

# Quantification of Spinal Cord Atrophy in Magnetic Resonance Images

Inauguraldissertation

zur  
Erlangung der Würde eines Dr. sc. med.

vorgelegt der  
Medizinischen Fakultät  
der Universität Basel

von  
Simon Pezold  
aus Langenbrettach, Deutschland

Basel, 2016

Genehmigt von der Medizinischen Fakultät  
auf Antrag von

Prof. Dr. Philippe C. Cattin, Universität Basel — *Fakultätsverantwortlicher*  
Prof. Dr. Till Sprenger, Universität Basel — *Korreferent*  
Dr. Xavier Bresson, École polytechnique fédérale de Lausanne — *externer Gutachter*

Basel, den 4. Oktober 2016

Prof. Dr. Thomas Gasser  
*Dekan*

***Quantification of Spinal Cord Atrophy in Magnetic Resonance Images***

*Inaugural dissertation to be awarded the degree of Dr. sc. med.  
presented at the Faculty of Medicine of the University of Basel  
by Simon Pezold from Langenbrettach, Germany; Basel, 2016*

*Approved by the Faculty of Medicine on application of  
Prof. Dr. Philippe C. Cattin, University of Basel — faculty representative  
Prof. Dr. Till Sprenger, University of Basel — secondary advisor  
Dr. Xavier Bresson, Swiss Federal Institute of Technology in Lausanne — external expert*

*Basel, October 4, 2016*

*Prof. Dr. Thomas Gasser — dean*

*Meiner Großtante  
Marta Simpfendörfer  
(1930-2014)*



# Contents

*Acknowledgments* vii  
*Summary/Zusammenfassung* ix

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation	1
1.2	Contribution	2
1.3	Outline	2
<b>2</b>	<b>Medical Background</b>	<b>3</b>
2.1	The Spinal Cord	3
2.2	Multiple Sclerosis	4
<b>3</b>	<b>Image Segmentation</b>	<b>7</b>
3.1	Manual Image Segmentation	7
3.2	Computer-based Image Segmentation	8
3.3	A General Binary Segmentation Energy	10
3.4	Graph Cut	14
3.5	Continuous Max Flow	18
<b>4</b>	<b>Segmenting the Spinal Cord</b>	<b>23</b>
<b>5</b>	<b>A Semi-Automatic Method for the Quantification of Spinal Cord Atrophy</b>	<b>27</b>
<b>6</b>	<b>Automatic Segmentation of the Spinal Cord Using Continuous Max Flow with Cross-sectional Similarity Prior and Tubularity Features</b>	<b>41</b>
<b>7</b>	<b>Reliable Volumetry of the Cervical Spinal Cord in MS Patient Follow-up Data with Cord Image Analyzer (<i>Cordial</i>)</b>	<b>55</b>
<b>8</b>	<b>Automatic, Robust, and Globally Optimal Segmentation of Tubular Structures</b>	<b>67</b>
8.A	Derivation of the Nonterminal Flow Equality	77
8.B	Derivation of an Update Rule for the Flow $p$	78
<b>9</b>	<b>Discussion and Conclusion</b>	<b>83</b>
	<i>Bibliography</i>	85
	<i>Curriculum Vitae</i>	89
	<i>Publications</i>	90



# Acknowledgments

First and foremost, I would like to thank Professor Dr. Philippe C. Cattin for his outstanding guidance, support, and motivational spirit throughout my dissertation. His great expertise and insight into every single project of the group continue to amaze me.

I would like to thank all current and former colleagues in the group for making and having made work at the Department of Biomedical Engineering a pleasure: Simon Andermatt, Dr. Patrik Arnold, Natalia Chicherova, Corinne Eymann-Baier, Dr. Beat Fasel, Dr. Ketut Fundana, Alina Giger, Antal Horváth, Dr. Christoph Jud, Dr. Silja Kiriyanthan, Dr. Tahir Majeed, Nadia Möri, Peter von Niederhäusern, Dr. Frank Preiswerk, Tiziano Ronchetti, Robin Sandkühler, Dr. Sebastian Scherer, Dr. Adrian Schneider, Jörg Wagner, and Stephan Wyder. Among these, my special gratitude goes to Antal and Ketut for letting me rely on their profound mathematical background many times, as well as to Adrian, Simon, and Stephan for offering me collaborations on exciting projects that led beyond my main scope of research.

I am grateful for the support of MIAC AG, Basel. Professor Dr. Ernst-Wilhelm Radü and Professor Dr. Till Sprenger provided me with superb advice and ideas from the medical side. I would like to extend my thanks to all our partners at MIAC AG and at the University Hospital Basel for working together in a trusting and fruitful atmosphere, in particular Dr. Michael Amann, Dr. Michaela Andělová, Charidimos Tsagkas, Dr. Katrin Weier, and Dr. Jens Würfel.

Thanks to all my friends for putting the life in my work–life balance. Finally, I cannot thank my family and especially my parents enough for too many things to mention here.





# Summary / Zusammenfassung

## Quantification of Spinal Cord Atrophy in Magnetic Resonance Images

Quantifying the volume of the spinal cord is of vital interest for studying and understanding diseases of the central nervous system such as multiple sclerosis (MS). In this thesis, which is motivated by MS research, we propose methods for measuring the spinal cord cross-sectional area and volume in magnetic resonance (MR) images. These measurements are used for determining neural atrophy and for performing both longitudinal and cross-sectional comparisons in clinical trials.

We present three evolutionary steps of our approach: In the first step, we use graph cut-based image segmentation on the intensities of T1-weighted MR images. In the second step, we combine a continuous max flow segmentation algorithm with a cross-sectional similarity prior and Hessian-based structural features, which we apply to T1- and T2-weighted images. The prior leverages the fact that the spinal cord is an elongated structure by constraining its cross-sectional shape to vary only slowly along one image axis. In conjunction with the additional features, the segmentation robustness is thus increased. In the third step, we combine continuous max flow with anisotropic total variation regularization, which enables us to direct the regularization of the cross-sectional shape of the spinal cord more flexibly.

We implement the proposed approach as a semi-automatic software toolchain that automatically segments the spinal cord, reconstructs its surface, and acquires the desired measurements. The software employs a user-provided anatomical landmark as well as hints for the location of the spinal cord and its surroundings. It accounts for the bending of the spine, MR-induced image distortions, and noise.

We evaluate the proposed methods in experiments on phantom, healthy subject, and patient data. Our measurement accuracy and precision are on par with the state of the art. At the same time, our measurements on MS patient data are in accordance with the medical literature.

## Rückenmarksatrophyquantifizierung in Magnetresonanztomographie-Bildern

Quantifizierungen des Rückenmarksvolumens sind für das Studium und zum Verständnis von Erkrankungen des Zentralnervensystems, wie zum Beispiel der multiplen Sklerose (MS), von grundlegendem Interesse. Die vorliegende Arbeit wird durch die MS-Forschung motiviert. Sie präsentiert Methoden, mit denen sowohl das Volumen als auch die Querschnittsfläche des Rückenmarks in Magnetresonanztomographie-(MR)-Bildern gemessen werden

können. Diese Messungen werden wiederum zur Bestimmung von Nervengewebsatrophie sowie für Vergleiche im Rahmen klinischer Längs- und Querschnittstudien herangezogen.

Der zugrunde liegende Ansatz durchläuft drei Entwicklungsschritte: Im ersten Schritt wird *Graph Cut* zur intensitätsbasierten Segmentierung von T1-gewichteten MR-Bildern verwendet. Ein zweiter Schritt kombiniert einen Segmentierungsalgorithmus, der auf *Continuous Max Flow* aufbaut, mit Strukturinformationen, die durch Auswertungen der Hesse-Matrix gewonnen werden. Weiterhin wird im zweiten Schritt zur Vorgabe gemacht, dass sich der Rückenmarksquerschnitt entlang einer Bildachse nur langsam verändern darf, wobei hier die längliche Form des Rückenmarks ausgenutzt wird. Dadurch, und durch die hinzugenommenen Strukturinformationen, wird die Robustheit der Segmentierung erhöht, welche nun auf T1- und T2-gewichteten Aufnahmen erfolgen kann. Im dritten Schritt wird *Continuous Max Flow* mit anisotroper Variations-Regularisierung (*anisotropic total variation regularization*) kombiniert, womit eine flexiblere Ausrichtung der Rückenmarksquerschnitts-Regularisierung erreicht wird.

Der Ansatz wird als teilautomatische Software implementiert, welche das Rückenmark segmentiert, seine Oberfläche rekonstruiert und die gewünschten Messungen durchführt. Als manuelle Eingaben genügen dabei die Markierung eines anatomischen Referenzpunktes sowie Hinweise auf die Lage des Rückenmarks und seiner Umgebung. Die Software berücksichtigt die Krümmung des Rückenmarks, MR-spezifische Verzerrungen sowie Bildrauschen.

Eine Evaluierung der vorgeschlagenen Methoden erfolgt auf Phantomdaten, auf Daten gesunder Freiwilliger sowie auf MS-Patientendaten. Die Richtigkeit und Präzision der Messungen erweist sich dabei etablierten Methoden gegenüber ebenbürtig. Gleichzeitig befinden sich die Messungen auf MS-Patientendaten im Einklang mit der medizinischen Fachliteratur.

# 1 Introduction

With more than two million patients, multiple sclerosis (MS) is one of the most common neurological disorders worldwide [45]. The patients, among which two thirds are women, usually show first symptoms in early adulthood and acquire different types of disability throughout their lifetime. A single causal explanation for MS is still unknown, and so is a cure. Consequently, a lot of research is directed towards understanding the mechanisms of MS and developing new drugs and therapies.

## 1.1 Motivation

Multiple sclerosis is a disorder of the central nervous system, that is, of the brain and spinal cord. It manifests in characteristic localized changes of the nervous tissue, so-called lesions. Identifying the lesions in magnetic resonance (MR) images has been done for a long time, using them as an aid or marker in MS diagnosis and for tracking the disease course. However, relating the lesion status to MS attacks and progression is ambiguous, which has come to be known as the “clinico-radiological paradox” [3] (see Section 2.2).

In the course of finding further markers for MS, another nervous tissue-specific property has gained attention more recently, namely *atrophy*, that is, generalized loss of nervous tissue. Especially regarding the patients’ locomotion, it has been suggested that *atrophy in the spinal cord* is strongly related to MS status and progression [29]. Similar to its usefulness in evaluating lesions, magnetic resonance imaging (MRI) has proven a viable modality also for atrophy assessment.

The desire to measure atrophy in MRI volumes motivates the development of adequate tools. Being part of a project of MS research through clinical trials, our main objective is to provide our clinical partners with the necessary methods and software tools for quantifying the spinal cord volume or cross-sectional area in patient MR images.

At present, it is still very common in MS research to segment the spinal cord and acquire quantitative measurements manually. The manual approach, however, is tedious and error-prone (see Section 3.1). Consequently, the developed tools should be easy to use and reliable, along with being capable of dealing with high amounts of data. In particular, they should help to track atrophy over time in longitudinal settings and enable groupwise comparisons in cross-sectional settings. Further motivation lies in a possible application of the developed methods to other pathologies, such as assessing the effects of rehabilitation in para- or quadriplegic trauma patients.

## 1.2 Contribution

We propose adaptations of established general image segmentation algorithms to the specific task of segmenting the spinal cord. Initially relying on graph cut [14], we use a continuous max flow segmentation algorithm [47] in later approaches. For segmentation robustness, we leverage the fact that the spinal cord is an inherently tubular and thus elongated structure: In our first version using max flow, we introduce a cross-sectional similarity prior, which constrains the spinal cord's cross-sectional area to vary only slowly along a predefined image axis. In a later version, we make this prior follow the cord's course itself, using anisotropic total variation regularization. Our choice of image features for the segmentation algorithm also exploits the tubularity assumption: Using only the image intensities in the beginning, we later integrate Hessian-based features such as vesselness, which increase segmentation robustness, enable us to apply our toolchain to a wider range of MR contrasts, and provide the necessary directional information for steering the anisotropic regularizer.

Based on the presented segmentation approaches, we build up a semi-automatic software toolchain for easy and large-scale acquisition of quantitative spinal cord measurements, both cross-sectional areas and volumes, in MR images. The software segments the spinal cord, reconstructs its surface, and measures the desired parameter. It employs initial hints for the location of the spinal cord and its surroundings along with an anatomical landmark provided by the user through a graphical user interface.

## 1.3 Outline

In Chapter 2, we make the reader familiar with the medical background of the thesis. Chapter 3 provides an introduction to image segmentation, in which we focus on describing and relating the segmentation approaches that we use in later chapters. Chapter 4 summarizes the peculiarities of segmenting the spinal cord in MR images and references to prior work.

Chapters 5–8 comprise the publications that resulted from our work to date. Chapter 5 presents our initial graph cut-based toolchain. Chapter 6 demonstrates a proof of concept using continuous max flow with the newly introduced cross-sectional similarity prior, along with adding Hessian-based image features. Chapter 7 fuses the approaches of the preceding two chapters by integrating max flow segmentation into the original toolchain and adds an extensive evaluation on both healthy subjects and MS patient follow-up data. Chapter 8 presents the next step of advancing our segmentation method, using anisotropic total variation to let the regularizer follow the actual cord direction.

We complete the thesis with a discussion and conclusion in Chapter 9.

## 2 Medical Background

**Outline.** This chapter provides a brief introduction to the two main medical aspects of the thesis: the *spinal cord* and *multiple sclerosis*. Section 2.1 describes the function, location, and inner structure of the spinal cord. Section 2.2 presents important aspects of multiple sclerosis, such as pathology, diagnosis, and especially the role of MRI in this context.

### 2.1 The Spinal Cord

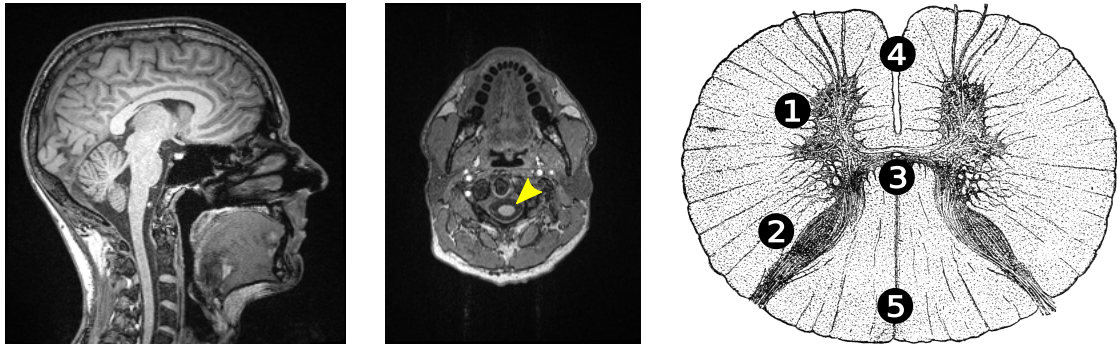
The spinal cord is a long tubular structure that represents the main nervous pathway connecting the brain with the peripheral nervous system. The brain and spinal cord together form the *central nervous system*.

**Function.** Besides controlling certain reflexes, the spinal cord mainly serves as a medium for transmitting information between the brain and the body periphery. Information flows in both directions: sensory input is provided from the receptors of the periphery to the brain, and motor commands are sent from the brain to the muscles of the body.

**Location.** Originating from the *medulla oblongata*, the spinal cord leaves the skull through the *foramen magnum* and traces its path down the spinal canal inside the vertebral column (see Fig. 2.1 on the following page, left). It stretches to the level of the first or second lumbar vertebra, thereby covering a length of approximately 42–45 cm in adults [20, p. 749]. At its inferior end, the spinal cord gradually narrows, forming the *conus medullaris*. Below this point, only a bundle of spinal nerves continue down the spinal canal, called the *cauda equina*.

Like the brain, the spinal cord is embedded in *cerebrospinal fluid* (CSF). The CSF circulates in the subarachnoid cavity, a space that is formed by two membranes: the *arachnoid mater* on the outside and the *pia mater* on the inside, which tightly encloses the nervous tissue.

**Appearance and Inner Structure.** The spinal cord consists of *gray matter* and *white matter*. In contrast to the brain, the white matter, containing the nervous tracts, is found on the outside, surrounding the gray matter, which comprises the nerve cell bodies. Together, the gray and white matter form an elliptical cross section, which is flattened along the dorsoventral axis (see Fig. 2.1, middle). For the most part, the spinal cord has a left–right diameter of roughly 8–10 mm and a dorsoventral diameter of roughly 5–6 mm. Both the degree of flattening and the diameter vary steadily along its course; see Kameyama et al. [23] for a detailed quantification of the excised spinal cord.



**Figure 2.1:** The spinal cord. *Left:* sagittal slice of a T1-weighted MR scan of the head; the spinal cord is the bright central vertical structure. *Middle:* transverse slice of the same MR scan on the C1/C2 level; the arrow indicates the bright spinal cord cross section; due to the choice of MR contrast, the gray and white matter parts appear homogeneous. *Right:* schematic spinal cord cross section (reproduction from Gray [20]; public domain); numbers denote the anterior column of the gray matter (1), the posterior column of the gray matter (2), the central canal (3, *bright elliptical structure*), the anterior median fissure (4), and the posterior sulcus (5).

The *gray matter* of the spinal cord has the shape of a butterfly or the letter H (see Fig. 2.1, right). Its two symmetric halves consist of a front part (*anterior column*) and back part (*posterior column*). The halves are joined by a commissure that contains the *central canal*, which is filled with CSF. The gray matter is mainly made up of nerve cell bodies, dendrites, and a smaller share of axons, with motor neurons residing in the anterior column and sensory neurons in the posterior column.

The surrounding *white matter* is divided by two depressions, the *anterior median fissure* and the *posterolateral sulcus* (see Fig. 2.1, right). It mainly consists of myelinated motor and sensory axons. The *myelin sheath* of the axons serves as an electrical insulator and thereby increases the speed of information propagation along the axon. The myelin, a fat-based substance, is responsible for the white matter's bright appearance that motivates its name.

## 2.2 Multiple Sclerosis

Multiple sclerosis (MS) is a chronic inflammatory and disabling disorder of the central nervous system [11]. With an estimated total number of 2.3 million patients in 2013, it is one of the most common neurological disorders worldwide [45]. Moreover, with a usual onset in the patients' twenties or thirties [46], it is in many countries the main reason for disability in young adults that has not been caused by trauma [45].

**Causes.** Different factors have been shown to play a role in the development of MS, among others: *gender*, with a worldwide female-to-male MS patient ratio of approximately two to one [45]; *genetics*, with an increased incidence in relatives of MS patients compared to the rest of the population [28, p. 2631]; *geography*, with the highest MS incidence in the temperate

climate zones, especially in large parts of Europe and Northern America [45], and the lowest around the equator [28, p. 2632]; *virus infections*, as seen in an increased MS risk following infection with Epstein–Barr virus [2]. It remains uncertain, however, which of these factors act as a cause and which act merely as a trigger for MS.

**Pathology.** On the cell level, the disease most prominently manifests in so-called plaques or *lesions*, especially in the white matter but also in the gray matter of the brain and spinal cord – the name *multiple sclerosis*, literally meaning *multiple hardening*, originates from these lesions. MS lesions are marked by varying degrees of inflammation, blood–brain barrier anomalies, as well as axonal demyelination and degeneration. Demyelination is believed to be caused either directly through attacks of the immune system’s T cells on the myelin sheath or indirectly by necrosis of oligodendrocytes, which produce the myelin sheath [28, p. 2649]. Demyelination impedes the transport of information in the form of electrical signals along the axons. Axonal loss stops the flow of information completely.

Apart from the more focal neurological changes due to lesions, generalized tissue loss or *atrophy*, mainly due to axonal damage, is happening in the gray and white matter both in the brain and in the spinal cord of MS patients.

**Symptoms and Disease Course.** As random as the distribution of the lesions in the central nervous system can be, as various are the symptoms in MS patients. The list of symptoms includes disturbance of vision, general sensory problems, motoric deficits, coordination and equilibrium impairment, physical weakness, bladder dysfunction, cognitive impairment like deficits in attention and reasoning, and many others [11].

After onset, most MS patients show recurrent symptomatic phases (*relapses*) that alternate with periods of stability and varying degrees of recovery (*remissions*). This is known as the *relapsing–remitting phase of MS* (RRMS). For most of these patients, their disease course ultimately turns into one of gradually increasing disability, called *secondary progressive MS* (SPMS). About 10–15 % of MS patients show such a gradual course from onset without relapses [28, p. 2629], known as *primary progressive MS* (PPMS). Various other types of MS exist, with *benign MS* and *fulminant MS* marking the boundaries of the severity spectrum.

Aspiring a quantitative criterion of their disease status, MS patients are often scored with respect to Kurtzke’s expanded disability status scale (EDSS) [25]. The scale assigns an accumulated score between zero (“normal neurologic exam” [25]) and ten (“death due to MS” [25]) based on the degree of disability in eight functional systems (“pyramidal, cerebellar, brain stem, sensory, bowel and bladder, visual, cerebral, other” [25]). As the scoring is done based on the judgment of a neurologist, a certain degree of subjectivity is implied.

**Treatment.** With an actual causal therapy of MS currently not being possible, administration of active substances focuses on inhibiting the *inflammatory processes* (interferon beta, glatiramer acetate) and on keeping the *blood–brain barrier* intact (natalizumab) in RRMS patients. Sustained *physical and cognitive training* may benefit the long-term prognosis of MS patients, exploiting the brain’s plasticity (i.e. its capabilities to adapt). Apart from this, patient-specific *symptomatic treatment* may be necessary. [28, pp. 2660–2664]

## 2 Medical Background

**Diagnosis.** In 2001, based on experience with previous diagnostic criteria for MS, an international expert panel proposed a set of new criteria that aimed at a high sensitivity (i.e. a high true positive rate) and specificity (i.e. a high true negative rate) for diagnosing MS. These criteria, known as the *McDonald criteria*, have since undergone two revisions with the latest one published in 2011 [37], and have gained wide appreciation and acceptance. The key principle of the McDonald criteria is “dissemination of lesions in space (DIS) and time (DIT)” [37]. This means that for diagnosis, the presence of MS-specific lesions has to be demonstrated in different parts of the central nervous system (DIS), along with the occurrence of new lesions over time or the simultaneous presence of lesions of different age (DIT). The criteria are designed for combining clinical evidence (such as the report or observation of an MS relapse) with findings from MRI (namely lesion detection in MR scans).

**The Role of MRI in MS.** In MR images, MS lesions are visible as dark or bright spots in the nervous tissue, depending on the chosen MR contrast and the use of contrast agents. MRI has thus traditionally been used mainly for the assessment of lesion formation and evolution. This has been done for diagnosis, for example with the McDonald criteria mentioned above, but also for tracking the disease course, especially in patients under treatment. For a long time, however, it has also been noted that the patients’ lesion status only weakly correlates with their clinical status [3, 19]. In particular, newly appearing lesions do not necessarily imply a clinical worsening of the disease and likewise, a constant number of lesions does not necessarily imply a stalled disease progression. This observation, which has been termed “*clinico-radiological paradox*” [3] or “*clinical/MRI paradox*” [19], has led to a search for other MS surrogate markers.

Despite the fact that MS is being considered an inflammatory-mediated demyelinating disease, there is increasing data that support *neurodegeneration* as the major cause of irreversible neurological disability in MS [46]. In connection to that, *brain atrophy* has been shown in recent years to be an alternative marker of MS that appears to be independent of lesion status [31]. Here, especially gray matter atrophy has been shown to be strongly associated with disease progression [19, 42]. Correlating brain atrophy with the patients’ disability score according to EDSS (see above) has turned out to be an ambiguous task, though. While some studies find significant correlation between atrophy and disability, others fail to find such a relationship [19].

As to the patients’ physical impairment, it has therefore been suggested to focus on *spinal cord atrophy* [15]. Indeed, an emerging body of literature demonstrates a relationship between spinal cord atrophy and MS status and progression; see for example Lukas et al. [29] and references therein for a recent overview. These observations, in turn, give rise to the demand for methods and tools to reliably measure spinal cord atrophy (see Chapter 4).



## 3 Image Segmentation

Very generally speaking, image segmentation is the process of dividing an image into non-intersecting parts, usually with the goal of gaining regions whose content shares distinct common properties and whose boundaries separate image parts where these properties differ.

A common special case is *binary image segmentation*, that is, the case of dividing the image into two regions, which we will exclusively address in the following. Binary image segmentation typically comprises the task of extracting an object from the image, which can be phrased as separating the *foreground* (i.e. the object of interest) from the *background* (i.e. everything else). The binary segmentation task can be achieved by either assigning a binary label to each point in the image, for example, by assigning 1 to the foreground points and 0 to the background points, or by delineating the boundary that separates the foreground from the background with an explicit parameterization, for example, by marking the object's outline with a polygonal chain in two-dimensional images.

**Outline.** This chapter is meant to provide background information on image segmentation as a foundation for the methodical chapters (Chapters 5–8). Section 3.1 briefly comments on manual image segmentation. Section 3.2 states the approaches to and challenges of computer-based image segmentation. Section 3.3 introduces a general model for computer-based binary image segmentation. Sections 3.4 and 3.5 present and relate two concrete manifestations of the model that can be used for segmenting images in practice and that we apply in subsequent chapters.

### 3.1 Manual Image Segmentation

The task of segmenting natural images, such as photographs or the output of medical imaging modalities, is often an easy one for human observers, who may outline the objects of interest intuitively or requiring only little training. However, two key problems arise with the manual approach. The first problem is that the task requires constant attention and may be experienced as tedious. This especially applies to the medical field with its requirement for accurate and precise data treatment and its large amount of (often three-dimensional) imaging data. The second one is the more systematic problem of *reproducibility*: It is very likely that two observers segmenting the same image will not produce exactly the same result. Moreover, even the same observer will most likely produce different segmentations if he or she segments the same image twice, due to variations in attention and judgment over time. If one wants to perform quantitative analyses of segmented images, as is the case in our problem setting, it may thus not be clear whether observed differences are merely caused by

these reproducibility issues or whether they are related to actual differences in the imaged objects.

Considering the problems with the manual approach, the interest in using computer-based image segmentation, is at least twofold: First, one wants to release human observers, which in the medical field are often highly trained medical experts, from their time-consuming task. Second, one wants to eliminate the reproducibility issues by using approaches that will always produce the same results when given the same image. Computer-based image segmentation, however, brings about its own challenges.

