Computational tools for objective assessment in Neuroimaging

by

Florian Weiler

A Thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Computer Science.

Approved Dissertation Committee:

Prof. Dr. Horst K. Hahn (Dissertation Advisor)
Jacobs University Bremen / Fraunhofer MEVIS

Prof. Dr. Lars Linsen
University of Münster / Jacobs University Bremen

Prof. Dr. Bernhard Preim
Otto-von-Guericke-Universität Magdeburg

Dr. Jan Klein
Fraunhofer MEVIS

Date of defense: January 29, 2019

Computer Science & Electrical Engineering
Statutory Declaration

<table>
<thead>
<tr>
<th>Family Name, Given/First Name</th>
<th>Weiler, Florian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matriculation number</td>
<td>20332106</td>
</tr>
<tr>
<td>Kind of thesis</td>
<td>PhD-Thesis</td>
</tr>
</tbody>
</table>

**English: Declaration of Authorship**

I hereby declare that the thesis submitted was created and written solely by myself without any external support. Any sources, direct or indirect, are marked as such. I am aware of the fact that the contents of the thesis in digital form may be revised with regard to usage of unauthorized aid as well as whether the whole or parts of it may be identified as plagiarism. I do agree my work to be entered into a database for it to be compared with existing sources, where it will remain in order to enable further comparisons with future theses. This does not grant any rights of reproduction and usage, however.

This document was neither presented to any other examination board nor has it been published.

**German: Erklärung der Autorenschaft (Urheberschaft)**


Diese Arbeit wurde noch keiner anderen Prüfungsbehörde vorgelegt noch wurde sie bisher veröffentlicht.

Bremen, 6. 7. 2020

Date, Signature
Abstract

MEDICAL IMAGING is nowadays capable of non-invasively displaying the human brain in manifold ways. A variety of different imaging techniques and modalities allow for displaying a broad spectrum of different properties of the brain. This includes both structural imaging, allowing for identification of anatomical features, changes, and variations, as well as functional imaging, allowing for characterization of aspects related to how the brain works, including physiologic or metabolic processes. In parts, these imaging techniques have entered routine clinical imaging, while others are almost exclusively utilized in research settings.

Computer assisted analysis of such data has become a highly active area of research. The past 20 years have seen tremendous efforts and advances in this field. Today, there exists a wealth of published methods addressing a broad spectrum of medical questions. Considering the maturity to which this field of research has grown by now raises expectations that most basic problems should actually be solved and even many of the more complex task should at least be manageable. However, looking at the role of image-based quantification in everyday clinical practice, draws a completely different picture. The de-facto standard approach in diagnostic radiology is a purely qualitative reading through human experts. Quantification, if used at all, is typically limited to simple measurements on single image-slices. Without debate, there exists a huge gap between what would be potentially possible as defined by the scientific state-of-the-art, and clinical reality.

This PhD thesis addresses this situation. It studies three general concepts aiming at objective image assessment: 1) quantification; 2) interactive segmentation; 3) interactive data-visualization. For each of these concepts, an exemplary application is chosen, and a novel method is proposed with a focus on fulfilling requirements that, if not met, would prevent integration into clinical workflows. First, I present a method for robust assessment of upper-spinal cord atrophy, a parameter which has been successfully correlated to several clinical markers in the context of multiple sclerosis. Second, I present two novel interactive tools for segmenting individual gyri of the brain and regions-of-interest in DTI data, both of which could build a foundation for integrating other, more advanced tools such as DTI fiber-tracking or cortical thickness measurements into clinical routine more smoothly. Finally, I present an example of how interactive 3D visualizations combined with efficient tools for exploration of image data can support the planning process of complex
neurosurgical interventions. This is demonstrated on the challenging task of pre-
surgical treatment planning for cerebral arteriovenous malformations (AVMs).
Acknowledgements

This thesis would not have been possible without the support of many people. At this point, I would like to express my thanks to those. First and foremost, I have to thank my supervisor Prof. Dr. Horst K. Hahn. He introduced me to the field of image processing and visualization in the context of neuroimaging. He accompanied my work from the very beginning at MeVis Research, always supportive, and always critical with a seemingly unlimited source of inspiring ideas. Discussions with him have always been encouraging and motivating.

Furthermore, I wish to thank the additional committee members Prof. Dr. Lars Linsen, Prof. Dr. Bernhard Preim, and Dr. Jan Klein for their input and comments to this work. Especially to Jan I need to express my gratitude for his patience and energy to keep pushing me forward repeatedly.

Parts of this work were conducted in the context of the German BMBF-funded project MS-Atrophie (FZ: 13GW0153C). I wish to thank all collaborators and partners involved in this project for their input and support, namely my colleagues at MEVIS Stefan Heldmann and Jan Klein, the team at Jung Diagnostics, and the team at mediri.

This work has been conducted at Fraunhofer MEVIS, and the enjoyable and inspiring work environment there has shaped this work fundamentally. This environment is created by every person working there, and I have to thank all of them. Special thanks go to a number of people from whom I have learned a lot during my years at MEVIS, namely Christoph Brachmann, Benjamin Geisler, Jochen Hirsch, Alexander Köhn, Jan-Martin Kuhnigk, Christian Rieder and Christian Schumann. All of you have in one way or another influenced this work.

Finally, my deepest thanks go to all of my family. To my parents, who always supported me on my path through life. To my wife Anja, for believing in me and keeping my head free during stressful and tiring times. And to my children Emily, Simon and Erik, for showing me every day, what really counts in life.
# Table of Contents

Abstract ................................................................. i  
Acknowledgements .................................................... iii  
Table of Contents ..................................................... iv  

1 Introduction ......................................................... 1  

2 Quantification of spinal cord atrophy in Multiple Sclerosis ........ 5  
2.1 Clinical background: Multiple Sclerosis ........................ 7  
2.2 Related Work ........................................................ 12  
2.3 Automated localization, segmentation and quantification of the cervical spinal cord ............................................... 18  
2.3.1 Template-based localization .................................... 18  
2.3.2 Heuristic watershed-based segmentation of the spinal cord ... 32  
2.3.3 Model-based quantification of spinal cord cross-sectional area 34  
2.4 Evaluation and Results ............................................. 39  
2.5 Discussion and Conclusion .......................................... 46  

3 Interactive tools for identification and delineation of brain structures 51  
3.1 Introduction .......................................................... 51  
3.2 2D-ROI generation from DTI data ................................. 52  
3.2.1 Motivation ....................................................... 53  
3.2.2 Method ........................................................... 55  
3.2.3 Evaluation and Results .......................................... 59  
3.2.4 Discussion ....................................................... 62
# Table of Contents

3.3 3D Gyrus segmentation ................................................. 64  
3.3.1 Related work ....................................................... 65  
3.3.2 Methods ............................................................. 66  
3.3.3 Results .............................................................. 70  
3.3.4 Discussion ........................................................... 73  
3.4 Conclusion .............................................................. 73  

4 Interactive Multi-Volume Visualization for Planning of Cerebral AVM Surgery ............................................. 77  
4.1 Introduction ............................................................. 77  
4.2 Medical background ................................................... 78  
4.3 Imaging Background ................................................... 80  
4.4 Related Work ........................................................... 82  
4.5 Preprocessing and Data Analysis .................................... 84  
4.5.1 Multimodal Registration .......................................... 86  
4.6 Volume Visualization for Surgery Planning ......................... 89  
4.7 Results and Discussion ............................................... 93  
4.7.1 Case 1 ............................................................... 93  
4.7.2 Case 2 ............................................................... 95  
4.7.3 Case 3 ............................................................... 97  
4.8 Discussion .............................................................. 98  
4.9 Conclusions ............................................................ 99  

5 Discussion and Conclusion ............................................. 101  
Bibliography ................................................................. 107
1 Introduction

“Begin at the beginning,” the King said, gravely, “and go on till you come to an end; then stop.”

(Lewis Carroll, Alice in Wonderland)

NEUROIMAGING describes the processes and techniques used for producing images of the human central nervous system. While not a distinct medical discipline in itself, it plays a central role in all those fields dealing with the brain or the nervous system, more specifically in neuroradiology, neurology and neurosurgery. It also provides elemental tools for other related areas of research, most prominently for neuroscience, but also for psychology or neurobiology.

The most relevant imaging techniques used in neuroimaging are magnetic resonance imaging (MRI), computed tomography (CT) ultrasound (US), and positron emission tomography (PET). In medicine, each of these modalities plays a different role, depending on the underlying clinical question. CT for example, is the modality
of choice in diagnosis and treatment decision for stroke, given its speed and good capabilities of displaying intra-cranial hemorrhage. PET is a vital tool in neurology with applications in the diagnosis of Parkinson’s or Alzheimer’s disease. MRI plays a special role for two reasons: First, it is capable of displaying many different properties of the brain through different MR-sequences. Second, there are no known risks or side effects associated with MRI, making it a powerful tool for repeated or extensive examinations in both patients or volunteers. Compared to MRI or CT, Ultrasound is a technically simpler modality that in consequence is also a lot cheaper, and often more easily applicable. It can however not compete with MRI and CT in terms of imaging details.

In all its applications, neuroimaging is merely a means to a specific end. In medical applications, its purpose is to support the process of differential diagnostics, either confirming or ruling out a suspected medical finding. Also, it can provide means for monitoring the evolution of a disease, both with or without therapy. Imaging can also be the decisive factor in selecting an optimal treatment for an individual patient. In neurosciences, it provides the source for proving or rejecting hypothesis on how specific processes in the brain act, for how structures in the brain work and interplay with each other, or ultimately for unveiling the many remaining mysteries behind what makes the brain as a whole act and function as it does.

The primary imaging modalities listed above have in the last four decades evolved to a state where each of them is capable of providing a much higher level of information than what could be grasped and interpreted qualitatively be means of visual reading of the images. Instead, a key to information extraction on such data lies in analyzing the data by means of computer assisted data analysis, visualization or quantification. The latter especially being a task, which computers handle by far more robust and reliable than humans could attempt to. However, there is no single mechanism for quantification of brain images. Obviously, the tools required will inevitably depend on both the question to be answered – or rather the property to be measured – and the underlying data this measurement shall be performed on.

Studying scientific literature of the last two decades reveals a wealth of publications focusing on the analysis of brain images. Topics addressed cover a broad spectrum from segmentation of anatomic structures (e.g. the brain, internal brain-structures, brain-vasculature, tumors or other kinds of lesions), quantification of relevant properties of such structures (i.e. volumes, lesion diameters, lesion counts, ...), monitoring of changes (i.e. for measuring therapy response) and many more. Often, studies are performed for a specific question to be answered on a specific type of data. Looking from afar, one could expect that many of the published tools have entered clinical routine, or have at least been embedded within software workstations in radiology departments to be available for selected cases. Reality
however draws a different picture. Clinical workflows in neuroradiology rely to a large extent on subjective visual assessment of black-and-white images through human readers.

There are two primary reasons for this situation:

- **Robustness of results** – Evaluations used for scientific publications are often performed on either small datasets, or on selected cohorts stemming from clinical studies, which do not represent everyday clinical data well. In consequence, the performance of many methods when applied to non-ideal, often widely heterogeneous clinical data can be rather poor. For clinical routine use however, robustness of results is a vital criterion.

- **Integration into workflows** – Clinical workflows in radiology are highly optimized in order to be able to generate a high throughput of cases. Time is the most critical factor. As a consequence, any tool utilized in this process needs to integrate tightly into these workflows. Any second in time consumed by an additional tool needs to be outweighed by the added value of the results. The critical component here is the time required by the operator. Cost intensive calculations are not prohibited per se, however they must not block the clinician from other tasks, but rather run automated in the background.

Any algorithmic component that aims at successful utilization in this field must in consequence suffice to those two criteria.

**Contributions of this thesis**

This thesis aims to contribute to this field by studying a selection of tasks motivated by clinically relevant, medical questions. For each of these tasks, I present a pragmatic solution that advances the possibilities of information extraction from medical images and gain of diagnostic insights. More precisely, this thesis addresses three aspects:

1. **Quantification of a segmented structure** – Direct quantification of properties of individual brain structures is a basic requirement that finds application in numerous scenarios, depending on the actual structure and the underlying clinical question. In this thesis, the quantification of the upper cervical spinal cord with respect to volumetry and calculation of mean cross-sectional areas at selected vertebral levels is studied. These measurements are highly relevant in staging and differentiating between different phenotypes of multiple sclerosis (MS), as well as monitoring progression of the disease.
2. **Interactive segmentation of brain structures** – There is a natural conflict between the strict requirement for robustness and the desire of methodological research for automation. Especially for segmentation, there are few entities that can robustly be segmented fully automatically. Typically, results degrade further when applied to data that deviates from the assumptions made during development of an algorithm – a prerequisite that will be rather common in clinical routine. In contrast, efficient, semi-automatic algorithms that incorporate knowledge from the user can in some cases create better results in less time. The critical point is how the interaction is realized. This question is studied with respect to segmentation of DTI regions-of-interest and the segmentation of cortical gyral structures of the brain.

3. **Interactive augmentation of data by means of visualization** – Visualization can be a powerful means to transform and augment information. Essentially, it can offer different views on the same data, thereby highlighting or attenuating parts of the information. Interactive visualization techniques can again incorporate the user and let him control how the information is to be presented. In the thesis, I present an application of multi-modal volume rendering applied to multi-spectral MRI data in the context of neurosurgical intervention planning.

The remainder of this thesis will dedicate one chapter to each of the topics above. A final concluding chapter will summarize the results and re-evaluate them with respect to the overall task addressed: Advancing the capabilities for objective assessment in neuroimaging.

Parts of this thesis have been previously published and presented at scientific conferences. At the end of each chapter, the corresponding publications are listed.
2 Quantification of spinal cord atrophy in Multiple Sclerosis

Life is trying things to see if they work.

(Ray Bradbury)

Measuring the size of anatomical structures of the human body is an elemental task required for many different processes in the field of medical image analysis. A prominent use-case would be the monitoring of the size of a tumor and its response to therapy. Another application is the detection and monitoring of atrophy. Atrophy describes the reduction in size of an organ, a part of an organ, or individual cells, that may be both of physiological or pathological origin. Regardless of the actual application, it may be equally important to obtain the absolute value of an individual measurement at a given timepoint as to assess the relative change of the measured property over time.
Depending on the dimension in which the measurement is taken, the size can relate to a variety of geometrical properties of a structure, such as the *length* for one-dimensional objects, *diameter*, *circumference* or *area* for two-dimensional measurements, or *volume* or *surface-area* for three-dimensional measurements.

Automation of these types of tasks is desirable primarily for two reasons: Efficiency, and accuracy. First, the task of measuring is time-consuming, and potentially scales exponentially with the number of dimensions involved. Specialized tools may cushion this aspect to some extent, nonetheless the relevance of processing time for a case must not be neglected, especially in the context of tight schedules in clinical radiology. The second reason may be considered of even bigger importance. Any measurement taken will inevitably be affected by a measurement error. This error in the context of human-made manual measures is evaluated as inter-rater variability, and is affected by many parameters, such as the users experience, the setting in which the user operates, the users current level of attention, but also their personal preferences and judgments. For a good part, these factors are completely unrelated to the measuring-task or the given data, but rather relate to the observer exclusively.

In contrast, the error produced by a computer algorithm will depend exclusively on two influencing factors: The implementation of the algorithm, and the nature of the given input data. In consequence, a well-trained computer algorithm will be able to reproducibly generate measures of constant accuracy given a constant quality of the underlying data itself.

In order to achieve full automation, three processing steps should be separated:

1. **Localization** – the process of localizing the structure to be measured within the data.

2. **Segmentation** – the process of separating a structure from surrounding objects.

3. **Quantification** – the process of deriving the required measurements from the segmented structure with the required level of precision.

The above given criteria are formulated on a general level that holds true for arbitrary measuring tasks in medical image analysis. In practice, requirements, accompanying constraints, and especially solutions to the individual steps involved will vary extensively. A general framework for solving these tasks is beyond the scope of this work. Instead, an exemplary solution to a selected measurement task from the field of neuroimaging is presented. In concrete terms, a fully automated method for localization, segmentation and volumetric quantification of the upper cervical spinal cord is given. A central clinical application for such a measurement is the detection and progress monitoring of long-term spinal cord atrophy in the context of Multiple Sclerosis (MS).
The chapter is organized as follows: First, a brief introduction of the medical background is given, motivating the clinical relevance and applicability of this method. Subsequently, an overview of related approaches is given. This is followed by a detailed description of the proposed method, individually addressing the three processing steps involved. Results are presented by means of an analysis of accuracy, as well as through a discussion of clinical studies and findings involving the method. A discussion of limitations and possible extensions concludes this chapter.

2.1 Clinical background: Multiple Sclerosis

Multiple sclerosis (MS) is the most common inflammatory disease of the central nervous system in the northern hemisphere. It is incurable, leads to an increasing neurological disability over time, and is one of the most frequent causes of premature retirement among young adults. It is one of the epidemiologically most relevant neurological diseases in Western Europe and North America. According to WHO estimates, 2.3 million people world wide are affected by the disease [11, 63, 96].

Clinical symptoms cover a broad spectrum of neurological conditions. Visual impairment is often an early symptom, as inflammation of the optic nerves (optic neuritis) often occurs as a precursor during the very early stages of the disease. Motor- and sensory issues are also a typical early symptom, which occur in conjunction with lesions in the spinal cord. Other symptoms include coordination and balance problems, emotional changes and depression, speech problems, muscle tremors, and difficulties to focus and stay concentrated, among others. Another typical symptom that is primarily associated with MS is fatigue, a feeling that patients describe as an overwhelming tiredness and exhaustion, independent of the actual state of rest or physical exertion.

Neurologists differentiate two phenotypes of MS: Relapsing-Remitting MS (RRMS) and Progressive MS (PMS). An older sub-differentiation of the progressive disease into primary and secondary progression (PPMS and SPMS) has been dismissed within the 2013 revision of MS characterization criteria [81, 82]. It is however still frequently found in publications. Progressive MS describes a continuously evolving progression of the disease accompanied by ongoing worsening of symptoms over a period of many years. In contrast, relapsing-remitting MS is characterized by short attacks of symptomatic disease activity that usually last between several hours to a few days, followed by partial or even complete recovery. These relapses are then followed by long periods of complete absence of symptoms, that last between a few month up to several years. Practically, a clear differentiation between these phenotypes is difficult, as both types can be mixed. Often RRMS will convert into PMS after several years. Also, a constant, progressive worsening may go along the RR phase.
Chapter 2. Quantification of spinal cord atrophy in Multiple Sclerosis

Figure 2.1: Typical temporal courses of MS: (a) Progressive MS: steady increase in disability without attacks. (b) Relapsing-remitting MS: Unpredictable attacks followed by periods of remission. (c) Relapsing-remitting MS, converting into progressive MS. (d) Progressive-relapsing MS: Steady decline since onset with superimposed attacks.

Clinical diagnosis of MS is nowadays made on the basis of the so-called McDonald Criteria. These formulate a set of requirements suited for an early diagnosis with a high degree of both specificity and sensitivity. They were initially defined in 2001 by the International Panel on MS Diagnosis [86] and later refined in 2005, 2010 and 2018 in order to incorporate current findings [102, 103, 125]. At its core, they rely on proving dissemination of the disease in space and time [116], in addition to exclusion of other disorders potentially mimicking MS. Besides clinical markers, MRI findings have become a crucial part of these criteria. In 1997, Barkhof et al. presented the first guidelines for differential diagnostics of MRI findings related to MS (the Barkhof criteria, [1]). Those, along with their later revisions [33, 34, 93] formulate requirements to the presence, size and location of focal white matter lesions, but also to cortical and spinal cord lesions. Figure 2.2 shows typical MRI findings of cerebral lesions in MS.

Several different diagnostic scales exist for staging the severity of the disease. The primary scale used for assessment is the Expanded Disability Status Scale (EDSS) introduced by Kurtzke in 1983 [72]. It provides a measure for the disease severity
2.1. Clinical background: Multiple Sclerosis

Figure 2.2: Typical MRI findings of MS: (a) and (b): Inflammatory white-matter MS lesions in Proton-density (PD)- and T2-weighted images. (c): Brain atrophy, as indicated by enlarged ventricles and prominently exposed sulcal spaces in T$_1$-weighted MRI.

on a scale from 0 to 10, with 0 meaning "normal neurologic exam" and 10 meaning "death due to MS". Evaluation of the EDSS score is based on a series of neurological tests, that address a total of eight different functional systems. Each functional system is graded individually. The functional systems considered are: Pyramidal, Cerebellar, Brain Stem, Sensory, Bowel & Bladder, Visual, Cerebral, and Other. The EDSS scale then defines rules on how the individual scores contribute to a total disability score.

The EDSS puts a focus on motor functions, while considering cognitive functions to a smaller extent. Therefore, an alternative scale has been proposed by the National MS Society Clinical Outcomes Assessment Task Force in 1999 [36]: The Multiple Sclerosis Functional Composite measure (MSFC). The MSFC individually covers three dimensions of clinical impairment in MS, namely leg function, arm/hand function, and cognitive function. Each dimension is assessed with an individual test: The Timed 25-Foot Walk (TWT) test for leg function. The Nine Hole Peg Test (9-HPT) for arm/hand function. And the Paced Auditory Serial Addition Test (PASAT) for cognitive function.

These scales are important not only in the staging of individual patients, but also for monitoring temporal changes or therapeutic responses in clinical trials. The correlation of objectively measurable biomarkers to these clinical scores is a fundamental step in the process of gaining a better understanding of the underlying physiological process of the disease. On the imaging side, the most frequently used biomarkers are the number, location and size of whiter-matter lesions and the global and local volume of the brains gray and white matter. More recently, spinal
Pathologically, MS consists of two primary components: Inflammatory lesions and atrophy. Inflammation occurs as an effect of an auto-immune reaction against healthy brain tissue. The cause of this reaction is up to this date not fully understood. Its effects however are well studied. Inflammation causes a damage to the myelin-sheaths responsible for protecting the white-matter axonal cells. It leads to a continuous process of de- and remyelination, which in the long run causes irreparable damage to the white matter axons and consequently cell loss. These inflammations typically manifest as lesions of varying number and size within the white-matter of both the cerebrum, cerebellum and the spinal cord [1, 84, 106, 141]. In later stages of the disease, cortical involvement and gray-matter lesions may also be observed [41, 62].

In addition to the inflammatory component, there is also a degenerative disease component that leads to the destruction of both cerebral and spinal cord tissues, so-called atrophy. Atrophy describes the general process of tissue loss. Brain atrophy is part of the natural physiological process of aging. It can be measured as brain volume loss per year (BVL/year). It is believed to begin in the early ages of adolescence with a very low yearly rate of 0.2% at the age of 35, and increases non-linearly to an average rate of 0.5% at the age of 70 years [115]. In healthy aging, brain atrophy affects both the gray- and white-matter of the brain comparatively, with gray-matter atrophy progressing slightly faster. In MS patients, the age-dependency of atrophy rates is absent. In a longitudinal study over a period of 13 years, De Stefano et al. examined brain volume loss in a cohort of 206 MS patients and compared that to 35 healthy controls. They reported an age-independent BVL/year of 0.51% on the patient cohort [27]. In consequence, the accumulated total brain volume loss in MS patients significantly increases with ongoing disease duration.
In addition to brain atrophy, atrophy of the spinal cord in MS has been studied separately during the last decades. Pathologic involvement of the spinal cord has been known since at least the 1970ies [98]. However, the amount of focus put to spinal cord atrophy in comparison to lesion load and brain atrophy has been surprisingly low. While brain atrophy is accepted as a validated outcome measure in clinical phase-3 trials, and has also been proposed to be used for assessment of individual patients, the same does not apply for spinal cord atrophy. This is all the more regrettable since recent studies have consistently shown that the degenerative changes are the most important morphological correlate of long-term disability.

