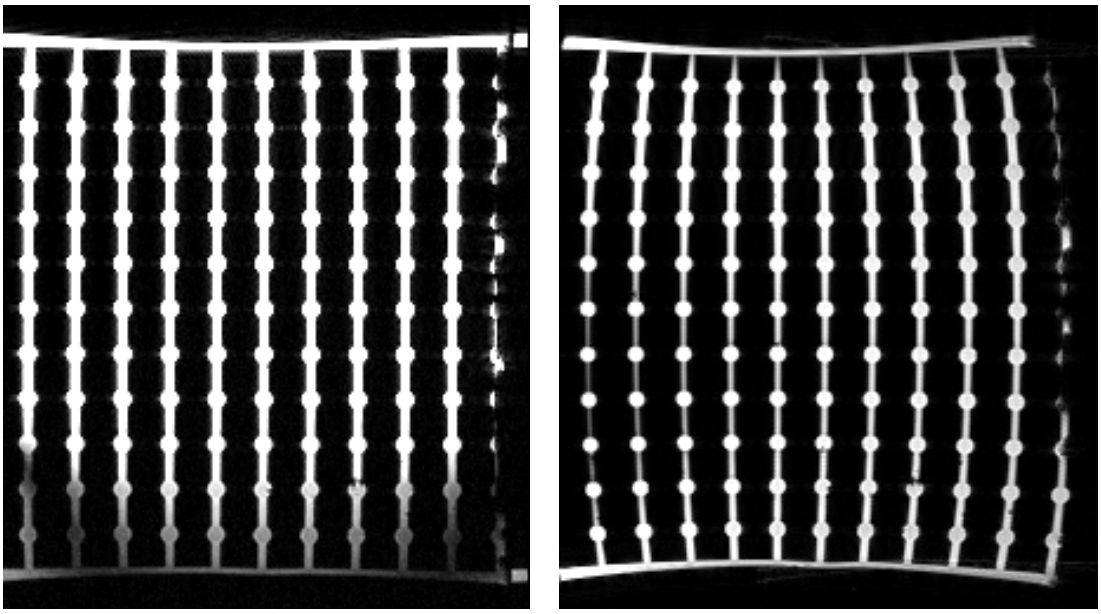


Figure 3.6: *Two slices of a phantom dataset acquired with the 1.5T Siemens Magnetom Avanto (left) and 3T Siemens Magnetom Trio (right). The interruptions in the bright bars are caused by inclusions of air. The intensity drop towards the lower end in the Avanto image arises from the fact that the phantom does not completely fit into the head coil. The difficulties encountered when attempting to separate spheres from bars are described in the text.*

Unfortunately, the phantom is perfectly symmetric which makes a fully automated evaluation impossible. The investigator has to know the phantom's position in the scanner, as there are several geometrical constellations that are indistinguishable when only evaluating the MR images.

Moreover, the phantom, though fitting perfectly inside the Trio head coil, does not completely fit inside the Avanto head coil due to this coil's converging cranial end. The 3T multi-channel head coils, that provide a far better image contrast, could not be used at all with this phantom since these generally are a lot narrower.

The most severe problems, however, become evident when taking a closer look at figure 3.6:

- The spherical elements' radius (4 mm) exceeds the bars' radius (2 mm) by only 2 voxels in a dataset with an the protocols specified in section 3.3.4.
- There is a smooth transition between spheres and bars, which additionally complicates their automatic segmentation.
- The distortions inherently change the shape of the spheres and bars in image space, so that it is not always trivial to identify the exact border between them.
- Air inclusions produce unpredictable variations in the spheres' and bars' shape such that not all spheres can be reliably detected in every dataset.

3.3.6 Phantom Data Evaluation

3.3.6.1 Analysis of Target Objects

The difficulties mentioned above make it impossible to separate the spheres from the interconnecting bars by a simple combination of morphological operations. Moreover, the idea to explicitly search for a reference model representation of these objects in the images is not the method of choice as the target objects' geometry slightly varies even within one dataset (see figure 3.7). As a direct consequence of the previously outlined imperfections in the MRI data acquisition process, the phantom's spheres are transformed to ellipsoids in image space, whereby their eccentricity is dependent on their spatial location. Furthermore, a closer look at these interconnected ellipsoids reveals that they show an inner structure as a result of Gibbs ringing, which is caused by incomplete digitization of the MRI-echo. Unfortunately, this inner pattern is not identical for all the ellipsoids since their position relative to the spatial discretization grid slightly varies even within one dataset.

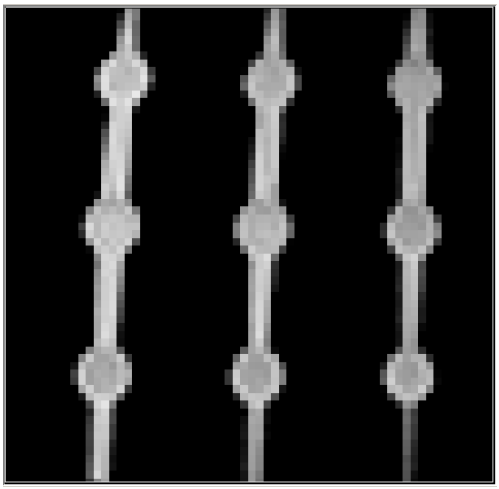


Figure 3.7: *As a direct consequence of imperfections in the MRI data acquisition process, the phantom's spheres are transformed to ellipsoids in image space. Their inner structure is a result of Gibbs ringing.*

3.3.6.2 Identification of Target Objects

From an abstract point of view, the algorithm for separating the ellipsoids from the bars they are connected by can be divided into two passes, the first one being a fast and robust identification of ellipsoids that grow out of cylinders, the second one being a computationally more demanding determination of these ellipsoids' centre points.

```
smooth image;
for each slice{
  binarize via Otsu separation [Otsu79];
  identify objects belonging to an ellipsoid;
}
do 3D connected component analysis;
```

Algorithm 3.1: *Pseudo-code representation of the algorithm for separating the ellipsoids from the bars they are connected by.*

For computational efficiency the initial steps of the first pass are 2D operations. The separation between bars and ellipsoids is done by eliminating all those pixels that are not part of a 2D object

- the bounding box of which is smaller than a user defined minimum (or larger than a user defined maximum)

- that consists of less than a user defined minimal number of pixels

These blobs are then filtered by position to eliminate irregular ellipsoids occurring as a result of wrap-around artefacts, and by shape (skewness, 3rd order moments) to eliminate degenerated objects resulting from air inclusions.

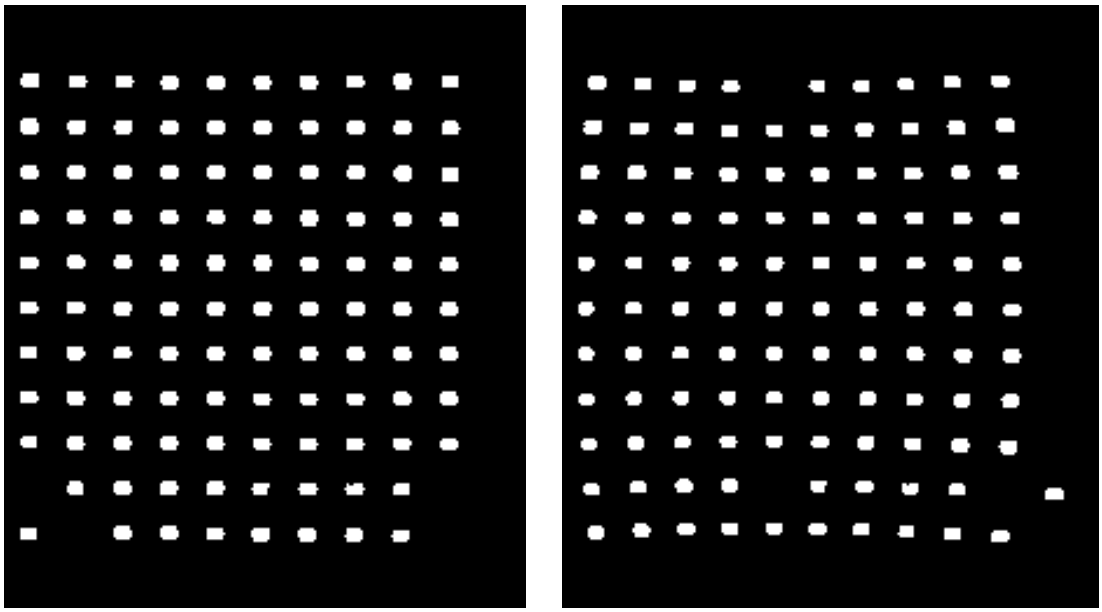


Figure 3.8: *Identified ellipsoids (identical slices as in figure 3.7)*

Figure 3.8 shows that this first pass robustly separates the ellipsoids from the bars they are connected by (all eliminated blobs showed significant asymmetries when analyzing them in 3D). Taking a closer look at the results, however, it becomes evident that in very few cases even adjacent blobs stretch across a different number of slices, which obviously does not correspond to the phantom's correct geometry.

The results are therefore refined in a second pass that consists of a computationally more demanding determination of the remaining ellipsoids' centre points (as illustrated in figures 3.9). Its principal idea is to determine the point-symmetry centre for each ellipsoid in the original image. To ease the gradient descent optimization strategy, candidate centre points need to be located on the interconnecting bar's centre line, which is determined via regression analysis. As a metric for symmetry the negative summed squared difference between each voxel and its mirrored counterpart (resulting from 3rd order B-spline interpolation) is used.

```
for each remaining blob{
  extract a small subimage (bounding box + margin);
  for each slice in subimage{
    determine intensity-weighted center (centerX,
          centerY) using the smoothed original image;
  }
  do regression analysis of centerX, centerY against z;
  determine that point on the regression axes for which
        the point-symmetry of the subimage is maximal;
  store this point as ellipsoid's center;
}
```

Algorithm 3.2: *Pseudo-code representation of the algorithm for determination of the ellipsoids' centre points.*

This refinement reliably eliminates all instabilities previously hinted at resulting from binarizing and spatially discretizing the target objects. In summary, the application of

the outlined algorithms 3.1 and 3.2 reduces a phantom dataset to a point set *G1*, whereby each point represents the centre of one of the phantom's spheres.

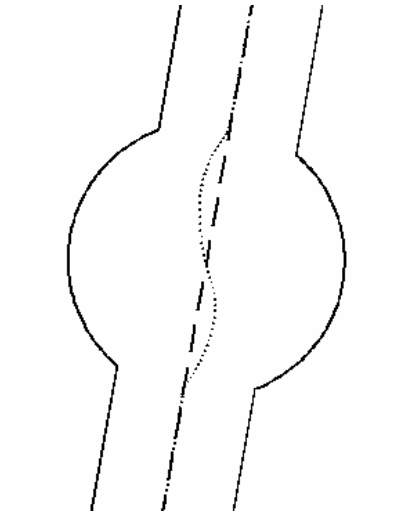


Figure 3.9: *Maximum symmetry detection along the regression axis (dashed line); the dotted line corresponds to the position of each slice's intensity weighted centre.*

3.3.6.3 Comparison to Reference Geometry

Since it was not known how well the phantom fitted to its formal specification, a new reference phantom model *G0* was created from the analysis of a CT scan (Siemens Somatom Sensation 64) which is generally accepted as being unaffected from geometrical distortions.

Due to the symmetric geometry of the phantom, the comparison of the determined point set *G1* against the reference geometry *G0* cannot be done fully automatically. The comparison therefore also is a two-step procedure, the first one solving ambiguities in phantom orientation and the second one carrying out the actual comparison.

Practically, the user first has to manually identify *n* homologous point-pairs (*n>2*) by assigning to *n* of the determined centre points the label of the corresponding sphere as it is named in the reference geometry *G0*. The point set *G1* is then rigidly transformed to best match *G0* by minimizing the mean squared error between the positions of the manually depicted homologous points. *G1* now is completely labeled by assigning each centre position the label of its nearest neighbour in *G1*, and no geometrical ambiguities remain.

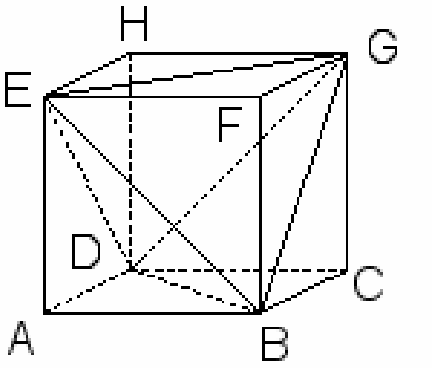
```
load centre positions detected in step1: G1;
load phantom's reference geometry: G0;
manually identify homologous points (n>2): P;
rigidly transform G1 to best match G0 using P;
for each point in G1{
     determine nearest neighbour in G0;
     transfer label;
}
rigidly transform G1 using central point-pairs;
for each point in G1{
     calculate geometrical deviation;
}
for each cell in G1{
     calculate volumetric deviation;
}
```

Algorithm 3.3: *Pseudo-code representation of the algorithm for comparing the determined grid G1 against the phantom's reference geometry G0.*

After being labeled, *G1* is transformed for a second time to best match *G0*, this time only taking the central 27 point-pairs into account, as these are least affected by geometrical distortions, since they are closest to the magnet's isocentre.

The resulting point set *G1* is then compared to the previously mentioned reference phantom model *G0*:

- The **positional comparison** is done straightforward and merely consists of a calculation of the absolute deviation in position for each sphere compared to its position according to the reference model *G0*.
- For **volumetric analysis** *G1* is subdivided into polyhedral cells by iteratively grouping eight adjacent centre points together in order to form one cell, the shape of which can be intuitively described as a "degenerated cuboid". For determination of the detected cells' volumes, the polyhedra are further divided into tetrahedra, the volumes of which can be easily calculated; finally, each cell's volume is compared to the volume of the corresponding cell in *G0*.



$$V = \frac{1}{6} \cdot \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \end{pmatrix}$$

$$\text{vertex n : = } \vec{x}_n = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}$$

Figure 3.10: *Division of a cube into 5 tetrahedra (ABDE, BCDG, BEFG, DEGH, BDEG) and formula for calculation of a single tetrahedron's volume [Jin93].*

3.3.6.4 Visualization

2D vector plots are generated for all three principal orientations (sagittal, coronal, axial) to ease visual inspection of the determined distortions. These vector plots do not only provide the viewer with an intuitive impression of the distortions' patterns, but also automatically reveal outliers, as geometrical distortions in MRI generally are relatively smooth.

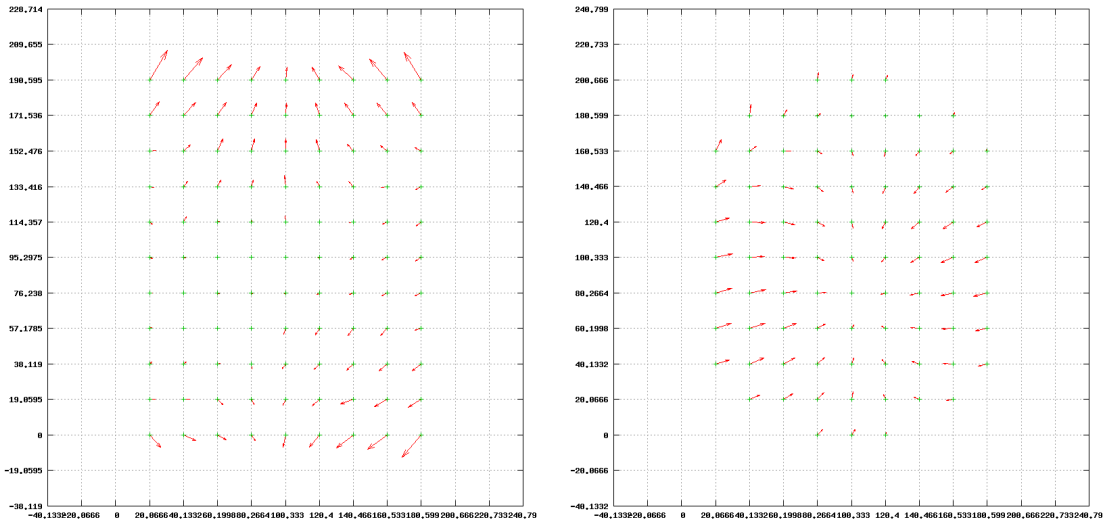


Figure 3.11: *Geometrical distortions (Siemens Magnetom Avanto, MPRAGE) visualized in a coronal (left) and axial (right) 2D vectorplot: green dots symbolize the correct position, red arrows point at the positions where the corresponding spheres were detected at in the MR image. For visualization purposes, the arrows are scaled with factor 10.*

These data are further analyzed with respect to the centre points respectively cells distance to the magnet's isocentre, resulting in diagrams visualizing the statistical distribution of the determined measures for spherical layers around the isocentre (for an example, see figure 3.14).

3.3.6.5 General Limitations of the Phantom Evaluation Method

The outlined method for quantification of geometrical distortions has got one major limitation when taking a more general look at it:
The labeling of the identified blobs by first identifying a few homologous point pairs, then rigidly transforming the complete point list *G1* to best match *G0* and finally transferring the label of each blob's nearest neighbour is error prone if geometrical distortions are as large as half the grid's spacing. If distortions are even larger, the proposed method will completely fail in these regions. However, this cannot happen when examining head images acquired with a head coil positioned in a reasonable distance to the iscocentre, as in these cases distortions are smaller than 19mm · 0.5 = 9.5mm. The proposed phantom analysis method, thus, is applicable to all phantom datasets that are acquired to reflect clinical practice in MRI neuroimaging.

3.3.6.6 Validation - Accuracy

Usually, phantom experiments allow for an experimental determination of an image analysis algorithm's accuracy, namely by comparing the algorithm's results to the actual values known from phantom specification. However, this validation strategy is based upon the assumption that the phantom – or at least its essential characteristic that is to be quantified by the algorithm – is unalteredly mapped into image space by data acquisition. Transferred to the phantom evaluation method outlined in the preceding sections, this means that the spheres' shape must not be modified by the imaging process. As this postulation is clearly violated in case of MRI, it is not possible to experimentally determine the outlined phantom evaluation method's accuracy with respect to the maximum error that can occur in the determination of the

spheres' centre points[16]. This is why the method's accuracy is investigated solely theoretically.

First of all, it is advisable to differentiate between the grid's directions: The accuracy orthogonal to the bars' direction is limited by noise and partial volume effects only. Potential errors resulting from these sources are minimized in the second pass of the target-object identification step by smoothing the input image and by subsequently determining each slice's centre of the extracted subimage by means of intensity weighted averaging.

As shown in figure 3.12, between 21 and 69 voxels contribute to the computation of each slice's centre (diameter range: 4-8mm; spatial resolution: 1mm isotropic). Taking into consideration the high CNR between the bright spheres and their dark surrounding, the maximum in-plane error therefore can be assumed to be less than 1/10 of the voxel spacing.

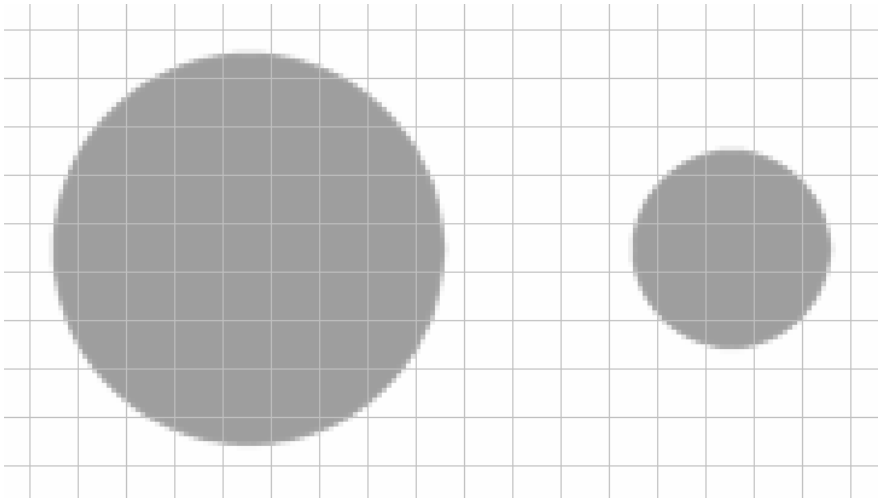


Figure 3.12: *Illustration of how many voxels contribute to the computation of each slice's intensity weighted centre (left: slice representing a sphere's centre; right: slice representing a bar).*

When looking at the accuracy of the proposed method in direction of the bars, it is helpful to look at each processing step individually: In section 3.3.6.2 it already has been mentioned that after the algorithm's first pass (sphere identification) few adjacent blobs stretch across a different number of slices, which does not correspond to the phantom's correct geometry. This error can be explained as an effect of the combination of binarization and subsequent spatial discretization of the detected ellipsoids: Slight differences in partial volume effects, and small variations in shape due to geometrical distortions and tiny air inclusions can lead to cases in which a 2D object is once considered as belonging to an ellipsoid, whereas it is considered a bar element in a adjacent ellipsoid, even though these segments nearly correspond to the same part of their target objects.

[16] In fact, the accuracy is different for each sphere because each sphere is differently affected by acquisition-determined geometrical distortions due to the difference in position in the scanner-based coordinate system. Consequently, the determination of the proposed algorithm's accuracy for "ideal" spheres connected by bars is not very meaningful. A realistic simulation of the impact of MRI data acquisition on the shape of the spheres on the other hand is almost impossible since this requires to adequately simulate all the sources of geometrical distortions described in chapter 3.2, including eddy currents and susceptibility artefacts.

These considerations explain the need for the outlined second pass. The maximum-symmetry detection described in section 3.3.6.2 eliminates this error. Moreover, a closer look at figure 3.7 reveals that this approach also is an appropriate way to handle significantly distorted spheres and Gibbs ringing, as these phenomena correspond to the assumption of point symmetry.

In summary, the accuracy of the phantom evaluation method can be assumed to be about 1/10 of the voxel spacing. Still, potential errors in the direction of the bars are slightly greater than the inplane error, because air inclusions tend to be predominantly located at the transition from spheres to bars.

3.3.6.7 Validation - Reproducibility

The phantom evaluation method's reproducibility can be determined by means of test-retest experiments: The phantom is thereby imaged twice with identical protocols and on identical scanners without repositioning and without any time gap between the scans. The only remaining differences, thus, are thermal expansion of the phantom due to the energy deposit during the first scan, slight changes in the scanner hardware due to heating, and signal noise in the acquired datasets.

Whereas the accuracy of the proposed phantom evaluation method is limited by noise and partial volume effects, the reproducibility of the detection of the spheres' centre points is additionally limited by temporal drifts in the scanners' resonance frequency during data acquisition and in between two scans. This error in frequency encoding results in a slightly varying discretization scheme in readout direction even within one data acquisition. This effect is best assessed by inspection of the difference image (see figure 3.13). It clearly shows a systematic pattern at object borders that looks similar to a shift in the direction of the bars. A detailed look at the difference images reveals that this effect is not exactly a shift in image space, but that there are minor variations in the difference patterns even between adjacent ellipsoids and, thus, relative morphological differences between these datasets. The magnitude of these differences is known to be small, but it cannot be quantified exactly, since it also depends on the heating of the scanner during data acquisition and on the cooling in between.

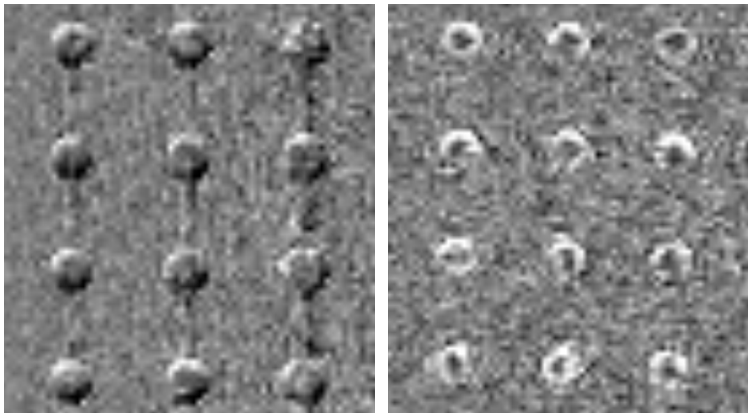


Figure 3.13: *Difference image between two datasets acquired in a test-retest experiment (left: sagittal view, right: axial view).*

The results of evaluating the MPRAGE datasets acquired with the Siemens Magnetom Avanto are shown in figure 3.14. For an image volume that is relevant for brain morphometry (spherical subvolume around the isocentre with a radius of 100mm), the maximal relative positional difference is less than 0.15mm, and the maximal relativ volume difference 0.35% (relative to the volume of a standard cell in *G0* of 7600 mm^3). The corresponding values for the Siemens Magnetom Trio are

0.11mm and 0.26%. The slightly better results for the Siemens Magnetom Trio most likely are a result of a higher signal-to-noise ratio at 3T.

Note that these numbers are the maximal errors in reproducibility of the complete experiment, consisting of data acquisition and evaluation. The reproducibility of the evaluation method alone is even higher, as there really are relative morphological changes in image space introduced by the scanning process.

The finding that relative geometrical distortions observed in this test-retest scenario are stronger for regions lying further away from the magnet's isocentre, corresponds well to scanner imperfections known to the manufacturers: Both the alternating mains voltage the scanner's gradients are (indirectly) supplied by and the movement of the cryopump corrupt the desired magnetic field. Effects from both error sources are known to increase linearly with the distance to the magnet's isocentre, a trend that is also evident in figure 3.14.

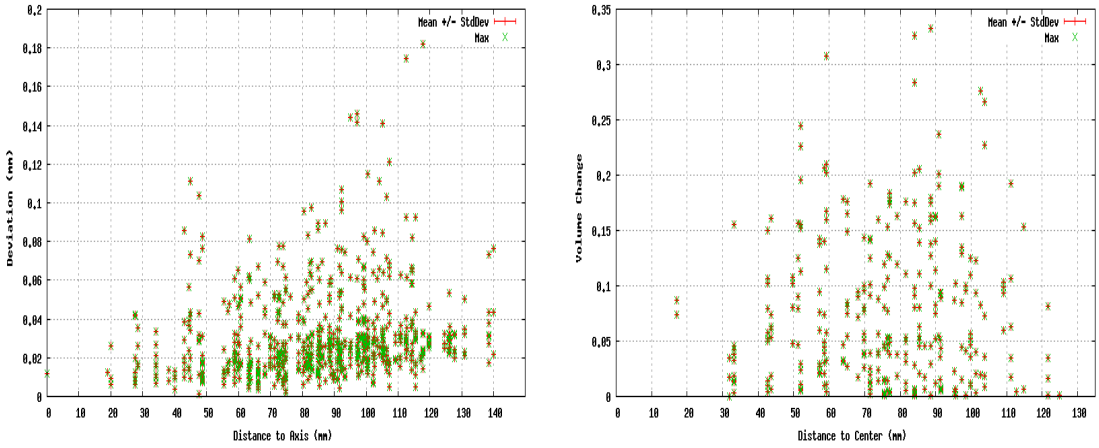


Figure 3.14: *Evaluation results: positional deviation of identical spheres (left), and relative volume change of identical cells (right) detected in two different datasets acquired with identical protocols in dependence of their distance to the isocentre (Siemens Magnetom Avanto, MPRAGE).*

3.3.6.8 Quantification of Repositioning Influences

Due to the nonlinear character of the distortion field, a different position of an otherwise physically unchanged object in the magnet is accompanied by relative morphological changes in image space. To get insight into the magnitude of morphological changes that are purely caused by minor errors in repositioning, the outlined evaluation was performed on two datasets, that, again, were acquired with identical protocols directly one after the other with a slight difference in the phantom's position in the scanner (rotation less than 5°, translation less than 2 mm). Figure 3.15 shows that this change produced positional deviations of up to 0.6 mm and relative volumetric changes of up to 1.2 % for the Siemens Magnetom Avanto. The corresponding numbers for the Siemens Magnetom Trio are 0.6 mm and 1.8%. Relative morphological changes are smaller for the Siemens Magnetom Avanto due to the smaller spatial gradients of the absolute distortion fields (see figure 3.6).

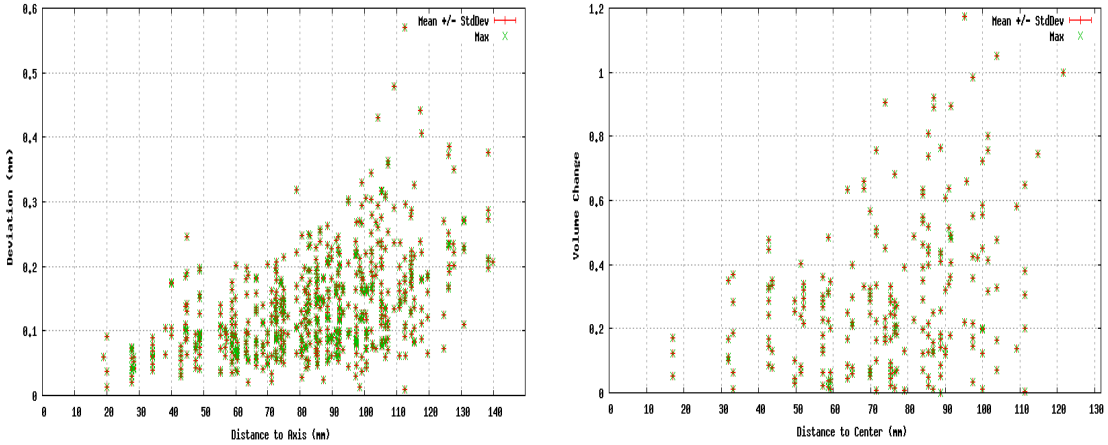


Figure 3.15: *Evaluation results for repositioning experiments: the presented geometrical changes in image space are purely caused by a slight difference in phantom positioning (Siemens Magnetom Avanto, MPRAGE).*

The explanation for the increase of the relative morphological changes with the distance to the isocentre becomes obvious when taking a closer look at figure 3.11: A slightly different position of the phantom results in greater positional deviations for spheres that lie in regions that are far away from the isocentre, as the spatial gradients of the distortions are stronger in these regions. Thus, identical positional differences in real space cause stronger morphological differences in image space for those regions that lie far away from the isocentre.

3.3.7 Head Image Evaluation

As stated above, it is advisable to evaluate the head images with a software package that already is accepted by the neuroimaging community as a valid tool for morphometric analyses, since this guarantees that the study's results are accepted by this community as valid, too. Practically, SPM2 was chosen as it is currently the most commonly used freely available toolbox for morphometrical analyses [Ashburner00].

3.3.7.1 Processing Pipeline

After correcting each dataset for intensity inhomogeneities, a template image is created for each data series (one template for each scanner) by rigidly aligning all datasets to the series' baseline scan. Having rigidly aligned each dataset to its template, GM is automatically segmented in each image with SPM2.

For regional volumetry, three regions of interest (ROIs) are defined in the template images (see Figure 3.16), and the GM volume is determined as the sum of GM density within these ROIs. Large box-shaped ROIs that stretch across the complete slice stack are chosen in order to minimize effects resulting from remaining imperfections of the rigid registration.

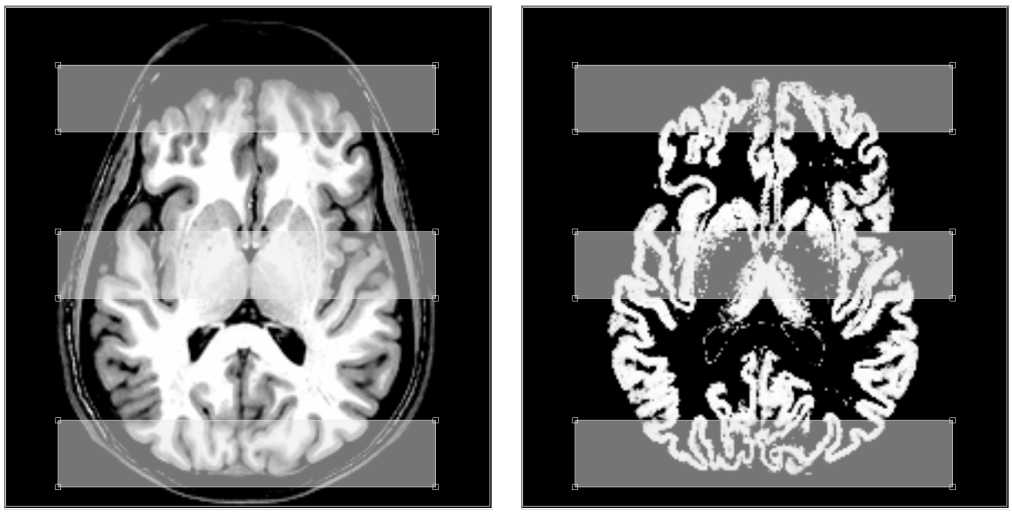


Figure 3.16: *ROI locations (anterior, medial, posterior) in the Avanto template image (left), and in a GM segmentation image (right).*

3.3.7.2 Reproducibility

The reproducibility of the outlined evaluation method is determined in the same way as in the phantom experiments. The volunteer is imaged twice with identical protocols on the same scanner without repositioning and without any time gap between the scans. The only remaining differences, thus, are slight changes in the scanner due to heating, noise in the acquired datasets, and differences in the volunteer's involuntary motion during scan.

A quantitative analysis of these datasets in the previously described way resulted in the following volumetric changes for the Siemens Magnetom Avanto: anterior ROI 0.50%, medial ROI 0.16%, and posterior ROI 0.35%. The results for the Siemens Magnetom Trio are: anterior ROI 0.25%, medial ROI 0.39%, and posterior ROI 0.12%. Note that the test-retest datasets, too, were aligned to the template datasets prior to segmentation.

3.3.7.3 Quantification of Repositioning Influences

Again, the magnitude of morphological variability introduced by imperfections in patient repositioning was estimated experimentally by acquiring two datasets directly one after the other with a slight difference in the volunteer's position in the scanner (rotation less than 5°, translation less than 2 mm, new selection of field of view, new shim).

These changes resulted the following volumetric changes for the Siemens Magnetom Avanto: anterior ROI 0.78 %, medial ROI 0.57 %, and posterior ROI 1.32 %. The corresponding numbers for the Siemens Magnetom Trio are: anterior ROI 0.95 %, medial ROI 0.43 %, and posterior ROI 0.16 %.

3.3.8 Results

Using the outlined methods, absolute morphological errors as well as acquisition-related morphological variability were investigated quantitatively. As mentioned earlier, this was done by scanning a phantom and a healthy human volunteer (28 years old, male) on a weekly basis for 6 months, alternating between a 1.5T Siemens Magnetom Avanto and a 3T Siemens Magnetom Trio, both scanners being marketing installations at Siemens MR in Erlangen, and by analyzing the resulting data series in the previously described way.

For the phantom data series the following measures were determined:

- absolute positional deviation from the phantom reference model *G0*
- absolute volumetric deviation from the phantom reference model *G0*
- variability of corresponding spheres' centre points (i.e. relative positional differences)
- variability of corresponding cells' volumes (i.e. relative volumetric differences)

Due to a missing ground truth of the volunteer's head's geometry, absolute morphological errors obviously cannot be determined for the head data series. The evaluation of the volunteer data therefore is restricted to the quantification of acquisition-related morphological variability. Moreover, only the MPRAGE data series was analyzed, as an additional analysis of the 3DFLASH data would not yield new knowledge: Differences in acquisition-related morphological variability that are caused by differences in the underlying pulse sequences[17] are best detected in the phantom data series, as these are not confounded by differences in patient movements during scan.

3.3.8.1 Acquisition-Related Absolute Morphological Errors (Avanto)

Figures 3.17 and 3.19 show the results obtained for the data series acquired with the 3DFLASH and MPRAGE protocols on the Siemens Magnetom Avanto. Within a volume relevant for brain morphometry (spherical subvolume around the isocentre with a radius of 100mm), the maximal absolute positional errors from the reference geometry are 2.0mm for both protocols. The maximum volumetric errors are 4.0% relative to a standard cell in *G0* of 7600mm^3.

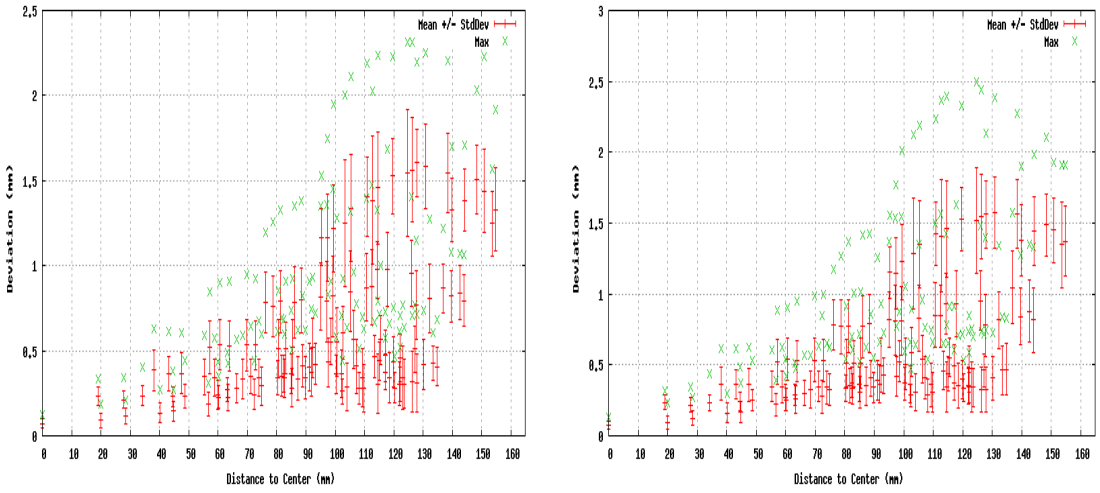


Figure 3.17: *Distribution of positional differences of corresponding spheres compared to the reference phantom model G0 in dependence of their distance to the isocentre: Siemens Magnetom Avanto, 3DFLASH (left), MPRAGE (right)*

Generally spoken, the absolute distortions are depending on the objects' distances to the magnet's isocentre whereby the difference in z-direction from the isocentre plays are more dominant role than the other directions (see figure 3.11). This pattern of the geometrical distortions also explains why volume changes of cells that have a similar

[17] Such differences could occur even when using identical hardware, e. g. if one sequence produces significantly stronger eddy currents, and the eddy current induced field distortions increase over time for both sequences due to scanner hardware imperfections and mechanical stress.

distance to the isocentre can differ dramatically (even in their signs): Figure 3.18 shows that cells lying close to the phantom's central axis are compressed, whereas the outmost cells are expanded.

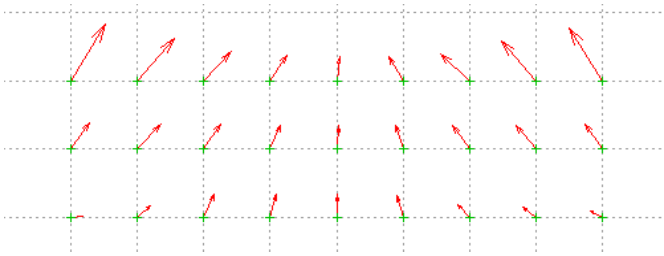


Figure 3.18: *Zoom into figure 3.11: coronal vectorplot of geometrical distortions (Siemens Magnetom Avanto, MPRAGE)*

As the magnitude of volumetric change is directly connected to the positional deviations in the subvolume under investigation, the change in volume generally is growing with the cell's distance to the isocentre, as shown in figure 3.19.

A plausible explanation for the increase of the distributions' standard deviations with the distance to the isocentre are slight differences in phantom positioning.

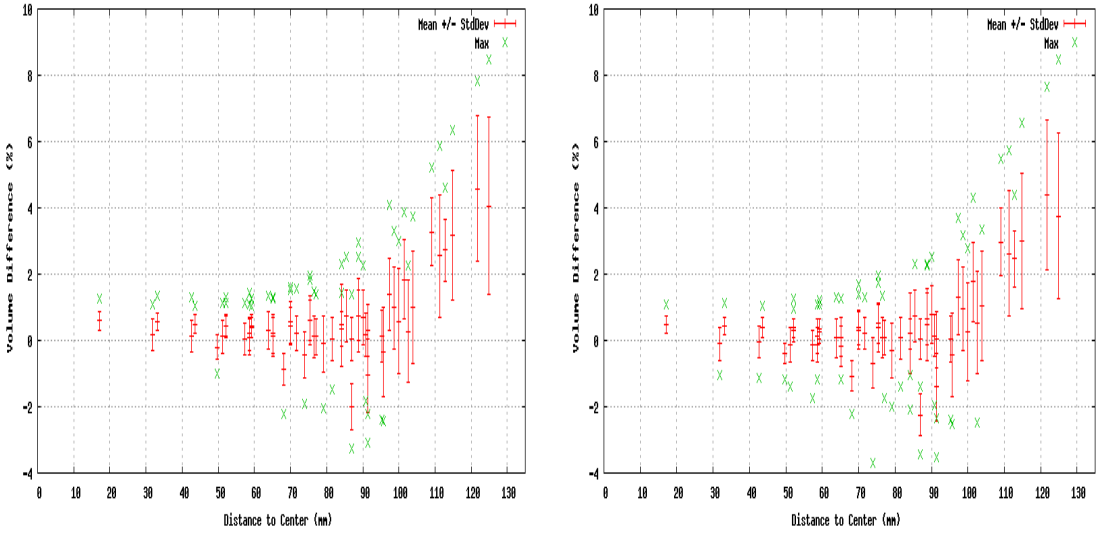


Figure 3.19: *Distribution of relative volume changes of corresponding cells compared to the reference phantom model G0 in dependence of their distances to the isocentre: Siemens Magnetom Avanto, 3DFLASH (left), MPRAGE (right)*

3.3.8.2 Acquisition-Related Absolute Morphological Errors (Trio)

The results for the Siemens Magnetom Trio are, in principle, very similar to the Avanto findings, only that the geometrical distortions are significantly larger in images acquired with the Trio, like it already was shown in figures 3.6 and 3.8. The maximal absolute positional errors from the reference geometry (within a volume relevant for brain morphometry) are 4.8mm, the maximal volumetric errors -16.0% relative to a standard cell in *G0*. Apart from the difference in magnitude, all major findings obtained in the analyses of the Avanto data also hold true for the Trio. Further details are given in figures 3.20 and 3.21.