## 3.2 Computer-based Image Segmentation

The prerequisite for computer-based image segmentation is framing the problem at hand in a way that is suitable for digital processing. Informally speaking, we can view this as the transfer of the necessary problem-specific knowledge to the computer. A plethora of computer-based image segmentation techniques have been proposed, a complete review of which is beyond the scope of this thesis. In the following, we present a rough and nonexhaustive categorization of them, before we focus on those techniques that we used in the later chapters.

**Basic Techniques.** Among the most *basic* or *ad hoc techniques*, we may classify methods such as thresholding, region growing, or watershed (see e.g. Pratt [39, pp. 579ff]). Taking the risk of overgeneralizing, one may say that while basic techniques are usually simple to understand and implement, they lack in robustness and flexibility for adjustment to particular segmentation tasks. In the most simple case of thresholding, for example, we assume that all image values of the foreground lie above a certain value, namely the threshold, and all background values lie below it, or vice versa. Thresholding therefore easily suffers from noise and illumination variation in the given images. Furthermore, it is inapplicable in cases where the assumed foreground–background value distribution is not fulfilled.

**Machine Learning Techniques.** If supervised *machine learning techniques* and especially *deep learning techniques* are applied to image segmentation, of which the latter more recently gained notable popularity (see e.g. Long et al. [26] and references therein), it is the goal to learn the segmentation problem by example instead of modeling it. To achieve this, already segmented images are fed into a framework that has certain learning capabilities (basically a system of equations, such as a neural network). Learning then means adjusting the parameters in the framework’s usually large parameter space in a way that the framework’s output for the shown images closely matches their known segmentations. At the same time, one tries to ensure that the parameterization generalizes well to unseen instances, that is, to unsegmented images. In practice, such techniques require substantial amounts of training data for good performance. This requirement, however, constitutes a major drawback for their use if such data are not available, like in our case.

**Energy-Based Techniques.** In *energy-based techniques* for image segmentation, the given segmentation problem is modeled “by hand”. In other words, we try to describe, relying on our

own expert knowledge, what are the properties of a good image segmentation, by means of a mathematical expression. In particular, we model a *cost function*, also called an *energy term*, which takes as its input an unsegmented image and a candidate solution, that is, a proposed segmentation, and maps them to a single value: the *cost* or *energy* of the candidate segmentation, given the image. The better the solution is according to our model (and thus according to our judgment of what makes a good segmentation) the smaller this value will be. Solving the problem therefore means finding in the solution space (i.e. in the set of all possible segmentations) a solution with the smallest possible cost (a global minimum of the cost function), or at least one with a cost that is small enough to signify an acceptable solution (a small local minimum, for example).

Naturally, the question arises as to how we can search the cost function's solution space in a meaningful way in order to minimize it. Brute force, that is, simply evaluating the cost function for all possible segmentations, is clearly not an option: Finding the foreground-background labeling of an image as small as 1000 pixels would already necessitate the evaluation of  $2^{1000} \approx 10^{301}$  possibilities. This number is beyond current computational capacities, where the word *beyond* is not meant in the sense of "a single desktop computer cannot handle the problem", but in the sense of "if every electron in the visible universe was a computer, together they still could not handle the problem"; see MacKay [30, p. 359] for a compelling thought experiment of such kind. Formulating an energy-based segmentation approach thus consists of at least two steps: modeling the problem – and devising a way to solve it.

Energy-based segmentation includes *edge-based techniques* such as snakes [24] and their various adaptations. Here, the segmentation boundary or snake is deformed to minimize a cost function that compromises between so-called external and internal forces. While the external forces draw the snake towards certain features in the image, such as edges, the internal forces ensure that the snake maintains certain shape properties. Snakes still find wide applicability in practice. However, they are prone to get stuck in local minima and thus require a good initialization in order to produce a good result.

In this thesis, we work with *region-based techniques*. In contrast to edge-based techniques, the energy of region-based techniques mainly depends on the content of the segmented regions rather than their boundaries. This allows for actually designing models that separate image regions sharing common properties. Recall that we formulated the latter as the goal of image segmentation above. Region-based segmentation techniques may be subdivided into discrete [4, 5, 7, 14, 21] and continuous [8, 9, 10, 33, 47] approaches. The structure of the listed approaches, both discrete and continuous, can be traced back to the Ising model [22] from the field of physics. The continuous approaches among them are all closely related to the piecewise constant Mumford–Shah model [32, p. 580].

In the following, we will introduce a particular case of the piecewise constant Mumford–Shah model and show its relation to the Ising model. Building upon the introduced model, we will formulate a general way to model the energy term for binary image segmentation. We will then present and relate the two approaches that we use for segmentation in the later chapters of this thesis: the discrete *graph cut* [7, 14] and the *continuous max flow* [47] approach. Graph cut provides global solutions to the discrete binary segmentation problem. It became popular in segmentation during the first decade of the century, especially with the development of fast and efficient graph cut algorithms tailored towards image segmentation. The continuous

### 3 Image Segmentation

max flow and related approaches overcome some of the deficits of graph cut, which we will discuss at the end of Sections 3.4 and 3.5. Additionally, since these continuous approaches are easily parallelizable, they gained popularity especially since general purpose computing on graphics processing units (GPUs) was facilitated with frameworks like CUDA<sup>1</sup> and OpenCL<sup>2</sup>.

### 3.3 A General Binary Segmentation Energy

Let  $\Omega \subset \mathbb{R}^d$  be a *bounded domain* of a  $d$ -dimensional image. In the context of medical images, usually  $d = 2$  (planar image),  $d = 3$  (volumetric image or time sequence of planar images), or  $d = 4$  (time sequence of volumetric images). In this thesis,  $d = 3$  for the most part, as we work with single instances of volumetric images. Let  $I : \Omega \rightarrow \mathcal{J}$  be the *image* function that maps points  $x \in \Omega$  to points in the space of image values  $\mathcal{J}$ , where for example  $\mathcal{J} \subset \mathbb{R}^c$  for a  $c$ -channel image or  $\mathcal{J} = [0, 1]$  for a normalized single-channel image. Finally, let  $u : \Omega \rightarrow \mathcal{U} = \{0, 1\}$  be the *binary segmentation* function that assigns a value of either 1 (foreground, i.e. object of interest) or 0 (background, i.e. everything else) to each point  $x \in \Omega$ .

We formulate the binary segmentation problem, as we have laid out above, in terms of a nonnegative energy functional  $E$  of  $u$ , given an image  $I$ , namely

$$E[u; I] = F[u, I] + R[u], \quad (3.1)$$

which we use for finding the best segmentation  $u^*$  given  $I$ , that is

$$u^* = \arg \min_{u : \Omega \rightarrow \mathcal{U}} E[u; I]. \quad (3.2)$$

In this form, the energy term  $E$  is the sum of two other terms that we assume to be nonnegative: the *fidelity term*  $F$  and the *regularization term*  $R$ .

*Bayesian perspective:* From a Bayesian point of view, we may interpret this formulation as the problem of maximizing the *posterior*  $P(u|I)$ , that is, the problem of finding the most probable segmentation  $u$  given the image  $I$ . From Bayes' theorem, we know that

$$P(u|I) \propto P(I|u) P(u). \quad (3.3)$$

Maximizing Eq. (3.3) with respect to  $u$  is equivalent to minimizing  $-\log P(u|I)$ :

$$\arg \max_u P(u|I) = \arg \min_u -\log P(u|I) = \arg \min_u -\log P(I|u) - \log P(u). \quad (3.4)$$

If we here define the *prior*  $P(u)$  as

$$P(u) := \exp(-R[u]) \quad \Leftrightarrow \quad -\log P(u) = R[u] \quad (3.5)$$

and the *likelihood*  $\mathcal{L}(u|I) = P(I|u)$  as

$$P(I|u) := \exp(-F[u, I]) \quad \Leftrightarrow \quad -\log P(I|u) = F[u, I], \quad (3.6)$$

<sup>1</sup><http://www.nvidia.com/cuda> (last accessed on November 3, 2016)

<sup>2</sup><https://www.khronos.org/opencl/> (last accessed on November 3, 2016)

we end up with the energy functional  $E$  of Eq. (3.1). In other words, we may relate the regularization term  $R$  to  $P(u)$  – the prior probability of the segmentation  $u$  – and the fidelity term  $F$  to  $\mathcal{L}(u|I)$  – the likelihood of  $u$  given an image  $I$ .

If we assume  $\Omega$  to be discrete, the Bayesian point of view motivates the use of energy terms that model the image and its segmentation as a Markov random field, which can be optimized with graph cut [21]. While not explicitly mentioned there, the graph described in Section 3.4 follows this idea.

*Modeling perspective:* From the perspective of modeling the segmentation problem, it is the task of the fidelity term  $F$  to make the segmentation  $u$  consistent with the given image information  $I$ , therefore  $F$  should have a small value if the segmentation agrees well with the image. On the other hand, it is the task of the regularization term  $R$  to ensure certain desired properties of the segmentation  $u$ , and  $R$  thus should be small if  $u$  actually fulfills these properties. The segmentation result represents a compromise between the two terms and their potentially conflicting tasks. Both  $F[u, I]$  and  $R[u]$  are often defined as integral operators, as we will see shortly. In the following section, we show what the rather abstract tasks of  $F$  and  $R$  mean in the concrete case of the binary piecewise constant Mumford–Shah model [32, p. 580].

**Piecewise Constant Mumford–Shah Model.** Chan and Vese’s version [10] of the piecewise constant Mumford–Shah model for binary segmentation of a single-channel image  $I : \Omega \rightarrow \mathbb{R}$  can be written as the energy term

$$E_{\tilde{M}}[\Gamma, c_0, c_1; I] = \int_{\text{in}(\Gamma)} (c_1 - I)^2 dx + \int_{\text{out}(\Gamma)} (c_0 - I)^2 dx + \lambda L(\Gamma), \quad (3.7)$$

where  $\text{in}(\Gamma)$  is a subset of  $\Omega$  with the boundary  $\Gamma$ ,  $\text{out}(\Gamma) = \Omega \setminus \text{in}(\Gamma)$ ,  $L$  is a boundary measure (e.g. arc length if  $d = 2$ ),  $\lambda \in \mathbb{R}_{\geq 0}$  is a weighting parameter, and  $c_0, c_1 \in \mathbb{R}$  are constants with  $c_0 \neq c_1$ . For our purposes, we rewrite the energy term in the notation introduced above. Furthermore, we assume  $c_0$  and  $c_1$  to be known beforehand, yielding

$$E_M[u; I] = F_M[u, I] + R_M[u] \quad (3.8)$$

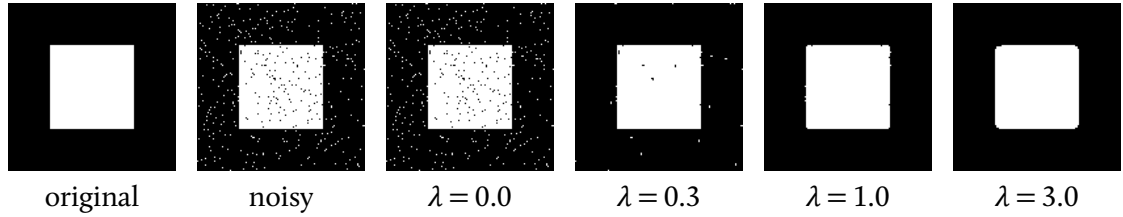
with

$$F_M[u, I] = \int_{\Omega} u (c_1 - I)^2 + (1 - u) (c_0 - I)^2 dx, \quad c_0, c_1 \in \mathbb{R}, \quad c_0 \neq c_1, \quad (3.9)$$

$$R_M[u] = \lambda L(\partial\{x \in \Omega \mid u(x) = 1\}), \quad \lambda \in \mathbb{R}_{\geq 0}, \quad (3.10)$$

where  $u = \mathbf{1}_{\text{in}(\Gamma)}$ , that is,  $u$  is an indicator function of  $\text{in}(\Gamma)$ , and  $\partial$  is the boundary (here, of the foreground).

In the model as formulated above, we assume that both background and foreground are regions of approximately constant intensities  $c_0$  and  $c_1$ , possibly perturbed by noise or other slight variations. If we recall the general definition of image segmentation at the beginning of this section, the property according to which we want to divide the image is here the common intensity of the pixels in each region. Consequently, the effect of the fidelity term  $F_M$  is



**Figure 3.1:** Influence of the regularization term  $R_M$ . Noise is added to a black-and-white image and the noisy image is segmented with increasing regularization weight  $\lambda$ , minimizing a relaxed version of the model in Eq. (3.8), where we fixed  $c_0$  and  $c_1$  to the ground truth values  $c_0 = 0$ ,  $c_1 = 1$ . See text for a more detailed description.

designed as follows: If for a certain point  $x$ ,  $I(x)$  is closer to  $c_1$ , then the expression under the integral is minimized by setting  $u(x) = 1$ , as  $(c_1 - I(x))^2 < (c_0 - I(x))^2$ . If  $I(x)$  is closer to  $c_0$ , then the expression is minimized by setting  $u(x) = 0$ , as  $(c_1 - I(x))^2 > (c_0 - I(x))^2$ .

The regularization term  $R_M$  is designed to minimize the foreground perimeter. While the idea behind the fidelity term may immediately make sense, the one behind the regularization term might be less obvious. Consider therefore the example of Fig. 3.1. Here, we added a small amount of noise to a binary (i.e. black-and-white) image of a square, keeping the resulting noisy image also binary. Now suppose we try to segment the image based solely on  $F_M$ , which we can achieve by setting  $\lambda = 0$ . What happens is that each bright pixel is assigned to one region and each dark pixel to the other, resulting in the best possible energy of  $E_M = 0$ . In other words, the segmentation with  $\lambda = 0$  is exactly the same as the noisy input or its negative image, depending on the choice of  $c_0$  and  $c_1$ . Here, the zero energy is only possible because our segmented image, like the segmentation, is binary.

By setting  $\lambda$  to larger values, we increasingly penalize the occurrence of foreground–background boundary segments, and therefore dark points in a bright neighborhood and vice versa. In other words, we achieve noise suppression by *penalizing an overly complex segmentation result*. Another effect becomes also visible, however: note the rounding of the square’s corners for  $\lambda = 3$ , which is due to penalizing the boundary length as well. This effect is sometimes called *shrinking bias*. Thus, we must always choose  $\lambda$  appropriately for a given problem.

At this point, it might not be clear how  $R_M$  in Eq. (3.10) is an integral operator, as we mentioned earlier: informally, we can think of calculating the perimeter as counting the points on the boundary, ignoring all other points in the domain.

**Relation to the Ising Model.** The Ising model [22] originally was designed to model the interactions of magnetic particles. It assumes particles in a regular grid that are exposed to an external magnetic field. The particles interact with each other, as well as with the external field. These interactions may be expressed in terms of an energy. In the model, each particle may assume two different states, *positive* or *negative*. The model assumes that only direct neighbors interact with each other, contributing to the setup’s energy if their states are opposing. A particle also contributes to the energy through the relationship of the external field’s local value to the particle’s state.

As Mumford and Shah point out [32, p. 581], their binary piecewise constant model may be seen as a continuous version of the Ising model. In particular, if we define a particle state configuration  $v : \Omega \rightarrow \{-1, 1\}$ , an external magnetic field  $I : \Omega \rightarrow \mathbb{R}$  on the discrete domain  $\Omega \subset \mathbb{N}^d$  with points  $x_k \in \Omega$ , neighborhoods  $\mathcal{N}_{x_k} \subset \Omega$  of  $x_k$ , and a weighting factor  $\tilde{\lambda} \in \mathbb{R}_{\geq 0}$ , then we may write the energy  $E_{\tilde{I}}[v]$  of the Ising model as

$$E_{\tilde{I}}[v; I] = \sum_{x_k \in \Omega} (v(x_k) - I(x_k))^2 + \tilde{\lambda} \sum_{\substack{x_k \in \Omega \\ x_l \in \mathcal{N}_{x_k}}} (v(x_k) - v(x_l))^2, \quad (3.11)$$

where the first sum models the interactions of the particles with the external field and the second sum models the interactions of the particles with their neighbors.

With appropriate choices of constants  $c_0, c_1 \in \mathbb{R}$  and  $\lambda \in \mathbb{R}_{\geq 0}$ , we may rewrite Eq. (3.11) as

$$E_{\tilde{I}}[u; I] = \sum_{x_k \in \Omega} u(x_k) (c_1 - I(x_k))^2 + (1 - u(x_k)) (c_0 - I(x_k))^2 + \lambda \sum_{\substack{x_k \in \Omega \\ x_l \in \mathcal{N}_{x_k}}} (u(x_k) - u(x_l))^2 \quad (3.12)$$

with  $u : \Omega \rightarrow \mathcal{U} = \{0, 1\}$ , which is just a discrete version of Eq. (3.8).

As an aside, a relation to the Ising model (or rather, to the Potts model [38], which generalizes the Ising model to more than two states per particle) has also been pointed out for the discrete graph cut approach (see Section 3.4), for example by Boykov et al. [6].

**Relaxing the Segmentation Energy.** Chan and Vese [10] were the first to solve Eq. (3.7) via level sets. However, their approach tends to converge to local minima and thus does not find globally optimal solutions in general. Chan et al. [9] showed that a *relaxed* version of Eq. (3.8) can be optimized globally. *Relaxing* means that the segmentation  $u$  is allowed to take on all values between 0 and 1, that is,  $u : \Omega \rightarrow \mathcal{U} = [0, 1]$ . In this relaxed setting, the regularization term  $R_M$  is approximated by

$$R_{\text{TM}}[u] = \lambda \int_{\Omega} |\nabla u| \, dx, \quad (3.13)$$

where  $\nabla$  denotes the gradient and  $|\cdot|$  is the  $l_2$  norm. The integral  $\int_{\Omega} |\nabla u| \, dx$  is called the *total variation* of  $u$ .

As Chan et al. also showed, we may gain an optimal binary segmentation  $u^\ell$  from the respective non-binary optimization result  $u^*$  simply by applying a threshold  $\ell \in (0, 1)$ , that is

$$u^\ell(x) = \begin{cases} 1, & u^*(x) > \ell, \\ 0, & u^*(x) \leq \ell. \end{cases} \quad (3.14)$$

In a sense, we thereby resort to the basic segmentation technique of thresholding that we mentioned at the beginning of Section 3.2. However, we do not apply a threshold to the image  $I$  itself, but we apply it to  $u^*$  instead, which can be seen as a sophisticated approximation of  $I$ .

Subsequent work of Bresson et al. [8] showed that the approach of relaxing and thresholding also works with replacing  $\lambda$  by a pointwise term under the integral (see next paragraph). Likewise, Olsson et al. [33] extended the approach to an anisotropic total variation regularizer. We will make use of both isotropic and anisotropic total variation regularization along with relaxation in Chapters 6–8.

**Generalizing the Model.** To allow for a more flexible choice of assumptions about the common property of foreground and background, we replace the factors  $(c_0 - I)^2$  and  $(c_1 - I)^2$  in Eq. (3.9) by general cost functions  $C_0, C_1 : \Omega \rightarrow \mathbb{R}_{\geq 0}$ , which still evaluate the image information and are designed to be small in points that are likely to belong to the respective region. Furthermore, we make  $\lambda$  in Eq. (3.13) a pointwise parameter, replacing it by a function  $C : \Omega \rightarrow \mathbb{R}_{\geq 0}$  under the integral [8].

We end up with a generalized version of the relaxed piecewise constant Mumford–Shah model, namely

$$E_{\text{gM}}[u; I] = F_{\text{gM}}[u, I] + R_{\text{gM}}[u], \quad (3.15)$$

with

$$F_{\text{gM}}[u, I] = \int_{\Omega} u C_1 + (1 - u) C_0 \, dx, \quad (3.16)$$

$$R_{\text{gM}}[u] = \int_{\Omega} C |\nabla u| \, dx. \quad (3.17)$$

As an example for choosing the costs in  $F_{\text{gM}}$ , Greig et al. [21] define  $C_0$  and  $C_1$  as

$$C_0(x) = \max\{I(x), 0\}, \quad (3.18)$$

$$C_1(x) = \max\{-I(x), 0\}, \quad (3.19)$$

$$\text{where } l(x) = \ln\left(\frac{p(I(x) \mid u(x) = 1)}{p(I(x) \mid u(x) = 0)}\right). \quad (3.20)$$

In this case, the criterion of foreground–background assignment is based on predefined probability density functions  $p(\cdot)$  of the image values.

An example for a pointwise cost  $C$  in  $R_{\text{gM}}$  is given in Chapter 8, where we define

$$C(x) = w \exp\left(-1/\zeta^2 |\nabla I(x)|^2\right) \quad \text{with } w, \zeta \in \mathbb{R}_{>0} \quad (3.21)$$

in order to guide the segmentation boundary towards image edges. Note that in the last example, the clear separation between  $F$  and  $R$  is somewhat weakened, as  $R$  here also evaluates image values rather than just values of  $u$ .

As a final adjustment of  $R$ , we can replace the total variation regularizer  $\int_{\Omega} |\nabla u| \, dx$  itself [33], preferably by a related term that better matches our assumptions on the segmented structure of interest’s shape. We demonstrate this in Chapter 8, where we propose to use an anisotropic total variation regularizer for tubular structure segmentation.

### 3.4 Graph Cut

If we discretize the generalized model of Eqs. (3.15)–(3.17), an optimal binary segmentation can be found by solving a *graph cut* problem, as was first proposed by Greig et al. [21]. In short, we imagine the sampling positions in the image as vertices in a graph. We link the vertices by edges whose weight depends on the similarity of the connected positions. We then try to separate the foreground from the background by severing preferably weak connections. Framing image segmentation as a graph cut problem enables us to use the algorithms that have been proposed for the problem in the field of graph theory.



**Definitions.** A graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  consists of a set of *vertices*  $\mathcal{V}$  and a set of *edges*  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ . Each edge  $e \in \mathcal{E}$  connects a pair of vertices  $v \in \mathcal{V}$ . Edges in a graph may be directed or undirected. A *directed edge*  $e = (v_a, v_b)$  with  $v_a, v_b \in \mathcal{V}$  provides a path from  $v_a$  to  $v_b$ , but not the other way round. An *undirected edge*  $e = \{v_a, v_b\}$  provides a path in both directions. As each undirected edge  $\{v_a, v_b\}$  can be equivalently represented by two directed edges,  $(v_a, v_b)$  and  $(v_b, v_a)$ , we only consider directed edges in the following. In a *weighted graph*, each edge is associated with a weight or cost  $C : \mathcal{E} \rightarrow \mathbb{R}_{\geq 0}$ , which we assume as nonnegative here. For our purposes, we define a *cut*  $\mathcal{C} \subset \mathcal{E}$  as a set of edges such that

$$\mathcal{C} = \{(v_a, v_b) \in \mathcal{E} \mid v_a \in \mathcal{V}_s, v_b \in \mathcal{V}_t\}, \quad (3.22)$$

where  $(\mathcal{V}_s, \mathcal{V}_t)$  is a partition of  $\mathcal{V}$ . This implies that after we sever the edges in  $\mathcal{C}$  (i.e. remove them from  $\mathcal{G}$ ), no more paths exist from any vertex in  $\mathcal{V}_s$  to any vertex in  $\mathcal{V}_t$ .<sup>3</sup> The *cut capacity*  $E(\mathcal{C})$  is defined as the sum of edge weights in  $\mathcal{C}$ , that is

$$E(\mathcal{C}) = \sum_{e \in \mathcal{C}} C(e). \quad (3.23)$$

For the purpose of solving the segmentation problem, we additionally define two *terminal vertices*: a special *source* vertex  $s \in \mathcal{V}$  with outgoing edges only and a *sink* vertex  $t \in \mathcal{V}$  with incoming edges only, along with their edge sets  $\mathcal{E}_s = \{(s, v_b) \in \mathcal{E} \mid v_b \neq s\}$  and  $\mathcal{E}_t = \{(v_a, t) \in \mathcal{E} \mid v_a \neq t\}$ . We call edges that are connected to either  $s$  or  $t$  *terminal edges* and all others *nonterminal edges*. We call  $\mathcal{C}$  an *s-t-cut* if  $s \in \mathcal{V}_s$  and  $t \in \mathcal{V}_t$ , meaning that after severing the cut's edges, no more paths from  $s$  to  $t$  exist.

**Building the Graph.** As mentioned above, solving the segmentation problem with graph cut is a discrete approach, which implies that we assume  $\Omega$  is a discrete domain (e.g.  $\Omega \subset \mathbb{N}^d$ ), made up of sampling positions  $x$ , which are known as the pixels ( $d = 2$ ) or voxels ( $d = 3$ ) of the image.

Following [21], we let  $\mathcal{V} = \Omega \cup \{s, t\}$ , that is, we let each sampling position  $x$  represent a vertex, to which we add a source and a sink. We then construct  $\mathcal{E}$  as follows (cf. Fig. 3.2 on page 19, which shows the graph after removing the cut edges): Each  $x$  is linked to  $s$  by an incoming edge of weight  $C_0(x)$ , which is designed to be large if  $x$  is likely to belong to the foreground and small if  $x$  is likely to belong to the background. This constitutes the set  $\mathcal{E}_s$  as defined above. Furthermore, each  $x$  is linked to  $t$  by an outgoing edge of weight  $C_1(x)$  having the opposite properties of  $C_0$ , that is, large for the background and small for the foreground, which constitutes  $\mathcal{E}_t$ . Note that Eqs. (3.18) and (3.19) fulfill these criteria for  $C_0$  and  $C_1$ , for example.

Finally, we create a set of *nonterminal edges*  $\mathcal{E}_n$ , such that  $\mathcal{E}_n \cup \mathcal{E}_s \cup \mathcal{E}_t = \mathcal{E}$ . We do so by adding pairs of opposing edges between each  $x_k \in \Omega$  and its spatial neighbors  $x_l \in \mathcal{N}_{x_k}$  (shown as undirected edges in Fig. 3.2). Here,  $\mathcal{N}_{x_k} \subset \Omega$  is the set of  $x_k$ 's neighbors, which could lie in a four-connected neighborhood in the two-dimensional case, for example. We assign a

<sup>3</sup>We follow Greig et al.'s definition of a cut here [21]. Alternatively, the word *cut* may refer to the partition  $(\mathcal{V}_s, \mathcal{V}_t)$  [5] and  $\mathcal{C}$  may be called a *cut set* [40, p. 115]. Note as an aside, that paths from  $\mathcal{V}_t$  to  $\mathcal{V}_s$  may still exist.