Spinal cord atrophy is typically assessed by determining the mean cross-sectional area (MCSA) at a given vertebral level. The spinal column is anatomically divided into three sections: The cervical-, thoracic-, and the lumbar spine. Within each section, vertebral bodies are numbered in the head-to-feet direction and prefixed with their respective section. This gives vertebral bodies C1-C7 for the cervical spine, T1-T12 for the thoracic spine, and L1-L5 forming the lumbar spine. In MS studies, the focus is mostly put to the upper cervical spinal cord, beginning with the top of C1 and continuing down to somewhere between C4 to C6. In the literature, in addition to MCSA a range of abbreviations for this measure are found and often used synonymously: Cervical cord Area (CCA), Cervical spinal cord Area (CSA), Upper cervical cord area (UCCA), Mean Upper Cervical Cord Area (MUCCA) are the commonly used names. For clarity reasons, this work will use the term MCSA throughout the following sections.

Very recently, Casserly et al. presented results from an extensive meta-analysis of 94 studies focusing on spinal cord atrophy in MS in the time period of 1993-2017 [17]. From those, they report pooled results of mean cervical spinal cord area of 69.72 mm$^2$ for progressive MS, 78.88 mm$^2$ for relapsing-remitting MS, and 80.87 mm$^2$ for healthy controls. In addition, they report a pooled annual atrophy rate of 1.78% from longitudinal studies, which is more than three times higher than that of brain atrophy in MS. They conclude that there is clear evidence for spinal cord atrophy in all types of MS, and that it is highly relevant to clinical disability in the disease. They further recommend an inclusion of MCSA as an outcome measure in clinical trials. Figure 2.3 shows two samples of spinal cord cross-sections, both for a healthy case, and for an MS-patient affected by spinal cord atrophy.

Routine MRI imaging is capable of visualizing lesions in the brain and spinal cord. It also plays an important role in the therapy monitoring with regard to both efficacy and safety-relevant aspects. Despite the fact that during the last twenty years a tremendous amount of scientific work has been invested into developing computer-assisted tools for lesion detection, counting, segmentation and quantification, the current clinical practice still relies heavily on purely manual, qualitative reading of MRI images in everyday routine. With respect to detection
Chapter 2. Quantification of spinal cord atrophy in Multiple Sclerosis

and counting of white matter lesions, the primary challenge lies in the required time and effort for the radiologist. The tasks themselves can be performed reliably and often more robust by a human reader, than they could be by current computer solutions. In contrast, precise manual measurement of atrophy is difficult and typically not feasible due to the high amount of effort paired with a low degree of achievable accuracy. Here, the need for computational tools becomes crucially important in order to successfully integrate atrophy-related measures into the process of diagnosis and disease-monitoring in MS.

The particular challenge in obtaining precise volumetric measurements especially in small objects is rooted in the nature of partial volume averaging as an inevitable effect of the digital imaging process with a limited spacial resolution. The partial volume effect describes the fact, that a single element within a digital image can only encode a single intensity value. In case of a single type of tissue covered by this element, this intensity value may directly relate to that tissue. However, when two or more different tissue-types are covered by the same element, this intensity level will equal a mixture of these different tissues. The task of reconstructing the individual contributions of each tissue class is called partial volume modeling. It is of crucial importance in order to obtain a precise measurement. Recall that the average value of MCSA in the upper cervical cord in healthy people varies roughly in the range of $0.7 - 0.9 \text{cm}^2$. Given the typical resolution of current MR imaging techniques of $1 \text{mm}^3$, this corresponds to an amount of 70-90 voxels covered by the spinal cord in a single cross-section. Considering further, that the amount of partial volume (PV) voxels for objects of this size is in the order of 30-50% (assuming an approximate circular cross-sectional shape), it becomes readily apparent, that precise and accurate estimation of the MCSA parameter is an extremely challenging task. Moderate changes caused by atrophy over time will almost exclusively manifest through changes in the distribution of PV voxels.

2.2 Related Work

Quantification of spinal cord atrophy has been studied and discussed in the literature before. An extensive overview of existing methods has been presented by DeLeener et al. in [26]. A selection of relevant publications is summarized here. These methods can be roughly categorized into three groups: Intensity-based, surface-based, and image-based techniques.

Local Intensity based methods

The first published method that focused on robust and reproducible quantification of spinal cord area is the work from Losseff et al. [80]. In their work, they identified
the partial volume effect (PVE) as the primary reason for poor reproducibility of earlier, manual measurements, especially in scan-rescan studies. In consequence, the authors proposed a manual process for measuring mean signal intensities for both pure spinal cord tissue and surrounding CSF structures. With these intensity values, they perform a threshold segmentation of the spinal cord using the center between mean tissue- and mean CSF-intensity as a threshold level. They evaluated their technique using two sample scan-rescan data from 15 subjects, for which they achieved a coefficient of variation of 0.79% and 1.61% for experienced and inexperienced raters respectively. Also, they demonstrated a significant correlation between measured cord area and EDSS scores on a collective of 60 MS patients.

ElMendili et al. [30] recently presented an extension of this method. They analyze T2-weighted images, on which a sequence of image processing operations including Otsu-thresholding [99], connected-components analysis and region-growing is performed. From this pipeline, an optimal threshold is automatically derived and quantification of cervical cord area (CCA) is performed equally to the method by Losseff et al. They evaluated their method on different subgroups created from a total dataset of 112 images (82 volunteers, 10 patients with amyotrophic lateral sclerosis (ALS), 10 patients with spinal muscular atrophy (SMA) and 10 patients with spinal cord injury (SCI). By comparing their results with the active surface method from Horsfield and colleagues [55] and the semi-automatic thresholding approach by Losseff [80], they were able to outperform both techniques. No scan-rescan analysis was performed.

An alternative approach was presented by Tench et al. [124]. The authors proposed a semi-automatic approach combining a Sobel-filter for edge detection with an operator initiated region-growing algorithm for segmentation. Identification of the spinal cord is done manual by the operator by placing a seed point within the cord, which is then used to initialize region-growing. They estimate partial volumes by analyzing intensity values around the border of the segmented cord. From these, they derive for each border-voxel the amount of cord tissue and CSF, by which the total cord area is then corrected. They evaluated their method on both phantom data, and MR images of 10 healthy volunteers.

Zivadinov et al. [146] presented a similar approach. They essentially combined a Sobel-edge detector with Gaussian smoothing, and a Canny-like edge thinning [15] using maximum suppression technique. Identification of the spinal cord is done manually by the user on each slice covered by the measured area. Unlike other authors, they did not consider handling of partial-volume effects in any way. They also extended their method to 3D and compared how cervical cord area relates to cervical cord absolute volume (CCAV). Further, they applied their method to both T1- and T2-weighted MRI data. Evaluation was performed on a collective of 66 MS patients and 19 controls. They were able to demonstrate a strong correlation
of CCAV with EDSS scores, most prominently in subgroups of SP- and PP-MS patients.

A slightly different technique has been published by Behrens [4]. They proposed a general framework for segmentation of tubular structures from medical images using a randomized Hough transform (RHT). Among other applications, they also demonstrated their method on a single MR image of the head where the spinal cord could be successfully segmented. However, they did not report results on volumetric accuracy of this approach.

**Surface-based methods**

In contrast to intensity based methods, surface based methods aim at reconstructing a surface model of a structure. Technically, such methods utilize algorithms such as active contours or surfaces, deformable models, active shape- or appearance models, or level-set methods.

An early publication following this approach comes from Coulon et al. [20]. They presented a semi-automatic technique based on an active surface. Manual initialization of the centerline of the spinal cord is required, using multiple markers on a sagittal plane. From those, a cylindrical cubic B-spline surface is fitted, by minimizing an energy-term consisting of internal (surface bending, torsion, elasticity), and external (intensity gradient) energies, along with some regularization. The optimization process is computationally expensive, with 4 to 10 hours computation time reported on hardware of that time. They performed both a phantom and a two-sample scan-rescan evaluation, with a reported coefficient of variation of 1.3% on the scan-rescan data.

Later, Horsefield et al. [55] presented a closely related approach using a simplified model for the spinal cord surface. Instead of the B-spline surface, they optimize evenly distributed radii along a user-defined centerline. Driving forces are the Sobel-derived intensity gradient and a regularization term enforcing a smooth surface. By that, they were able to reduce calculation times down to 1 to 2 minutes. They evaluated their method using inter- and intra-rater analysis, but no scan-rescan comparison. For this, they achieved slightly better results than those reported by Losseff et al.[80]. Later, their method was used by several groups ([111, 127] to correlate cord atrophy with clinical markers in MS.

McIntosh et al. [89] presented a concept called ‘spinal crawlers’, an extension of an earlier method by the same group called *deformable organisms (DefOrg)* ([52, 88]). They model an organism’s body as a spring-mass system, that gets equipped with sensory modules, behavioral routines and decision making strategies calibrated to detect and follow the tubular structure of the spinal cord. The user needs to provide two seed-points to initialize the algorithm. The reported computation time is 10 minutes. Originally, they evaluated their method on a set
of four MRI images by measuring the Hausdorff distance between the automatic results and a fully manual segmentation. Later, they extended this approach with a live-wire technique to reduce segmentation errors [90]. This was then evaluated on 20 T1-weighted MRI images from 10 healthy volunteers and 10 MS patients, again evaluating the Hausdorff-distance from a manually delineated ground-truth.

Kawahara et al. [61] later presented another extension of this method. They incorporate a probabilistic shape-model as external energy force. The shape model was trained using a principal components analysis on 2D axial cross-sections of the spinal cord. Their model incorporates partial volume estimation. For evaluation, they presented a modified formula of the Jaccard index to account for the PVE, as well as an area-overlap. Reported mean Jaccard is 0.82, mean volume-overlap is 95.6%.

A technique based on 2D level-sets was proposed by Sonkova and colleagues [119]. Their method requires the user to define a region-of-interest on an automatically detected mid-sagittal plane, and place markers for initialization inside the spinal cord. Then they perform a 2D level-set clustering on axial slices, propagating the results iteratively. To reduce the impact of noise, they perform anisotropic diffusion filtering [100] as a preprocessing step. They evaluated their methods on 24 3D T1-weighted MR images of MS patients and age- and sex-matched healthy controls.

Koh et al. [66] presented an active-contours method using gradient vector flow fields as external forces for segmenting the spinal cord on 2D sagittal slices. Later, they extended their method by incorporating a saliency map derived from a Gabor filterbank [67]. From that, they derive initial contours for the active-contours optimization. This extension allows them to perform a fully-automatic segmentation on sagittal images. They evaluate their method on a dataset of 60 MS patients, for which they achieve a mean DICE coefficient of 0.71 in comparison to a manually generated reference-segmentation.

Law et al. [73] published a method for spinal cord centerline extraction and segmentation over a long segment of the cervical, thoracic and lumbar spinal cord. Their method requires a manual definition of the start- and end-point of the measurement. Using a combination of a gradient competition descriptor for orientation detection along with an anisotropic speed function for centerline detection, the algorithm then finds an optimal connection between the given points. Segmentation is then carried out by clustering voxels based on intensity similarities, smoothness, and a connectivity constraint. They evaluated their technique using the Dice similarity coefficient with reported mean result of 0.81 over 10 T1-weighted images, and 0.89 over 15 T2-weighted images. A clinical evaluation has not been reported.

DeLeener et al. presented another fully automatic spinal cord segmentation pipeline based on a multi-resolution deformable surface model [24]. For initialization,
they implemented a processing chain applying an elliptical Hough transform on a vesselness filtered image [39] with subsequent filtering of false positives based on several assumptions about the shape and localization of the spinal cord. From these ellipses, a tube-shaped triangle mesh is fitted to the image using gradient and shape information. This fit is performed in two steps, first using a low-resolution mesh for coarse optimization, second using a high-resolution mesh for local refinement. Later [23], they coupled their segmentation technique with a method for template-based automated vertebral labeling [126]. The whole processing chain has been integrated into the open-source software package *Spinal Cord Toolbox (SCT)* [25].

**Image-based methods**

A third category of algorithms comprises those approaches that involve analysis of an image as a whole. This group covers algorithms based on neighborhood relations such as graph cuts or Watersheds, as well as those techniques involving template-based techniques and the use of atlas registrations. Similar to surface-based methods, these techniques define a segmentation task as either a local or global optimum to some minimization problem. In contrast to the previous category, the result of this process is not a surface but rather set of voxels defining the solution.

A fully automated method using an atlas-based localization of the spinal cord was presented by Carbonell-Caballero et al. [16]. They perform a slice-wise 2D atlas registration using a correlation-coefficient based similarity measure and a brute force optimization approach to find a rigid template transformation. From this template, they analyze mean tissue intensities for spinal cord and CSF on the slice to derive a threshold-based segmentation. The precise area of the cord is then obtained using an expectation-maximization algorithm to fit a mixture of normal distributions to the local histogram of the cord. They evaluated their method by comparing the automated results to multiple fully manual segmentations on T2-weighted MRI, with a reported mean correlation of 0.982%.

Bergo et al. [6] presented a software toolbox providing both a semi-automated spinal cord segmentation tool and a manual correction method. The automatic segmentation is based on tree-pruning on an Image Foresting Transform (IFT) [31], a method closely related to graph-cuts. Here, the user needs to provide samples defining an object and the background, from which the algorithm finds an optimal segmentation. From this segmentation, atrophy parameters are derived as geometric parameters such as flatness and longest diameter of an automatically fitted ellipse on the cross-sectional spinal cord. Also, the area of the cross-section is evaluated. Where the automatic segmentation fails, a manually drawn ellipse can be used to measure these parameters. Their method does not explicitly address partial volume effects. Instead, they evaluated the effect of upsampling the original
image data with a factor of up to four, and evaluated the effect on accuracy. Evaluation was performed on a dataset of 30 images involving healthy volunteers and patients. They evaluated the inter- and intra-rater reproducibility between three raters performing a total of 360 individual measurements with a mean CoV of 2.91% for intra-rater and 4.06% for inter-rater agreement.

A related approach was presented by Fonov et al. [38]. They proposed to use graph cuts to perform spinal cord segmentations on axial slices. The user is required to place one seed point at the center of a spinal cord cross-section, from which the first segmentation is derived. The result is then propagated to adjacent slices using a linear registration, followed by an individual segmentation on that new slice. They have used their technique as basis for creation of a generic spinal cord template. Pezold et al. [101] presented a similar approach that combines a graph cut based pre-segmentation with an edge-based refinement process. Following the segmentation of the cord, they also extract the centerline and calculate cross-sectional views orthogonal to this.

An approach relying exclusively on an atlas registration has been developed by Stroman et al. [122]. They have created a spinal cord atlas from fMRI data of the spinal cord using images from 24 healthy people. They implemented a semi-automatic registration technique, that relies on manual user initialization through landmarks, followed by a linear transformation and 3D spatial normalization step using the user-provided curvature of the spine. They evaluated their method on fMRI images from eight volunteers by measuring distances between 1000 reference points spanning the whole spinal cord and brain stem. On average, 93% of these points were below $2\,\text{mm}$ distant from the template positions. The maximum reported error is $5\,\text{mm}$.

Later, Chen et al. [18] presented a fully automatic segmentation technique relying primarily on an atlas segmentation. Their method makes no assumptions on the field of view of the input data, and is capable of segmenting arbitrary segments of the cervical and thoracic spinal cord. To achieve this, they couple their template image with topological meta-information including probabilities for the presence of CSF and spinal cord tissue. They evaluated their approach on both T1- and T2*-weighted MRI data. Using a qualitative assessment of automatic results, they report a success-quote 96.6% on a total of 238 images. Additionally, they analyzed the correlation of EDSS scores and cervical cord atrophy in a cohort of MS patients, reporting results in agreement to a similar study by Horsefield et al.[55].
Chapter 2. Quantification of spinal cord atrophy in Multiple Sclerosis

2.3 Automated localization, segmentation and quantification of the cervical spinal cord

The aim of the method presented in the following section is to provide a fully automated workflow for measuring the volume and mean cross-sectional area of the upper cervical spinal cord. This process consists of three main steps: localization, segmentation and quantification. In order to achieve robustness for the complete pipeline, each individual step needs to perform robustly on its own, as each step relies on the results of the previous.

For localization, a template based method using a customized registration workflow is proposed. The registration is supported by a dedicated preprocessing pipeline that aims at homogenizing the input data in favor of achieving a stable registration result. A systematic evaluation of the main parameters of this pipeline with respect to their influence on the overall registration accuracy and robustness is presented. Segmentation is carried out using an interactive watershed segmentation. In order to achieve full automation, a heuristic approach for basin selection is introduced. For quantification, a Gaussian mixture model is fitted to the intensity distribution of the segmented spinal cord. The mixture model provides both a classification of tissue type within the object, as well as a precise volume measurement. This builds the basis for measuring the both the precise length, as well as mean cross-sectional area. A brief overview of the process as a whole is given in Figure 2.4.

2.3.1 Template-based localization

Template matching is a widely used approach for identification of structures in medical images. It builds upon the fundamental assumption that if a basic prior knowledge about the location, shape or size of a structure of interest exists, this can be located within an individual dataset by matching an appropriate template to it. The tools for achieving such a match can be found in the area of image registration. If the registration is successful, knowledge from the template can be transferred and structures of interest can be located. Such an approach requires two components: First, a template image suited to fit the structure of interest. Second, a registration pipeline capable of matching the images. In this work, both of these elements are presented with respect to the task of cervical spinal cord analysis.

Basics of image registration

The image registration problem can be generally phrased as the task of finding "a reasonable transformation such that a transformed version of a template image is
Figure 2.4: Overview of the image-processing pipeline: Template-based localization of the spinal cord is followed by watershed-based heuristic segmentation. The segmented object is then quantified using a Gaussian mixture model.

Similar to a reference image” [92]. The solution to this problem can be formulated as an optimization problem.

$$J[y] = D[T[y], R] + \alpha S[y] \rightarrow \min,$$

Here, $T$ and $R$ denote a template and a reference image, $y$ is the transformation, $D$ is the distance measure, and $S$ is a regularization term. $\alpha$ is a scalar parameter controlling the strength of the regularization. Finally, $J$ is the joint functional to be minimized over the transformation $y$. From this general definition, two elements can be identified as crucial for the overall result of the registration, namely the choice of an appropriate distance measure $D$, and the type and strength of the regularization term $\alpha S$.

The distance measure defines what is considered similar within an image. It needs to be chosen in accordance with the given registration problem. A simple approach is to directly measure the energy of the difference image $T[y] - R$. For optimization purposes, the $L_2$-norm is evaluated as the sum of squared differences, \textit{(SSD)}. SSD works reasonably well if the input images expose a similar signal range (i.e. the difference between a point $T(x_1)$ and its reference point $R(y(x_1))$ equals 0). In medical imaging, this assumption is rarely valid. While it holds for CT images, it is typically not given for MR images, where no absolute reference for measured signal intensities exist. Further, it is invalid by definition for multimodal imaging, where different images basically show different properties of the same object, which in consequence leads to different image intensities. Another disadvantage of SSD is its lack of robustness. As the difference is squared, the error rises quadratically if outliers are present in either one or the other image.
Chapter 2. Quantification of spinal cord atrophy in Multiple Sclerosis

For multi-modal image registration, alternative similarity measures have been developed. Cross-correlation aims at overcoming the dependency on direct intensity correspondence as found in SSD by maximizing the correlation between two images. Mutual information (MI) exploits the idea, that if two images show similar structures at corresponding positions, they will exhibit more or less clearly isolated peaks in their joint intensity distributions, regardless of the actual intensities and even their relations to each other. This can be measured as the entropy of the joint images. In consequence, image registration based on mutual information can be achieved by maximizing the entropy of the images joint intensity distributions.

Although mutual information has been widely applied as a distance measure for multi-modal registration, it exposes several critical drawbacks. For one, the joint density of a discretized image can only be approximated, which is typically done using either histograms or Parzen windows. These, however, need to be parameterized (i.e. bin-size of the histogram, width of the Parzen window) and the resulting registration is highly sensitive to the chosen parameters. Even more crucial, however, is the fact, that mutual information will typically expose many local maxima that the optimization process may run into, distracting from the actually desired global optimum.

As an alternative distance measure, normalized gradient fields (NGF) have been proposed by Haber and Modersitzki [45]. Here, image-similarity is defined as follows: "two images are considered similar, if intensity changes occur at the same locations." Intensity changes can be measured by calculating the gradient of an image. However, as the magnitude of the gradient directly depends on the dynamic range of the input images, the normalized gradient field is used instead of the gradient magnitude. When comparing two images, the angles formed by the two vectors at each relating image position can be evaluated using either the cross- or dot-product.

The potential of NGF as a distance measure especially for non-linear registrations has been demonstrated in numerous applications. As such, the focus in this work is put exclusively on NGF as a distance measure.

The normalized gradient field of an image is defined as follows:

\[ n_\varepsilon(I, x) := \frac{\nabla I(x)}{\|\nabla I(x)\|_\varepsilon}, \quad \|\nabla I(x)\|_\varepsilon := \sqrt{\nabla I(x)^T \nabla I(x) + \varepsilon^2} \quad (2.2) \]

Essentially, the gradient vector \( \nabla I \) for each image element \( x \) is normalized. This approach completely discards the gradient strength, and leaves only the gradient direction at each position within an image. In order to suppress very small gradients originating from image noise, and also to obtain differentiability for those positions where no intensity change occurs at all, a parameter \( \varepsilon \) is introduced to the normalization quotient. For gradients significantly larger than \( \varepsilon \), the effect of \( \varepsilon \) on the normalization can be neglected. However, in regions with \( \varepsilon \) larger than the
2.3. Automated localization, segmentation and quantification of the cervical spinal cord

gradients, \( n_\varepsilon(I, x) \) comes close to 0. Illustratively speaking, \( \varepsilon \) acts as a dampening filter on small gradients. Haber and Modersitzki originally proposed an automatic estimate of \( \varepsilon \) based on the overall signal-to-noise level of the input image. In this work, however, \( \varepsilon \) is one of the three main parameters to be evaluated with respect to the overall registration accuracy.

Based on the definition of \( n_\varepsilon(I, x) \), the similarity between two images can now be measured by evaluating the angles formed by \( n_\varepsilon(R, x) \) and \( n_\varepsilon(T(y), x) \) using either the cross- or the dot-product of the two. For reasons of simplification, the following formalization is limited to the use of dot-product, which relates to the cosine formed by the two angles. From an optimization point of view, the only difference between using the cross- or the dot-product is that square of the dot-product needs to be maximized, whilst the square of the norm of the cross-product needs to be minimized.

In summary, the distance-measure \( D_{NGF} \) can be defined as

\[
d(T, R) = \langle n(R, x), n(T, x) \rangle^2; \quad D_{NGF}(T, R) = -\frac{1}{2} \int_{\Omega} d(T, R) dx \quad (2.3)
\]

This builds the foundation for the registration pipeline presented in the next section.