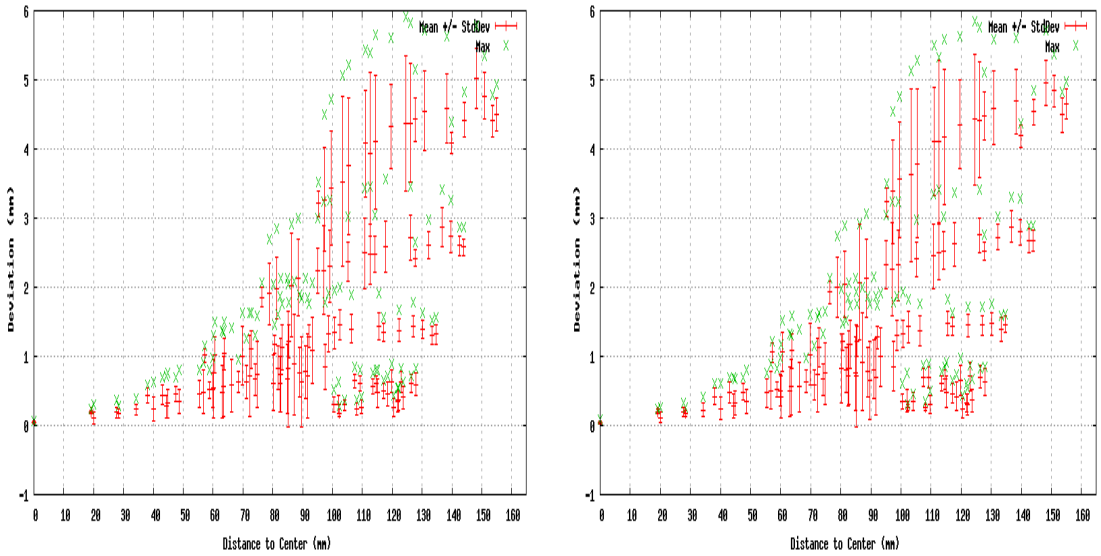


Figure 3.20: *Distribution of positional differences of corresponding spheres compared to the reference phantom model G0 in dependence of their distance to the isocentre: Siemens Magnetom Trio, 3DFLASH (left), MPRAGE (right)*

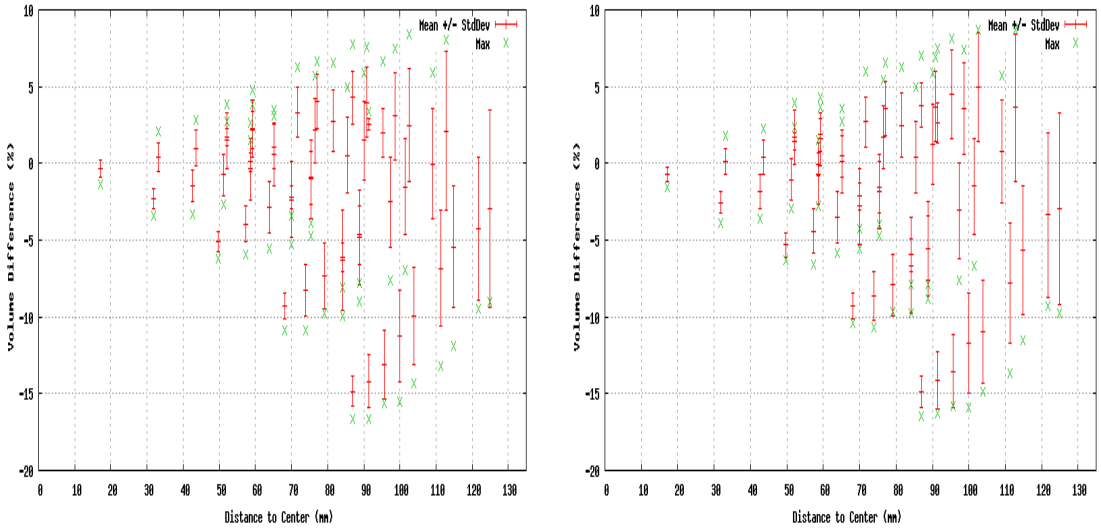


Figure 3.21: *Distribution of relative volume changes of corresponding cells compared to the reference phantom model G0 in dependence of their distances to the isocentre: Siemens Magnetom Trio, 3DFLASH (left), MPRAGE (right)*

3.3.8.3 Acquisition-Related Morphological Variability (Phantom, Avanto)

Figures 3.22 and 3.23 show the results obtained for the phantom data acquired with the 3DFLASH and MPRAGE protocols specified in section 3.3.4 on the Siemens Magnetom Avanto.

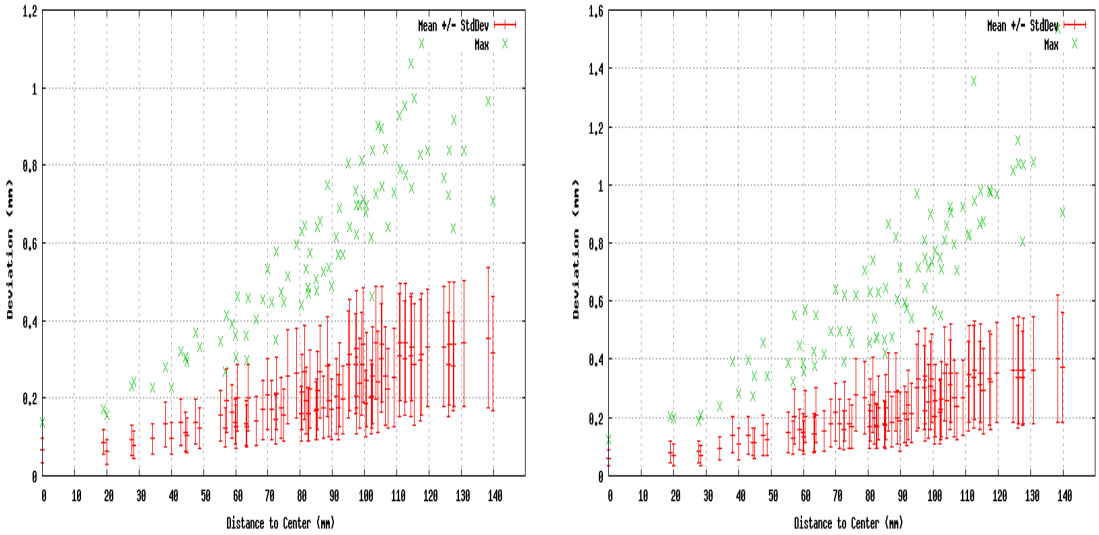


Figure 3.22: *Distribution of relative positional differences of corresponding spheres in dependence of their distance to the isocentre: Siemens Magnetom Avanto, 3DFLASH (left), MPRAGE (right)*

As expected, the magnitude of acquisition-related morphological variability is smaller than the absolute errors shown in the preceding sections. Moreover, the diagrams reveal that acquisition-related morphological variability is very similar in magnitude for the protocols under investigation. Within a volume relevant for brain morphometry, the maximum relative morphological differences are 1.0mm in position and 2.0% in volume (relative to a standard cell in *G0* of 7600 mm^3).

The reason for the increase of acquisition-related morphological variability with the distance to the isocentre, again, are slight differences in phantom positioning. As the underlying nonlinear distortion fields' spatial gradients are stronger in regions that are further away from the isocentre, the same amount of repositioning error manifests itself as a larger morphological change for those parts of the object being further away from the isocentre.

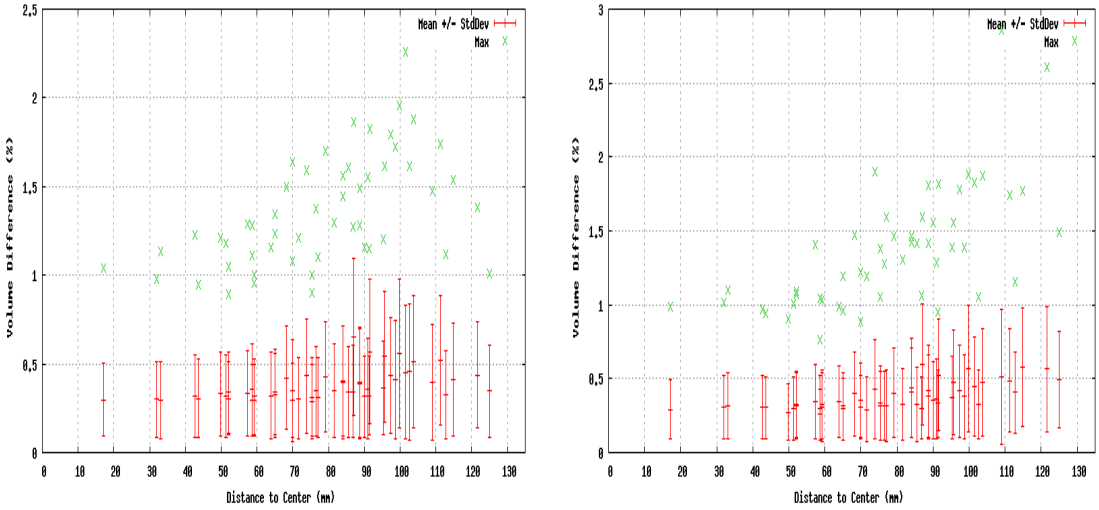


Figure 3.23: *Distribution of relative volumetric differences of corresponding cells in dependence of their distances to the isocentre: Siemens Magnetom Avanto, 3DFLASH (left), MPRAGE (right)*

3.3.8.4 Acquisition-Related Morphological Variability (Phantom, Trio)

As stated above, the same evaluations were made for the images acquired with the Siemens Magnetom Trio, and again, the observations made based on the Avanto

evaluation can be directly transferred to the Trio data. In summary, acquisition-related morphological variability is even slightly smaller in magnitude than in the Avanto data series, despite its larger absolute geometrical distortions. Details are given in figures 3.24 and 3.25.

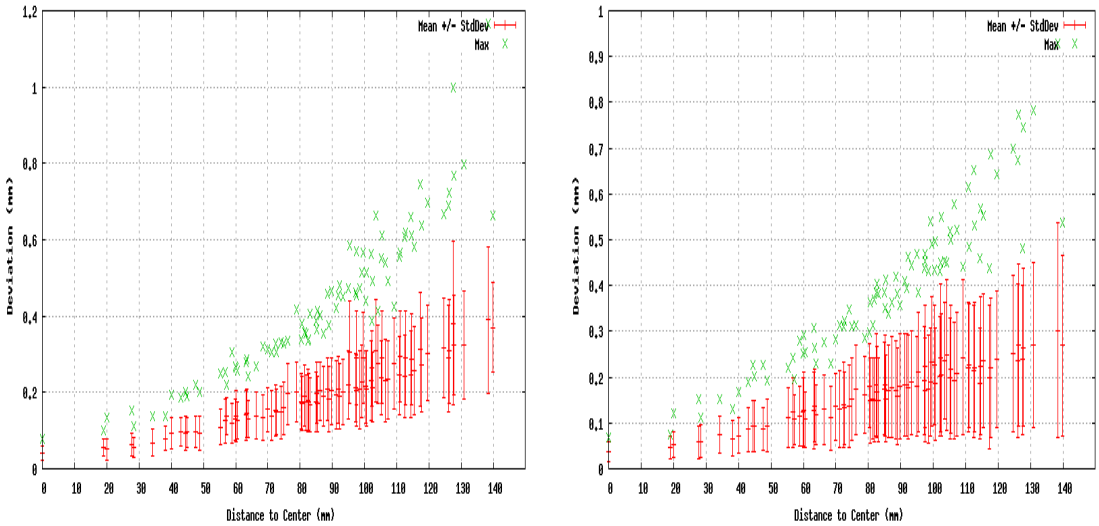


Figure 3.24: *Distribution of relative positional differences of corresponding spheres in dependence of their distance to the isocentre: Siemens Magnetom Trio, 3DFLASH (left), MPRAGE (right)*

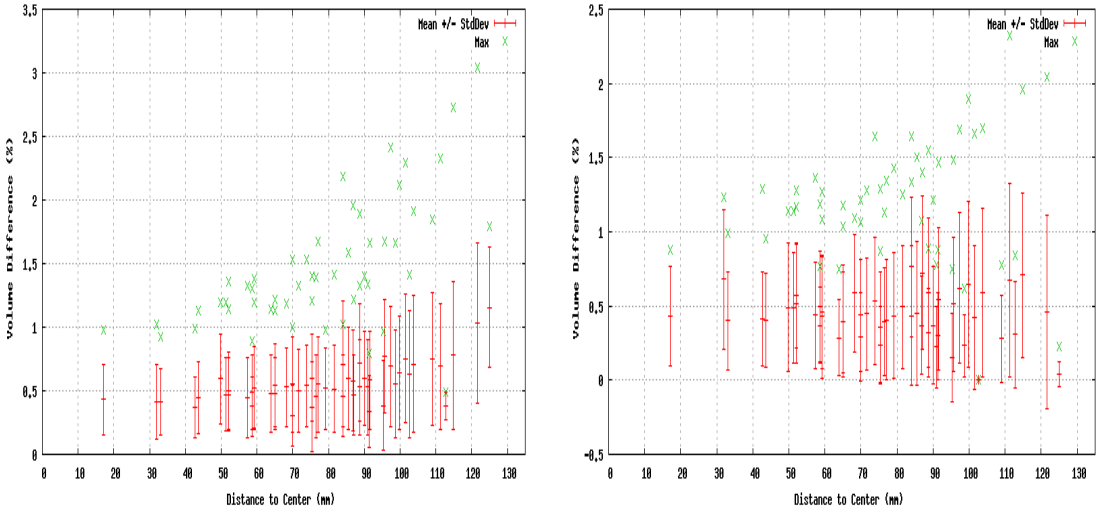


Figure 3.25: *Distribution of relative volume changes of corresponding cells compared with each other in dependence of their distances to the isocentre: Siemens Magnetom Trio, 3DFLASH (left), MPRAGE (right)*

3.3.8.5 Acquisition-Related Morphological Variability (Volunteer, Avanto, Trio)

The results of the phantom experiments are well reflected in the evaluations of the head images (see table 3.3). The overall morphological variability as determined with the outlined evaluation procedure is up to 5%, including differences in the volunteer's unintended movement during scan (see section 3.2.2.2). As the anterior and posterior ROI are lying further away from the isocentre, the acquisition-related variability in local GM volumetry is higher for them compared to the medial ROI.

	Avanto	Trio
anterior ROI	2.88 % (of 40773.8 mm^3)	4.74 % (of 29693.4 mm^3)
medial ROI	2.69 % (of 187359.2 mm^3)	2.60 % (of 184959.6 mm^3)
posterior ROI	5.24 % (of 68907.8 mm^3)	4.78 % (of 86514.3 mm^3)

Table 3.3: *Results of head image evaluation: acquisition-related variability in local GM volumetry estimated by range divided by mean (given in brackets)*

3.3.9 Discussion

In this chapter, it was experimentally quantified to what extent acquisition-related morphological variability affects the detection limit of image-based morphometric applications aiming at the detection of subtle changes in intra-subject longitudinal MRI data.

In contrast to prior work, not only the confounding effect of gradient coil nonlinearity was analyzed, but rather the impact of all sources of acquisition-determined geometrical distortions. Moreover, these were investigated for the first time in a long-term study in order to also account for imperfect hardware stability and for maintenance activities like scanner calibration or software updates. Finally, the presented study extends prior research in terms of the way acquisition-related morphological variability was evaluated, namely locally, in the complete inner volume of the head coil, and with respect to both positional as well as volumetric variability. The study's findings therefore enable medical professionals to directly infer to what degree the morphometric measures presented in chapter 2.2 typically are confounded by acquisition-related morphological variability, no matter whether deep brain structures or cortex regions are morphometrically investigated.

Regarding the presented method for analyzing the phantom data, it can be stated that it allows for a reliable quantification of acquisition-related morphological variability. Its reproducibility as determined in test-retest experiments is higher than effects expected from patient movement during scanning. In detail, the maximal relative positional difference was 0.2mm, and the maximal change in volume 0.35% (including effects caused by the scanners' resonance frequency drift). Analogous test-retest experiments for validation of the head image evaluation method resulted in volumetric changes of up to 0.50%.

As outlined in chapter 3.2, an important influencing factor on the magnitude of acquisition-related morphological variability in longitudinal MRI are imperfections in patient repositioning. Even though it is difficult to generally quantify this magnitude due to its strong dependency on the scanner, the imaging sequence, and on the difference in position itself, the results of the presented repositioning experiments can be regarded as a useful estimation. Phantom experiments yielded up to 0.6mm relative positional deviation and up to 1.2% relative volumetric change; analogous human volunteer experiments resulted in variability in local GM volume of up to 1.3%.

For detection of neurodegenerative diseases, the brain is obviously the only organ of interest, which is why the empirically determined distributions of acquisition-related morphological variability for regions the distance of which to the isocentre exceeds 100 mm can be neglected. In detail, the maximal relative morphological differences were 1.0mm positional and 2.0% volumetric change on a 7600 mm^3 cuboid.

The absolute morphological deviations from the reference phantom model *G0* are by far greater in datasets acquired with the Siemens Magnetom Trio than in the Avanto

data. The maximal absolute deviations measured were 2.0mm positional distortion and 4.0% volume mismatch for the Avanto, and 5.0mm respectively -16% for the Trio. These uncertainties have to be taken into account when clinically interpreting absolute volumetric measures determined from MRI images. The reason for this enormous difference between the Avanto and the Trio data are the different gradient coils integrated in these scanners; the Avanto gradient coils are more linear in a region relevant for brain morphometry.

The latter is also the reason why application of the 3D gradient nonlinearity correction software offered by Siemens does not reduce acquisition-related morphological variability in the Avanto data series as much as in the Trio data: Whereas the volumetric variability could be reduced from 2.0 % to 1.4 % in the Trio data, the reduction in the Avanto data series was rather marginal (from 2.0 % to 1.8 %). The superiority of the Trio over the Avanto after gradient nonlinearity correction can have several reasons, among them that its hardware could be more stable over time, e.g. more resistant to mechanical stress, or that the active shimming procedure could be more reproducible in patient-based coordinates. This question cannot be answered based on the data acquired in the course of this study.

When comparing the two sequences (3DFLASH, MPRAGE) against each other, it turns out that they behave very similarly in terms of geometrical distortions. Not only the magnitude of morphological variability, but also the shape of the distributions is very similar. This can be interpreted as high similarity in the underlying 3D distortion fields. Moreover, this similarity in terms of geometrical distortions means that the decision, which of these sequence to choose for brain morphometry, can be made solely based upon signal-to-noise (SNR) and contrast-to-noise (CNR) measurements.

The results of the phantom experiments are well reflected in the evaluations of the head images: Especially cortical areas that are relatively far away from the magnet's isocenter are affected by geometrical distortions. As the volunteer was positioned with great care, and since he was very cooperative, the numbers given in table 3.3 (up to 5% variability in local GM volume determination) have to be considered a realistic estimation of reproducibility in state-of-the-art MRI based morphometry of cortex regions in terms of a lower limit.

Summarizing, it can be stated that that acquisition-related morphological variability affects the detection limit achievable with any post processing application: Diseases that are accompanied by gradual morphological changes cannot be reliably detected with imaged-based techniques until the actual anatomical change exceeds the previously stated acquisition-related morphological variability in image space. Transferred to Alzheimer's diseases this means that a portion of about 2% of neurons in functionally highly important parts of the medial temporal lobe are irreversibly destroyed until therapeutic treatment can be started with.

Recapitulating chapter 2.1, it can be concluded that these 2% are too much to be considered irrelevant for MRI-based early AD detection. As mentioned earlier, Thompson et al report a maximum annual loss of local GM of 5% [Thompson03]. Simplifying (in assuming linear AD progression), this means that acquisition-related morphological variability would delay clinically valid image-based AD detection by 4.8 months.

In fact, state-of-the-art methods for segmentation of deep brain structures report a maximal reproducibility in absolute volumetry of about 3-4 % in healthy volunteers despite application of 3D gradient coil nonlinearity correction [Quinn05]. Due to an increase in patient movement as a result of dementia patients' reduced ability to cooperate and because of a decreased tissue contrast in dementia patients' image data, these values very likely will be even larger in a clinical setting. As a substantial portion of the reported 3-4% can be assumed to be a result of acquisition-related morphological variability, it is necessary to think about ways of further reducing morphological variability in image space that does not correspond to an actual change in a patient's anatomy, but rather is a direct consequence of imperfections in MRI data acquisition.

3.4 Managing Morphological Variability

3.4.1 Introduction

Recapitulating the investigations presented in this chapter so far, it can be stated that longitudinal data series contain morphological variability in image space for a physically unchanged object even if it is scanned with identical hardware using always the same imaging protocol. Moreover, it is evident that the outlined sources of acquisition-related morphological variability will not be eliminated by scanner manufacturers for several years to come. And since MRI scanners – as a result of their enormous purchase costs – often are used for some years even after they have become outdated, it can be concluded that the image processing community as well as medical professionals have to deal with acquisition-related morphological variability for quite a long time.

It was shown in section 2.2 that most image processing and –analysis groups working in the field of high-accuracy morphometry so far seem to neglect the limitations imposed on their application-specific detection limit by acquisition-related morphological variability, and that they rather concentrate on analyzing each dataset individually as accurately as possible. For cross sectional studies with sufficiently large populations, this strategy will still yield statistically valid results as both groups can be assumed to be similarly affected by acquisition-related morphological variability. In an intra-subject scenario, however, it has to be accepted that following this approach, local morphological changes smaller than the previously given detection limits simply cannot be considered significant. In fact, the experimental study on quantifying acquisition-related morphological variability revealed that this phenomenon would lead to a time delay in image-based AD detection of about 5 months.

Currently, state-of-the-art methods for segmentation of deep brain structures report a maximal reproducibility in absolute volumetry of about 3-4 % in healthy volunteers despite application of 3D gradient coil nonlinearity correction [Quinn05]. Due to an increase in patient movement as a result of dementia patients' reduced ability to cooperate and because of a decreased tissue contrast in dementia patients' image data, these values very likely will be even larger in a clinical setting. Taking into account the results of section 3.3, these considerations are strong arguments for the necessity to develop methods for overcoming the limitations imposed on MRI-based morphometry by imperfections in the data acquisition process.

Therefore, in this subchapter possible strategies for correcting MRI data for acquisition-related morphological variability that go beyond 3D gradient coil nonlinearity correction are investigated, and they are evaluated with respect to their suitability for spatial normalization of intra-subject MRI data series in a clinical setting.

3.4.2 Spatial Normalization Using Extrinsic Markers

3.4.2.1 Principal Idea

The only technique known from literature that is explicitly dedicated to the elimination of acquisition-determined geometrical distortions is to determine the overall 3D distortion field by first acquiring phantom data and subsequently comparing the phantom's geometry in image space to its geometrical specification [Breeuwer00][Mallozzi04][Wang04b]. As illustrated in figure 3.26, the resulting nonlinear distortion field can then be used to correct subsequently acquired patient data for geometrical distortions.

The variants of this technique differ mostly in the design of the phantom used[18]. However, it is refrained here from explicitly describing their technical details, since these are not the determining factors for an evaluation of the potential and limitations of phantom-based spatial normalization of MRI data. For the latter, it is sufficient to look at these procedures from a more general point of view.

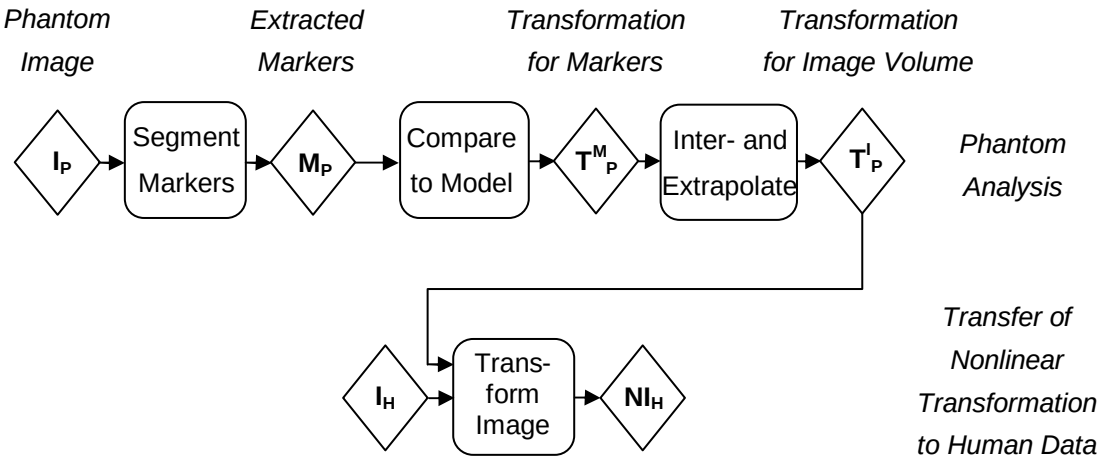


Figure 3.26: *Schematic illustration of phantom-based spatial normalization.*

3.4.2.2 Limitations

With regard to phantom-based approaches, it is important to take into consideration how reproducible the active shimming procedure works for a specific phantom (see section 3.3.4). Typically, 3D distortion correction phantoms have a relatively complex internal structure, and some of them do not allow for a highly reproducible active shimming. To avoid the transfer of this variability to patient data, phantom-based approaches ideally would include the correction of the determined distortion fields for susceptibility artefacts specific to the phantom used to determine them. One possible approach to do this is the inversion of the pulse sequence's readout polarity. As illustrated in figure 3.27, susceptibility induced geometrical distortions are reversed in

[18] In fact, the method for analyzing phantom image data presented in section 3.2.6 can be regarded as one representative of this class of techniques.

image space for an inversion of the readout polarity[19]. So far, however, none of the phantom-based approaches for spatial normalization explicitly addresses this problem. In fact, all these methods neglect phantom-specific geometrical distortions, which implies that these are directly transferred to subsequently acquired patient data.

Moreover, normalization strategies that are built upon phantom data are unable to correct data for (a) differences in the patient's movements during scan and (b) for morphological changes resulting from differences in susceptibility induced geometrical distortions (see section 3.2.2).

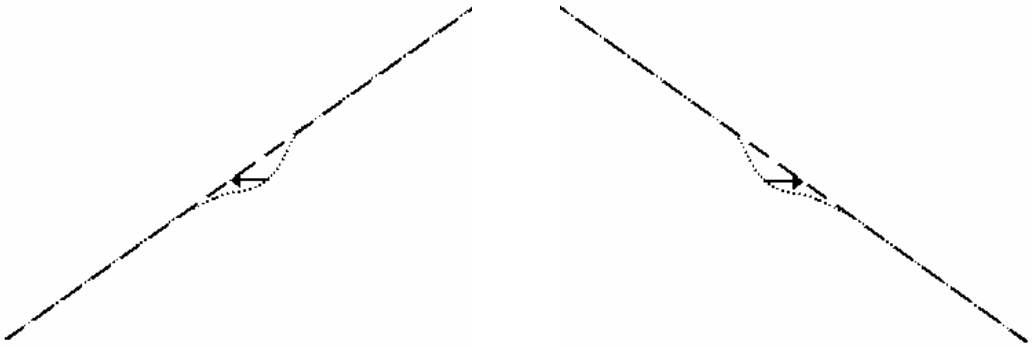


Figure 3.27: *Geometric distortions due to differences in magnetic susceptibility in dependence on the readout gradient; the readout gradient is represented by the dashed line, the effective gradient (readout gradient plus local susceptibility influences) is visualized as the dotted line; the arrow indicates a shift in image space (in frequency encoding direction) due to local differences in magnetic susceptibility*

Probably even more important is the workflow-related drawback that phantom-based approaches involve. As a result of imperfect hardware stability due to mechanical stress, the scanners' distortion fields would have to be determined regularly, for example in a quality assurance procedure every morning. Importantly, these fields ideally would have to be determined separately for every imaging protocol potentially used for morphometrical examinations on that day. In short, phantom-based spatial normalization implies a lot of additional work (consisting of data acquisition, analysis, storage and application) that needs to be done very carefully in order to be effective in a clinical setting.

3.4.3 Spatial Normalization Using Intrinsic Marker

3.4.3.1 Principal Idea

From a general point of view, all the outlined problems with phantom-based strategies are direct concomitants of the use of extrinsic markers for spatial normalization. They would, thus, vanish if it was possible to utilize structures intrinsically contained in the patients' image data.

Practically, this approach requires an accurate and reproducible segmentation of suitable markers, followed by a nonlinear alignment of theirs. As illustrated in figure 3.28, it is furthermore necessary to appropriately interpolate between the resulting locally defined transformations to generate a nonlinear deformation field that is defined over the complete image volume in order to be able to anatomically align the image data. If implemented correctly, this processing pipeline would nonlinearly normalize longitudinal datasets such that acquisition-related morphological

[19] It should be mentioned that eddy current induced distortions are not affected by an inversion of the readout polarity as eddy currents are reversed as well, and such are the dynamic field distortions they cause.

differences are eliminated while differences corresponding to true anatomical changes are preserved.

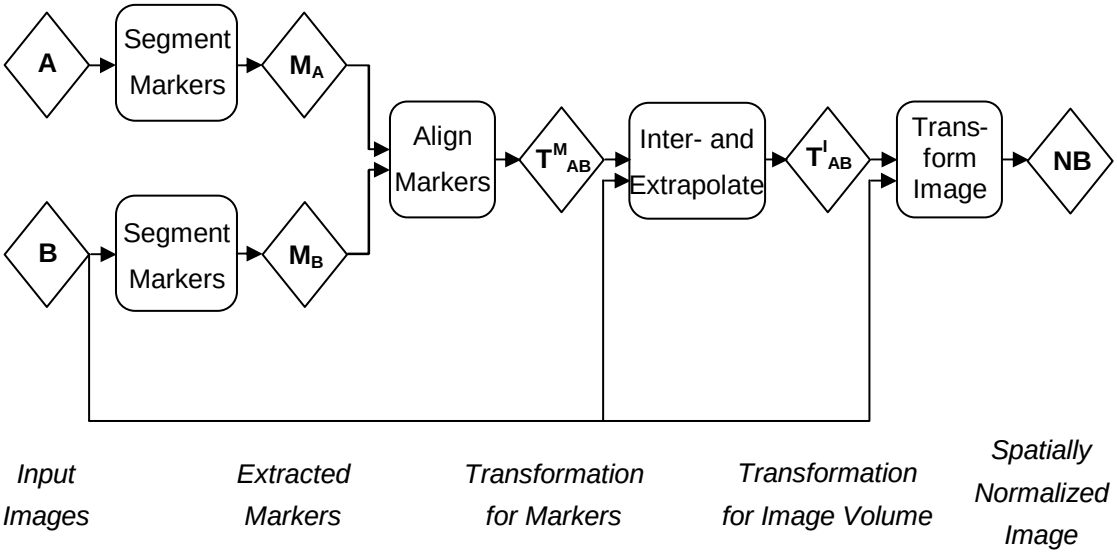


Figure 3.28: *Schematic illustration of spatial normalization using intrinsic markers.*

3.4.3.2 Identification of Suitable Markers

Abstracting from the indicated technical challenges, the crucial factor of this approach is the identification of suitable markers, as these have to satisfy strong requirements:

- They need to be unaffected by the pathology under investigation, but affected by the outlined sources of acquisition-related morphological variability in image space.
- They need to be densely distributed within the imaged volume.
- They need to be located closely to the anatomical structures of interest.

At first glance, it seems plausible in longitudinal neuroimaging to use the skull as an intrinsic means for spatial normalization, as it can be considered an almost perfectly stable structure over time in most cases. In T1-weighted structural MRI data, however, the skull is not really visible. Practically, the skull is therefore identified in these images to a large extent by delineation of subcutaneous fatty tissue, the position of which in image space is noticeably affected by the chemical shift artefact (see section 3.2.2). Consequently, some variability in the segmentation of the skull has to be expected, especially in the presence of imperfections in patient repositioning (see section 3.2.3), and spatial normalization of longitudinal brain images using the patient's skull would be confounded by these errors. Another important factor that argues against the skull as intrinsic marker is its comparatively large distance to deep brain structures of interest to morphometry.

When recapitulating section 2.1, however, it turns out that WM is a superior candidate marker for spatial normalization of structural MRI head images of dementia patients, as long as these are in an early stage of the disease: As the volume of a neuron's cell body exceeds its axon's volume by far, atrophy is becoming noticeable in MRI images in GM a lot earlier than in WM. In other words, if atrophy has become large enough to be just noticeable in GM, it is definitely not noticeable in WM, yet. For image-based *early* detection of neurodegenerative diseases it is therefore a valid assumption to consider WM as morphologically unaffected by the disease. Moreover,

WM naturally is always anatomically neighbouring GM throughout the complete brain which assures that it is equally affected by acquisition-related morphological variability. In summary, all the previously stated requirements are met.

3.4.3.3 Representatives

So far no post-processing methods have been published that explicitly aim at eliminating acquisition-related morphological variability from MRI brain images based upon the spatial alignment of well-suited markers intrinsically contained in the original datasets.

However, the tools developed for cross-sectional analyses of cortical thickness (see section 2.2.4) reduce acquisition-related morphological variability - even though they do not seem to have been developed with the intention to do so. This characteristic is a by-product of these methods' necessity to establish a uniform surface-based coordinate system in order to make possible subsequent statistical comparisons across different datasets.

Practically, all variants of this approach make use of a spherical coordinate system: Given two datasets *A* and *B*, and the corresponding surface models M_A and M_B representing the border between GM and WM in these datasets, M_A and M_B are transformed (by a nonlinear transformation *S*) into a common spherical coordinate system. Since a comparison of cortical thickness requires to align the surface models to each other, $S(M_B)$ is reparameterized such that its similarity to $S(M_A)$ is maximized - which implicitly corresponds to a deformation of M_B into the shape of M_A. There are two different approaches for determining the nonlinear transformation *S*:

- The first method iteratively inflates each vertex of the input surface driven by local convexity or concavity until a sphere emerges, whereby the integral of the normal movement of a vertex during inflation is used to characterize the folding pattern of the surface. This folding pattern is subsequently also used as input data for the maximization of surface similarity [Fischl99b].
- In the second technique, *S* is intrinsically determined during segmentation, as M_A and M_B are determined by iteratively deforming a spherical mesh into the shape of the border between GM and WM. The correspondence problem is then solved either by restricting the number of vertices during the deformation procedure (and, thus, guaranteeing a 1-1-mapping for each surface vertex) [MacDonald00] or by determining the spatial correspondence of ridges on major gyri and fissures of important sulci that occur consistently in normal subjects [Thompson01a]. Again, the surfaces' folding pattern - this time characterized by local surface curvature - can be utilized to constrain the surface reparameterization driven by manually identified crest lines [Davatzikos97].

Abstracting from their concrete notation, these methods have in common that they nonlinearly align the reconstructed WM surface representations to each other. Thus, ideally all acquisition-related morphological variability contained in WM is eliminated by them.

3.4.3.4 Limitations

For all its general advantages, spatial normalization using intrinsic markers – no matter which markers are used - does not allow for the determination of the overall

3D distortion field due to the missing ground truth of a patient's exact geometry. In other words, intrinsic markers can only be used to eliminate acquisition-related morphological variability, but not to correct image data for geometrical distortions caused by imperfections in data acquisition.

Another drawback of these techniques is their specificity. Depending on the choice of markers, a normalization procedure using intrinsic markers will only produce valid results for a limited number of pathologies. For example, WM-based normalization approaches – despite being suitable for the case of early Alzheimer's disease – apparently are not applicable to WM diseases.

Regarding the outlined techniques for cross-sectional analyses of cortical thickness, it can be stated that they reduce acquisition-related morphological variability, but only in cortex regions. More explicitly, these methods are unable to spatially normalize subcortical midbrain and deep brain structures, as these simply do not fit into the concept of cortical thickness. In fact, these structures are excluded from cortical thickness analyses, as an inclusion of them would completely corrupt the spherical topology of the WM/GM-border that all the presented techniques enforce [Fischl99a]. Moreover, it has to be questioned whether the outlined surface-based nonlinear registration strategy in a common spherical coordinate system is as accurate as a registration technique that does not require a substantial distortion of the inherent geometry of WM. Whereas mapping the original representations of patients' brain surfaces into a sphere is inevitable to compare surface-based morphometric measures in cross-sectional studies, it is not necessarily required for elimination of acquisition-related morphological variability in image space.

Finally, it is obvious that spatial normalization strategies using intrinsic markers depend on their reliable segmentation in the acquired image data. However, this does not necessarily mean that the tissue contrast needs to be identical in the images that are to be brought into alignment. Taking the outlined WM-based approach as an example again, it rather requires WM to show a sufficiently strong contrast to neighbouring GM – no matter whether the underlying image data were acquired with T1- or T2-weighting.

3.4.4 Extrinsic versus Intrinsic Markers

Having outlined the two possible approaches for reducing acquisition-related morphological variability in longitudinal MRI brain data series, they are now evaluated with respect to their suitability for spatial normalization in the context of early detection of neurodegenerative disease in a clinical setting.

Phantom-based approaches at first glance seem to be technically rather straight forward, but when looking at them in detail, it turns out that their correct implementation requires a great deal of work that needs to be smoothly integrated into the clinical workflow:

- First of all, the distortion fields determined by analyzing additional phantom scans ideally need to be corrected for phantom-specific geometrical distortions in order to avoid the transfer of potential morphological variability specific to the phantom scans to human data. This necessity arises especially from imperfect reproducibility in the active shimming procedure when analyzed in the imaged object's coordinate system (see section 3.2.3).

- Secondly, the smooth integration of phantom-based normalization strategies into the clinical workflow is a rather complex task: Due to imperfect hardware stability, the scanners' distortion fields would have to be measured regularly, and they ideally would have to be measured separately for every imaging protocol potentially used for morphometrical examinations (because of differences in the underlying 3D deformation fields resulting from different eddy current induced magnetic field distortions or from different readout bandwidths).
- Finally, this strategy is generally unable to correct neither for differences in the patient's movements during scanning nor for morphological changes resulting from differences in susceptibility induced geometrical distortions in the original patient data.

As all the outlined problems are direct consequences of the use of extrinsic markers, it is attractive to find ways of utilizing structures intrinsically contained in the patients' image data for spatial normalization. Based upon the analysis of the manifestation of early Alzheimer's disease in image space given in section 2.1, the only structure meeting all requirements that these intrinsic markers need to fulfil has been shown to be WM. Despite all its general advantages, however, WM-based spatial normalization also has got some drawbacks:

- It can only be used for elimination of acquisition-related morphological variability, but not for correcting image data for geometrical distortions caused by imperfections in data acquisition due to a missing ground truth.
- Its application is specific to pathologies that do not affect WM. Regarding Alzheimer's disease this means that WM-based normalization may only be used for images of patients that are at a very early stage of the disease.
- Finally, WM obviously needs to show a sufficient signal intensity contrast to GM in the images to be normalized.

Concluding, it can be stated that the two approaches have their strengths in different scenarios: When aiming at correcting image data for geometrical distortions, e.g. in data pooling in large cross-sectional multi-centre multi-vendor studies, phantom-based approaches clearly are to be preferred. In longitudinal intra-subject scenarios like early detection of Alzheimer's disease, however, spatial normalization using WM as an intrinsic marker theoretically is superior even though it is technically more demanding. In short, this is a direct consequence of being - in principle - capable of fully eliminating acquisition-related morphological variability in image space while preserving differences corresponding to real anatomical changes – without any additional scans or clinical workflow changes.

3.4.5 Discussion

The previous chapters have shown that acquisition-related morphological variability affects the detection limit achievable with any post processing application. For early detection of neurodegenerative diseases, it is therefore necessary to develop methods capable of further reducing this variability in order to fully exploit the potential of image-based morphometry.

From a technical point of view, this reduction of acquisition-related morphological variability is a question of spatial normalization. In practice, it can be tackled by

utilizing either extrinsic or intrinsic markers, whereby the intrinsic structure most suitable for spatial normalization of longitudinal MRI data series of early Alzheimer's patients has been shown to be WM.