### 3 Image Segmentation

symmetric weight  $C_n(x_k, x_l)$  to all  $e \in \mathcal{E}_n$ , which is designed to be large if  $x_k$  and  $x_l$  are likely to belong to the same region and small if they are likely to belong to different regions. An example for  $C_n$  is given by Boykov et al. [7] as

$$C_n(x_k, x_l) = \begin{cases} w \exp\left(-\frac{1}{2\sigma^2} (I(x_k) - I(x_l))^2\right), & x_l \in \mathcal{N}_{x_k}, \\ 0, & x_l \notin \mathcal{N}_{x_k}, \end{cases} \quad (3.24)$$

which can be seen as a discrete version of Eq. (3.21), and which we use in Chapter 5. As a consequence, we can write  $C(e)$  as

$$C(e) = \begin{cases} C_0(x_l), & e \in \mathcal{E}_s, \\ C_1(x_k), & e \in \mathcal{E}_t, \\ C_n(x_k, x_l), & e \in \mathcal{E}_n, \end{cases} \quad \text{with } e = (x_k, x_l). \quad (3.25)$$

**Minimum Cut.** It is our goal to find a *minimum s-t-cut* of  $\mathcal{G}$ , that is, a  $\mathcal{C}$  of minimum cut capacity  $E(\mathcal{C})$  that separates the source  $s$  from the sink  $t$ . In order to achieve such a cut in the described setup, we observe that we have to sever either the terminal edge in  $\mathcal{E}_s$  or the terminal edge in  $\mathcal{E}_t$  for each  $x \in \Omega$ . Furthermore, at the boundaries of the partition  $(\mathcal{V}_s, \mathcal{V}_t)$ , we must sever the nonterminal edge  $(x_k, x_l)$  if  $x_k \in \mathcal{V}_s$  and  $x_l \in \mathcal{V}_t$ . As a consequence, we can rewrite the *s-t-cut* capacity as

$$E(\mathcal{C}) = F(\mathcal{C}) + R(\mathcal{C}), \quad (3.26)$$

with

$$F(\mathcal{C}) = \sum_{x \in \Omega} u(x) C_1(x) + (1 - u(x)) C_0(x), \quad (3.27)$$

$$R(\mathcal{C}) = 1/2 \sum_{x_k, x_l \in \Omega} C_n(x_k, x_l) |u(x_k) - u(x_l)|, \quad (3.28)$$

where  $F(\mathcal{C})$  encodes the severing of the terminal edges,  $R(\mathcal{C})$  encodes the severing of the nonterminal edges, and  $u : \Omega \rightarrow \{0, 1\}$  gives the desired segmentation as before. Division by two is necessary in  $R(\mathcal{C})$  as only one of the two edges in each non-terminal edge pair is severed whereas both are counted in  $R(\mathcal{C})$ 's double sum.

While the segmentation  $u$  turns up in Eqs. (3.27) and (3.28), it is not directly represented in the graph itself. We get its values, once we have found a minimum *s-t-cut*, by assigning 1 to each position  $x$  whose connection to  $s$  has not been severed and 0 to the remaining positions, that is, the set of  $x$  that are still connected to  $t$  (again, see Fig. 3.2 on page 19).

**Relation to Mumford–Shah Model.** A striking similarity between the cut capacity terms of Eqs. (3.26)–(3.28) and the generalized segmentation model of Eqs. (3.15)–(3.17) becomes apparent (see Table 3.1 on page 21 for a side-by-side view). As it turns out,  $F(\mathcal{C})$  and  $R(\mathcal{C})$  are discrete equivalents of  $F_{\text{gM}}[u, I]$  and  $R_{\text{gM}}[u]$ , at least if we allow only direct neighbors in  $\Omega$  to be connected via nonterminal edges.

For the  $F$  terms, their only difference lies in replacing the integral with a sum. For  $R(\mathcal{C})$ , we notice that the contribution  $C_n(x_k, x_l) |u(x_k) - u(x_l)|$  of each pair of neighbors  $x_k$  and  $x_l$  can be seen as the contribution of a point  $\chi$  that lies halfway between  $x_k$  and  $x_l$ . In particular,  $|u(x_k) - u(x_l)|$  can be interpreted as a central finite difference estimate of the gradient magnitude  $|\nabla u(\chi)|$  for the center point  $\chi$ , with the gradient  $\nabla u$  directed along the edge pair that connects  $x_k$  and  $x_l$ . The weight  $C_n(x_k, x_l)$ , which depends on both  $x_k$  and  $x_l$ , can be interpreted as the pointwise cost  $C(\chi)$  for the center point  $\chi$ .

The terminal part of the cut capacity,  $F(\mathcal{C})$ , therefore represents a discretized version of the fidelity term  $F_{\text{gM}}[u, I]$  on  $\Omega$ 's sampling positions. The nonterminal part,  $R(\mathcal{C})$ , however, can be seen as a discretization of the regularization term  $R_{\text{gM}}[u]$  between these sampling positions. The meaning of the word *between* here depends on the chosen neighborhood system. This described similarity is no coincidence, of course: recall that we modeled  $\mathcal{G}$  in order to represent the segmentation problem of Eqs. (3.15)–(3.17) in the first place.<sup>4</sup>

**Maximum Flow.** To find an actual solution, that is, a minimum  $s$ - $t$ -cut, we can tackle the problem from another perspective: As was first described by Ford and Fulkerson [16], finding a minimum cut corresponds to determining the *maximum flow* through a graph. In this context, we see  $\mathcal{G}$  as a network of pipes, where edges  $\mathcal{E}$  serve as pipe segments, the edge weight  $C(e)$  describes the capacity of pipe segment  $e$ , that is, the amount of flow that may pass, and vertices  $\mathcal{V}$  provide connections between pipe segments. Furthermore, we interpret the source  $s$  as a flow emitter (a fountain or spring, for example) and the sink  $t$  as a flow consumer (a drain, for example) – in fact, choosing the names *source* and *sink* is based on this interpretation.

We can see that a substance flowing through  $\mathcal{G}$  has to obey certain constraints: The amount of *source flow*,  $p_s : \Omega \rightarrow \mathbb{R}_{\geq 0}$ , which leaves  $s$  through an edge  $(s, x) \in \mathcal{E}_s$ , and the amount of *sink flow*  $p_t : \Omega \rightarrow \mathbb{R}_{\geq 0}$ , which arrives in  $t$  through an edge  $(x, t) \in \mathcal{E}_t$ , may not exceed their respective edges' capacities  $C_0(x)$  and  $C_1(x)$ , thus

$$0 \leq p_s(x) \leq C_0(x), \quad (3.29)$$

$$0 \leq p_t(x) \leq C_1(x). \quad (3.30)$$

A similar constraint applies for the flow  $p : \Omega \times \Omega \rightarrow \mathbb{R}$  between nonterminal neighbors  $x_k$  and  $x_l$ , which may not exceed the respective edge pair's capacity  $C_n(x_k, x_l)$  either. In this case, however, we have to consider that for each node pair, flow in both directions is possible. We therefore allow both negative and positive values in  $p$ , with negative values signifying flow that arrives in  $x_k$  (leaving  $x_l$ ) and positive values signifying flow that leaves  $x_k$  (arriving in  $x_l$ ), which implies that  $p(x_k, x_l) = -p(x_l, x_k)$ . The respective constraint is

$$|p(x_k, x_l)| \leq C_n(x_k, x_l). \quad (3.31)$$

---

<sup>4</sup>From a historical standpoint, the presented train of thought is not quite accurate, however; in particular, because Greig et al.'s work of applying graph cut to images [21] and Mumford and Shah's work that introduces their model [32] address different, albeit related, problems and use different cost functions, as we have already shown. Moreover, both approaches were published in 1989, without one referencing the other. It is perhaps more correct to say that both papers drew inspiration from a 1984 publication by Geman and Geman [17], in which the two brothers proposed a method for image restoration, which they related to the Ising model [22]. This shared inspiration of [21] and [32], in turn, seems to have led to such compatible results.

### 3 Image Segmentation

Furthermore, as all vertices beside  $s$  and  $t$  (i.e. all  $x \in \Omega$ ) serve purely as passive connections, meaning they neither emit nor consume flow, the sum of incoming flows, signified by a negative value, must equal the sum of outgoing flows, signified by a positive value, for each  $x_k \in \Omega$ . This is equivalent to stating that the sum of incoming and outgoing flows must be zero:

$$-p_s(x_k) + p_t(x_k) + \sum_{x_l \in \mathcal{N}_{x_k}^r} p(x_k, x_l) = 0. \quad (3.32)$$

Ford and Fulkerson showed that the maximum possible total amount of flow from  $s$  to  $t$  has exactly the same value as a minimum  $s$ - $t$ -cut's capacity [16], which is known as the *max-flow min-cut theorem*. Because all flow is emitted by  $s$ , we can measure the total current amount of flow by summing all  $p_s(x)$  or equivalently by summing all  $p_t(x)$ , as all flow is consumed by  $t$ . The problem of finding the maximum flow can thus be written as

$$\max_{p_s} \sum_{x \in \Omega} p_s(x), \quad (3.33)$$

so that constraints (3.29)–(3.32) hold. At the same time, the total current amount of source and sink flow is also equivalent to the total amount of flow across the cut, as all paths from  $s$  to  $t$  lead through  $\mathcal{C}$ .

Once we have found a solution for this maximization problem, a result for the original problem is given by an  $s$ - $t$ -cut that severs *saturated edges* only, that is, edges in which the amount of flow exactly matches the respective capacity. Note that neither the maximum flow configuration nor the minimum cut solution have to be unique, but all possible solutions produce the same cut capacity  $E(\mathcal{C})$ .

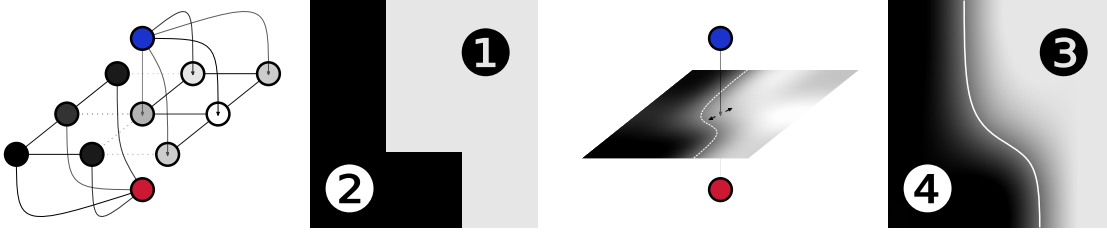
A general method that finds a minimum cut through calculating the maximum flow was described by Ford and Fulkerson [16]. Later approaches that were developed specifically for image segmentation took advantage of the regular graph structure resulting from the typically regular sampling of images, which led to less memory requirement and shorter running times. In Chapter 5, we use an implementation<sup>5</sup> of Delong and Boykov's approach [14] that adapts the push-relabel algorithm [18] to regular grids.

**Grid Bias.** The downside of explicitly modeling connections between pixels or voxels via edges is the consequence that the segmentation has to follow these edges as well. The so-called *grid bias* can lead to artefacts in the segmentation result, sometimes called *metrication errors* [47], which manifest in straight segmentation boundaries or corners in places where the segmented object is actually curved (see Fig. 8 in [4], for example). Increasing the neighborhood, that is, connecting each voxel to more neighbors, reduces the artefacts [4], albeit at the expense of increased memory requirement and running time.

## 3.5 Continuous Max Flow

Strang [43] was the first to show that the concept of finding minimum cuts via flow maximization can be applied to a continuous domain as well. In the following, we will derive a

<sup>5</sup><https://github.com/samehkhamsi/RegionPushRelabel> (last accessed on November 3, 2016)



**Figure 3.2:** Analogy between graph cut and continuous max flow in two dimensions. *Left:* Using graph cut for image segmentation, we model each pixel as a graph vertex (gray circles), which we connect to its neighbors, to a source (blue), and to a sink (red); after the cut, the segmentation foreground consists of the nodes that remain connected to the source (1), and the background consists of the nodes that remain connected to the sink (2). *Right:* Using continuous max flow, we assume a continuous image domain, in which each point is connected to a source (blue) and to a sink (red); the net flow within the domain here can have an arbitrary direction (black arrows) and the cut is given by an isoline (white line) in the segmentation result; segmentation foreground values (3) lie above the isoline value, background values (4) lie below it.

continuous flow maximization formulation for the generalized segmentation model of Section 3.3, and we will compare it with the discrete flow maximization of graph cut. In our derivation we largely follow Yuan et al. [47] but take the opposite direction, as we start from the minimum cut formulation and arrive at the flow maximization formulation. Formulating image segmentation in this way avoids the problem of grid bias, along with having some other benefits (see below).

**Minimum Cut–Maximum Flow Duality.** Recall the minimization problem of the generalized segmentation model in Eqs. (3.15)–(3.17), which we may compactly write as

$$\min_{u(x) \in [0,1]} \int_{\Omega} u C_1 dx + \int_{\Omega} (1 - u) C_0 dx + \int_{\Omega} |\nabla u| C dx. \quad (3.34)$$

As already mentioned, we assume the bounded image domain  $\Omega \subset \mathbb{R}^d$  to be continuous here. Like in the graph cut approach, we may interpret the cost functions  $C$ ,  $C_0$ , and  $C_1$  as capacity constraints for a flow through  $\Omega$ . We will now successively replace the summands in Eq. (3.34) with terms that introduce flow variables in order to construct the equivalent flow maximization or *max flow* formulation.

*First summand:* We introduce a sink flow variable  $p_t : \Omega \rightarrow \mathbb{R}$  and show that the following equality holds for any given values  $u(x) \in [0, 1]$ :

$$\int_{\Omega} u C_1 dx = \max_{p_t(x) \leq C_1(x)} \int_{\Omega} u p_t dx. \quad (3.35)$$

We can see this by distinguishing two cases in each point  $x$ : If  $u(x) = 0$ , then any value for  $p_t(x) \leq C_1(x)$  can be chosen, and the value of  $u p_t$  will always be 0. If  $u(x) > 0$ , the product will be maximized by  $p_t(x) = C_1(x)$ , reaching a maximum value of  $u(x) C_1(x)$ .

### 3 Image Segmentation

*Second summand:* We introduce the source flow  $p_s : \Omega \rightarrow \mathbb{R}$ , leading to the equality

$$\int_{\Omega} (1 - u) C_0 \, dx = \max_{p_s(x) \leq C_0(x)} \int_{\Omega} (1 - u) p_s \, dx, \quad (3.36)$$

which can be shown following the same argument as for the first summand.

*Third summand:* We introduce a vector-valued variable  $p : \Omega \rightarrow \mathbb{R}^d$  for the nonterminal flow, that is, for the flow within the image domain. Then we show that the following equality holds:

$$\int_{\Omega} |\nabla u| C \, dx = \max_{|p(x)| \leq C(x)} \int_{\Omega} u \operatorname{div} p \, dx. \quad (3.37)$$

We derive the respective equality in Appendix 8.A for the more general case of anisotropic total variation, where setting the matrix  $S$  to the identity matrix gives the result for Eq. (3.37).

*Max flow formulation:* If we now insert Eqs. (3.35)–(3.37) into Eq. (3.34), we can rewrite the segmentation as a saddle point problem, namely

$$\max_{p_s, p_t, p} \min_{u(x) \in [0,1]} \int_{\Omega} u p_t + (1 - u) p_s + u \operatorname{div} p \, dx \quad (3.38)$$

$$= \max_{p_s, p_t, p} \min_{u(x) \in [0,1]} \int_{\Omega} p_s + u (-p_s + p_t + \operatorname{div} p) \, dx, \quad (3.39)$$

subject to the capacity constraints

$$p_s(x) \leq C_0(x), \quad p_t(x) \leq C_1(x), \quad |p(x)| \leq C(x). \quad (3.40)$$

Furthermore, we can view  $u$  in Eq. (3.39) as a Lagrange multiplier that incorporates the flow conservation constraint

$$-p_s + p_t + \operatorname{div} p = 0, \quad (3.41)$$

which enables us to equivalently state the segmentation problem of Eqs. (3.34) and (3.39) as

$$\max_{p_s, p_t, p} \int_{\Omega} p_s \, dx, \quad (3.42)$$

such that the capacity constraints (3.40) and the flow conservation constraint (3.41) hold.

From an implementation point of view, Eq. (3.39) is better suited than Eq. (3.42) as the basis for an actual algorithm, because Eq. (3.39) explicitly contains the relaxed segmentation  $u$ , in contrast to Eq. (3.42). Consequently, the algorithms employed and presented in Chapters 6–8 are based on Eq. (3.39). For completeness, we note that some approaches also optimize Eq. (3.34) (e.g. Chan et al. [9]), which were not used in the context of this thesis though.

**Parallels to Graph Cut.** The conceptual parallels between graph cut and continuous max flow are visualized in Fig. 3.2 on the previous page. An overview of the equations in both approaches is given in Table 3.1 on the facing page for convenience.

In continuous max flow, like in the discrete graph cut approach, we assume each point in  $\Omega$  to be connected to a flow-emitting source  $s$  and a flow-consuming sink  $t$ . We also assume

**Table 3.1:** Energy minimization and flow maximization equations in graph cut and continuous max flow. See text for variable definitions and further detail.

Discrete setting (graph cut)		Continuous setting (max flow)	
Minimize		Minimize	
	$E(\mathcal{C}) = F(\mathcal{C}) + R(\mathcal{C}), \quad (3.26)$		$E_{\text{gM}}[u; I] = F_{\text{gM}}[u, I] + R_{\text{gM}}[u], \quad (3.15)$
with		with	
	$F(\mathcal{C}) = \sum_{x \in \Omega} u(x) C_1(x) + (1 - u(x)) C_0(x), \quad (3.27)$		$F_{\text{gM}}[u, I] = \int_{\Omega} u(x) C_1(x) + (1 - u(x)) C_0(x) dx, \quad (3.16)$
	$R(\mathcal{C}) = 1/2 \sum_{x_k, x_l \in \Omega} C_n(x_k, x_l)  u(x_k) - u(x_l) . \quad (3.28)$		$R_{\text{gM}}[u] = \int_{\Omega} C(x)  \nabla u(x)  dx. \quad (3.17)$
Maximize		Maximize	
	$\sum_{x \in \Omega} p_s(x), \quad (3.33)$		$\int_{\Omega} p_s(x) dx, \quad (3.42)$
subject to		subject to	
	$0 \leq p_s(x) \leq C_0(x), \quad (3.29)$		$p_s(x) \leq C_0(x), \quad (3.40)$
	$0 \leq p_t(x) \leq C_1(x), \quad (3.30)$		$p_t(x) \leq C_1(x), \quad (3.40)$
	$ p(x_k, x_l)  \leq C_n(x_k, x_l), \quad (3.31)$		$ p(x)  \leq C(x), \quad (3.40)$
	$-p_s(x_k) + p_t(x_k) + \sum_{x_l \in \mathcal{N}_k} p(x_k, x_l) = 0. \quad (3.32)$		$-p_s(x) + p_t(x) + \text{div } p(x) = 0. \quad (3.41)$

the flow to be constrained by three capacities:  $C_0(x)$  for the source flow  $p_s$ ,  $C_1(x)$  for the sink flow  $p_t$ , and  $C(x)$  for the nonterminal flow  $p$ , as stated in Eq. (3.40). Likewise, we assume flow conservation for the points in  $\Omega$ , as stated in Eq. (3.41). On page 16, we have already commented on the similarities between Eqs. (3.26)–(3.28) from graph cut and Eqs. (3.15)–(3.17). While the latter equations originate from the generalized Mumford–Shah model, they may just as well be seen as a continuous minimum cut formulation. Now also note the similarities between Eqs. (3.29)–(3.33) from graph cut and Eqs. (3.40)–(3.42) from max flow. Apart from replacing the sum in Eq. (3.33) with an integral in Eq. (3.42) and dropping the lower bounds in the nonterminal flow constraints of Eq. (3.40), their difference lies in the divergence  $\text{div } p$  of Eq. (3.41), which takes on the role of the summed nonterminal edge flows of Eq. (3.32), and which is defined as

$$\text{div } p := \sum_{i=1}^d \frac{\partial p_i}{\partial x_i} \quad \text{with} \quad p = (p_1, \dots, p_d)^\top, \quad x = (x_1, \dots, x_d)^\top. \quad (3.43)$$

A negative value in  $\text{div } p$  signifies incoming net flow and a positive sign signifies outgoing net flow within the domain for the current point. Just as with  $R(\mathcal{C})$ , the graph cut formulation may here be interpreted as a partial derivative estimate of the continuous formulation. More specifically, in graph cut, the sum of flows for each opposing pair of neighbors (e.g. left and

### 3 Image Segmentation

right neighbor) can be seen as an estimate of the corresponding partial derivative in  $\text{div } p$ .<sup>6</sup>

**Benefits.** Formulating image segmentation as a continuous max flow problem has at least three benefits. First, compared to the discrete graph cut approach, we avoid the problem of grid bias, as we do not explicitly model the neighborhood relationship between neighboring points in the domain and, as a consequence, do not constrain the segmentation to follow any particular choice of neighborhood model.

Second, as can be seen from our description of relaxing the segmentation model in Section 3.3, the *cut* is now given by the level set of value  $\ell$  (i.e. the  $\ell$ -isoline in  $d = 2$  or the  $\ell$ -isosurface in  $d = 3$ ) in the relaxed segmentation  $u$ . As a consequence, the segmentation is thus not bound to the sampling positions, as is the case in graph cut.

Third, the continuous max flow formulation can be optimized with comparatively simple and straightforward algorithms. This is true, for example, for the algorithms proposed by Yuan et al. [47], which we employ and adapt in Chapters 6–8. Additionally, they are very well suited for implementation on massively parallel devices like GPUs (i.e. standard computer graphics cards), which may significantly reduce their runtime compared to largely sequential implementations.

---

<sup>6</sup>Recall that in the graph cut case, we have  $p(x_k, x_l) = -p(x_l, x_k)$ , which means that the sign, and thus the direction, of the flow is to be seen relative to the current point rather than absolute (i.e. relative to the coordinate system), as is the case in the continuous max flow formulation. As a consequence, the sum of flows to opposing neighbors in the graph cut case corresponds to a differential expression (namely the respective partial derivative) in the max flow case.



## 4 Segmenting the Spinal Cord

The clinical interest of our project, as has been motivated in Sections 1.1 and 2.2, lies in quantifying spinal cord volumes and cross-sectional areas, along with their changes. More specifically, we aim at measuring atrophy in the spinal cord in the same patient over time (longitudinal setting), and we are interested in comparing spinal cord measurements between different groups (cross-sectional setting), such as between healthy subjects and MS patients or between patient groups of different MS subtypes.

**General Idea.** The task of measuring atrophy in a longitudinal setting can be framed as an image segmentation approach, or at least as an approach that requires image segmentation, by proceeding as follows. First, we segment the same part of the spinal cord in multiple images of the same subject, using one of the previously described image segmentation techniques, for example. We then quantify the segmented part by measuring its cross-sectional area in square millimeters or its volume in cubic millimeters. Measurements in real world units are usually possible, because medical imaging modalities provide the necessary information about sampling distances in the metadata of the image files. Finally, by calculating differences between images that were acquired at different points in time, we get a measure of *absolute volume loss* for the subject, from which we can infer secondary measures like *annualized atrophy rate* (see Chapter 7) or *percentage volume loss* [29]. Similar approaches can be chosen in a cross-sectional setting, for example by comparing mean spinal cord cross-sectional areas of the different groups (see Chapter 5, for example).

**Challenges.** While the devised approach appears straightforward, it brings about several challenges that we have to tackle (cf. Fig. 4.1 on page 25).

*Image quality:* MR images are prone to noise, certain types of artefacts, and intensity inhomogeneities, which prevents the use of basic segmentation techniques like thresholding and calls for more robust approaches.

*Image resolution:* In current clinical routine, the resolution of spinal cord MR images is usually on the order of 1 mm. With a cord diameter on the order of 10 mm (see Section 2.1), this renders segmentation approaches that are bound to the voxel grid unfeasible.

*Image distortion:* MR image distortion due to gradient nonlinearity makes direct volume quantification and comparison difficult and requires distortion correction (see the experiments in Chapter 5).

*MR contrast:* Different MR sequences bring about different appearances of the imaged tissue, which may make adjustments of the segmentation approach necessary.