Registration pipeline for template matching of the upper cervical spinal cord

In [35], Fischer and Modersitzki demonstrate the ill-posed nature of image registration. They show both mathematically and illustratively that even for very simple examples, it is impossible to find a unique transformation that maps a template to a reference. In the case of matching an atlas to an individual image, this observation becomes immediately evident as well: since a unique mapping of one to the other does not exist, it becomes impossible to find such a solution. Thus, the actual goal is to find a suitable transformation, that maps a subset of the information carried by the atlas to the target image, with sufficient accuracy. In terms of the optimization problem defined in Equation 2.2, this can be formulated as finding a suitable – not necessarily global – optimum without getting trapped by an unwanted solution.

To achieve this, Fischer and Modersitzki recommend the following ideas:

- Start with a proper pre-alignment
- Add regularization or penalties, to suppress unwanted solutions
- Explore multi-level / multi-scale techniques, to successively guide the registration from coarse to fine structures.
Further, they suggest the integration of user-knowledge as a step towards well-posedness.

The proposed method follows these ideas. It is designed to operate on 3D T\textsubscript{1}-weighted MRI images of the head. Such images, when acquired in clinical routine, typically cover the upper section of the cervical spinal cord which has been demonstrated to be the relevant section with respect to detecting pathological changes in MS. The typical image resolution is $\sim 1\text{mm}^3$ on both 1.5T and 3T MRI scanners. The field of view covers approximately the upper 25 cm of the head. Depending on the actual anatomy of a patient, this will include the spinal cord up to the vertebral discs between C3 and C5. These are the underlying prerequisites forming the basis for the following pipeline.

In the following section, a basic processing pipeline consisting of template selection, image pre-processing, image pre-alignment and finally combined linear- and non-linear registration is discussed. Subsequently, the performance of this pipeline under variation of several critical parameters will be evaluated in terms of robustness and accuracy, hinting towards a suggested, potentially optimal parameterization to this specific problem.

**Template selection:** The proposed method is built upon using a single individual MR image as an atlas. Alternative approaches, using averages of multiple images from different individuals have been neglected after achieving promising results using a single image approach. Aiming at a robust solution, the atlas image should be selected in a way that it closely resembles the expected patient images by means of MR acquisition parameters, field-of-view, and ideally also the scanner used. Further, it should exhibit a high level of image quality, as defined by a high signal-to-noise ratio coupled with an absence of typical MR imaging artifacts. For the atlas, it is advisable to choose an image with a large field-of-view, i.e. covering a large number of vertebral bodies, as only that information can be mapped to the patient’s image that is already present in the atlas.

**Pre-processing:** Data pre-processing aims at homogenizing the atlas and the patient image, based on user provided prior knowledge about the given problem. The goal is, to provide the registration algorithm with the best possible starting point. This is the basis for tuning registration parameters in a way that the algorithm is enabled to operate robustly on a broad range of input data. A schematic overview of the pre-processing pipeline is given in Figure 2.5. To this end the following steps are performed:

- **Cropping** – All non-head areas of the input images are cropped. This step serves primarily to reduce computation times by limiting the region-of-interest (ROI) to the head itself. To this end, a head-mask is extracted using a
2.3. Automated localization, segmentation and quantification of the cervical spinal cord

Figure 2.5: Overview of the pre-processing steps applied: (a) Cropping of the input image. (b) Intensity normalization to $i_{\text{median}} = 100$. (c) Inhomogeneity correction using N3 algorithm. (d) Anisotropic filtering using the sticks-filter.

combination of an Otsu-threshold [99] and basic morphological operations (i.e. closing, connected-component analysis). The resulting mask allows for measuring the height of the displayed head, which in the proposed registration pipeline can be used to pre-adjust the field-of-view between the template and the reference image prior to registration.

- **MR inhomogeneity correction** – low-frequency MR inhomogeneities stemming from $B_0$ magnet inhomogeneities are suppressed using the well-established N3 (non-parametric non-uniform intensity normalization) algorithm [118].

- **Signal normalization** – The intensity range of the patients image is re-scaled to match the input dynamic of the atlas image. Here, the median of the intensity distribution (i.e. the 50%-quantile) is normalized to (an arbitrarily chosen value of) 100. Note that this normalization step has a direct effect on the expected gradient magnitude of the image. This again interacts with the value of $\varepsilon$ of the NGF distance measure.
• **Anisotropic diffusion filtering** – Finally, the images are smoothed using an edge-preserving filter. As the goal here is to enhance straight, elongated structures such as the spinal cord, the use of a sticks filter is suggested. This approach has originally been developed by Czerwinski et al. for enhancing linear structures within speckle-noise, as found in ultrasound images [21]. The underlying idea is to analyze all lines passing through the center of a discrete filter kernel, and assign a value based on different properties accumulated over all sticks to the currently filtered voxel. By choosing the **mean of minimum variance** mode, the desired edge-enhancing effect is produced. In contrast to the standard anisotropic diffusion filter by Perona-Malik [100], this has the advantage of being straight-forward to parameterize, as the kernel size and orientation are the only parameters to set. In this application, the sum of two 2d-kernels along the sagittal and coronal image plane are used in order enhance edges along the main direction of the spinal cord, while avoiding smoothing within the axial plane.

It should be noted, that this smoothing step is to some extend concurring with the suppression of small gradients as performed by the parameter $\varepsilon$ of the distance measure. The effect of this, as well as the overall interplay between these two approaches will be evaluated.

**Atlas initialization:** Prior to the actual process of image registration, a deterministic pre-alignment of the images can be performed. Both images are assumed to display the head and upper neck as acquired in a mostly standardized orientation in clinical routine imaging. Such images are acquired in sagittal direction, with field-of-view of $25-30cm$. Initialization consists of determining both an initial rotation and translation, that can be used to compose a rigid transformation matrix and fed into the registration algorithm as an initial guess. The rotational part is obtained by realigning the orthonormal basis of both images. The translational part is determined by the vector spanned between the center-of-gravity of both images.

**Registration sequence:** For the actual registration process, a sequence of multiple separate steps is suggested. The rational behind this is to try to successively refine the registration result. A widely used intuitive approach is to concatenate a linear affine registration with a subsequent non-linear registration. In this work, an extension of this sequence is proposed. Instead of a single non-linear registration stage, a concatenation of multiple non-linear steps with subsequently relaxing regularization strength is suggested. The benefit of this is, that the risk of unwanted deformation caused by low regularization is reduced by initially starting with strong regularization, effectively creating a stiff deformation. This, however, will fail to
handle larger differences that are to be expected in atlas matching. Consequently, relaxing the regularization strength on a previously achieved acceptable stiff result will allow the deformation to iteratively refine itself.

This technique can be combined with a multi-level strategy. In order for the stiff registration to do the coarse work, it needs to start calculation on a coarse level, focusing on the overall geometry and coarse structures of the images. With relaxing regularization strength, the initial level needs to be raised to allow the similarity measure to focus on finer structures.

A multi-level representation of an image is obtained by successively downsampling the image by a constant factor, i.e. 2. By this, with each step the number of voxels, as well as the information contained in them gets halved for each image dimension. For 3D-images, this happens in a cubic manner, reducing the amount of information by a factor of $\frac{1}{8}$ per level. Note that from an information theoretic point of view, downsampling an image directly correlates to low-pass filtering. Consequently, the lower the scale gets, the lower the cutoff for the low-pass gets, removing high-frequent, i.e. small structures.

**Experimental setup for parameter tuning**

For parameter tuning and evaluation of registration performance, a dataset consisting of 111 $T_1$-weighted MRI images was used. The images stem from the MEISE study [113]. Images were acquired on a 3.0T MRI scanner (Siemens Skyra, Erlangen, Germany). MPRAGE sequence with TR=1900, TE=2.43 was used for the evaluation.

For each evaluation dataset, a reference segmentation was created using a modified, interactive version of the segmentation technique described in this work. In short, the spinal cord was segmented manually using an interactive watershed transformation. On the resulting object, tissue classification was done using the model described in detail in Section 2.3.3. The centerline was then extracted by sub-dividing the segmented spinal cord into multiple, thin slabs, and calculating the center-of-gravity of each slab. Finally, a B-spline was fitted through these points. This method has been previously published in [133].

Additionally, multiple markers were placed on the centerline indicating where each vertebral disc would subdivide the spinal cord. Since this subdivision is not anatomically clearly defined, the following approach was chosen: For each division, a plane running orthogonally to the centerline was aligned in such a way that it would exactly cross the center of the associated vertebral disk between two adjacent vertebral bodies on a sagittal plane crossing the center of the cord. The intersection of that plane with the centerline was then used as a landmark. As an exception, for the topmost landmark, the plane was aligned in such a way that it covers with the top of vertebral body C1. Figure 2.6 illustrates this in detail. The resulting
landmarks and spinal cord segmentations build the reference data for the evaluation of both registration accuracy as well as automatic segmentation results.

As a first step, a template image was chosen from the 111 available images. This was done as follows: First, 10 candidate images were selected based on manual, visual assessment, following the criteria listed above (cf. paragraph Template Selection). Each of these 10 candidates was registered to the remaining 110 images of the study, using an initial parameterization of the developed pipeline ($\alpha = 1$, $\varepsilon = 1$, anisotropic filtering=off). The results were then assessed visually, and the number of clear failures were counted, i.e. those cases, in which the registration sequence had clearly converged into an implausible result. The aim of this process was to identify one image as a template that would be best suited to fit the remaining images of the study. At this stage, the accuracy of the succeeding cases has not been evaluated, in favor of focusing solely on the success-rate. Consequently, the one candidate yielding the smallest amount of failed cases was chosen as template image. Using the default parameterization, the best performing image succeeded in 108 out of 110 cases.

Next, a systematic evaluation of three parameters of the registration pipeline has been performed by fully exploring the parameter space spanned by a range of plausible values for each. For evaluating the performance of a single parameterization, both the centerline of the spinal cord and the landmarks for vertebral subdivisions were analyzed separately. For the vertebral section markers, both the mean and maximum errors were analyzed. Here, the goal was to detect both the general agreement as well as critical registration errors as indicated by outliers. For the centerline, the Hausdorff distance for line-segments was chosen as a measure as described by Wirtz et al. [142]. To eliminate the bias of potentially shifted start- and end-points, evaluation of this measure was limited to the vertical section covered by both centerlines from the reference and the template image, i.e. the vertical overlap (cf. Figure 2.6). Here, the goal has been to identify how well the centerline has been matched with respect to its axial plane. Detection of a vertical shift along the spinal cords primary axis is covered by the vertebral disc markers. The following parameters were assessed:

**Regularization strength $\alpha$:** In non-linear image registration, the strength of the regularization term can intuitively be grasped as the stiffness, or rigidity of the deformation. It is controlled by parameter $\alpha$ from Equation 2.1. The impact of this parameter has been evaluated in two ways: First, by performing only a single registration step, successively increasing its value ($\alpha \in \{1, 10, 500\}$). Second, the effect of successively relaxing regularization strength has been studied, by starting off with a high $\alpha$-value on a coarse level, and lowering its value for the finer levels.
2.3. Automated localization, segmentation and quantification of the cervical spinal cord

Figure 2.6: (a) Landmarks defined for each image within the evaluation dataset: the centerline of the spinal cord, and markers indicating each section of a vertebral body. (b): Volume-rendering of the semi-automatically segmented spinal cord as reference.

**Edge-parameter $\varepsilon$:** For the normalized gradient field distance measure (NGF), parameter $\varepsilon$ defines a lower bound at which gradients are discarded from the measure. This serves two purposes: First, it reduces the effect of noise, and second it allows to steer the focus of the distance measure towards a controllable minimum gradient strength. Recall that NGF considers the gradient direction only, while completely discarding the gradient strength. As such, it is not capable of distinguishing whether a single gradient corresponds to a clearly visible edge, or to a small intensity shift caused by noise. For this evaluation, $\varepsilon \in \{0.1, 1, 10, 20\}$ has been evaluated.

**Anisotropic filtering:** Although a sufficiently high value for $\varepsilon$ will remove small gradients as caused by noise, the effect is not identical to actually de-noising the data prior to registration. Noise effects the data as a whole, while the edge-parameter effectively acts as a soft threshold, filtering only small gradients. In consequence, noise present on top of relevant edges is completely unaffected by this parameter. In contrast, anisotropic-filtering can effectively reduce noise in the data, while enhancing edges at the same time. This effect has been evaluated in a binary fashion, i.e. the whole pipeline has been individually assessed with and without filtering.
Chapter 2. Quantification of spinal cord atrophy in Multiple Sclerosis

Figure 2.7: Distribution of Hausdorff distances (for centerlines) and absolute distances (for vertebral disc landmarks) for different regularization strengths ($\alpha \in \{1, 10, 500\}$). Top: Three individual registrations. Bottom: Sequential registration with subsequently relaxing stiffness. Note the different scale on the x-axis on both plots.

Results of parameter evaluation

Figure 2.7 shows the distribution of localization errors for the two different registration strategies. The upper diagram displays errors for three independent registrations using $\alpha$ values of 500, 10 and 1. With respect to the Hausdorff distances of the centerlines, it can be observed that the mean error improves with decreasing $\alpha$, however, at the same time, the number of outliers increases significantly. Looking at the absolute distances of the vertebral disc landmarks, the lowest mean error is observed for $\alpha = 10$. However, here the number of outliers, as well as the overall width of the error distribution increases drastically. From these observations, it can be concluded that strong regularization helps with respect to stability of the result, while weak regularization can increase accuracy - if a stable
2.3. **Automated localization, segmentation and quantification of the cervical spinal cord**

Figure 2.8: Effect of anisotropic filtering in combination with different $\varepsilon$-values ($\varepsilon \in \{0.1, 1, 10, 20\}$), logarithmic scale on x-axis. Although the effect on accuracy is negligible, anisotropic-filtering significantly reduces the number of complete failures.

result is found. Looking at the outliers in detail reveals that here, the registration failed completely, i.e. there was no plausible match of corresponding anatomical structures.

In contrast, looking at the sequential registration, where each step in the pipeline builds upon the result of the previous step, it can be observed that both the mean error as well as the number of outliers decreases for each step. This is well in line with what would be expected. Strong regularization prevents strong local deformations, which is beneficial for initial alignment on a coarse level. This improves stability at the cost of accuracy. Providing such a coarse result to a slightly less stiff subsequent step allows for further relaxing. At the same time, the risk of converging towards a false result is reduced by the better initialization. This step can be repeated multiple times, although in this evaluation, the improvement between the second and third step was minimal.

Next, the effect of the edge-parameter $\varepsilon$ has been evaluated. Recall that for the similarity measure, this parameter acts as a weak threshold for suppression of small gradients. This parameter will be evaluated in companion with the anisotropic-diffusion filtering that aims at the same issue from a different direction. Figure 2.8 compares Hausdorff distances for different $\varepsilon$-values ($\varepsilon \in \{0.1, 1, 10, 20\}$), both with and without anisotropic filtering activated. Two things can be seen here. First, there is a slight improvement in terms of accuracy achieved with anisotropic filtering active. However, this effect is rather small and does not strongly motivate filtering. Much more interesting is the effect on outliers. Within the data used for evaluation, there were four cases that would completely fail without anisotropic
Chapter 2. Quantification of spinal cord atrophy in Multiple Sclerosis

Figure 2.9: Example of a case for which the registration completely failed without anisotropic filtering (a), while generating a proper result with filtering (b). The patient image is used as base-image, the gradient of the template image shown as blue overlay.

filtering. Those cases achieved acceptable results with filtering activated. Figure 2.9 shows such an example.

The effect of $\varepsilon$ alone is plotted in Figure 2.10. The plot shows the distribution of Hausdorff distances in the upper four bars, and the absolute landmark distances in the lower bars. Looking at the centerlines alone, a value of $\varepsilon = 10$ appears to be optimal, although the difference in terms of both median accuracy as well as distribution of errors is very comparable to that for $\varepsilon \in \{0.1, 1\}$. Larger values, however, clearly worsen performance and introduce a significant number of failing cases. With respect to vertebral disc landmarks, the picture is less clear. Here, $\varepsilon = 0.1$ yields the lowest median error, however it also gives more prominent outliers. In terms of outliers, again $\varepsilon = 10$ performs best.

A note on the chosen range of $\varepsilon$: As $\varepsilon$ acts as a soft threshold on the gradient magnitude, its range needs to be adjusted to the expected gradient strengths relevant to the registration task. The gradient magnitude again depends linearly on the overall dynamic range of the input image. In consequence, $\varepsilon$ also depends linearly on the input image’s dynamic range. During the preprocessing stage, this range is normalized to a value of 100 for the median of the intensity distribution. The observed optimal value of $\varepsilon = 10$ thus corresponds to 10% of the median input intensity.
2.3. Automated localization, segmentation and quantification of the cervical spinal cord

![Graph showing impact of edge-parameter ε](image)

Figure 2.10: Comparison of different ε-values \( \varepsilon \in \{0.1, 1, 10, 20\} \), \texttt{anisotropicFiltering=true} for centerline Hausdorff-distances and vertebral disc landmark distances. Higher values slightly increase accuracy.

In summary, the following recommendations for finding optimal parameters can be given. Note that these recommendations need to be treated with care. In this context, one should keep the ill-posed nature of image registration in general in mind. While these parameters gave optimal results on the dataset used for this evaluation, results may differ for different datasets.

- **Regularization:** A positive effect of iteratively relaxing the regularization strength can be stated. Starting with a stiff, coarse alignment with successively relaxing \( \alpha \) clearly helped improving accuracy while keeping the overall registration stable.

- **Anisotropic filtering:** Anisotropic filtering had a significant positive effect of reducing the number of complete failures. It also increased overall accuracy on a moderate scale. In this work, a sticks-filter was chosen as it provides robust results with little need for manual parameter adjustments. However, other filters with similar properties should have a comparable effect.

- **Edge parameter \( \varepsilon \):** An interval in the range of approximately \( \varepsilon \in [1\%, 10\%] \) of the median input intensity proved to yield optimal results. Lowering the value had minor effect (likely being compensated by the filtering). Increasing the value started introducing failing registrations, likely because relevant image gradients were being cut-off.
2.3.2 Heuristic watershed-based segmentation of the spinal cord

For segmentation of the spinal cord, the use of a watershed-transformation (WT) is proposed. The watershed-transformation has become a widely used basic segmentation algorithm, capable of providing robust and meaningful segmentation results. It is based on partitioning an image into small regions of similar intensities, with subsequent merging of selected regions to form the final segmentation result. The core of this concept is the interpretation of an image as a topographic relief with image intensities corresponding to topographic heights. Such a relief is composed of a set of adjacent basins, where each local minimum corresponds to the bottom of a basin, while local maxima separate neighboring basins. A segmentation is obtained by virtually filling the basins with water. Wherever one basin merges into another, a watershed is formed. With an increasing level of water, more and more basins will merge, eventually resulting in a single basin. At the same time, a hierarchy of watersheds is formed, which allows extraction of a specific segmentation by clustering all basins controlled by a specific watershed.

A standard algorithm for implementation of the watershed transformation on digital images has been published by Vincent and Soille [129]. It builds the basis for computationally efficient application of the WT in digital image processing. Hahn and Peitgen [49] introduced an extension of this algorithm in shape of the interactive watershed transformation (IWT). They extended the original algorithm by organizing all atomic basins in a tree-structure. From this, individual segmentations can be extracted efficiently by selecting subtrees and clustering all corresponding leaves (i.e. atomic basins). They further introduced the preflooding height parameter, which allows to control the minimum basin depth. This provides an effective means to control clustering sensitivity and prevent over-segmentation. Finally, their algorithm allows for interactive control of the segmentation result by selecting individual branches within the tree structure, which from a users perspective can be realized by simply placing markers within the image.

With respect to fully automatic segmentation of the spinal cord, the IWT appears to be a promising algorithm candidate. Its suitability for interactive segmentation of structures in medical images has been demonstrated in numerous cases, such as for skull-stripping in MRI [50, 117], segmentation of cerebral ventricles [42, 47], segmentation of the lung and lung-lobes from CT [71], to name but a few examples. It has also been used for semi-automatic segmentation of the spinal cord from MR previously [22, 83, 133], relying on manual localization and marker placement.

The proposed algorithm for automatic marker placement attempts to simulate the way a human operator would place markers. The general process would be to repeatedly place a single include- or exclude-marker, evaluate the result, and repeat this process until the resulting segmentation meets the expected outcome. Optimal
2.3. Automated localization, segmentation and quantification of the cervical spinal cord

Figure 2.11: (a) The inner and the outer segmentation boundaries derived from the eroded / dilated reference segmentation. (b) Initialization using include-marker from the reference centerline. The orange area outside the outer boundary is identified as over-segmentation. (c) Final segmentation result, after placing a single exclude-marker to a randomly chosen bright voxel within the over-segmented area.

positions for markers are within the "deep" basins of the watershed-transformation, as here, they mark larger sections of the basin-tree. As the IWT is performed on an inverted image, deep basins correspond to high image intensities. Practically speaking this means that markers should ideally be placed within local maxima of the input image.

As a prerequisite, the segmentation algorithm requires a proper localization of the spinal cord through the template registration described in Section 2.3.1. Based on this, the segmentation algorithm works as follows:

1. From the template annotation, the centerline, the vertebral disc marker, and the spinal cord segmentation mask are transformed to the patient image using the deformation field $y$.

2. A horizontal slab is extracted along the transformed centerline. This defines both start- and end-points for the segmentation. It is primarily important to define an upper limit and prevent the segmentation to extend into the brainstem or further.

3. The transformed centerline is then sampled in $1cm$ steps. The resulting points are used as initial include-markers for the segmentation.
4. Starting from this initialization, an iterative process of marker placement starts, where in each step, the current segmentation agreement with an expected result is evaluated, and include- and exclude-markers are placed until this requirement is met.

5. For evaluation of the segmentation result, two mask objects are created based on the template’s spinal cord segmentation. The inner and the outer bound. These masks are generated using morphologic erosion and dilation respectively. Segmentation markers are generated, whenever a bright structure is either present outside the outer bound, or absent within the inner bound. Bright relates to the mean intensity of the transformed reverence spinal cord within the patient image. To estimate this value, the mean ($\mu_{SC}$) and standard-deviation ($\sigma_{SC}$) of the intensity distribution of the transformed, eroded spinal cord mask is used, and the threshold for bright structures set to $\mu_{SC} - 3\sigma_{SC}$.

For each iteration, only a single marker is generated. This is done only for the largest cluster of voxels above the threshold. The exact position of the marker is picked randomly from the 25% brightest voxels within this cluster, again corresponding to a relatively deep basin. The randomness introduced here makes the whole process non-deterministic, which helps in case the algorithm doesn’t terminate after a set limit.

The process terminates as soon as either no more outlying clusters are detected, or a pre-defined maximum number of iterations is reached ($n = 10$). This may happen, if the actual positions for the chosen markers is conflicting in a way that the interplay between include- and exclude markers prevents the segmentation from succeeding. To overcome this issue, the non-deterministic nature of the algorithm is exploited. In case of termination without a satisfying segmentation result, the whole process is restarted, this time yielding different positions for the correction marker. Only if the algorithm as a whole fails repeatedly, the segmentation is considered to be failed. Algorithm 1 describes the proposed algorithm in pseudo code.