An evaluation of these two general approaches' suitability for spatial normalization of longitudinal intra-subject data series in a clinical setting revealed that WM-based spatial normalization theoretically is superior. As mentioned above, this is a direct consequence of it being - in principle - capable of fully eliminating acquisition-related morphological variability while preserving differences corresponding to real anatomical changes – without any additional scans or clinical workflow changes.

The only post-processing methods known from literature that utilize WM as a normalizing structure are algorithms developed for cross-sectional analyses of cortical thickness. As described in section 3.4.3.3, these methods nonlinearly align the reconstructed WM surface representations to each other in a common spherical coordinate system to allow for statistical comparisons of cortical thickness across different datasets. However, these methods are unable to spatially normalize subcortical midbrain and deep brain structures, as these simply do not fit into the concept of cortical thickness. Moreover, it is questionable whether the outlined surface-based nonlinear registration strategy in a common spherical coordinate system is as accurate as registration techniques that do not require a substantial distortion of the inherent geometry of WM.

3.5 Conclusion

Even though a lot of effort has been spent over the past decades to develop innovative image processing and –analysis algorithms and to improve existing methods in terms of accuracy, reproducibility and computational efficiency, relatively few research was undertaken to find out to what extent the validity of results obtained with these methods is limited by inherent imperfections of the input images. In fact, section 2.2 revealed that all these methods implicitly and tacitly assume that the data acquisition process does not introduce any morphological variability into image space.

In this chapter, it was shown that this assumption is not correct for MRI-based morphometry. Imperfections in the MRI data acquisition process, along with practical limitations in the reproducibility of patient positioning, and imperfect hardware stability were shown to introduce morphological variability into an object's representation in image space even for a physically unchanged object. Regarding the complexity of this variability, it can be stated that it corresponds to a nonlinear spatial transformation of the object's morphology in image space, whereby the underlying deformation field is dependent on the scanner, the imaging sequence, its parameterization and the imaged object itself.

In order to answer the - from a clinical point of view - crucial question to what extent the MRI data acquisition process quantitatively affects the detection limit of existing morphometric applications, an experimental study was designed and carried out that extends prior research in several aspects:

- Not only was the confounding effect of gradient coil nonlinearity analyzed, but rather the impact of all sources of acquisition-determined geometrical distortions.

- These were investigated for the first time in a long-term study in order to also account for imperfect hardware stability and for maintenance activities like scanner calibration or software updates.
- Acquisition-related morphological variability was evaluated locally, in the complete inner volume of the head coil, and with respect to both positional as well as volumetric variability.

The study's findings therefore for the first time enable medical professionals to directly infer to what degree the morphometric measures presented in section 2.2 typically are confounded by acquisition-related morphological variability, no matter whether deep brain structures or cortex regions are morphometrically investigated.

In practice, the study revealed that diseases that are accompanied by gradual morphological changes cannot be reliably detected with imaged-based techniques until the actual anatomical change exceeds a volume change of at least 2%.[20] Recapitulating section 2.1, it can be concluded that this is too much to be considered irrelevant for MRI-based early AD detection. As mentioned earlier, Thompson et al report a maximum annual loss of local GM of 5% [Thompson03]. Simplifying (in assuming linear AD progression), this means that acquisition-related morphological variability delays clinically valid image-based AD detection by nearly 5 months.

In order to fully exploit the potential of MRI-based morphometry, it is therefore required to develop methods capable of further reducing acquisition-related morphological variability. From a technical point of view, this reduction is a question of spatial normalization that can be tackled by utilizing either extrinsic or intrinsic markers. An evaluation of these two general approaches' suitability for spatial normalization of longitudinal intra-subject data series in a clinical setting revealed that spatial normalization using intrinsic markers generally is superior a long as an appropriate intrinsic marker can be identified. For early detection of AD, this was shown to be WM.

The only post-processing methods known from literature that utilize WM as a normalizing structure are algorithms developed for cross-sectional analyses of cortical thickness[21]. However, these methods are unable to spatially normalize subcortical midbrain and deep brain structures, as these simply do not fit into the concept of cortical thickness. Moreover, it is questionable whether the outlined surface-based nonlinear registration strategy in a common spherical coordinate system is as accurate as registration techniques that do not require a substantial distortion of the inherent geometry of WM.

The following chapter therefore further explores the idea of WM-based spatial normalization of intra-subject MRI brain data series, whereby it is put special attention to the identification of ways to overcome the outlined limitations of existing techniques.

[20] For further details regarding the magnitude of acquisition-related morphological variability, please refer to section 3.3.8.

[21] However, these algorithms do not seem to have been developed with the intention to reduce acquisition-related morphological variability. This characteristic merely is a by-product of these methods' necessity to establish a uniform surface-based coordinate system in order to make possible subsequent statistical comparisons across different datasets.

4 Development of a Software Prototype for Elimination of Acquisition-Related Morphological Variability by Means of WM-Based Spatial Normalization

4.1 Introduction

The investigations presented in the previous chapters revealed that MRI-based high-accuracy morphometry needs to take into consideration the data acquisition process, since the latter introduces variability into an imaged object's morphological representation in image space in longitudinal data series even for a physically unchanged object. Recapitulating, this is a consequence of MRI-specific acquisition-determined geometrical distortions being not perfectly reproducible in patient-based coordinates due to practically unavoidable imperfections in patient repositioning, patient movements during scanning and imperfect hardware stability.

The clinical relevance of this acquisition-related morphological variability for MRI-based high-accuracy morphometry was shown by means of an experimental long-term study designed to quantify the impact of the MRI data acquisition process on morphometric analyses. Essentially, these experiments revealed that changes in a patient's morphology cannot be reliably detected with MRI until the actual change of an anatomical entity's volume exceeds at least 2%. Taken the findings of Thompson et al as a basis for the progress of AD [Thompson03], this value corresponds to a delay in MRI-based early detection of AD of nearly 5 months. From a clinical point of view, this has to be even considered a conservative estimate of the detection limit achievable with state-of-the-art morphometric applications, as the 2% do not include morphological variability caused by differences in patient movement during scanning. In fact, state-of-the-art methods for segmentation of deep brain structures report a maximal reproducibility in absolute volumetry of about 3-4% in healthy volunteers despite application of 3D gradient coil nonlinearity correction [Quinn05]. Due to an increase in patient movement as a result of dementia patients' reduced ability to cooperate and because of a decreased tissue contrast in dementia patients' image data, these values very likely will be even larger in a clinical setting.

In order to fully exploit the potential of MRI-based morphometry, it is therefore required to develop methods capable of correcting intra-subject MRI data series for acquisition-related morphological variability. The analysis of possible strategies for this task resulted in phantom-based methods being generally inferior to methods utilizing intrinsic markers. For this thesis's medical context, i.e. MRI-based early detection of AD, WM was shown to be the optimal intrinsic marker. In total, these considerations resulted in the proposal of a novel concept for elimination of acquisition-related morphological variability from intra-subject MRI head data series, which can be characterized as WM-based spatial normalization.

This chapter presents this novel concept in terms of a software prototype implemented in order to be able to experimentally test the correctness of the previously presented system analysis. It is thereby divided into three parts:

- Firstly, it is given a focused review of the state-of-the-art in image registration and WM segmentation in order to identify the most promising image processing

methods to be used for a prototypical implementation of WM-based spatial normalization.

- Subsequently, this general specification is refined by a discussion of the technical details of the prototype implemented in this thesis.
- Finally, experimental results obtained with the proposed method are presented, whereby special emphasis is put on controlled validation experiments that allow for a profound discussion of both the method's general potential, and its limitations in its prototypical implementation. The latter also serves as the foundation for suggestions for future work.

For better readability, each of these subchapters, again, is self-contained insofar as they start with an introduction that briefly outlines their internal structure as well as their position within the chapter's overall context. Moreover, each of them is completed by a summary of its main results and of the resulting conclusions.

4.2 WM-Based Spatial Normalization - Focused Technical Review

4.2.1 Introduction

As outlined in section 3.4.3.3, the only post-processing methods known from literature that utilize WM as a normalizing structure are algorithms developed for cross-sectional analyses of cortical thickness, e.g. those proposed by [Fischl99b] and [Thompson01a]. Abstracting from their concrete mathematical notation, these post-processing techniques have in common that they nonlinearly align the reconstructed WM surface representations to each other in a common spherical coordinate system, such that a large portion of acquisition-related morphological variability contained in WM is eliminated by them. However, these methods have got one fundamental limitation, namely their disability to spatially normalize subcortical midbrain and deep brain structures, as these do not fit into the concept of cortical thickness. Furthermore, it is questionable whether surface-based nonlinear registration in a common spherical coordinate system is as accurate as registration techniques that do not require a substantial distortion of the inherent geometry of WM.

This subchapter takes a more general look at the idea of WM-based spatial normalization of intra-subject MRI data series in order to identify ways of overcoming the outlined limitations of existing techniques. The aim of this review is the specification of a software prototype designed to experimentally test whether acquisition-related morphological variability can be further reduced with this normalization strategy given a WM/GM-CNR that is achievable in state-of-the-art MRI today. In practice, this means that it is investigated which well-established image processing algorithms are most promisingly combined into one processing pipeline to technically realize WM-based spatial normalization of intra-subject MRI brain data series. Consequently, the review of prior research in image-registration and -segmentation is focused on this specific scenario, as an exhaustive survey covering every aspect of these problems and every algorithmic solution proposed by the image processing community so far would be beyond the scope of this thesis.[22]

[22] For example, inter-subject or inter-modality registration is not covered, and neither are segmentation algorithms that require significant user interaction.

Regarding the structure of this subchapter, first a general outline of the complete processing pipeline necessary for WM-based spatial normalization is given. Subsequently, algorithmic solutions to the main technical challenges of this task, i.e. WM segmentation and registration are discussed. These reviews are focused in the just explained way, but are self-contained insofar as they include a short general introduction into the problem for the non-expert reader. Since WM-based spatial normalization essentially is a question of nonlinear registration, whereas WM segmentation – in comparison – rather is a preprocessing step, the review does not follow the processing pipeline's sequence of execution, but rather starts with the most important component, i.e. image registration.

Each of these separate reviews is concluded by a discussion of the presented methods' suitability for the given task. Finally, the concluding design of the prototype is presented.

4.2.2 General Overview

WM-based spatial normalization of brain images can be considered a special case of the previously outlined general concept of spatial normalization using intrinsic markers. Customizing the schematic illustration given in section 3.4.3.1 to this specific task, results in figure 4.1.

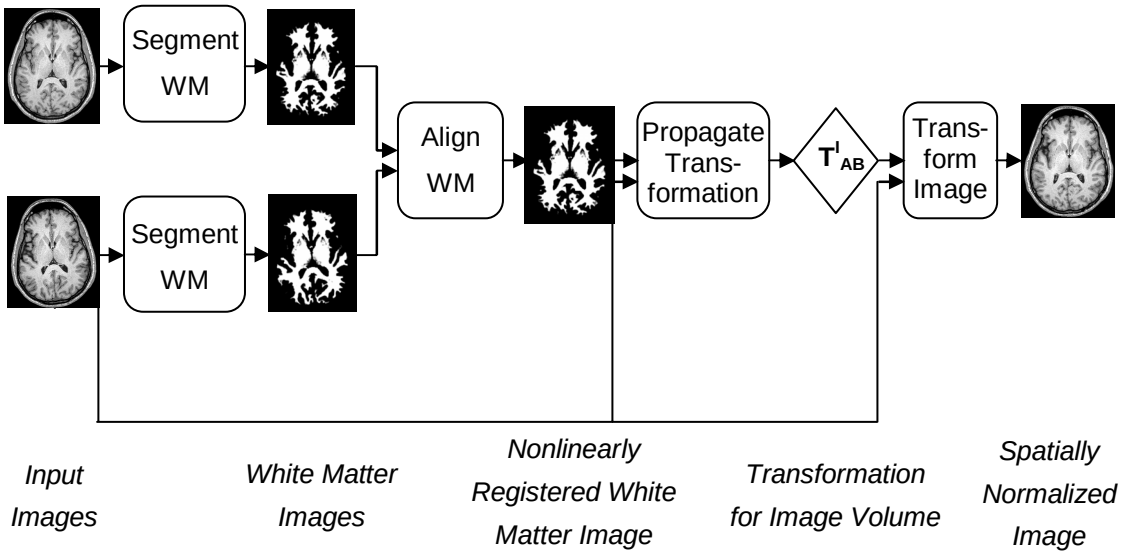


Figure 4.1: *Schematic illustration of WM-based spatial normalization of longitudinal brain images.*

The normalization procedure is best described as a processing pipeline that starts with WM segmentation, is continued with a nonlinear registration of these WM representations, and that is completed by a propagation of the local transformations to surrounding non-WM voxels such that the complete image volume can be spatially normalized. It is thereby desirable to overcome the previously outlined limitations of cortical-thickness approaches with respect to spatial normalization of subcortical GM regions. If implemented correctly, the resulting processing pipeline is capable of eliminating all acquisition-related morphological variability (contained in WM) while preserving differences corresponding to true anatomical changes (in GM), as long as WM itself is not noticeably affected by the pathology under investigation. It was

already shown in section 3.4.3.2 that this condition is fulfilled in early Alzheimer's disease.

4.2.3 Nonlinear Registration

4.2.3.1 Description of the Problem

Simplifying, the term medical image registration can be understood as the determination of a transformation that spatially aligns two images *A* and *B* such that the location of an anatomical feature in one dataset is mapped to the location of the corresponding feature in the other dataset.

A more exact definition requires formulating the relation of the actual physical object to its representation in image space. So far this relation has been referred to as data acquisition process, but from a mathematical point of view it can be considered a mapping of points in the object within its field of view Ω to intensities. Thus, *A* and *B* become

$$A : \vec{x}_A \in \Omega_A \mapsto A(\vec{x}_A) \tag{4.1}$$

$$B : \vec{x}_B \in \Omega_B \mapsto B(\vec{x}_B) \tag{4.2}$$

Based upon this, a transformation *T* from image *B* to *A* can be written as:

$$T_{AB} : \vec{x}_B \mapsto \vec{x}_A \Leftrightarrow T_{AB}(\vec{x}_B) = \vec{x}_A \tag{4.3}$$

The task of a registration procedure can now be described as determination of the transformation T_{AB} which – when applied to *B* – makes sure that $A(\vec{x}_A)$ and $B^T(\vec{x}_A)$ represent the same location in the imaged object.

When thinking about the nature of this transform T_{AB}, it is helpful to recapitulate the specific characteristics of MRI explored in chapter 3.2: In section 3.2.2, it was shown that the MRI data acquisition process does not preserve the morphological properties of an imaged object, but rather introduces geometrical distortions, whereby the underlying spatial transformation is nonlinear. In other words, the mappings *A* and *B* are – when thinking of them as spatial transformations from real space to image space – not adequately representable by a rigid-body or affine transformation.

However, image registration is not aiming at recovering the transformation from real space to image space, but rather at determining the transformation T_{AB} from image *B* to *A*. Thus, T_{AB} could well be correctly described by a rigid-body transformation, if the geometrical distortions introduced into the images by the MRI data acquisition process were perfectly reproducible in the imaged object's coordinate system. Recapitulating section 3.2.3, however, it can be stated that this reproducibility is not given even for a physically unchanged object. This is why T_{AB}, too, can only be correctly modelled as a nonlinear transformation.

4.2.3.2 Review

Image registration has been a field of active research over decades, and it is widely used in a great variety of different applications. Several surveys have been published that classify the various algorithmic solutions proposed by the image processing community according to criteria like the nature and domain of the transformation, the dimensionality of the patient image data, the modalities involved or the subject of the registration process, among them [Maintz98], [Zitova03] or [Pluim03]. This review is focused on intra-subject registration of intra-modal brain data series that do not

contain large (relative) deformations. Moreover, it is restricted to a discussion of techniques the implementation of which is practicable in the context of the development of a prototype.

When classifying image registration algorithms according to their approach towards the correspondence-problem, two classes remain:

- Feature-based techniques make use of image points, curves or surfaces representing the same physical or anatomical entity in *A* and *B*. As these features can be considered as sets of image points, they are here written as $M_A = \{\vec{x}_A^i\}_{i=1..n}$ respectively $M_B = \{\vec{x}_B^i\}_{i=1..m}$. Practically, T_{AB} is determined by first segmenting these features in both images and afterwards computing the transformation T_{AB}^M that brings them into best possible alignment[23]. Since T_{AB}^M is defined only for M_B, an interpolation step is required to yield T_{AB} that is defined for the entire domain Ω_B.
- In contrast to feature-based approaches, voxel-based registration methods do not explicitly delineate structures, but rather use statistical properties of the intensity distributions of anatomically corresponding regions. Generally spoken, they can be thought of as an iterative maximization of an intensity-based similarity measure computed over *A* and B^T in their common overlap region $\Omega_{A,B}^T$.

$$\Omega_{A,B}^T = \{\vec{x}_A \in \Omega_A \mid T_{AB}^{-1}(\vec{x}_A) \in \Omega_B\} \tag{4.4}$$

  Gradient-based approaches can also be regarded as voxel-based approaches as they usually evaluate the images' spatial derivatives at every voxel location in $\Omega_{A,B}^T$.

4.2.3.2.1 Landmark-Based Registration

A large number of registration methods have been published that are based upon the determination of well-suited image points in *A* and *B*, generally referred to as landmarks or homologous points. These landmarks do not necessarily have to be anatomically meaningful [Woerz03] [Frantz00], they can also be chosen purely depending on their geometrical characteristics [Thirion96]. However, they are always considered homologous point pairs, meaning that there is an intrinsic 1-to-1-mapping between them and that – following the notation given above – *n* equals *m*.

Linear registration of landmarks is well understood and widely used in image-guided surgery. Technically it is an optimal fitting problem in a least square sense [Arun87]: Following the notation given above, it is searched for that transformation T_{AB}^M which minimizes $D(T_{AB}^M)$:

$$D(T_{AB}^M) = \left\| T_{AB}^M(M_B) - M_A \right\|^2 \tag{4.5}$$

In nonlinear landmark-based registration, T_{AB}^M is implicitly determined in the segmentation step due to the intrinsic 1-to-1-mapping of the identified landmarks. When their segmentation is completed, registration, thus, becomes a question of

[23] The degree of alignment thereby strongly depends on the number of degrees of freedom of T_{AB}^M; it does not necessarily have to be spatial identity.

adequate interpolation. To prevent T_{AB} from containing unrealistic bending and folding, the interpolation typically is regularized such that no discontinuities in T_{AB} and its first order spatial derivates occur [Bookstein91] [Evans91].

Landmark-based registration methods are often used in constellations that impose high requirements on processing time. In these cases, few homologous points are identified, and the complete image volume is then mapped according to the transformation between these points.

4.2.3.2.2 Curve-Based Registration

A logical extension of the preceding approach is to segment anatomically corresponding curves rather than points. Historically, these curves used to be tissue borders identified in 2D slices, which is not the best basis for volume registration due to differences in contours merely resulting from differences in patient positioning and partial-volume effects. State-of-the-art registration methods relying on curve identification, thus, rather delineate prominent curves on 3D surfaces. As this implicitly requires that these surfaces are reconstructed beforehand, it is - in intra-subject registration - usually preferred to utilize the complete surfaces for determination of T_{AB} rather than relying on a few selected curves only, which is why they are not further considered here.

It seems to be worth mentioning, however, that curve-based techniques are used in inter-subject brain mapping in a procedure often referred to as cortical pattern matching: By determining the spatial correspondence of ridges on major gyri and fissures of important sulci that occur consistently in normal subjects, the cortices of different subjects are spatially aligned in consideration of their cortical variability [Thompson01a].

4.2.3.2.3 Surface-Based Registration

Surface-based techniques require to reconstruct the 3D cortical surface representations M_A and M_B from the subject's scans *A* and *B* prior to registration. Two different approaches for this task are known from literature:

- The first deforms a 3D surface with a fixed topology, in brain mapping typically a spherical mesh, into the shape of the anatomical entity of interest [Davatzikos96] [MacDonald00] [Thompson01a]. Usually, the deforming surface iteratively evolves to match a specified isointensity value in *A*, whereby it is simultaneously forced to fulfil additional constraints regarding its elasticity, rigidity, twisting (see section 4.2.4.2.2 for details), and topology. Most importantly, it is prevented from self-intersecting.
- The second technique identifies the anatomical structure first as a set of voxels, i.e. a result of a hard WM segmentation procedure, and then performs a surface tessellation on it [Fischl99a]. In contrast to the first approach, this technique does not implicitly guarantee that the topology of the resulting surface is identical to that of the template. If a spherical topology is required, this method therefore either involves a refinement of the reconstructed surface to remove handles and holes [Fischl01], or a manipulation of the segmentation result image to remove these topological "defects" [Bazin05].

Linear registration of surfaces is very similar to linear landmark-based registration, the main differences being the semantic connection between the vertices of each

surface and that M_A and M_B do not necessarily have to have the same number of elements. Consequently, the spatial distance between M_A and M_B cannot be computed using the distance measure given in section 4.2.3.2.1, but requires a different approach. The most popular technique for linear registration of surfaces is the so-called iterative closest point algorithm (ICPA) [Besl92]. It is a two stage procedure: First it is identified, for each point in M_A, the spatially closest point in M_B, and afterwards the linear transformation that best aligns this point set M_B^* with M_A in a least square sense is determined analogously to section 4.2.3.2.1. The algorithm terminates when the change in the mean squared error falls below a threshold. For triangulated surfaces, the closest point is often found by interpolating between the facets, such that M_B^* does not necessarily contain only members of M_B.

Nonlinear surface registration methods can be divided into two classes, namely purely surface-based approaches and hybrid techniques:

- The methodological principles of purely surface-based approaches have been presented in section 3.4.3.3 already which is why they are only briefly summarized here: These algorithms align surfaces in their 2D parametric domain, and therefore, require a uniform surface-based coordinate system. In brain mapping, this is usually chosen to be a sphere. Practically, the two surfaces M_A and M_B are transformed (by a nonlinear transformation *S*) into a common spherical coordinate system, and there $S(M_B)$ is reparameterized such that its similarity to $S(M_A)$ is maximized - which implicitly corresponds to a deformation of M_B into the shape of M_A. Further details as well as key publications on this topic can be found in section 3.4.3.3.
- Alternatively, M_B can be considered a deformable model and the registration process a special case of active contour-based segmentation. In practice, this strategy implies that M_A is first linearly transformed into *B* by applying T_{LAB}^{-1} (the inverse of the transformation resulting from linear registration of *B* to *A*). Subsequently, this linearly transformed surface is iteratively refined into the shape of the corresponding anatomical entity in *B* as described in section 4.2.4.2.2. If the number of surface vertices is kept constant during this deformation process, the differences in position between each pair of corresponding vertices form a 3D deformation field, which transforms each member of M_B into the anatomically corresponding location in *A*.

4.2.3.2.4 Similarity-Based Registration

As outlined above, similarity-based registration methods do not explicitly delineate well-suited features in image space, but rather use statistical properties of the intensity distributions of anatomically corresponding regions. Intuitively, these properties can also be interpreted as a mathematical characterization of image similarity.

There are various similarity measures both for intra- as well as for inter-modality scenarios, the most important of which are briefly introduced below. Further metrics can be found in [Holden00].

- Summed squared difference considers image similarity between *A* and *B* to be maximized, if their average grey value difference is minimized in $\Omega_{A,B}^{T}$. It is suitable only for intra-modality registration.

$$SSD = \frac{1}{N} \cdot \sum_{\vec{x}_i \in \Omega_{A,B}^{T}} \left(A(\vec{x}_i) - B^T(\vec{x}_i) \right)^2 \tag{4.6}$$

- Normalized cross-correlation understands image similarity as degree of positive linear dependence of intensities in the common overlap region. According to this measure, image similarity is maximized in case of positive linear dependence (*CC=1*) and minimized in case of negative linear dependence (*CC=-1*). Though not requiring the same anatomical unity to show identical signal intensity in *A* and *B*, it is also applicable only to intra-modality registration.

$$CC = \frac{\sum_{\vec{x}_i \in \Omega_{A,B}^{T}} \left(A(\vec{x}_i) - \overline{A} \right) \cdot \left(B^T(\vec{x}_i) - \overline{B^T} \right)}{\sqrt{\sum_{\vec{x}_i \in \Omega_{A,B}^{T}} \left(A(\vec{x}_i) - \overline{A} \right)^2 \cdot \sum_{\vec{x}_i \in \Omega_{A,B}^{T}} \left(B^T(\vec{x}_i) - \overline{B^T} \right)^2}} \tag{4.7}$$

- Mutual information is derived from information theory and considers image registration of two images *A* and *B* as the maximization of that information which is shared by both images [Wells96][Maes97]. According to Shannon, the information of a signal can be quantified by its entropy *H*, whereby p_i is meant to be the probability of the occurrence of event *i*:

$$H = -\sum_i p_i \cdot \log(p_i) \tag{4.8}$$

  Applied to the concept of mutual information of two images, this means that these images are independent from each other with respect to the information they contain, if their common entropy $H(A, B^T)$ equals the sum of their individual entropies $H(A)+H(B)$. In other words, the spatial transformation which maximizes the difference between $H(A, B^T)$ and $H(A)+H(B)$ can be considered the best possible reduction of the total amount of information in the combined image by elimination of redundancies. Thus, it can also be understood a maximization of the images' similarity. Mutual information can formally be written as

$$MI(A, B^T) = \sum_a \sum_b p_{AB^T}(a,b) \cdot \log \frac{p_{AB^T}(a,b)}{p_A(a) \cdot p_{B^T}(b)} \tag{4.9}$$

  Practically, its computation is based upon histograms: the marginal probability density functions $p_A(a)$ and $p_B(b)$ are estimated by the images' individual histograms, and their joint probability density function by dividing their bivariate histogram $p_{AB^T}(a,b)$ by the number of voxels in $\Omega_{A,B}^{T}$. Due to its very general understanding of image similarity, mutual information has become the most popular similarity measure in inter-modality registration.

Linear similarity-based registration of images is done by iteratively maximizing the similarity between *A* and B^T in their common overlap region $\Omega_{A,B}^{T}$, until the gain in similarity between two iterations falls below a threshold. Additional to the choice of a suitable similarity measure, it involves the implementation of both an appropriate optimization algorithm and an interpolation strategy.

Nonlinear similarity-based registration still is a field of active research. The most popular techniques known from literature are block matching and grid-based techniques:

- The basic idea behind block matching is to approximate the nonlinear deformation field by piecewise linear transformations [Roesch00] [Hufnagel05]: Practically, this concept involves the identification of those subvolumes in *B*, that are very likely to be accurately detectable in *A* due to their intensity distribution. Afterwards it is searched for the corresponding subvolumes in *A* allowing locally rigid or affine transformations [Roesch01]. Finally, these numerous local transformations T_{AB}^{M} are propagated over the image volume by either filtering with Gaussian-like kernels or by interpolation with spline functions, the most popular ones being thin-plate- and B-splines. An important implication of this approach is that information-theoretic similarity measures cannot be used, as the metric is evaluated locally.
- Grid-based approaches model the transformation T_{AB} as linear combinations of smooth basis functions placed on a regular grid of control points. The nonlinear registration process then becomes an iterative optimization of the basis functions' coefficients until the similarity between *A* and B^T (according to the implemented similarity measure) reaches its maximum. Since the grid spacing determines the number of degrees of freedom, its parameterization strongly depends on the nature of the process that is to be recovered by image registration: If it involves high-frequency deformations, the spacing needs to be small, otherwise a rather coarse grid is sufficient. The question which type of basis function to choose is rather an implementational one, but still of great practical importance: As the image similarity measure is evaluated in the entire overlap regions $\Omega_{A,B}^{T}$, basis functions with local support are computationally advantageous in constellations that require a small grid spacing, and thus, the iterative optimization of a large number of coefficients. This also explains the frequent use of B-splines [Rueckert99] [Studholme01] [Rohlfing01] and radial basis functions [Rohde03] in applications aiming at recovering every detail of the spatial deformation between *A* and *B*. If only coarse morphological differences are to be compensated for, it is a viable alternative to use basis functions not providing local support [Ashburner99].

If required these approaches can be further adapted to allow for rigid structures within deformable tissue through fixation of inter-control point distances [Little97] [Tanner00].

4.2.3.2.5 Gradient-Based Registration

Gradient-based approaches evaluate the images' spatial derivatives in order to bring the datasets into anatomical alignment. They are mostly used in an intra-modality context; for inter-modality registration a preceding intensity transformation is required [Guimond01].

The most popular representative of gradient-based registration is the optical flow method [Barron94][Periaswamy03]. It models the differences between two images *A* and *B* as fluid particles that flow through the image volume, whereby it assumes that no compression or expansion of structures has happened within the local aperture.

Mathematically, optical flow can be expressed as a movement of each voxel in the direction of its local spatial gradient, until the local time gradient vanishes.

$$dg(x, y, z, t) = \frac{\partial g}{\partial x} \cdot dx + \frac{\partial g}{\partial y} \cdot dy + \frac{\partial g}{\partial z} \cdot dz + \frac{\partial g}{\partial t} \cdot dt = 0 \qquad (4.10)$$

Gradient-based methods usually do not concentrate on a set of control points, but are designed to determine movements wherever it is possible with respect to the images' intensity structure; thus, they intrinsically generate comparatively densely filled deformation fields. However, they still require an interpolation strategy for homogeneous image regions, the local spatial intensity structure of which do not allow for a direct determination of the local transformation. Like other iterative procedures, gradient-based methods are usually terminated when the change in difference between A and B^T falls below a threshold.

4.2.3.3 Discussion

Having outlined the underlying principles of the most important approaches in medical image registration, these are now assessed with respect to their suitability for WM-based spatial normalization of brain images.

Unfortunately, the experimental computer vision community has not succeeded yet in establishing reference-procedures and -image-databases for a systematic and unbiased validation of the various nonlinear registration algorithms, to the effect that it is not yet possible to evaluate the presented registration approaches on the basis of an objective "benchmark". Their suitability for the intended prototypical implementation of WM-based spatial normalization is therefore assessed based upon theoretical considerations. Implementational aspects and performance issues are thereby not equally strongly weighted as considerations regarding (a) the methods' general capability to satisfy this application's high demand for accuracy, and (b) potential risks that are specific to a certain technique, and thus, avoidable when choosing a different approach.

Because of its high demand for accuracy, nonlinear registration of WM as proposed in section 4.2.2 requires a direct determination of acquisition-related morphological changes present in WM in the complete brain. For landmark-based techniques this implies that a large number of homologous points would have to be accurately segmented in A and B to avoid errors in the overall deformation field resulting from interpolation. This is because the reliability of a local transformation $\vec{u} \in T_{AB}$ at the location $\vec{x} \in \Omega_B$ that is computed by interpolation between a subset of transformations $U \subseteq T_{AB}^M$ determined for the support points $X \subseteq M_B$ generally decreases with the spatial distance of $\vec{x}$ to the elements of X.[24] Since it is very difficult to accurately segment such a large number of landmarks in medical datasets in the presence of image noise, partial volume effects, and differences in spatial discretization due to patient repositioning, landmark-based techniques are not very promising in the context of WM-based spatial normalization.

Registration techniques relying on the introduction of a common surface-based spherical coordinate system do not seem to be the methods of choice for eliminating acquisition-related morphological variability from longitudinal intra-subject MRI data, since they substantially distort the inherent geometry of the original WM surface

[24] Of course, this is a very general view on reliability in the context of interpolation. Practically, there are a lot more influencing factors among them the number of support points and the choice of the interpolation basis function. The spatial distance to the support points often is incorporated into the strategy of determining weighting factors for the base functions.

representations. As stated earlier, it is questionable whether such approaches – in intra-subject scenarios - yield as accurate results as those obtained with registration techniques that do not require an intermediate mapping of the original WM surface to a sphere. The presented alternative surface-based registration technique via deformable surface segmentation evidently overcomes these drawbacks. However, it still bears a risk that all surface-based techniques are accompanied by due to the nature of surface reconstruction. Though often considered a purely computational challenge, it is all but trivial when used in applications that impose high requirements on accuracy and reproducibility. As pointed out in section 2.2.7 already (and illustrated in figure 4.2 again), different interpretations of voxels containing more than one tissue class can seriously affect the shape of a reconstructed surface[25]. In longitudinal imaging scenarios, slight differences in partial volume effects, e.g. caused by a change in patient positioning, in the choice of the imaging field of view, or in acquisition-determined geometrical distortions, are very likely to additionally complicate accurate and reproducible surface reconstruction. In total, these considerations lead to the conclusion to avoid the additional risk that surface-based approaches bear.

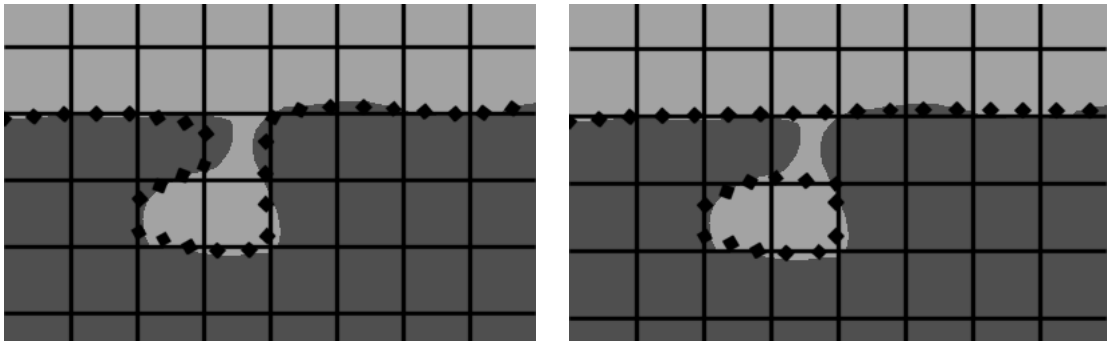


Figure 4.2: *Different interpretations of voxels containing more than one tissue class can seriously affect the shape of a reconstructed surface (here: dotted line). In longitudinal imaging scenarios, slight differences in partial volume effects additionally complicate accurate and reproducible surface reconstruction.*

Both similarity- and gradient-based techniques work directly on the original images respectively on the segmented images, and thus, are free from potential risks resulting from a transfer of theirs into a different representation. Moreover, they are superior to landmark- and curve-based approaches in terms of the achievable registration accuracy, as they do not rely on a few selected image features only.

Within this group of voxel-based methods none of the outlined approaches can be called superior to the others with respect to their suitability for WM-based spatial normalization, as – from a general point of view – all of them locally evaluate change in image similarity. This is obvious for block matching and optical flow techniques, but holds true also for grid-based approaches when looking at them in more detail. It was pointed out earlier that grid-based nonlinear registration methods aiming at capturing the details of a complex spatial transformation model the underlying deformation field as a linear combination of basis functions with local support. Changing one of the basis functions' coefficients, thus, becomes a local adaption of

[25] This argument is supported by the fact that in neuroimaging surface reconstruction techniques often require additional processing steps to correct for "defects" in the resulting surfaces' topology. From an abstract point of view, the methods for removing these incorrect handles and holes – no matter whether it is done surface-based [Fischl01], or by a manipulation of the segmentation result images [Bazin05] - do not truly improve segmentation results, but rather manipulate them for purely topological reasons.

the deformation field, the effect of which on global image similarity is evaluated locally only, i.e. in the respective basis function's support region.

The decision which voxel-based method to choose, thus, more or less seems to be a question of implementation. For this thesis it was decided to exemplarily implement the nonlinear registration step within WM-based spatial normalization as a process of optical flow because of the following reasons:

- In order to meet this application's strong demand for accuracy it is important to directly determine the local transformation within $\Omega_{A,B}^{T}$ wherever possible rather than unnecessarily relying on the appropriateness of an interpolation between a set of control points. Optical flow does this wherever the images' local spatial intensity structure allows for a computation of the local transformation.
- Due to its mathematical simplicity and ease of implementation, image registration based on optical flow can be tracked and understood comparatively easily. This is helpful in the course of development especially in the previously outlined context of a prototypical implementation of WM-based spatial normalization of intra-subject brain images.

Still, it is important to point out that a different voxel-based approach, in principle, is equally suitable for this task.

4.2.4 WM Segmentation

4.2.4.1 Description of the Problem

In medical imaging, segmentation is generally considered as the delineation of anatomical structures or tissue types in patient image data, whereby it is often distinguished between hard and soft segmentation:

Hard segmentation means delineation in a strict sense: the partitioning of an image A into a number of non-overlapping regions A_i such that each voxel belongs to exactly one region which in the ideal case represents one object:

$$A = \bigcup_{i=0}^{n} A_i \tag{4.11}$$

Soft segmentation algorithms extend this concept of disjoint image partitions and aim at determining - for each voxel location $\vec{x}_j$ - the probability p_{ji} of it belonging to an object respectively class C_i, whereby for each voxel, equation 4.12 must hold:

$$\forall \vec{x}_j \in A : \sum_{i=0}^{n} p_{ji} = 1 \tag{4.12}$$

Soft segmentation techniques, thus, automatically take into consideration partial volume effects that arise from a voxel containing more than one anatomical structure. For accurate image-based volumetry, soft segmentation methods therefore generally are superior to hard segmentation techniques particularly in the case of anisotropic voxels.

4.2.4.2 Review

Since segmentation is one of the most common tasks in clinical practice and medical research, a great variety of different algorithms have been proposed, which is why an exhaustive review covering every variant of the underlying concepts is beyond the scope of this work. As mentioned above, this review is therefore focused on those segmentation methods that are well suited for the previously outlined prototypical

implementation of WM-based spatial normalization: This implies that only those concepts are discussed that are well established in tissue segmentation of MRI brain images and that do not require substantial user interaction.

Furthermore, the review concentrates on the essence of segmentation and does not include an inspection of additional preprocessing steps that are required to correct for imaging artefacts that otherwise would corrupt the segmentation procedure. Most importantly in this context, it is assumed that the MRI specific image nonuniformity artefact already has been compensated for, e. g. by using a method such as [Sled98b], [Styner00] or [Jellus05].

Generally spoken, all segmentation techniques have in common that they assume an anatomical entity to be represented in image space as a homogeneous region that is distorted primarily by signal noise. However, they differ strongly in the image features they are guided by.

4.2.4.2.1 Region-Based Segmentation

Region-based techniques segment an image by directly identifying homogenous regions within it. The most popular representatives of region-based segmentation algorithms are region-growing, the watershed-transformation and histogram-based methods:

- Region-growing usually begins with the localisation of seed points in distinct image regions. Starting from these, it is iteratively grown into their direct neighbourhoods whereby only those voxels are added to a region that meet the region's membership condition. In its simplest form, this membership is encoded as a static intensity interval, but it can also be implemented more flexibly as a dynamic evaluation of the voxel's intensity with respect to the region's current intensity distribution: If a voxel's probability of class membership is sufficiently large, it is considered a new member, added to the region, and the class-specific intensity distribution is updated [Pohle01]. The segmentation process ends if either no unclassified voxels remain or if none of the regions' neighbouring voxels fulfils their membership conditions any more.

  Transferred to WM segmentation in MRI brain images, region-growing requires robust automated seed point identification. Since WM is spatially connected to the eyes, and thus, to the skull base by the optical nerve which – in T1-weighted images – shows similar intensity values as WM, region-based WM segmentation requires additional processing to prevent it from leaking into non-WM parts of the head with similar intensity. A widely used solution to this problem is skull-stripping that is typically done by means of deformable models [Smith02b] (see section 4.2.4.2.2), or by application of the watershed-transform (see below).

- The watershed-transformation [Roerdink01] considers an image derived from the original dataset *A*, usually either its gradient magnitude image [Sijbers97] or its inversion [Andrade97], as a topographic relief where the intensity of each voxel is interpreted as its physical elevation. The segmentation procedure can intuitively be described as first piercing holes at each local minimum of the topographic relief, and then immersing the surface into water to the effect that each basin is flooded starting with the global minimum. Whenever two basins begin to merge, a barrier is built between them, and thus, the overall procedure results in a partitioning of the image in many basins separated by barriers respectively watersheds. It is intuitively clear that the watershed-transform generally yields an

oversegmentation in the presence of image noise. The final segmentation therefore requires to subsequently merge the atomic basins produced by the watershed-transform into anatomically meaningful semantic regions either in a semi-automatic manner or by specification of merging rules.

- Histogram-based segmentation algorithms partition the image into distinct regions without explicitly considering spatial connectivity, but rather by classifying each voxel solely based upon its intensity. In the simplest case, histogram-based segmentation is interval thresholding. A popular algorithm for automatic threshold determination is maximization of the inter-class variance V as seen in equation 4.13, where p_i is the probability of class C_i, μ_i its intensity mean, and μ the image's total intensity mean [Otsu79].

$$V = \sum_{i=0}^{n} p_i \cdot (\mu_i - \mu)^2 \tag{4.13}$$

  More elaborate approaches model the image's intensity distribution as a mixture of several Gaussian distributed probability density functions $G_i(\mu_i, \sigma_i)$ as given in equation 4.14, where each component's portion of the total intensity distribution is given by p_i. In the analysis of brain images p_M typically incorporates three Gaussian distributed components, one for each WM, GM and CSF.

$$p_M = \sum_{i=0}^{n} p_i \cdot G_i(\mu_i, \sigma_i) \tag{4.14}$$

  The segmentation process then becomes an optimization of the parameter vector $\Theta = \{\mu_0, \sigma_0, p_0; \ldots; \mu_n, \sigma_n, p_n\}$, such that the resulting model p_M best explains the image's true intensity distribution. The most commonly used estimation strategy is the so-called expectation maximization (EM) algorithm first proposed by [Dempster77]. It maximizes the log-likelihood of Θ in an iterative manner, whereby each iteration consists of two processing steps: Roughly spoken, the E-step computes the class probability of each voxel based on the current estimate of Θ^t, whereas the M-step computes the new expectation of Θ^{t+1} based on the results of the previous E-step.