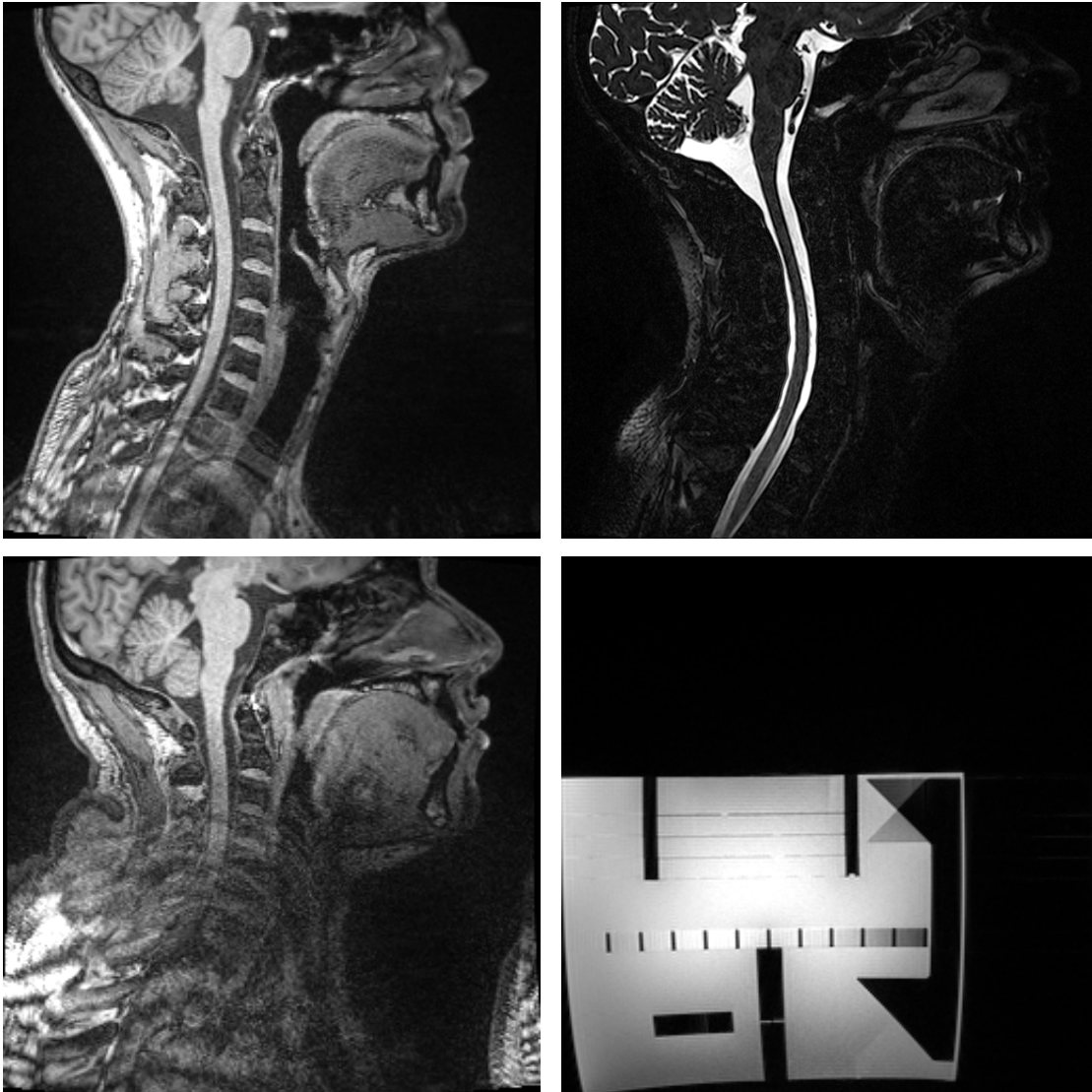
#### 4 Segmenting the Spinal Cord

*Physiological variability:* Among different subjects, the spinal cord size itself, the size of the enclosing spinal canal, and the amount of CSF may vary. Additionally, depending on different bending of the spinal cord, the nervous tissue may or may not touch the surrounding vertebrae in certain positions. Robust spinal cord segmentation approaches have to be capable of handling these degrees of variability.

*Subject positioning:* As the spine is a flexible structure, its bending may vary between scans, depending on how a subject is placed in the MR scanner. This makes it difficult to measure the same segment of the spinal cord repeatedly and makes it necessary to account for the bending.

*Comparability:* In addition to the previous point, the more general question arises as to what is the *same segment* of the spinal cord. In a longitudinal setting, on top of the bending, the spinal cord may also slide in the spinal canal, which brings about the need for finding reliable landmarks to relocate a previously measured spinal cord segment. In a cross-sectional setting, in addition to that, certain types of measurement normalization may be necessary to enable meaningful comparisons between different subjects.

**Proposed Methods.** Starting from the pioneering work of Losseff et al. [27], a number of methods for quantitative measurements of the spinal cord have been proposed. Their main differences lie in the type of measurement (cross-sectional area or volume), the choice of landmarks and the dealing with spine bending for maintaining measurement comparability, the degree of user interaction, the applicability to different kinds of MR sequences, the measurement target (the whole spinal cord or its inner gray and white matter structures), and finally the chosen image segmentation approach itself. A brief summary of important methods along with a list of references is given in the introductory section of our publication in Chapter 7. De Leener et al. [13] give a more extensive and complete overview in their 2016 review article. The methods proposed in the following chapters constitute our contribution to the literature to date.



**Figure 4.1:** Challenges of spinal cord segmentation in MR images. *Upper row:* Example of different MR contrasts; both images show the same subject in approximately the same position, using a T1-weighted MPRAGE sequence (*upper left*) and a T2-weighted 3D TSE sequence (*upper right*); note the different appearance of the spinal cord and its immediate surroundings. *Lower left:* Example of a very noisy image; the spinal cord is barely visible in the lower half of the image. *Lower right:* Example of MRI distortions; the image shows a perspex phantom whose outline is actually rectangular; note how the lower corners bend inward considerably.



## 5 A Semi-Automatic Method for the Quantification of Spinal Cord Atrophy

The publication presented in this chapter describes the initial version of our toolchain for spinal cord atrophy quantification. In a first, interactive presegmentation step, which uses *graph cut*, we acquire a binary mask of the spinal cord. In a second step, the cord surface is delineated more precisely using edge detection in one-dimensional intensity profiles, which enables us to create a *surface reconstruction* of the selected spinal cord segment. In a last step, the surface is reformatted so that either *cross-sectional area* or *volume measurements* perpendicular to the spinal cord centerline can be acquired.

The toolchain is evaluated on phantom images as well as images of both healthy subjects and multiple sclerosis patients. Phantom experiments show accurate cross-sectional area measurements that slightly but consistently overestimate the true cross-sectional area. Results of measurement reproducibility in healthy subjects compare favorably with reported results of state-of-the-art spinal cord segmentation approaches. Patient data experiments confirm previously observed correlations between spinal cord volume and disability status, which suggests that the presented toolchain is actually suited for application to clinical data.

**Publication.** The proposed approach was presented at the workshop *Computational Methods and Clinical Applications for Spine Imaging* (CSI 2013) in conjunction with the *16th International Conference on Medical Image Computing and Computer Assisted Intervention* (MICCAI), September 2013, Nagoya, Japan. It was published<sup>1</sup> as part of the workshop proceedings [34].

---

<sup>1</sup>[http://dx.doi.org/10.1007/978-3-319-07269-2\\_13](http://dx.doi.org/10.1007/978-3-319-07269-2_13) (last accessed on November 3, 2016)

# A Semi-automatic Method for the Quantification of Spinal Cord Atrophy

Simon Pezold, Michael Amann, Katrin Weier, Ketut Fundana,  
Ernst W. Radue, Till Sprenger and Philippe C. Cattin

**Abstract** Due to its high flexibility, the spinal cord is a particularly challenging part of the central nervous system for the quantification of nervous tissue changes. In this paper, a novel semi-automatic method is presented that reconstructs the cord surface from MR images and reformats it to slices that lie perpendicular to its centerline. In this way, meaningful comparisons of cord cross-sectional areas are possible. Furthermore, the method enables to quantify the complete upper cervical cord volume. Our approach combines graph cut for presegmentation, edge detection in intensity profiles for segmentation refinement, and the application of starbursts for reformatting the cord surface. Only a minimum amount of user input and interaction time is required. To quantify the limits and to demonstrate the robustness of our approach, its accuracy is validated in a phantom study and its precision is shown in a volunteer scan–rescan study. The method’s reproducibility is compared to similar published quantification approaches. The application to clinical patient data is presented by comparing the cord cross-sections of a group of multiple sclerosis patients with those of a matched control group, and by correlating the upper cervical cord volumes of a large MS patient cohort with the patients’ disability status. Finally, we demonstrate that the geometric distortion correction of the MR scanner is crucial when quantitatively evaluating spinal cord atrophy.

## 1 Introduction

Multiple sclerosis (MS) is a chronic inflammatory disorder of the central nervous system that causes both motor disability and cognitive impairment. So far, the diagnosis and disease monitoring has been based on characteristic patterns of lesions in the central nervous system that evolve during the disease progression. In recent years, however, it has been shown that neurodegenerative processes play a central role in

---

S. Pezold (✉) · M. Amann · K. Weier · K. Fundana · E. W. Radue · T. Sprenger · P. C. Cattin  
University Hospital Basel, Basel, Switzerland  
e-mail: simon.pezold@unibas.ch

J. Yao et al. (eds.), *Computational Methods and Clinical Applications* 143  
*for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 17,  
DOI: 10.1007/978-3-319-07269-2\_13, © Springer International Publishing Switzerland 2014

and may be a key to the development of disability [10]. A hallmark of neurodegeneration is atrophy; that is, the loss of nervous tissue. Atrophy can be investigated on a macroscopic scale by using magnetic resonance imaging (MRI), and it has been shown to correlate better with clinical disability than lesion patterns [10]. More specifically, spinal cord (SC) atrophy has been suggested as a biomarker for disease progression, due to the critical role of the SC in motor control [8].

During the last decade, several approaches have been proposed and applied to measure SC cross-sectional areas (CSAs) and volume (e.g., see Miller et al. [9] and Bakshi et al. [2] for methodological overviews), including manual tracing of the SC border as well as semi-automated, intensity-based tracing and subsequent measurement of the resulting CSAs. Common to all these approaches, however, is their requirement for significant user input. A higher degree of automation is therefore desirable to reduce the amount of user input and, with it, the amount of time needed for the usually tedious tasks of manual measurements. More recently, two approaches have been introduced that automatically reconstruct the surface of a manually selected SC section and then successively straighten the result by reformatting it with respect to the SC centerline [4, 6]. In this way, they make it possible to simultaneously assess SC volume changes in a larger region compared to previous approaches.

In this publication, we present a semi-automatic technique that reconstructs the cervical section of the SC surface and then either reformats it to slices perpendicular to the SC centerline or measures the volume of a perpendicularly clipped SC segment. Our method requires only little user input: two small sets of labeled voxels marking both the SC and the background, and one user-provided anatomical landmark to indicate a starting point for the reformatting process. We evaluate the accuracy of our method via a phantom structure of known dimensions. Furthermore, its precision is assessed by analyzing scan-rescan datasets of healthy volunteers. Finally, the applicability to clinical data is shown by comparing the mean CSA of a group of MS patients with an age-matched group of healthy subjects and by referring to a study where our method was successfully applied to correlate upper cervical cord volume with MS disability status. In contrast to the above mentioned methods [4, 6], we put a special focus on how MRI-specific image distortions influence measurements by showing that a distortion correction routine may improve the reproducibility of measurements.

## 2 Materials

To assess the accuracy of our method, a cylindrical perspex phantom filled with copper sulfate-doped water was scanned on a 1.5T whole-body MR scanner (Avanto, Siemens Medical, Germany) with a T1-weighted MPRAGE sequence (TR/TI/TE/ $\alpha = 2.1$  s/1.1 s/3.1 ms/15°); 192 slices in sagittal orientation were acquired with an in-slice resolution of 0.98 mm  $\times$  0.98 mm and a slice thickness of 1 mm. The phantom was scanned in 11 different  $z$ -positions relative to the magnetic field center (−50 to 50 mm in increments of 10 mm). The manufacturer's three-dimensional

distortion correction routine was applied to the data to see the effects of image distortions induced by gradient non-linearity [7]. Both original and corrected datasets were reconstructed.

To assess the scan–rescan reliability, 12 healthy volunteers (3 female, 9 male, mean age 32.4 y, range 26–44 y) were scanned on a 3T whole-body MR scanner (Verio, Siemens Medical, Germany) with a T1-weighted MPRAGE sequence (TR/TI/TE/ $\alpha$  = 2.0 s/1.0 s/3.4 ms/8°); 192 slices in sagittal orientation parallel to the interhemispheric fissure were acquired with an isotropic resolution of 1 mm<sup>3</sup>. Both original and distortion-corrected datasets were reconstructed.

To show the applicability to clinical data, 12 relapsing-remitting MS patients (8 female, 4 male, mean age 32.2 y, range 21–46 y; mean disease duration 8.2 y, range 1–17 y, median EDSS 3.0, range 1–4) and 12 age-matched controls (6 female, 6 male, mean age 31.6 y, range 22–48 y) were scanned on a 3T whole-body MR scanner (Verio, Siemens Medical, Germany) with a T1-weighted MPRAGE sequence (TR/TI/TE/ $\alpha$  = 1.6 s/0.9 s/2.7 ms/9°); 192 slices in sagittal orientation parallel to the interhemispheric fissure were acquired with an isotropic resolution of 1 mm<sup>3</sup>. Distortion-corrected datasets were reconstructed.

### 3 Method

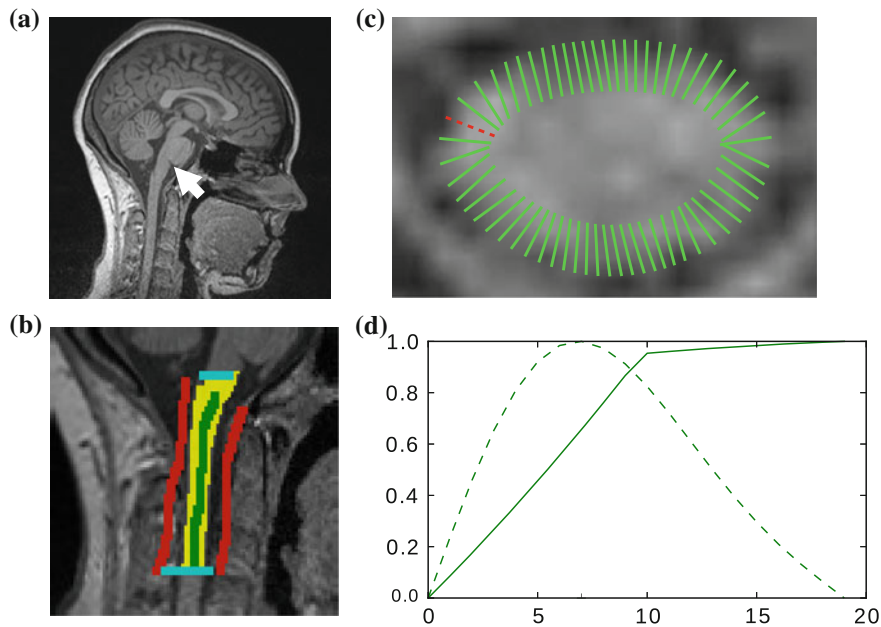
The proposed method can be broken down into four distinct steps, which we refer to as *presegmentation*, *segmentation refinement*, *surface reconstruction*, and *reformatting*. Of these steps, only presegmentation and reformatting need manual intervention while the others run in a completely automated manner. In this way, the user interaction time lies in the order of two to five minutes per scan.

#### 3.1 Presegmentation

The aim of the presegmentation step (see Fig. 1b) is to gain a binary voxel mask that roughly separates the SC section of interest from the background; that is, from the surrounding cerebrospinal fluid (CSF) and all non-cord tissue. While in principle any kind of thresholding technique could be applied, we use graph cuts [3] because of their flexibility and speed. To compensate for intensity differences caused by field inhomogeneities, we apply a bias field correction [11] to the image volumes beforehand. Furthermore, we normalize the image intensities to the [0, 1] interval. We then build a six-connected graph from the voxels around the region of interest (which the user may sketch in transverse, sagittal, and coronal projections of the image volume).

The t-link weights are calculated based on a naive Bayes classifier via the intensity distributions of a set of foreground and background seed points; that is, a selection of voxels labeled by the user as definitely belonging either to the SC or its surroundings.





**Fig. 1** **a** Location of the *cisterna pontis*. **b** Presegmentation (blue region of interest, red background seeds, green cord seeds, yellow result). **c+d** Refinement: **c** intensity profile locations, **d** single intensity profile (solid) with smooth derivative estimate (dashed), both normalized for display

We model the foreground as a univariate normal distribution and the background as a mixture of four Gaussians. More specifically, we calculate the weights  $w_{fg}(x)$  and  $w_{bg}(x)$  for the t-links that connect voxel  $x$  to the foreground and background terminal, respectively, as

$$w_{fg}(x) = \begin{cases} \infty & \text{if } x \in \mathcal{F} \\ 0 & \text{if } x \in \mathcal{B} \\ 1 - w_{bg}(x) & \text{else} \end{cases} \quad \text{and} \quad w_{bg}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{F} \\ \infty & \text{if } x \in \mathcal{B} \\ \frac{p_{bg}(I(x))}{p_{bg}(I(x)) + p_{fg}(I(x))} & \text{else} \end{cases},$$

where  $\mathcal{F}$  and  $\mathcal{B}$  are the sets of foreground and background seed points,  $I(x)$  is the intensity of voxel  $x$ , and  $p_{fg}$  and  $p_{bg}$  are the probability density functions that we estimated from the foreground and background seed point intensities.

The n-link weights  $w(x_a, x_b)$  between neighboring voxels  $x_a$  and  $x_b$  are calculated as  $w(x_a, x_b) = \kappa \exp(-0.5\zeta^{-2}(I(x_a) - I(x_b))^2)$ , where  $I(x_a)$  and  $I(x_b)$  are the respective voxel intensities,  $\kappa$  is a weighting factor, and  $\zeta$  determines the spread of the Gaussian-shaped function (with smaller values for  $\zeta$  leading to a faster decrease of  $w(x_a, x_b)$  for increasing differences  $I(x_a) - I(x_b)$ ). The presegmentation is concluded by connected-component labeling, assuring that only the region that includes foreground seeds is retained.

### 3.2 Segmentation Refinement

Let  $I: \Omega \rightarrow \mathbb{R}$  denote the preprocessed (i.e., bias field corrected) image, and let  $\mathbf{x} = (x_1, x_2, x_3)^\top \in \Omega$  denote the voxel indices in the image domain  $\Omega \subset \mathbb{N}^3$ . Furthermore, let  $M \subset \Omega$  denote the set of foreground mask voxel indices from the pre-segmentation step. To reduce noise in  $I$ , we apply the `GradientAnisotropic-DiffusionImageFilter` of ITK<sup>1</sup> with the conductance parameter fixed to 3.0, the time step fixed to 0.05, and the number of iterations fixed to 20 and 5 for the image used in the first and second pass, respectively. In this way, we yield the denoised images  $\hat{I}_1(\mathbf{x})$  and  $\hat{I}_2(\mathbf{x})$ .

In the first pass, for each transversal slice, we determine the mask boundary voxel indices  $B_z$  as

$$B_z = M_z \setminus (M_z \ominus S_4), \quad (1)$$

where  $\ominus$  denotes the morphological erosion operator,  $S_4$  is the two-dimensional structuring element representing four-connectivity, and  $M_z = \{(x_1, x_2, x_3)^\top \in M: x_3 = z\}$  is the subset of mask voxels for the  $z$ -th transversal slice.

We then fit a periodic smoothing B-spline [5]  $\mathbf{s}_z(\tau)$  of degree three through the ordered voxel indices  $\mathbf{b}_i^z \in B_z$ . We distribute the spline's knots  $t_i \in [0, 1]$  ( $i = 1, \dots, |B_z|$ ) according to

$$t_i = \frac{t_i^*}{|B_z|} \quad \text{with } t_1^* = 0 \quad \text{and} \quad t_i^* = t_{i-1}^* + \|\mathbf{b}_i^z - \mathbf{b}_{i-1}^z\|, \quad (2)$$

and constrain the spline smoothness by a smoothing parameter  $s$  via

$$\|\mathbf{b}_i^z - \mathbf{s}_z(t_i)\|^2 \leq s. \quad (3)$$

The order of the boundary voxels  $\mathbf{b}_i^z$  is determined by calculating an estimate of their centroid as  $\hat{\mathbf{c}}_z = \frac{1}{|B_z|} \sum_i \mathbf{b}_i^z$  and then sorting them according to the angles formed by the  $x_1$ -axis and the vectors  $\mathbf{b}_i^z - \hat{\mathbf{c}}_z$ . Once we have  $\mathbf{s}_z(\tau)$ , we divide it into  $n_1$  sections of equal arc length, yielding  $n_1$  new vertices  $\mathbf{u}_j^z \in \mathbf{s}_z(\tau)$  ( $j = 1, \dots, n_1$ ) at the section endpoints. For each  $\mathbf{u}_j^z$ , we then extract a one-dimensional intensity profile (see Fig. 1c+d)  $P_j^z(x)$  with  $x \in [0 \dots k_1 - 1]$  as

$$P_j^z(x) = \hat{I}_1(\mathbf{v}_j^z + \delta(x)) \quad \text{with } \delta(x) = d_1 \cdot \left(x - \frac{k_1-1}{2}\right) \mathbf{n}_j^z \quad \text{and} \quad \mathbf{v}_j^z = \mathbf{u}_j^z - o \mathbf{n}_j^z, \quad (4)$$

where  $\mathbf{n}_j^z$  denotes the unit normal vector pointing inside the spline curve at  $\mathbf{u}_j^z$ , the resampling distance is given via  $d_1 \in \mathbb{R}$ , the number of profile samples via  $k_1 \in \mathbb{N}$ , and  $o \in \mathbb{R}$  is an offset to control the profile centering with respect to  $\mathbf{u}_j^z$ , yielding offset-corrected vertices  $\mathbf{v}_j^z$ . As our approach to calculate  $B_z$  systematically underestimates the mask boundary by half a voxel, we set  $o = 0.5$ . To get the

<sup>1</sup> <http://www.itk.org/>.

intensity values at the resampling positions  $x$ , we use bilinear interpolation in the  $z$ th transversal slice of  $\hat{I}_1$ .

Given that the extracted profiles point to the inside of the spline curve and knowing that in T1-weighted images the inside (i.e., the SC tissue) typically appears brighter than its immediate surroundings (i.e., the CSF), we try to refine the  $\mathbf{v}_j^z$  by means of edge detection in the profiles  $P_j^z$ . We thus calculate their derivatives as  $P_j^{z'}(x) = G'(\sigma) * P_j^z(x)$ , where  $G'(\sigma)$  is the spatial derivative of a Gaussian kernel with standard deviation  $\sigma$  and zero mean. We then search for all local maxima  $x_{m,j}$  in  $P_j^{z'}$  and calculate the new boundary estimate  $\mathbf{w}_j^z$  as

$$\mathbf{w}_j^z = \mathbf{v}_j^z + \delta(\hat{x}_{m,j}) \quad \text{with} \quad \hat{x}_{m,j} = \arg \min_{x_{m,j}} \|\delta(x_{m,j})\|. \quad (5)$$

To be less susceptible to noise, we dismiss all  $x_{m,j}$  with  $P_j^{z'}(x_{m,j}) < c \cdot P_j^{z'}(x_{\max,j})$  beforehand, where  $c \in [0, 1]$  serves as a threshold and  $x_{\max,j}$  is the global maximum position of  $P_j^{z'}$ . If no valid maxima are retained, which may be the case if the global maximum value is negative, we set  $\mathbf{w}_j^z = \mathbf{v}_j^z$ .

A second pass of boundary estimation follows, similar to the first pass, starting with the  $\mathbf{w}_j^z$  as initial estimate. The only differences are the following: first, to ensure a homogeneous distribution of the boundary estimates, particularly with regard to the surface reconstruction step (Sect. 3.3), the fitted spline is now resampled  $n_2$  times at equal *angular* distances, using the spline center as point of reference, yielding redistributed estimates  $\mathbf{y}_j^z$  ( $j = 1, \dots, n_2$ ). Second, the new intensity profiles  $Q_j^z(x)$  with  $x \in [0..k_2 - 1]$  are extracted as

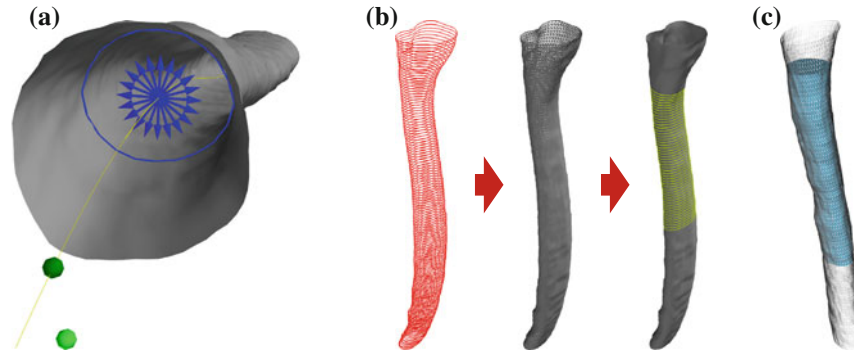
$$Q_j^z(x) = \hat{I}_2 \left( \mathbf{y}_j^z + d_2 \cdot \left( x - \frac{k_2-1}{2} \right) \frac{\nabla \hat{I}_2^z(y_1, y_2)}{\|\nabla \hat{I}_2^z(y_1, y_2)\|} \right) \quad \text{with} \quad (y_1, y_2, z)^\top = \mathbf{y}_j^z; \quad (6)$$

that is, no offset is added and the normalized gradient vectors of the in-slice intensities  $\hat{I}_2^z$  at  $\mathbf{y}_j^z$  replace the spline normal vector. The boundary re-estimation is calculated analogous to Eq. (5), leading to the final boundary position  $\mathbf{z}_j^z$ .

As a result of the refinement procedure, we now have  $n_2$  vertices  $\mathbf{z}_j^z$  for all of those transversal slices that contain foreground mask voxels. These vertices represent the slice-wise SC contour at subvoxel precision, provided that the profile resampling distances  $d_1$  and  $d_2$  were chosen sufficiently small.

### 3.3 Surface Reconstruction

We transform the  $\mathbf{z}_j^z$  to their locations  $\mathbf{a}_j^z$  in the metric world coordinate system by means of a transformation matrix determined from the image source's meta data. We then connect each  $\mathbf{a}_j^z$  to  $\mathbf{a}_j^{z+1}$ ,  $\mathbf{a}_{j+1}^z$ , and  $\mathbf{a}_{j+1}^{z+1}$ , which results in strips of  $2n_2$



**Fig. 2** **a** Reformatting (yellow centerline, green landmark and anchor point, blue simplified starburst and new contour). **b** Schematic reformatting steps: contour stacking, surface reconstruction, reslicing. **c** Perpendicular volume clipping (gray original surface, blue clipped volume segment)

triangles between each pair of successive slices and thus in a complete SC surface reconstruction  $S$  for the initially masked region of interest (see Fig. 2b).

### 3.4 Reformatting

To compare different image volumes, spatial correspondence has to be established between them. In our setting we have to make sure that anatomically corresponding SC locations are compared. This is not straightforward for two reasons: first, the SC can be bent rather differently in the head and neck area between scans; and second, the SC slides along the spinal canal during this bending, making landmarks such as the vertebrae or intervertebral discs unsuitable.

We identified the *cisterna pontis*, a distinct indentation at the caudal *pons*, as a landmark that may easily be spotted and manually marked (see Fig. 1a). Moreover, the landmark is a structure that is part of the nervous tissue and, as such, stays in a fixed position relative to the SC. This may be a benefit compared to other features located on the intervertebral discs [6, 8] or on bone structures such as the *foramen magnum* [4], which are more likely to be susceptible to relocation due to bending. To the best of our knowledge, the *cisterna pontis* has not been described as a landmark in the context of SC surface reconstructions.