2.3.3 Model-based quantification of spinal cord cross-sectional area

The primary parameter for quantification of spinal cord tissue integrity is the mean-cross-sectional area (MCSA) of the upper cervical spinal cord. This can be precisely calculated by measuring the volume of a sub-section and normalizing for its length. With a labeling of vertebral bodies available, the MCSA can also be estimated section-wise for each vertebral body.
2.3. Automated localization, segmentation and quantification of the cervical spinal cord

Algorithm 1 Pseudo code for automatic marker placement. $S$ denotes a set of voxels forming a segmentation mask. $M$ denotes a set of markers. The function segmentSpinalcord takes the source image $I$, the reference segmentation mask $S_{\text{Ref}}$ and a set of initialization markers $M_{\text{Ref}}$ as input. In case of successful termination, it returns the spinal cord segmentation of input image $I$.

```plaintext
function segmentSpinalcord(I, S_{\text{Ref}}, M_{\text{Ref}})
    S_{\text{Outer}} ← dilate(S_{\text{Ref}}, x = 7, y = 7)  # Get dilated reference mask
    S_{\text{Inner}} ← erode(S_{\text{Ref}}, x = 5, y = 5)  # Get eroded reference mask
    M_{\text{Include}} ← M_{\text{Ref}}  # Initialize segmentation marker lists
    M_{\text{Exclude}} ← \emptyset
    \mu_{\text{SC}} ← mean(I(S_{\text{Ref}}))
    \sigma_{\text{SC}} ← stddev(I(S_{\text{Ref}}))
    t ← \mu_{\text{SC}} - 3\sigma_{\text{SC}}  # Threshold for outlier intensities
    initializeWatershed(I)  # Initialize IWT

    for i ← 1, maxIterations do
        S ← updateSegmentation(M_{\text{Include}}, M_{\text{Exclude}})  # Get segmentation for current marker configuration
        T_{\text{inside}} ← \{x | x \notin S \land x \in S_{\text{Inner}}\}
        T_{\text{outside}} ← \{x | x \in S \land x \notin S_{\text{Outer}} \land I(x) > t\}  # Calculate intersection with inner- and outer boundaries
        if T_{\text{inside}} \neq \emptyset then
            M_{\text{inside}} ← generateMarker(I, T_{\text{inside}})  # Generate marker for non-empty sets
        if T_{\text{outside}} \neq \emptyset then
            M_{\text{outside}} ← generateMarker(I, T_{\text{outside}})
        if M_{\text{inside}} \cup M_{\text{outside}} = \emptyset then
            return S  # Terminate if no markers were added
        return \emptyset  # Terminate if maxIterations is reached

function generateMarker(I, M)
    t ← quantile(I(M), 75)  # Calculate 75% quantile of masked image
    return random(\{x | x \in M \land I(x) > t\})  # Return random voxel-position with intensity above t
```

For measuring the volume, a Gaussian mixture model is used. Mixture models are based on the assumption, that individual tissue types within an image are
represented by a certain intensity distribution. For the given task, this requires the segmentation mask to consist exclusively of two distinct tissue types, spinal cord tissue and surrounding cerebrospinal fluid (CSF). Further, it requires sufficient samples of each tissue type available in order to obtain a robust estimate of each tissue’s intensity distribution. Both of these requirements are met by the result of the watershed segmentation. Figure 2.12 shows a typical segmentation result along with the corresponding intensity histogram.

Finally, the mixture model requires the parameters of the expected distribution to be constant over the whole set of voxels analyzed. For a Gaussian mixture model, this relates to the mean and variance values of each tissue class. This assumption is typically not guaranteed for MR images of the head. An inherent property of MRI is a low-frequent spatial fluctuation of image intensities caused by inhomogeneities of the magnetic field. In 3D T<sub>1</sub>-weighted images, this typically manifests as decreasing intensities towards the image borders, as compared to higher intensities within the center of the head. Apparently, the spinal cord being located in the lower part of the field of view, is especially affected by this. To overcome this, a simple inhomogeneity correction is proposed, that builds upon the existing segmentation of the spinal cord.

**Intensity normalization**

To overcome this problem, the segmented image is analyzed along the primary axis of the spinal cord. As the goal is to obtain a constant mean intensity of the spinal cord tissue regardless of spatial position, the image under the segmentation mask is analyzed in slabs of 5 mm thickness. For each slice, the histogram of the underlying segmented image is evaluated. From this, the intensity value of the right-most peak is determined, corresponding to an approximate mean intensity of the cross-section of the spinal cord in the given slab. Iterating over the full extent of the segmented object yields a vector of mean tissue intensities. This vector is interpolated with a cubic B-spline, which finally gives a continuous function for the spatial distribution of mean tissue intensities of the spinal cord. With this, the image can then be corrected for signal inhomogeneities by dividing each slice of the input region by its estimated mean tissue intensity. The resulting image contains the input spinal cord area, with a now constant average mean intensity of the spinal cord.

**Mixture model based quantification**

The mixture model used in this work utilizes a bimodal Gaussian distribution for the two pure tissue classes spinal cord and CSF, and an additional dedicated class modeling the partial volume distribution. Modeling spinal cord tissue with a single Gaussian is a simplification based on the assumption that gray- and white matter of
2.3. Automated localization, segmentation and quantification of the cervical spinal cord

Figure 2.12: Segmentation mask and corresponding intensity histogram resulting from automatic watershed segmentation.

the spinal cord can be represented by the same tissue type. This is largely valid for 1.5T or 3T MR imaging, it might however become an issue with respect to higher resolution imaging at high-field MRI. At 7T MRI, the internal structure of the spinal cord becomes visible and should consequently be considered when applying a mixture model. However, as the method proposed here aims at application on clinical data stemming from 1.5T or 3T MRI scanners, this is practically not an issue. Also, it has been shown by Gudbjarsson and Patz [Gudbjarsson1995] that the noise distribution in MRI data follows a Rice-distribution instead of a Gaussian. This aspect is neglected here, as for low signal-to-noise ratios, the Gaussian-distribution closely approximates the Rice-distribution.

The partial volume model used is the one proposed by Hahn [51]. It follows the concepts for partial volume modeling introduced by Santiago and Gage [114], which assumes that for two adjacent tissue types, the partial volume effect is both symmetric with respect to both classes, and spatially uniform across the whole sample. Spatial uniformity is achieved using the intensity normalization discussed above. Symmetry can be assumed as given as soon as the size of the imaged objects significantly exceeds the voxel size. This is also valid for the spinal cord. For two distinct tissue classes \(\alpha\) and \(\beta\), Hahn proposed the following model for the partial volume class:

\[
\Phi_{\alpha\beta}(x) = \frac{\Phi_\alpha(x) - \Phi_\beta(x)}{\mu_\beta - \mu_\alpha} \quad \text{and} \quad \Phi_\alpha(x) = \int_{x'=-\infty}^{x} \varphi_\alpha(x') dx' \tag{2.4}
\]

Here, \(x\) is a gray value and \(\varphi_\alpha(x)\) is a normalized Gaussian distribution with parameters \(\mu_\alpha\) and \(\sigma_\alpha\), corresponding to the pure tissue of class \(\alpha\). This essentially
models the partial volume distribution as a smooth rectangular function, bound at both sides by the integrals of each pure tissue’s distribution. For an arbitrary number of classes $\alpha \in \mathbb{N}$, the full model resolves to:

$$f(x) = \sum_{\alpha} A_{\alpha} \cdot \varphi_{\alpha}(x) + \sum_{\alpha, \beta | \mu_{\alpha} < \mu_{\beta}} A_{\alpha\beta} \cdot p_{\alpha\beta}(x) \quad (2.5)$$

Here, $A_{\alpha}$ is the amplitude of tissue class $\alpha$. The model is fitted to the local histogram of the segmented spinal cord object using a least-square minimization. Assuming that the source histogram follows the expected distributions, the model can be initialized robustly by setting initial values $\mu_{CSF} = 0.25$ and $\mu_{SC} = 0.75$ with respect to the relative range of histogram.

From a successful fit, the volume $V_{SC}$ of the spinal cord is obtained by taking the sum of the total volume of the pure tissue-class plus half of the volume of the PV-class:

$$V_{SC} = \int A_{SC} \cdot \varphi_{SC} + \frac{1}{2} \int A_{PV} \cdot p_{PV,CSF} \quad (2.6)$$

Figure 2.13 shows the result of the classification process. The classification result is displayed as color overlay using a traffic-light color scheme for CSF, PV, and spinal cord.

**Figure 2.13**: Left: Result of the tissue classification. The dotted curve represents the intensity distribution of the input region. The red, yellow, and green curves represent the individual tissue and PV classes. The gray area represents the summation of the fitted curves and shows good agreement with the input data. Middle+Right: Tissue classification as color-overlay over the original data.
2.4 Evaluation and Results

Centerline extraction

As a last step, towards determination of the MCSA, the length of the analyzed section needs to be measured. For this, the centerline of the segmented spinal cord is required. Extraction of the centerline is achieved by identifying a number of representative support points along the path of the spinal cord, and fitting a B-spline function through the given points. As a first step, each voxel of the input segmentation is labeled using its probability of containing pure spinal cord tissue as a label. The result is a weighted mask in the range $[0, 1]$. This corresponds to the green area in Figure 2.13. Next, this weighted mask is processed in slabs of 15mm thickness, equally distributed along the total cord. For each slab, the center-of gravity is calculated and added as a support point of the B-spline. For the start- end end-point, only the first and last slice are analyzed. Fitting a B-spline through these points yields a representation of the centerline, allowing to precisely determine the spinal cord’s length $L_{SC}$. Figure 2.14 illustrates this algorithm. With the length of the centerline we can now calculate the mean cross-sectional area by normalizing the measured volume to the section length:

$$MCSA = \frac{V_{SC}}{L_{SC}}$$

(2.7)

The overall method proposed here has been presented at the yearly conference of the European committee for treatment and research in Multiple Sclerosis (ECTRIMS) in 2017 [135].

2.4 Evaluation and Results

This method has been evaluated under two aspects: First, in terms of overall achievable accuracy of the quantification method for MCSA. Second, in terms of segmentation quality of the fully automated approach compared to a user supervised semi-automatic method. Both evaluations have been performed on different datasets.

Evaluation of accuracy

For evaluation of accuracy, a scan- / re-scan scenario was used. Images were acquired from 5 volunteers, each being scanned 5 times at one day, giving a total of 25 images. After every single image acquisition, subjects were asked to leave and re-enter the scanner enforcing a realistic repositioning between scans. All images were acquired on a 3T Skyra-MR scanner (Siemens, Germany), using a 3D MP-RAGE sequence in axial direction with isotropic voxel-size of $1mm^3$ (TR=1900, TE=2.43). The field of view was centered around the upper 8 vertebral bodies of the spine.
Chapter 2. Quantification of spinal cord atrophy in Multiple Sclerosis

Geometry distortion correction was applied on the scanner to correct for volumetric errors caused by gradient field inhomogeneities. We chose an axial acquisition scheme to minimize for motion artifacts caused by breathing or swallowing during the acquisition.

Analysis of these images was performed using the quantification method described in Section 2.3.3, combined with a manual localization and semi-automatic, supervised segmentation of the spinal cord. The manual approach has been a predecessor of the proposed method and has been described in detail in [133]. In short, the method works as follows: Localization is performed manually on a 2D rendering of the data. From an oblique-plane aligned through the center of the spinal cord, a vector is spanned defining start- and end-points of the measurement. Segmentation of the spinal cord is then done with an IWT, however, marker placement is done completely manual.

From each of these images, the mean cross-sectional area for three sections of the cord was evaluated: first, the topmost section ranging from C1-C2, second the section between C3-C4, and third the complete section from C1-C4, yielding a total of 75 measurements. For the intra- and inter-rater analysis, the same measurements were repeated five times by 4 individual raters in randomized order for only a single MR image of each volunteer. All measurements were performed by technicians who underwent a short training from an experienced neuro-radiologist to learn how to define the anatomical sections.

Table 2.1 shows the results of the scan-/re-scan analysis. The mean coefficient of variation achieved over the 75 individual measurements was 0.62% (min=0.094%, max=0.956%, median=0.570%). For an expected average MCSA value of $80\text{mm}^2$ in healthy people, this equals an area of $0.49\text{mm}^2$.

Figures 2.15 and 2.16 summarize the results of the inter- and intra-rater analysis.
### 2.4. Evaluation and Results

Table 2.1: Results of individual measurements (a=C1-C2, b=C3-C4, c=C1-C4) and individual scans (Scan1-Scan5), along with mean, std.-dev. and coefficient of variation.

<table>
<thead>
<tr>
<th>ID</th>
<th>Scan1</th>
<th>Scan2</th>
<th>Scan3</th>
<th>Scan4</th>
<th>Scan5</th>
<th>Mean</th>
<th>Std.-Dev.</th>
<th>CoV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>0.853</td>
<td>0.862</td>
<td>0.855</td>
<td>0.863</td>
<td>0.86</td>
<td>0.859</td>
<td>0.00439</td>
<td>0.51</td>
</tr>
<tr>
<td>1b</td>
<td>0.908</td>
<td>0.915</td>
<td>0.906</td>
<td>0.916</td>
<td>0.926</td>
<td>0.914</td>
<td>0.00789</td>
<td>0.86</td>
</tr>
<tr>
<td>1c</td>
<td>0.878</td>
<td>0.886</td>
<td>0.890</td>
<td>0.899</td>
<td>0.897</td>
<td>0.890</td>
<td>0.00851</td>
<td>0.96</td>
</tr>
<tr>
<td>2a</td>
<td>0.842</td>
<td>0.840</td>
<td>0.844</td>
<td>0.839</td>
<td>0.844</td>
<td>0.842</td>
<td>0.00228</td>
<td>0.27</td>
</tr>
<tr>
<td>2b</td>
<td>0.880</td>
<td>0.874</td>
<td>0.889</td>
<td>0.895</td>
<td>0.884</td>
<td>0.884</td>
<td>0.00808</td>
<td>0.91</td>
</tr>
<tr>
<td>2c</td>
<td>0.861</td>
<td>0.852</td>
<td>0.866</td>
<td>0.870</td>
<td>0.857</td>
<td>0.861</td>
<td>0.00712</td>
<td>0.83</td>
</tr>
<tr>
<td>3a</td>
<td>0.767</td>
<td>0.760</td>
<td>0.756</td>
<td>0.765</td>
<td>0.761</td>
<td>0.762</td>
<td>0.00432</td>
<td>0.57</td>
</tr>
<tr>
<td>3b</td>
<td>0.787</td>
<td>0.787</td>
<td>0.770</td>
<td>0.783</td>
<td>0.786</td>
<td>0.783</td>
<td>0.00723</td>
<td>0.92</td>
</tr>
<tr>
<td>3c</td>
<td>0.778</td>
<td>0.761</td>
<td>0.766</td>
<td>0.777</td>
<td>0.77</td>
<td>0.770</td>
<td>0.00723</td>
<td>0.94</td>
</tr>
<tr>
<td>4a</td>
<td>0.872</td>
<td>0.872</td>
<td>0.871</td>
<td>0.873</td>
<td>0.874</td>
<td>0.872</td>
<td>0.00114</td>
<td>0.13</td>
</tr>
<tr>
<td>4b</td>
<td>0.905</td>
<td>0.899</td>
<td>0.901</td>
<td>0.896</td>
<td>0.896</td>
<td>0.899</td>
<td>0.00378</td>
<td>0.42</td>
</tr>
<tr>
<td>4c</td>
<td>0.891</td>
<td>0.884</td>
<td>0.890</td>
<td>0.886</td>
<td>0.885</td>
<td>0.887</td>
<td>0.00311</td>
<td>0.35</td>
</tr>
<tr>
<td>5a</td>
<td>0.889</td>
<td>0.888</td>
<td>0.889</td>
<td>0.89</td>
<td>0.89</td>
<td>0.889</td>
<td>0.00084</td>
<td>0.09</td>
</tr>
<tr>
<td>5b</td>
<td>0.981</td>
<td>0.969</td>
<td>0.961</td>
<td>0.981</td>
<td>0.968</td>
<td>0.972</td>
<td>0.00877</td>
<td>0.90</td>
</tr>
<tr>
<td>5c</td>
<td>0.938</td>
<td>0.932</td>
<td>0.925</td>
<td>0.936</td>
<td>0.936</td>
<td>0.933</td>
<td>0.00518</td>
<td>0.55</td>
</tr>
</tbody>
</table>

respectively. For intra-rater analysis, the plot shows a single column for each measurement performed (C1-C2, C3-C4, C1-C4), on a single, selected image from each volunteer (5 images), performed by the 4 raters, totaling 60 entries in the plot. Within each column, each marker type represents on of five measurements. The y-axis displays for each such measurement the difference between a single a single measurement of this run, and the mean over all five measures. For inter-rater results, a similar plot is given comparing the mean of each individual measurement-set of a single rater, against the mean obtained by all raters.

It may be noted that for each subject, the measured section between C3 and C4 is significantly larger than the section between C1 and C2. In consequence, the total section between C1-C4 lies somewhere in between. Anatomically, one would expect this to be either roughly equal, or even the other way around, since the spinal cord gradually becomes thinner the further downward it goes. The reason for this lies in the acquired data. The axial acquisition direction creates a slight pin-cushion like geometric distortion, which narrows the geometry towards the borders of the image. Since the data has not been distortion corrected on the MR scanner, this effect manifests as larger MCSA measurement within the middle of the image, as opposed to the borders.

In summary, an excellent agreement between repeated measurements over differ-
Chapter 2. Quantification of spinal cord atrophy in Multiple Sclerosis

Figure 2.15: Plot showing the deviation of individual, repeated measurements from a single rater and the mean of those five measurements. Each column represents one of three measures, on a single volunteer’s image, by each of the four raters. Nearly all measurement-points are within an interval of \([-0.005mm^2, 0.005mm^2]\), except for a small number of outliers.

Figure 2.16: Here, the difference between the mean of each raters 5 measurements is compared to the mean value of all 4 raters. Again, all measurements are within an interval of \([-0.005mm^2, 0.005mm^2]\).

ent observers was reached with respect to both intra- and inter-observer reproducibility. For the intra-rater comparison, a CoV of 0.39% (min=0.12%, max=1.34%) was obtained. For the inter-rater comparison, a CoV of 0.28% (min=0.12%, max=0.53%) was achieved. The maximum CoV observed through all measurements was 1.34%, which can easily be credited to a small number of outliers found for single measurements. Averaging over multiple repetitions could be a possibility to address this issue.
2.4. Evaluation and Results

Figure 2.17: Plot (a) shows boxplots of the differences between a single rater’s measurement, and the mean measurement of all raters, effectively showing the amount of agreement between the 4 raters. Plot (b) shows the distribution of Dice-coefficients between segmentation masks generated by the fully automatic segmentation technique and manually generated reference masks.

Evaluation of segmentation quality

Evaluation of the automatic segmentation quality was performed on the MEISE dataset, that has also been used for parameter tuning of the registration pipeline [113]. Two approaches were used: First, a comparison of mask-images of the segmented spinal cord (after applying a voxel-classification based on the quantification result) between the automated method and masks generated using the above mentioned semi-automatic approach by a single rater. For the resulting masks, the Sørensen-Dice coefficient was calculated [28, 120]. As a second approach, the quantified values obtained for MCSA at levels C1-C3 were compared between the fully automated segmentation and three human raters, again using the semi-automatic approach. For evaluation of the latter, a similar approach to the one followed for accuracy evaluation was chosen: In absence of a known ground-truth, the mean of all 4 raters was chosen as a reference. Based on this, the absolute deviation of each rater’s single measurement was assessed.

As a first result, it should be noted that fully automatic segmentation succeeded for all images present in the study. Succeeded means, that the template based initialization yielded sufficient accuracy to initialize and guide the proposed segmentation algorithm.

Figure 2.17 summarizes the result of both approaches. Figure 2.17(a) shows boxplots for the distribution of deviations from the mean measurement for all raters. It can be seen, that the algorithm performed worst with respect to both
overall variance of the result, as well as in terms of outliers. However, the difference between the automated method and human raters A and B are minimal. Rater C in contrast achieved an overall better performance. Also, it should be noted that the median disagreement between the different raters was 0.44mm, 0.31mm, 0.44mm and 0.29mm respectively.

Figure 2.17(b) shows the distribution of Dice coefficients. Here, a median score of 0.983 was achieved, underlining the excellent agreement of the automatic segmentation result with those generated interactively under expert supervision. These results have been presented at the CARS conference 2018 [136].

Clinical evaluations

The quantification method presented here has been used in several clinical publications to demonstrate the relevance of spinal cord atrophy in the context of MS. This section summarizes recent publications based on this method.

Lukas et al. conducted a study on 440 MS patients from two centers [84]. They evaluated the relevance of spinal cord atrophy in comparison to whole brain atrophy, brain lesions, spinal cord lesions, and diffuse spinal cord abnormalities, with the aim of determining whether MCSA offers additional diagnostic and clinical information. Spinal cord analysis was performed on 3D T1-weighted MPRAGE images of the cervical cord. The section C1-C2 was chosen for measurement. The cohort of the study consisted of 311 patients with relapsing-remitting MS (RRMS) and 129 patients with progressive MS (PMS) (37 primary progressive (PPMS), 92 secondary progressive (SPMS). No controls were included in the study. Clinical disability was scored using the Expanded Disability Status Scale (EDSS), the timed 25-foot walk test (TWT) and the nin-hole peg test (9-HPT). They found significantly reduced MCSA in patients with PMS compared to RRMS. Further, they were able to show a highly significant inverse correlation between MCSA and EDSS, TWT, and 9-HPT test findings. MCSA and the number of T1-hypointense brain lesions were the strongest MR imaging parameters for explaining physical disability. They concluded that spinal cord abnormalities in general, and especially upper cervical spinal cord atrophy are important determinants of clinical disability in MS.

Daams et al. performed a similar study on a collective of 196 patients with long standing MS [22]. Inclusion criterion was a minimum disease duration of 10 years (mean 19.94 years). The patient cohort consisted of 125 patients with RRMS, 49 patients with SPMS and 22 patients with PPMS. Also, they included 55 healthy, age-matched controls (HCs) for comparison. Similar to Lukas et al., they compared MCSA, number of spinal cord lesions, brain volume and brain lesion volume with established clinical scores (EDSS, TWT, 9-HPT). For MCSA, the section C1-C2 was chosen for measurement. They found MCSA to be significantly lower throughout the whole patient cohort compared to HCs. Further, they found MCSA
in progressive MS to be significantly lower than in RRMS. Also they report an association between MCSA and EDSS, 9-HPT, TWT, and overall disease duration.

Liu et al. studied differences in atrophy patterns between MS and neuromyelitis optica (NMO) [77]. They recruited a cohort of 35 patients with NMO, 35 patients with MS (all RRMS), and 35 healthy controls. In addition to upper cervical MCSA, they measured brain volume (brain parenchymal fraction, gray matter fraction and white matter fraction) as well as lesion load in the brain and the spinal cord. They found MCSA to be significantly reduced in NMO, with no or very low amount of brain atrophy. In contrast, in the MS patient group, brain atrophy was significantly higher compared to HCs, with MCSA being less significant, even though reduced.