In summary, region-growing and watershed-based segmentation result in a set of non-overlapping components which makes them typical representatives of hard segmentation. Given the high requirements that WM-based spatial normalization imposes on accuracy and reproducibility, these methods would require further processing steps that explicitly evaluate the WM border region with respect to partial volume effects. In contrast, the EM algorithm in conjunction with modelling the image's intensity distribution as a mixture of Gaussian-distributed probability density functions (EM-GMM) implicitly produces a soft segmentation.

4.2.4.2.2 Edge-Based Segmentation

Whereas region-based segmentation relies on a mathematical description of image homogeneity, edge-based techniques are guided by the image's local intensity gradient. Simplifying, these approaches can be characterized as methods aiming at the identification of anatomically meaningful edges. In the presence of image noise, however, pure edge detection often results in additional contour-sections that do not reflect a true anatomical boundary. Therefore, edge detection is often combined with

either isotropic [Marr80] [Canny86], or anisotropic [Perona90] spatial smoothing. If the CNR between the target object and neighbouring tissue is not sufficiently strong, edge detection algorithms in general yield incomplete contours that need to be completed in a further processing step

Another concept to overcome the outlined difficulties is the so-called active contour technique first proposed by [Kass88]. Its basic idea is to iteratively deform a set of 2D-contours or a 3D-surface [Cohen91] into the shape of the object of interest, whereby the deformation process is modelled as the minimization of an energy functional that is determined by the antagonistic interplay of external and internal forces (see equation 4.15): Whereas the former – usually determined by evaluating the image's intensity gradient magnitude (see equation 4.17) – locally attract the surface, the latter define its mechanical properties, i.e. to what extent it is allowed to stretch (w_{10}, w_{01}), bend (w_{20}, w_{02}) and twist (w_{11}) (see equation 4.16). The internal force, thus, acts as a smoothness term in this regularization context. The deformation process stops when the energy functional reaches a local minimum (which ideally is also its global minimum). A more exact description requires to describe the surface $\vec{v}$ in its 2D parametric domain $s,r \in [0,1]$, i.e. $\vec{v}(s,r) = (x(s,r), y(s,r), z(s,r))$ [Cohen91].

$$E(\vec{v}) = \alpha \cdot E_{ext}(\vec{v}) + \beta \cdot E_{\text{int}}(\vec{v}) \rightarrow \min \tag{4.15}$$

$$E_{\text{int}}(\vec{v}) = \int_0^1 \int_0^1 w_{10} \left|\frac{\partial \vec{v}}{\partial s}\right|^2 + w_{01} \left|\frac{\partial \vec{v}}{\partial r}\right|^2 + w_{11} \left|\frac{\partial^2 \vec{v}}{\partial s \partial r}\right|^2 + w_{20} \left|\frac{\partial^2 \vec{v}}{\partial s^2}\right|^2 + w_{02} \left|\frac{\partial^2 \vec{v}}{\partial r^2}\right|^2 dsdr \tag{4.16}$$

The external force couples the contour to the image and usually contains a term directing $\vec{v}$ to strong edges with maximal gradient magnitude.

$$E_{ext}(\vec{v}) = \int_0^1 \int_0^1 - w \cdot \nabla \left[A(\vec{v}(s,r))\right]^2 dsdr \tag{4.17}$$

The accuracy of active contour methods is known to be sensitive to both a good starting position and the shape of the initial contour. Its starting state, thus, is often defined by incorporating a probabilistic atlas (see section 4.2.4.2.3). For WM segmentation, however, the biggest challenge is an adequate formulation and weighting of the driving forces, because the anatomical WM boundary contains a lot of high curvature parts that are difficult to accurately capture with active contour models, which in general favour globally smooth contours as solutions due to the built-in smoothness term.

4.2.4.2.3 Atlas-Based Segmentation

Atlas-based methods are closely linked to image registration. In short, they consist of a – usually nonlinear – transformation of the patient image data that is to be segmented onto a reference atlas that already has been segmented and classified such that the location of every anatomical feature in the first dataset is mapped to the location of the corresponding feature in the atlas. Applying the inverse of that transform leads to an anatomical labelling of the previously unsegmented dataset.

In order to be able to cope with anatomical variability, a certain amount of variability is usually incorporated in the reference dataset by modelling it as a probabilistic atlas.

Still the initial segmentation results often require an additional refinement after registration of the atlas to *A*:

- One common strategy of correction is the application of active contour models that use the initial labelling as their starting position and iteratively refine the segmentation as described in section 4.2.4.2.2.
- Another popular method consists of the incorporation of spatial neighbourhood relations by making use of Markov random fields. The atlas proposed by Fischl and colleagues for segmentation of deep brain structures [Fischl02], e. g. encodes for each atlas location not only (a) the probability for each anatomical structure, but also (b) the intensity distribution of each anatomical structure, and (c) the joint probability that an anatomical structure is the neighbour at the topical atlas position given its current anatomical label. Given these distributions in atlas space the segmentation is performed by first linearly mapping the dataset onto the probabilistic atlas and then iteratively maximizing the probability of the current labelling by modelling it as a non-stationary Markov random field.

4.2.4.3 Discussion

The assessment of the outlined approaches' suitability for WM segmentation in the context of WM-based spatial normalization of MRI brain images is a difficult task, because state-of-the-art brain segmentation often incorporates several of the presented methods to overcome specific weaknesses of a single method. Recently, several different combinations have been used in large clinical studies in a fully automated way, and an evaluation of theirs with respect to accuracy and reproducibility has shown that they provide very similar results [Fernandez05]. Nevertheless, what was said on validation in the context of nonlinear registration holds true also for image segmentation: No reference-procedures and -image-databases have been established yet to allow for a systematic and unbiased validation of the various segmentation algorithms, to the effect that it is not yet possible to evaluate them on the basis of an objective "benchmark". Consequently, this review, again, is especially focused on potential risks that the different approaches impose on the validity of results obtained with them due to their distinguishing features. Additionally, it is taken into consideration how much user-interaction the presented methods require, since the segmentation is desired to work fully automated in the context of the targeted prototypical implementation of WM-based spatial normalization.

Active contours respectively deformable surfaces are a powerful technique to deal with variability in human anatomy, especially in combination with probabilistic atlases as outlined in 4.2.4.2.2 and 4.2.4.2.3. The shape of WM, however, is very complex due to the cortical folding, and thus, the weighting of the external and internal forces becomes a very difficult task. As pointed out earlier, active contour techniques generally favour globally smooth contours as solutions due to the built-in smoothness term; hence, it is questionable whether they are capable of adequately capturing narrow gyri without simultaneously smoothing neighbouring sulci in the context of automated high-accuracy WM segmentation. In addition to the previously stated considerations with respect to potential risks that surface-based approaches bear as a result of the discrete nature of medical images (see section 4.2.3.3), this suggests to choose an alternative WM segmentation technique in the context of WM-based spatial normalization of MRI brain images.

It was mentioned that atlas-based approaches usually comprise an additional processing step to refine the initial segmentation as resulting from the transformation

of the unsegmented dataset to the template atlas – typically the application of either deformable models or Markov random fields. The risks accompanied by the use of deformable models in WM segmentation have been discussed above already. Markov random fields on the other hand are effective especially in segmentation of deep brain structures that show poor CNR in image space, because a segmentation of theirs then greatly benefits from incorporation of model knowledge in terms of neighbourhood relations. However, for a prototypical implementation of WM-based spatial normalization of brain images, it is a viable strategy to request a CNR between WM and GM in the test datasets that is technically achievable with state-of-the-art scanners and imaging protocols. Hence, it does not seem to be required to apply the comparatively complex combination of nonlinear-atlas registration and subsequent Markov random field optimization, especially if simpler methods are available.

Out of the presented region-based segmentation methods, watershed-based segmentation seems to be least suitable for the given task because of its sensitivity to signal noise. Given the task's high requirements on accuracy, it can be assumed that the correct merging of the atomic basins produced by the watershed-transform into one region, which accurately represents WM, requires either significant user interaction or elaborate knowledge-based approaches (e.g. [Koethe06]).

In contrast, region-growing and the EM-GMM method can be implemented more easily to segment WM in a fully automated fashion. In practice, these approaches merely require an effective skull-stripping procedure prior to segmentation in order to prevent the segmentation from being confounded by non-brain tissue[26]. As WM usually can be regarded as one spatially connected structure after skull-stripping, region-growing would only require the automatic identification of a single seed point that is guaranteed to be located within WM, and the EM-GMM method would merely need a reasonable initialisation of the parameter vector Θ. The latter even offers the possibility to incorporate prior knowledge about both the classes' probabilities and - if the image's contrast characteristics are known - their intensity distributions [Pohl02].

Another advantage of the EM-GMM method over region-growing is its reduced sensitivity to ringing-like motion artefacts in MRI data. As shown in figure 4.3, these phenomena can destroy the spatial connectivity of WM in image space, which would require the interactive identification of one additional seed point for each spatially disconnected WM region if utilizing region-growing. As EM-GMM does not require spatial connectivity, it can be assumed to be more robust in the presence of motion artefacts[27].

[26] It was mentioned before that in original T1-weigthed MRI data, WM is spatially connected to the eyes, and thus, to the skull base by the optical nerve that shows intensity values similar to those of WM.

[27] One can argue, again, that it is a viable strategy to assume patient image data for a prototypical implementation of WM-based spatial normalization, which are not affected by motion artefacts. However, these artefacts are very likely to be present in clinical datasets of dementia patients due to their reduced ability to cooperate during imaging. Therefore, it might be unrealistic to require clinical data series without motion artefacts.

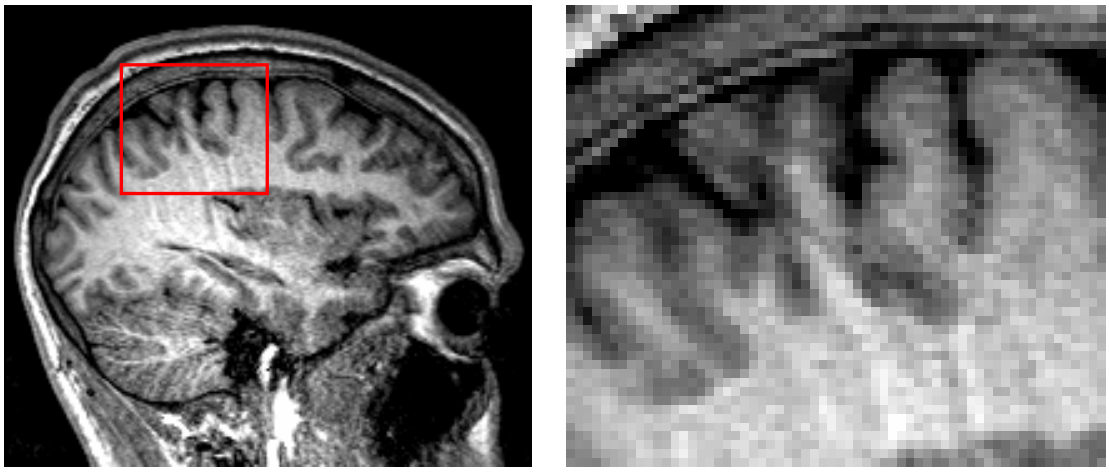


Figure 4.3: *Motion artefacts can destroy - in image space - the spatial connectivity of partially thin structures like WM (the highlighted window is enlarged on the right-hand side). The dark ringing-like motion artefact separates the WM portion of the gyrus in the window's centre into two spatially disconnected regions.*

In summary, this makes EM-GMM the most suitable method for WM segmentation for the targeted prototypical implementation of WM-based spatial normalization of MRI data: If combined with a skull-stripping procedure, it can be implemented as a fully automated soft segmentation method that does not impose any smoothness constraints on the WM surface, is comparatively robust to motion artefacts and has proven to be comparable in the achievable accuracy to other state-of-the-art segmentation procedures [Fernandez05].

4.2.5 Discussion

In this subchapter, it was taken a general look at the idea of WM-based spatial normalization of intra-subject MRI data series with the aim to technically specify a prototypical implementation of this method that overcomes the previously outlined limitations of cortical-thickness-based approaches. In practice, it was investigated in what way which well-established image processing algorithms are most promisingly combined into one processing pipeline to technically realize WM-based spatial normalization of single-subject MRI brain data series.

The review of prior research in image-registration and -segmentation was focused on this specific scenario and resulted in the selection of methods that were identified as well-suited when being combined as outlined in figure 4.1. Concretely, the EM algorithm in conjunction with modelling the brain's intensity distribution by means of a Gaussian mixture model was chosen for WM segmentation, and optical flow for nonlinear registration. A detailed reasoning for this choice is given in the respective discussions (see sections 4.2.3.3 and 4.2.4.3) and is not repeated here. The resulting processing pipeline can be regarded as an extension of cortical-thickness-based approaches since this method neither is restricted to a normalization of cortical GM regions, nor requires a distortion of the original morphology of WM into a sphere.

Instead of summarizing the individual reviews, this section builds upon them and draws conclusions for the complete processing pipeline. In doing so, it refines the so-far rather coarse schematic illustration of the sequence of processing steps, e.g. by incorporating those preprocessing steps that have been shown to be indispensable for EM-GMM-based WM segmentation.

In practice, the segmentation procedure itself also has to be a processing pipeline consisting of the sequential execution of intensity non-uniformity correction, skull-

stripping, brain segmentation and tissue classification (see figure 4.4), whereby the order of the first two processing steps is dependent on the actual method used for bias field correction.

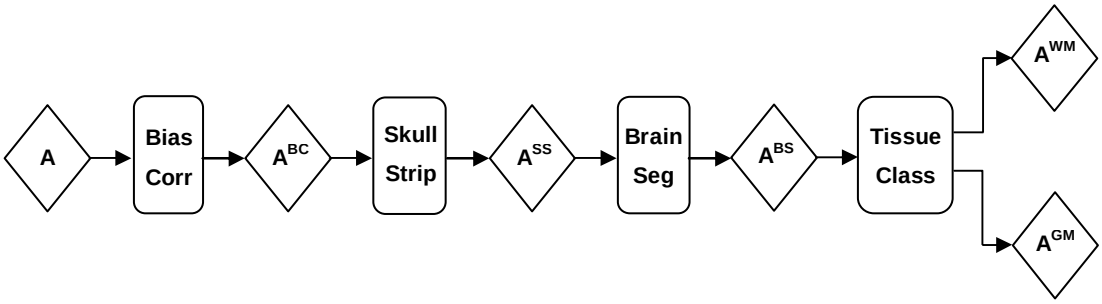


Figure 4.4: *Schematic illustration of the sequence of processing steps in WM segmentation (details are given in the text).*

As mentioned before, correction of intensity non-uniformity is required because region-based segmentation assumes that a tissue type shows homogenous intensity in image space - and a significant difference in intensity with respect to other tissue types. This condition is not perfectly met in original MRI data [Sled98a]. The two most prominent factors that lead to this low-frequency intensity drift in image space are:

- Inhomogeneous radio-frequency receiver coil profiles present themselves as a dark centre and bright periphery pattern, as the individual coil elements are more sensitive to signal from regions that are spatially close to their position on the head-coil.
- In contrast, dielectric resonances generally lead to a bright centre and dark periphery pattern in the head; they occur in 3T data only, as the radio-frequency wavelength at 3T is half that at 1.5T, and thus, in the range of the head's extension, so that standing waves can arise.

In the outlined segmentation pipeline it is explicitly distinguished between skull-stripping and brain segmentation, indicating that skull-stripping, here, is understood in a strict sense: It eliminates large portions of non-brain tissue in image space, e.g. the eyes, facial muscles, skin and subcutaneous fat, but not necessarily completely. Especially the meninges are usually not reliably removed from MRI head images by skull-stripping, as they are spatially very close to GM and show very similar signal intensity. In order to prevent histogram-based segmentation techniques from being confounded by non-brain tissue, an additional brain segmentation step is required.

As mentioned before, the segmentation procedure is completed by tissue classification according to the EM-GMM algorithm. As indicated in figure 4.4, this classification is understood in a soft manner; hence, one result image is created for each tissue type separately, here GM and WM.

In the context of WM-based spatial normalization, these WM images serve as input data for the subsequent nonlinear registration procedure. For both stability as well as performance reasons, it is advisable to incorporate a linear registration in this nonlinear spatial normalization (see figure 4.5).

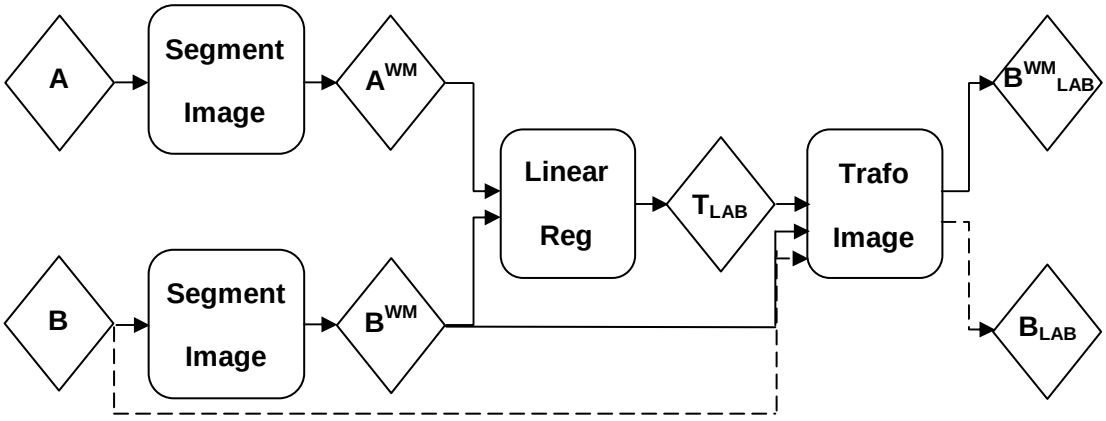


Figure 4.5: *Schematic illustration of the data flow in linear registration; the original dataset B is linearly registered by application of the same transform T_{LAB} as determined for the WM images.*

Regarding this linear alignment, it is important to point out that the outlined strategy ideally gives the best result achievable with linear registration, as it relies solely on WM. Both non-brain tissue like the eyes, as well as GM that might have undergone a morphological change due to a beginning neurodegenerative disease would corrupt the correctness of the transform if the original images were used as input data.

Having rigidly aligned the images, their nonlinear registration follows. As optical flow evaluates the images' local intensity derivatives, homogenous image regions do not provide any useful information for local determination of movements, and thus, can be excluded from the region of interest: Therefore, optical flow is evaluated only at the border of WM to GM. In order not to favour one dataset over the other, the region of interest M^I_{AB} is determined as the unification of the border regions as identified in each image separately (see figure 4.6).

As mentioned before, the locally determined translations finally need to be propagated to surrounding non-WM voxels such that the complete image volume can be spatially normalized.

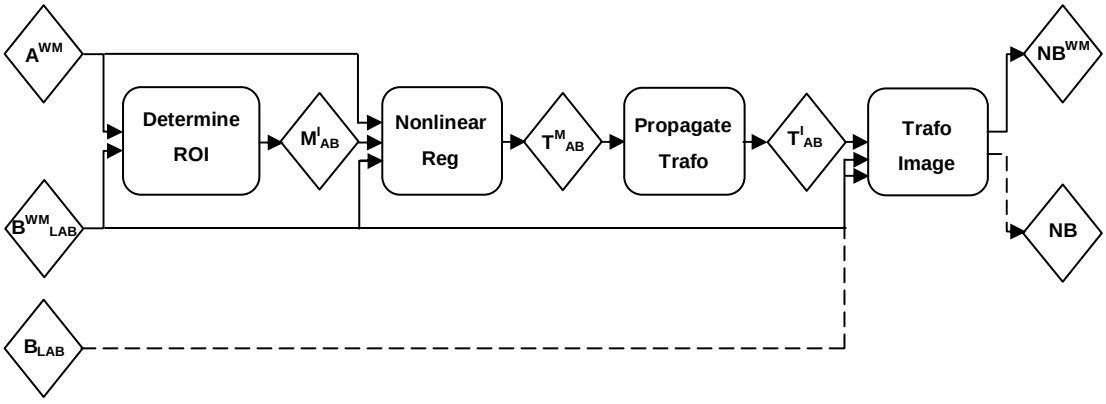


Figure 4.6: *Schematic illustration of the data flow in nonlinear registration: The nonlinear alignment is based upon the spatial intensity structure in the border region of WM and GM. Having detected morphological changes in this region of interest, the respective transformation T^M_{AB} is propagated to allow for a nonlinear registration of the complete image volume. By concatenating the previously determined linear transformation T_{LAB} with T^I_{AB}, the original dataset B is nonlinearly spatially normalized based upon WM only.*

In the following subchapter, the general specification of the targeted software prototype for WM-based spatial normalization will be completed by a description of its technical details.

4.3 WM-Based Spatial Normalization – Technical Details

4.3.1 Introduction

In the previous subchapter, a focused look was taken at the state of the art in image registration and WM segmentation in order to identify suitable methods for a prototypical implementation of WM-based elimination of acquisition-related morphological variability from longitudinal MRI brain data series. In detail, the expectation-maximization algorithm in conjunction with modelling the brain's intensity distribution by means of a Gaussian mixture model (EM-GMM) was chosen for brain tissue segmentation, and optical flow for nonlinear registration of the segmented WM representations. Based upon this selection, the concluding design of the overall prototype was presented in terms of a detailed processing pipeline including all required additional preprocessing steps.

In the following sections, this general specification is refined by a discussion of the prototype's technical details. Since some of the processing steps are practically solved by means of well-established standard procedures, these are not presented here to the last detail, but still profoundly enough to allow for a solid understanding of the methods' functionality. Further details can be easily found out about by studying the referenced original work. Due to its central importance for the complete normalization strategy, the main focus in this subchapter is set on nonlinear registration.

4.3.2 Segmentation

As shown in the previous chapter's conclusions, WM segmentation based upon the EM-GMM algorithm requires several preprocessing steps, and thus, can be considered a processing pipeline itself consisting of the sequential execution of skull-stripping, intensity non-uniformity correction, brain segmentation and tissue classification (see figure 4.7).

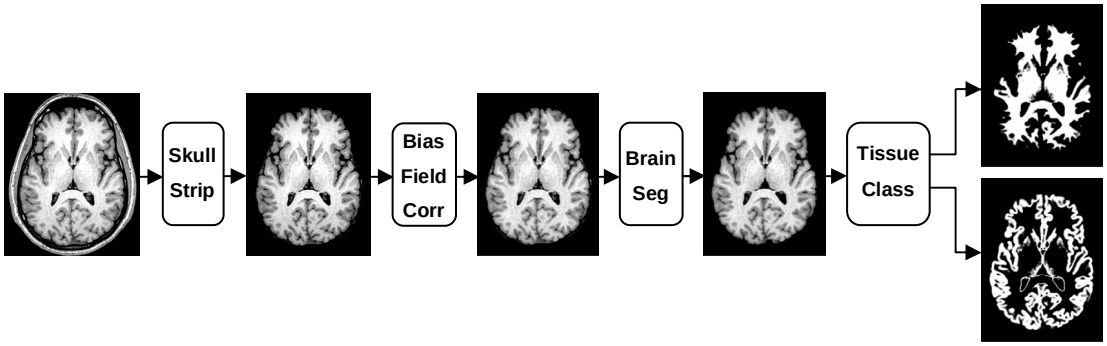


Figure 4.7: *Sequence of processing steps in WM segmentation (details given in the text).*

4.3.2.1 Skull-Stripping

The skull-stripping problem is solved by means of the method proposed by [Hahn00]. It is a variant of the watershed-transform (see section 4.2.4.2.1), which takes the inverted original image as topological relief, and is additionally tailored such that neighbouring catchment basins are merged only if their volume and their difference in minimum height do not exceed user-defined thresholds. Finally, the brain is identified as the large central catchment basin. The thresholds can be chosen to be robust against changes in the MRI acquisition protocol, as long as the general image contrast is not altered. For T1-weighted images acquired with the protocols specified in section

3.3.4, a volume threshold of 5% of the complete image volume, and a height-difference of 10% of the input image's intensity range yield results as shown in figure 4.8.

Figure 4.8: *Representative result of watershed-based skull-stripping as proposed by [Hahn00]: original image (left), catchment basins after watershed transform (middle), original image masked with large central basin (right)*

4.3.2.2 Intensity Non-Uniformity Correction

Since EM-GMM-based segmentation assumes that a tissue type shows homogenous intensity in image space – along with a significant difference in intensity to other tissue types - MRI data needs to be corrected for intensity non-uniformities prior to segmentation [Sled98a]. Several methods dedicated to the solution of this problem are known from literature, among them those proposed by [Sled98b], [Styner00] or [Jellus05]. In the prototypical implementation of WM-based spatial normalization at hand, it is made use of a variant of the latter approach which generally can be described as homomorphic filtering that makes special arrangements to prevent the filtering from being confounded by image regions with very low signal intensity: First, image regions with low signal intensity are segmented by means of Otsu-separation [Otsu79] and added to the background as identified by skull-stripping beforehand. The remaining coarse WM mask is then dilated whereby each background voxel is assigned the mean value of n neighbouring voxels ($20<n<80$) showing sufficiently high intensity. Remaining background voxels are assigned the image's global intensity mean to avoid sharp intensity changes at borders. The subsequent homomorphic filtering is parameterized to suppress low frequency components while not altering high frequency portions, such that the ratio of the filtered and the original image can be used as non-uniformity correction field (see figure 4.9). Since the bias field has only low-frequency components, the outlined procedure can be combined with downsampling without compromising the method's effectiveness. The low pass filtering itself is done in the frequency domain, and thus, involves the application of discrete Fourier transformation (DFT). The cutoff-frequency was determined empirically ($f_c = 25$ for a downsampled image size of 64^3 voxels).

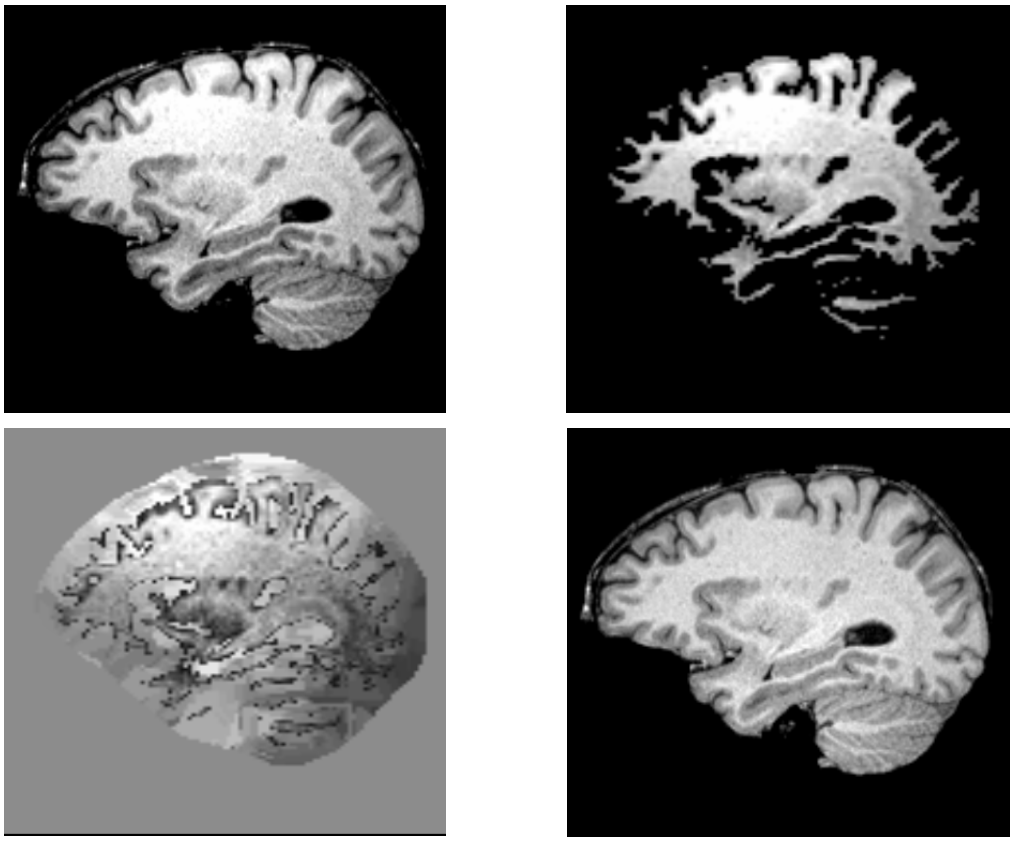


Figure 4.9: *Intensity non-uniformity correction by homomorphic filtering aims at eliminating low frequency intensity drifts while not altering high frequency components. As outlined in the text, the method is tailored to use WM as normalizing structure: skull-stripped original image (upper left), coarse white matter mask (upper right), dilated white matter mask (lower left), result image (lower right).*

4.3.2.3 Brain Segmentation

The necessity for additional brain segmentation after skull-stripping arises from the fact that the outlined method for skull-stripping tends to incorrectly classify the meninges as brain tissue, since these, too, are forming deep catchment basins lying inside the large basin that is bounded by the CSF-filled gap surrounding the brain. These structures need to be explicitly eliminated in order to restrict the subsequent tissue classification to brain tissue. In practice, this brain segmentation starts with a coarse identification of WM in the bias-field-corrected skull-stripped image by Otsu-separation [Otsu79]. This WM mask is then iteratively dilated whereby only those voxels are added, the intensity of which lies within an intensity interval that reflects the signal range of GM and WM. Moreover, the dilation is spatially restricted to eliminate all voxels that have an unreasonably large distance to WM ($d < 6mm$). Since the mask is constrained to form one spatially compact structure through 3D connected component analysis, the meninges are eliminated, as shown in figure 4.10. Finally, the remaining mask is dilated one more time, this time adding only voxels that show intensity significantly lower than GM to get CSF back into the brain mask. This heuristic proved to be robust in all clinical image data of dementia patients ($n=12$), since these datasets are typically accompanied by a broadened gap between GM and the meninges due to the beginning neurodegeneration itself. With respect to image data of young and healthy volunteers, the outlined method always improved brain segmentation, but did not yield perfect results in case of a smooth transition between GM and the meninges (see figure 4.10). However, for a prototypical implementation of WM-based spatial normalization, it was considered sufficient to

ensure that bright voxels in the meninges and in subcutaneous fatty tissue are removed from the preliminary brain mask. Voxels in the refined brain mask that anatomically correspond to the meninges and show the same signal intensity as GM were considered as uncritical, since these do not confound the subsequent segmentation of WM.

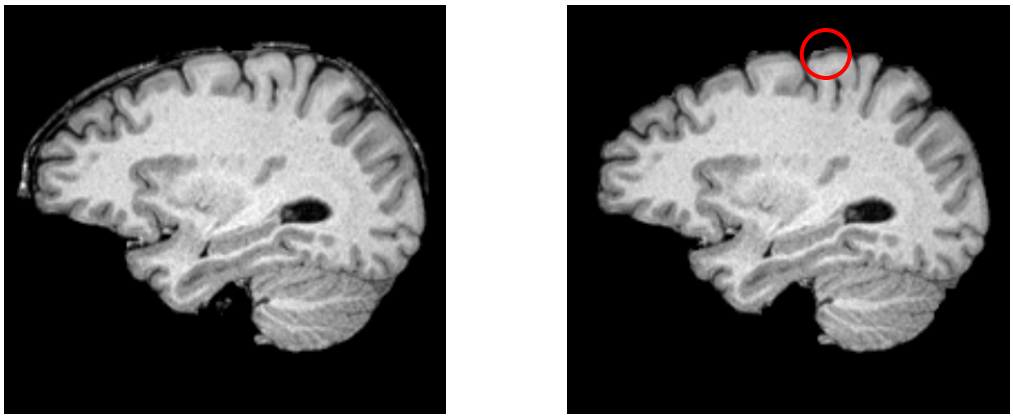


Figure 4.10: *The meninges are largely eliminated by an additional brain segmentation step: original skull-stripped image (left), segmented brain (right). In case of image data of young and healthy volunteers, residual voxels that anatomically correspond to the meninges can remain due to a smooth transition between GM and the meninges (see red circle). However, these are uncritical for WM-based spatial normalization (explanation in the text).*

4.3.2.4 Tissue Classification

As mentioned before, the segmentation procedure is completed by tissue classification according to expectation maximization. Having executed the previously described brain segmentation, the intensity histogram of the remaining image is modelled as a mixture of three Gaussian distributed probability density functions $G_i(\mu_i, \sigma_i)$, one for each WM, GM and CSF, as given in equation 4.18, where each component's portion of the total intensity distribution is encoded by p_i. Further background classes are not required since all non-brain voxels have been eliminated beforehand.

$$p_M = \sum_{i=0}^{n} p_i \cdot G_i(\mu_i, \sigma_i) \qquad (4.18)$$

The segmentation process then becomes an optimization of the parameter vector $\Theta = \{\mu_0, \sigma_0, p_0; ...; \mu_n, \sigma_n, p_n\}$, such that the model p_M best explains the image's true intensity distribution. As mentioned in section 4.2.4.2.1, the expectation-maximization algorithm maximizes the log-likelihood of Θ in an iterative manner, whereby each iteration consists of two processing steps [Dempster77]:

- In the E-step, the posteriori probability p_{ji} that given $a_j = A(\vec{x}_j)$ (the intensity of voxel $\vec{x}_j$ in image A), $\vec{x}_j$ belongs to class C_i is determined based on the current parameter vector Θ^t at iteration t, whereby $G(a_j \mid \mu_i^t, \sigma_i^t)$ is meant to be the probability of intensity a_j given its corresponding voxel $\vec{x}_j$ is a member of class C_i:

$$p_{ji}^t = \frac{p_i \cdot G(a_j \mid \mu_i^t, \sigma_i^t)}{\sum_{i=0}^{n} p_i \cdot G(a_j \mid \mu_i^t, \sigma_i^t)} \qquad (4.19)$$

- The M-step then updates the parameter vector Θ^{t+1} based upon the results of the previous E-step, meaning that it uses this data as if it was measured data:

$$\mu_i^{t+1} = \frac{\sum_{j=\min A(\vec{x})}^{\max A(\vec{x})} p_{ji}^t \cdot a_j}{\sum_{j=\min A(\vec{x})}^{\max A(\vec{x})} p_{ji}^t} \tag{4.20}$$

$$\sigma_i^{t+1} = \sqrt{\frac{\sum_{j=\min A(\vec{x})}^{\max A(\vec{x})} p_{ji}^t \cdot (a_j - \mu_i^{t+1})}{\sum_{j=\min A(\vec{x})}^{\max A(\vec{x})} p_{ji}^t}} \tag{4.21}$$

$$p_i^{t+1} = \frac{\sum_{j=\min A(\vec{x})}^{\max A(\vec{x})} p_{ji}^t}{\max A(\vec{x}) - \min A(\vec{x}) + 1} \tag{4.22}$$

Expectation maximization requires reasonable starting values for the parameter vector that reflects the classes' distributions. These starting estimates, again, are computed by Otsu-separation [Otsu79]: It is searched for those two thresholds that separate WM, GM and CSF from each other such that their overall inter-class variance *V* is maximized.

$$V = \sum_{i=0}^{n} p_i \cdot (\mu_i - \mu)^2 \tag{4.23}$$

The resulting parameter vector is then used as a coarse initial guess for the subsequent expectation maximization. The distributions' parameters are optimized until the algorithm either converges or a maximum number of iterations is reached (*n=100*). Finally, the resulting parameter vector is used for a soft tissue classification that encodes for each voxel the probability of it belonging to either of the candidate classes (see equation 4.24, *a* is meant to be the intensity of voxel $\vec{x}$, i.e. $A(\vec{x})$). Practically, one soft segmentation image is created for each tissue type separately, as illustrated in figure 4.11.

$$p_{WM}(a) = \frac{G(a \mid \mu_{WM}, \sigma_{WM})}{p_{WM}(a) + p_{GM}(a) + p_{CSF}(a)} \tag{4.24}$$

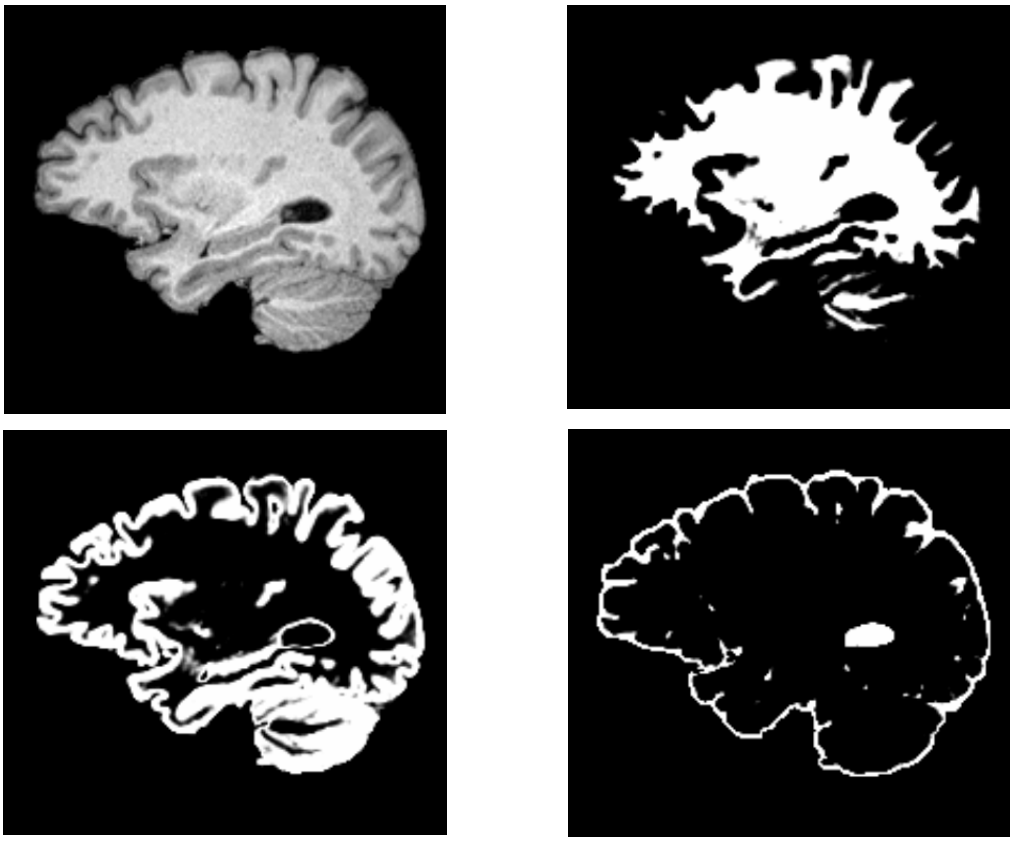


Figure 4.11: *Representative result of soft tissue classification according to the expectation-maximization algorithm: original brain image (upper left), WM image (upper right), GM image (lower left), CSF image (lower right).*

4.3.3 Registration

Having executed the outlined segmentation pipeline for both input images, the images are spatially normalized based upon WM, whereby the registration is a two-stage procedure: first the WM segmentation result images are linearly registered, and afterwards nonlinear registration is carried out to correct for acquisition-related morphological variability. As the registration is purely relying on WM, its result is neither confounded by non-brain tissue like the eyes nor by potential pathological changes in GM.

4.3.3.1 Linear Registration

Linear registration uses the complete WM result images as region of interest and – in short – iteratively minimizes the images' summed squared difference *SSD* in their common overlap region $\Omega_{A,B}^{T}$ [Holden00].

$$SSD = \tfrac{1}{N} \cdot \sum_{\vec{x}_i \in \Omega_{A,B}^{T}} \left(A(\vec{x}_i) - B^{T}(\vec{x}_i) \right)^2 \qquad (4.25)$$

The minimization strategy thereby follows the principle of gradient descent [Press93]. Starting with the transformation T_{AB}^{0} respectively parameter vector Θ^0, *SSD* is iteratively minimized by proceeding from Θ^0 in direction of the local gradient of *SSD* at Θ^0 in negative direction. The stepwidth is proportional to the local gradient and can be further adjusted by the parameter γ (see equation 4.26). This iterative procedure is repeated until either a local minimum is reached or if the gain in image

similarity between two iterations falls below a user defined threshold. 3rd order B-spline interpolation is used for computation of intensity values at non-grid locations.

$$\Theta^{i+1} = \Theta^{i} - \gamma \cdot \nabla SSD(\Theta^{i}) \qquad (4.26)$$

To increase the method's robustness, the images are linearly registered in a multi-resolution scheme [Dengler88] [Bajcsy89]: At each resolution stage *r*, the original images are first Gaussian filtered with variance $(r \cdot 0.5)^2$ and then downsampled by factor *r* in each direction, such that r^3 voxels are merged into one (see figure 4.12). Subsequently, image similarity is maximized for each resolution stage separately, starting with the result of the preceding resolution level.