Concerning the bending, we propose a similar approach as previously described by Coulon et al. [4]. Let  $\mathbf{c}_z$  describe the centroid of the polygon formed by the vertices  $\mathbf{a}_j^z$ ; let  $z = 1, \dots, m$  here serve as an index variable that consecutively numbers all  $m$  transversal slices containing parts of the surface reconstruction  $S$ . We then define the centerline  $\mathbf{c}(\tau)$  of  $S$  as a smoothing B-spline of degree three through the  $\mathbf{c}_z$  (Sect. 3.2). Let  $\mathbf{p} \in \mathbb{R}^3$  denote the position of the *cisterna pontis* landmark in world coordinates. The anchor point  $\mathbf{c}_0 = \mathbf{c}(\tau_0)$  for reformatting is then determined by the condition

$(\mathbf{p} - \mathbf{c}_0) \cdot \mathbf{c}'_0 = 0$ , where  $\mathbf{c}'_0$  is a spline tangent vector in  $\mathbf{c}_0$ ; in other words,  $\mathbf{c}_0$  is an orthogonal projection of  $\mathbf{p}$  onto the spline curve. We extrapolate the superior spline end if necessary to find a  $\tau_0$  and  $\mathbf{c}_0$  that satisfy the condition. Depending on whether one intends to measure reformatted CSAs or the volume of a SC section, either the reslicing steps or the volume measurement steps described below follow.

**Reslicing.** For measuring reformatted CSAs, a total of  $n_3$  reslicing positions  $\mathbf{c}_i = \mathbf{c}(\tau_i)$  ( $i = 1, \dots, n_3$ ) are determined by solving

$$\int_{\tau_0}^{\tau_i} \left\| \frac{d\mathbf{c}}{d\tau} \right\| d\tau = \omega + (i - 1)d_3 \quad (7)$$

with respect to  $\tau_i$ ; in other words, the  $\mathbf{c}_i$  are calculated at intervals of equal arc length  $d_3$  along the spline curve, starting from  $\mathbf{c}_0$ , with an offset  $\omega$ .

Based on the  $\mathbf{c}_i$ , we want to reformat the surface reconstruction  $S$  to slices that lie perpendicular to the centerline  $\mathbf{c}$ . Formally speaking, we thus require for each  $\mathbf{c}_i$  the set of surface points  $C_i \subset S$  that satisfy  $C_i = R_i \cap S$  with  $R_i = \{\mathbf{r}: (\mathbf{r} - \mathbf{c}_i) \cdot \mathbf{c}'_i = 0\}$ ; that is, we require the points that lie in the intersections of  $S$  with the planes  $R_i \perp \mathbf{c}$  through the reslicing positions  $\mathbf{c}_i$ .

In practice, we calculate approximations of these intersections. For this, we build  $n_3$  bundles of rays, hereafter referred to as ‘starbursts’. Each starburst (see Fig. 2a) consists of  $n_4$  rays  $\mathbf{q}_j^i$  ( $j = 1, \dots, n_4$ ) given by

$$\mathbf{q}_j^i(\lambda) = \mathbf{c}_i + \lambda \mathbf{r}_j^i \quad \text{with } \lambda \geq 0, \quad \|\mathbf{r}_j^i\| = 1, \quad \text{and } \mathbf{r}_j^i \cdot \mathbf{c}'_i = 0. \quad (8)$$

The direction vectors  $\mathbf{r}_j^i$  are directed at equal angular intervals around  $\mathbf{c}_i$ ; that is,  $\forall j: \mathbf{r}_j^i \cdot \mathbf{r}_{j+1}^i = \cos(\frac{2\pi}{n_4})$ . The actual reslicing procedure amounts to a series of ray–triangle intersections of all  $\mathbf{q}_j^i$  with the triangles that form  $S$ . As a result, we get a new set of  $n_3 \cdot n_4$  vertices  $\mathbf{b}_j^i \in C_i$ , which for each of the  $n_3$  positions may be connected to a polygon serving as a contour representation for the respective reformatted slice. These contours are finally suitable for CSA measurements that make inter-scan comparisons possible.

**Volume Measurement.** For measuring the volume of an SC section of length  $l$ , the surface reconstruction  $S$  is clipped by two planes that are located at arc lengths  $\omega$  and  $\omega + l$  measured along the centerline (see Eq. 7) and that lie perpendicular to the respective centerline tangent vectors. We then close the ends of the clipped section and calculate the volume of the resulting closed surface (see Fig. 2c) based on the divergence theorem [1].

## 4 Results

To assess the performance of our proposed algorithm, we conducted experiments on phantom, scan–rescan, and real patient data. For all these experiments, the following parameter settings were applied:  $\kappa = 0.4$ ,  $\zeta = 0.5$ ,  $s = 3.0$  voxels (splines in the refinement step),  $s = 3.5$  mm (centerline splines),  $n_{1,2,4} = 60$ ,  $k_{1,2} = 20$ ,  $d_1 = 0.2$  voxels,  $d_2 = 0.1$  voxels,  $c = 0.3$ ,  $\sigma = 0.5$  voxels,  $d_3 = 0.25$  mm. If not stated differently, reslicing took place in superior–inferior direction over a length of 50 mm, resulting in  $n_3 = 201$  new slices.

**Phantom Evaluation.** Among other structures, the used phantom contains a solid cylindrical structure, surrounded by a liquid-filled cavity, of 60 mm length and 25 mm diameter (corresponding to a CSA of 490.9 mm<sup>2</sup>), which was roughly aligned with the scanner’s  $z$ -axis during the scans. As a substitute for the SC landmark, we placed a marker at the most posterior point of the structure’s boundary in the most superior slice where its CSA was still completely visible and set  $\omega = 2$  mm.

In the uncorrected scans, the mean CSA was  $504.7 \pm 1.8$  mm<sup>2</sup>, thus the true CSA was overestimated by approximately 2.8 %. In the corrected scans, the mean CSA was  $503.0 \pm 1.3$  mm<sup>2</sup>, thus the true CSA was overestimated by approximately 2.5 %.

**Scan–Rescan Evaluation.** For the scan–rescan evaluation, the twelve subjects were scanned three times in a row (scans S1, S2, S3). Between S1 and S2 they were asked not to move so that the SC location and bending would be as similar as possible in both scans. Between S2 and S3 the subjects had to exit the scanner and were repositioned afterwards. One subject was excluded because the *cisterna pontis* was not visible in all scans. To measure the intra-observer reproducibility, the distortion-corrected S1 scans were presegmented twice by one observer, with a period of more than one month between associated presegmentations. To measure the inter-observer reproducibility, the distortion-corrected S1 scans were presegmented independently by a second observer.

We focused on two different aspects for interpreting the results: the effects of the distortion correction routine, and the comparison to published SC segmentation methods.

In a first experiment (E1), we set the reslicing offset to  $\omega = 50$  mm and measured the CSAs in superior–inferior direction over a length of 15 mm (resulting in 61 new slices), approximately covering the caudal end of the C2 vertebra. We then averaged the CSA for the reformatted contours of each scan and calculated the coefficient of variation (CV; i.e., the sample standard deviation over the mean) for comparisons of S1 and S2, S1 and S3 (both in the corrected and the uncorrected scans), as well as for the intra-observer and inter-observer comparisons. This setup was chosen in order to compare our results with the method described by Losseff et al. [8], who measure the average CSA of the caudal C2 end over five slices of 3 mm thickness that were reformatted to lie perpendicular to the SC orientation.

In a second experiment (E2), in order to compare our results with the methods described by Coulon et al. [4] and Horsfield et al. [6], we set the reslicing offset

**Table 1** Coefficients of variation, distortion-corrected versus uncorrected scans (%)

	E1: C2		E2: C1–C3		E2: C1–C3		E3: C1–C3	
	Avg. CSA		Avg. CSA		Volume		Slice-wise CSA	
	S1–S2	S1–S3	S1–S2	S1–S3	S1–S2	S1–S3	S1–S2	S1–S3
Uncorrected	0.42	1.55	0.36	1.90	0.35	1.89	1.16	2.48
Corrected	0.42	<b>1.15</b>	<b>0.33</b>	<b>0.94</b>	<b>0.33</b>	<b>0.93</b>	<b>1.10</b>	<b>1.76</b>

to  $\omega = 25$  mm and segmented over our default length of 50 mm. In this way, we covered a wider region of the cervical SC, namely approximately the section between the cranial end of the C1 and the caudal end of the C3 vertebra. For each scan, we then calculated the mean CSA over the reformatted contours of the complete region as well as the region's perpendicularly clipped volume. The CVs were calculated for the same combinations as described in the first experiment.

In a third experiment (E3), we wanted to see whether our method is applicable to not only measure a mean CSA, but also to measure the CSA on specific levels along the SC reliably. We thus swapped the order of averaging and CV calculation: first, we calculated a CV for the contours on the same level (e.g., the CV for the first reformatted contour in the S1–S2 comparison for subject one), then we averaged over the CVs for each comparison.

*Effects of the distortion correction routine.* A comparison of the mean CVs over all subjects for the distortion-corrected and uncorrected scans is shown in Table 1. The results conform to our expectations. As the S1–S2 position change was minimal, the image distortions have little influence here. This is because they affect both scan and rescan similarly, and thus the S1–S2 CVs are on the same level for the corrected and uncorrected scans. After repositioning (i.e., S1–S3), however, the uncorrected data CVs become distinctly worse than the corrected data CVs, showing the benefit of the correction routine as soon as scan conditions are not perfectly similar anymore. Furthermore, the S1–S3 corrected data CVs are also worse than their S1–S2 equivalents, which suggests that the correction routine does not account for all geometric distortions in the scans. Partial volume effects, for example, may be another contributor to variations in the reconstructions and thus in the subsequent measurements.

*Comparison with published methods.* For all experiments, the mean CVs over all subjects are reported in Table 2, making use of the distortion-corrected scans for our method. The results of Experiments 1 and 2 compare favorably with the values reported by Losseff et al. [8], Coulon et al. [4], and Horsfield et al. [6]. The low intra- and inter-observer CVs show the strength of our method in that the outcome is very robust given different presegmentation inputs, whether produced by the same observer or different observers.

The results of Experiment 3 indicate that even measurements on single-contour level may produce reasonable outcomes with our method. To put these values into perspective: given a realistic CSA of  $75 \text{ mm}^2$ , changes of  $0.91 \text{ mm}^2$  (1.96 standard deviations) on a specific slice should be detectable in distortion-corrected scans with

**Table 2** Coefficients of variation, comparison with published methods (%)

	Scan-rescan		Intra-observer	Inter-observer
	S1-S2	S1-S3		
E1 Our method, C2 average CSA	0.42	1.15	<b>0.15</b>	<b>0.14</b>
Horsfield et al. [6], C2 average CSA	–	–	0.59	1.36
Losseff et al. [8], experienced observer	–	<b>0.79</b>	0.73	–
Losseff et al. [8], inexperienced observer	–	1.61	1.03	–
E2 Our method, C1–C3 average CSA	0.33	<b>0.94</b>	<b>0.28</b>	<b>0.36</b>
Horsfield et al. [6], C2–C5 average CSA	–	–	0.44	1.07
Coulon et al. [4], average CSA	–	1.31	0.77	–
Our method, C1–C3 volume	0.33	<b>0.93</b>	<b>0.26</b>	0.35
Coulon et al. [4], volume	–	1.35	1.36	–
E3 Our method, C1–C3 slice-wise CSA	1.10	1.76	0.62	0.70

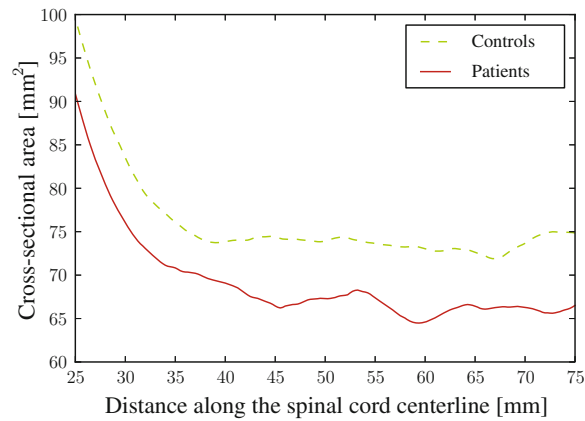
95 % confidence by the same observer.<sup>2</sup> Assuming a CSA of circular shape, this corresponds to a change in radius of 0.03 mm.

**Preliminary Evaluation on Patient Data.** One distortion-corrected scan per subject was used for the cross-sectional clinical data evaluation. The SC surfaces were reconstructed and resliced with an offset of  $\omega = 25$  mm, and the SC volume was measured over the default 50 mm section. One patient was excluded afterwards because the reconstruction showed substantial spike artefacts in the inferior part caused by some refinement line profiles erroneously capturing the edge of the surrounding vertebra. The mean SC volume was  $3,464 \pm 592$  mm<sup>3</sup> for the patient group and  $3,811 \pm 444$  mm<sup>3</sup> for the control group. For both the patient group and the control group, we also calculated the mean slice-wise CSAs. As can be seen in Fig. 3, the patient CSA is smaller than that of the healthy controls throughout all slices. Comparing the mean slice-wise CSAs on each level in a paired-samples t-test showed a statistically significant difference ( $p < 0.01$ ) between the two groups.

**Evaluation on a Large MS Patient Cohort.** In a recent publication [12], we applied our method to a cohort of 172 MS patients. Perpendicularly clipped SC volumes were measured as described above, with an offset of  $\omega = 20$  mm. In a hierarchical multiple linear regression analysis including demographic factors as well as volumetric measures and MS lesion load, the SC volume was shown to be one of the strongest predictors ( $p < 0.001$ ;  $\beta = -0.28$ ) of the Expanded Disability Status Scale (EDSS) score, which signifies the clinically determined degree of disability. In this way, both similar results of other studies were confirmed (such as the ones mentioned in [10]) and the applicability of our method to clinical data was demonstrated.

<sup>2</sup> Note that this is *not* the same value as the  $0.67$  mm<sup>2</sup> reported by Horsfield et al. [6] in a similar argument for the C2–C5 region, as they describe the CV of average CSAs, while we describe an average CV over slice-wise CSAs here. If we do the same calculation for our method with the C1–C3 average CSA (see Table 2, E2), assuming a CSA of  $78$  mm<sup>2</sup> as reported in [6], the detectable change even drops to  $0.43$  mm<sup>2</sup>.





**Fig. 3** Mean cross-sectional area along the spinal cord centerline for the MS patient and control group, distances measured relative to the reformatting anchor point

## 5 Conclusion

We presented a novel semi-automatic method for the reconstruction and reformatting of the spinal cord surface that enables reliable comparisons of both the complete upper cervical cord area and the cord cross-sectional areas over the range of several vertebrae, even if different degrees of spine bending occur between scans. The accuracy of our method was demonstrated by measurements on phantom data. We could also show that a minimum amount of interaction time (two to five minutes) and user-provided input (two sets of labeled voxels and one anatomical landmark) are sufficient to acquire highly reproducible results. These results reach a comparable if not superior precision level with respect to similar approaches [4, 6] (coefficient of variation values  $<1\%$  for mean CSAs and volumes, values in the  $1\text{--}2\%$  range for slice-wise CSAs). Furthermore, we showed that the application of a correction routine to account for geometric distortions induced by gradient non-linearity increases the degree of reproducibility. The latter appears to us to be an important aspect that nevertheless has been neglected (or has at least not been mentioned explicitly) by similar cord reconstruction approaches (cf. [4, 6]), so far.

Finally, the method's applicability to clinical data was demonstrated by a comparison of the slice-wise cross-sectional areas of a group of MS patients with an age-matched control group. We could later confirm these preliminary results by successfully correlating SC volumes with disability scores in a large cohort of MS patients, thus showing the suitability of our approach for everyday use.

**Acknowledgments** This work was supported by the MIAC Corporation, University Hospital Basel, Switzerland.

## References

1. Alyassin, A.M., Lancaster III, J.L., Downs, J.H., Fox, P.T.: Evaluation of new algorithms for the interactive measurement of surface area and volume. *Med. Phys.* **21**(6), 741–752 (1994)
2. Bakshi, R., Dandamudi, V.S.R., Neema, M., De, C., Bermel, R.A.: Measurement of brain and spinal cord atrophy by magnetic resonance imaging as a tool to monitor multiple sclerosis. *J. Neuroimaging* **15**, 30S–45S (2005)
3. Boykov, Y.Y., Jolly, M.P.: Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images. In: Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings, vol. 1, pp. 105–112 (2001)
4. Coulon, O., Hickman, S.J., Parker, G.J., Barker, G.J., Miller, D.H., Arridge, S.R.: Quantification of spinal cord atrophy from magnetic resonance images via a b-spline active surface model. *Mag. Reson. Med.* **47**(6), 1176–1185 (2002)
5. Dierckx, P.: Algorithms for smoothing data with periodic and parametric splines. *Comput. Graph. Image Process.* **20**(2), 171–184 (1982)
6. Horsfield, M.A., Sala, S., Neema, M., Absinta, M., Bakshi, A., Sormani, M.P., Rocca, M.A., Bakshi, R., Filippi, M.: Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: application in multiple sclerosis. *NeuroImage* **50**(2), 446–455 (2010)
7. Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., Kennedy, D., Schmitt, F., Brown, G., MacFall, J., Fischl, B., Dale, A.: Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *NeuroImage* **30**(2), 436–443 (2006)
8. Losseff, N.A., Webb, S.L., O’Riordan, J.I., Page, R., Wang, L., Barker, G.J., Tofts, P.S., McDonald, W.I., Miller, D.H., Thompson, A.J.: Spinal cord atrophy and disability in multiple sclerosis. *Brain* **119**(3), 701–708 (1996)
9. Miller, D.H., Barkhof, F., Frank, J.A., Parker, G.J.M., Thompson, A.J.: Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. *Brain* **125**(8), 1676–1695 (2002)
10. Reynolds, R., Roncaroli, F., Nicholas, R., Radotra, B., Gveric, D., Howell, O.: The neuropathological basis of clinical progression in multiple sclerosis. *Acta Neuropathol.* **122**(2), 155–170 (2011)
11. Tustison, N., Avants, B., Cook, P., Zheng, Y., Egan, A., Yushkevich, P., Gee, J.: N4ITK: improved n3 bias correction. *IEEE Trans. Med. Imaging* **29**(6), 1310–1320 (2010)
12. Weier, K., Pezold, S., Andelova, M., Amann, M., Magon, S., Naegelin, Y., Radue, E.W., Stipich, C., Gass, A., Kappos, L., Cattin, P., Sprenger, T.: Both spinal cord volume and spinal cord lesions impact physical disability in multiple sclerosis. *Multiple Sclerosis J.* **19**(Suppl), 188–189 (2013)

## 6 Automatic Segmentation of the Spinal Cord Using Continuous Max Flow with Cross-sectional Similarity Prior and Tubularity Features

In the publication presented in this chapter, we demonstrate, in terms of a proof of concept, how algorithmic adjustments and the integration of additional features, apart from the image intensities, may increase segmentation robustness. In particular, we adapt the *continuous max flow* algorithm for our goal of spinal cord segmentation by introducing a *similarity prior* along the imaged subject's main body axis, which penalizes cross-sectional area changes of the segmented object and thus helps to guide the segmentation in noisy or low-contrast image regions. Furthermore, we integrate two Hessian-based structural features into the segmentation model's cost functions: a *vesselness* feature, which responds to tube-like structures like the spinal cord, and a *csfness* feature, which responds to both tube- and plate-like structures like the cerebrospinal fluid (CSF).

We compare automatic segmentations of the spinal cord both with and without the cross-sectional similarity prior, as well as with and without the *csfness* feature, showing that both proposed adaptations improve the segmentation results.

**Publication.** The proposed approach was presented at the *2nd MICCAI Workshop and Challenge Computational Methods and Clinical Applications for Spine Imaging (CSI 2014)* in conjunction with the *17th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, September 2014, Boston, Massachusetts, USA. It was published<sup>1</sup> as part of the workshop proceedings [35].

---

<sup>1</sup>[http://dx.doi.org/10.1007/978-3-319-14148-0\\_10](http://dx.doi.org/10.1007/978-3-319-14148-0_10) (last accessed on November 3, 2016)

## Automatic Segmentation of the Spinal Cord Using Continuous Max Flow with Cross-sectional Similarity Prior and Tubularity Features

Simon Pezold, Ketut Fundana, Michael Amann, Michaela Andelova,  
Armanda Pfister, Till Sprenger and Philippe C. Cattin

**Abstract** Segmenting tubular structures from medical image data is a common problem; be it vessels, airways, or nervous tissue like the spinal cord. Many application-specific segmentation techniques have been proposed in the literature, but only few of them are fully automatic and even fewer approaches maintain a convex formulation. In this paper, we show how to integrate a cross-sectional similarity prior into the convex continuous max-flow framework that helps to guide segmentations in image regions suffering from noise or artefacts. Furthermore, we propose a scheme to explicitly include tubularity features in the segmentation process for increased robustness and measurement repeatability. We demonstrate the performance of our approach by automatically segmenting the cervical spinal cord in magnetic resonance images, by reconstructing its surface, and acquiring volume measurements.

---

S. Pezold (✉) · K. Fundana · P.C. Cattin  
Department of Biomedical Engineering, University of Basel, Basel, Switzerland  
e-mail: simon.pezold@unibas.ch

K. Fundana  
e-mail: ketut.fundana@unibas.ch

P.C. Cattin  
e-mail: philippe.cattin@unibas.ch

M. Amann · M. Andelova · A. Pfister · T. Sprenger  
University Hospital Basel, Basel, Switzerland  
e-mail: michael.amann@usb.ch

M. Andelova  
e-mail: michaela.andelova@usb.ch

A. Pfister  
e-mail: armanda.pfister@usb.ch

T. Sprenger  
e-mail: till.sprenger@usb.ch

© Springer International Publishing Switzerland 2015  
J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical  
Applications for Spine Imaging*, Lecture Notes in Computational Vision  
and Biomechanics 20, DOI 10.1007/978-3-319-14148-0\_10

107

## 1 Introduction

The segmentation of oriented tubular structures in the body is a common task in medical applications. Examples include measuring functional vessel volumes in patients of cardiovascular diseases, or quantifying spinal cord atrophy (i.e., the loss of nervous tissue) in a variety of neurodegenerative diseases. Multiple sclerosis (MS) is a prominent example among the latter diseases. Clinical MS studies have shown relationships between the degree of cord atrophy and both the strength of disease [1] and disease duration [2]. Therefore, in recent years, assessing spinal cord atrophy has become a highly active topic of research, resulting in a number of methods that were specifically tailored towards the segmentation of the spinal cord (see e.g. the recently published segmentation approaches of Asman et al. [3], De Leener et al. [4] and the methods referenced therein, or the earlier review of Miller et al. [5]). Only few of these methods, however, make extensive use of the fact that the spinal cord is an inherently tubular structure.

In this paper, we present an automated method that aims at the more general goal of segmenting tubular structures in image volumes. Manual intervention on the target data is reduced to placing a landmark if the segmentation result is ambiguous. As a proof of concept, we successfully demonstrate the practicability of our method by segmenting the spinal cord in magnetic resonance (MR) images and acquiring volume measurements from surface reconstructions of the segmentation results.

We adjust Yuan et al.'s continuous max-flow framework [6] to include a cross-sectional similarity prior. This prior exploits the fact that an oriented elongated structure shows only little change in shape along its orientation. Thus, the prior may guide the segmentation in regions where image information is missing or ambiguous. A related approach of including a similarity prior is pursued by Qiu et al. [7]. Due to their different problem setting (they aim for axial symmetry), they formulate parts of the problem in a discrete setting, while our formulation is continuous. We also propose a way to include tubularity features in the segmentation process. Specifically for the segmentation of the spinal cord, we furthermore introduce the new *csfness* feature, which is designed to improve discrimination between the spinal cord and the cerebrospinal fluid (CSF) that immediately surrounds it.

## 2 Method

In the following subsections, we introduce our adaptation of the max-flow approach and define the flow capacity functions together with the features that we use in experiments. We present an algorithm to solve the adapted problem, and we conclude the section by proposing a scheme to reconstruct the surface from the segmentation result, which we use for quantitative measurements.

**Notation.** Let  $I : \Omega \rightarrow \mathcal{I}$  denote the intensity non uniformity corrected image [8] with intensities in the normalized intensity space  $\mathcal{I} = [0, 1]$ , where

$x = (x_1, x_2, x_3)^T \in \Omega$  are the coordinates in the continuous image domain  $\Omega \subset \mathbb{R}^3$ . Throughout the whole section, we furthermore assume that the tubular structure of interest is roughly oriented parallel to the  $x_3$  axis. Figuring out the orientation should be straightforward for most clinical applications, as the subject's orientation with respect to the image can be determined from the image's meta data for most clinical imaging modalities.