Later, the same group published a comparison of MCSA as measured on dedicated 3D MPRAGE images of the spinal cord and those measured on the upper cervical cord on typical MRI scans of the head, which typically includes the spinal cord section of vertebral bodies C1 to C4 or C5 [78]. 97 subjects were recruited from three clinical centers, including 60 MS patients and 37 controls. All data was acquired on 3T MRI scanners, including scanners from the three major vendors Siemens, Philips and GE. For all subjects, standard head images including only the upper cervical cord as well as cervical images including the whole cervical spinal cord were acquired. For a subset of 11 patients, additional contrast enhanced images after intravenous gadolinium admission were included and analyzed. Geometric distortion correction was applied on the scanners. They report an excellent agreement of MCSA measure taken on either cervical or head images, with an interclass correlation coefficient of 0.987. Comparing the results on head images with and without gadolinium enhancement obtained an ICC of 0.991, suggesting that there is little to no effect on the robustness of the method.

Bellenberg et al. analyzed a possible relation of MCSA with regional brain atrophy [5]. In a group of 132 MS patients (71 RRMS, 61 PMS) and 45 healthy subjects, they compared individual MCSA values measured around C1-C2 with local brain volumes using voxel-based morphometry. They found a positive correlation exclusively between MCSA and brain volumes in the infratentorial brain regions in RRMS patients. In contrast, in PMS, although considerably affected by both spinal cord and brain atrophy, no significant correlation between spinal cord area and specific brain regions was found.

Hagström et al. studied the relevance of cervical spinal cord atrophy in a 2-year follow-up study of patients with clinically isolated syndrome (CIS) and early MS [46]. Their study included 110 therapy-naive patients consisting of 53 patients with CIS and 57 patients with clinically definite MS. Additionally, they included 34 age- and gender-matched controls for comparison. During the course of the study, they further divided the group of CIS patients into those that converted into clinically definite MS and those that did not convert. Using this distinction, they
were able to show that converting patients had a significantly lower MCSA already at disease onset, in contrast to the non-converting group. No significant differences were found between the control group and the non converting group. Patients with definite MS also showed significantly lower MCSA compared to the controls. They conclude that spinal cord atrophy in MS is present from the very beginning of the disease. Additionally, they suggest that MCSA may be a predictive marker for conversion of patients with CIS into clinically definite MS.

All studies described above were conducted using the quantification technique presented in this work.

### 2.5 Discussion and Conclusion

In this chapter, a fully automated workflow for measuring the mean cross-sectional area of the upper cervical spinal cord has been presented. It is based on a combination of template based localization of the spinal cord, a heuristic watershed based segmentation, and a robust histogram based quantification. Each processing step involved in the pipeline has been evaluated in terms of robustness (for registration and segmentation) or accuracy (for the quantification). [44]

The following contributions to advancing the state-of-the-art have been made:

**Template registration:** A dedicated processing pipeline for the template based localization of the spinal cord, has been proposed and evaluated. This pipeline consists of a series of pre-processing steps including anisotropic filtering, MR inhomogeneity correction, and intensity normalization in order to homogenize the data. Then, a combination of geometric pre-alignment and linear affine registration is used to generate an optimal starting point for non-linear refinement. And finally, a multi-step non-linear registration based on normalized gradient fields as similarity measure, with iteratively relaxing regularization strength. Especially this last step helped to significantly improve registration accuracy, without sacrificing overall robustness. All central parameters involved in this pipeline have been evaluated and optimized.

**Heuristic segmentation:** For generating a segmentation mask of the spinal cord, an intuitive extension to the interactive watershed transformation has been developed. In short, the implemented method simulates how a human operator would manually place markers to obtain a segmentation. Exploiting the nature of watershed-based segmentations, an algorithm has been implemented that iteratively identifies good positions for marker placement, and subsequently compares the achieved segmentation result with an expected outcome, derived form the template segmentation. Should the algorithm fail to terminate after a set number of steps,
all markers placed are discarded and the process starts over. Since the positions of the markers are generated in a non-deterministic way, such a repetition will lead to a new, eventually better result.

**Quantification:** The quantification used in this method is equal to that proposed by Hahn. However, in order to become applicable for the given task, two additional processing steps became necessary: First, the correction of intensity inhomogeneities in the input image. Such effects corrupt the underlying assumptions of the quantification model used. In the context of spinal cord analysis from conventional MR images of the head, this effect is often even more pronounced, as a signal drop is stronger towards the borders of the image. Second, for precise estimation of MCSA not only the volume is required, but also the length of the measured section. Since the spinal cord is a curved object, a simple approximation using a line connecting the start- and end-point would introduce an avoidable error and reduce overall accuracy. In consequence, a method for extraction of the spinal cord centerline is proposed. These additions, in combination with the histogram based analysis provide the basis for the overall high accuracy of the method.

One aspect to note is that the tissue modeling is done using a Gaussian mixture model, whereas the actual distribution of noise in MRI is rician. For future research, it would be interesting to assess the impact of this aspect to overall accuracy of the model fit.

The pipeline as a whole has not been evaluated clinically. However, all individual steps involved have been published or presented at international conferences. Using the dataset from the MEISE study, the overall method achieved a success-rate of 100%. This rate relates to the overall robustness of the template registration, and to the heuristic watershed segmentation. Over these cases, a median Dice score of 0.983 compared to segmentations generated semi-automatically using the supervised, interactive workflow was achieved. Comparing measured MCSA values obtained from the fully automatic pipeline to those generated by three human operators showed an excellent agreement, well below 0.5mm$^2$ variation.

A potential limitation of the method is that lesions potentially present in the spinal cord are not specifically handled. This is relevant with respect to T1-hypointense lesions, corresponding to older lesions. In the images, such lesions will show reduced signal as compared to healthy spinal cord tissue. In the histogram, such lesions will typically be found in the upper quartile of the partial volume class. As such, the effect of a lesion would be a slight over-estimation of the partial volume class, at the cost of the pure tissue class. Evaluation of the impact of this effect could be an aspect of future work in the field.
Chapter 2. Quantification of spinal cord atrophy in Multiple Sclerosis

Related publications as primary author


Relevant publications as co-author


48
3 Interactive tools for identification and delineation of brain structures

At its heart, engineering is about using science to find creative, practical solutions. It is a noble profession.

(Queen Elizabeth II)

3.1 Introduction

Although large aspects of the functioning of the human brain remain still unknown, its overall structure on a macroscopic scale, individual components and also many aspects of the interplay and core processes between those components are well understood. This is especially true for the gyral structures of the brains cortex, as well as for many of the larger pathways forming the white matter inside of the brain. Consequently, many clinical questions specifically address certain pathways, gyral cortical structures, or the relations and connections between them. To support
Chapter 3. Interactive tools for identification and delineation of brain structures

such tasks with respect to image based measurements, it is a necessary prerequisite to identify and segment those structures.

For this purpose, two interactive approaches that facilitate such tasks are presented and evaluated in this chapter. First, an interactive contouring tool for diffusion tensor imaging data (DTI) data, that allows for clustering seed-voxels for DTI fiber tracking using a single click with subsequent, interactive adaption of a parameterization appropriate to the selected structure. Second, an interactive, 3D segmentation tool, that allows for precise delineation of gyral structures from a 3D-rendering of the human brain based on T1-weighted MRI data. Both tools could be individually used to address specific questions related to white matter fiber tracts, or structural properties of functional areas on the brain’s surface. But they could also be combined to explore the connections between functional areas and white matter tracts in individual human subjects.

3.2 2D-ROI generation from DTI data

Diffusion tensor imaging is a type of MR imaging technique capable of providing knowledge about the internal structure and organization of the axonal connections inside the white matter of the brain [2]. It is based on measuring the diffusion strength of water molecules, also known as Brownian motion, along a specific direction defined in the measurement. By repeating such a measurement for different directions distributed evenly around a sphere, the generated data allows for approximation of a diffusion tensor, a $3 \times 3$ matrix representing the primary diffusion direction and strength at any given location inside the measured volume. Since it is known that inside the brain Brownian motion is stronger along the directions of white matter axonal pathways, the measured diffusion properties can serve as a surrogate for the tissue orientation inside the otherwise mostly homogeneous appearing white matter tissue.

Fiber tracking is a specifically interesting and clinically well-established analysis technique for DTI data [94]. It facilitates the reconstruction of anatomically known white matter structures by tracing trajectories through the tensor field obtained from an individual scan. A crucial step when using this technique is the placement and shape of regions-of-interest (ROIs) to identify the structures in question. Typically, free-hand contours or simple geometric shapes like rectangles are placed in regions, where a given structure can be identified using the color coded DTI representation. However, such approaches result in a high variability of the resulting tracts and usually require additional filtering and placement of multiple ROIs. Also, the generation of accurate ROIs using a freehand tool can require a significant amount of interaction time, depending on the desired degree of accuracy. In this work, a method is presented which allows for interactive generation of anatomically
meaningful ROIs for DTI fiber tracking, based on geometric similarities of the underlying tensor field. The method works similar to the magic-wand tool known from image editing software tools to create reasonable, fully image based ROIs using a single mouse-click.

3.2.1 Motivation

DTI allows insight into the structural properties of the white matter of the human brain [95]. A number of analysis techniques for DTI data exist, among them the quantitative analysis of image parameters such as fractional anisotropy (FA), axial or radial diffusivity (AD, RD) or the apparent diffusion coefficient (ADC), as well as the reconstruction of anatomically known axonal structures using fiber tracking algorithms [3].

The most commonly used 2D-visualization of DTI images uses a color coding scheme, where the principal diffusion direction and the degree of anisotropy is encoded as a three component vector, which is subsequently mapped into RGBA color space [104]. This representation allows identification of known structures based upon their color and intensity at certain positions inside the brain. By convention, left-right is mapped to red, anterior-posterior is mapped to green, and up-down is mapped to the blue color channel. Following this color-scheme, the corticospinal tract which runs mostly perpendicular along the head-to-feet axis connecting brain stem, internal capsule and spreading into the sensorimotor cortical areas, appears mostly blue in the color coded scheme, while the corpus callosum which connects the left and the right hemisphere appears red on a mid-sagittal plane. This color coded representation has two properties, which are of relevance to the physician analyzing the data: First, areas with similar diffusion direction will receive similar color values, and second, areas with a similar degree of anisotropy are represented by similar color intensities. Consequently, anatomical structures with a large number of axons running mostly parallel through a region of the brain appear as blobs of the same color, which makes them easy to identify for the physician.

Nowadays, a great selection of software tools for dealing with DTI data is available. A comprehensive overview of different approaches and tools is given in [91]. One aspect of the author’s conclusion is that currently, there are no fully satisfying interactions schemes offered for dealing with DTI data.

Of all clinical applications of DTI, fiber tracking is certainly the clinically most wide-spread, with applications in neurosurgery, neurology and neuroradiology. When trying to reconstruct a known white matter tract on a patient’s individual DTI dataset, a physician first needs to locate an area in the brain, where the structure of interest can be identified clearly. This process is supported by overlaying the color-coded DTI data over a registered, anatomical MR image. The desired
structure can then be identified based on the color contrast of the DTI image. Afterwards, a ROI is created covering this region. Subsequently, all fibers passing through this ROI can be reconstructed using a fiber tracking algorithm [3, 40].

Typically, the resulting tract will be over-inclusive in some parts, while other parts of the structure might be missing. This is an inevitable property of DTI based white matter fiber tract reconstructions, which is rooted in the imaging method, the underlying diffusion model as well as the reconstruction methods and tracking algorithms utilized. The pragmatic approach to deal with this issue is to reconstruct a tract iteratively by putting together a number of different ROIs to control inclusion and exclusion of individual fibers [8]. However, a negative side-effect of this is that in theory almost arbitrary tracts can be reconstructed from any DTI dataset by just using enough ROIs. An alternative approach for selecting white matter structures from a DTI dataset is to perform tractography on the whole dataset (whole-brain fiber tracking, WB-FT), and subsequently group the resulting fibers by geometric similarities. This can be done either automatically or interactively [19, 58, 65].

A number of approaches for interactive real-time fiber tracking have been published. Such techniques can be divided into two groups. The first group covers GPU-based implementations of both deterministic [68, 70] and probabilistic [87] tracking algorithms. The high degree of parallelism offered by modern GPUs allows for a massive acceleration allowing for reconstruction of large fiber bundles at interactive rates. A second group focuses on optimized data-structures for efficient search queries on fiber structures [64]. Here, tracking for a full DTI dataset is performed in a pre-processing step. The resulting fibers are stored in a tree-based data-structure, which allows for time-efficient querying and interactive selection of those fibers that pass through an arbitrary ROI.

Such tracking methods appear very promising, since they potentially allow for a more explorative approach towards DTI analysis, as opposed to the time-consuming and systematically error-prone method of multi-ROI composition. However, interaction with such techniques is not trivial. Typically, a spherical or cuboid ROI can be moved through the dataset with interactive updates of the fibers passing through. While this certainly does create a wow-effect to the user, it does not necessarily yield anatomically plausible fiber-bundles.

To overcome such issues, it is desirable to have a tool at hand that allows for reproducible creation of anatomically meaningful ROIs, accompanied with a standardized description of how specific white matter structures such as the corticospinal or the optical tract should be reconstructed. To ensure clinical applicability, such a tool should also be easy to use with minimal interaction requirements. The tool presented in this work, aims towards reaching such a goal.
Contributions

The novel contributions of this work can be summarized as follows:

- Introduction of a similarity measure for diffusion tensors that allows for clustering of voxels with similar diffusion properties within a DTI image. The similarity measure combines the two most relevant properties of the diffusion tensor, namely the principal diffusion direction and the level of anisotropy. Inclusion of a weighting parameter makes it particularly useful for interactive segmentation approaches.

- Formulation of an efficient interaction concept that facilitates the generation of anatomically meaningful ROIs with a single mouse-click. The presented techniques significantly reduce the time and precision required to create a ROI in comparison to conventional drawing approaches. This becomes especially interesting in the context of interactive whole-brain fiber tracking, a technique for which sufficiently satisfying interaction techniques are currently an open question.

In the following section, the mathematical model behind the similarity estimation of the diffusion tensors is formulated. Furthermore, an intuitive concept for interaction with the tool is proposed. Subsequently, results from a multi-user evaluation study will be presented. Finally, potential benefits as well as limitations of the method are discussed.

3.2.2 Method

The general strategy for creating ROIs for DTI fiber tracking is to identify an area in the dataset that consists of voxels with similar diffusion direction and fractional anisotropy, corresponding to the location and orientation of an anatomically known white matter structure. In the color coded DTI representation commonly used in radiology to identify white matter structures, these two parameters correspond to the color and intensity of the overlay. For the method described herein, the color encoding becomes superfluous, since these two parameters can be extracted directly from the tensor itself. Consequently, the proposed method technically boils down to an interactive segmentation approach on tensor fields, based on the tensors main eigenvalue and eigenvector. Different approaches for segmentation of tensor fields have been proposed in the literature [59, 130]. However, these approaches mostly focus on the reconstruction of 3D-structures, while the aspect of interaction is disregarded.

In contrast, the proposed method explicitly focuses on an efficient interaction scheme. It is based on the concept of clustering voxels with similar diffusion
properties with respect to a user-defined reference position. Therefore, two steps are required: First, a reference point needs to be selected using a click with the mouse. This reference point defines the diffusion properties to which similar points are added in order to create the ROI. Afterwards, with the mouse button still pressed, moving the mouse along its x- and y-axis allows to interactively manipulate the two main parameters for the contouring algorithm. Depending on these parameters, a contour can be calculated and updated in real-time. Once the mouse-button is released, the contour is finalized.

To facilitate this, two similarity maps are calculated for the tensor field, based on the reference point defined interactively by the user. These are the angular similarity map $\varphi$, measuring the similarity in diffusion direction, and the magnitude similarity map $m$, measuring the similarity in fractional anisotropy. The angular similarity map is acquired by calculating the angle between each tensors main eigenvector and the reference direction, convolved with a Gaussian shaped envelope:
3.2. 2D-ROI generation from DTI data

Figure 3.2: Examples of different contours generated using different weighting parameters on the similarity map. All contours have been generated from the exactly same point-of-reference.

\[ \varphi(v(x); v_{\text{ref}}, \sigma_{\varphi}) = e^{-\frac{\arccos(v(x)^T v_{\text{ref}})^2}{\sigma_{\varphi}^2}} \] (3.1)

Here, \( v(x) \) defines the diffusion tensors largest eigenvector at any position \( x \), and \( v_{\text{ref}} \) defines the largest eigenvector of the reference point. The resulting angle is convolved with a Gaussian shaped function which maps all angles into the range \([0, 1]\).

Similarly, the magnitude similarity map \( m \) is acquired by calculating the absolute difference of the largest eigenvalues:

\[ m(\lambda(x); \lambda_{\text{ref}}, \sigma_m) = e^{-\frac{|\lambda(x) - \lambda_{\text{ref}}|^2}{\sigma_m^2}} \] (3.2)

In analogy to the angular similarity map, \( \lambda(x) \) defines the largest eigenvalue at position \( x \) and \( \lambda_{\text{ref}} \) defines the largest eigenvalue of the reference tensor. For both maps, \( \sigma_{\varphi} \) and \( \sigma_m \) control the width of the Gaussian envelope.

Finally, the weighted similarity map \( w \) is calculated by multiplying \( m \) and \( \varphi \), while interpolating the width of the Gaussian envelopes in such a way that it is either small for the angular and large for the magnitude similarity, vice-versa, or something in-between. This allows to continuously adjust the weighting between angular and magnitude similarity, thereby allowing to control the shape of the generated contour. Consequently, the blending parameter \( \sigma \) can be more intuitively described as a shape-parameter. The constants \( F_{\varphi} \) and \( F_m \) are used to scale the Gaussian envelopes to the individual domains of \( \varphi \) and \( m \).

\[ w(\varphi, m; v_{\text{ref}}, \lambda_{\text{ref}}, \sigma) = \varphi(v_{\text{ref}}, \sigma * F_{\varphi}) \times m(\lambda_{\text{ref}}, (1 - \sigma) * F_m) \] (3.3)

Figure 3.1 illustrates this on the concept of a simple, 2-dimensional, spherical DTI phantom. Once the similarity map has been calculated, a contour line can
be extracted using a marching-squares algorithm, which allows for generating a 2D iso-line inside an image for a given threshold $t_{iso}$. In theory, more than one iso-line may be found within an image, however for the proposed algorithm only the one which includes the reference position of the initial mouse-click is calculated. Since the similarity map covers a range between 0 and 1, the threshold for the iso-line needs to be within this range as well. A value close to 1 means that a very high similarity is required, which will likely result in small contours, while on the other hand a low value allows for less similarity, resulting in larger contours. Consequently, the threshold parameter for the marching-squares algorithm can also be interpreted as a size-parameter.

Since the result of this tool is a 2D-contour, all calculations can be carried out on a single two-dimensional slice. This greatly reduces computational complexity and allows for good interactive update rates on standard PC hardware. This slice is selected before generation of the contour. It can be either a standard orthogonal representation of the data or an arbitrarily oriented slice through the dataset. All subsequent calculations are carried out on this slice only. Figure 3.2 shows several contours that can be generated from a single reference position, using different values for $\sigma$ and $t_{iso}$. All contours have been created for the same reference point, using different combinations of threshold and similarity weight. The purpose of the phantom is to demonstrate how the underlying similarity criteria affect the shape of the resulting contours.

**Interaction scheme**

The algorithm relies on three parameters that need to be adjusted interactively by the user:

- A reference point for which the similarity measure shall be calculated.
- A weighting parameter $\sigma$, to adjust between angular and magnitude similarity.
- A similarity threshold $t_{iso}$, to define the required similarity level for the resulting ROI.

These three parameters can be nicely mapped onto an interaction scheme for a conventional computer mouse. The reference point is defined by clicking at the desired position in a 2D-viewer. The remaining two parameters can then be mapped to the horizontal and the vertical axis of the mouse. As long as the user keeps the mouse button pressed, the free parameters can be adjusted interactively. While doing so, the resulting ROI is updated and shown on the fly. The ROI is finalized once the mouse button is released. During this exploration phase, it is necessary to re-evaluate equations 3.1, 3.2 and 3.3. Also, with any change in parameters, a new
3.2. 2D-ROI generation from DTI data

contour needs to be calculated using the marching squares algorithm. However, keeping in mind that this operation only needs to be performed on a single 2D slice of the image, interactive performance is not an issue on current computing hardware.

3.2.3 Evaluation and Results

The method has been evaluated with respect to interaction time, as well as segmentation quality in terms of its ability to generate clinically feasible ROIs for DTI fiber tracking. For this, a user study was performed among three users. Participants were a radiologist and two computer scientists with moderate to extensive experience in DTI analysis. Participants were asked to outline a defined set of four different ROIs (three of them both for the left- and right hemisphere, giving a total of seven ROIs per dataset) from DTI data, using both a conventional freehand drawing tool, and the semi-automatic tool proposed in this work. Evaluation was performed on a subset of images from the MEISE study (cf. Section 2.4, [113]). DTI data was acquired on a Siemens Skyra 3T MR scanner (Siemens, Erlangen, Germany), scanning parameters were $2 \text{mm}^3$ voxel size, $\text{TR}=7600$, $\text{TE}=90$, 30 diffusion directions, two repetitions. For interaction purposes, the color-coded DTI data was overlaid onto the rigidly registered $T_1$-weighted image, acquired during the same scanning-session.

Figure 3.3 shows the regions used for this evaluation. These consist of:

- The internal capsule, left and right, on an axial plane. This is a relevant structure for reconstructing the corticospinal tract (CST) which passes through the internal capsule’s posterior limb. A common approach for CST reconstruction is to perform filtering at both the internal capsule and the spinal cord level. It can be identified as a blueish- to purple structure with a bean-like shape.

- The corpus callosum (CC) on the mid-sagittal plane. The CC is the central connecting fiber bundle connecting both the left- and the right hemispheres. It can be easily identified by its clear red color representation on a mid-sagittal plane.

- The superior longitudinal fasciculus (SLF), left and right, an association-bundle connecting functional areas from the temporal-, occipital-, parietal- and frontal-lobes. It appears green on a central, coronal plane, running more or less orthogonal to the CST.

- The cingulum, left and right. Another association fiber that is part of the limbic system. It can be identified as two green spots lying directly on top of the CC.
Chapter 3. Interactive tools for identification and delineation of brain structures

Figure 3.3: Cross-sections of the regions used for the evaluation. The indicated regions are commonly used for reconstruction of the described pathways: (a) The corticospinal tract at the posterior part of the internal capsule. (b) The corpus callosum on the mid-sagittal plane. (b) The superior longitudinal fasciculus and the cingulum on a coronal cut.

The study setting was defined as follows: Each participant was asked to first draw a contour for each structure manually, using a simple freehand drawing tool without any "intelligent" drawing support. Afterwards, the drawn contour was rated on a scale from 1 to 6, mimicking German school grades. Since the manual drawing defines the reference in terms of contour quality, a grade worse than 2 was considered unacceptable, and the contour was discarded and completely redrawn. Subsequently, the same contour was generated using the semi-automatic ROI tool presented here. Again, interaction-time, as well as the number of re-tries was recorded. Also, users were again asked to rate the resulting contour on the same scale as previously. Here of course, any grade was considered valid.

The evaluation itself was performed on 20 randomly selected datasets. In total, each user had to delineate 140 structures both manually, and interactively. The results are plotted in Figures 3.4 and 3.5. Figure 3.4 gives a comparison of the distribution of the time in seconds required for contouring each structure, by each participant. Structures that were outlined on each hemisphere are grouped together, since no significant difference for these structures was expected.