Figure 4.12: *Illustration of multi-resolution image pyramid: The linear registration procedure starts at resolution level r=4 (left), is continued with downsampling factor 2 (middle) and finished at full resolution (right). For visualization purposes, the downsampled images are scaled with their downsampling factors to match the original image in size.*

Prior to this multi-resolution coarse-to-fine registration, the translational component of the parameter vector Θ^0 is initialized by the difference in the WM result images' centres of mass.

4.3.3.2 Nonlinear Registration

After linear registration, the images are in reasonable, but not in best-possible alignment because of the previously investigated nonlinear acquisition-related morphological variability in longitudinal MRI data series. By compensating for the remaining morphological differences in the spatial distribution of WM in image space and transferring this transformation to non-WM voxels, the original images can be spatially normalized without eliminating pathological changes contained in GM, provided that WM is not affected by the pathology. It was shown in section 3.4.2.2 that this condition is met in early AD.

The most important approaches in nonlinear image registration were presented and evaluated in section 4.2.3, and optical flow was selected as methodological basis for this exemplary implementation of WM-based spatial normalization of brain images.

4.3.3.2.1 Optical Flow – General Remarks

Optical flow can be described as a technique that models temporal intensity changes between two images *A* and *B* as fluid particles flowing through the image volume, whereby it assumes that no compression or expansion of structures has happened within the local aperture. The mathematical basis of optical flow is the motion constraint equation, which is derived in 3D below (following [Hata00]):

Assuming that *A(x,y,z,t)* is the intensity of a centre voxel in a small spatial neighbourhood, which moves by *dx, dy, dz* in the time interval *dt* to *x+dx, y+dy, z+dz,* one can write

$$A(x, y, z, t) = A(x + dx, y + dy, z + dz, t + dt) \quad (4.27)$$

Provided that both the local translations *dx, dy, dz* and the temporal sampling interval *dt* are small, it is possible to perform a 1st order Taylor series expansion about *A(x,y,z,t)*, given in equation 4.28, where *R* are residual higher order terms that are assumed to be small enough to be ignored.

$$A(x+dx, y+dy, z+dz, t+dt) = A(x, y, z, t) + \frac{\partial A}{\partial x}dx + \frac{\partial A}{\partial y}dy + \frac{\partial A}{\partial z}dz + \frac{\partial A}{\partial t}dt + R \quad (4.28)$$

These two equations can be combined into one, which results in

$$\frac{\partial A}{\partial x}dx + \frac{\partial A}{\partial y}dy + \frac{\partial A}{\partial z}dz + \frac{\partial A}{\partial t}dt = 0\text{, or} \quad (4.29)$$

$$\frac{\partial A}{\partial x}\frac{dx}{dt} + \frac{\partial A}{\partial y}\frac{dy}{dt} + \frac{\partial A}{\partial z}\frac{dz}{dt} + \frac{\partial A}{\partial t}\frac{dt}{dt} = 0\,. \quad (4.30)$$

Writing optical flow as $\vec{u} = \begin{pmatrix} \frac{dx}{dt} & \frac{dy}{dt} & \frac{dz}{dt} \end{pmatrix}^T$, the spatial intensity gradient as $\nabla A = \begin{pmatrix} \frac{\partial A}{\partial x} & \frac{\partial A}{\partial y} & \frac{\partial A}{\partial z} \end{pmatrix}^T$, and the temporal intensity derivative as $A_t = \frac{\partial A}{\partial t}$, results in the 3D motion constraint equation[28]:

$$\nabla A \cdot \vec{u} = -A_t \quad (4.31)$$

This equation describes a plane in 3D space, and thus, can be rewritten as:

$$\vec{u} \cdot \vec{n} = u_n\,, \quad (4.32)$$

where $\vec{n} = \frac{\nabla A}{\|\nabla A\|}$ is the surface normal of the plane, and $u_n = \frac{-A_t}{\|\nabla A\|}$ the 3D plane normal velocity magnitude [Barron94].

Taking a closer look at equations 4.31 and 4.32, it becomes evident that within a local aperture, the motion constraint equation cannot be solved unambiguously. In practice, only the flow in direction of the local intensity gradient can be recovered, but not the velocity along the plane (see figure 4.13).

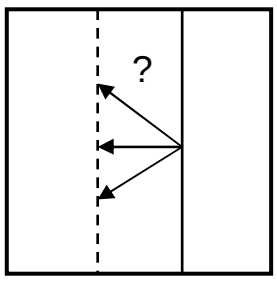


Figure 4.13: *2D-illustration of the aperture problem: the object's movement within a local aperture cannot be unambiguously determined, but only the component along the intensity gradient.*

[28] To avoid misunderstanding, it is here explicitly stated that ^{T} is to be taken as transposition operator in the context of matrices and vectors. It is not to be confused with a spatial transformation *T* as used in other parts of this chapter.

Simplifying, it can be stated that 3D optical flow describes the translation of a plane (at an imaged object's border) through image space with constant velocity in direction of its surface normal.

For nonlinear image registration, optical flow is computed locally for relatively small spatial neighbourhoods. These independently determined local translations are usually brought into spatial relation to each other by incorporating an additional smoothness constraint which is designed to ensure that no unrealistic spatial discontinuities in the overall deformation field occur for each moving object. If the analysed images contain multiple objects, which might have moved in different directions, the formulation of a suitable smoothness constraint can become very difficult, since it must eliminate discontinuities within each object, while allowing them between different objects; however, for WM-based spatial normalization, this is not the case, which is why a global smoothness constraint was considered to be sufficient for this prototypical implementation (see section 4.3.3.2.3).

4.3.3.2.2 Local Least-Squares Computation of Optical Flow

Several algorithms for computation of optical flow have been developed by the image analysis community so far [Barron94]. Methodologically, these differ mainly in the regularization constraints incorporated to yield appropriately smooth flow fields, but not in the way the initial local normal velocities are computed. The latter is usually done by means of a least-squares fit of local first-order constraints to a constant model for $\vec{u} = \begin{pmatrix} u_x & u_y & u_z \end{pmatrix}^T$ in each small spatial neighbourhood *W* by minimizing

$$\sum_{\vec{x} \in W} \left(\nabla A(\vec{x}, t) \cdot \vec{u} + A_t(\vec{x}, t) \right)^2 \mapsto \min \tag{4.33}$$

The solution to equation 4.33 is given by

$$\vec{u} = -\left(G^T \cdot G \right)^{-1} \cdot G^T \cdot C \,, \tag{4.34}$$

where *G* and *C* are explained below for a neighbourhood *W* consisting of *n+1* voxels. In order to be more robust against signal noise, the WM segmentation images are smoothed with a 3x3x3 Gaussian kernel before the spatial derivatives are computed by means of finite differences.

$$G = \begin{bmatrix} \frac{\partial A}{\partial x}(\vec{x}_0) & \frac{\partial A}{\partial y}(\vec{x}_0) & \frac{\partial A}{\partial z}(\vec{x}_0) \\ \frac{\partial A}{\partial x}(\vec{x}_1) & \frac{\partial A}{\partial y}(\vec{x}_1) & \frac{\partial A}{\partial z}(\vec{x}_1) \\ \frac{\partial A}{\partial x}(\vec{x}_n) & \frac{\partial A}{\partial y}(\vec{x}_n) & \frac{\partial A}{\partial z}(\vec{x}_n) \end{bmatrix} \tag{4.35}$$

$$C = \begin{pmatrix} \frac{\partial A}{\partial t}(\vec{x}_0) \\ \frac{\partial A}{\partial t}(\vec{x}_1) \\ \cdots \\ \frac{\partial A}{\partial t}(\vec{x}_n) \end{pmatrix} \tag{4.36}$$

Equation 4.34 reveals that the resulting translation $\vec{u} = \begin{pmatrix} u_x & u_y & u_z \end{pmatrix}^T$ is only determinable if $\left(G^T \cdot G\right)^{-1}$ is invertible. As a measure of numerical stability, typically the smallest eigenvalue of $\left(G^T \cdot G\right)$ is used [Barron94]. In order to adequately reflect the segmented WM images' SNR, this confidence threshold – in this thesis's implementation – is formulated in dependence of the CNR between WM and GM according to the result of the preceding tissue segmentation (see section 4.3.2.4); if the smallest eigenvalue is not larger than $k \cdot \frac{1}{CNR} \approx k \cdot \frac{0.5 \cdot \left(\sigma_{WM} + \sigma_{GM}\right)}{\mu_{WM} - \mu_{GM}}$, $\vec{u}$ is not accepted as reliable, but instead marked as invalid for subsequent processing steps[29]. In practice, a suitable value for *k* was determined empirically as 100.

Due to the strong accuracy requirements that WM-based spatial normalization imposes on registration, it is not assumed normal velocity in neighbourhoods, the intensity structure of which does not allow for a reliable computation of full optical flow in the outlined way, e.g. by modelling $\vec{u}$ as a linear combination of the eigenvectors corresponding to eigenvalues larger than the stability threshold [Spies02]. In practice, this implies that no local translations are computed

- in homogenous regions (where, neglecting signal noise, all partial spatial derivates are zero, $\left(G^T \cdot G\right)$ is the zero matrix, and thus, not invertible);
- in border regions that do not contain any local curvature, e.g. a plane of a cuboid (where, neglecting signal noise, the spatial gradients point in the same direction for all voxels in the aperture, such that all rows of *G* are linearly dependent from each other, *G*'s rank equals one, and the determinant of $\left(G^T \cdot G\right)$ is zero);
- if the aperture captures two of these non-curvature borders (and nothing else), e.g. an edge of a cuboid (where, neglecting signal noise, *G*'s rank is two, and thus, $\left(G^T \cdot G\right)$ is not invertible either).

For computational efficiency, it is sensible to exclude homogeneous regions from the region of interest M_{AB}^I for that optical flow is evaluated in. Moreover, it is advisable to determine M_{AB}^I as the unification of the border regions identified in each image separately in order not to favour one dataset over the other.

In this thesis's prototype, the individual borders M_A^I and M_B^I are extracted as the difference between the binarized WM segmentation images and their binary erosion

[29] The true CNR is usually approximated as the ratio of the contrast ($\mu_{WM} - \mu_{GM}$) to the noise standard deviation [Haacke99]. For stability reasons, the latter is here approximated as the average WM and GM noise standard deviation.

(see equations 4.37-4.39; a typical value for c is 50% WM probability). M_B^I, of course, is computed based upon the linearly registered WM image B_{LAB}^{WM}.

$$\hat{A}^{WM}(\vec{x}_A) = \begin{cases} 1 & A^{WM}(\vec{x}_A) \geq c \\ 0 & A^{WM}(\vec{x}_A) < c \end{cases} \tag{4.37}$$

$$M_A^I(\vec{x}_A) = \hat{A}^{WM}(\vec{x}_A) - \left(Erosion \otimes \hat{A}^{WM}\right)(\vec{x}_A) \tag{4.38}$$

$$M_{AB}^I(\vec{x}_A) = \begin{cases} 1 & M_A^I(\vec{x}_A) \vee M_B^I(\vec{x}_A) \\ 0 & else \end{cases} \tag{4.39}$$

4.3.3.2.3 Insufficient Sampling in Time

From the preceding analysis it can be concluded that optical flow-based methods determine local translations wherever it is possible with respect to an image's spatial intensity structure. However, this does not necessarily mean that every 3D normal velocity computed by optical flow is reliable even if the local neighbourhood's spatial intensity structure allows for a numerically stable solution of equation 4.34. In addition, the sampling interval *dt* has to be appropriately small.

If *dt* is too large for the "motion" captured in the input data series, the datasets contain large morphological changes, and optical flow can give incorrect results, as illustrated in figure 4.14. At first glance, this does not seem to be relevant for WM-based spatial normalization, as acquisition-related morphological variability in longitudinal MRI data is comparatively small, but this consideration is misleading, as not the absolute amount of morphological change is the parameter to look at, but rather the change in relation to the "moving" object's size. In detail, incorrect translations result if the object's movement between two time-points exceeds half the object's size, as illustrated in figure 4.14. Due to the highly convoluted shape of the cortex, WM images usually contain thin strung-out folds, which bear the risk of violating this condition even if the absolute magnitude of acquisition-related morphological variability is small.

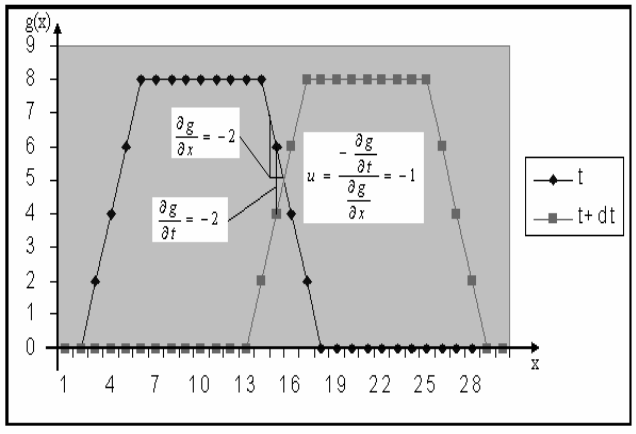


Figure 4.14: *If the sampling interval in the time domain is inappropriately large, the input data series will contain large morphological changes, and optical flow can give incorrect results.*

One possible approach known from literature to deal with insufficient sampling in the time domain is the implementation of optical flow in a multi-resolution framework [Hata00] [Periaswamy03]. However, since downsampling of the input datasets does not always preserve the thin strung-out WM folds (see figure 4.12), which are the

only structures that inappropriate sampling in time is relevant for in this thesis's context, no spatial coarse-to-fine strategy was implemented. Instead, two different arrangements are taken that do not require to reduce spatial resolution:

First of all, it is introduced an additional consistency check that accepts the computed full 3D normal velocities as reliable only if the mean spatial gradient vectors of the regions captured by the aperture *W* in A^{WM} and B^{WM}_{LAB} do not differ strongly from each other. This makes sure that optical flow is computed only for those structures that already are in reasonable alignment. In practice, the consistency check consists of a component-wise comparison of the gradient vector's signs, and thus, ensures that optical flow is evaluated only if the spatial gradients of the regions captured by the aperture in A^{WM} and B^{WM}_{LAB} are in the same octant of the image-based coordinate system. In order to exclude as few locations as possible from a direct computation of optical flow, the outlined gradient-control consistency check is combined with an iterative scheme to determine the size of the local aperture *W*: Starting with a minimum kernel size s_{min}, the aperture's size *s* is increased until it either passes the gradient-control check or a user-defined maximum kernel size s_{max} is reached[30]. Typical values for s_{min} and s_{max} are 5 and 11.

Secondly, this thesis's implementation of optical flow-based registration does not attempt to optimally align the input datasets at once. Instead, the complete nonlinear registration procedure, ranging from the initial local least-squares fit to the step of extrapolation (see next section), is iteratively repeated, whereby at each iteration the deformation is constrained to not exceed a user-defined maximum flow magnitude. By doing so, the datasets are iteratively brought into better spatial alignment, which increases the reliability of the computed local translations. Ideally, this iterative registration ultimately also brings the described thin strung-out WM folds into an alignment so well that a direct evaluation of optical flow becomes possible also for them. This strategy was implemented because it is always favourable to directly compute optical flow rather than relying on the correctness of an extrapolation. Since it is known from the analyses presented in section 3 that acquisition-related morphological variability in position does not exceed a value of 2.5 mm in an image region relevant for whole-brain morphometry, the maximum number of required iterations *iterMax* can be automatically computed as

$$iterMax = floor\left(\frac{2.5}{flowMax} + 0.5\right). \qquad (4.40)$$

4.3.3.2.4 Completion of Intermediate Deformation Fields

In the previous sections it was shown that it is not possible to reliably compute optical flow for every image location in the region of interest M^{I}_{AB}, such that the necessity for appropriately filling gaps in the intermediate forward and reverse deformation fields $T^{M}_{AB^T}$ and $T^{M}_{B^T A}$ emerges[31]. In this prototype of WM-based spatial normalization,

[30] For better readability, the notation of the kernel size is simplified in this section. A size *s* is meant to represent a kernel consisting of $s \times s \times s$ voxels.

[31] $T^{M}_{AB^T}$ is the nonlinear transformation from the current WM image B^{WM}_{T} to A^{WM} defined in the region of interest M^{I}_{AB}.

this is done in a two-step procedure, whereby the first step is obligatory and the second is an option for the user:

- Firstly, remaining gaps are filled by a component-wise weighted mean extrapolation (see below).
- Optionally, the deformation field is subsequently regularized by application of the global smoothness constraint proposed by [Horn81].

In detail, the first part is an extrapolation that assigns each voxel in M_{AB}^{I}, which normal velocity is not known for, the mean velocity $\vec{u}_m$ computed from its neighbours, whereby each translational vector $\vec{u}$ contributing to the resulting mean is weighted component-wise by the reciprocal value of its directional variance $\vec{\sigma}_u^2$, the computation of which is derived below.

The correctness of the assumption made in the initial computation of optical flow, i.e. that of a constant model for $\vec{u} = \begin{pmatrix} u_x & u_y & u_z \end{pmatrix}^T$ in a small spatial neighbourhood *W*, can be assessed by evaluation of the variance σ^2 in *W*. When again assuming *W* to consist of *n+1* voxels, σ^2 can be written as:

$$\sigma^2 = \frac{1}{(n+1)-3} \cdot R \,, \tag{4.41}$$

where *R* is the residual error term:

$$R = \sum_{\vec{x} \in W} \left(\nabla A(\vec{x}, t) \cdot \vec{u} + A_t(\vec{x}, t) \right)^2 . \tag{4.42}$$

The components of the directional variance $\vec{\sigma}_u^2$ correspond to the diagonal elements of the covariance matrix *Q*:

$$Q = \sigma^2 \cdot \left(G^T \cdot G \right)^{-1} , \tag{4.43}$$

In order to favour those components of a local translation for which the variance according to the outlined error-analysis is small, each translation $\vec{u}$ contributing to the mean is weighted with

$$\vec{w} = \begin{pmatrix} w_x = 1/\left(\sigma_{ux}^2 + \varepsilon\right) \\ w_y = 1/\left(\sigma_{uy}^2 + \varepsilon\right) \\ w_z = 1/\left(\sigma_{uz}^2 + \varepsilon\right) \end{pmatrix} \tag{4.44}$$

Experiments showed that it is even advisable to refine the initially computed normal velocities by the described weighted mean computation, if the kernel size *s* defining the neighbourhood used for mean computation is chosen appropriately small. A typical value for *s* is 5.

Having filled all remaining gaps in the intermediate forward and reverse deformation fields, the user has the option to additionally smooth $T_{AB^T}^{M}$ and $T_{B^T A}^{M}$ according to the extension of the regularization proposed by Horn and Schunck to 3D [Horn81]:

The latter consists of the incorporation of a global smoothness term *S* into the computation of a flow field *T* by minimizing

$$\sum_{\vec{x}} \left(\nabla A(\vec{x},t)\cdot T(\vec{x}) + A_t(\vec{x},t)\right)^2 + \alpha^2 \cdot S \mapsto \min \ , \qquad (4.45)$$

where α reflects the relative influence of the smoothness term, which – in practice – is the squared sum of the first order spatial derivatives of T

$$S = \left(\frac{\partial T_x}{\partial x}(\vec{x})\right)^2 + \left(\frac{\partial T_x}{\partial y}(\vec{x})\right)^2 + \left(\frac{\partial T_x}{\partial z}(\vec{x})\right)^2 + \left(\frac{\partial T_y}{\partial x}(\vec{x})\right)^2 + \ldots + \left(\frac{\partial T_z}{\partial z}(\vec{x})\right)^2 . \qquad (4.46)$$

Equation 4.45 may be solved by the calculus of variations [Horn81]. The Euler-Lagrange equations yield:

$$\Delta T(\vec{x}) = \frac{\nabla A(\vec{x},t)\cdot T(\vec{x}) + A_t(\vec{x},t)}{\alpha^2} \cdot \nabla A(\vec{x},t) , \qquad (4.47)$$

where $\Delta A(\vec{x},t)$ denotes the Laplacian operator, which - following Horn and Schunk - is approximated as the difference between the local translation $\vec{u} = T(\vec{x})$ and the Gaussian weighted mean $\hat{\vec{u}} = \hat{T}(\vec{x})$ in its direct neighbourhood:

$$\Delta T(\vec{x}) \approx \hat{T}(\vec{x}) - T(\vec{x}) \qquad (4.48)$$

Combining these equations and isolating $T(\vec{x})$ leads to:

$$T(\vec{x}) = \hat{T}(\vec{x}) - \frac{\nabla A(\vec{x},t)\cdot \hat{T}(\vec{x}) + A_t(\vec{x},t)}{\left|\nabla A(\vec{x},t)\right|^2 + \alpha^2} \cdot \nabla A(\vec{x},t) \qquad (4.49)$$

The iterative solution of this system of linear equations is given by the Gauss Seidel method:

$$T(\vec{x})^{i+1} = \hat{T}(\vec{x})^i - \frac{\nabla A(\vec{x},t)\cdot \hat{T}(\vec{x})^i + A_t(\vec{x},t)}{\left|\nabla A(\vec{x},t)\right|^2 + \alpha^2} \cdot \nabla A(\vec{x},t) , \qquad (4.50)$$

By taking a closer look at equation 4.50 it becomes evident that this iterative procedure replaces the current velocity by its local mean in image regions, which the spatial intensity gradient is small for; in regions with a strong gradient, however, the local mean is modified according to the optical flow term $\nabla A(\vec{x},t)\cdot \hat{T}(\vec{x})^i + A_t(\vec{x},t)$ that has not already been considered yet. Both the weighting factor α, as well as the number of iterations are user-defined parameters; typical values are 1 and 100, respectively.

4.3.3.2.5 Consistency

So far, registration has been solely formulated as the determination of that transformation T_{AB} from image *B* to image *A* which brings the images into best possible anatomical alignment[32]. The direction of this mapping, however, is rather arbitrary and could just as well be the other way around. In other words, the registration of *A* to *B* ideally should result in the transformation $T_{BA} = T_{AB}^{-1}$ (and vice versa).

[32] Because of this statement's generality, the problem of consistency between forward and backward registration is described by means of a general notation, i.e. the input images are referred to as *A* and *B* rather than as A^{WM} and B^{WM}_{LAB} again.

Two approaches for enforcing this consistency are known from literature:

- Thirion developed a heuristic approach that iteratively estimates the forward transformation T_{AB}, the reverse transformation T_{BA} and the residual transformation $T_R = T_{AB} \circ T_{BA}$. At each iteration, half of the residual T_R is added to T_{AB} and the other half is mapped through T_{AB} and added to T_{BA} [Thirion98].
- Christensen and Johnson do not only approximate the correspondences between the forward and reverse transformations, but explicitly compute them by introducing a special consistency term into their optimization functional [Christensen01].

Compared to Thirion's approach, the latter method is more appropriate in presence of large morphological changes between *A* and *B*. Under such circumstances, it is imaginable that the transferal of the estimated residual transformation $T_R = T_{AB} \circ T_{BA}$ partly corrupts the individual transformations and further complicates the registration procedure. If morphological changes between *A* and *B* are small, however, both techniques can be considered equally well suited for ensuring consistency in elimination of acquisition-related morphological variability. Due to its implementation being rather straight forward, Thirion's approach was chosen for this thesis's prototypical implementation of WM-based elimination of acquisition-related morphological variability.

4.3.3.2.6 Schematic Overview

Figure 4.15 provides an overview of the sequence of processing steps for one iteration[33]. B_T^{WM} thereby represents the linearly registered WM segmentation image B_{LAB}^{WM} after application of the current overall nonlinear transformation T_{AB}^M resulting from the directly preceding iteration. The spatial and temporal gradients required for the evaluation of optical flow as outlined in section 4.3.3.2.2 are determined for this intermediate dataset as well as for the original WM segmentation image A^{WM}. B_T^{WM} is always computed based upon B_{LAB}^{WM} to avoid unnecessary loss in spatial information resulting from repeated interpolation at the end of each iteration.

Focussing solely on WM border regions (M_{AB}^I) for this evaluation is motivated by the aperture problem outlined in section 4.3.3.2.1. Consistency between the forward and backward transformation T_{AB}^M respectively T_{BA}^M is guaranteed by enforcing consistency between the intermediate transformations $T_{AB^T}^M$ and $T_{B^TA}^M$ by means of the strategy outlined in section 4.3.3.2.5. The processing step here referred to as "FillGaps" represents the extrapolation and smoothing mechanisms described in section 4.3.3.2.4.

[33] Implementing optical flow-based registration in an iterative scheme was motivated in section 4.3.3.2.3.

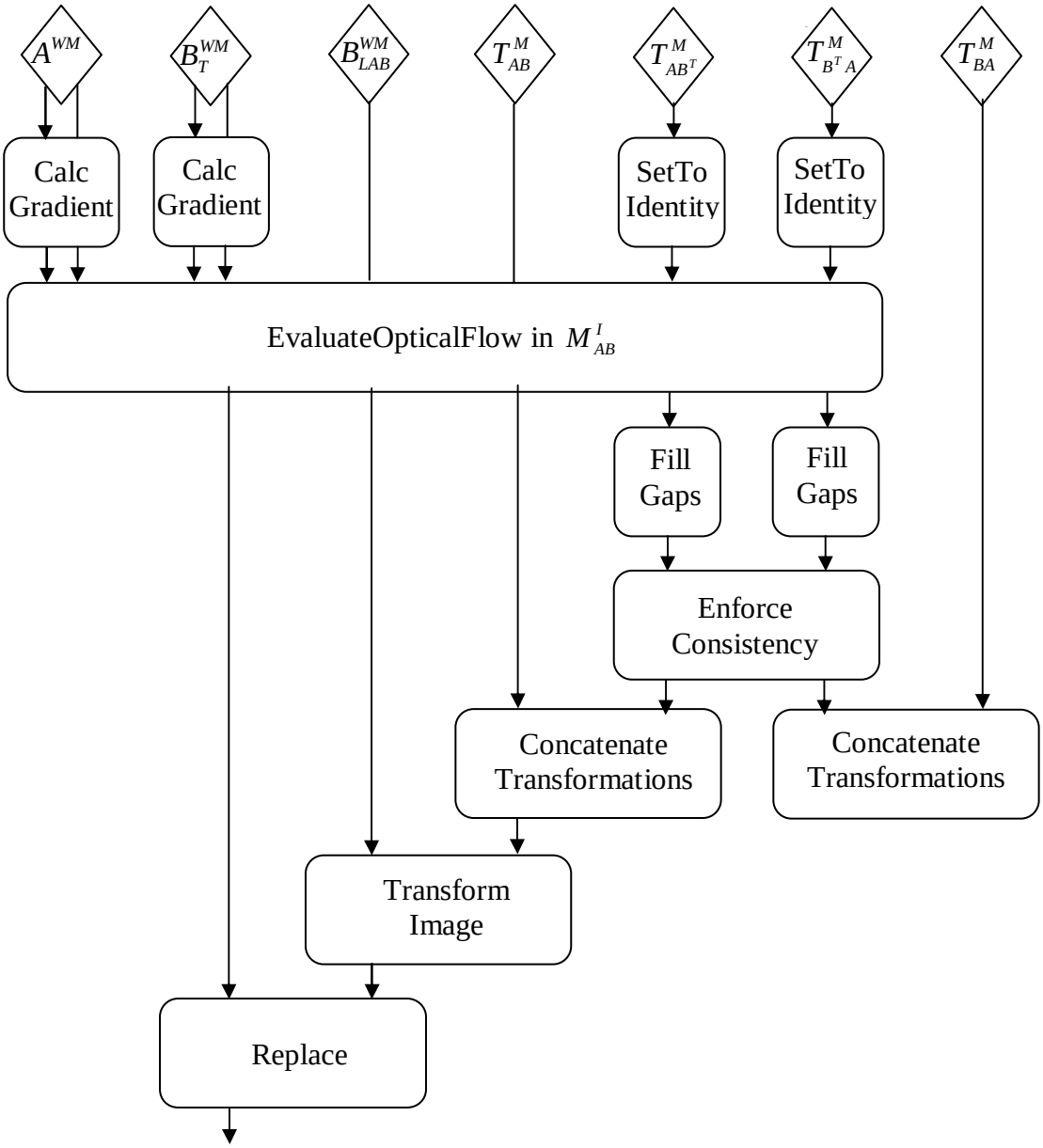


Figure 4.15: *Schematic illustration of the sequence of processing steps for one iteration (explanation in the text).*

4.3.3.2.7 Propagation to Non-WM Voxels

The overall nonlinear transformation T_{AB}^{M} determined by the iterative procedure presented in the preceding sections - ideally - brings the WM representations of the datasets *A* and *B* into perfect spatial alignment, and thus, compensates for all morphological variability introduced into the images by imperfections in the data acquisition process. In order for it to be useful for early detection of neurodegenerative diseases, however, the transformation determined for WM needs to be propagated to cortical as well as subcortical GM regions. As these are spatially adjacent to WM, it does not seem to be necessary to carry out elaborate extrapolation strategies to yield T_{AB}^{I}, but it is rather sufficient to purely transfer the WM transformation to GM.

The propagation therefore simply consists of an iterative dilation of the WM border regions M^{I}_{AB}, whereby each non WM-voxel, which the translation is not known for, is assigned the mean translation computed from its neighbours. This iterative mean propagation is continued until all voxels of the previously determined brain mask (see section 4.3.2.3) are assigned a translational vector.

After this processing step, the resulting transformation T^{I}_{AB} is defined for the whole brain. It can now be used to spatially normalize dataset *B* such that – ideally – all acquisition-related morphological differences to *A* are eliminated without compromising changes in GM corresponding to a true anatomical change (see figure 4.16).

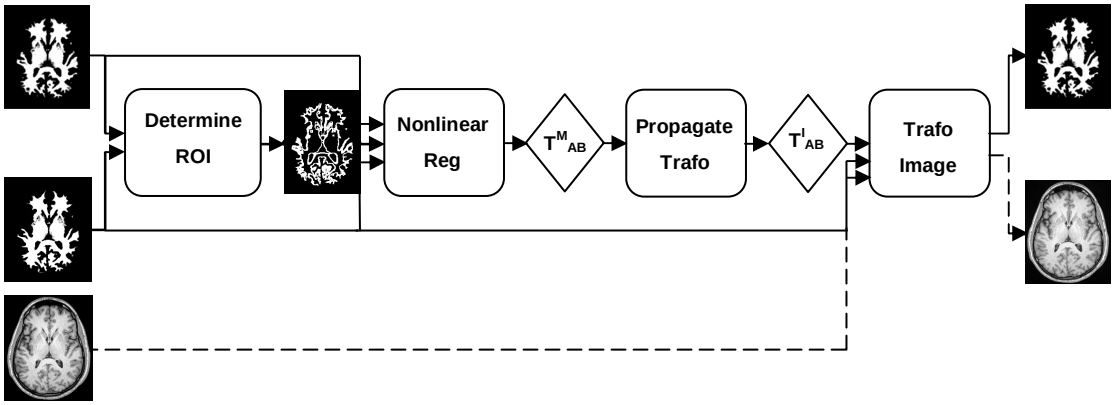


Figure 4.16: *Schematic illustration of the data flow in nonlinear registration: As a result of the aperture-problem, nonlinear alignment is based upon the intensity structure in the border region of WM and GM. Having detected morphological changes in this region of interest, the respective transformation T^{M}_{AB} is propagated to allow for a nonlinear registration of the complete image volume. By concatenating the previously determined linear transformation T_{LAB} with T^{I}_{AB}, the original dataset B is nonlinearly spatially normalized based upon WM only.*

4.3.4 Discussion

In the preceding sections, the mathematical and conceptual aspects of the processing pipeline abstractly outlined in the previous subchapter were discussed in detail: Concretely, the technical solution for each partial processing step as realized in this thesis's prototypical implementation of WM-based elimination of acquisition-related morphological variability from longitudinal MRI brain data series was presented. This discussion comprised the processing pipeline required for tissue classification by means of the EM-GMM method, but was mainly focused upon optical flow-based nonlinear registration.

In the analysis of the mathematical foundation of optical flow, it was put special emphasis on the question of the locally computed normal velocity's reliability: By introduction of a CNR-dependent confidence threshold, it is made sure that only numerically stable solutions to the local least-squares fit are used for subsequent processing steps [Barron94]. With respect to potentially insufficient sampling in time, a coarse-to-fine multi-resolution approach was considered suboptimal due to fact that spatial downsampling cannot guarantee to preserve the thin strung-out WM-folds, which are the only structures that insufficient sampling in time is relevant for in the context of acquisition-related morphological variability. Therefore, alternative strategies were implemented that work on the datasets' original resolution. In detail, these are

- the presented “gradient control” consistency check, which ensures that optical flow is evaluated only for regions that already are in reasonable spatial alignment;
- the implementation of nonlinear registration as an iterative process, whereby the deformation field is constrained at each iteration to not exceed a user-defined maximum flow magnitude; by doing so, the datasets are iteratively brought into better spatial alignment, which increases the reliability of the computed flow vectors.

The component-wise weighted local mean extrapolation of sufficiently reliable flow vectors to neighbouring voxels, which the initial least-squares fit did not give reliable results for, is based upon an error analysis that systematically assesses the correctness of assuming a constant model for the unknown translation within the local aperture. Generally spoken, this extrapolation favours those components of a flow vector contributing to the mean computation for which the variance according to this error analysis is small. The outlined weighting, thus, also serves to increase the reliability of the computed overall deformation field. Optionally, the latter can additionally be regularized by incorporation of a global smoothness constraint following the method proposed by [Horn81].

Finally, the overall deformation field’s reliability is supported by ensuring that the forward and backward transformation are consistent with each other in terms of each of them being the inverse of one another [Thirion98].

To ensure future extensibility, the prototype was implemented as a processing pipeline consisting of easily replaceable software modules. By this the flexibility is provided to not only extend existing modules, but also to easily explore alternative methods for e.g. intensity non-uniformity correction, nonlinear registration or deformation field extrapolation.

4.4 Experimental Results

4.4.1 Introduction

Based upon a detailed analysis of the characteristics of both the MRI data acquisition process and the physiological processes that Alzheimer’s disease is accompanied by, a novel concept for eliminating acquisition-related morphological variability in image space from intra-subject MRI brain data series of AD patients was proposed that does not require additional phantom imaging, but rather can be described as WM-based nonlinear spatial normalization (see chapter 3). In this chapter, so far the mathematical and conceptual aspects of a prototypical realization of this general idea were discussed.

For the remainder of this chapter, the experimental results obtained with this software prototype are presented. It starts with a discussion of validation experiments carried out with simple software phantoms with the intention to (a) find suitable parameter sets for more complex test cases, and (b) to find out whether the implemented prototype is, in principle, capable of dealing with a CNR between WM and GM that is achievable in state-of-the-art clinical MRI-based morphometry today. The experiences gained from these tests were subsequently used in more realistic validation experiments in which different patterns of GM atrophy are artificially introduced into a human volunteer MRI data series in order to allow for a profound discussion of the method’s general suitability for eliminating acquisition-related

morphological variability from intra-subject MRI data series; finally, results obtained by applying the prototype to clinical data of potential early AD patients are discussed in order to both understand the method's clinical potential, but also its limitations in its current prototypical implementation. The latter also serves as the foundation for suggestions for future work.

Since the overall sequence of processing steps involved in the implemented software prototype has already been illustrated in section 4.3 by means of intermediate result images for one representative test case, not all intermediate results are given in this section again. However, the images shown allow for a full understanding of the functioning of the prototype.

4.4.2 Synthetic Software Phantoms

4.4.2.1 Model Description – Volumetric Primitives

In the software phantoms used for basic validation, acquisition-related morphological variability in longitudinal intra-subject image data of early dementia patients is modelled in a very simplified way by means of volumetric primitives. These are arranged in image space to reflect the phenomena expected in clinical data series, only in a much simpler form:

- The datasets contain three distinct compartments, representing WM, GM and CSF
- SNR and CNR are identical for baseline and follow-up data, but the signal noise is different for each dataset.
- There are small morphological changes in the WM compartments between baseline and follow-up data reflecting acquisition-related morphological variability.
- In addition to morphological differences present in WM, the GM compartments are regionally reduced in volume to reflect atrophy, whereby the magnitude of atrophy varies from region to region.

Cylinders, ellipsoids and boxes were used as volumetric primitives, different spatial arrangements were made, and it was experimented with different noise levels with the intention to identify the minimum required WM/GM-CNR which is necessary in order for the implemented prototype to be able to give satisfying results.

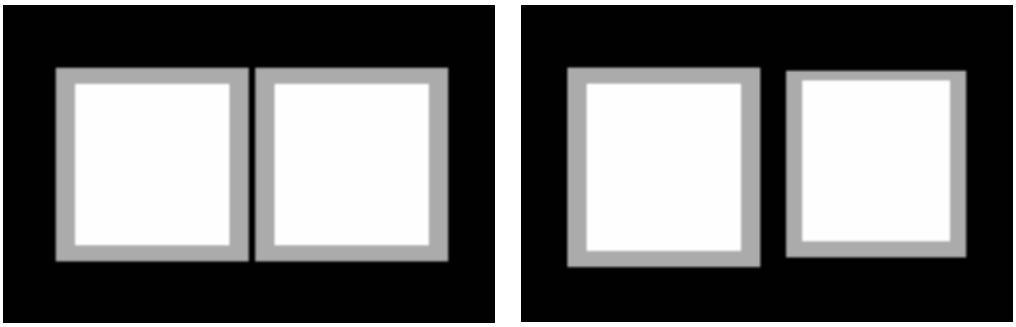


Figure 4.17: *Illustration of one representative of the type of software phantom used for basic validation. Both acquisition-related morphological variability as well as regionally varying atrophy between baseline (left) and follow-up data (right) are modeled by means of volumetric primitives, here boxes. Further details are given in the text.*

A typical example of such a phantom is given in figure 4.17. The bright inner boxes representing WM are positioned differently in the baseline (left) and the follow-up dataset (right), and they are slightly varied in size: Whereas the left box's centre was

shifted down-left, the right box's centre was moved up-right. In addition, the left box's size was increased in y-, and the right box's size decreased in x-direction. The outer "GM-boxes" follow this simplified "acquisition-related" morphological variability, and the right "GM-box" is additionally asymmetrically shrunk, which is meant to represent regionally varying atrophy. No atrophy is introduced into the left "GM-box".

Whereas the mean intensities of WM, GM, and CSF were kept constant throughout these experiments ($\mu_{WM} = 400$, $\mu_{GM} = 300$, $\mu_{CSF} = 100$), different levels of signal noise were added in order to simulate different CNRs (zero-mean Gaussian noise, $20 \leq \sigma \leq 50$)[34]. Figure 4.18 shows example pictures of the baseline dataset of the same software phantom as in figure 4.17 with different CNRs.[35]

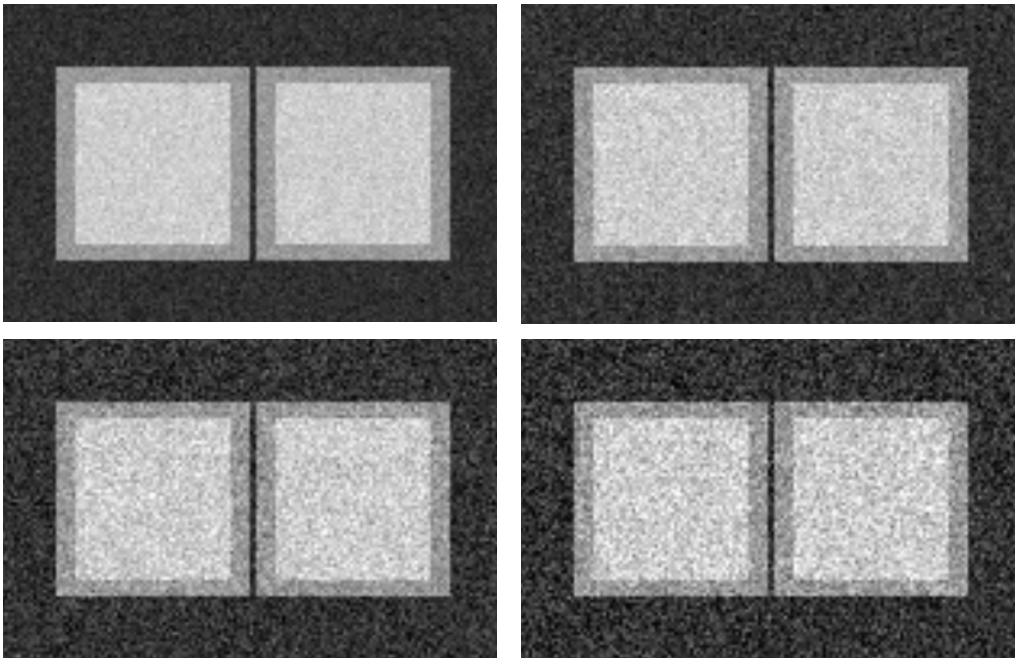


Figure 4.18: *Baseline dataset of the same software phantom as in figure 4.17, but with different WM/GM-CNRs (upper left: 5, upper right: 3 1/3 , lower left: 2.5, lower right: 2).*

4.4.2.2 Validation Method

In the software phantoms described, acquisition-related morphological variability as well as the regional signal loss reflecting GM atrophy are exactly known: When not introducing any "acquisition-related" morphological variability in WM, but keeping the GM-changes, the images shown in figure 4.19 emerge. Abstracting from signal noise, the presented prototype ideally should produce the same result.[36]

---

[34] These values were chosen to adequately reflect the contrast characteristics in the datasets acquired with the MPRAGE protocols given in section 3.3.4 for the 1.5T Siemens Magnetom Avanto. The tissue distributions were estimated by means of the tissue segmentation procedure described in section 4.3.2.4. The signal noise standard deviation of 20 thereby represents the most homogeneous tissue (WM); the noise standard deviation for GM and CSF usually was between 30 and 35. The mean intensities were determined as about 380, 290, and 90 respectively.