**Original max-flow formulation.** A general formulation for the continuous max-flow problem with spatial flow  $p(x)$ , source flow  $p_s(x)$ , sink flow  $p_t(x)$ , and corresponding flow capacities  $C(x)$ ,  $C_s(x)$ ,  $C_t(x)$  is stated by Yuan et al. [6] as

$$\max_{p_s, p_t, p} \int_{\Omega} p_s dx, \quad (1)$$

subject to the flow capacity constraints

$$p_s(x) \leq C_s(x), \quad p_t(x) \leq C_t(x), \quad \|p(x)\| \leq C(x) \quad (2)$$

and the flow conservation constraint

$$\operatorname{div} p(x) - p_s(x) + p_t(x) = 0. \quad (3)$$

## 2.1 Cross-Sectional Similarity Prior

Following our goal to impose a cross-sectional similarity prior on the segmentation, we split the spatial flow  $p(x)$  into an in-slice component  $q : \Omega \rightarrow \mathbb{R}^2$  and a through-slice component  $r : \Omega \rightarrow \mathbb{R}$  with respect to slices that lie perpendicular to the  $x_3$  axis (see Fig. 1a). The resulting continuous max-flow problem can then be written as follows:

$$\max_{p_s, p_t, q, r} \int_{\Omega} p_s dx, \quad (4)$$

subject to the new flow capacity constraints

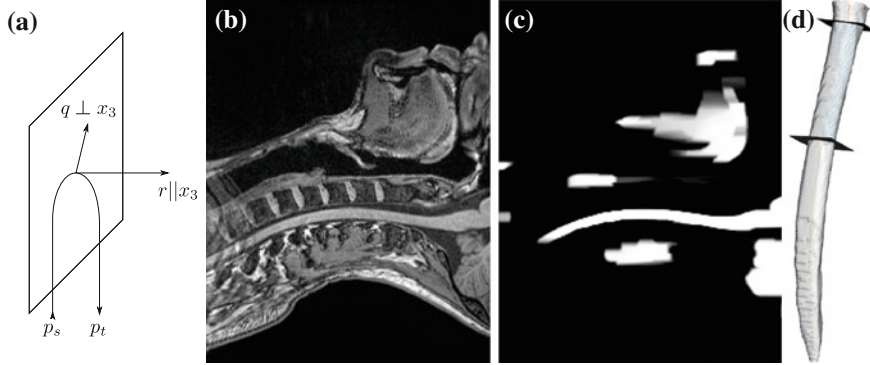
$$p_s(x) \leq C_s(x), \quad p_t(x) \leq C_t(x), \quad \|q(x)\| \leq \alpha(x), \quad |r(x)| \leq \beta(x) \quad (5)$$

and the new flow conservation constraint

$$\operatorname{div}_{12} q(x) + r'(x) - p_s(x) + p_t(x) = 0, \quad (6)$$

where  $\operatorname{div}_{12} q$  denotes the divergence of  $q$  perpendicular to the  $x_3$  axis and  $r'$  denotes the derivative of  $r$  along the  $x_3$  axis.

The flow formulation now possesses the desired property of having the spatial flow capacity  $C(x)$  of [6] represented by two separate terms, namely the in-slice flow capacity  $\alpha(x)$  and the through-slice flow capacity  $\beta(x)$ . The latter capacity,



**Fig. 1** Method overview. **a** Proposed flow configuration: the spatial flow is split into an in-slice component  $q$ , perpendicular to the axis along which the tubular structure is oriented, and a through-slice component  $r$ , parallel to the axis. **b** Sample sagittal slice of one of the images used for evaluation. **c** Segmentation result. **d** Surface reconstruction with cutting planes for volume measurement

$\beta(x)$ , represents the cross-sectional similarity prior that allows for precise control over the through-slice flow behavior: For example, we may choose an edge-based cost function for  $\alpha(x)$  that drives the segmentation towards edges in  $I$ , while setting  $\beta(x) = \beta_0$  to enforce constant similarity throughout all slices. Or we may calculate  $\beta(x) = \beta(x_3)$  as a slice wise cost-function that, for each slice, adjusts the similarity prior to the in-slice noise level (reinforcing the similarity prior if the noise level is high and relaxing it if the noise level is low). Other combinations are possible, of course: note that both  $\alpha$  and  $\beta$  may be formulated pointwise.

**Dual formulation.** Introducing the Lagrange multiplier  $u = u(x)$  and following the steps in [6], the max-flow problem can be reformulated as the equivalent primal-dual model

$$\max_{p_s, p_t, q, r} \min_u \int_{\Omega} p_s dx + \int_{\Omega} u \cdot (\text{div}_{12} q + r' - p_s + p_t) dx \quad (7)$$

subject to the capacity constraints (5). The equivalent dual model representing a relaxed min-cut problem then becomes

$$\min_{u \in [0,1]} E(u) := \int_{\Omega} \{(1-u)C_s + uC_t + \alpha |\nabla_{12} u| + \beta |u'|\} dx. \quad (8)$$

Here,  $\nabla_{12} u$  denotes the in-slice gradient and  $u'$  denotes the through-slice derivative of  $u$  with respect to the  $x_3$  axis, similar to the definitions of  $\text{div}_{12} q$  and  $r'$  above. It can be shown that each level set function  $u^\ell(x)$ ,  $\ell \in (0, 1]$  given by

$$u^\ell(x) := \begin{cases} 1, & u^*(x) > \ell \\ 0, & u^*(x) \leq \ell \end{cases} \quad \text{with } u^* := \underset{u}{\text{argmin}} E(u) \quad (9)$$

is a global binary solution of the adapted problem stated in Eq. (4).

## 2.2 Tubularity Features

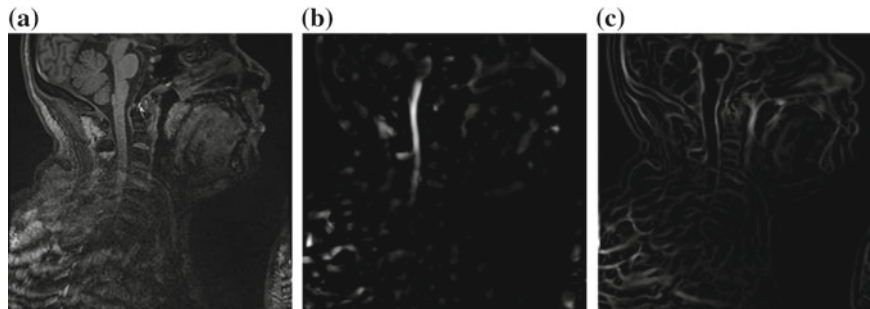
As our goal is to segment tubular structures in the image, it appears natural to include tubularity features in the flow capacity calculations. A well-known tubularity feature is Frangi's measure of *vesselness* [9] (see Fig. 2b),  $v^*(x) = \max_{\xi \in S_v} v(x; \xi)$ , where, for each scale  $\xi$  in the predefined set of scales  $S_v$ , the vesselness  $v(x; \xi)$  of bright tubular structures on dark background is

$$v(x; \xi) = \begin{cases} 0, & \lambda_2 \geq 0 \vee \lambda_3 \geq 0 \\ \left(1 - \exp\left(-2\frac{\lambda_2^2}{\lambda_3^2}\right)\right) \exp\left(-2\frac{\lambda_1^2}{\lambda_2\lambda_3}\right) \left(1 - \exp\left(-\frac{\sum_{i=1}^3 \lambda_i^2}{2h^2}\right)\right) & \text{else,} \end{cases} \quad (10)$$

with  $\lambda_i = \lambda_i(x)$  denoting the ordered eigenvalues ( $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$ ) of the point-wise Hessian matrices that result from convolving the input image  $I$  with Gaussian derivatives of standard deviation  $\xi$ . We define  $h$  as half of the maximum Hessian norm at the current scale as suggested by Frangi [9].

In our experiments on segmenting the spinal cord, we decided to include another feature that specifically describes the background that immediately surrounds the target structure. The spinal cord is embedded in cerebrospinal fluid (CSF), which appears dark in the used MR sequences. As the CSF also appears largely elongated, but exhibits both tube-like and plate-like properties, we adapt Frangi's vesselness feature to a *csfness* feature  $w^*(x)$  (see Fig. 2c) that discriminates between blob-like structures and non-blobs. We do so by replacing the eigenvalue ratio terms of  $v^*$  with an equivalent term composed of  $\lambda_1$  and  $\lambda_3$ , as it is the latter ratio that discriminates both vessels and plates from blobs in Hessian eigenvalue analysis [9]. Consequently, we define  $w^*(x) = \max_{\xi \in S_w} w(x; \xi)$  for dark non-blobs on bright background in the scales  $S_w$  with

$$w(x; \xi) = \begin{cases} 0, & \lambda_3 \leq 0 \\ \exp\left(-2\frac{\lambda_1^2}{\lambda_3^2}\right) \left(1 - \exp\left(-\frac{\sum_{i=1}^3 \lambda_i^2}{2h^2}\right)\right) & \text{else.} \end{cases} \quad (11)$$



**Fig. 2** Features used in segmentation. **a** Image intensities. **b** Vesselness response. **c** Csfness response



**Combining the features.** Let  $\mathcal{V} = [0, 1] \ni v^*$ ,  $\mathcal{W} = [0, 1] \ni w^*$  be the vesselness and csfness feature spaces, let  $\mathcal{Y} = \mathcal{I} \times \mathcal{V} \subset \mathbb{R}^2$  and  $\mathcal{Z} = \mathcal{I} \times \mathcal{V} \times \mathcal{W} \subset \mathbb{R}^3$  be two combined feature spaces, let  $I_2 : \Omega \rightarrow \mathcal{Y}$ ,  $I_3 : \Omega \rightarrow \mathcal{Z}$  be two new image functions that map to the combined feature spaces, and let  $y \in \mathcal{Y}$ ,  $z \in \mathcal{Z}$  be the coordinates in the combined feature spaces.

Furthermore, let  $Y_f = \{y_f^i\}_{i=1}^M$ ,  $Y_b = \{y_b^j\}_{j=1}^N$  be two sets holding samples of  $\mathcal{Y}$  with known foreground and background membership, respectively. Based on these training sets, we propose to calculate the capacities for the terminal flow constraints (5) using kernel density estimates:

$$C_s(y) = C_s(I_2(x)) = \frac{\frac{1}{M} \sum_{i=1}^M K_{\Sigma_f}(y - y_f^i)}{\frac{1}{M} \sum_{i=1}^M K_{\Sigma_f}(y - y_f^i) + \frac{1}{N} \sum_{j=1}^N K_{\Sigma_b}(y - y_b^j)}, \quad (12)$$

$$C_t(y) = C_t(I_2(x)) = 1 - C_s(y), \quad (13)$$

where  $K_{\Sigma}$  is a Gaussian kernel with zero mean and diagonal covariance matrix  $\Sigma$ , holding variances  $\sigma_d^2$  for the feature dimensions  $d$  as diagonal elements. Terminal capacities for the feature space  $\mathcal{Z}$  may be calculated in a similar way. For the sake of simplicity, we choose the non-terminal capacities as constants in our experiments:  $\alpha(x) = \alpha_0$ ,  $\beta(x) = \beta_0$ .

### 2.3 Algorithm

In accordance with the original max-flow approach, we propose to find a global solution to our adapted formulation by setting up the respective augmented Lagrangian equation as

$$L_c(p_s, p_t, q, r, u) := \int_{\Omega} p_s dx + \int_{\Omega} u \cdot (\text{div}_{12} q + r' - p_s + p_t) dx - \frac{c}{2} \|\text{div}_{12} q + r' - p_s + p_t\|^2, \quad (14)$$

and iteratively optimizing it using Algorithm 1, based on the algorithm in [6].

### 2.4 Surface Reconstruction

As can be concluded from Eq. (9), reconstructing the surface of the segmented structure amounts to finding the isosurface of level  $\ell \in (0, 1]$  in the segmentation result  $u^*$  (see Fig. 1c, d). We propose to extract the isoline as a polygon of  $m$  vertices for each slice along the  $x_3$  axis and successively connect the resulting dots in space.

---

**Algorithm 1** Augmented Lagrangian based max-flow algorithm.

---

Arbitrarily initialize  $p_s^1, p_t^1, q^1, r^1, u^1$ ; initialize  $C_s, C_t, \alpha, \beta$ ; specify a tolerance  $\hat{\varepsilon}$  and a step size  $c$ ; set the iteration count  $k = 1$ . For each iteration, perform the following:

- Optimize  $q$ , fixing the other variables:

$$q^{k+1} \leftarrow \underset{\|q\| \leq \alpha}{\operatorname{argmax}} L_c = \underset{\|q\| \leq \alpha}{\operatorname{argmax}} -\frac{c}{2} \left\| \operatorname{div}_{12} q + r'^k - p_s^k + p_t^k - \frac{u^k}{c} \right\|^2,$$

using a projected gradient ascent step of step size  $\gamma_q$ , as suggested in [10]:

- update  $q$ :  $q^{k+1} \leftarrow q^k + \gamma_q \cdot \nabla_{12}(\operatorname{div}_{12} q^k + r'^k - p_s^k + p_t^k - \frac{u^k}{c})$ ,
- project  $q$ :  $q^{k+1} \leftarrow \begin{cases} \frac{q^{k+1}}{\|q^{k+1}\|} \cdot \min\{\|q^{k+1}\|, \alpha\}, & q^{k+1} \neq 0 \\ 0, & q^{k+1} = 0 \end{cases}$ .

- Optimize  $r$ , fixing the other variables:

$$r^{k+1} \leftarrow \underset{|r| \leq \beta}{\operatorname{argmax}} L_c = \underset{|r| \leq \beta}{\operatorname{argmax}} -\frac{c}{2} \left\| \operatorname{div}_{12} q^{k+1} + r' - p_s^k + p_t^k - \frac{u^k}{c} \right\|^2,$$

using a projected gradient ascent step of step size  $\gamma_r$ :

- update  $r$ :  $r^{k+1} \leftarrow r^k + \gamma_r \cdot \frac{\partial}{\partial x_3}(\operatorname{div}_{12} q^{k+1} + r'^k - p_s^k + p_t^k - \frac{u^k}{c})$ ,
- project  $r$ :  $r^{k+1} \leftarrow \operatorname{sgn}(r^{k+1}) \cdot \min\{|r^{k+1}|, \beta\}$ .

- Optimize  $p_s$  and  $p_t$  pointwise:

- i)  $p_s^{k+1} \leftarrow \frac{1}{c} - \frac{u^k}{c} + \operatorname{div}_{12} q^{k+1} + r'^{k+1} + p_t^k$ , ii)  $p_s^{k+1} \leftarrow \min\{p_s^{k+1}, C_s\}$ ,
- iii)  $p_t^{k+1} \leftarrow \frac{u^k}{c} - \operatorname{div}_{12} q^{k+1} - r'^{k+1} + p_s^{k+1}$ , iv)  $p_t^{k+1} \leftarrow \min\{p_t^{k+1}, C_t\}$ .

- Calculate the pointwise error  $\varepsilon$ :  $\varepsilon^{k+1} \leftarrow c \cdot (\operatorname{div}_{12} q^{k+1} + r'^{k+1} - p_s^{k+1} + p_t^{k+1})$ .
  - Update  $u$ :  $u^{k+1} \leftarrow u^k - \varepsilon^{k+1}$ .
  - Terminate if  $\frac{1}{|\Omega|} \int_{\Omega} |\varepsilon^{k+1}(x)| dx < \hat{\varepsilon}$ , otherwise update  $k \leftarrow k + 1$  and continue.
- 

This provides us with the slicewise contours of the segmentation at no additional cost, which then facilitates estimating the centerline, namely as a curve fit through the centroids of the contours. A centerline estimate, in turn, may be useful to acquire quantitative measurements from the reconstruction (see Sects. 3, 4).

If there are multiple foreground regions in  $u^*$ , a point of reference may be used to choose the region closest to it. Likewise, heuristic criteria like sudden jumps of the centroid or a threshold on the contour line's convexity may be used to determine a cutoff for the tubular structure of interest. In the spinal cord segmentation experiments below, we define a point of reference by an anatomical landmark, and we define two cutoff criteria as finding either a distance  $> d$  between the centroids of two consecutive slices or finding a contour line with convexity  $< t$ . As a measure of convexity, we employ the ratio of the contour line's area and the area of its convex hull.

### 3 Materials

Applicability of our approach is shown by segmenting the spinal cord in MR images of healthy volunteers (Figs. 1b, 2a) and MS patients.

To assess accuracy and reproducibility, 11 healthy volunteers (3 female, 8 male, mean age 32.7 year, range 26–44 year) were scanned on a 3 T whole-body MR scanner (Verio, Siemens Medical, Germany) with a T1-weighted MPRAGE sequence (TR/TI/TE/ $\alpha = 2.0$  s/1.0 s/3.4 ms/8°); 192 slices in sagittal orientation parallel to the interhemispheric fissure were acquired with an isotropic resolution of 1 mm<sup>3</sup>. Image volumes were corrected for gradient nonlinearity distortions using the scanner manufacturer’s correction routine.

To show applicability to clinical data, we used follow-up data of 32 MS patients (21 female, 11 male, mean age 47.1 year, range 22–60 year; 22 patients with relapsing-remitting MS, 10 patients with primary progressive MS, mean disease duration 13.8 year, range 3–31 year, median EDSS 3.0, range 1.5–6.0). The patients were scanned on a 1.5 T whole-body MR scanner (Avanto, Siemens Medical, Germany) with a T1-weighted MPRAGE sequence (TR/TI/TE/ $\alpha = 2.08$  s/1.1 s/3.93 ms/15°); 160 slices in sagittal orientation parallel to the interhemispheric fissure were acquired with an in-slice resolution of 0.98 mm  $\times$  0.98 mm and a slice thickness of 1 mm. Scans were acquired at two points in time approximately 5 years apart (mean 5.04 year, range 4.55–5.41 year); demographic data above is given with respect to the earlier scan. Distortion correction was applied to the surface reconstructions using the method of Janke et al. [11].

For the calculation of the terminal flow constraints  $C.(y)$  and  $C.(z)$ , sample sets were acquired on 150 separate scans of MS patients. The training patients were scanned with the same MPRAGE sequence as the 32 MS patients above. Foreground/background membership of the training samples was determined using a graph cuts-based [12] semi automated method described as *presegmentation* in [13]. To speed up calculations, features were discretized to 50 bins in the  $[0, 1]$  interval in each feature dimension. Silverman’s rule of thumb with  $\sigma_d = 4 \frac{1}{d+4} (n(D+2))^{-\frac{1}{d+4}} \hat{\sigma}_d$  provided a  $\sigma_d^2$  estimate, where  $n$  is the number of samples,  $\hat{\sigma}_d$  is the sample standard deviation in  $d$ , and  $D$  is 2 for  $C.(y)$  and 3 for  $C.(z)$ . To avoid zero bins, a small additive constant of  $0.0001k$  was added to the resulting bin values, where  $k$  is the maximum value of all bins.

For all experiments, the following parameters were applied:  $\alpha_0 = 0.5$ ,  $\beta_0 = 2.5$ ,  $S_v = [2 \text{ mm}, 4 \text{ mm}]$  (16 values),  $S_w = [1 \text{ mm}, 2 \text{ mm}]$  (8 values) for the flow capacities;  $\ell = 0.5$ ,  $m = 60$ ,  $d = 10 \text{ mm}$ ,  $t = 0.95$  for the surface reconstruction. Likewise for all reported volume measurements, the volume of a spinal cord surface segment of 50 mm centerline length, which was clipped by planes perpendicular to the centerline and which was located approximately 25 mm inferior of a manually marked landmark, was evaluated as described in [13].

## 4 Results

**Scan–rescan evaluation.** In two experiments on scan–rescan data, we evaluated the accuracy and reproducibility of our method. To show the benefits of splitting the spatial flow into an in-slice component and a through-slice component, we repeated all experiments using Yuan et al.’s original formulation [6] for segmentation, setting the spatial flow capacity  $C(x)$  (see Eq. (2)) to a value of  $C(x) = 1.0$ , as this parameter choice provided the highest number of successful surface reconstructions in the second experiment below. For the experiments, we scanned the 11 healthy volunteers (see Sect. 3) three times in a row (scans S1, S2, S3), without repositioning between S1 and S2 and with repositioning between S2 and S3, resulting in 33 scans altogether.

*Accuracy:* As we work with human *in-vivo* data, it was not possible to acquire quantitative ground truth measurements, for example, via histologic specimen. We therefore used manual segmentations of the image data as a gold standard for comparison in the first experiment. To make such manual measurements feasible, a semiautomated approach that allows for human feedback in the segmentation process seemed appropriate. We thus segmented all scan–rescan datasets with the method described as *presegmentation* in [13], placing foreground/background seeds manually and adjusting them in an iterative manner until we acquired a satisfying binary segmentation. We then compared the overlap of this gold standard segmentation with the binarized results of the automated segmentation for a 50 mm cord segment, located 25 mm inferior of the manually marked landmark. As a measure of overlap agreement, we calculated Dice coefficients for the region overlaps.

With our approach, we gained a mean Dice coefficient of 0.88 using the proposed feature combination  $\mathcal{Z}$  (i.e., intensity + vesselness + csfness) and 0.82 using feature combination  $\mathcal{Y}$  (i.e., intensity + vesselness). With Yuan et al.’s approach, we gained a mean Dice coefficient of 0.86 using  $\mathcal{Z}$  and 0.79 using  $\mathcal{Y}$ . Therefore, our approach proves superior in the given problem setting. Furthermore, it can be seen that including the csfness feature into the segmentation process improves the segmentation accuracy.

*Reproducibility:* In the second experiment, we assessed the reproducibility of our method. The cervical spinal cord was segmented using feature combinations  $\mathcal{Y}$  and  $\mathcal{Z}$ , its surface was reconstructed, and the volume of the 50 mm cord surface segment (see Sect. 3) was compared between scans and rescans. As a measure of reproducibility, we calculated the coefficients of variation (CV; i.e., the sample standard deviation over the mean) of the measured volumes for all possible S1–S2 comparisons (i.e., without repositioning) and S1–S3 comparisons (i.e., with repositioning). An overview of the mean CVs is given in Table 1.

For our proposed segmentation approach, the subsequent reconstruction of the complete surface segment succeeded for 30 out of 33 scans using  $\mathcal{Y}$  and 32 out of 33 scans using  $\mathcal{Z}$ . All failures happened for the same subject, whose scans showed an extremely low signal-to-noise ratio upon visual inspection. For Yuan et al.’s segmentation approach, the surface reconstruction succeeded for 25 out of 33 scans using  $\mathcal{Y}$  and 28 out of 33 scans using  $\mathcal{Z}$ .

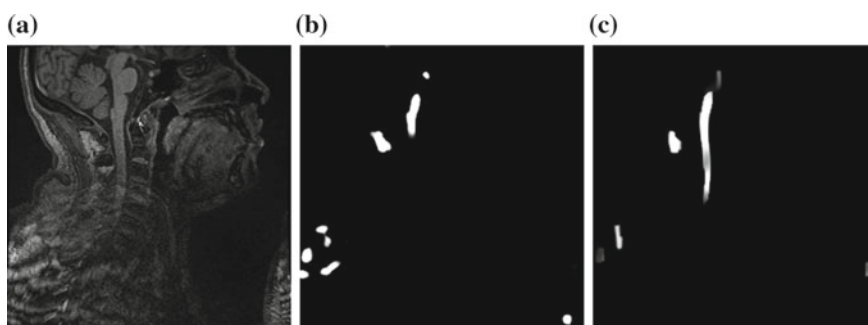
**Table 1** Coefficients of variation (%) for spinal cord segment volume measurements, using feature combinations  $\mathcal{Y}$  and  $\mathcal{Z}$  with Yuan et al.'s and our segmentation method

	$\mathcal{Y}$		$\mathcal{Z}$	
	S1 versus S2	S1 versus S3	S1 versus S2	S1 versus S3
Ours	<b>2.02</b>	<b>5.81</b>	<b>2.13</b>	4.90
Yuan et al.	2.54	6.84	2.85	<b>4.42</b>

As one could expect, CVs are lower for the S1–S2 comparison, due to the fact that the subjects were not repositioned. Furthermore, including the csfness feature makes the segmentation more robust (more successful surface reconstructions) while at the same time having beneficial effects on the reproducibility (substantially lower CVs for the more realistic S1–S3 comparisons). Similar statements on improved robustness and reproducibility can be made when comparing our adapted max-flow formulation with the original formulation: in both aspects, our method proves largely superior. And while for feature combination  $\mathcal{Z}$  the S1–S3 CV of the original approach is better than ours, one has to keep in mind that ours is calculated on a higher number of successful reconstructions, including the more challenging ones on which the original approach failed.

An exemplary case where the surface reconstruction failed for the original max-flow formulation while succeeding for our adapted formulation is shown in Fig. 3. As can be seen, the segmentation stops early for the original formulation while it extends further down into the noisy image regions for ours. Relaxing  $C(x)$  in this case would possibly enable the original formulation to also extend further down; however, this would come at the price of an overall higher susceptibility to noise. By contrast, controlling  $\alpha(x)$  and  $\beta(x)$  separately in our approach enables us to largely circumvent this tradeoff.

On the whole, the CVs we obtained by our method are higher than those of established methods that are actually used in MS research (most notably, Losseff et al. [1] and the methods compared in [13]). On the one hand, however, one should keep

**Fig. 3** Comparison of the segmentation approaches on noisy, low-contrast case. **a** Input image. **b** Segmentation result using the original max-flow formulation [6]. **c** Segmentation result using our max-flow formulation

in mind the substantially higher amount of manual intervention in these approaches. On the other hand, we see our presented framework in its current state more as a proof of concept than as a tool that is ready for clinical use.

**Evaluation on patient data.** A comparison of the five-year follow-up MS patient data (see Sect. 3), using the proposed max-flow formulation for segmentation, showed a mean yearly atrophy of  $25.4 \text{ mm}^3$  in the 50 mm cord surface segment (maximum loss:  $194.3 \text{ mm}^3$ , maximum gain:  $53.4 \text{ mm}^3$ ). The mean yearly percentage loss was 0.9 % (maximum loss: 7.0 %, maximum gain: 2.0 %). These measurements agree well with the observation of cord atrophy during MS progression reported in the literature [2]. Nevertheless, due to the high variability, our measurements should again be interpreted as a proof of concept for our segmentation method rather than as hard clinical data.

**Computational performance.** As we implemented the max-flow segmentation on the GPU based on code provided by the authors of [6, 10], results can be acquired extremely fast, namely in the order of seconds. Other parts of the implementation also show a high parallelization potential in that they are mainly pointwise (such as the feature calculation and the surface extraction). We therefore assume that the complete chain of steps from feature calculation to quantitative measurements could be optimized to run in less than a minute per subject.

## 5 Conclusion

We presented a new segmentation algorithm based on continuous max flow that was specifically tailored towards the segmentation of elongated structures: a cross-sectional similarity prior was introduced, which guides the segmentation in regions of missing or contradictory image information. We showed how tubularity features may be used in the flow capacity constraints to increase segmentation robustness and measurement repeatability. Finally, we successfully demonstrated the clinical applicability of our method by segmenting the spinal cord in both healthy volunteers and multiple sclerosis patients.

**Acknowledgments** We would like to thank Ernst-Wilhelm Radue and the MIAC AG, University Hospital Basel, Basel, Switzerland, for their support.