What strikes out, is the massive difference in contouring time taken for the corpus callosum. Here, two unrelated effects contribute to this observation. First, due to its large size and relatively complex geometry, the corpus callosum is difficult to outline manually. Especially if a precise delineation is required, manual drawing has to performed carefully and meticulously, which results in relative high interaction times observed for all three raters using the freehand tool. In contrast, the CC is particularly simple to outline for the semi-automatic approach, since it consists of
mostly identically oriented diffusion vectors while at the same time being clearly separated from neighboring structures such as the cingulum. As a result, the CC has often been segmented perfectly with a single click, without any need for parameter adaption.

For all other structures, a significant reduction in interaction time can be observed as well, although not nearly as prominent as for the CC. Comparing the three users with each other, no relevant difference can be found. In addition to these absolute values, Figure 3.5 shows the relative differences measured for each structure. Here, any value larger than 1 corresponds to an acceleration gained by the novel tool over the freehand tool. Again, the massive acceleration for the CC becomes obvious. However, it can also seen more clearly now how the interactive tool performed for the other structures. For users A and C, we find a median acceleration of 2 for each structure, with nearly no cases of deceleration. For user B, the effect was...
Chapter 3. Interactive tools for identification and delineation of brain structures

<table>
<thead>
<tr>
<th></th>
<th>CC</th>
<th>CST</th>
<th>Cingulum</th>
<th>SLF</th>
</tr>
</thead>
<tbody>
<tr>
<td>User A</td>
<td>8.38</td>
<td>2.04</td>
<td>2.43</td>
<td>1.95</td>
</tr>
<tr>
<td>User B</td>
<td>12.46</td>
<td>1.39</td>
<td>1.67</td>
<td>1.43</td>
</tr>
<tr>
<td>User C</td>
<td>9.05</td>
<td>2.24</td>
<td>2.12</td>
<td>1.99</td>
</tr>
</tbody>
</table>

Table 3.1: Median acceleration factor grouped by user and structure.

Table 3.1 lists the median acceleration factors for each user and structure.

Another interesting aspect to look at is how the different users rated the semi-automatically generated structures in comparison to the manually drawn. To assess this, the difference in the given scores has been analyzed. Table 3.2 groups these results for each structure and user. The comparison was performed by just looking at whether the two contours generated per structure received the same score, or whether one or the other was rated better. It can be seen that especially for the corpus callosum and for the cingulum, the large majority of contours generated with the novel tool was rated better than the manually drawn one. The CST and the SLF was rated equally good segmented for a majority of contours by all raters. The only case, where one structure was rated worse for a significant number of samples was the CST as rated by user A. Considering that this was only observed for a single user in this clarity, allows to draw the conclusion that a subjective assessment may be a relevant factor. Nonetheless, it also shows that the spectrum of possible contours that can be generated by the interactive tool is limited to what can be reconstructed based on the tensor-similarity assumption. In consequence, there may be cases where the tool will inevitably fail to produce a specific contour.

<table>
<thead>
<tr>
<th></th>
<th>CC</th>
<th>CST</th>
<th>Cingulum</th>
<th>SLF</th>
</tr>
</thead>
<tbody>
<tr>
<td>User</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>Better</td>
<td>12</td>
<td>20</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>Equal</td>
<td>6</td>
<td>0</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>Worse</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 3.2: Assessment of segmentation quality, grouped by user and structure. Numbers represent the number of contours where the semi-automatically generated contour was rated better, equal or worse than the manually drawn one.

3.2.4 Discussion

The key idea of the algorithm described here is that ROIs for performing fiber tracking on DTI data are typically placed around clusters of voxels exhibiting similar diffusion properties. When reconstructing fiber tracts using tractography
Figure 3.5: This plot shows the distribution of time ratios required for generating each contour, i.e. the fraction of time required by the semi-automatic tool with respect to the manual freehand tool. Any value larger than 1 corresponds to an acceleration achieved.

algorithms, radiologists typically choose anatomical regions for which the principal direction of the axonal structures passing are known. The directional color coding used in DTI visualizations is the key to identify such regions. The fractional anisotropy which relates to the intensity of the color is used to identify the extent of the bundle in question. Consequently, anatomically meaningful ROIs will usually be placed around blobs of a certain color associated with a known white matter structure. As long as the generation of ROIs is guided by these two properties, the proposed algorithm is capable of generating useful ROIs with a minimum of interaction effort. In comparison to manual delineations the novel technique is also significantly more accurate, since the generated contours are intrinsically bound to voxels with similar diffusion properties. This is especially helpful when defining ROIs in the vicinity of orthogonally touching fiber bundles, such as e.g. the corpus callosum and the cingulum, as shown during the evaluation performed.

A critical step of the algorithm is the placement of the reference position. Ideally, this should be chosen as the average direction of the fiber bundle at the given location. However, in practice this position may not always be easily determined.
As a result, it may happen that the algorithm fails to produce the ROI that the user has in mind. The most pragmatic way to deal with this would be to repeatedly place the reference position and discard the contour if the initial click was located improperly. An alternative would be to adjust the position of the reference point iteratively based on average values of intermediate similarity clusters.

Another point that needs to be kept in mind is that the above mentioned assumptions for similarity criteria do not always hold. In pathological cases like brain tumors, fiber bundles may be significantly displaced and deformed. In such situations it might become necessary to include a broader range of fiber directions within the ROI. For such cases, a fall-back strategy to a fully manual drawing tool or interactive refinement of generated contours should be provided within a DTI software package.

\section*{3.3 3D Gyrus segmentation}

The cortical surface of the human brain is a highly convolved structure consisting of gyri and sulci. Both in neuroscience, and also in clinical neurology, the gyri are of special interest, as they are considered to be the functional modules of the cortex. Many gyri are associated with well identified functions, such as movement control, sensing, hearing, language processing and word generation, or the processing of the visual system. Although the precise pattern and distribution of these functional areas varies considerably across individual humans, many of the low-level functions can typically be located reliably within individual subjects. Prominent examples for such gyri are e.g. the primary-motor cortex, located on the pre-central gyrus, or the somatosensory cortex on the post-central gyrus.

There is a natural interest in the development of techniques facilitating the delineation and subsequent identification of individual gyri. In neuroscience, such techniques are required e.g. for mapping results of functional experiments to anatomical structures, in the context of brain connectivity mapping and analysis of functional networks. In clinical applications, it is often required to quantify parameters such as cortical thickness or local cortical volume for brain structures involved with certain pathologies. Focal cortical dysplasia, a common cause of epilepsy manifests in local thickening of single gyri. In neurosurgery, understanding the interplay between white matter fiber tracts and functional gray matter areas in the vicinity of a tumor can by crucial for choosing the right treatment approach.

The classical approach to identification of functional areas in the brain goes back to the parcellation proposed by Korbinian Brodmann in 1909 who suggested to divide the brain into 52 areas \cite{brodmann1909}. His parcellation that has later been refined to include further areas, still builds the basis to localizing functional areas of the brain in current neuroanatomy. It also builds the basis for several digital brain...
atlases. In order to be able to associate Brodmann’s areas to e.g. an MRI scan of an individual brain, it is necessary to map such an atlas to the image of the brain, and to adjust the mapping to the individual parcellation of that brain. The process thus consists of two mostly independent steps, the parcellation, and the mapping. The work presented here focuses solely on the parcellation of the brain, forming the basis for e.g. manual labeling of functional areas, or the creation or adaption of digital atlases.

3.3.1 Related work

In recent years, different approaches for automatic parcellation of the brain into either sulcal basins, or into individual gyri have been presented. Lohmann et al. [79] proposed a method to calculate sulcal basins from a $T_1$-weighted MRI image using a sequence of morphologic operations. Their method is computational efficient and allows for fully automated parcellation based purely on morphologic properties of the input data. Fischl et al. [37] proposed a surface based approach to parcellation. Building upon a computationally expensive explicit surface representation of the cortical surface, they register a pre-generated sulcal atlas on the topologically spherical cortical surface. The registration is guided by local curvature, representing sulcal basins and gyral folds. Li et al. [75] presented an atlas-free approach which also operates on the reconstructed cortical surface. They apply a tracking algorithm on a principal direction flow field to obtain a geometrically consistent sulcal parcellation. Also operating on a given surface representation and a set of sulcal basin lines, Cachia et al. [13] proposed a method cortical labeling using geodesic Voronoi diagrams. They claim to achieve improved correspondence to actual anatomic structures.

Another general approach to gyral labeling relates to using atlas based segmentations. Attributing to the large anatomical variations of the brain over individual subjects, multi-atlas approaches have been shown to outperform single atlas techniques. Two recent methods utilizing such techniques have been developed by Wang et al. [54] and Wu et al. [143]. Recently, fully convolutional networks have also been trained and applied to the task [32]. Depending on available training data and specific network architectures, highly accurate labelings have been achieved.

In general, these methods have in common that they make a number of assumptions on the input data, in favor of achieving a fully automated result. While especially newer methods reported qualitatively accurate results, none of the methods above could potentially be embedded into current clinical workflows, considering the amount of processing time and partially also the need for expert supervision and correction of results. In contrast, the approach presented here is an interactive method for the segmentation of individual gyri focusing on robustness and interaction efficiency, at the cost of automation. This makes this technique suitable
Chapter 3. Interactive tools for identification and delineation of brain structures

for application in actual clinical environments, where robustness to a wide variety of input data is typically a more critical requirement as opposed to automation. Also, it is computationally fast and does not rely on time consuming pre-processing. It works directly on the volumetric image data and does not require and explicit reconstruction of the brains surface. Gyri are defined interactively by drawing contours on a 3D volume-rendering of the brain. From these contours, corresponding gyri are then derived by clustering all voxels connected to the same fraction of white matter tissue.

3.3.2 Methods

The proposed method exploits one of the most fundamental properties of the anatomical structure of the brain: the connectivity between cortical gray matter and white matter of the brain. Cortical gray matter forms a ribbon of more or less constant thickness around the white matter. Anatomically, this is represented by a specific layered structure of cells in the cortex, which at the lowest layer connects to the axons of the white matter. In an MR image of the brain, this can be modeled in a simplified manner by identifying for any voxel of the cortical gray matter the closest corresponding white matter voxel. This mapping can further be improved, by constraining the distance measure used in such a way as to prevent skipping neighboring gray matter structures. To facilitate a mapping between these sets of corresponding voxels, a gray-matter association map (GMA map) is computed as a first step. With this map available, connected clusters of gray-matter voxels can be identified by clustering smaller areas of white matter voxels, and consequently labeling all gray matter voxels attached.

The underlying concept of this approach is based on the observation, that on a macroscopic scale in the brain all gray matter is somewhere connected to white matter, and that consequently for every gray matter voxel, it is possible to identify a corresponding white matter voxel to which it is most closely connected. This idea is exploited, by performing a segmentation that first identifies all associated WM voxels for a small sample of GM voxels, and then clusters all GM voxels attached directly and indirectly to the WM voxels in proximity to those initially selected.

Pre-requisites

As input data, this method requires a segmentation of the brain from an MR image, as well as a classification of all brain voxels into its primary tissue classes, white matter (WM), gray matter (GM), or cerebrospinal fluid (CSF). Both of these tasks have been studied extensively in the past, and a number of approaches for solving them are available. Here, an interactive approach for the segmentation of the brain using a watershed transformation is used, followed by a correction for MR intensity.
in inhomogeneities using the N3 algorithm [118] and a tissue classification step using multi Gaussian mixture-modeling [48].

**Gray matter association map**

From the maximum likelihood classification of the brain, the largest set of connected white matter voxels is computed using a connected component analysis. Then, the GMA map is calculated, which encodes for every GM voxel inside the brain mask the location of the WM voxel to which it is connected. To do so, as a first step a unique index value is assigned to each border voxel within the WM mask. This index will later be used to address a look-up table associating indices to WM voxel coordinates. This in consequence allows to locate any WM voxel spatially using just its ID. Subsequently, these WM indices are propagated over all voxels classified as either pure- or partial- gray matter. Figure 3.6 illustrated this using a color-coded representation for the indexes. After the propagation, each GM voxel can be associated with exactly one of the WM surface indices.

A Euclidean distance transformation with its corresponding Voronoi division on the WM index image would be a simple way to compute an approximation of the GMA map. This, however, would lead to a mapping based purely on the Euclidean distance, which might be invalid for many configurations of the cortex. Consider a configuration where two neighboring gyri are of significantly different thickness. The Voronoi division, based on an Euclidean distance measure would separate these
two gyri in their exact middle, effectively cutting part of the thicker-gyrus off and attaching it to the thinner.

To overcome this problem, it is beneficial to analyze the intensity profile found in such locations (cf. Figure 3.7). It can be seen that the intensity drops continuously the further it goes away from the white matter, down to a local minimum where the two gyri are in contact. From there on, intensity rises again continuously, until reaching the opposing white matter. Consequently, the separation problem can be solved by propagating WM indices in an iterative manner using a weighted function of WM distance and voxel intensity. More precisely, WM-indices are iteratively propagated into a subset of the GM voxels until all GM voxels have been processed. This subset is formed as a selection of voxels being both

1. adjacent to the current front of voxel propagation, e.g. all voxels neighboring either WM voxels, or neighboring GM voxels added in a previous iteration
2. and an element of the 10% brightest quantile of all unprocessed voxels. This causes the algorithm to always process brighter voxels prior to darker ones.

The inclusion of the brightness criteria essentially prevents the swapping of adjacent gyri. The 10% quantile has been chosen empirically and is essentially used for performance optimization. A lower value would lead to a higher number of iterations, while a higher number might cause darker voxels to be added too quickly to the WM component. This approach guarantees that brighter voxels are always added prior to darker ones, preventing the propagation over local minima as described in the example above. The algorithm terminates as soon as all GM labeled voxels have been processed. Figure 3.7 illustrates this. The whole map can be calculated in less than one minute on current generation standard PC hardware.
3.3. 3D Gyrus segmentation

Figure 3.8: Top-left: Initial contour drawn on the brains surface. Top-right: Resulting gyrus segmentation. Bottom-left: Volume rendering of the WM-mask only. The voxels corresponding to the selected gyrus are highlighted. Bottom-right: Results of the constrained dilation after 5 iterations.

Interactive gyrus definition

The GMA-map provides a bijective mapping function between WM- and GM-voxels, i.e. it allows

- for any GM-voxel: to identify the associated WM-voxel and
- for any WM-voxel: to identify the set of GM-voxels associated to it.

Mapping between corresponding voxels can be achieved efficiently by selecting labels from the GMA map. Each label within the gray-matter area of the input image, will include exactly one voxel in the white matter. One voxel of white matter will typically include a set of gray-matter voxels. To use this map for parcellation
of individual gyri, one can now simply cluster voxels on the surface of the white matter blob, and combine all included indices to a single label in the result image. This facilitates the following workflow: The user can depict a gyrus, by delineating a small sample of GM voxels. This can be achieved by e.g. drawing a 3D contour on a volume-rendering representation of the brain, or by marking a portion of a gyrus in a conventional 2D rendering of the original MRI data. This defines a set of GM voxels to be used as input for the clustering, which is done using a three-step approach:

1. First, the set of input-voxels is used to lookup all included WM voxels from the GMA-map. The results defines the set of WM voxels in closest proximity to the sample of GM voxels given for the selected gyrus. However, this will typically not cover the complete gyrus, but rather the footprint of the drawn contour on the white matter.

2. Next, an iterative constrained morphologic dilation step is performed, which is achieved by iteratively dilating the seed voxels with a $3 \times 3 \times 3$ kernel, and subsequently masking the result with the WM component, to prevent inclusion of WM voxels from neighboring structures. This yields a local segmentation of the white-matter portion of the gyrus. The number of iterations required depends on the input image resolution and chosen dilation kernel size. It relates to the average distance of neighboring gyri. For a typical input resolution of $1\text{mm}^3$, five iterations are suggested. Figure 3.8 illustrates this process.

3. Finally, a back-projection of all selected WM voxels over the segmented brain is calculated using the lookup-table of the GMA-map. This clusters all voxels connected to the interactively selected section of WM, thus identifying all voxels belonging to the selected gyrus. The whole process takes about 1-2 seconds to compute on a current generation consumer PC with Intel Core-i7 CPU @ 2.6 GHz, depending on the size of the selected gyrus.

The resulting segmentation includes all gray-matter voxels corresponding to the segmented portion of the underlying white matter blob. This provides the basis for labeling either individual gyri of interest, or a whole brain.

### 3.3.3 Results

This algorithm has been implemented in a prototypical software tool in MeVisLab. A workflow has been designed that allows to load a $T_1$-weighted MR image from DICOM data and perform a watershed based interactive segmentation of the brain. During brain segmentation, both hemispheres as well as the cerebellum are labeled.
3.3. 3D Gyrus segmentation

Figure 3.9: Top: An annotated drawing of the human brain, as seen from the outside. Taken from the classical medical textbook *Anatomy of the Human Body* by Henry Gray [43], License: Public domain. Bottom: The same annotation applied to a segmented brain from a T1-weighted MRI scan, using the method proposed in this work.

individually, to allow for individual treatment per hemisphere. Subsequently, the image is corrected for intensity inhomogeneities using the N3 algorithm [118]. Tissue classification is performed on the corrected image, and the resulting maximum likelihood maps are used as input to the algorithm. The interactive part can typically be performed in a few minutes. This is followed by computation of the GMA-map, which depending on the size of the input segmentation requires between 5-15 seconds. As a result, the proposed method is easily applicable also in a clinical environment, where both manual processing, as well as automated computation times are commonly encountered constraints.

The segmented brain is then displayed in a 3D direct volume renderer, as well as on a slice-based standard 2D image viewer. In the 3D viewer, the user is
Chapter 3. Interactive tools for identification and delineation of brain structures

Figure 3.10: Top: An annotated drawing of the human cerebrum, as seen from the inside. Taken from the classical medical textbook *Anatomy of the Human Body* by Henry Gray [43], License: Public domain. Bottom: The same annotation applied to a segmented brain from a T1-weighted MRI scan, using the method proposed in this work.

provided with a drawing tool, that allows to depict individual gyri by clicking on the desired position. Since a direct volume renderer is used, the picking operation can incorporate the customly adapted transfer function, thereby allowing to control the depth of markers placed. A double click finalizes the current selection. After each new landmark defined, the GM-clustering is re-triggered in a background process. Finally, the resulting segmentation can be assigned a name and stored.

The method has been evaluated by labeling a complete MR scan of a volunteer based on a basic anatomical description of mayor gyri. Figures 3.9 and 3.10 show both the template, taken from the classic medical textbook *Anatomy of the Human Body* by Henry Gray [43], and the resulting labeling using the method presented here. Interaction time was approximately 10 minutes per hemisphere. The main
difficulty was the identification of corresponding gyri. Once a gyrus had been identified, delineation took only a few seconds for locating proper landmarks.

3.3.4 Discussion

The proposed method has been shown to be capable of robustly segmenting individual gyri on a given dataset. The method exploits the intrinsic property of gray- and white-matter connectivity of the brain, and the way these tissue types are represented in T1-weighted MR images. As pre-processing it requires an accurate brain segmentation, along with a tissue classification. Both pre-requisites are relatively easy to compute, as they have been studied extensively in the past. With this provided, the method facilitates user-controlled interactive segmentation of gyri with on-the-fly updates of segmentation results. Potential applications of such a tool could be e.g. for generating cortex-based seed-regions for DTI fiber tracking, for mapping fMRI results to individual gyri, or for the measurement of cortical thickness of selected structures. A potentially interesting application would be in the field of neurosurgery. Here, however, the assumption of tissue connectivity might be violated, depending on the location and type of a tumor. Another potential limitation in this field could be the requirement of a non-contrast enhanced image of the brain, which is sometimes not acquired in favor of saving total time in the MR session. In neurology, however, the tool could be easily integrated into workflows specifically analyzing properties of certain structures of interest. Here, it provides a time-efficient and robust alternative to complex, template based method of whole brain labeling.

3.4 Conclusion

In this chapter, two approaches for interactive segmentation of clinically relevant structures based on MRI data were presented. Although technically unrelated, both approaches share the same underlying philosophy: Instead of aiming at full automation, they rather rely on combining the users capability of identifying structures of interest with the computer’s ability to perform segmentation or clustering based on precisely quantifiable properties, such as geometric similarities or distance based measures. As a result, both tools provide highly interactive control over the offered task, while automating the process in a meaningful manner. As such, they can be considered as pragmatic solutions to complex tasks, that by themselves provide basic input to a range of clinically relevant problems. Instead of relying on complex assumptions for a given problem, they are based on robust engineering with a focus on practical applicability.
The first method targets generation of regions-of-interest for DTI fiber tracking. It exploits the idea that a meaningful ROI should include points with similar diffusion properties, thereby allowing to depict anatomically meaningful axonal structures. A simple to use interaction scheme has been proposed, which allows the user to control both the shape and the size of the generated ROI. The tool has been evaluated in a small user-study, where both the interaction time, as well as the quality of the resulting ROIs has been assessed. All in all, the tool outperformed a manual freehand tool in both aspects, speed, and accuracy, providing both a better reproducibility and a reduced interaction effort at the same time. Such an interaction tool appears highly promising, since it allows improved exploration and interaction with DTI data. In combination with real-time whole brain fiber tracking, it may become a vital element in strengthening the role of DTI in clinical routine.

The second method addressed the task of delineating individual gyri of the human brain. Here, a novel method for performing fast and intuitive interactive segmentations was developed. The approach is based on clustering of GM voxels connected to a single WM structure. It consists of a robust processing pipeline composed from standard morphological image processing operators, making only a single assumption on the structure of the input data: the connectivity between white matter and gray matter voxels. As such it behaves very robust even on input data of heterogeneous quality. In contrast to previously presented techniques, the method does not require an explicit reconstruction of the brains surface, which is computationally expensive to compute in a robust fashion. The method favors user interaction over automation. Thus, it provides a tool that allows for robust segmentation of structures of interest, based solely on geometrical properties while leaving the knowledge about the actual parcellation of structures in question in the hands of the (expert-) user.

Related publications as primary author

The methods presented in this chapter have been presented a three international conferences. The corresponding papers are:

3.4. Conclusion

All methods presented here have been developed independently by myself. Dr. Jan Klein and Dr. med. Benjamin Geisler assisted in the user study for evaluation of the interactive DTI ROI tool.
4 Interactive Multi-Volume Visualization for Planning of Cerebral AVM Surgery

They always say time changes things, but you actually have to change them yourself.

(Andy Warhol)

4.1 Introduction

As a third major topic of how computer assistance can support the process of interpretation and understanding of medical images, visualization is addressed. Computer Visualization is a huge and complex topic that nowadays has become an indispensable component in nearly every area of human life, be it entertainment, education, industrial design, and of course, medicine. As such, a tremendous amount of research has been invested into this field, with constant progress achieved since
the early days in the 1980ies. In the field of medical imaging, dedicated visualization techniques have evolved addressing specific questions relating to either the image acquisition used (e.g. MRI, CT or US), or the clinical question driving the task (e.g. surgical planning, virtual endoscopy, blood-flow analysis, DTI fiber tracking, and many more). For many tasks, advanced visualization techniques have matured to a degree where they have been integrated into medical workstations in radiology departments. Nonetheless, there remains a large number of challenges associated with applicability of advanced visualization techniques in clinical routine, and in consequence, the impact of many techniques to everyday clinical routine is in many areas limited. This chapter will address this topic, by demonstrating how interactive multi-volume visualization can significantly support the process of presurgical planning, demonstrated on the clinical issue of treatment planning for cerebral arterio-venous malformations (AVMs). Evaluation has been performed by means of three clinical case-studies of retrospectively analyzed cases.