[35] Following [Haacke99], the true CNR is approximated as the as the ratio of the contrast ($\mu_{WM} - \mu_{GM}$) to the noise standard deviation

[36] Due to the difference of these software phantoms to MRI head datasets, the proposed method was reduced to its essence in these validation experiments, i.e. no skull-stripping, bias-field correction or linear registration was applied, but only tissue segmentation and WM-based nonlinear registration.

Figure 4.19: *When not introducing any "acquisition-related" morphological variability in WM into the software phantoms presented in figure 4.17, but keeping the GM-differences, the left image results. On the right, the difference image between the baseline dataset and this "ideal" follow-up dataset is shown.*

The evaluation of the accuracy of the experimentally spatially normalized datasets was done by quantifying the mean intensity differences between these normalized datasets and their corresponding baseline dataset. In order to be able to also detect potential "false positives", the changes were quantified both in the region affected by "GM-atrophy" and in the region only affected by "acquisition-related morphological variability". Ideally, the mean intensity of the former region would be identical to the intensity difference between GM and CSF, i.e. -200, whereas the mean difference in the latter region would be zero.

4.4.2.3 Results

A suitable parameter set was identified empirically based on the presented box-phantom data with a noise level of $\sigma = 20$ (WM/GM-CNR = 5). These parameters were subsequently applied to all other experiments presented in this section. In detail, the following values were chosen:

Parameter Description	Value
maximum number of iterations in the EM-algorithm	100
intensity threshold for segmenting WM-borders	50 % WM probability
size of Gaussian kernel for smoothing the WM segmentation images in order to stabilize the computation of spatial derivatives by finite differences	3 x 3 x 3 voxels
maximum allowed flow velocity at each iteration number of nonlinear registration iterations	0.5 mm 6
minimum aperture size maximum aperture size	5 x 5 x 5 voxels 11 x 11 x 11 voxels
confidence threshold for matrix inversion	$100 \cdot \frac{0.5 \cdot (\sigma_{WM} + \sigma_{GM})}{\mu_{WM} - \mu_{GM}}$
weighting factor of global smoothness constraint number of iterations for solving equation 4.45	1 100

Table 4.1: *Empirically determined parameter set for validation experiments.*

The progress of WM-based spatial normalization by means of the implemented prototype to the exemplarily presented dataset with the given parameters is presented in figure 4.20: The follow-up dataset's WM representation is iteratively aligned to the WM in the baseline dataset until practically all "acquisition-related" morphological variability contained in WM is eliminated.

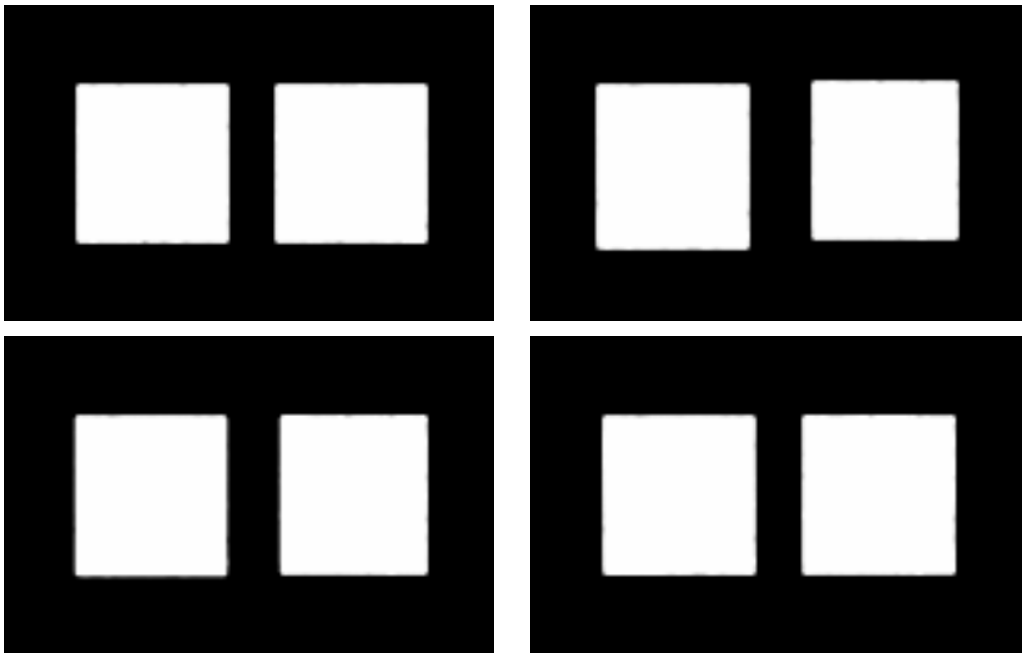


Figure 4.20: *Iterative progress of WM-based spatial normalization (upper left: baseline WM, upper right: follow-up WM, lower left: follow-up WM after iteration 3, lower right: follow-up WM after iteration 6).*

The WM/GM-CNR was expected to be an important influencing factor on the accuracy of the obtained results. However, taking a look at the difference images between the baseline image and the corresponding spatially normalized follow-up image of the previously described box-shaped software phantom (see figure 4.21), it becomes obvious that the prototype can cope quite well even with noise levels that are significantly higher than it can be expected in state-of-the-art clinical MRI data.

This visual impression is supported by the results of the quantitative evaluation given in table 4.2.[37] For a WM/GM-CNR higher than 2½, the implemented prototype almost perfectly eliminates "acquisition-related" morphological variability, while not compromising "disease-related" differences (GM atrophy regions) or incorrectly introducing additional differences. At a CNR lower than 2½, however, the accuracy of the results cannot be considered sufficient any more.

[37] Table 4.2 reports the results of the validation experiments with the presented box-phantoms. These results are representative for all the synthetic software phantoms experimented with, no matter which magnitude of "acquisition-related" morphological variability was introduced into the datasets and no matter whether boxes, ellipsoids, cylinders or combinations of these volumetric primitives were used.

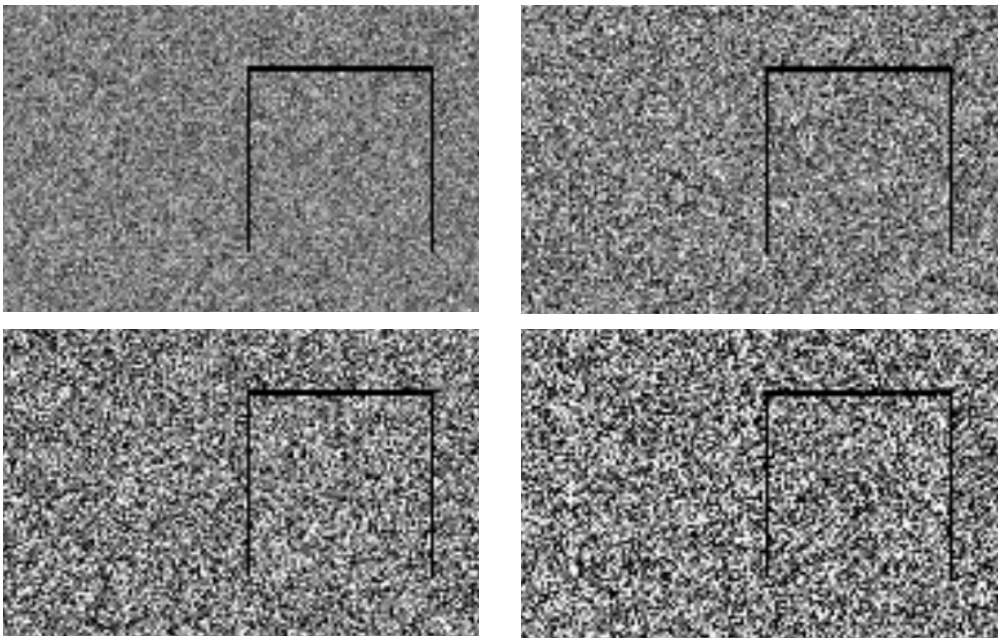


Figure 4.21: *Difference images between baseline dataset and the corresponding spatially normalized follow-up dataset for different WM/GM-CNRs (upper left: 5, upper right: 3⅓, lower left: 2½, lower left: 2). The images reflect WM-based spatial normalization without application of the global smoothness constraint according to [Horn81].*

		CNR = 5		CNR = 3 ⅓		CNR = 2 ½		CNR = 2	
		weighted mean	Horn	weighted mean	Horn	weighted mean	Horn	weighted mean	Horn
GM-atrophy	Mean	-199,84	-200,64	-199,19	-198,02	-199,03	-193,10	-194,82	-182,49
	StdDev	28,42	28,42	42,38	46,08	55,62	63,29	69,29	82,75
acquisition-related changes	Mean	-0,46	-0,68	-0,54	-0,74	-0,43	-0,78	-0,38	-0,86
	StdDev	28,33	28,51	42,31	42,68	55,74	56,17	67,98	68,49

Table 4.2: *Quantitative results of applying the proposed prototype for WM-based spatial normalization to the box-shaped software phantom presented earlier. The normalization's accuracy is strongly dependent on the WM/GM-CNR in the input images, which has to be at least 2½ to allow for a reliable elimination of acquisition-related morphological variability. The additional incorporation of the global smoothness constraint as proposed by [Horn81] reduces the accuracy especially at higher noise levels. Analogous results were obtained with other variants of the synthetic software phantoms the generation of which was described in section 4.4.2.1.*

Table 4.2 also shows that the additional incorporation of the global smoothness constraint as proposed by [Horn81] did not improve the results; for a poor CNR, it even deteriorated them. However, this is not to be misunderstood as a sign for the smoothness constraint's general inappropriateness, but rather as a consequence of a suboptimal implementation in terms of an uncritical use of the local intensity gradient $\nabla A(\vec{x},t)$ in the iterative Gauss Seidel equation 4.50: In contrast to the local least-squares computation of optical flow presented in section 4.3.3.2.2, the re-computation of optical flow according to equation 4.50 neither incorporates a systematic error analysis, nor a term that takes into consideration how reliable the locally determined spatial gradient is with respect to the images' noise level. Evidently, the lower the WM/GM-CNR is in the input datasets, the less reliable is a locally computed spatial

gradient in the corresponding WM segmentation images, and thus, its uncritical use can corrupt the flow field.

4.4.3 Hybrid Phantoms

4.4.3.1 Model Description - Artificial Atrophy

In this section, the previously presented validation experiments with synthetic software phantoms are extended by a discussion of experiments with more realistic input data, namely longitudinal MRI data series of a healthy human volunteer, into which different patterns of GM atrophy were artificially introduced. However, since neither the factors initializing AD, nor the molecular processes that control the progression of the disease are clear yet (see section 2.1), it is not possible to truly simulate AD. The best thing that can be done to enable controlled validation experiments is to reasonably reproduce the target phenomenon, i.e. GM atrophy.

In this thesis, this was achieved by a combination of standard image processing operations: Atrophy is artificially introduced into arbitrary MRI brain datasets by application of a specifically parameterized kernel-based grey-value erosion that in short can be described as follows: The intensity $a_i = A(\vec{x}_i)$ of a voxel $\vec{x}_i$ falling inside the grey-value interval representing the transition between CSF and GM is decreased proportionally to its difference in intensity to the minimum intensity in its direct spatial neighbourhood N (see equations 4.51 and 4.52). The proportionality factor $k(\vec{x}_i)$ is chosen to be directly dependent on the voxel's Euclidean distance to the spatially closest voxel that is either fully occupied by GM or partially filled with GM and WM ($k(\vec{x}_i) = k' \cdot minDist(\vec{x}_i)$). To avoid unrealistically strong intensity gradients between CSF and GM, the transition between these tissues in image space is additionally smoothed by means of an isotropic Gaussian kernel (see figure 4.22).

$$N(\vec{x}_i) = \left\{ \vec{x} \quad \middle| \quad \left\| \vec{x} - \vec{x}_i \right\| < r \right\} \tag{4.51}$$

$$A^{AT}(\vec{x}_i) = \begin{cases} a_i - k(\vec{x}_i) \cdot (a_i - \min(A(N(\vec{x}_i))) & \mu_{CSF} < a_i < t \cdot (\mu_{GM} - \mu_{CSF}) + \mu_{CSF} \\ a_i & else \end{cases} \tag{4.52}$$

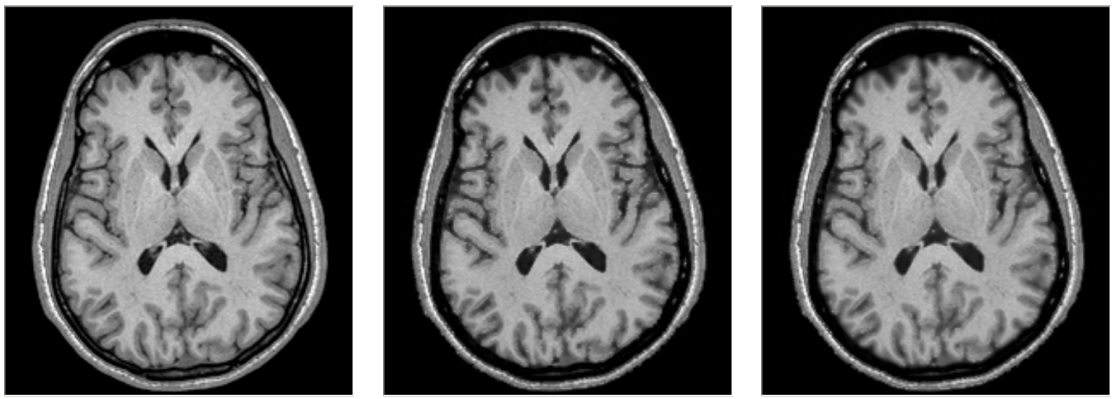


Figure 4.22: *Intermediate results of the presented method for artificially introducing GM atrophy into a MRI brain dataset of a healthy volunteer (k'=0.5, t=0.8; left: original image, middle: image after the described grey-value erosion; right: image after additional Gaussian smoothing of image regions reflecting the transition between GM and CSF).*

The proportionality factor k' and the intensity threshold t determine the amount of atrophy that is artificially introduced into the dataset. Different patterns of atrophy were generated by choosing these parameters differently for distinct image regions: For example, whole brain atrophy is obtained by spatially restricting this operation to voxels representing brain tissue (see figure 4.23). For regional atrophy, the operation is restricted to a user-defined region of interest (see figure 4.24).

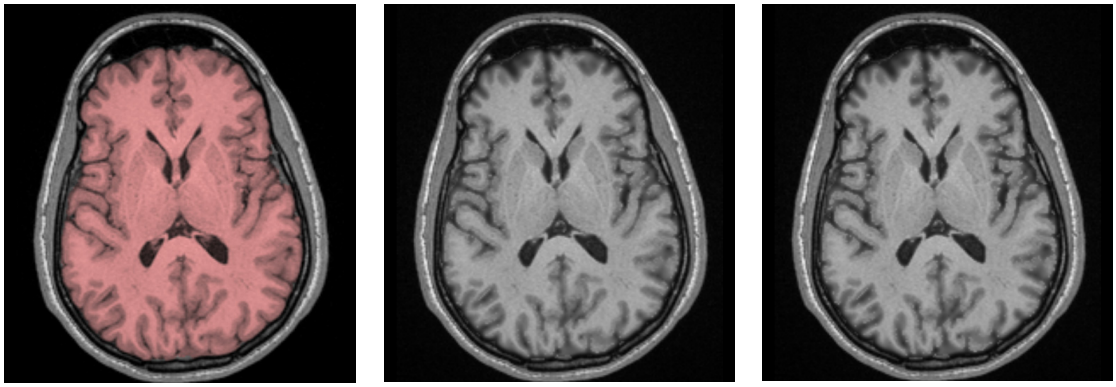


Figure 4.23: *Sample pictures of datasets into which whole-brain atrophy was artificially introduced by means of the operation outlined in the text (left: original MRI brain image of a healthy volunteer, middle: subtle GM atrophy (k'=0.5; t=0.8); right: slightly more pronounced GM atrophy (k'=0.8; t=0.8)). The operation is restricted to voxels representing brain tissue (which are segmented as described in section 4.3.2.3, and here, semi-transparently overlaid over the original image). In order to reasonably reproduce beginning neurodegeneration, the operation is parameterized to introduce subtle morphological changes.*

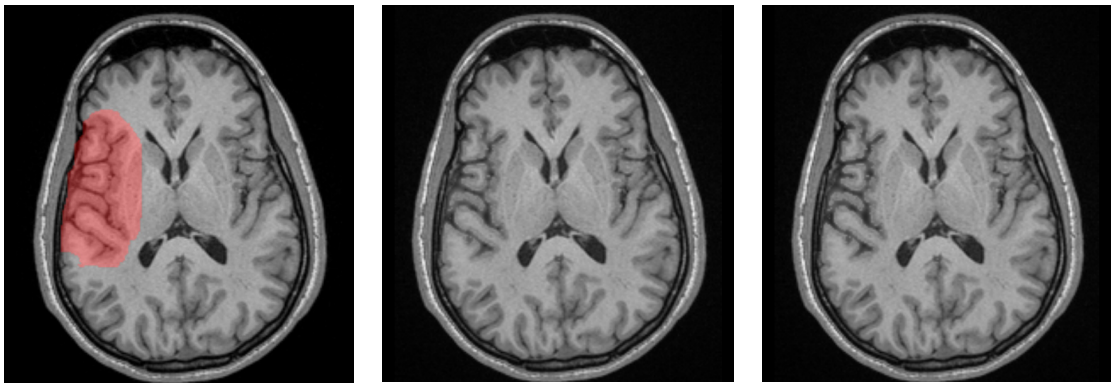


Figure 4.24: *Sample pictures of datasets into which regional atrophy was artificially introduced by means of the operation outlined in the text (left: original MRI brain image of a healthy volunteer, middle: subtle GM atrophy (k'=0.5; t=0.8); right: slightly more pronounced GM atrophy (k'=0.8; t=0.8)). The operation is restricted to a user-defined ROI, which is here semi-transparently overlaid over the original image.*

4.4.3.2 Validation Scenario

The presented operation for introducing artificial atrophy into an MRI brain dataset was used for validating the proposed method for WM-based elimination of acquisition-related morphological variability from longitudinal intra-subject MRI datasets in the following way:

- Several datasets of a healthy volunteer's head were acquired at different positions in the scanner. In order to guarantee that a noticeable amount of acquisition-related morphological variability was contained in the data, the volunteer was translated by several centimetres inside the MRI scanner, and he was advised to additionally rotate his head by a few degrees between the scans. No gradient nonlinearity correction was applied such that acquisition-related morphological variability can be assumed to be larger than in clinical

datasets that are acquired following a standardized patient repositioning protocol.

- Artificial atrophy was introduced into all but one dataset, which was subsequently considered to be the baseline dataset.
- The proposed prototype for WM-based spatial normalization was applied to this data series using the parameter set given in table 4.1.

Besides a detailed visual analysis of the results, the prototype's effectiveness was also assessed quantitatively by comparing the amount of originally introduced GM-atrophy to the atrophy present after WM-based spatial normalization. Atrophy was thereby measured as the difference between brain voxels' GM portions, which again were estimated using the probabilistic tissue classification method outlined in section 4.3.2.4. In order to make sure that these results are not confounded by signal differences in vascular structures (see figure 4.29), only those voxels were included in the quantification of atrophy the intensity of which falls inside the interval corresponding to the transition between CSF and GM ($\mu_{CSF} \leq a \leq \mu_{GM}$). Following this strategy, GM-atrophy was quantified for:

- the follow-up dataset which GM atrophy was artificially introduced into in relation to the original follow-up dataset (originally introduced atrophy AT_{orig})
- the spatially normalized follow-up dataset, into which GM atrophy was artificially introduced, in relation to the baseline dataset (observed atrophy $AT_{observed}$)
- the two times 3rd order B-spline resampled follow-up dataset[38], into which GM atrophy was artificially introduced, in relation to the original follow-up dataset (ideally observable atrophy $AT_{observable}$)

The quantification of $AT_{observable}$ was done in order to take into consideration the impact of resampling involved in the normalization procedure: Since both the intermediate linear registration step, as well as the subsequent nonlinear spatial transformation imply resampling, the follow-up dataset is slightly smoothed in the course of spatial normalization. Since this smoothing was considered a potential confounder for the comparison of the amount of GM-atrophy, $AT_{observed}$ was compared to $AT_{observable}$ rather than AT_{orig}.

4.4.3.3 Results

Figure 4.25 shows two datasets of such an intra-subject data series acquired at significantly different positions in the scanner. GM-atrophy was introduced into the region highlighted in figure 4.24.

[38] To reasonably reproduce the effect caused by resampling, the dataset was forward and backward transformed using the linear transformation determined as described in section 4.3.3.1.

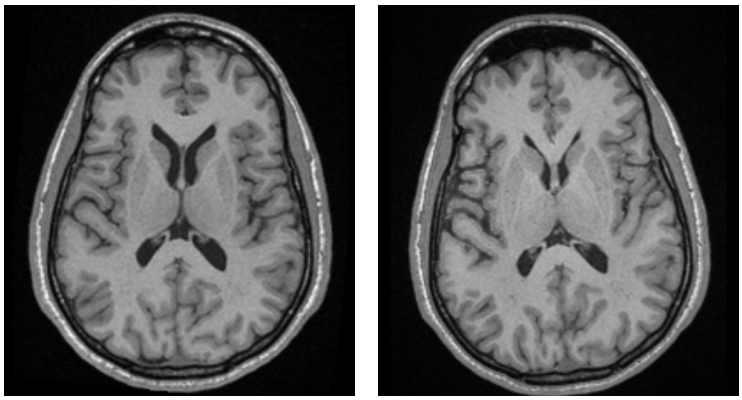


Figure 4.25: *Example input data for validation experiments (left: baseline dataset, right: follow-up dataset containing moderate regional artificial atrophy). Due to the different positions of the volunteer in the scanner, the datasets additionally contain acquisition-related morphological variability.*

The magnitude of acquisition-related morphological variability becomes evident in figure 4.26, which presents the result of these datasets' linear registration (using the method outlined in section 4.3.3.1). Atrophy is clearly noticeable in the difference images, but there are also differences in other image regions, such that – in a clinical setting – it would be difficult for a physician to reliably tell which differences reflect true anatomical changes in the patient and which ones are merely resulting from imperfections in data acquisition.

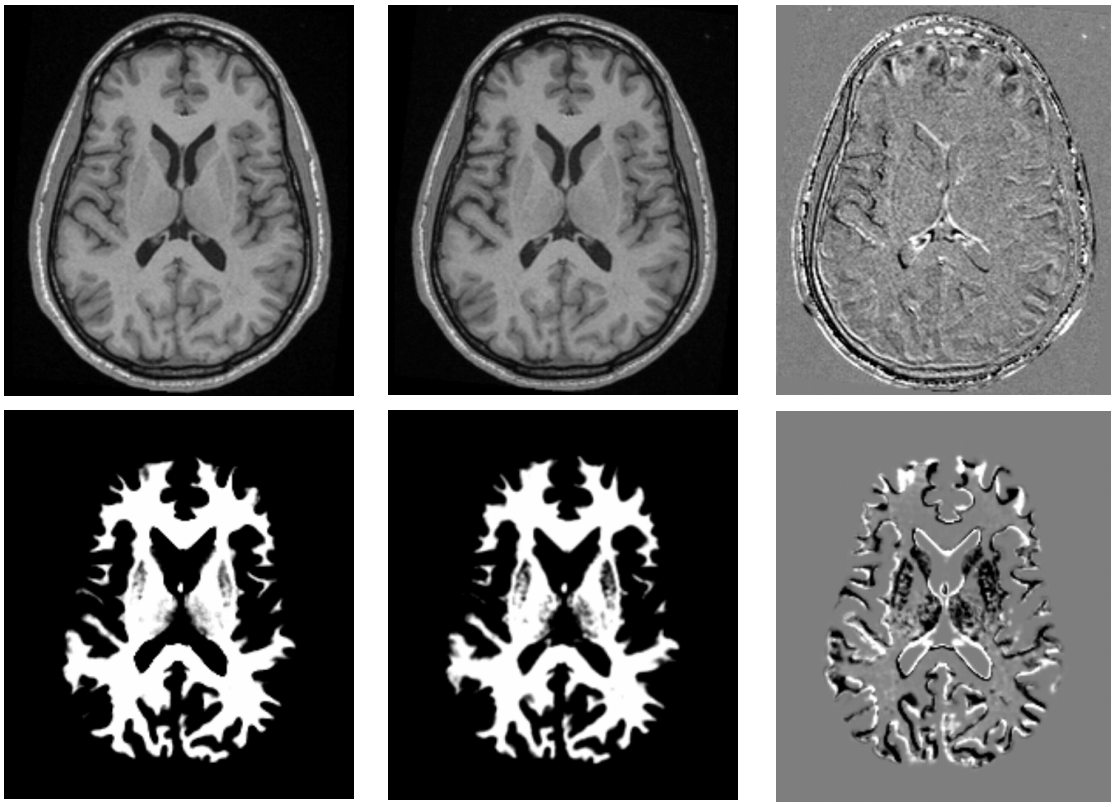


Figure 4.26: *Result of linear registration of the input data shown in figure 4.25 (top row: left: baseline dataset, middle: linearly registered follow-up dataset, right: difference image; bottom row: corresponding WM segmentation images). Acquisition-related morphological variability is superimposed on artificially introduced atrophy.*

The progress of WM-based spatial normalization is illustrated in figure 4.27, the final result in figure 4.28: Acquisition-related morphological variability in the brain is iteratively eliminated, whereas artificially introduced regional atrophy in the follow-

up data is preserved. Figure 4.28 also illustrates the finding that the additional incorporation of the global smoothness constraint according to [Horn81], like in the experiments with synthetic software phantom, did not improve the results: More specifically, not all artificially introduced GM-atrophy is preserved, and not all acquisition-related morphological variability is eliminated. The reason for this has been already discussed in section 4.4.2.3.

Figure 4.27: *Iterative progress of WM-based spatial normalization using the prototype with weighted mean interpolation only (upper left: baseline WM, upper right: linearly registered follow-up WM, lower left: follow-up WM after iteration 3, lower right: follow-up WM after iteration 6).*

Due to the spatial restriction of this normalization procedure to voxels representing brain tissue (see section 4.3.3.2.7), morphological differences in non-brain image regions are not reduced. With respect to the way the "detected atrophy" shows up in the difference image, it can be stated that this fully corresponds to the way it is introduced into the follow-up dataset, i.e. signal loss is significant only for voxels representing the transition between CSF and GM.

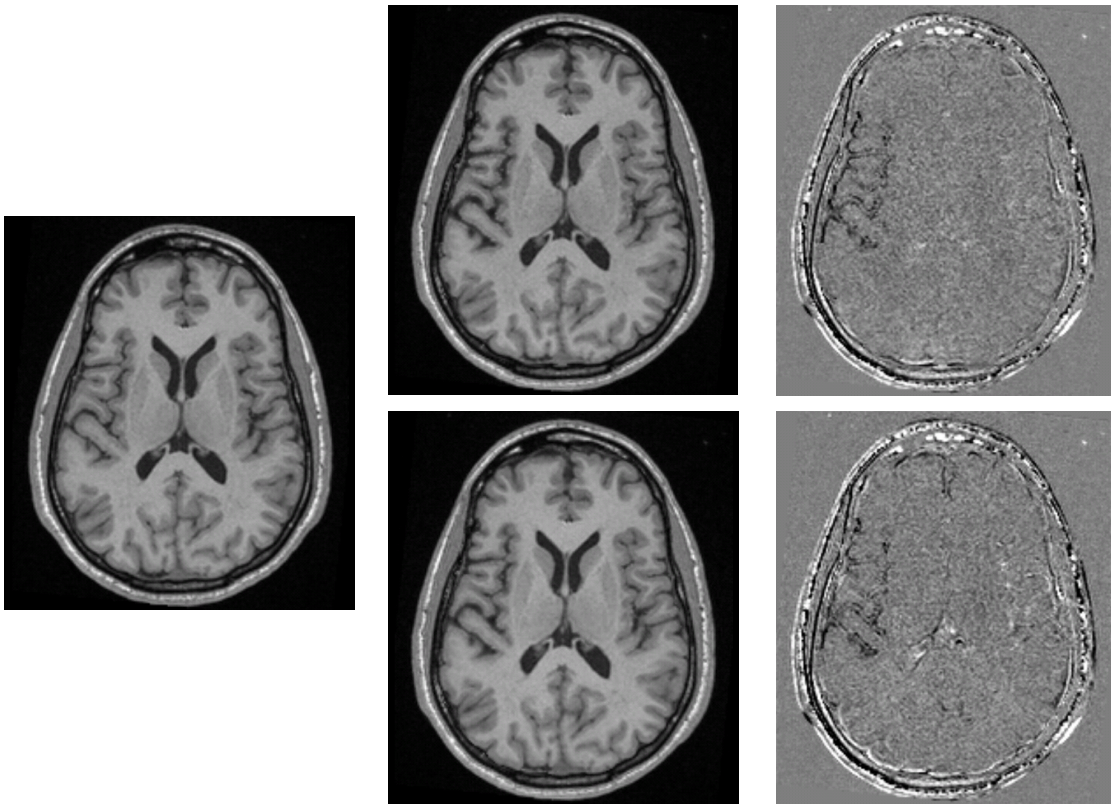


Figure 4.28: *Result of applying the proposed method for elimination of acquisition-related morphological variability from intra-subject MRI brain data to the input data shown in figure 4.25 (left: baseline dataset; top row middle: normalized follow-up dataset applying the prototype with weighted mean extrapolation only; top row right: difference image; bottom row middle: normalized follow-up dataset with additional incorporation of the smoothness constraint proposed by [Horn81], bottom row right: difference image).*

Typically, a few residual differences between the datasets remain despite WM-based spatial normalization. Based upon their characteristics in the difference image, they can be divided into two groups: (a) small bright or dark spots, and (b) dark/bright ribbons at the exterior border of the brain.

A close look at figure 4.29 shows that the spots anatomically correspond to either small vessels or the choroid plexus that give different signal in the baseline and the follow-up dataset even though they are spatially accurately registered (see red rectangles). Thus, these phenomena are not artefacts introduced by the normalization method. Most likely they are caused by vascular flow respectively CSF pulsation.

The comparison of the dark/bright ribbon highlighted in figure 4.29 by a red ellipse shows that it anatomically corresponds to the volunteer's meninges. As pointed out in section 4.3.2.3, the meninges are excluded from the brain mask, and thus, they are not spatially normalized by the implemented prototype for elimination of acquisition-related morphological variability. Consequently, these residual differences can also be explained by the prototype's technical implementation.

Most importantly, however, the outlined typical residual differences between baseline and follow-up datasets are very unlikely to confuse a physician in the diagnosis of potential neurodegeneration because neither vessels nor the meninges are meaningful with regard to such disorders. They can, thus, be considered clinically uncritical.

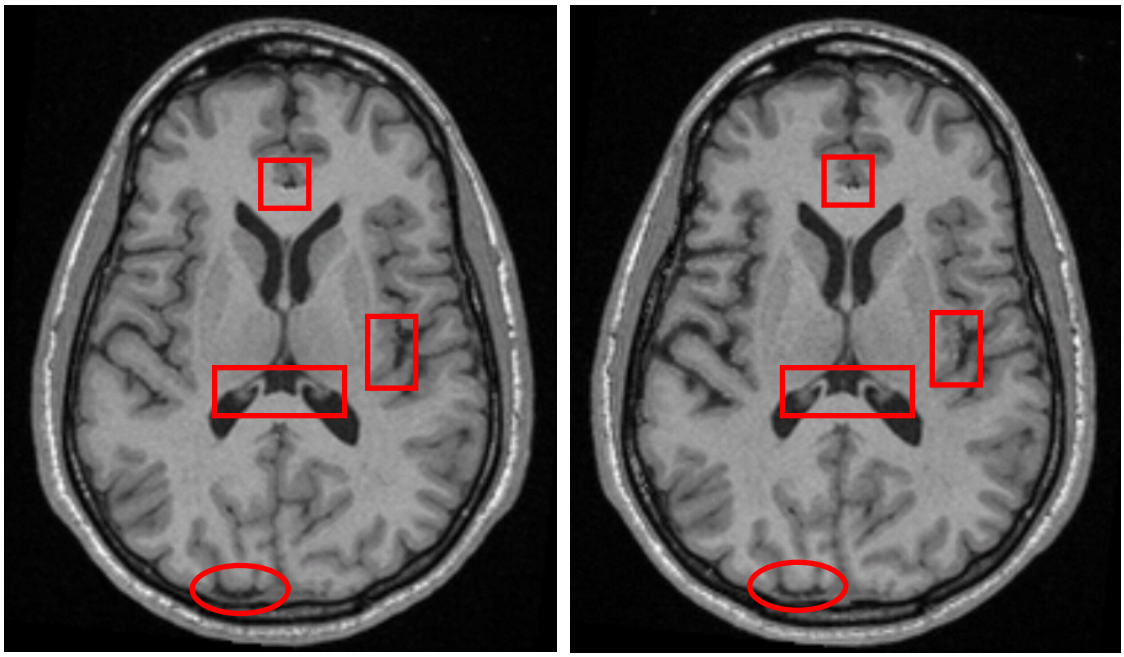


Figure 4.29: *Enlarged baseline (left) and WM-based normalized follow-up image (right). The red rectangles and ellipse mark regions that show typical residual differences between the datasets the origin of which are either small vessels, the choroid plexus or the meninges (for further explanation see text).*

The results of the quantitative validation are given in table 4.3.[39] They show that - abstracting from resampling influences - the proposed method is capable of reliably eliminating acquisition-related morphological differences from the hybrid phantom data series used, while preserving artificially introduced GM atrophy (difference between $AT_{observable}$ and $AT_{observed}$ less than 1%). Table 4.3 also proves that resampling systematically decreases the amount of atrophy originally contained in a dataset. If used in a clinical setting, the method proposed in section 4.3 therefore would have to be modified in such a way that linear and nonlinear registration are concatenated to the effect that only one resampling step is involved in the normalization procedure.

	#changed voxels	AT_{orig} (ml)	$AT_{observable}$ (ml)	$AT_{observed}$ (ml)
test case 1	15036	4348.56	4263.46	4231.10
test case 2	15941	4820.30	4696.21	4672.50
test case 3	16665	5310.23	5118.20	5102.89
test case 4	17301	5825.41	5553.66	5544.04
test case 5	17794	6369.57	5969.51	5958.48
test case 6	18201	6845.67	6362.43	6339.73

Table 4.3: *Results of the quantitative evaluation of validation experiments with "hybrid phantoms": Different amounts of artificial atrophy were introduced into the follow-up dataset. The number of voxels the intensity of which was reduced is listed in the first column. This change corresponds to the GM atrophy given in the second column. The observed atrophy after WM-based nonlinear spatial normalization is almost identical to the theoretically observable GM atrophy (taking into consideration the smoothing effect of resampling). Test case 4 corresponds to the one presented throughout this section in figures 4.24-4.29; all other test cases were created analogously, but with a different parameterization (follow-up dataset, ROI, parameters) of the operation used for introducing artificial atrophy (see section 4.4.3.1).*

[39] As mentioned above, it was obvious from pure inspection of the normalized images that the additional incorporation of the smoothness constraint proposed by [Horn01] detoriated the results in its current implementation. The quantitative evaluation, therefore, was not additionally carried out for datasets normalized with this option.

4.4.4 Clinical Data

4.4.4.1 Focus

Although the experimental results presented so far can be considered a proof-of-principle for the concept of eliminating acquisition-related morphological variability by means of WM-based spatial normalization, they are not sufficient to allow for a comprehensive evaluation of the implemented prototype in a clinical setting. Evidently, clinical patient data can be accompanied by phenomena not adequately modelled in the controlled validation datasets; in the context of image-based early detection of AD, the most relevant differences to the outlined hybrid phantom data series presumably are additional motion artefacts and slightly different contrast characteristics due to ageing. In order to still assess the method's clinical potential, this section discusses results obtained by applying the implemented prototype to clinical patient data. These experiments do not claim to be a full clinical validation, but they rather aim at identifying the current implementation's strengths and weaknesses with respect to phenomena that might not have been realistically modelled in the validation datasets presented in the previous sections.

4.4.4.2 Validation Strategy – Theoretical and Practical Problems

Clinical validation of image processing and –analysis methods in general is a very difficult task, since the underlying ground truth is usually unknown. A common way to deal with this lack of information is comparing an image analysis method's performance against results generated by medical experts in a manual or interactive manner. However, this strategy is impractical for validating the implemented prototype for eliminating acquisition-related morphological variability from data series of early AD patients, because the target phenomenon, i.e. GM atrophy, is too subtle to be reliably detectable by visual inspection in the presence of acquisition-related morphological variability. As illustrated in figure 4.26, it is almost impossible to differentiate between true morphological changes in the patient's anatomy and changes caused by imperfections in the data acquisition process, if the impact of these phenomena on the patient's representation in image space is in the same order of magnitude. Under these circumstances, the best validation strategy was thought to be relating image-based findings to changes in the patient's cognitive state assessed by neuropsychological tests like the MMSE. If the observed atrophy correlates with the measured cognitive decline, it is reasonable to consider the image-based findings as correct.

Unfortunately however, it proved to be very difficult to get access to intra-subject data series that comprise both high-quality MRI image data as well as neuropsychological measures of potential AD patients' cognition over a time interval of one to two years. In practice, either the CNR of the data was insufficient or the clinical state (e.g. estimated by neuropsychological measures) of the patient was unknown. In fact, it seems that such data series are just in the process of being acquired, and thus, not available for research groups not directly working at clinics participating in these studies. Therefore, the validation of experiments with clinical image data of potential early AD patients had to be reduced to a rather "soft" assessment of the obtained results' plausibility: Clearly, any GM gain, or GM loss in brain regions known to not be affected by early AD, e.g. sensory and motor areas, can be considered very implausible, whereas no atrophy or GM atrophy in the medial temporal lobe would be plausible findings.

4.4.4.3 Retrospective Analysis of Data Series of Confirmed AD Patients

As stated above, none of the intra-subject data series of confirmed AD patients available for this thesis provided sufficient image quality for WM-based elimination of acquisition-related morphological variability. More concretely, the CNR between WM and GM was too low to allow for a reliable segmentation of WM.

An example of such a data series is given in figure 4.30: The time interval between each image pair was reported to be about one year. For better comparison, the datasets are linearly registered by means of a modified variant of the method outlined in section 4.3.3.1: As a result of the poor WM/GM-CNR, the skull-stripped original images are used as input data for the registration procedure rather than the segmented WM representations.

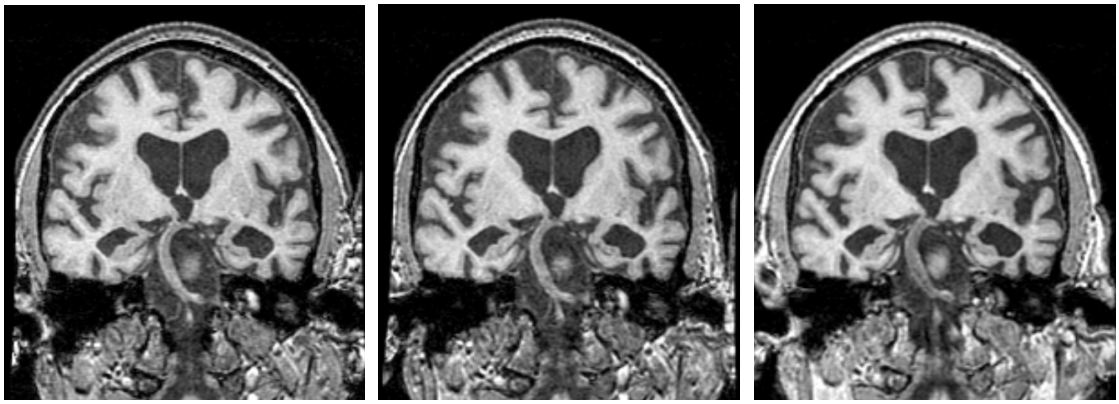


Figure 4.30: *Example data series of an AD patient (left: baseline dataset; middle: follow-up dataset acquired one year later; right follow-up dataset acquired one additional year later). The datasets are linearly registered to ease visual comparisons. Data are corrected for gradient coil nonlinearities.*

Morphological changes that occurred in this 2-year time interval become instantly apparent when looking at the difference images shown in figure 4.31: Most striking are (a) strong hippocampal atrophy and (b) significant enlargement of the ventricles. However, there are additional changes in other brain regions that could be caused equally well by the disorder or by imperfections in the data acquisition process. Regarding the clinical use of these findings, it has to be stated that atrophy is obvious in the baseline dataset already, such that the presented case cannot be considered an example of early detection of AD, but rather of image-based confirmation of pronounced AD.