## References

1. Losseff, N.A., Webb, S.L., O’Riordan, J.I., Page, R., Wang, L., Barker, G.J., Tofts, P.S., McDonald, W.I., Miller, D.H., Thompson, A.J.: Spinal cord atrophy and disability in multiple sclerosis. *Brain* **119**(3), 701–708 (1996)
2. Rashid, W., Davies, G.R., Chard, D.T., Griffin, C.M., Altmann, D.R., Gordon, R., Thompson, A.J., Miller, D.H.: Increasing cord atrophy in early relapsing-remitting multiple sclerosis: a 3 year study. *J. Neurol., Neurosurg. Psychiatry* **77**(1), 51–55 (2006)

3. Asman, A., Smith, S., Reich, D., Landman, B.: Robust GM/WM segmentation of the spinal cord with iterative non-local statistical fusion. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*. Lecture Notes in Computer Science, vol. 8149, pp. 759–767. Springer, Heidelberg (2013)
4. De Leener, B., Kadoury, S., Cohen-Adad, J.: Robust, accurate and fast automatic segmentation of the spinal cord. *NeuroImage* **98**, 528–536 (2014)
5. Miller, D.H., Barkhof, F., Frank, J.A., Parker, G.J.M., Thompson, A.J.: Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. *Brain* **125**(8), 1676–1695 (2002)
6. Yuan, J., Bae, E., Tai, X.C.: A study on continuous max-flow and min-cut approaches. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2217–2224 (2010)
7. Qiu, W., Yuan, J., Ukwatta, E., Sun, Y., Rajchl, M., Fenster, A.: Fast globally optimal segmentation of 3d prostate mri with axial symmetry prior. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*. Lecture Notes in Computer Science, vol. 8150, pp. 198–205. Springer, Heidelberg (2013)
8. Tustison, N., Avants, B., Cook, P., Zheng, Y., Egan, A., Yushkevich, P., Gee, J.: N4ITK: improved n3 bias correction. *IEEE TMI* **29**(6), 1310–1320 (2010)
9. Frangi, A., Niessen, W., Vincken, K., Viergever, M.: Multiscale vessel enhancement filtering. In: Wells, W., Colchester, A., Delp, S. (eds.) *Medical Image Computing and Computer-Assisted Intervention—MICCAI'98*. Lecture Notes in Computer Science, vol. 1496, pp. 130–137. Springer, Heidelberg (1998)
10. Yuan, J., Bae, E., Tai, X.C., Boykov, Y.: A study on continuous max-flow and min-cut approaches. Tech. Rep. CAM 10–61, UCLA, CAM, UCLA (2010)
11. Janke, A., Zhao, H., Cowin, G.J., Galloway, G.J., Doddrell, D.M.: Use of spherical harmonic deconvolution methods to compensate for nonlinear gradient effects on MRI images. *Magn. Reson. Med.* **52**(1), 115–122 (2004)
12. Boykov, Y.Y., Jolly, M.P.: Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In: *Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings*, vol. 1, pp. 105–112 (2001)
13. Pezold, S., Amann, M., Weier, K., Fundana, K., Radue, E., Sprenger, T., Cattin, P.: A semi-automatic method for the quantification of spinal cord atrophy. In: Yao, J., Klinder, T., Li, S. (eds.) *Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics, vol. 17, pp. 143–155. Springer International Publishing (2014)





## 7 Reliable Volumetry of the Cervical Spinal Cord in MS Patient Follow-up Data with Cord Image Analyzer (*Cordial*)

In the publication presented in this chapter, we integrate the continuous max flow-based segmentation approach of Chapter 6 into the toolchain of Chapter 5. As a result of replacing graph cut with continuous max flow, the segmentation becomes faster and more memory-efficient, and thus allows for the handling of larger images. As a result of integrating the structural features, an image preprocessing step can be omitted and less user input is required in the interactive presegmentation. Additionally, the handling of more MR sequences becomes possible, with the previous toolchain only working on T1-weighted images and the proposed version also working on T2-weighted images, as demonstrated in the experiments.

Like the previous version, the adapted toolchain is evaluated both on healthy subjects and multiple sclerosis patients. This time, we use longitudinal patient data, that is, MR images of the same patients acquired at different points in time. Again, measurement reproducibility in the healthy subjects is high. Spinal cord atrophy rates in patient subgroups of different forms of MS largely agree with previously published findings.

**Publication.** The publication was written in joint first authorship by Michael Amann and Simon Pezold, with S. Pezold developing and describing the proposed toolchain and M. Amann focusing on experimental design and evaluation. It was accepted by the *Journal of Neurology* (J Neurol) in April 2016 and published<sup>1</sup> as part of volume 263, issue 7 of the journal [1].

---

<sup>1</sup><http://dx.doi.org/10.1007/s00415-016-8133-0> (last accessed on November 3, 2016)



## Reliable volumetry of the cervical spinal cord in MS patient follow-up data with cord image analyzer (*Cordial*)

Michael Amann<sup>1,2,3</sup> · Simon Pezold<sup>4</sup> · Yvonne Naegelin<sup>2</sup> · Ketut Fundana<sup>4</sup> ·  
Michaela Andělová<sup>2</sup> · Katrin Weier<sup>2</sup> · Christoph Stippich<sup>3</sup> · Ludwig Kappos<sup>2</sup> ·  
Ernst-Wilhelm Radue<sup>1</sup> · Philippe Cattin<sup>4</sup> · Till Sprenger<sup>2,5</sup>

Received: 8 February 2016/Revised: 1 April 2016/Accepted: 13 April 2016/Published online: 9 May 2016  
© Springer-Verlag Berlin Heidelberg 2016

**Abstract** Spinal cord (SC) atrophy is an important contributor to the development of disability in many neurological disorders including multiple sclerosis (MS). To assess the spinal cord atrophy in clinical trials and clinical practice, largely automated methods are needed due to the sheer amount of data. Moreover, using these methods in longitudinal trials requires them to deliver highly reliable measurements, enabling comparisons of multiple data sets of the same subject over time. We present a method for SC volumetry using 3D MRI data providing volume measurements for SC sections of fixed length and location. The segmentation combines a continuous max flow approach with SC surface reconstruction that locates the SC boundary based on image voxel intensities. Two cutting planes perpendicular to the SC centerline are determined based on predefined distances to an anatomical landmark, and the cervical SC volume (CSCV) is then calculated in-between these boundaries. The development of the method focused on its application in MRI follow-up

studies; the method provides a high scan–rescan reliability, which was tested on healthy subject data. Scan–rescan reliability coefficients of variation (COV) were below 1 %, intra- and interrater COV were even lower (0.1–0.2 %). To show the applicability in longitudinal trials, 3-year follow-up data of 48 patients with a progressive course of MS were assessed. In this cohort, CSCV loss was the only significant predictor of disability progression ( $p = 0.02$ ). We are, therefore, confident that our method provides a reliable tool for SC volumetry in longitudinal clinical trials.

**Keywords** Spinal cord · Segmentation · Continuous max flow · Volumetry · Reliability · Multiple sclerosis · Atrophy · Progression

### Introduction

MR imaging of the spinal cord (SC) is a valuable part of the diagnostic work-up in patients with multiple sclerosis (MS) and other neurological disorders. Currently, mainly, signal intensity changes within the cord are considered in the clinical management of disorders of the central nervous system. However, SC atrophy has been suggested as an additional important contributor to the accrual of physical disability, and the measurements of SC volume over time may deliver valuable information on disease progression and treatment effects.

Although the benefit of MR imaging of the SC has been demonstrated for decades [8, 19, 21], the segmentation of the SC in MR images is still demanding: the SC is a long, thin anatomical structure, making MR images of the SC prone to partial volume effects. Moreover, the SC is relatively mobile in the spinal canal depending on the subject's

M. Amann and S. Pezold contributed equally.

✉ Michael Amann  
Michael.Amann@usb.ch

<sup>1</sup> Medical Image Analysis Center (MIAC) AG, Mittlere Strasse 83, 4054 Basel, Switzerland

<sup>2</sup> Department of Neurology, University Hospital Basel, Petersgraben 4, 4031 Basel, Switzerland

<sup>3</sup> Division of Diagnostic and Interventional Neuroradiology, Department of Radiology, University Hospital Basel, Petersgraben 4, 4031 Basel, Switzerland

<sup>4</sup> Department of Biomedical Engineering, University of Basel, Gewerbestrasse 14, 4123 Allschwil, Switzerland

<sup>5</sup> Department of Neurology, DKD Helios Klinik Wiesbaden, Aukammallee 33, 65191 Wiesbaden, Germany

posture, which complicates the reproduction of localized measurements.

Manual SC segmentation is very time-consuming and biased by operator-dependent factors; therefore, semi- or fully automatic volumetric methods have been proposed, with pioneering work reaching back as far as 1996 [16]. Since then, a wide range of different segmentation approaches have been employed, such as active contours or surfaces [4, 10, 11, 20], level sets [30], partial volume modeling [2], gradient vector flow [14], atlas-based approaches [3, 6], and tubular deformable models [5]. Likewise, the amount of required user interaction has varied to a great degree, reaching from the manual identification of the SC centerline [11] to the identification of multiple [10] or single [4, 6] anatomical landmarks, with completely automated approaches presented only recently [31]. While many of these techniques have been shown to provide precise and reproducible measurements of either CS volume or cross-sectional area (CSA) in healthy subjects or cross-sectional trial settings, only few methods have been validated and/or evaluated on patient follow-up data demonstrating the applicability in longitudinal trial settings until recently. Notable exceptions are the method of Losseff et al. [16], which was evaluated on 1-year follow-up data of 28 MS patients [28], the active surface approach [11], which was evaluated on 2.3-year follow-up data of 35 MS patients by Valsasina and colleagues [29], and the work of Yiannakas et al. [31], who evaluated their method on 1-year follow-up data of 30 MS patients.

In this work, we present an extended evaluation of *cordial*, the “cord image analyzer”<sup>1</sup>. The method has been developed with longitudinal studies in mind. Measurements are provided for SC sections of fixed length and location and, therefore, allow for a reliable measurement of localized SC volume and its comparison in the same subject over time. The SC can slide within the spinal canal [9, 25] and *cordial*, therefore, relies on natural landmarks on the SC itself, such as the medullopontine sulcus, allowing reproducible measurements of comparable spinal cord segments. Manual input to our method is limited to region of interest selection as well as the placement of the landmark and two sets of SC labels and background labels; apart from that, the workflow is completely automated.

We here tested the reliability of *cordial* in a scan–rescan scenario with healthy volunteers, evaluating two different 3D MR contrasts: a T1-weighted (T1w) MPRAGE and a T2-weighted (T2w) 3D turbo spin echo (TSE) sequence. To demonstrate the applicability on longitudinal clinical data, we processed 3-year follow-up MPRAGE images of 48 MS patients, assessing the relation between cervical SC atrophy and disability.

## Methods

### Data analysis

The method that we use for volumetry in *cordial* is largely a combination of methodological developments previously described in [23] and [24]. The approach can be divided into three distinct steps, which we refer to as *presegmentation*, *segmentation refinement*, and *volumetric measurement* in the following. While the presegmentation step is interactive and requires a user interaction of about 2–5 min per scan, the subsequent steps are completely automated.

The most important deviation from the pipeline described in [23] is the refinement of the presegmentation step. Previously, we used a graph cut segmentation approach that solely relied on image intensity information. In the current setup, we use a continuous max flow segmentation approach that additionally relies on the structural *vesselness* and *csfness* features described in [24]. This change has two advantages: first, replacing graph cut with continuous max flow segmentation allows for a faster and more memory-efficient presegmentation, which especially facilitates the handling of high-resolution data sets. Second, integrating the vesselness and *csfness* features creates more robust segmentation results, making the preprocessing step of image intensity bias field correction obsolete in [23] and requiring fewer SC and background labels to be provided by the user.

In the following subsections, we give a brief recapitulation of the steps to make the paper as self-contained as possible. We point out further differences to our previous work where necessary. As the latter mainly affects the *Presegmentation* section, it contains most technical detail. Readers not interested in technicalities may safely skip this section. Suffice to say, the output of the presegmentation step is a voxel mask with two different labels for the SC and its surroundings.

### Presegmentation

In the presegmentation step, a binary mask is produced that roughly separates the SC section of interest from the surrounding cerebrospinal fluid (CSF) and non-SC tissue (called “background” in the following). Via a graphical user interface (GUI), the following user input has to be provided: a delineation of the box-shaped region of interest, a set of  $M$  voxels labeled by the user as definitely belonging to the SC, defining the labeled SC area  $\Omega_s$ , a similar set of  $N$  voxels for the background, defining the labeled background area  $\Omega_b$ , and an anatomical landmark that indicates the location of medullopontine sulcus. The

<sup>1</sup> <https://github.com/spezold/cordial>

landmark is later used as a point of reference in the volumetric measurement step (Fig. 1). For the actual presegmentation, the method relies on the continuous max flow algorithm with cross-sectional similarity prior described in [24], which is based on Yuan et al. work [32]. In contrast to [24], we now use a supervised version of the algorithm that utilizes the local information we can gain from the manual user input; that is, each point labeled as SC by the user has to lie in the SC region of the segmentation result, and each background point has to lie in the result's background. In brief, the approach models the segmentation problem in terms of a network of flows, where each point  $x = (x_1, x_2, x_3)^T \in \Omega$  in the image domain  $\Omega \subset \mathbb{R}^3$  is thought to be connected to a flow-emitting source  $s$ , a flow-consuming sink  $t$ , and its neighboring points. Furthermore, each connection is supposed to have a certain flow capacity, which is denoted as  $C_s(x)$  for the connection to the source,  $C_t(x)$  for the connection to the sink,  $\beta$  for the connection to neighbors along the  $x_3$  axis (assuming that the SC is roughly aligned with this axis), and  $\alpha$  for the connection to the remaining neighbors. The algorithm tries to maximize the flow emitted by the source while adhering to the constraints imposed by the flow capacities and by the manual labeling. The segmentation result is given by the surface that cuts through all saturated connections and in this way separates the SC from the background.

In particular, the algorithm aims at maximizing the functional

$$\max_{p_s, p_t, q, r} \int_{\Omega} l_b p_s dx - \int_{\Omega} l_c p_t dx, \tag{1}$$

subject to the flow capacity constraints

$$p_s(x) \leq C_s(x), \quad p_t(x) \leq C_t(x), \quad \|q(x)\| \leq \alpha, \quad |r(x)| \leq \beta \tag{2}$$

and to the flow conservation constraint

$$\text{div}_{12}q(x) + r'(x) - p_s(x) + p_t(x) = 0,$$

where  $p_s$  denotes the flow emitted by  $s$ ,  $p_t$  denotes the flow consumed by  $t$ ,  $q$  denotes the flow in slices perpendicular to the  $x_3$  axis,  $r$  denotes the flow along the  $x_3$  axis,  $\text{div}_{12}q$  denotes the divergence of  $q$  perpendicular to the  $x_3$  axis, and  $r'$  denotes the derivative of  $r$  along the  $x_3$  axis. The labeling functions

$$l_c = l_c(x) := \begin{cases} 1, & x \in \Omega_c \\ 0, & x \notin \Omega_c \end{cases} \text{ and } l_b = l_b(x) := \begin{cases} 0, & x \in \Omega_b \\ 1, & x \notin \Omega_b \end{cases}$$

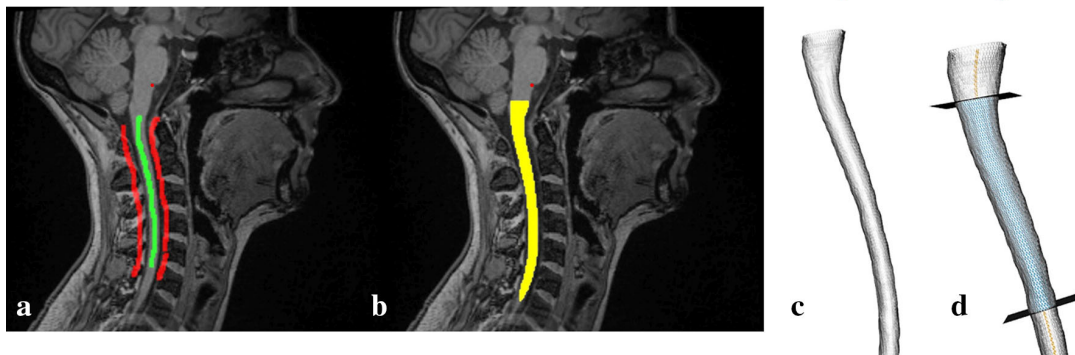
integrate the knowledge of the positions of the manually selected points. Introducing the Lagrange multiplier  $u = u(x)$ , Eq. (1) can be reformulated as:

$$\begin{aligned} \max_{p_s, p_t, q, r} \min_u E := & \int_{\Omega} l_b p_s dx - \int_{\Omega} l_c p_t dx \\ & + \int_{\Omega} u \cdot (\text{div}_{12}q + r' - p_s + p_t) dx \end{aligned}$$

subject to the flow capacity constraints (2). Thresholding  $\arg \min_u E$  such that values  $\leq 0.5$  become 0 and values  $> 0.5$  become 1 gives the desired presegmentation result.

The capacity constraint  $\beta$  represents the desired cross-sectional similarity prior: the higher its value is chosen, and the less the SC's cross-sectional shape is allowed to change in its course along the  $x_3$  axis. In this way, the prior increases the segmentation robustness.

The values of the source and sink capacities,  $C_s$  and  $C_t$ , are based on a pointwise classification of the voxels in the region of interest, using a combination of three features: their intensity normalized to the [0,1] interval, their response to Frangi's *vesselness* feature [7], and the response to our custom *csfness* feature [24]. We chose to use the vesselness feature as the SC is a tubular structure.



**Fig. 1** Spinal cord (SC) segmentation with ‘cord image analyzer’ (cordial). **a** In a first step, a subset of SC voxels (green) and background voxels (e.g., cerebrospinal fluid, vertebral bodies; red) are selected manually. In addition, an anatomical landmark—the

medullopontine sulcus—is defined manually (red dot). **b** Presegmentation result of the SC (yellow). **c** SC surface reconstruction. **d** Definition of the cutting planes for SC volume calculation

Likewise, we integrated the csfness feature as the SC is immediately surrounded by the CSF. The latter feature is derived from the vesselness feature, but responses are high for both tube-like and plate-like structures. We define it as  $w_x^* := \max_{\xi \in S_w} w(x; \xi)$  for dark structures on bright background in the predefined scales  $S_w$  with

$$w(x; \xi) = \begin{cases} 0, & \lambda_3 \leq 0 \\ \exp\left(-2 \frac{\lambda_1^2}{\lambda_3^2}\right) \left(1 - \exp\left(-\frac{\sum_{i=1}^3 \lambda_i^2}{2h^2}\right)\right) & \text{else} \end{cases}$$

where  $\lambda_i = \lambda_i(x)$  denote the ordered eigenvalues ( $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$ ) of the pointwise Hessian matrices that result from convolving the input image with Gaussian derivatives of standard deviation  $\xi$ , and  $h$  is defined as half of the maximum Frobenius norm of all Hessian matrices at scale  $\xi$ , as suggested by Frangi [7]. Classification is achieved using kernel density estimates. Let  $\mathcal{I} = [0, 1] \ni i^*$ ,  $\mathcal{V} = [0, 1] \ni v^*$ , and  $\mathcal{W} = [0, 1] \ni w^*$  be the intensity, vesselness, and csfness feature spaces with the respective image functions  $I : \Omega \rightarrow \mathcal{I}$ ,  $I_v : \Omega \rightarrow \mathcal{V}$ , and  $I_w : \Omega \rightarrow \mathcal{W}$ . Furthermore, let  $W_c = \{w_c^i\}_{i=1}^M$  and  $W_b = \{w_b^j\}_{j=1}^N$  be the csfness values of the manually determined SC and background samples, respectively. We calculate  $p_{c,w^*} := p(w^*|c) = \frac{1}{M} \sum_{i=1}^M K_{\sigma_c}(w^* - w_c^i)$ , where  $K_{\sigma_c}$  is a Gaussian kernel with zero mean and standard deviation  $\sigma_c$ . Analogously, we determine  $p_{b,w^*} := \frac{1}{N} \sum_{j=1}^N K_{\sigma_b}(w^* - w_b^j)$  as well as  $p_{c,v^*}$ ,  $p_{b,v^*}$ ,  $p_{c,i^*}$ , and  $p_{b,i^*}$ . We then calculate the pointwise source and sink capacities as  $C_s(i_x^*, v_x^*, w_x^*) = \frac{p_{c,i^*} p_{c,v^*} p_{c,w^*}}{p_{c,i^*} p_{c,v^*} p_{c,w^*} + p_{b,i^*} p_{b,v^*} p_{b,w^*}}$ ,  $C_t = 1 - C_s$ . Note that, in contrast to [24], we now use separate kernel density estimates for the features rather than a kernel density estimate for a combined feature space here, which improves the runtime while achieving comparable classification results.

*Segmentation refinement*

In the segmentation refinement step, a surface reconstruction of the presegmented part of the SC is produced that delineates the boundary between the nervous tissue and its surroundings. The processing steps have been previously described as *Segmentation Refinement* and *Surface Reconstruction* in [23], except for the fact that the robustness of the proposed feature combination allows us to omit the previously required bias field correction of the input images. In short, we first apply an anisotropic diffusion filter to the input images. We then determine the boundary of the presegmentation mask in each transverse slice of the region of interest. For an evenly distributed set of boundary points, we

extract a short 1D intensity profile and determine its point of maximum intensity change as refined boundary coordinate. This approach is based on the observation that in the MRI sequences that we use, the SC and the immediately surrounding CSF have opposed intensities. If the refined boundary coordinate cannot be determined (which is the case, for example, if the intensities on the profile are constant or change in the wrong direction), we resort to the initial contour position. Second, similar pass of profile extraction follows, which eventually leads to a stack of slicewise contours, whose vertex coordinates are independent of the input image’s sampling positions. The surface reconstruction is completed by connecting the vertices of neighboring contours in a zigzag pattern.

As previously pointed out [23], it is necessary to account for gradient non-linearity distortions for being able to measure volumes reliably in MR images. For this reason, we correct the surface reconstructions with the method described by Janke et al. [12] before acquiring measurements if required. If the DICOM header indicates that an input image has already been corrected on the MR scanner (employing the manufacturer’s respective correction routines), we omit the correction on the surface reconstructions.

*Volumetric measurement*

For measuring the volume of an SC section of length  $l$  that has a distance  $d$  to the manually placed landmark (see “Presegmentation”), we determine the centerline of the reconstructed surface segment by fitting the third degree smoothing B-spline through the centroids of its slicewise contours. The initial point for the measurement is determined by intersecting the centerline with a sphere of radius  $d$  around the landmark. The end point is defined by the centerline point that lies at a length of  $l$  inferior of the initial point, measured along the centerline. Through both points, planes are fitted perpendicular to the centerline. We then measure the volume of the surface segment that is enclosed by the planes. This volume measurement approach again largely follows the corresponding approach described in [23], with the only exception being the determination of the initial point. In preliminary scan-rescan experiments, we found the approach that uses the sphere to produce a lower coefficient of variation as the one described in [23], which we ascribe to the fact that with this procedure, no extrapolation of the centerline is necessary.

**Application to MRI data**

All human studies in this work have been approved by the appropriate ethics committee and have, therefore, been performed in accordance with the ethical standards laid

down in the 1964 Declaration of Helsinki and its later amendments.

#### Reliability assessment

To assess the reliability of *cordial*'s SC segmentation, 24 healthy controls (HC, 12 women, mean age 33.3 years, ranges 24–49 years) were scanned on a 3.0 T whole-body MR scanner (Magnetom Verio, Siemens Medical, Erlangen, Germany) with a T1w MPRAGE sequence. The sequence parameters were repetition time (TR) of 2.0 s, inversion time (TI) of 1.0 s, echo time (TE) of 3.4 ms, flip angle ( $\alpha$ ) of  $8^\circ$ , and field-of-view (FOV) of  $256 \times 256 \text{ mm}^2$ . One hundred ninety-two sagittal slices with an isotropic resolution of  $1 \text{ mm}^3$  were centered onto the chin and aligned to the cervical SC. The acquisition time for the MPRAGE was 4:08 min (parallel imaging; GRAPPA with acceleration factor 2).

In a subgroup of 18 HCs (10 women, mean age 35.0 years, ranges 24–49 years), a 3D T2w Turbo Spin Echo (TSE) sequence was additionally scanned (see Fig. 2). The sequence parameters of the T2w 3D TSE were TR = 1.5 s, TE = 226 ms, turbo factor of 99, echo spacing 4.6 ms, and FOV =  $240 \times 240 \text{ mm}^2$ . Fifty-six sagittal slices were acquired with a spatial resolution of  $0.75 \times 0.75 \times 0.75 \text{ mm}^3$  (interpolated to  $0.375 \times 0.375 \times 0.75 \text{ mm}^3$ ). These slices were also centered onto the chin and aligned to the cervical SC. The acquisition time for the T2w 3D TSE was 5:39 min (parallel imaging; GRAPPA with acceleration factor 2, 1.5 averages).

Both scans were acquired with a combination of a 12-element head matrix coil and a 4-element neck matrix coil. Both the MPRAGE and 3D TSE were each repeated three times. Between the first and the second scan block, the subject remained in the scanner (back-to-back scan—best-case scenario). Before the third block, the subject left

the scanner and was then repositioned and rescanned. This third scan was performed to assess the additional variances introduced by positioning and B0 shim, as both effects are relevant in follow-up scans.