4.2 Medical background

Arteriovenous malformations (AVM) of the brain are vascular disorders that affect roughly 0.01 - 0.5% of the population. They are characterized by the presence of direct connections between arteries and veins which bypass the capillary bed normally responsible for deoxygenation of arterial blood and the transfer into the venous system. This vascular short-circuit is comprised of numerous coiled and convolved connections, that form the nidus of the lesion. The nidus may be fed by one or more arteries, the feeders, and drains into one or more veins. To complicate things, additional ”en passage” vessels may be involved, supplying both the AVM as well as important functional areas of the brain. Ligation of such feeders would cause postoperative neurological deficits and must therefore be avoided. Clinical staging of these lesions is obtained following the Spetzler-Martin scale [121], which grades the severity of an AVM on a scale form 1 – 5. Criteria for grading are the size, proximity to eloquent functional areas, and the type of venous drainage present.

Treatment options for AVMs include stereotactic radiotherapy, endovascular embolisation, and microsurgical excision. For any of these therapeutic approaches, a precise understanding of the underlying angioarchitecture is a necessary prerequisite. Especially for a neurosurgical approach, it is crucial to identify all feeding arterial branches beside the draining veins. During surgery, the first step after locating the nidus is to ligate the main feeding arteries. Afterwards, the nidus can be carefully dissected from the surrounding brain tissue before the draining veins can be ligated and the nidus removed. Failure to identify the feeders or ligation of the draining veins prior to obliteration of all feeders may cause rupture of the nidus, resulting
in intraoperative hemorrhage and brain swelling, which may lead to serious patient impairment. Ligation of “en-passage” arteries would cause insufficient oxygen supply of connected parts of the brain accompanied by the risk of postoperative neurological deficits.

Considering these conditions, it becomes apparent that a complete understanding of the structure of the AVM prior to surgery is crucial for successful dissection. This, however, is a challenging task. First, there is no single imaging modality that could provide all required information about the neurovascular structures. Instead, a number of techniques exist that are capable of identifying individual parts of the relevant anatomy. Second, it is difficult to create a mental 3D model of the complex vascular configurations found in cerebral AVMs. At this point, the conventional 2D slicewise representation of tomographic images performs particularly poor in supporting the neurosurgeon with the generation of such a mental 3D model.

To overcome these difficulties, a computational framework is presented that allows for interactive exploration of multiple tomographic images in a 3D multi-volume rendering system. Images derived from MRI-sequences are combined and preprocessed in a semi-automatic manner. The resulting datasets can be visualized simultaneously using individual look-up tables and clipping planes per dataset. A volume renderer that implements a modular shader concept which allows for application of individual shading effects like illumination, clipping, or other custom effects per volume is utilized. Also, each volume can be of arbitrary voxel size and can be transformed by the renderer at runtime. The method has been integrated into a prototypical medical software assistant that has been installed in the 3D imaging laboratory at Lahey Clinic Medical Center in Burlington, MA, USA. As a demonstration of the potential usefulness of this method, results are presented from a retrospective analysis of the planning for three clinical cases that underwent microsurgical AVM resection at Lahey clinic and correlate them with intraoperative findings.

This work has been strongly motivated by real-world clinical requirements. Considering the rare occurrence of cerebral AVMs in conjunction with the substantial technical challenge posed to the intervening neurosurgeon, the tremendous potential of presurgical support by means of a 3D patient-individual data workup becomes apparent. In the software prototype, state-of-the-art visualization methods are carefully combined with new visualization and interaction techniques that were requested by medical experts to facilitate intuitive exploration of multimodal volume data. Implementation was done in close collaboration and active discussions with neurosurgeons and neuroradiologists. Consequently, techniques were chosen that enable the visualization of complex medical data with simple interaction requirements focusing on the ability to quickly recognize and identify critical structures of interest.
Chapter 4. Interactive Multi-Volume Visualization for Planning of Cerebral AVM Surgery

The main contributions of this work are:

- **Pre-processing pipeline for data analysis:** An image processing pipeline is proposed, that prepares datasets from a routine clinical AVM imaging protocol to be used in an interactive multi-volume rendering application in a stream-lined, semi-automatic fashion. The focus is on displaying vascular structures with the ultimate goal of allowing a perceptually easy yet accurate identification of arteries and veins.

- **Multi-volume visualization of vascular structures:** A multi-volume rendering pipeline is proposed, that allows for rendering of an arbitrary number of datasets, each with individual shading and clipping options. It achieves interactive framerates on current-generation GPUs using the clinical data commonly used in preoperative AVM surgery planning.

- **Demonstration of how computer-based visualization solves real-world medical challenges:** The framework has been integrated into a prototypical medical software assistant that has been installed within the radiology department of the hospital. Planning based on preoperative imaging data was retrospectively compared with intraoperative findings in three cases. Results of this evaluation are presented at the end of this chapter.

4.3 Imaging Background

The key challenge in preoperative planning of microsurgical AVM resection is the identification of the primary feeding arteries and draining veins. A number of imaging techniques can be utilized to support this task. However, no single modality is capable of answering all relevant clinical questions at the same time. Therefore, the standard clinical imaging protocol typically comprises a combination of several tomographic and projection-based image acquisitions. The highest level of detail can be achieved with a digital subtraction angiography (DSA), in which a contrast agent is injected directly into the affected arterial branch through an endovascular catheter. During the inflow of contrast agent, X-ray projection images are acquired sequentially and subtracted from an initial reference image. This allows dynamic imaging of the distribution of the contrast agent through cerebral vasculature at very high spatial and temporal resolution. Although DSA is capable of displaying vessels much smaller than any other current imaging technique, it comes with the huge drawback of being based on projections, which cannot easily acquire and portray depth information of the visualized structures. Therefore, DSA imaging is not suited to be incorporated within this work.
Contrast-enhanced CT angiography (CTA) is a tomographic technique capable of generating high-resolution 3D images with voxel size in the order of 0.3mm$^3$ and below. CTA of the head delivers high contrast of vascular structures compared to background tissue or partly enhanced brain parenchyma and surrounding structures. In combination with the high spatial image resolution, it is capable of displaying both the arterial and the venous tree with a high level of detail. CTA images can potentially be used for volume rendering of the vascular tree, however, the skull bone which overlaps with the intensity range of the vessels occludes the view inside the brain. Consequently, further image processing steps for removal of the bone become mandatory prior to visualization.

Magnetic Resonance Imaging provides a variety of different contrasts for displaying intra-cerebral structures. Without doubt, it is the most flexible and powerful imaging technique for the brain and has become an essential tool for both neuroradiology and neurosurgery. The primary MRI imaging technique for displaying vascular structures that is currently established in clinical routine is the time-of-flight (TOF) angiography, which can be used to selectively visualize arteries or veins depending on the applied saturation pulse and parameterization. It is based on the measurement of blood flow velocities and requires no contrast agent. An arterial TOF can be used to acquire images with a voxel size in the order of 0.5 – 1mm$^3$, which allows visualizing relatively thin vessels compared to other MR-angiography sequences. Venous TOF images require different acquisition parameters due to the lower flow velocities of venous blood. Typically, a 2D acquisition scheme is used resulting in increased voxel size in the order of 1.5mm$^3$. Consequently, it is most useful for displaying larger veins. The venous TOF is also referred to as MR Venography (MRV). Although the different TOF sequences predominantly display either the arterial or the venous vascular tree, one needs to be aware of the fact that a certain signal overlap is found in both images. An arterial TOF may contain parts of the venous vessel tree, and vice versa.

Besides TOF MR-angiography, contrast-enhanced T$_1$-weighted imaging (CE-T$_1$) can be used to display cerebral vessels. Technically speaking, the same MRI sequence as for conventional T$_1$-weighted imaging is used, after intravenous injection of a paramagnetic contrast agent. As a result, CE-T$_1$ images exhibit contrast characteristics similar to native T$_1$-weighted images with additional high intensity levels in vascular structures. For both native as well as CE-T$_1$ images, a 3D acquisition scheme can be used, which allows acquiring images with an isotropic voxel size of 1mm$^3$. In neuroradiology, such images are used to sensitize disorders of the blood brain barrier, tumor pathologies, as well as vascular disorders like aneurysms or AVMs. However, due to the high contrast of non-vascular tissues, they cannot be compared directly with conventional angiography images. Some image preprocessing steps become necessary before they may be used in MIP or
direct volume renderings useful for showing vessels.

For this work, the goal was to provide an analysis tool that allows combining the various angiographic imaging techniques in a flexible manner, in order to facilitate an understanding of the underlying anatomy in the best possible way. The idea is to allow simultaneous volumetric renderings of the different datasets to provide the neurosurgeon with an efficient way to combine and integrate the information obtained from the individual datasets, without the difficulty of performing a mental fusion of the complex and partly redundant information. This is achieved by presenting different datasets simultaneously using direct volume rendering with individually colored transfer functions. Explicit image segmentation for reconstruction of vascular trees is avoided in favor of presenting the datasets in a mostly unmodified manner, thereby avoiding the risk of discarding information from the images that would be too small to be detected by existing segmentation algorithms. Instead, the transfer function for each dataset can be manipulated interactively using a well-established window/level function. Thus, the final interpretation of structures visible inside each image lies within the responsibility of the medical user. Following this approach, the risk of image misinterpretation caused by limitations of image processing algorithms is reduced.

4.4 Related Work

Multimodal visualization for surgical planning has become a field of intense research in recent years. A thorough introduction to the field can be found in [104] and [105]. A recent survey paper by Lawonn et al. gives a detailed overview of the current state-of-the-art in the field [74]. Here, a brief overview of a number of relevant approaches with respect to neurosurgical planning is summarized.

Beyer et al. [7] introduced an application for neurosurgical planning and presented methods for the visualization of multimodal data for neurosurgery planning. To remove the cranial bone, skull peeling was proposed as an extension of the opacity peeling algorithm introduced by Rezk-Salama et al. [107] using registered CT and MRI data. In [56], methods for illustrative visualization of fMRI data combined with anatomy of the brain are presented. Rößler et al. [112] described a multi-volume framework for the visualization of functional brain images using the graphics hardware to allow interactive visual exploration of the data. Hong et al. [53] presented compositing functions for multimodal volume fusion. Interactive visualization techniques for combining multimodal datasets, such as functional data, are discussed in [9] to assist neurosurgical planning. Köhn et al. presented an application for neurosurgical planning and assessing of risk structures in [69]. In their work, they visualize vessels as well as functional data such as fMRI activation areas and fiber tracts from DTI. Rieder et al. presented in [109] a prototype
for visualization of multimodal data for neurosurgical tumor treatment. They describe methods to enhance important functional data and visualize these data combined with anatomical data along a virtual access path. [110] addresses the issue that fused multimodal visualizations typically complicate the recognition of anatomical details of the brain and pathological tissue at the same time without loss of information. Furthermore, in the volume rendering, important structures as well as suspicious high-intensity signals from multiple sequences are enhanced using a fuzzy clustering approach. In 2010, the IEE Visualization Contest addressed the question of multimodal visualization for neurosurgical planning. The winning entry by Diepenbrock et al. [29] effectively combined direct and indirect volume rendering with a tools for access-path planning and a tumor map for risk assessment of a chosen path.

Interactive clipping techniques are powerful and general utilities for the visualization of volume data [138, 139]. Manssour et al. [85] verified that cutting tools can reveal additional diagnostic information for multimodal volume rendering. Wu et al. [144] presented fast parallel slice cutting and partial exposing algorithms which are used for real-time volume rendering to support image-guided surgery and therapy planning. Tappenbeck et al. [123] presented distance as a second dimension for transfer function definition. They compute a distance volume by slice-based selection of distance ranges. Zhou et al. [57] presented focal region-guided enhancement of features in the volume rendering. In [145], the proposed method is extended by a distance measure as the main factor to control the volume features in the context region. Lindholm et al. [76] proposed a novel approach for boundary aware object reconstructions, allowing precise differentiation between adjacent structures using a single lookup-table.

In the field of intracranial aneurysm visualization, Higuera et al. [128] make use of a 2D transfer functions based on measured values and gradient magnitudes extracted from the CTA data to enhance the 3D visualization of intracranial aneurysms. Nishihara et al. [97] reported in a clinical study the usefulness of volume-rendered visualizations using computed tomographic angiography for surgical planning of aneurysms. Joshi et al. [60] proposed a vesselness coefficient incorporated within the classification stage of the rendering pipeline, which in contrast to pure vesselness filtering enhances visibility at branching structures. Bullit et al. [12] used vessel trees to visualize the relationship between an AVM lesion in three dimensions. The proposed methods include information about the hierarchical vasculature to allow color coding and interactive occlusion of subtrees. Furthermore, visualization results from two AVM cases were compared with findings during surgery.
Figure 4.1: Demonstration of the vesselness filter applied to an arterial TOF image. Figure (a) shows a volume rendering of the original TOF image masked with the reformatted brainmask. Figure (b) shows a volume rendering of the result of the vesselness filters. Transfer functions for both renderings have been adapted such that no non-vascular tissue obstructs the view to the vessels.

4.5 Preprocessing and Data Analysis

The proposed visualization framework requires datasets to pass through a preprocessing pipeline, prior to being suited for volume rendering. Required preprocessing steps include registration of the multiple images into a common coordinate system, segmentation of the brain from native T\textsubscript{1}-weighted images, and an enhancement step of both the CE-T\textsubscript{1} as well as the TOF images. Since the quality of the preprocessing results is crucial for the clinical review of the data, a supervised semi-automatic approach for brain segmentation and image registration was preferred over a fully automatic approach. The workup for each case can be performed by a medical assistant and requires at most 10 minutes for a trained user. All additional processing steps are calculated fully automatic. The final preprocessing results are saved and can be imported into the viewer application within several seconds. The following section presents a detailed description of the required processing. Figure 4.2 gives a schematic overview of the process.
Brain Segmentation and Tissue Classification

As a first preprocessing step, a segmentation of the brain from the skull is performed. This procedure, often referred to as skull stripping or brain extraction, is carried out semi-automatically using an interactive watershed transformation on a T1-weighted MRI image [50]. The algorithm exploits the connectivity of the brain through the white matter. It is a marker-based approach that allows controlling the segmentation mask using markers for inclusion and exclusion of individual areas of the segmentation mask. Typically it requires between 1 and 10 markers to be placed within the image in order to obtain a proper brain segmentation mask. For the purpose of vascular visualization, it is important to include the large sinus veins on the brain’s surface within the segmentation mask as well as the carotid arteries. Later, all other datasets will be masked with the brain segmentation mask obtained from this step, consequently, vessels missing from this mask can not be visualized in the later stage.

Following the brain segmentation, a tissue classification of the brain is performed using a multi Gaussian fit on the brain histogram [48]. The model uses a mixture of three pure Gaussian classes for cerebral gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF), and additional partial volume classes for CSF/GM,
Chapter 4. Interactive Multi-Volume Visualization for Planning of Cerebral AVM Surgery

Figure 4.3: Schematic illustration of the morphological brain peeling: The upper diagram shows a horizontal profile through the distance-transformed brainmask. The curve increases for positions deeper inside the brain. The lower diagram shows resultant weighting functions after multiplication with a Gaussian. The different curves represent different sigmas.

CSF/WM and GM/WM mixtures, respectively. The resulting tissue classification can then be adapted to automatically generate a transfer function for the brain, mapping voxel intensities for each tissue type to appropriate color values.

4.5.1 Multimodal Registration

Image registration is a crucial step for any system incorporating multimodal imaging data. It describes the process of transforming different images into a common coordinate system. For the problem considered here, a precise matching of the vascular structures from the different MRI sequences is mandatory. Since there is a partial overlap of vessels visible within the different images, even a small mismatch could lead to irritating artifacts that could ultimately lead to misinterpretation of the final visualization by the surgeon. Here, a supervised rigid registration process is employed, which incorporates an automatic registration algorithm based on normalized mutual information (NMI) similarity measure [140], along with the option to manually refine the automatically obtained results. Such a refinement may
again be used as initialization to another iteration of the automatic registration, and so on. Ultimately, the medical assistant supervising the preprocessing needs to accept the final registration result.

The contrast-enhanced T\textsubscript{1}-weighted image is used as reference image, for all other images. It has a brain tissue contrast similar to the native T\textsubscript{1}-weighted image, which allows for accurate registration of these two images. At the same time, it displays both the arterial as well as the venous vessels, which provides the required contrast for registration of the arterial as well as the venous TOF data. Non-linear registration has not been considered, as the MR sequences involved don’t typically suffer from significant geometric distortion as e.g. EPI sequences.

**Virtual MR Subtraction Angiography**

In the CE T\textsubscript{1}-weighted image, vascular structures appear significantly brighter than the white matter. However, intensity values for smaller vessels typically overlap the intensity range of the white matter. Consequently, it becomes necessary to attenuate non-vascular tissue from such images prior to using it for volume rendering, virtually creating an MR subtraction angiography. This is done using the native T\textsubscript{1}-weighted image, which exhibits a similar contrast between gray and white matter tissue. A simple subtraction of the native from the contrast-enhanced T\textsubscript{1}-weighted image results in suppression of most brain tissue, while keeping even thin vessels well visible.

An additional scaling of the CE T\textsubscript{1}-weighted image may become necessary, because in MRI there are no absolute intensity values for individual tissue classes. This can easily be computed using the available tissue classification from the native T\textsubscript{1}-weighted image. The white matter probability map can be utilized to calculate a weighted histogram of the contrast enhanced image. The mean value of this histogram is considered to be the mean white matter intensity of the contrast enhanced image. With the WM mean of the native T\textsubscript{1} being known from the Gaussian mixture fit, the ratio of the two mean values can be applied as the scale factor to the CE T\textsubscript{1}-weighted image.

**Vessleness Filtering**

To enhance vascular structures within the images, the use of a multi-scale vessleness filter as originally proposed by Frangi et al. [39] becomes necessary. It is based on an analysis of the Eigenvalues of the images Hessian matrix at multiple scales. The Hessian captures structural properties of a local neighborhood of an image by evaluating the second order derivative along all three image dimensions. Image differentiation is obtained by convolution with derivatives of Gaussians, where the width of the Gaussian kernel controls the scale of the resulting Hessian. The resulting
structure tensor can afterwards be evaluated, by analyzing the configuration of its Eigenvalues, which differs for tubular-, sheet- or blob-like structures. Following this observation, a vesselness measure can be calculated, which describes for every image voxel its similarity to a tube like structure at a given scale.

The vesselness filter yields excellent results in improving contrast of the vascular tree, especially for the arterial TOF data. For the clinical evaluation, vesselness analysis has been applied on three scales with a Gaussian sigma ranging from 0.3 to 1.4 mm on the arterial TOF images. The effect is visually demonstrated in Figure 4.1.

**Morphological Brain-peeling**

For the arterial TOF images, it becomes necessary to attenuate bright structures close to the brain’s surface. Such structures result from the vesselness filter’s response to sulcal structures on the brain surface. At larger scales of the filter, these will generate some similarity to tubular structures, consequently producing a slight filter response. Visually, this results in a halo-like artifact around the surface of the brain, which partly occludes the view to internal structures of the brain when performing volume rendering on the data.

To handle such artifacts, a morphological brain-peeling technique has been implemented, that directly modifies the affected datasets using a combination of elemental morphological image processing operators. As a first step, a Euclidean distance transformation is calculated on the brainmask obtained from the skull-stripping step. The resulting image encodes an approximated shortest distance to the surface of the mask in \text{mm} for any voxel within the mask. Since the goal is to attenuate voxels close to the surface of the brain, the distance map is utilized to evaluate a Gaussian function with mean value zero, and a width that corresponds to the depth of the attenuation. The resulting image is inverted, such that its values are transferred into a range between [0..1] and multiplied with the original image, which causes structures within the set attenuation range to be gradually attenuated, instead of being clipped away sharply. Note that actual vessels will be affected to a smaller degree than artifacts, because they appear much brighter than the halo artifact. For the clinical data available, a Gaussian sigma value of 1.5 (\text{mm}) was determined empirically. Effectively, this steps acts as a soft masking of the input image with the provided brainmask. Figure 4.3 illustrates the transformation of the mask image. The upper diagram shows an intensity profile of the original distance-transformed brainmask. The lower diagram shows resulting intensity profiles after the Gaussian transformation for different sigmas.
**4.6 Volume Visualization for Surgery Planning**

Direct multi-volume rendering is used to visualize the three-dimensional shape of the AVM and the surrounding angio-architecture to support the neurosurgeon in the planning process. To facilitate efficient, interactive exploration of the different datasets, three approaches are combined:

1. **Independent clip-planes** are offered for each dataset. This allows control over possibly occluding structures between the multiple datasets. While the orientation of all clip-planes is synchronized, the viewer allows to set individual offsets for each clip-plane.

2. **Independent transfer-functions** per volume are set. This allows controlling the transparency and contrast of each volume individually, while at the same time providing the option of using a different base-color for each volume. Like this, it becomes possible to color arteries and veins differently, which is a central element in understanding the interplay between both structures.

3. A **focus of attention** rendering is applied, allowing to highlight a spherical area of variable size, while attenuating surrounding structures.

Technically, a dynamic volume rendering framework based on a scene graph is used. Every data set can be appended to the volume rendering at run-time.
of the application. Corresponding shading parameters, clip planes and transfer functions are independently defined for each data set and can be controlled by the user through the GUI of the application. Figure 4.5 illustrates the scene graph for multi-volume rendering. For every additional volume, a subgraph with child nodes of the corresponding visualization parameters is connected to the main volume renderer. Thus, every data set can be toggled on or off and manipulated independently.

![Figure 4.5: Illustrative description of the multi-volume rendering scene graph. Every additional volume has corresponding child nodes for clipping, transfer function and the volume texture.](image)

**Shaded Multi-Volume Rendering**

For the combined multi-volume rendering, various datasets, such as the T1-weighted main volume, the MR subtraction angiography and the time-of-flight datasets need to be visualized simultaneously. The additional datasets can be independently classified with separate transfer functions using interactive adaption by the medical expert. Using the pre-processed registration information, all datasets can be transformed to spatially correlate with the main volume. For that, the additional datasets with the corresponding transformation matrix are loaded into the volume renderer. The transformation is applied per fragment on the graphics hardware, using GLSL shader programming. Finally, the fused voxel information is composed, utilizing the max operator in the shader [14].
4.6. Volume Visualization for Surgery Planning

Additionally, the extracted brain mask is used to separate the voxels of the T₁-weighted main volume into brain and skull. For that, a two-dimensional transfer function is generated, concatenating an individual transfer function for both brain and skull. To visually separate the grey and the white brain matter, the brain transfer function is automatically derived from the histogram analysis performed during pre-processing (cf. Section 4.5).

In order to enhance perception of the spatial relations of the different datasets involved, volume shading is utilized in the visualization. Thus, for every dataset, the corresponding gradients have to be calculated and stored in a volume texture. In the shader program, the gradient textures are used to calculate Phong shading individually per dataset. Additional shading effects such as toon shading or boundary enhancement [108] could be additionally enabled per volume.