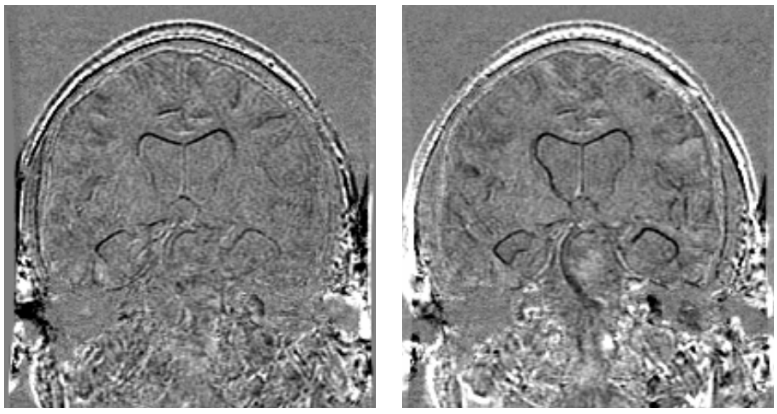


Figure 4.31: *Difference images of the datasets shown in figure 4.30 (left: differences between the first follow-up dataset and the baseline dataset; right: differences between the second follow-up dataset and the baseline dataset).*

It has been shown in section 4.4.2, that WM-based elimination of acquisition-related morphological variability relies heavily on a sufficiently strong CNR between GM and WM in the original brain datasets. Datasets that do not fulfil the minimum requirement of a WM/GM-CNR higher than 2½, cannot be spatially normalized with the presented prototype and are preferably analyzed by means of the boundary shift integral (see section 2.2.6).

4.4.4.4 Prospective Analysis of Data Series of Potential AD Patients

In order to assess the implemented prototype's clinical potential, 5 high-quality MRI data pairs of potential AD patients were analyzed, that were reported to have been acquired at a time interval of about 6 months (image data were kindly provided by the ADNI-committee as exemplary test cases). However, no neuropsychological measures of the patients' cognitive development were available, and no information was available regarding the question whether these potential AD patients developed AD or whether they did not. Therefore, the results obtained by applying the implemented prototype to these data series could not be compared against the clinical state of the patients, and the validation had to be reduced to a qualitative assessment of the obtained results' plausibility.

Generally spoken, applying the prototype lead to two different types of results: (a) normalized datasets that did not show any significant differences to the respective baseline dataset (see figure 4.32), and (b) datasets that contained residual signal loss in brain regions that are known to be typically affected by early AD (see figure 4.33). Taking into consideration the rather short time interval of 6 months and the clinical characteristics of the investigated population (potential AD patients), both result types are reasonable findings.

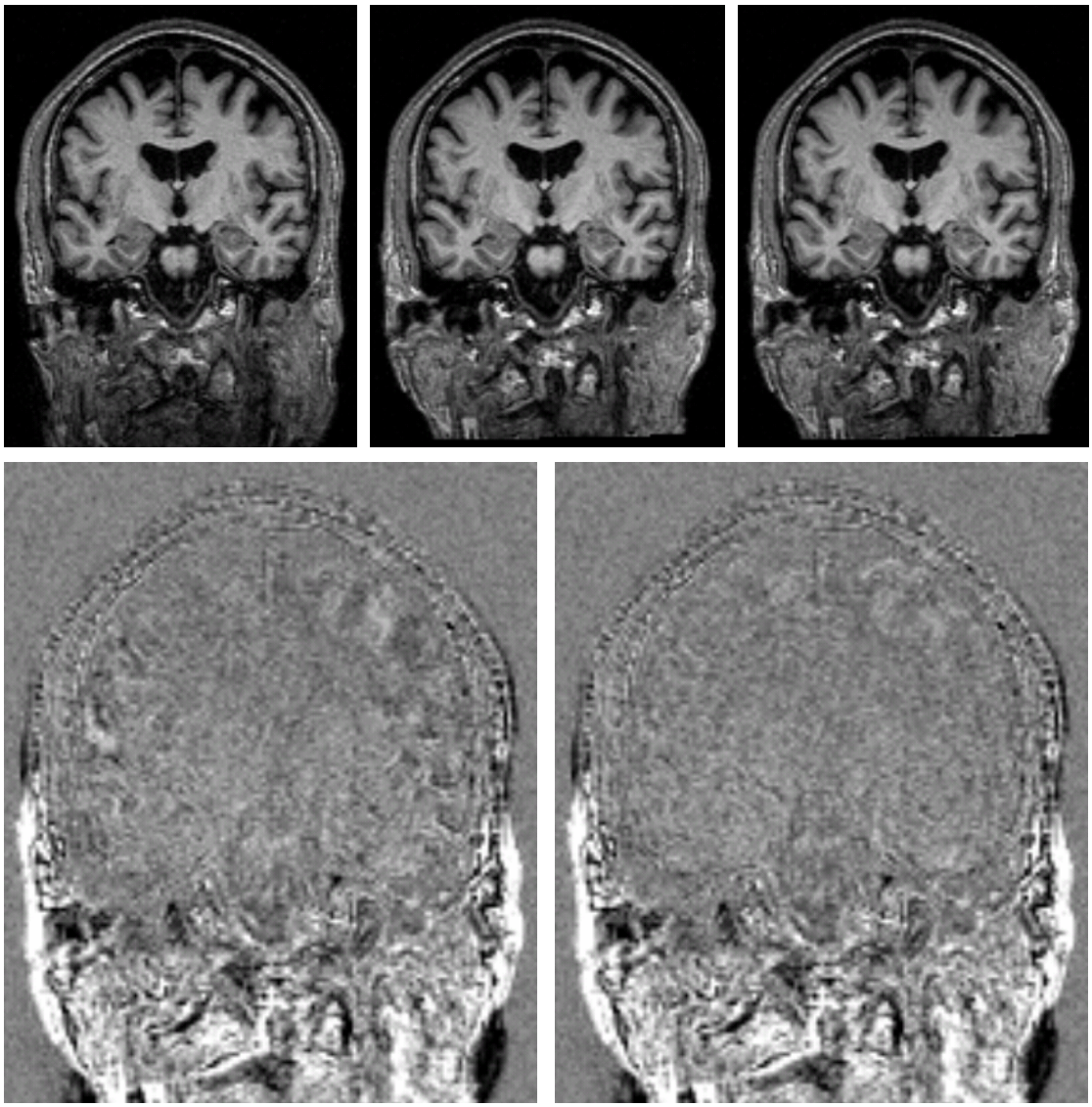


Figure 4.32: *Example data series of a potential early AD patient (top row: left: baseline dataset; middle: linearly aligned follow-up dataset acquired 6 months later; right: result of additional WM-based nonlinear spatially normalization; bottom row: corresponding difference images). No significant morphological differences remain after WM-based spatial normalization.*

In all cases the implemented prototype efficiently removed morphological variability from the datasets that was obviously not related to neurodegeneration. In the exemplary test data presented in figures 4.32 and 4.33, these are the prominent differences in the lateral (figure 4.32), and in the upper parts of the brain (figure 4.33). Importantly, figure 4.33 also illustrates that the prototype additionally eliminated more subtle changes that are not identifiable as acquisition-related changes at first glance (highlighted by red ellipses). In contrast, differences around the hippocampus (marked by red rectangles) do not vanish in figure 4.33, but they are even more pronounced in the spatially normalized images. Recapitulating section 2.1, this is a remarkable finding, since these regions are known to be affected by GM atrophy from the very beginning of the disease.

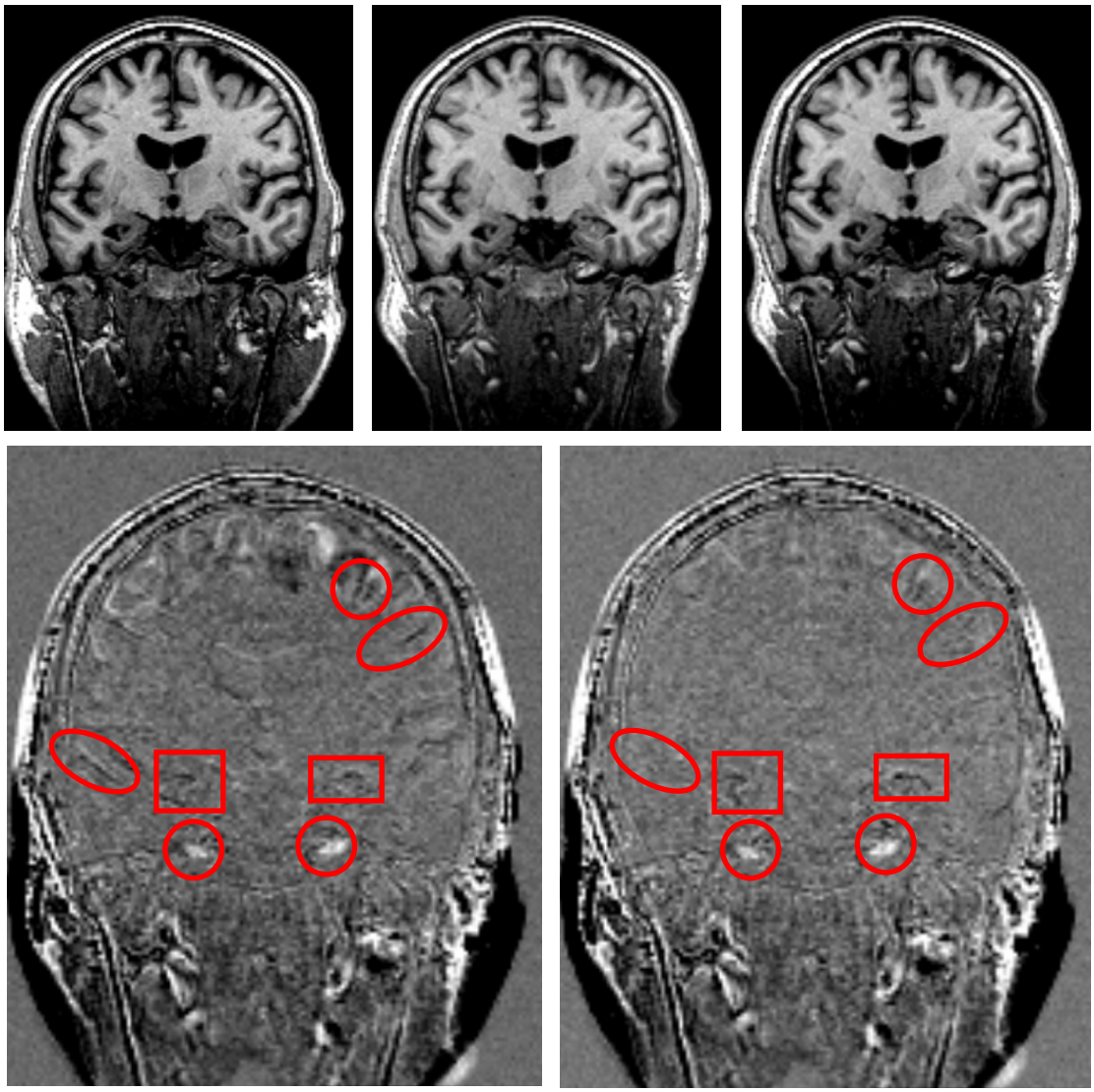


Figure 4.33: *Example data series of a potential early AD patient (top row: left: baseline dataset; middle: linearly aligned follow-up dataset acquired 6 months later; right: result of additional WM-based nonlinear spatially normalization; bottom row: corresponding difference images). Further explanations are given in the text.*

However, the difference image in figure 4.33 also contains residual morphological changes that cannot be explained by beginning AD. One source of these "false positives" are vascular structures (see lower red circles), that already have been identified as rather "uncritical confounders" in section 4.4.3.3. Instead, the difference in the topmost red circle is clinically relevant. Taking a close look at this region in the difference images and in the enlarged rigidly aligned follow-up image in figure 4.34, a ringing-like motion artefact becomes apparent. Due to this ringing the gradient control consistency check described in section 4.3.3.2.3 classifies the spatial intensity structure of these regions as not reliable for evaluation of optical flow. In subchapter 4.3, it was assumed that regions that do not pass this test are brought into better alignment by subsequent component-wise weighted mean extrapolation, and that - after a few iterations of the nonlinear registration procedure - they would be in spatial alignment good enough to allow for a direct evaluation of optical flow. The presence of ringing clearly can violate this assumption.

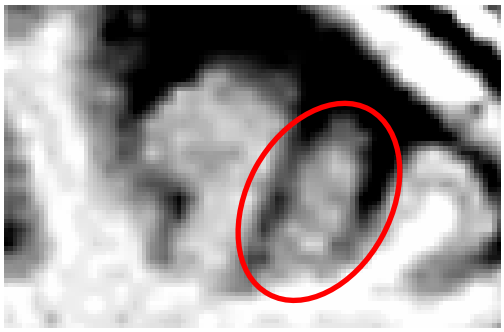


Figure 4.34: *Enlargement of the rigidly aligned follow-up image reveals the presence of ringing due to patient motion during scanning.*

4.4.5 Discussion

This chapter presented experimental results obtained with the previously presented software prototype for elimination of acquisition-related morphological variability in image space from longitudinal intra-subject MRI brain data.

Practical experiments with synthetic software phantoms that model the phenomena of interest, i.e. acquisition-related morphological variability and GM atrophy, by means of volumetric primitives, proved that the implemented prototype is capable of reliably solving this task in a simplified setting up to a WM/GM-CNR of 2½. This also implies that its application neither incorrectly introduced additional morphological differences, nor reduced the magnitude of true morphological changes. Experiments with different signal noise levels showed that the presented weighted mean interpolation is more robust to signal noise than the incorporation of the global smoothness constraint as proposed by [Horn81]. However, this is not to be misunderstood as a sign for the smoothness constraint's general inappropriateness, but rather as a consequence of a suboptimal implementation in terms of an uncritical use of the local intensity gradient $\nabla A(\vec{x},t)$ in the iterative Gauss Seidel equation 4.50. For a general evaluation of these two approaches with respect to their suitability for accurately filling gaps in flow fields in the presence of signal noise, alternative gradient-computation algorithms would have to be considered.

The same conclusions can be drawn from more complex validation experiments with intra-subject MRI brain data of a healthy volunteer, into which beginning GM atrophy was artificially introduced by means of a post-processing method that can be characterized as specifically parameterized kernel-based grey-value erosion. Using this operation, different amounts of regional GM atrophy were reasonably reproduced in longitudinal data series that contain acquisition-related morphological variability in image space. Quantitative evaluations of these data series showed that the implemented prototype efficiently eliminated acquisition-related morphological variability while not compromising artificially introduced GM atrophy. All additional residual differences between datasets despite WM-based spatial normalization could be explained by the prototype's technical implementation: Anatomically, these differences either corresponded to structures outside the brain, e.g. the meninges, or to vascular structures that showed different signal intensities despite having been accurately aligned. These phenomena, thus, are not artefacts introduced by the normalization method. Most importantly, they can be considered clinically uncritical, since it is very unlikely that they would confuse a physician in the diagnosis of potential neurodegeneration.

In order to assess the prototype's strengths and weaknesses with respect to phenomena that might have not been realistically modelled in the described hybrid phantom images, it was also applied to clinical data series of potential AD patients. Unfortunately, however, it proved to be very difficult to get access to intra-subject data series that comprise both high-quality MRI image data as well as neuropsychological measures of potential AD patients' cognition over a time interval of one to two years. Therefore, the validation of experiments with clinical image data of potential early AD patients had to be reduced to a rather "soft" assessment of the obtained results' plausibility.

In practice, the experiments with clinical data of potential early AD patients revealed that the implemented prototype removed morphological variability that was not related to neurodegeneration if the quality of the input datasets was sufficiently high. The most important criteria for image quality in this context are a strong CNR between WM and GM, and the absence of image artefacts like ringing. Whereas the former can be rather easily controlled by utilizing high-quality imaging protocols like those given in section 3.3.4, ringing has to be considered a phenomenon that is not avoidable, and in fact, likely to be present in clinical image data, since potential dementia patients are not necessarily capable of keeping their head still during data acquisition. Further research is therefore necessary before this implementation of WM-based spatial normalization can be applied in clinical routine. Nevertheless, the results obtained by applying the prototype to clinical patient data show the great potential that the general idea of eliminating acquisition-related morphological variability by means of nonlinear spatial alignment of WM has in high-accuracy MRI-based morphometry. Abstracting from the reported difficulties in the presence of ringing, no implausible results were created by the implemented prototype, i.e. neither unrealistic GM gain, nor GM loss in brain regions that are known to be not affected by early AD. Instead, the normalized datasets either showed no significant morphological changes or a small amount of GM atrophy around the hippocampus, which is known to be the brain region where GM atrophy starts in AD. Unfortunately, however, these findings could not be compared to the patients' clinical state in terms of neuropsychological measures like the MMSE-score, which is why they cannot be considered confirmed clinical findings.

The prototype's reported difficulty to appropriately deal with ringing artefacts already indicates the direction for future research. However, ringing caused by patient motion during scanning generally is considered a problem of data acquisition rather than of post-processing. In fact, novel techniques for prospective motion correction have been published which suggest that this problem can be solved by MRI physicists in the near future [Kouwe04]. The problem of WM-based spatial normalization of brain images in the presence of ringing artefacts therefore most promisingly is approached in close cooperation between MRI physicists and image processing professionals.

4.5 Conclusion

In this chapter, a novel concept for eliminating acquisition-related morphological variability from intra-subject MRI brain data series of early dementia patients was presented in terms of a software prototype that can be coarsely described as WM-based spatial normalization.

The review of prior research in image registration and -segmentation was focused on this specific scenario and resulted in the selection of (a) the EM algorithm in

conjunction with modelling the brain's intensity distribution by means of a Gaussian mixture model for WM segmentation, and of (b) 3D optical flow for nonlinear registration of the resulting WM representations. A detailed reasoning for this choice is given in the respective discussions (see sections 4.2.3.3 and 4.2.4.3).

In practice, the segmentation procedure itself was implemented as a processing pipeline consisting of the sequential execution of skull-stripping [Hahn00], intensity non-uniformity correction [Jellus05], brain segmentation, and tissue classification by means of the EM algorithm [Dempster77] [Pohl02].

Optical flow-based nonlinear registration was implemented following [Barron94] as the local least-squares fit of first-order constraints to a constant model for the flow velocity vector in each small spatial neighbourhood. By introducing a CNR-dependent confidence threshold, it was made sure that only numerically stable solutions to this fit are used for subsequent processing steps. Remaining gaps in the overall flow fields are subsequently filled by a component-wise weighted mean extrapolation of sufficiently reliable flow vectors to neighbouring voxels, which the initial least-squares fit did not give reliable results for. The weights are thereby determined based upon an error analysis that systematically assesses the correctness of assuming a constant model for the unknown translation within the local aperture. The overall deformation field's reliability is additionally supported by ensuring that the forward and backward transformation are consistent with each other in terms of each of them being the inverse of one another [Thirion98].

To ensure future extensibility, the prototype was implemented as a processing pipeline consisting of easily replaceable software modules. By this the flexibility is provided to not only extend existing modules, but also to easily explore alternative methods for e.g. intensity non-uniformity correction, nonlinear registration or deformation field extrapolation.

Controlled validation experiments with synthetic software phantoms as well as with human volunteer data, into which GM atrophy was artificially introduced, showed that the outlined prototypical implementation of WM-based spatial normalization is capable of eliminating acquisition-related morphological variability while preserving GM atrophy in datasets showing a WM/GM-CNR that is achievable in clinical morphometry today. Experiments with clinical data revealed that the implemented prototype removed morphological variability that was not related to neurodegeneration if the quality of the input datasets was sufficiently high, but that the method's current implementation is sensitive to ringing artefacts caused by patient movement during scanning.

Concluding, the results obtained with the software prototype presented in this chapter can be considered a proof-of-principle for the concept of eliminating acquisition-related morphological variability by means of WM-based spatial normalization. The prototype shows both convincing results in controlled validation experiments, and promising results in tests with clinical patient data, but needs further improvements regarding its stability towards ringing artefacts in the input data caused by patient movement during scanning. Irrespective of these remaining imperfections it can serve as the methodological basis for future developments aiming at fully exploiting the clinical potential of MRI-based morphomotry despite the investigated phenomenon of acquisition-related morphological variability.

5 Future Perspectives – Multi-Contrast Analyses

5.1 Introduction

In the last decade MRI-based high-accuracy morphometry was mostly done by exclusively analyzing T1-weighted images, as the only existing reasonably fast 3D acquisition techniques providing sufficiently high spatial resolution were gradient echo sequences giving T1 contrast. In the last years, however, new 3D acquisition schemes have been developed providing different contrasts in clinically practicable acquisition times, among them SPACE (sampling perfection with application optimized contrasts using different flip angle evolutions) for T2- and PD-weighted image data, and susceptibility-weighted imaging (SWI) that has been shown to be capable of highlighting local iron accumulation. Moreover, the ongoing technical progress in diffusion-weighted imaging (DWI) might make it possible in the future to analyze neurodegenerative diseases like Alzheimer's also regarding their impact on WM structure. Finally, it is well known that MRI also provides techniques for assessing the brain's function which might be useful to support morphometric analyses in image-based detection of neurodegenerative diseases. In total, it can be assumed that in the near future, post-processing applications are extended to allow for analyses of multiple contrasts in a combined way, since these contain complementary information.

This chapter therefore tries to identify which contrasts to concentrate on in the near future. This is done by inspecting both the contrasts' general potential with respect to early detection of neurodegenerative diseases, but also by analyzing their current technical limitations. Special attention is thereby paid to the problem of acquisition-related morphological variability: Due to sequence-specific differences in acquisition-determined geometrical distortions, the elimination of this variability becomes even more complex if post processing applications are built upon combined multi-contrast analyses of MRI data series. Finally, it is evaluated how well the method for elimination of acquisition-related morphological variability proposed in this thesis is applicable also to spatial normalization of longitudinal intra-subject inter-contrast data.

5.2 SPACE

5.2.1 Contrast Mechanism

As mentioned above, high-accuracy brain morphometry in the last decade was clearly dominated by the analysis of T1-weighted images, as the only existing reasonably fast 3D acquisition techniques providing sufficiently high spatial resolution were gradient echo sequences giving T1 contrast. Until recently, 3D turbo spin echo acquisition schemes providing T2 contrast required rather long measurement time, as the total number of echoes in the echo train was limited and the echo spacing of selective refocusing pulses was rather long. Therefore to achieve a desired contrast, intolerably long acquisition schemes were necessary or – with longer echo trains – degradation in contrast had to be accepted. SPACE solves this problem both by using non-selective refocusing pulses to shorten the echo spacing and by varying the flip-angle over the echo train [Mugler00]. In combination, this allows for the 3D acquisition of high-resolution datasets with different contrast characteristics, among them T2- and PD-

weighted images in acquisition times comparable to those known from standard gradient echo pulse sequences. In fact, variation of the refocusing pulse flip angle leads to significant contribution of T1 contrast in the image. Nevertheless, until now, neither the image appearance nor the lesion-conspicuity was reported to be significantly altered.

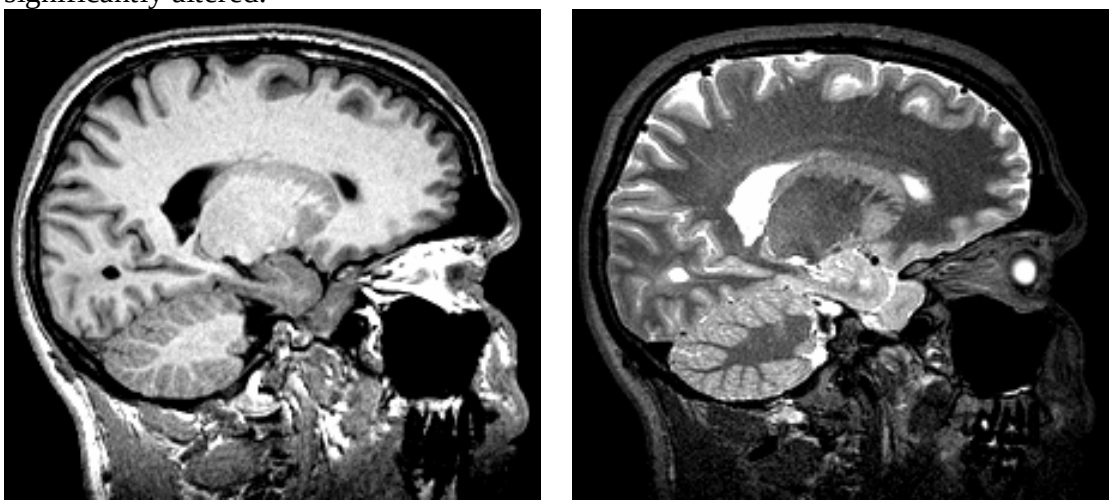


Figure 5.1. *Two corresponding slices from the same volunteer taken from a T1-weighted MPRAGE- (left) and a T2-weighted SPACE-dataset (right).*

5.2.2 Potential for Image-Based Detection of Neurodegenerative Diseases

Due to the unavailability of imaging techniques providing T2 or PD contrast at high isotropic spatial resolution until recently, not much research has been done yet regarding the systematic evaluation to what extent the additional analysis of such contrasts increases accuracy and reproducibility in the detection of subtle morphological changes in a patient's anatomy[40]. However, there already are clear indications that high-resolution T2- and PD-weighted data provide additional information on neurodegenerative processes that is not contained in T1-weighted datasets: Graham et al used T2-weighted images to investigate regional difference in brain iron deposition in patients with Parkinson's disease to healthy control subjects, and identified reduced iron content in the striatum and in the putamen [Graham00]. In MRI-based volumetry of multiple sclerosis (MS) lesions, T2- and PD-weighted data already have been integrated in image analysis procedures for years to increase the validity of results obtained with them - even though these datasets typically have an interslice resolution of 3mm only [Leemput01] [Zijdenbos02]. Thus, it is justified to assume that the additional use of T2- and PD-weighted data can help to increase the reliability of applications aiming at MRI-based early detection of neurodegenerative diseases.

5.2.3 Acquisition-Related Morphological Variability and Its Elimination

As shown in figure 5.1, T2-weighted SPACE datasets provide sufficient WM/GM contrast to allow for a reliable segmentation of WM. Consequently, the proposed

[40] SPACE was introduced by Siemens MR as commercially available product imaging sequence at the RSNA 2004.

method for WM-based elimination of acquisition-related morphological variability can be applied to SPACE data as well. In fact, the proposed method cannot only be used to spatially normalize longitudinal intra-subject SPACE data series, but also to normalize high-resolution SPACE- and MPRAGE- respectively 3DFLASH data for subsequent multi-contrast analyses. In other words: it is also a valuable technique for inter-contrast normalization.

To illustrate the magnitude of loss in effective spatial resolution in multi-contrast analyses due to sequence-specific differences in acquisition-determined geometrical distortions, relative morphological differences between two datasets have been quantified, one of them acquired with a standard MPRAGE protocol providing T1 contrast, and the other one with a typical SPACE-protocol giving T2 contrast (MPRAGE: TR 2300 ms, TI 1100 ms, TE 3.93 ms, flip angle 12°, matrix 224 x 256, bandwidth 130 Hz/pixel; SPACE: TR 3200 ms, TE 507.0 ms, variable flip angle, matrix 222 x 256, bandwidth 445 Hz/pixel, parallel imaging acceleration factor 2). The same phantom and analysis procedure as outlined in chapter 3 was used[41].

The maximum relative positional difference detected in a subvolume relevant for brain morphometry (spherical subvolume around the isocentre with a radius of 100 mm) was 0.6 mm and the maximum volumetric difference was 1.4 % (relative to a standard cell in *G0* of 7600 mm^3). Note that these values represent acquisition-determined effects only, since the data were acquired directly one after the other without repositioning the phantom. Obviously, some of this morphological inter-contrast variability is resulting from the protocols' difference in readout bandwidth (see section 3.2.2). However, clinical SPACE-protocols generally show a higher readout bandwidth, and transferring this to MPRAGE significantly reduces the SNR in the MPRAGE data, which is why the detected differences do not represent an academic worst-case scenario, but rather clinical practice.

To get an idea of the magnitude of loss in effective spatial resolution introduced by differences in readout bandwidth, the experiment was repeated with one additional data acquisition, using a modified MPRAGE protocol with a bandwidth of 440 Hz/Pixel. This ensures that susceptibility induced geometrical distortions are identical for the modified MPRAGE and the SPACE data. Practically, the maximum positional difference was reduced to 0.4 mm and the maximum volumetric difference to 1.0 % (relative to a standard cell in *G0* of 7600 mm^3). These residual differences are the sum of all remaining uncompensated sequence-specific effects, predominantly differences in eddy current induced distortions.

5.3 Susceptibility Weighted Imaging (SWI)

5.3.1 Contrast Mechanism

SWI basically utilizes local differences in magnetic susceptibility between neighbouring tissue types to enhance the contrast between them. As outlined in section 3.2.2.2, differences in magnetic susceptibility lead to slight changes in the local magnetic field strength, and thus, to regional differences in the spins' precession

[41] As a similar analysis of human volunteer data would have been confounded by differences in the volunteer's movements during scan, the exemplary investigation of effective spatial resolution in multi-contrast analyses has been purely done based on phantom data.

frequency. Consequently, signal from substances with different magnetic susceptibilities compared to their neighbouring tissue will become out of phase with these tissues at sufficiently long echo times. In SWI, this tissue-dependent difference in phase is used to create contrasts complementary to those known from standard structural imaging. Simplifying, this is done by multiplying the MRI magnitude images with the corresponding phase images – once those have been corrected for phase-wrapping artefacts [Haacke04].

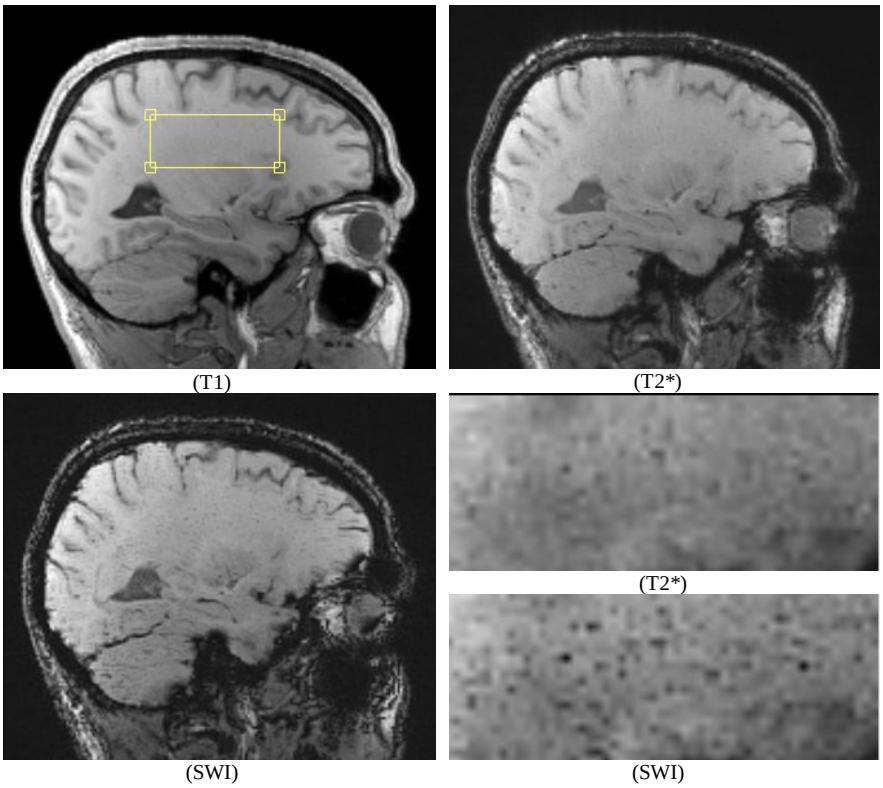


Figure 5.2. *As outlined in section 5.3.3, SWI can be integrated in a 3D multi-echo FLASH sequence, that allows for simultaneous acquisition of high-resolution T1- (upper left), T2*- (upper right) and susceptibility weighted data (lower left). The lower right quadrant shows enlarged T2*-weighted and SWI images of the area indicated in the T1-weighted data: Small vessels are clearly more pronounced in the SWI image than in the T2* image due to increased iron concentration in the vessels.*

5.3.2 Potential for Image-Based Detection of Neurodegenerative Diseases

SWI has been used to investigate vascular abnormalities, brain neoplasms and microbleeds, whereas its application to MRI-based detection of neurodegenerative diseases so far has been proposed theoretically only [Sehgal05]. However, its practical exploration in the field of dementia research rather seems to be a question of time: As pointed out in section 2.1, several neurodegenerative diseases are accompanied by increased iron accumulation, among them AD, and these local iron deposits clearly will be emphasized by SWI. The question whether SWI is even capable of highlighting amyloid plaques or neurofibrillary tangles remains to be answered. It is not completely unrealistic that these substances also differ in magnetic susceptibility from normal brain tissue such that they would become visible in SWI data. However, this can only be clarified by dedicated clinical experiments that would have to include comparisons to histological examinations of the imaged tissue. In

short, SWI may represent a valuable link between structural imaging and the molecular bases of neurodegenerative diseases.

5.3.3 Acquisition-Related Morphological Variability and Its Elimination

The proposed method for elimination of acquisition-related morphological variability by nonlinear registration of WM segmentation results can be directly applied to SWI data, if SWI is integrated into an appropriate gradient echo pulse sequence. As indicated in [Haacke04] already, SWI can even be implemented in terms of a 3D multi-echo FLASH approach that simultaneously provides T1, T2* and SWI contrast if the echo times are chosen appropriately. Acquisition-related morphological variability can then be determined based on the T1-weighted images showing strong WM/GM contrast. If the readout bandwidths are chosen to be identical for all echoes, the resulting nonlinear spatial transformation can be subsequently applied to the T1-weighted as well as the SWI dataset.

In order to demonstrate the feasibility of this 3D multi-echo FLASH approach, it has been prototypically implemented in this thesis: Five gradient echoes were acquired at TE=2.2 ms, 5.2 ms. 8.2 ms, 20.0 ms, and 25.4 ms (TR = 30ms; flip angle = 25°, isotropic resolution $1.3mm^3$). Images from the first three echoes were combined to obtain T1 contrast, whereas the images from the last two echoes were used to calculate T2* and SWI images as described above. The results are shown in figure 5.2.

In summary, the outlined multi-echo scheme allows for (a) elimination of acquisition-related morphological variability (based on T1-weighted data), (b) for standard morphometrical analysis (also based on T1-weighted data), and (c) for the analysis of local intensity abnormalities due to increased iron concentration (based on SWI data). Thus, it seems to be a very promising approach for multi-contrast analyses of subtle changes in a patient's anatomy.

5.4 Diffusion-Weighted Imaging

5.4.1 Contrast Mechanism

DWI provides insight into the microscopic structure of tissue. It is based upon the idea that in biological tissues, free diffusion of water molecules is hindered by barriers like cell membranes or myelin sheaths so that the diffusion of these water molecules is no longer unordered, but rather directed. This anisotropy of the observed diffusion is interpreted as the structural characteristic of the underlying tissue. In MRI, diffusion-weighting is achieved by first dephasing spins with a magnetic field gradient pulse and subsequently rephasing them with a second pulse. Any movement of the spins between these pulses results in incomplete rephasing, and thus, to signal loss, which is stronger the more diffusion has taken place in direction of the applied diffusion gradient. In order to be able to adequately characterize a voxel's diffusion profile, it is required to repeat the MRI scans, applying diffusion gradients in different directions (and with different magnitudes) [Bammer03].

Clearly, macroscopic motion, e.g. caused by respiration or CSF pulsation, can seriously corrupt the validity of DWI data, which is why in DWI, pulse sequences are preferred that allow for the acquisition of images with very high temporal resolution.

The most commonly used imaging technique in DWI today is echo planar imaging (EPI).

5.4.2 Potential for Image-Based Detection of Neurodegenerative Diseases

As outlined in section 2.2, image-based morphometry so far is targeting at quantifying the macroscopic morphology of anatomical entities. In contrast, DWI offers the opportunity to non-invasively investigate the microscopic structure of tissues. This includes processes like demyelinisation, violations of the blood-brain barrier and changes in the WM architecture due to neural death, all of which are common phenomena in neurodegeneration.

Regarding early detection of AD, it can be stated that there already has been found experimental evidence for the outlined hypothesis that DWI can provide additional information for early detection of neurodegeneration. Some research groups have reported decreased anisotropy in WM regions connecting brain regions affected in AD [Takahashi02], others found increased diffusivity in the medial temporal lobe in AD patients [Kantarci05]. Moreover, these changes were shown to be different from structural WM-changes in nondemented ageing [Head04].

Despite these convincing indications for a general suitability of DWI for early detection of neurodegenerative diseases, DWI is not used in this context as routinely as structural imaging techniques because of additional technical challenges it is accompanied by. Above all, the spatial resolution in DWI usually is significantly lower (typically by a factor of 2 to 3) than in structural imaging, which implies that the signal attenuation always is a result of a large amount of water protons diffusing in potentially different directions due to crossing or touching nerve fibers. Moreover, the nominal resolution in DWI does not truly reflect its effective spatial resolution: As outlined above, DWI requires the application of strong diffusion gradients in different directions, which inevitably leads to relative differences in eddy current induced geometrical distortions in the corresponding individual diffusion images. Consequently, diffusion measures derived from these individual images always imply a certain amount of “hidden” spatial averaging. Finally, it has to be stated that dementia patients are a specifically challenging population for DWI analyses because of the previously mentioned sensitivity of DWI to macroscopic motion [Jones05]; in dementia patients, not only CSF pulsation and respiratory motion have to be dealt with, but also with motion due to dementia patients’ inability to keep their head still during data acquisition.

In total, it can be stated that DWI is a very promising technique, but it does not yet seem to be reliable enough to allow for clinical longitudinal intra-subject high-accuracy applications. Significant technical progress seems to be required until it can be applied to dementia patients in a clinical setting.

5.4.3 Acquisition-Related Morphological Variability and Its Elimination

The considerations regarding the technical implementation of DWI make it obvious that acquisition-related morphological variability is significantly larger for DWI- than for structural imaging-series:

- Firstly, the additional diffusion gradient pulses are a lot stronger and changed more rapidly such that eddy current induced geometrical distortions are larger. Keeping in mind the influence of imperfections in patient-repositioning, it becomes apparent that acquisition-related morphological variability has to be larger, as well.
- EPI is an imaging technique that is very sensitive to local magnetic field distortions due to the low bandwidth in phase-encoding direction (typically around 30 Hz/Pixel). As mentioned in section 3.2.2.2, local differences in magnetic susceptibility can lead to an off-resonance frequency of up to 80 Hz in brain regions that are located close to air-tissue transitions, causing dislocations of 2-3 voxels in phase-encoding direction (applying equation 3.2). In the presence of patient repositioning errors these geometrical distortions obviously result in significant acquisition-related morphological variability.

Regarding the elimination of this acquisition-related morphological, the proposed WM-based method – in principle – is applicable to DWI data as well. For WM segmentation, inversion-recovery diffusion-weighted EPI would have to be used in order to get sufficient WM/GM-contrast in the datasets [Hadjiprocopis05]. However, in order for DWI to be effective in early detection of neurodegenerative diseases in a clinical setting, its seems to be necessary to better control signal variability due to macroscopic patient motion, and to increase the effective spatial resolution (e.g. by utilizing parallel imaging techniques and readout segmentation in EPI [Porter06], or by implementing non-Cartesian k-space trajectories like spirals [Li99]).

5.5 Further Imaging Contrasts – Beyond Morphometry

All MRI contrasts discussed so far provide information on tissue structure. However, it is well known that MRI also allows for the assessment of brain function: Whereas functional MRI (fMRI) aims at characterizing brain activation by measuring the rate of loss of the spins' phase coherence resulting from regional deoxygenation [Logothetis04][42], perfusion imaging quantifies cerebral blood flow either by application of contrast agent or by magnetically labelling arterial spins by means of specifically designed RF-pulses and subsequently measuring their signal contribution in brain regions they flow into [Calamante99]. Both techniques assume that a local increase in brain activation is accompanied by a regional increase in oxygenated arterial blood delivery.

Only very little is known about the effect of neurodegenerative diseases on the brain's functional organization: fMRI studies regarding this topic have reported decreased brain activation in cognitively normal individuals at high risk for AD in parts of the temporal lobe [Smith99], and the activation of a greater extent of the parahippomcampal gyrus in clinically presymptomatic AD patients [Dickerson04]. The latter finding is interpreted as a sign of compensation for regional loss of neurons caused by the disease. First perfusion imaging studies have revealed regional hypoperfusion in AD patients compared to normal controls in the right inferior

[42] For better readability, this is a rather simplified description of the fMRI contrast mechanism. In fact, the so-called BOLD (blood oxygenation level dependent) contrast depends not only on blood oxygenation but also on the cerebral blood flow and volume. Further details can be found in [Logothetis04].

parietal lobe, which is reported to be consistent with similar PET and SPECT studies [Johnson05].