Independent Clip Planes

A general challenge in handling multi-modal volume rendering is the occlusion problem. Varying non-transparent anatomical structures may overlay other important structures. For instance, intracranial vessels are not visible because of their location inside of the brain. To address this issue with respect to facilitating the exploration of the vascular structures, independent clip planes per dataset are included in the volume rendering. A single clip plane, defined by the plane normal and distance to origin, can be activated for each volume. Technically, in the rendering shader, the distance between the current sample and an active clip plane is calculated and the fragment is discarded in case of a negative distance. In the viewer application, all clip planes can be independently toggled on or off and transformed by the user in the two-dimensional slice views.

An issue arising from the combination of clip-planes and shaded volume rendering is the occurrence of artifacts on the clip planes surface, caused by undefined gradients in homogeneous tissue such as the gray or white brain matter. In order to cut a shaded volume correctly, the inverted plane normals are used as gradients at the clip plane’s boundary, as proposed in [138]. Hence, the neurosurgeon is able to cut away skull and brain to visually explore the AVM nidus with respect to the underlying white and grey brain matter.

Focus Of Attention

While clip planes can be an efficient tool for cutting objects along a planar direction, they are by definition not well suited for attenuation of arbitrarily shaped structures such as vessels. To address this challenge, application of a focus of attention rendering is suggested. The focus of attention defines for a user defined position inside the dataset the area of interest around which structures can be either
Chapter 4. Interactive Multi-Volume Visualization for Planning of Cerebral AVM Surgery

Figure 4.6: The GUI of the viewing application. An MR subtraction angiography and a time-of-flight data set are loaded to visually explore the draining veins (blue) and feeding arteries (red).

attenuated or enhanced in the volume rendering. For that, the focus point can precisely be defined in a two-dimensional slice rendering, preferably offering multiple orthogonal reformations. Based on this, for each voxel, the distance to the focus point is evaluated in a shader, and both the color saturation and alpha value can be manipulated using a customizable distance measure [109]. Thus, structures out of interest, such as vessels, can be attenuated interactively by the medical expert. Intuitively speaking, the focus of attention can be considered as a spotlight into the data, guiding the focus onto a specific structure, while fading – without necessarily masking away – external, possibly distracting structures.

Overall implementation

The overall system has been implemented as a standalone medical software assistant, running on conventional PC or Mac hardware with a current generation GPU processor. Interactive framerates were achieved using an Intel Core i7 6700HQ processor at 2.6 GHz with a nVidia GeForce GTX 960M GPU. The viewer application allows to load the preprocessed datasets and directly start analysis of a case without any further manual processing requirements. Cross-sectional slice views in the three standard planar reformations can be used to navigate through
the datasets, to control the position and alignment of the primary clipping plane, and also to define the position of a focus of attention. A panel on the right gives access to all control elements required for interaction. For each dataset, the user is able to individually toggle rendering on or off, define clipping parameters, and also adjust the base color of the transfer function. Further, the radius, de-saturation, and opacity parameters of the focus of attention can be controlled here. The main viewer of the application is reserved for the volume rendering of the data. Here the camera is controlled, as well as the transfer function of the selected dataset, using a standard mouse interaction scheme. Figure 4.6 illustrates the GUI of the viewer application.

4.7 Results and Discussion

The tool has been installed in the radiology department of our clinical partner, where it has been used to retrospectively analyze three AVM cases that have undergone surgery during the years 2008 to 2010. Analysis of these cases was done by the same neurosurgeon who had performed the actual neurosurgical procedures. In the following section, the potential usefulness of our application for preoperative planning is discussed by summarizing the findings based on the proposed tool, accompanied with documentations of how they correlated with intraoperative findings.

4.7.1 Case 1

Case 1 consists of a AVM classified as Spetzler Martin Grade III lesion [121], which was located in the frontal opercular regions on the dominant hemisphere. The main surgical challenges in this case related to speech and motor function of the brain as well as the overall anatomical location and the visibility of the nidus. The following surgical issues were specified by the neurosurgeon:

- Opening the Sylvian fissure in order to visualize the surface representation of the AVM and identify the main feeders, isolating and disconnecting them.

- Identification of feeders that are “en passage”. This is critical in this location, as disrupting a vessel that is merely supplying branches to the AVM but continuing on to supply critical motor or speech regions would result in poor neurological outcome.

- Following the main venous drainage to the AVM nidus and then dissecting the AVM away from the surrounding brain without injuring the adjacent eloquent cortex.
Figure 4.7: (a) Volume visualization of Case 1 using T1, MR-Venography and MR-Angiography datasets. Utilizing the cut planes, the Sylvian fissure (right arrow) as well as an underlying arterial feeder (left arrow) can be illustrated. (b) Detailed rendering of Case 1. The AVM nidus is enhanced using the focus of attention technique. Arterious feeders (white arrow) and draining veins (black arrow) are visible in the vicinity of the AVM.

The visualization application provided a critical understanding of the precise location of the AVM. MR-Angiography and MR-Venography datasets were used for the identification of the venous drainage pattern and the exact location of the feeding vessels. The focus of attention proved helpful to reduce visual clutter of unimportant vessels. Furthermore, the topography of the feeding arteries in relation to the venous drainage and cortical anatomy was visualized using the clip plane technique. Essentially, the visualization allowed the surgeon to visualize exactly what he would have encountered during surgery, especially regarding the location of the feeders, and most importantly, identification of a single but critical “en passage” vessel.

Figure 4.7 (a) shows an annotated overview visualization of Case 1. The clip plane is applied to the brain, allowing a clear view of the Sylvian fissure and the underlying feeding vessel at the brain’s surface. Furthermore, the draining vein and its pathway are clearly visible in the volume rendering. In Figure 4.7 (b), a detailed view of the vessels in the vicinity of the AVM is presented. The brain and the skull are completely clipped away. The focus of attention is set to the AVM center, resulting in the attenuation of unimportant vascular structures and hereby
enhancing the nidus. Following the path of the feeding arteries, the spatial location of the “en passage” vessel can be recognized in the center of the lesion.

The neurosurgeon concluded, that in this particular case, the volume visualization would have been invaluable for surgery planning. It allowed to visualize what to expect without the need for “mental gymnastics” when trying to fuse all the 2D data mentally. During surgery, the neurosurgeon would have been able to correlate the vessel that he thought was the “en passage” with the volume rendering and follow its course, sealing the branches to the AVM and preserving the main vessel. Also, understanding the topography of the lesion would have been significantly simplified. The offered visualization allowed him to explore precisely the gliotic plane\(^1\) which is important to avoid straying into the normal adjacent cortex and white matter.

### 4.7.2 Case 2

The AVM in case 2 was less complex and risky than the AVM of case 1. It was a Spetzler Martin Grade II located in the insula region of the right hemisphere. Again, the main challenges related to motor function due to middle cerebral artery branch involvement. This AVM was located deep in the Sylvian fissure and had no obvious surface representation. One would need to open the Sylvian fissure to identify the AVM within a myriad of middle cerebral artery (MCA) branches and venous tributaries. Surgical issues included:

- Opening the fissure while dealing with arterialized veins. Several veins formed a network which would not allow complete opening without closing of a branch. In such cases, it is critical to know which branch could be taken safely. Taking the wrong one would result in immediate venous hypertension, brain swelling, and hemorrhage.

- With the fissure opened, following the main vein to the AVM was undertaken. However, due to the multitude of overlapping MCA branches, it was difficult to determine which were the true feeders, the “en passage” vessels and other simply uninvolved MCA branches in the fissure.

- Once feeders were secured, resecting the AVM without straying into the deep white matter was important in this case, as the AVM extends into the claustrum and extreme capsule. Delving too deep into this location could damage the internal capsule and result in hemiparesis or plegia.

\(^1\)A gliotic plane develops between the lesion and cerebral white matter, allowing for relatively safe removal.
Chapter 4. Interactive Multi-Volume Visualization for Planning of Cerebral AVM Surgery

Figure 4.8: Volume rendering of Case 2 using MR-Angiography and subtraction angiography. Figure (a) gives an overview of the case along with an anatomical reference. The AVM is located deep in the Sylvian fissure (white arrow). The presence of the draining vein (black arrow) in both datasets confused the neurosurgeon in identifying the vessel. Figure (b) gives a detailed visualization of the AVM. The main two feeders (white arrows) straddle the main draining vein (black arrow).

For this case, the volume rendering allowed the neurosurgeon to gain an understanding of most of the topography and representation of the AVM as it relates to the insula and frontotemporal-Sylvian regions prior to opening the fissure. He was virtually able to determine which vein could have been sacrificed to open the fissure. The visualization clearly revealed that the main two feeders were straddling the main draining vein, allowing confirmation and isolation of these two vessels. One difficulty was determining the course of the MCA branch that was “en passage”. The rendering gave the impression that a superiorly located vessel extending beyond the AVM was an “en passage” MCA branch when in fact, during surgery, it revealed as a draining vein on the brain surface. This impression has been caused by the fact, that this particular vessel was enhanced in both, the TOF-angiography as well as the MR venography. Additionally, because of an above-average flow speed of the blood in that vessel during image acquisition, the draining vein was also enhanced in the MR angiography. A more extensive exploration of the multiple datasets by modification of the individual transfer functions finally helped to confirm the venous nature of this vessel.
4.7. Results and Discussion

Figure 4.8 (a) shows the volume rendering of Case 2. The brain is clipped away so that the AVM can be located deep within the Sylvian fissure (white arrow). MR-angiography and MR-venography datasets are used to visualize the feeding arteries and the draining veins. The posterior draining vessel (black arrow) is visible in both datasets which confused the neurosurgeon. In Figure 4.8 (b), a close-up volume rendering of the AVM nidus is presented. Using the focus of attention, the neurosurgeon was offered a clear view to explore and understand the shape of the arterial feeders and the draining vessels in detail. The main two feeders (white arrows) straddle the main draining vein (black arrow).

4.7.3 Case 3

The AVM of case 3 was quite small and of relatively low risk. This was a Spetzler Martin grade II AVM. The only significant risk and challenge was given by its location in the sensory strip and proximity to the motor cortex. This leads to an increased risk of postoperative lower extremity paralysis or paresis. Furthermore, the AVM was located deeply inside a sulcus so there would be no apparent surface representation other than the draining vein. Understanding the topography and relation to the motor strip was critical in this case.

The provided volume visualization allowed the neurosurgeon to visualize the sulci and general topography of the region without the need for imprecise mental fusing of the MRI and conventional CT-Angiography. This allowed him to explore the draining vein leading to the sulci. This would have helped determining the appropriate sulci to open and follow it in. Once identified, the feeding vessels were clear, and the volume rendering was of no additional value in that aspect. However, one difficulty found with the visualization was precisely estimating the extent of the nidus using the viewer application. It appeared that the volume rendering underestimated the size of the AVM nidus in this case. It was difficult to visualize the exact extent of the nidus on the model. If one followed the main vein on the rendering, the AVM became apparent, but its full extent and borders were not clearly visualized. This issue is likely due to both, the small size of the AVM and relatively low image resolution of the MR-Venography dataset.

Figure 4.9 (a) shows the volume rendering of Case 3. A clip plane is set on top of the AVMs location in order to visualize the draining veins (black arrows) with respect to the gyri and sulci of the brain. Furthermore, a feeding artery following the brain’s surface to the AVM is visible (white arrow). In (b), the brain is completely clipped away and the focus is set to the AVM’s nidus. The image shows an additional feeding artery passing through the brain as well as the draining veins at the brain’s surface. In comparison to the multi-volume visualization, a conventional maximum intensity projection (MIP) of the MRV is shown in (c). Although the same vascular structures are present in both renderings,
Chapter 4. Interactive Multi-Volume Visualization for Planning of Cerebral AVM Surgery

Figure 4.9: Volume rendering of Case 3. In (a), the location of the draining veins (black arrows) can be explored, incorporating the brain’s anatomy. Image (b) shows the feeding arteries (black arrows) and the draining veins. In (c), a conventional MIP rendering of the MR-Venography is presented.

it is significantly more difficult to grasp the vascular configuration of the lesion from this conventional presentation.

4.8 Discussion

In contrast to the traditional MIP rendering of vascular structures, the volume visualization developed here allows visualizing fused anatomical structures such as the gyri and sulci simultaneously with various additional vascular datasets. This multimodal visualization allows a neurosurgeon to understand the three-dimensional anatomy of the AVM nidus. Moreover, the extent and spatial position of the lesion can be estimated with respect to the brain’s anatomy. For instance, the evaluation of the lesion’s relation to eloquent areas is of high importance for surgical planning. Furthermore, feeding arteries and draining veins can be identified utilizing the combined rendering of MR-venography with time-of-flight MR-angiography datasets. In contrast, with a traditional MIP rendering of such datasets the neurosurgeon has to mentally fuse varying visualizations of the vascular structures, which is challenging, particularly if a complex angioarchitecture is present. Also, identification of “en-passage” vessels is critical during surgery planning, because disruption of a vessel that is merely supplying branches to the AVM but continuing on to supply critical motor or speech regions would result in poor surgery outcome (cf. Case 1 and Case 2).

A potential drawback of the method is caused by the circumstance that partially ambiguous presentation of parts of the vascular tree in arterial and venous images could potentially lead to misinterpretation of the data, as it was observed for case 2. This, however, rather relates to the image acquisition technique as opposed to
the visualization technique applied. It needs to be addressed by making careful exploration of all available datasets mandatory during planning, in order to specify the precise nature of the vessels surrounding the AVM nidus (cf. the feeding vein in Case 2). Another aspect relating to the necessity for careful data exploration is the fact that depending on the window settings of transfer function of a dataset in combination with the limited image resolution, small vessels could possibly become invisible in the volume rendering and thus not be considered for surgical planning. More critically, the impression of the lesions size will also depend on the settings of the transfer function, causing a potential risk of underestimating the lesion’s true extent. The latter points underline the requirement for careful exploration of all datasets, which needs to be assured by the neurosurgeon performing the preoperative planning.

4.9 Conclusions

In this chapter, a data-processing pipeline along with an interactive volume-rendering framework for visualization of cerebral vascular structures based on multiple MRI datasets has been developed. This work has been primarily motivated by the clinical challenge posed to neurosurgeons during the preoperative analysis and risk assessment of AVMs. Specifically, this challenge is defined by the difficulty of understanding the typically complex angio-architecture of an AVM, which includes identifying feeding arteries, draining veins and arteries “en passage”. Also, understanding the spatial relation of the nidus to eloquent areas of the brain, which strongly affects the possible access path to the lesion may be a difficult task in certain cases. Exclusively relying on conventional 2D slice representations of tomographic images results in a complicated mental process of generating a 3D model of the lesion.

This problem has been addressed by combining multiple MR-based angiographic datasets within a multi-volume rendering framework that allows for individual shading and clipping for each volume. Vascular trees are presented without performing an explicit segmentation in order to avoid the risk of unintentionally hiding information from the data caused by insufficiently accurate segmentation results. Instead, the approach allows the neurosurgeon to explore the given vascular anatomy interactively, allowing him to fuse his prior knowledge with patient-individual findings from the available images. The possibility of manipulating individual transfer functions of each dataset especially facilitates the mental classification of affected vessels as either arterial or venous.

To demonstrate the usefulness of this technique, results from three clinical cases that have undergone neurosurgical AVM resection were presented. These cases were re-analyzed retrospectively using the method described in this work. As a
subjective assessment, the performing neurosurgeon repeatedly stated how much the possibilities of the 3D visualization would have helped him plan and perform the resection. For selected cases, he rated the visualizations as invaluable, emphasizing that he would not want to miss such techniques in the future.

**Publications as primary author**

The work presented in this chapter has been originally presented at the *Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)* in Norrköping, Sweden in 2012. This work contains the following contributions by colleagues or co-authors: Christian Rieder implemented the focus of attention shader, and assisted with implementation of the clip-plane handling in the shaded multi-volume rendering. Carlos A. David, MD, is the neurosurgeon who performed the retrospective case-study analysis. Christiane Engel assisted with the pre-processing of the data used for the evaluation. All data used in this chapter is courtesy of Lahey Clinic Medical Center, 41 Mall Road, Burlington, MA 01805, USA.


© 2012 The Author(s).


Reproduced by kind permission of the Eurographics Association.
5 Discussion and Conclusion

Arriving at one goal is the starting point to another.

(John Dewey)

Analysis of medical images has become a cornerstone in modern medicine. Ever since Wilhelm Conrad Röntgen produced the first images revealing internal properties of the human body using X-Ray imaging, it has continuously increased in relevance up to a point where medicine has become unthinkable without it. This role has increased exponentially with every new technique introduced to the field during the second half of the 20th century. With the ever advancing possibilities of image acquisition techniques like computed tomography, ultrasound and magnetic resonance imaging, the amount and complexity of the data to be analyzed has increased to a point where pure qualitative interpretation...
can merely scratch the surface of what is possible in terms of information extraction. As such, computer assisted analysis has become equally important as the process of generating such images itself.

The importance of drawing precise and reliable conclusions from images becomes readily apparent when considering the impact caused by the decisions drawn. Wherever a diagnosis and the resulting treatment options for an individual patient rely even partially on insights gained from images, the necessity of validity and correctness of such decisions becomes immediately evident. As such, quantification can be considered the ultimate goal to reach for. However, quantification is not the only means to reach this goal. Visualization is an equally important element. Presenting data in a manner appropriate to both the data and the underlying clinical question can be the decisive factor towards drawing the right conclusion. Also, the option to interactively explore data can enable a physician to extract exactly the information required for making a diagnosis or treatment decision. These aspects can by summarized under the concept of *objective assessment*: Making decisions based on objective, validated information.

This thesis studied a limited aspect of this concept. The focus lies in the field of neuroimaging, more specifically it deals with the analysis of MR images of the human brain. And again within this vast field, only three selected examples have been studied: First, the automated volumetric quantification of spinal cord atrophy, a clinically highly relevant parameter in the context of MS. Second, the development of specialized, interactive tools for explorative analysis of specific structures of the human brain, namely the extraction of meaningful ROIs from DTI data, and the precise segmentation of individual gyri of the brain from patient individual images. And finally, the use of interactive visualization techniques that significantly simplify the integration of partial information carried by different images to a larger picture revealing the relations and interplay of arteries and veins in the context of pre-surgical planning for AVM treatment.

Research on image analysis and quantification in neuroimaging is a vast and active field. However, it is extremely rare that results from such endeavors are successfully transferred into clinical practice. Leaving aside potential legal, economic or merely organizational aspects, there remain two primary reasons for this situation: In order to become successfully applicable in clinical routine, a software based tool needs to be equally robust and time efficient. In the end, potential short-comings in one or both of these requirements must be outweighed by the added value gained by the tools. In light of these constraints, simple and robust tools can in some situations be preferable over more complicated and advanced solutions, assuming that they suffice the afore mentioned criteria.

The tools developed and discussed in this thesis have the potential to do so. All methods have been developed with a focus on both robustness towards variations of
the input data, and efficiency in either computation- or interaction time. As such, each method by itself carries the potential of pushing image analysis in clinical neurosciences a small step further from qualitative interpretation, towards objective assessment.

**Contributions of this thesis**

Each of the main chapters of this thesis addressed one of the three concepts listed above: quantification, interactive exploration, and visualization. They all have in common that they offer pragmatic, clinically applicable solutions for specific problems when dealing with analysis and interpretation of images of the human brain.

**Quantification of spinal cord atrophy**

For measuring atrophy of the spinal cord, a robust fully automated workflow has been developed and evaluated. It addresses three steps required in order to solve the underlying measurement: Localization, segmentation and quantification of the spinal cord. For localization, a template based registration pipeline has been developed and evaluated against a large number of input images. As part of this evaluation, the impact and relevance of different critical parameters of the registration algorithm have been analyzed systematically. This allows for adjustment of these parameters based on profound knowledge, as opposed to heuristic tuning as it is commonly done when performing image registration.

Building on the template based localization of the spinal cord, a robust segmentation algorithm has been implemented. It uses a watershed transformation that is steered by landmarks obtained from the template, in combination with incorporation of expert knowledge about the expected shape of the structure to be segmented.

Finally, a process for quantification of the volume and length of the segmented section of the spinal cord has been implemented and evaluated. It is based on analysis of the intensity histogram, to which a bimodal Gaussian mixture model is fitted. The result of the fit provides a precise estimate of the spinal cord volume. By finally measuring the exact length of the spinal cord’s section, the mean cross-sectional area can be calculated, which is a clinically established parameter for assessing spinal cord pathology in MS.

The robustness of the automation of the method has been evaluated on a dataset of 111 images, yielding a success-rate of 100%. The quantification approach of this method has been used in a number of clinical studies examining the role of spinal cord atrophy in different phenotypes and stages of MS. These results underline
the potential of the developed method to robustly transfer a fully automatic quantification tool into clinical workflows.

**Interactive segmentation**

Addressing the field of interactive exploration of data, two technically independent techniques for interactive segmentations have been developed. Both techniques allow for extraction of structures of interest without requiring complex processing or parameterizations. First, a technique for interactive generation of anatomically meaningful ROIs from DTI data has been developed, that potentially allows for interactive exploration of DTI data using real-time fiber tracking. By this, the relationship between a seed-region used for tracking and the resulting fiber bundle can be explored interactively. Especially in the context of DTI, where a fundamental problem lies in the difficulty of selecting appropriate seed or filter regions, such a tool can provide an option to explore a given dataset, and by this better understand and objectively judge the information contained in the data.

The second method addresses the task of segmentation of individual brain gyri. Although fully automated segmentation and labeling techniques can be found in the literature, they are computationally complex, and robustness to anatomical variations found in clinical images is not guaranteed. In consequence, such labeling techniques currently do not exist in a clinically applicable form. However, there are situations where a full labeling of the brain is not required. Instead, availability of a precise segmentation of single gyri might be sufficient, e.g. for mapping results from fMRI experiments, or for quantifying tissue parameters for selected structures only. The proposed method addresses exactly this task. It does provide a means for precise segmentation of individual gyri within seconds. In consequence, it opens the door for currently unavailable image analysis techniques in clinical settings, effectively increasing objectivity in medical image interpretation.

**Interactive multi-volume visualization**

The potential of visualization for gaining insights into images has been studied in the context of neurosurgical planning for AVM treatment. Especially in neurosurgery, there are multiple scenarios in which significant additional information can be gained from fusing multiple datasets. While different imaging techniques can each reveal a specific aspect of a case, crucial additional information is gained by combining the data. In AVM surgery, this relates to the identification of arteries and veins from different imaging techniques, and especially the way they interact inside a nidus. In other applications of neurosurgical planning, fiber tracts extracted from DTI, eloquent areas measured using fMRI, or tumors extracted from anatomical sequences are equally important. A key element here is, that in the end, the whole
is more than just the sum of its parts. Combining different images by means of visualization ultimately becomes the decisive element for determining how to approach surgery.

**Final conclusion**

From the vast field of image interpretation tasks existing in the context of neuroimaging, this thesis studied three selected, specific examples. Each of them addressed a different aspect of how data analysis can become more objective, by providing computational tools that fulfill basic requirements for transfer into clinical workflows. All of those methods have in common that they focus on a combination of robustness for automated analysis, and if necessary the incorporation of user knowledge through interaction techniques where automation is not feasible. The potential of clinical applicability has been successfully demonstrated for each of them.
Bibliography


Bibliography


<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title and Details</th>
</tr>
</thead>
</table>