Since both fMRI as well as perfusion imaging requires high temporal resolution, their implementation typically is based upon an EPI readout module. Thus, the resulting drawback regarding the magnitude of magnetic susceptibility induced geometrical distortions, as well as the limitation in spatial resolution – both effects described in the discussion of DWI already - hold true for these imaging techniques as well. Preferably, these problems should be solved prior to an integration of fMRI or perfusion data into high-accuracy analyses of longitudinal intra-subject data series in a clinical setting.

5.6 Conclusion

In the last years new 3D acquisition schemes have been developed providing contrasts other than T1-weighting, among them SPACE for T2- and PD-weighted image data, and SWI that highlights local iron accumulation. Since both techniques allow for the acquisition of isotropic high-resolution data in clinically practicable acquisition times, it is close at hand to incorporate these contrasts in morphometric analyses targeting at early detection of neurodegenerative diseases.

In this chapter, it has been shown that SWI can be smoothly integrated in a 3D multi-echo FLASH sequence simultaneously providing T1, T2* and SWI contrast. If the readout bandwidths are chosen to be identical, there are no inter-contrast differences in acquisition-determined geometrical distortions between these datasets. Therefore, acquisition-related morphological variability in longitudinal inter-contrast intra-subject data series can be eliminated solely based on the T1-weighted data by WM-based spatial normalization. Since the outlined multi-echo acquisition scheme allows for standard morphometric analyses as well as for the evaluation of regional intensity abnormalities due to local iron accumulation, it can be considered a promising approach for the detection of subtle changes in a patient's anatomy based upon multi-contrast analyses of MRI data.

SPACE offers the possibility to extend such analyses by high-resolution T2- and PD-weighted data. However, it has been shown that a naive combination of SPACE- and 3DFLASH respectively MPRAGE data without taking into consideration acquisition-determined morphological inter-contrast differences bears the risk of lowering an application's accuracy and reproducibility compared to a standard single contrast analysis. The outlined exemplary quantification of relative morphological differences between a T1-weighted MPRAGE- and a T2-weighted SPACE-dataset acquired with clinical protocols revealed that these differences have to be taken into consideration in high-accuracy applications. In detail, local volumetric differences of up to 1.4% were measured in a subvolume relevant for brain morphometry. Note that this inter-contrast variability was purely caused by differences in the pulse sequences, as no repositioning or re-shimming was done. Since T2-weighted SPACE datasets show high contrast between GM and WM, the method for elimination of acquisition-related morphological variability from intra-subject data series proposed in this thesis is also suitable for eliminating the described inter-contrast variability.

Regarding DWI, fMRI and perfusion imaging, it was shown that these techniques have a great potential for image-based early detection of neurodegenerative diseases, but that they do not yet seem to be reliable enough to allow for high-accuracy

analyses of longitudinal intra-subject data series of dementia patients in a clinical setting. Above all, this is a consequence of insufficient spatial resolution, of significantly stronger acquisition-determined geometrical distortions, and – in case of DWI – of an increased sensitivity to macroscopic patient motion. However, once these technical problems have been solved, the idea of eliminating acquisition-related morphological variability based upon nonlinear registration of WM portions can be applied to these data as well, since the underlying pulse sequences can be extended by an inversion recovery module to allow for automated WM segmentation.

Summarizing, it can be stated that the method for elimination of acquisition-related morphological variability proposed in this thesis cannot only be applied to the image contrasts that are relevant for high-accuracy morphometry today, but it is also compatible to probable future developments in DWI, fMRI and perfusion imaging.

6 Final Summary and Conclusions

Over the past decades a great variety of different computational methods to quantitatively describe anatomy based upon the analysis of medical image data have been proposed by the image processing and –analysis community. As shown in chapter 2, however, none of the various approaches explicitly takes into consideration potential morphological inaccuracies inherently contained in the input images. Instead, all these methods implicitly and tacitly assume that the data acquisition process does not introduce any morphological variability into image space.

In chapter 3, it was shown that this assumption is not correct for MRI-based morphometry. Imperfections in the MRI data acquisition process, along with practical limitations in the reproducibility of patient positioning, and imperfect hardware stability were shown to introduce morphological variability into an object's representation in image space even for a physically unchanged object. The analysis of the underlying MRI physics revealed that this acquisition-related morphological variability corresponds to a nonlinear spatial transformation of the imaged object's morphology in image space, whereby the underlying deformation field is dependent on the scanner, the imaging sequence, its parameterization and the imaged object itself. As a direct consequence of this complexity, the - from a clinical point of view - crucial question to what extent the MRI data acquisition process quantitatively affects the detection limit of existing morphometric applications cannot be answered purely theoretically, but only by means of an experimental study. The study carried out in this thesis was specifically designed to enable medical professionals for the first time to directly infer to what degree morphometric measures typically are confounded by acquisition-related morphological variability.

In practice, the study revealed that diseases that are accompanied by gradual morphological changes cannot be reliably detected with imaged-based techniques until the actual anatomical change exceeds a volume change of at least 2%. With regard to this thesis's medical context, Alzheimer's disease (AD), this is too much to be considered irrelevant for MRI-based early AD detection. Taking the findings of Thompson et al as a basis for the progress of AD [Thompson03], this value corresponds to a delay in MRI-based early detection of AD of nearly 5 months.

A direct consequence of this study, therefore, is the necessity to develop methods capable of further reducing acquisition-related morphological variability. From a technical point of view, this reduction is a question of spatial normalization that can be tackled by utilizing either extrinsic or intrinsic markers. An evaluation of these two general approaches' suitability for spatial normalization of longitudinal intra-subject data series in a clinical setting revealed that spatial normalization using intrinsic markers generally is superior a long as an appropriate intrinsic marker can be identified. For early detection of AD, this was shown to be WM. In total, this analysis resulted in the proposal of a novel concept for eliminating acquisition-related morphological variability from intra-subject MRI brain data series of early dementia patients that can be coarsely described as WM-based spatial normalization.

In chapter 4, this concept was presented in terms of a software prototype: In short, the WM segmentation procedure was implemented as the sequential execution of skull-stripping [Hahn00], intensity non-uniformity correction [Jellus05], brain segmentation, and tissue classification by means of the EM algorithm [Dempster77] [Pohl02]. Nonlinear registration of the resulting WM representations was

implemented as a process of optical flow following the work of [Barron94]. To ensure future extensibility, the prototype was implemented as a processing pipeline consisting of easily replaceable software modules. By this the flexibility is provided to not only extend existing modules, but also to easily explore alternative methods for e.g. intensity non-uniformity correction, nonlinear registration or deformation field extrapolation.

In summary, the results obtained with this software prototype can be considered a proof-of-principle for the concept of eliminating acquisition-related morphological variability by means of WM-based spatial normalization. The prototype shows both convincing results in controlled validation experiments, and promising results in tests with clinical patient data, but needs further improvements regarding its stability towards ringing artefacts in the input data caused by patient movement during scanning. As it was also shown to be applicable to imaging contrasts other than T1-weighting, it can – irrespective of the remaining imperfections – serve as the methodological basis for future developments aiming at fully exploiting the clinical potential of MRI-based morphomotry despite the investigated phenomenon of acquisition-related morphological variability.

Taking a more general look at this thesis, it can be seen as an example for the necessity in medical image processing and -analysis to always investigate the complete clinical workflow including the image formation process. Without a profound understanding of the latter, it is not possible to provide physicians with any post-processing application's limits in a clinical setting, and thus, with a reliable statement on the clinical validity of results obtained with this application.

Appendix A - Glossary

Accuracy — *Accuracy* describes the degree of conformity of a *measured* quantity to its true value.

Acquisition-Related Variabiltiy — *Acquisition-related variability* is the *variability* of the *measurement* that is closely related to, but not solely *determined* by the data acquisition process. It also includes effects caused by the patient's movement during scanning or by imperfections in the *reproducibility* of patient positioning.

Acquisition-Determined Variability — *Acquisition-determined variability* is the *variability* in the *measurement* that is solely *determined* by the data acquisition process. It does not include effects caused by the patient's movement during scan or by imperfections in the *reproducibility* of patient positioning.

Active Shimming — *Active shimming* describes the process of adjusting the magnetic field homogeneity with the imaged object in the scanner in order to minimize inhomogeneities caused by differences in magnetic susceptibility. It consists of an automatic determination of the field homogeneity immediately before the scanning process starts. This magnetic field map is then used to automatically adjust the setting of the dedicated shimming coils.

Approximation — An *approximation* is an inexact representation of something that is still close enough to be useful.

Area — *Area* is a physical quantity expressing the size of a part of a surface. In the study of brain function, however, the term is also used to describe regions that are known to be responsible for a certain cognitive ability. In accordance to this practice, this thesis names a cortical region an *area* whenever it is defined by its function rather than its anatomical location.

Determination — *Determination* describes the ascertaining or fixing of the magnitude of a quantity.

Estimation — *Estimation* is the calculated *approximation* of a result which is usable even if input data may be incomplete, uncertain or noisy.

Image Analysis — *Image analysis* is the extraction of information from images. In image-based *morphometry*, the term is closely related to both *determination* and *estimation*, as every result *determined* by the application of *image analysis* algorithms implicitly is an *estimation* of the true magnitude of the anatomical quantity.

Image Processing — *Image processing* is a modification of the image, e.g. to make it more suitable for subsequent *image analysis* tasks or for visualization purposes.

Measurement In a strict sense, *measurement* means the act of comparing the magnitude of a quantity to the magnitude of a unit. In this thesis, however, the term is often also used in a more general way comprising the complete experiment, including patient positioning, data acquisition, *image processing* and *image analysis*.

Morphometry The term *morphometry* is derived from the Greek words morpho - shape and metron - measure, and is generally used in medicine to describe the discipline of quantitative anatomy. Though being – in principle – also applicable to the 2D analysis of anatomical structures, the term is used in this thesis as a synonym for 3D image-based quantitative anatomy by means of *image processing* and -*analysis* techniques.

Passive Shimming *Passive shimming* describes the process of adjusting the magnetic field homogeneity without the imaged object in the scanner. It aims at minimizing static field inhomogeneities which are practically unavoidable in the construction of whole-body magnets.

Reconstruction see *surface-reconstruction*

Registration Medical image *registration* is the *determination* of a transformation that spatially aligns two images such that the location of an anatomical feature in one dataset is mapped to the location of the corresponding feature in the other dataset.

Repeatability *Repeatability* is the variation in *measurements* taken on the same item under the same conditions within a short period in time.

Reproducibility *Reproducibility* is the variation in *measurements* taken on the same item under the same conditions, but allowing a substantial time interval in between the *measurements*.

Segmentation Medical image *segmentation* is generally considered the delineation of anatomical structures or tissue types in previously acquired patient data, whereby it is often distinguished between *hard* and *soft segmentation*.

Hard segmentation interprets delineation in a strict sense, meaning that it understands *segmentation* as the partitioning of an image into a number of non-overlapping regions such that each voxel belongs to exactly one region which represents one object.

Soft segmentation algorithms extend this concept of disjoint image partitions insofar as they estimate for each voxel the probability of class membership, whereby the candidate classes need to fully explain each voxel.

Surface-reconstruction *Surface-reconstruction* describes the process of constructing a surface out of a binary voxel set resulting from *segmenting* an anatomical structure in an image. The resulting surface often is also referred to as surface-model to indicate that it is merely an *approximation* of the true surface.

Validity If a result is both *accurate* and *reproducible*, it is called *valid*.

Variability The tendency of a *measurement* process to produce slightly different results for the same test item. *Variability* is a generalization of both *repeatability* and *reproducibility*.

Appendix B - Acronyms

3DFLASH	3-dimensional fast low angle shot
AD	Alzheimer's disease
AP	amyloid plaques
APP	amyloid precursor protein
BOLD	blood oxygenation level dependent
BSI	boundary shift integral
CC	correlation coefficient
CNR	contrast-to-noise ratio
CSF	cerebrospinal fluid
CT	computed tomography
DBM	deformation-based morphometry
DWI	diffusion-weighted imaging
EPI	echo-planar imaging
fMRI	functional magnetic resonance imaging
FoVp	field of view in phase encoding direction
FoVr	field of view in readout direction
SPGR	fast spoiled gradient refocused acquisition into steady state
FTD	fronto-temporal dementia
GM	grey matter
iPAT	integrated parallel acquisition techniques
MCI	mild cognitive impairment
MGH	Massachusetts General Hospital
MI	mutual information
MMSE	mini mental state examination
MPRAGE	magnetization-prepared rapid gradient echo
MRI	magnetic resonance imaging
MS	multiple sclerosis
NFT	neurofibrillary tangles
PD	proton density
ROI	region of interest
RSNA	Radiological Society of North America
SNR	signal-to-noise ratio

SPACE	sampling perfection with application optimized contrasts using different flip angle evolutions
SSD	summed squared difference
SWI	susceptibility-weighted imaging
T1-FFE	T1-weighted fast field echo
TE	echo time
TI	inversion time
TR	repetition time
VAD	vascular dementia
VBM	voxel-based morphometry
VBR	ventricle-brain ratio
WM	white matter

Appendix C - Bibliography

[Andrade97] M.C. de Andrade, G. Bertrand, A.A. Araujo. Segmentation of Microscopic Images by Flooding Simulation: A Catchment Basins Merging Algorithm; Proceedings of the SPIE, 3026, 1997, 164-175

[Arun87] K.S. Arun, T.S. Huang, S.D. Blostein. Least-Squares Fitting of Two 3-D Sets; IEEE Transactions on Pattern Analysis and Machine Intelligence, 9, 1987, 698-700

[Ashburner99] J.T. Ashburner, K.J. Friston. Nonlinear Spatial Normalization Using Basis Functions; Human Brain Mapping, 7, 1999, 254-266

[Ashburner00] J.T. Ashburner. Voxel-Based Morphometry – The Methods; NeuroImage, 11, 2000, 805-821

[Bajcsy89] R. Bajcsy, S. Kovacic. Multiresolution Elastic Matching; Computer Vision, Graphics, and Image Processing, 46, 1989, 1-21

[Bammer03] R. Bammer. Basic Principles of Diffusion-Weighted Imaging; European Journal of Radiology, 45, 2003, 169-184

[Barnes04] J. Barnes, R.I. Scahill, R.G. Boyesa et al. Differentiating AD from Aging Using Semiautomated Measurement of Hippocampal Atrophy Rates; NeuroImage, 23, 2004, 574-581

[Barron94] J.L. Barron, D.J. Fleet, S.S. Beauchemin, Performance of Optical Flow Techniques; International Journal of Computer Vision, 12, 1994, 43-77

[Bartzokis00] G. Bartzokis, T.A. Tishler. MRI Evaluation of Basal Ganglia Ferritin Iron and Neurotoxicity in Alzheimer's and Huntington's Disease; Cellular and Molecular Biology, 46, 2000, 821-833

[Basser00] P.J. Basser S. Pajevic, C. Pierpaoli et al. In Vivo Fiber Tractography Using DT-MRI Data, Magnetic Resonance in Medicine, 44, 2000, 625-632

[Bazin05] P.L. Bazin, D.L. Pham. Topology Correction Using Fast Marching Methods and ist Application to Brain Segmentation, Proceedings of the MICCAI, (2), 2005, 484-491

[Besl92] P.J. Besl, N.D. McKay. A Method for Registration of 3-D Shapes; IEEE Transactions on Pattern Analysis and Machine Intelligence, 14, 1992, 239-256

[Bookstein91] F.L. Bookstein. Thin-Plate Splines and the Atlas Problem for Biomedical Images; In: A.C.F. Colchester, D.J. Hawkes. Information Processing in Medical Imaging, Springer-Verlag, 1991, 326-342

[Braak91] H. Braak, E. Braak. Neuropathological Stageing of Alzheimer-Related Changes; Acta Neuropathologica, 82, 1991, 239-259

[Breeuwer00] M.M. Breeuwer, M. Holden, W. Zylka. Detection and Correction of Geometric Distortion in 3D MR Images; Proceedings of the SPIE, 4322, 2001, 1110-1120

[Calamante99] F. Calamante, D.L. Thomas, G.S. Pell et al. Measuring Cerebral Blood Flow Using Magnetic Resonance Imaging Techniques; Journal of Cerebral Blood Flow and Metabolism, 19, 1999, 701-735

[Canny86] J. Canny. A Computational Approach to Edge Detection; IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 187-217

[Chung01] M.K. Chung, K.J. Worsley, T. Paus et al. A Unified Approach to Deformation-Based Morphometry; NeuroImage, 14, 2001, 595-606

[Chung03] M.K. Chung, K.J. Worsley, S. Robbins et al. Deformation-Based Surface Morphometry Applied to Gray Matter Deformation; NeuroImage, 18, 2003, 198-213

[Cohen91] I. Cohen, L.D. Cohen, N. Ayache. Introducting New Deformable Surfaces to Segment 3-D Images; Proceedings of the CVPR, 1991, 738-739

[Cortechs04] http://www.cortechs.net/deepgray.htm; 09.03.2004

[Davatzikos96] C. Davatzikos, R.N. Bryan. Using a Deformable Surface Model to Obtain a Shape Representation of the Cortex; IEEE Transactions on Medical Imaging, 15, 1996, 785-795

[Davatzikos97] C. Davatzikos. Spatial Transformation and Registration of Brain Images Using Elastically Deformable Models; Computer Vision and Image Unterstanding, 66, 1997, 207-222

[Davatzikos01] C. Davatzikos, A. Genc, D. Xu et al. Voxel-Based Morphometry Using the RAVENS Maps: Methods and Validation Using Simulated Longitudinal Atrophy; NeuroImage, 14, 2001, 1361-1369

[Dempster77] A. Dempster, N. Laird, D. Rubin. Maximum Likelihood from Incomplete Data Via the EM Algorithm; Journal of the Royal Statistical Society, B, 39, 1977, 1-38

[Dengler88] J. Dengler, M. Schmidt. The Dynamic Pyramid – A Model for Motion Analysis with Controlled Continuity; International Journal of Pattern Recognition and Artificial Intelligence, 2, 1988, 275-286

[Dickerson04] B.C. Dickerson, D.H. Salat, J.F. Bates et al. Medial Temporal Lobe Function and Structure in Mild Cognitive Impairment; Annals of Neurology, 56, 2004, 27-35

[Elias00] M.F. Elias, A. Beiser, P.A. Wolf et al. The Preclinical Phase of Alzheimer Disease: A 22-Year Prospective Study of the Framingham Cohort; Archives of Neurology, 57, 2000, 808-813

[Evans91] A.C. Evans, W. Dai, L. Collins et al. Warping of a Computerized 3-D Atlas to Match Brain Image Volumes for Quantitative Neuroanatomical and Functional Analysis; Proceedings of the SPIE, 1445, 1991, 236-246

[Fernandez05] M. Martin-Fernandez, S. Bouix, L. Ungar et al. Two Methods for Validating Brain Tissue Classifiers; Proceedings of the MICCAI, 2005, 515-522

[Ferrant01] M. Ferrant, A. Nabavi, B. Macq et al. Registration of 3D Intraoperative MR Images of the Brain Using a Finite Element Biomechanical Model; IEEE Transactions on Medical Imaging, 20, 2001, 1384-1397

[Fischl99a] B. Fischl, M.I. Sereno, A.M. Dale. Cortical Surface-Based Analysis I. Segmentation and Surface Reconstruction; NeuroImage, 9, 1999, 179-194

[Fischl99b] B. Fischl, M.I. Sereno, A.M. Dale. Cortical Surface-Based Analysis I. Inflation, Flattening , and a Surface-Based Coordinate System; NeuroImage, 9, 1999, 195-207

[Fischl00] B. Fischl, A.M. Dale. Measuring the Thickness of the Human Cerebral Cortex From Magnetic Resonance Images; Proceedings of the NAS, 97, 2000, 11050-11055

[Fischl01] B. Fischl, A. Liu, A.M. Dale. Automated Manifold Surgery: Constructing Geometrically Accurate and Topologically Correct Models of the Human Cerebral Cortex; IEEE Transactions on Medical Imaging, 20, 2001, 70-80

[Fischl02] B. Fischl, D.H. Salat, E. Busa et al. Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain; Neuron, 33, 2002, 341-355

[Fischl04] B. Fischl, D.H. Salat, A.J. van der Kouwe et al. Sequence-Independent Segmentation of Magnetic Resonance Images; NeuroImage, 23 Suppl 1, 2004, S69-64

[Forstl96] H. Forstl, C. Besthorn, F. Hentschel et al. Frontal Lobe Degeneration and Alzheimer's Disease: A Controlled Study on Clinical Findings, Volumetric Brain Changes and Quantitative Electroencephalography Data; Dementia, 7, 1996, 27-34

[Frantz00] S. Frantz, K. Rohr, H.S. Stiehl. Localisation of 3D Anatomical Point Landmarks in 3D Tomographic Images Using Deformable Models. Proceedings of the MICCAI, 2000, 492-501

[Freeborough97] P.A. Freeborough, N.C. Fox. The Boundary Shift Integral: An Accurate and Robust Measure of Cerebral Volume Changes from Registered Repeat MRI; IEEE Transactions on Medical Imaging, 16, 1997, 623-629

[Frisoni99] G.B. Frisoni, M.P. Laakso, A. Beltramello et al. Hippocampal and Entorhinal Cortex Atrophy in Frontotemporal Dementia and Alzheimer's Disease; Neurology, 52, 1999, 91-100

[Gaser01] C. Gaser. Deformationsbasierte Morphometrie – eine Methode zur Analyse hirnstruktureller Veränderungen; PhD Thesis, University of Magdeburg, Institute of Natural Sciences, 2000

[Gaser04] C. Gaser, I. Nenadic, B.R. Buchsbaum et al. Ventricular Enlargement in Schizophrenia Related to Volume Reduction of the Thalamus, Striatum, and Superior Temporal Cortex; American Journal of Psychiatry, 161, 2004, 154-156

[Graham00] J.M. Graham, M.N. Paley, N. Hoggard et al. Magnetic Resonance Imaging of Brain Iron Deeposition in Patients with Idiopathic Parkinson's Disease Using the PRIME Sequence; Proceedings of the ISMRM, 2000, 1264

[Guimond01] A. Guimond, A. Roche, N. Ayache et al. Three-Dimensional Multimodal Brain Warping Using the Demons Algorithm and Adaptive Intensity Corrections; IEEE Transactions on Medical Imaging, 20, 2001, 58-69

[Hagemann99] A. Hagemann, K. Rohr, H.S. Stiehl et al. Biomechanical Modeling of the Human Head for Physically Based Nonrigid Image Registration; IEEE Transactions on Medical Imaging – Special Issue on Model-Based Analysis of Medical Imaging, 18, 1999, 875-884

[Haacke99] E.M. Haacke. Magnetic Resonance Imaging – Physical Principles and Sequence Design; Wiley, 1999

[Haacke04] E.M. Haacke, Y. Xu, Y.C.N. Cheng et al. Susceptibility Weighted Imaging (SWI); Magnetic Resonance in Medicine; 52, 2004, 612-618

[Hadjiprocopis05] A. Hadjiprocopis, W. Rashid, P.S. Tofts. Unbiased Segmentation of Diffusion-Weighted Magnetic Resonances Images of Brain Using Iterative Clustering; Magnetic Resonance Imaging, 23, 2005, 877-885

[Hahn00] H.K. Hahn, H.O. Peitgen. The Skull Stripping Problem in MRI Solved by a Single 3D Watershed Transform, Proceedings of the MICCAI, 2000, 134-143

[Hahn01] H.K. Hahn, M.G. Lentschig, B. Terwey et al. Clinical MRI Based Volumetry: The Cerebral Ventricles; Proceedings of the MICCAI, 2001, 1291-1292

[Hata00] N. Hata, A. Nabavi, W.M. Wells et al. Three-Dimensional Optical Flow Method for Measurement of Volumetric Brain Deformation from Intraoperative MR Images; Journal of Computer Assisted Tomography, 24, 2000, 531-538

[Head04] D. Head, R.L. Buckner, J.S. Shimony et al. Differential Vulnerability of Anterior White Matter in Nondemented Aging with Minimal Accerlation in Dementia of the Alzheimer Type: Evidence from Diffusion Tensor Imaging; Cerebral Cortex, 14, 2004, 410-423

[Holden00] M. Holden, D.L.G. Hill, E.R.E. Denton et al. Voxel Similarity Measures for 3D Serial MR Brain Image Registration; IEEE Transactions on Medical Imaging, 19, 2000, 94-102

[Horn81] B.K.P. Horn, B.G. Schunck. Determining Optical Flow; Artificial Intelligence, 18, 1981, 185-204

[Hufnagel05] H. Hufnagel, X. Pennec, G. Malandin et al. Non-Linear 2D and 3D Registration Using Block-Matching and B-Splines; Proceedings of the BVM, 2005, 325-329

[Jack97] C.R. Jack, R.C. Petersen, Y.C. Xu et al. Medial Temporal Atrophy on MRI in Normal Aging and Very Mild Alzheimer's Disease; Neurology, 49, 1997, 786-794

[Janke00] A.L. Janke, G. de Zubicaray, S.E. Rose et al. Characterising Cortical Changes in Alzheimer's Disease; Proceedings of the ISMRM, 2000, 580

[Janke01] A.L. Janke, G. de Zubicaray, S.E. Rose et al. 4D Deformation Modeling of Cortical Disease Progression in Alzheimer's Dementia; Magnetic Resonance in Medicine, 46, 2001, 661-666

[Jellus05] V. Jellus, B. Kiefer. Optimization of the Homomorphic Filter for Bias Field Correction; Proceedings of the ISMRM, 2005, 2248

[Jobst92] K.A. Jobst, A.D. Smith, C.S. Barker et al. Association of Atrophy of the Medial Temporal Lobe With Reduced Blood Flow in the Posterior Parietotemporal Cortex in Patients With a Clinical and Pathological Diagnosis of Alzheimer's Disease; Journal for Neurology, Neurosurgery and Psychiatry, 55, 1992, 190-194

[Jones05] D. Jones, C. Pierpaoli. Contribution of Cardiac Pulsation to Variability of Tractography Results; Proceedings of the ISMRM, 2005, 222

[Jovicich04a] J. Jovicich, D. Greve, E. Haley et al. Multi-Site Structural MRI Studies: An Evaluation of Image Distortions and Image Intensity Reproducibility; Proceedings of the ISMRM, 2004, 136

[Jovicich04b] J. Jovicich, E. Haley, D. Greve et al. Reliability in Multi-Site Structural MRI Studies: Effects of Gradient Non-Linearity Correction on Volume and Displacement of Brain Subcortical Structures; HBM, 2004

[Kantarci05] K. Kantarci, R.C. Petersen, B.F. Boeve et al. DWI Predicts Future Progression to Alzheimer Disease in Amnestic Mild Cognitive Impairment; Neurology, 64, 2005, 902-904

[Karas04] G.B. Karas, P. Scheltens, S.A. Rombouts, et al. Global and Local Gray Matter Loss in Mild Cognitive Impairment and Alzheimer's Disease; NeuroImage, 23, 2004, 708-716

[Kass88] M. Kass, A. Witkin, D. Terzopoulos. Snakes: Active Contour Models; International Journal of Computer Vision, 1988, 321-331

[Kitagaki98] H. Kitagaki, E. Mori, S. Yamaji et al. Frontotemporal Dementia and Alzheimer Disease: Evaluation of Cortical Atrophy With Automated Hemispheric Surface Display Generated With MR Images; Radiology, 208, 1998, 431-439

[Koethe06] U. Koethe, P. Stelldinger, H. Meine. Probably Correct Edgel Linking and Subpixel Boundary Reconstruction; Proceedings of the DAGM, 2006, 81-90

[Kouwe04] A.J. van der Kouwe, A.M. Dale. Rapid Real-Time Prospective Rigid Body Motion Correction During Imaging Using Clover-Leaf Navigators; Proceedings of the ISMRM, 2004, 95

[Leemput01] K. Van Leemput, F. Maes, D. Vandermeulen et al. Automated Segmentation of Multiple Sclerosis Lesions by Modeloutlier Detection; IEEE Transactions on Medical Imaging, 20, 2001, 677-688

[Lehtovirta00] M. Lehtovirta, M.P. Laakso, G.B. Frisoni et al. How Does the Apolipoprotein E Genotype Modulate the Brain in Ageing and in Alzheimer's Disease? A Review of Neuroimaging Studies; Neurobiololgy of Ageing, 21, 2000, 293-300

[Li99] T.Q. Li, A.M. Takahashi, T. Hindmarsh et al. ACD Mapping by Means of a Single-Shot Spiral MRI Technique with Application in Acute Cerebral Ischemia; Magnetic Resonance Imaging, 41, 1999, 143-147

[Little97] J.A. Little, D.L.G. Hill, D.J. Hawkes. Deformations Incorporating Rigid Structures; Computer Vision and Image Understanding, 66, 1997, 223-232

[Logothetis04] N.K. Logothetis, B.A. Wandell. Interpreting the BOLD Signal; Annual Review of Physiology, 66, 2004, 735-769

[Lucas81] B.D. Lucas, T. Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision; Proceedings of the International Joint Conference on Artificial Intelligence, 1981, 674-679

[MacDonald00] D. MacDonald, N. Kabani, D. Avis et al. Automated 3-D Extraction of Inner and Outer Surfaces of Cerebral Cortex from MRI; NeuroImage, 12, 2000, 340-356

[Maes97] F. Maes, A. Collignon, D. Vandermeulen et al. Multimodality Image Registration by Maximization of Mutual Information; IEEE Transactions on Medical Imaging, 16, 1997, 187-198

[Maguire00] E.A. Maguire, D.G. Gadian, I.S. Johnsrude et al. Navigation-Related Structural Change in the Hippocampi of Taxi Drivers; Proceedings of the NAS; 97, 2000**,** 4398-4403

[Maintz98] J. Maintz, M. Viergever. A Survey of Medical Image Registration; Medical Image Analysis, 2, 1998, 1-36

[Mallozzi04] R.P. Mallozzi, D.J. Blezek, C.P. Ward et al. Phantom-Based Geometric Distortion Correction for Volumetric Imaging of Alzheimer's Disease; Proceedings of the ISMRM, 2004, 1302

[Marr80] D. Marr, E. Hildreth. Theory of Edge Detection; Proceedings of the Royal Society of London, B, Biological Sciences, 1980, 187-217

[Martin05] M. Martin-Ferndandez, S. Bouix, L. Ungar et al. Two Methods for Validating Brain Tissue Classifiers; Proceedings of the MICCAI, (1), 2005, 515-522

[Menon99] R.S. Menon, S.G. Kim. Spatial and Temporal Limits in Cognitive Neuroimaging with fMRI; Trends in Cognitive Sciences, 3, 1999, 207-216

[Mesulam00] M.M. Mesulam. A Plasticity-Based Theory of the Pathogenesis of Alzheimer's Disease; Annals of the New York Academy of Sciences, 924, 2000, 42-52

[Minoshima97] S. Minoshima, B. Giordani, S. Berent et al. Metabolic Reduction in the Posterior Cingulate Cortex in Very Early Alzheimer's Disease; Annals of Neurology; 42, 1997, 85-94

[Mugler00] J.P. III. Mugler, B. Kiefer, J. Brookeman. Three-Dimensional T2-Weighted Imaging of the Brain Using Very Long Spin-Echo Trains; Proceedings of the ISMRM, 2000, 687

[NIA00] National Institute on Aging 1999-2000 Alzheimer's Report

[NIA02] National Institute on Aging 2001-2002 Alzheimer's Report

[Nitz99] W.R. Nitz. MR Imaging: Acronyms and Clinical Applications; European Radiology, 9, 1999, 979-997

[O'Brien97] J.T. O'Brien, P. Desmond, D. Ames et al. Temporal Lobe Magnetic Resonance Imaging Can Differentiate Alzheimer's Disease From Normal Ageing, Depression, Vascular Dementia and Other Causes of Cognitive Impairment; Psychological Medicine, 27, 1997, 1267-1275

[Ong05] W.Y. Ong, A. Farooqui. Iron, Neuroinflammation, and Alzheimer's Disease; Journal of Alzheimer's Disease, 8, 2005, 183-200

[Otsu79] N. Otsu. A Threshold Selection Method from Grey-Level Histograms; IEEE Transactions on System, Man, and Cybernetics, 9, 1979, 62-66

[Paviour06] D.C. Paviour, S.L. Price, M. Jahanshahi et al. Longitudinal MRI in Progressive Supranuclearpalsy and Multiple System Atrophy: Rates and Regions of Atrophy; Brain, 129, 2006, 1040-1049

[Periaswamy03] S. Periaswamy, H. Farid. Elastic Registration in the Presence of Intensity Variations; IEEE Transactions on Medical Imaging, 22, 2003, 865-874

[Perona90] P. Perona, J. Malik. Scale-Space and Edge Detection Using Anisotropic Diffusion, IEEE Transaction on Image Processing, 1990, 629-639

[Pluim03] J.P.W. Pluim, J.M. Fitzpatrick. Image Registration; IEEE Transactions on Medical Imaging, 22, 2003, 1341-1343

[Pohl02] K.M. Pohl, W.M. Wells, A. Guimond et al. Incorporating Non-Rigid Registration into Expectation Maximization Algorithm to Segment MR Images; Proceedings of the MICCAI, (1), 2002, 564-571

[Pohle01] R. Pohle, K.D. Tönnies. Segmentation of Medical Images Using Adaptive Region Growing; Proceedings of the SPIE, 4322, 2001, 1337-1346

[Porter04] D.A. Porter, E. Mueller. Multi-Shot Diffusion-Weighted EPI with Readout Mosaic Segmentation and 2D Navigator Correction; Proceedings of the ISMRM, 2004, 442

[Porter06] D.A. Porter, R.M. Heidemann. Multi-Shot Diffusion-Weighted EPI at 3T Using Readout-Segmented EPI and GRAPPA; Proceedings of the ISMRM, 2006, 1046

[Press93] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery. Numerical Recipes in C: The Art of Scientific Computing; Cambridge University Press, 2nd Edition, 1993

[Quinn05] B. Quinn, E. Fenstermacher, X. Han et al. Test-Retest Reliability Assessment for Longitudinal MRI Studies: A Comparison of the Effects of Different T1-weighted Protocols, Scanner Platforms, and Field Strengths on Semi-Automated Hippocampal Volume Measures; HBM, 2005

[Reiman96] E.M. Reiman, R.J. Caselli, L.S. Yun et al. Preclinical Evidence of Alzheimer's Disease in Persons Homozygous for the Epsilon 4 Allele for Apolipoprotein E; The New England Journal of Medicine, 334, 1996, 752-758

[Rohde03] G.K. Rohde. A. Aldroubi, B.M. Dawant. The Adaptive Bases Algorithm for Intensity-Based Nonrigid Image Registration; IEEE Transactions on Medical Imaging, 22, 2003, 1470-1479

[Roerdink01] J.B.T.M. Roerdink, A. Meijster. The Watershed Transform: Definitions, Algorithms and Parallelization Strategies; Fundamenta Informaticae, 41, 2001, 187-228

[Roesch00] P. Rösch, T. Netsch, M. Quist et al. Robust 3D Deformation Field Estimation by Template Propagation; Proceedings of the MICCAI, 2000, 656-663

[Roesch01] P. Rösch, T. Mohs, T. Netsch et al. Template Selection and Rejection for Robust Non-Rigid 3D Registration in the Presence of Large Deformations; Proceedings of the SPIE, 4322, 2001, 545-556

[Rohlfing01] T. Rohlfing, C.R. Maurer Jr. Intensity-Based Non-Rigid Registration Using Adaptive Multilevel Free-Form Deformation With an Incompressibility Contraint; Proceedings of the MICCAI, 2001, 111-119

[Rueckert99] D. Rueckert, L.I. Sonoda, C. Hayes et al. Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images; IEEE Transactions on Medical Imaging, 18, 1999, 712-721

[Sastre05] J. Sastre-Garriga, G.T. Ingle, D.T. Chard et al. Grey and White Matter Volume Changes in Early Primary Progressive Multiple Sclerosis: A Longitudinal Study; Brain, 128, 2005, 1454-1460

[Scahill02] R.I. Scahill, J.M. Schott, J.M. Stevens et al. Mapping the Evolution of Regional Atrophy in Alzheimer's Disease: Unbiased Analysis of Fluid-Registered Serial MRI; Proceedings of the NAS, 99, 2002, 4703-4707

[Sehgal05] V. Sehgal, Z. Delproposto, E.M. Haacke et al. Clinical Applications of Neuroimaging With Susceptibility-Weighted Imaging; Journal of Magnetic Resonance Imaging; 22, 2005, 439-450

[Shen02] D. Shen, S. Moffat, S.M. Resnick et al. Measuring Size and Shape of the Hippocampus in MR Images Using a Deformable Shape Model; NeuroImage, 15, 2002, 422-434

[Sijbers97] J. Sijbers, M. Verhoye, P. Scheunders et al. Watershed-Based Segmentation of 3D MR Data for Volume Quantification; Magnetic Resonance Imaging, 15, 1997, 679-688

[Sled98a] J.G. Sled, G.B. Pike. Understanding Intensity Non-uniformity in MRI; Proceedings of the MICCAI, 1998, 614-622

[Sled98b] J.G. Sled, P. Zijdenbos, A.C. Evans. A Nonparametric Method for Automatic Correction of Intensity Nonuniformity in MRI Data; IEEE Transactions on Medical Imaging, 17, 1998, 87-97

[Smith99] C.D. Smith, A.H. Andersen, R.J. Kryscio et al. Altered Brain Activation in Cognitively Intact Individuals at High Risk for Alzheimer's Disease; Neurology, 53, 1999, 1391-1396

[Smith02a] A.D. Smith. Imaging the Progression of Alzheimer Pathology Through the Brain; Proceedings of the NAS, 99, 2002, 4135-4137

[Smith02b] S. Smith. Fast Robust Automated Brain Extraction; Human Brain Mapping, 17, 2002, 143-155

[Spies02] H. Spies, B. Jaehne, J.L. Barron. Range Flow Estimation; Computer Vision and Image Understanding, 85, 2002, 209-231

[Starck02] G. Starck, M. Borga, M. Friberg et al. Fully Automatic Segmentation of the Hippocampus in MR Images; Proceedings of ESMRMB, 2002

[Styner00] M. Styner, C. Brechbuehler, G. Szekely et al. Parametric Estimate of Intensity Inhomogeneities Applied to MRI; IEEE Transactions on Medical Imaging, 19, 2000, 153-165

[Studholme01] C. Studholme, V.A. Cardenas, M.W. Weiner. Multiscale Image and Multiscale Deformation of Brain Anatomy for Building Average Brain Atlases; Proceedings of the SPIE, 4322, 2001, 557-568

[Takahashi02] S. Takahashi, H. Yonezawa, J. Takahashi et al. Selective Reduction of Diffusion Anisotropy in White Matter of Alzheimer Disease Brains Measured by 3.0 Tesla Magnetic Resonance Imaging; Neuroscience Letters, 322, 2002, 45-48

[Thacker02] N.A. Thacker, A.R. Varma, D. Bathgate et al. Dementing Disorders: Volumetric Measurement of Cerebrospinal Fluid to Distinguís Normal from Pathologic Findings – Feasibility Study"; Radiology, 224, 2002, 278-285

[Thirion96] J.-P. Thirion. New Feature Points Based on Geometric Invariants for 3D Image Registration. International Journal of Computer Vision, 18, 1996, 121-137

[Thompson01a] P.M. Thompson, A.W. Toga. Cortical Pattern Matching; in Handbook of Medical Imaging (Editor I. N. Bankman), 2000, 581-587

[Thompson01b] P.M. Thompson, M.S. Mega, R.P. Woods et al. Cortical Change in Alzheimer's Disease Detected with a Disease-Specific Population-Based Brain Atlas; Cerebral Cortex; 11, 2001, 1-16

[Thompson02] P.M. Thompson, A.W. Toga. A Framework for Computational Anatomy; Computing and Visualization in Science, 5, 2002, 13-34

[Thompson03] P.M. Thompson, K.M. Hayashi, G. de Zubicaray et al. Dynamics of Gray Matter Loss in Alzheimer's Disease; The Journal of Neuroscience, 23, 2003, 994-1005

[Wang04a] D. Wang, W. Strugnell, G. Cowin et al. Geometric Distortion in Clinical MRI Systems Part I: Evaluation Using a 3D Phantom; Magnetic Resonance Imaging, 22, 2004, 1211-1221

[Wang04b] D. Wang, W. Strugnell, G. Cowin et al. Geometric Distortion in Clinical MRI Systems Part II: Correction Using a 3D Phantom; Magnetic Resonance Imaging, 22, 2004, 1223-1232

[Wells96] W.M. Wells III, P. Viola, H. Atsumi et al. Multi-Modal Volume Registration by Maximization of Mutual Information; Medical Image Analysis, 1, 1996, 35-51

[Woerz03] S. Woerz, K. Rohr. Localization of Anatomical Point Landmarks in 3D Medical Images by Fitting 3D Parametric Intensity Models; Proceedings of the BVM, 2003, 6-10

[Wu02] Y. Wu, K. Pohl, S.K. Warfield et al. Automated Segmentation of Cerebral Ventricular Compartments; Proceedings of the ISMRM, 2003

[Zijdenbos02] A.P. Zijdenbos, R. Forghani, A.C. Evans. Automatic "Pipeline" Analysis of 3-D MRI Data for Clinical Trials: Application to Multiple Sclerosis; IEEE Transactions on Medical Imaging, 21, 2002, 1280-1291