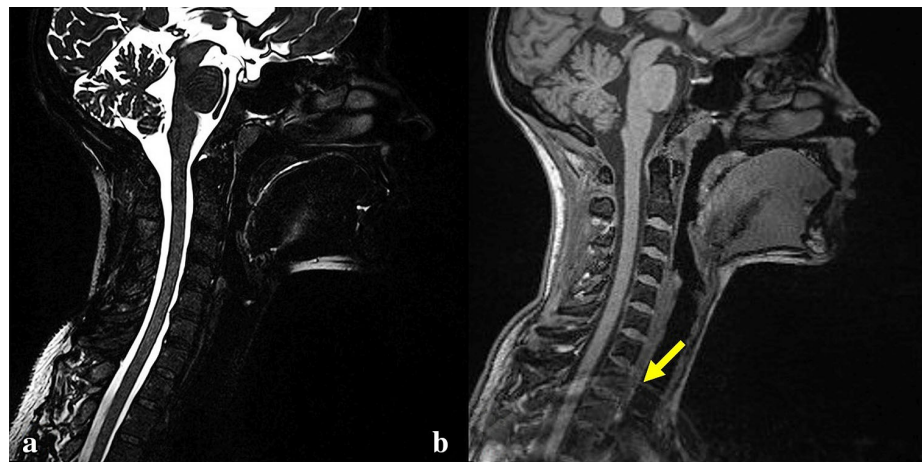
For each measurement, the cervical SC volume (CSCV) was calculated. With the definition summarized in “volumetric measurement”,  $d$  (the distance between the superior cutting plane of the SC segment and the medullopontine sulcus) was defined to be 30 mm. The length  $l$  of the SC segment was defined to be 50 mm for the MPRAGE sequence. In our data, this was the maximum feasible length as in more inferior parts of the SC aliasing artifacts (shoulders, etc.) severely degraded the image quality (Fig. 2b). In the 3D TSE sequence, a slab-selective excitation pulse was applied. Therefore, the 3D TSE images had minor fold-over artifacts. For this reason, the CSCV for the 3D TSE images was calculated for  $l = 50 \text{ mm}$  and  $l = 90 \text{ mm}$ .

The reproducibility of CSCV calculation was assessed using the coefficient of variation (COV) independently for each MRI sequence. COV was defined as the percentage of the standard deviation (SD) normalized to the mean CSCV:

$$\text{COV} = 100 \cdot \frac{\text{SD}(\text{CSCV})}{\text{mean}(\text{CSCV})} \%$$

Two experienced scientists (M. A., S. P.) performed the presegmentation. Segmentation refinement and volumetric measurement were automatically processed by *cordial*. COV for scan–rescan reliability was calculated separately for each rater between the first and the second as well as between the first and the third scan. Each data evaluation was performed twice, with an interval of at least 48 h between the two evaluations (runs). Intrarater reliability was calculated for each rater as the COV between run one and run two. Interrater reliability was calculated as the COV between the two raters, separately for run one and run two. All available data sets were used in this evaluation

**Fig. 2** Sample MRI data applicable to segmentation with *cordial*. **a** T2-weighted 3D Turbo Spin Echo (T2w 3D TSE) of the SC. The T2w 3D TSE image demonstrates a high contrast between the SC cord and the cerebrospinal fluid. **b** Same healthy control scanned with a T1-weighted MPRAGE sequence. In the lower part of the field of view, an aliasing artifact of the shoulder is apparent (yellow arrow). This impairs cord segmentation in the respective area



(i.e., data sets of 24 HC for MPRAGE and data sets of 18 HC for 3D TSE). In addition, we tested for possible significant dependencies of CSCV and demographic factors (age and gender).

#### Evaluation of patient data and clinical correlations

To evaluate the applicability of *cordial* to clinical MRI data with a focus on longitudinal settings, we segmented the MRI data of an existing cohort of patients with progressive MS in whom 3-year clinical and MRI follow-up data were available (two points in time). All patients were participants in an ongoing study on the phenotype–genotype characterization of MS [1]. They were treated with disease modifying immunomodulatory treatments at the discretion of the treating physician. The patients underwent a detailed clinical neurological assessment including Expanded Disability Status Scale (EDSS) scoring by a neurostatus certified neurologist (<http://www.neurostatus.net>). Informed consent was obtained in writing from all participating patients, in accordance with the local ethics committee approval and the declaration of Helsinki. Forty-eight MS patients were included in this analysis (24 women). Mean age of the patients at baseline was 52.1 years (ranges 22–67 years); mean disease duration was 14.0 years with a range of 1–50 years. The median baseline EDSS was 4.5 (ranges 2.5–6.5). Thirty of the included patients had a secondary progressive (SPMS) disease course; 18 patients had primary progressive MS (PPMS). Additional clinical and demographic characteristics of the patients are summarized in Table 1.

The patients were scanned on a 1.5 T whole-body MR scanner (Avanto, Siemens Medical, Germany) using a 12-element head matrix coil. The scanning protocol encompassed 3D MPRAGE head scans with TR = 2.08 s, TI = 1.1 s, TE = 3.93 ms, and  $\alpha = 15^\circ$  (no parallel imaging). One hundred sixty sagittal slices parallel to the interhemispheric fissure were acquired with an in-plane

resolution of  $0.98 \times 0.98 \text{ mm}^2$  and a slice thickness of 1 mm. The measurements were performed at two points in time  $\sim 3$  years apart (mean follow-up period 3.1 years and ranges 2.9–3.6 years) using an identical MRI protocol and the same scanner.

For each point in time, CSCV was computed; then, annualized CSCV change ( $\Delta\text{CSCV}$ ) was calculated for each patient according to

$$\Delta\text{CSCV} = \frac{\text{CSCV}(\text{BL}) - \text{CSCV}(\text{FU})}{\Delta t},$$

where  $\Delta t$  is the period between baseline (BL) and follow-up (FU).

Differences in CSCV between baseline and follow-up were compared using Wilcoxon sign-rank test, whereas  $\Delta\text{CSCV}$  was compared between SPMS and PPMS patients by means of Mann–Whitney U test for independent samples. To determine correlations between cervical SC atrophy and disability progression, we tested two different models using a multiple linear regression analysis with EDSS change over 3 years ( $\Delta\text{EDSS}$ ) as dependent variable. In the first model,  $\Delta\text{CSCV}$  was included as independent variable; in the second model, baseline CSCV was included instead. For both models, we calculated whole brain atrophy using SIENA [26, 27] and included annualized percent brain volume change as independent variable into the MLR model. In addition, the demographic factors age, disease duration, disease course, and gender were also included in both models.

## Results

### Data processing times

The manual intervention (delineation of the region of interest, manual selection of a subset of SC, and a subset of background voxels, and definition of the medullopontine

**Table 1** Demographic and clinical characteristics of the patients

	PPMS ( <i>N</i> = 18)	SPMS ( <i>N</i> = 30)	<i>p</i> value
Gender	9 M/9 F	15 M/15 F	
Age	46.5 years (21–62 years)	55.4 years (39–67 years)	0.002
Disease duration	9.5 years (1–29 years)	20.5 years (4–47 years)	<0.001
EDSS (BL)	4 (2.5–6)	4.5 (2.5–6.5)	0.3
EDSS (FU)	4.5 (2.5–6.5)	6 (3.0–7.5)	0.2
$\Delta\text{EDSS}$	0.5 (–1.0–2.5)	0.5 (–0.5–3.5)	0.9
$\Delta t$	3.1 (2.9–3.4)	3.1 (2.9–3.4)	0.2

PPMS primary progressive multiple sclerosis, SPMS secondary progressive MS, BL baseline, FU follow-up, EDSS Expanded Disability Status Scale,  $\Delta\text{EDSS}$  change of EDSS between BL and FU,  $\Delta t$  period between BL and FU measurement, mean (range) values are stated for age, disease duration, and  $\Delta t$ , median (range) values are given for EDSS. Statistical comparisons between the groups were calculated by Mann–Whitney U tests for independent samples

sulcus landmark) and the automated presegmentation can be performed in approximately 2–5 min per data set including the visual review of the presegmentation result. The processing time depends on the spatial resolution of the respective data set and on the region of interest size. The remaining, fully automated steps of surface reconstruction, distortion correction, and volumetric measurement, which are largely implemented in Python code, take another 2–8 min processing time on a standard CPU (Intel i7-2600, 8 × 3.4 GHz). The largest part of the processing time is used up by the surface reconstruction, and again strongly depends on the spatial resolution and on the region of interest size.

**Reliability assessment in healthy subjects**

For the MPRAGE data (*l* = 50 mm), mean CSCV for the healthy controls was 3852 mm<sup>3</sup> (ranges 3117–4447 mm<sup>3</sup>); for the subgroup scanned both with MPRAGE and T2w 3D TSE, mean CSCV was 3819 mm<sup>3</sup> (ranges 3117–4447 mm<sup>3</sup>). In the 3D TSE measurements, mean CSCV was 4248 mm<sup>3</sup> (ranges 3570–4967 mm<sup>3</sup>) for *l* = 50 mm, which was significantly higher (*p* < 0.001; Wilcoxon sign-rank test) than in the MPRAGE data. However, the intra-class correlation coefficient for consistency (two-way mixed model with subjects as random factor and the MR techniques as fixed factor) was 0.949, indicating very strong consistency between the two techniques. The CSCV results, also including values for the 3D TSE data with *l* = 90 mm, are summarized in Table 2.

The coefficients of variation for the different sequences are listed in Tables 3, 4, and 5. For all sequences and comparisons,

the COV were less than 1 %. The intra- and interrater COV were in the range of 0.1–0.2 %. The scan–rescan COV was higher between scan one and three (subject was repositioned in-between: 0.90–0.95 % for MPRAGE; 0.41–0.53 % for TSE) than between scan one and two (no repositioning: 0.47–0.53 % for MPRAGE; 0.32–0.43 % for TSE).

No significant CSCV differences were found between male and female controls, neither in the T1w data sets (men: 3941 ± 329 mm<sup>3</sup>, women: 3763 ± 351 mm<sup>3</sup>; *t* test: *p* = 0.22) nor in the T2w data sets (*l* = 50 mm; men: 4406 ± 362 mm<sup>3</sup>, women: 4123 ± 367 mm<sup>3</sup>; *p* = 0.12); (*l* = 90 mm; men: 8109 ± 714 mm<sup>3</sup>, women: 7627 ± 699 mm<sup>3</sup>; *p* = 0.17). In addition, no significant correlation (Pearson correlation) between CSCV and age was observed [T1w: *p* = 0.30; T2w(50 mm): *p* = 0.33; T2w(90 mm): *p* = 0.36].

**Evaluation of patient data and clinical correlations**

The PPMS patients in our cohort were significantly younger and had a significantly shorter disease duration than the SPMS patients (see Table 1). Mean annualized cervical cord atrophy rates of 0.5 % were found (PPMS: 0.9 %; SPMS: 0.2 %). In the whole cohort as well as in PPMS, significant CSCV differences between baseline and follow-up were found, but not in SPMS (cohort: *p* = 0.004; PPMS: *p* = 0.008; SPMS: *p* = 0.11). We did not observe significant CSCV differences between PPMS and SPMS neither at baseline nor at 3-year follow-up. However, the volume loss per year ( $\Delta$ CSCV) was by trend higher in PPMS than in SPMS (*p* = 0.053). The CSCV and  $\Delta$ CSCV values of the MS patients (mean ± standard deviation) are summarized in

**Table 2** Cervical spinal cord volume (CSCV) in healthy controls as measured using different MRI sequences

	MPRAGE ( <i>l</i> = 50 mm); <i>N</i> = 24	MPRAGE ( <i>l</i> = 50 mm); <i>N</i> = 18	T2w 3D TSE ( <i>l</i> = 50 mm); <i>N</i> = 18	T2w 3D TSE ( <i>l</i> = 90 mm); <i>N</i> = 18
Mean CSCV	3852	3819	4248	7841
Range	3117–4447	3117–4447	3570–4967	6546–5159

*N* number of subjects and *l* length of the spinal cord section. Mean CSCV and range are given in mm<sup>3</sup>. The 18 subjects in column 2 are the same as in columns 3 and 4

**Table 3** Coefficients of variation (COV) for the MPRAGE measurements with a cord segment length of *l* = 50 mm (in brackets: 95 % confidence interval)

MPRAGE, <i>l</i> =50 mm			Intrater COV (%)	run #1 / run #2
scan–rescan COV (%)	scan #1 / scan #2	scan #1 / scan #3	rater #1	0.12 (0.09–0.14)
rater #1, run #1	0.53 (0.35–0.71)	0.95 (0.60–1.30)	rater #2	0.08 (0.05–0.11)
rater #1, run #2	0.53 (0.35–0.70)	0.95 (0.60–1.29)	<b>Interrater COV (%)</b>	rater #1 / rater #2
rater #2, run #1	0.46 (0.29–0.63)	0.90 (0.55–1.24)	run #1	0.17 (0.12–0.23)
rater #2, run #2	0.47 (0.30–0.64)	0.90 (0.55–1.25)	run #2	0.14 (0.10–0.18)

Subjects were scanned three times, between scan #2 and scan #3, they were repositioned. Segmentation was performed twice by two different raters (for details see text)



**Table 4** Coefficients of variation (COV) for the T2w 3D TSE measurements with a cord segment length of  $l = 50$  mm (in brackets: 95 % confidence interval)

T2w 3D TSE, $l=50$ mm			Intrarater COV (%)	run #1 / run #2
scan-rescan COV (%)	scan #1 / scan #2	scan #1 / scan #3	rater #1	0.06 (0.05-0.08)
rater #1, run #1	0.34 (0.20-0.47)	0.50 (0.32-0.68)	rater #2	0.05 (0.03-0.06)
rater #1, run #2	0.32 (0.19-0.45)	0.47 (0.30-0.64)	Interrater COV (%)	rater #1 / rater #2
rater #2, run #1	0.33 (0.20-0.46)	0.50 (0.34-0.66)	run #1	0.07 (0.05-0.08)
rater #2, run #2	0.33 (0.20-0.45)	0.48 (0.31-0.64)	run #2	0.06 (0.05-0.08)

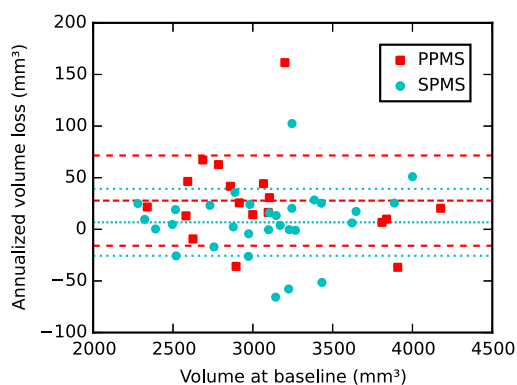
**Table 5** Coefficients of variation (COV) for the T2w 3D TSE measurements with a cord segment length of  $l = 90$  mm (in brackets: 95 % confidence interval)

T2w 3D TSE, $l=90$ mm			Intrarater COV (%)	run #1 / run #2
scan-rescan COV (%)	scan #1 / scan #2	scan #1 / scan #3	rater #1	0.15 (0.09-0.22)
rater #1, run #1	0.43 (0.29-0.57)	0.48 (0.30-0.66)	rater #2	0.16 (0.09-0.23)
rater #1, run #2	0.36 (0.20-0.52)	0.41 (0.27-0.56)	Interrater COV (%)	rater #1 / rater #2
rater #2, run #1	0.43 (0.28-0.57)	0.45 (0.26-0.64)	run #1	0.22 (0.15-0.30)
rater #2, run #2	0.36 (0.21-0.51)	0.53 (0.26-0.80)	run #2	0.16 (0.09-0.23)

**Table 6** Cervical spinal cord volume (CSCV) and annualized volume loss ( $\Delta$ CSCV) in patients with progressive MS (mean  $\pm$  standard deviation)

CSCV ( $\text{mm}^3$ )	BL	3-year FU	$\Delta$ CSCV	$p$ value (between FU and BL)
All patients	3072 $\pm$ 466	3028 $\pm$ 482	14.7 $\pm$ 38.0	0.004
PPMS	3082 $\pm$ 521	2999 $\pm$ 561	27.8 $\pm$ 43.7	0.008
SPMS	3066 $\pm$ 439	3045 $\pm$ 437	6.8 $\pm$ 32.4	0.11
$p$ value (between groups)	0.7	0.04	0.05	

Differences in CSCV between FU and BL were calculated with Wilcoxon signed rank test; whereas differences in  $\Delta$ CSCV between the two subgroups were calculated with Mann–Whitney U test for independent samples



**Fig. 3** Annualized cervical spinal cord volume loss ( $\Delta$ CSCV) for the individual multiple sclerosis (MS) patients. Each spot represents a patient; red squares mark primary progressive MS (PPMS) patients; cyan circles mark secondary progressive MS (SPMS) patients. Lines indicate the means and standard deviations of the  $\Delta$ CSCV measurements; dashed red lines are used for the PPMS patient group and dotted cyan lines for the SPMS group. Note that positive  $\Delta$ CSCV values mean volume loss (atrophy)

Table 6. Figure 3 shows the  $\Delta$ CSCV values for the individual patients.

The two different models with  $\Delta$ EDSS as dependent variable yielded the following results: in the multiple linear regression analysis with  $\Delta$ CSCV as independent variable, the annualized CSCV loss was the only significant predictor of EDSS change ( $p = 0.02$ ). Neither annualized percent brain volume change nor any of the demographic factors reached significance. In the model with baseline CSCV as independent variable, age ( $p = 0.001$ ), disease course ( $p = 0.03$ ), and baseline CSCV ( $p = 0.01$ ) were significant predictors. Again, annualized percent brain volume change was not significant.

### Discussion

In this work, we present *cordial*, a mostly automated tool for SC volumetry, optimized for the analysis of follow-up MRI data. The manual input to *cordial* is limited to the

selection of a few voxels belonging to the SC and the background, the delineation of a box-shaped region of interest (selection of three corner marks) and the definition of an anatomical landmark. The presegmentation algorithm works well even if SC and background voxels are selected in only one slice within the region of interest. Voxel selection is straightforward, drawing strokes for SC, and background by clicking and dragging with the mouse. For the majority of data sets, one stroke for marking parts of the SC and about two to five strokes that partially cover the CSF, vertebrae, and other non-cord tissue are sufficient. The strokes within the SC do not need to outline or be close to the outer border of the SC (for example, see Fig. 1a). User interactions, including visual inspection and potential corrections of the presegmentation, are performed in less than 5 min. While we already demonstrated a completely automated version of the underlying segmentation approach in [24], we believe that the proposed set-up provides a good compromise between manual interaction time and the possibility to easily correct segmentation errors. All subsequent steps of *cordial* are completely automated and can be executed in batch mode. Therefore, *cordial* allows the processing of large data set in clinical practice or clinical trial data sets.

In all data sets processed, no presegmentation mistakes were identified unless the selected region of interest encompassed the cerebellar peduncle. With both MRI sequences used in this work, the cerebellar tissue is iso-intense to the medulla oblongata and the SC; therefore, *cordial* is not able to separate the cerebellum from the SC in these sequences. This segmentation failure can be easily corrected by an inferior positioning of the respective corner mark for the region of interest. After segmentation refinement, no segmentation failures were identified neither in healthy subjects nor in MS patients, even if spinal cord lesions were present.

In *cordial*, the surface and centerline of the SC are reconstructed. In combination with 3D MRI techniques providing high spatial resolution, this allows a flexible choice of which level and volume of the SC is to be studied. Frequently in SC imaging, the cross-sectional area at a specific anatomical landmark is calculated and used in statistical evaluations. We decided to use the volume of an SC section instead. As the SC is a relatively flexible structure, its inclination is dependent on the subject's positioning. We believe that the volume of a specific SC section is less influenced by this effect than a cross-sectional area at a certain level relative to the spine. However, further evaluation will be necessary to provide a quantitative justification. In this work, we also propose an anatomical landmark on the tissue of interest itself (the medullopontine sulcus in this work) instead of a bony landmark. We believe that this choice reduces the effect of

spinal cord movements relative to the landmark due to flexion differences between scans. Again, further experiments will be necessary for quantifying these effects. If SC parts other than the cervical region are evaluated, other reference landmarks could be chosen. In a yet unpublished study of the lumbar spine, for example, we used a landmark on the tip of the conus medullaris instead and reached similar scan–rescan COV results of  $1.4 \pm 0.8\%$ . Regarding both cross-sectional area measurements and the use of other anatomical landmarks, we would like to point out that as a side effect of reconstructing the surface of a whole SC segment and its centerline, the measurement of cross-sectional areas at any point along the section is possible at virtually no additional cost [4]. In our work, we defined a constant distance between the segmentation volume and the landmark as well as a constant segment length. We believe that this choice is appropriate in studies focused on longitudinal changes like the patient study presented here. In cross-sectional studies, individual variations of anatomy could be further minimized by normalization as proposed by Oh and colleagues [22].

In the healthy control subjects, the calculated CSCV was significantly higher in the T2w TSE data compared to the T1 MPRAGE data. This difference is not surprising, as the tissue contrast is very different in both techniques (see Fig. 2), providing differing intensity profiles. Other groups have also reported similar sequence-dependent effects [13]. Despite these apparent differences, the volumes measured with both techniques are highly correlated with an intra-class correlation coefficient of 0.949.

In the reliability assessment, the coefficients of variation of *cordial* were below 1 % in all comparisons with intra- and interrater reliability COV in the range of 0.1–0.2 %. Scan–rescan reliability was slightly less accurate, indicating some slight impact of patient positioning and posture on the segmentation results. The COV between the first and the third scans were higher (subject was completely repositioned before scan #3) than those between the first and the second scans. This effect was especially apparent when using the MPRAGE data and was less pronounced in the 3D TSE scans (without repositioning 0.47–0.53 % for MPRAGE; 0.32–0.43 % for TSE with repositioning 0.90–0.95 % for MPRAGE; 0.41–0.53 % for TSE). There are two probable reasons for the higher variance of the MPRAGE volumes in our study. First, the real spatial resolution of the 3D TSE images is 2.4 times higher than that of the MPRAGE images. Therefore, the MPRAGE data are more sensitive to partial volume effects. Second, the T2w 3D TSE images have an approximately eight times higher contrast-to-noise ratio between SC and adjacent CSF. For this reason, the intensity profile gradient on the 3D TSE images is much steeper, and slight differences in the profile have less influence on the determination of the

intensity change maximum. Despite the higher COV of the MPRAGE data sets, our results are comparable or superior to those of other published methods. Calculating cross-sectional areas, Losseff et al. [16] reported scan–rescan COV of 0.79 %, interrater COV of 0.83 % and intrarater COV of 0.73 % on 3D T1w MR data with a resolution of  $0.98 \times 0.98 \times 1 \text{ mm}^3$ . Segmenting T1w data with the same resolution, Coulon and colleagues [4] reported intrarater COV of 0.77 % for cross-sectional areas and 1.36 % for volumes, and scan–rescan COV of 1.31 % (areas) and of 1.35 % (volumes). In the work of Horsfield et al. [11], the COV of 0.44–0.59 % are reported for intrarater variability and 1.07–1.36 % for interrater variability, segmenting MPRAGE data with a spatial resolution of  $1.09 \times 1.09 \times 1.0 \text{ mm}^3$ . Calculating the mean cross-sectional spinal cord area over a section of 50 mm thickness in MPRAGE data, Lukas and colleagues [17] found intrarater COV of 0.49 %, interrater COV of 1.24 %, and scan–rescan COV of 1.33 %.

Applying *cordial* to the MRI data of a clinical follow-up study, we found mean annualized cervical cord atrophy rates ( $\Delta\text{CSCV}$ ) of 0.5 % in patients with progressive MS (PPMS 0.9 %; SPMS 0.2 %). This is in line with the work of Laule and colleagues [15], which reported a mean SC volume atrophy rate of 0.8 % at the level C2–C3 for PPMS patients. Lukas et al. [18] have reported higher atrophy rates in progressive MS patients (PPMS 1.9 %, SPMS 2.0 %, measured also at level C2–C3).

In our patient study, the  $\Delta\text{CSCV}$  was by trend ( $p = 0.053$ ) higher in the PPMS patients compared to the SPMS patients. However, both mean age and mean disease duration were significantly longer in the SPMS patients. If adjusted for age and disease duration, the  $\Delta\text{CSCV}$  differences between the two patient groups disappeared ( $p = 0.8$ ). One might speculate that age- and disease duration-correlated effects could mimic an apparent effect of disease course onto cervical SC atrophy. More interestingly, the change in EDSS over 3 years significantly correlates with  $\Delta\text{CSCV}$ , independent of brain atrophy and any demographic factor included in our multiple linear regression model. In the alternative MLR model, also baseline CSCV was a significant predictor of EDSS change. This emphasizes the impact of the upper SC to clinical disability and the need for having appropriate methods to study SC changes over time.

In conclusion, we could demonstrate that *cordial* allows a highly reliable volumetry of the SC. Most importantly, we demonstrated the clinical applicability of our method in longitudinal settings by segmenting the SC in follow-up data sets of progressive multiple sclerosis patients. The segmentation in *cordial* is largely automated, minimizing operator bias and cost. Because of the graphical user interface, even non-experienced users can easily provide

the manual input to *cordial*. The mainly automated processing steps in combination with the possibility of volume reconstructions and the definition of a landmark defined on the tissue of interest itself imply that the new method is not only reliable, but also highly flexible with potential applications not only in MS, but also generally in diseases affecting the SC, also beyond the cervical SC.

#### Compliance with ethical standards

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest regarding the content of this article.

**Ethical standards** All human studies in this work have been approved by the appropriate ethics committee and have therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

#### References

1. Bendfeldt K, Kuster P, Traud S, Egger H, Winkhofer S, Mueller-Lenke N, Naegelin Y, Gass A, Kappos L, Matthews PM, Nichols TE, Radue EW, Borgwardt SJ (2009) Association of regional gray matter volume loss and progression of white matter lesions in multiple sclerosis—a longitudinal voxel-based morphometry study. *NeuroImage* 45:60–67
2. Carbonell-Caballero J, Manjon JV, Marti-Bonmati L, Olalla JR, Casanova B, de la Iglesia-Vaya M, Coret F, Robles M (2006) Accurate quantification methods to evaluate cervical cord atrophy in multiple sclerosis patients. *Magma* 19:237–246
3. Chen M, Carass A, Oh J, Nair G, Pham DL, Reich DS, Prince JL (2013) Automatic magnetic resonance spinal cord segmentation with topology constraints for variable fields of view. *NeuroImage* 83:1051–1062
4. Coulon O, Hickman SJ, Parker GJ, Barker GJ, Miller DH, Arridge SR (2002) Quantification of spinal cord atrophy from magnetic resonance images via a B-spline active surface model. *Magn Reson Med* 47:1176–1185
5. De Leener B, Cohen-Adad J, Kadoury S (2015) Automatic segmentation of the spinal cord and spinal canal coupled with vertebral labeling. *IEEE Trans Med Imaging* 34:1705–1718
6. Fonov VS, Le Troter A, Taso M, De Leener B, Leveque G, Benhamou M, Sdika M, Benali H, Pradat PF, Collins DL, Callot V, Cohen-Adad J (2014) Framework for integrated MRI average of the spinal cord white and gray matter: the MNI-Poly-AMU template. *NeuroImage* 102(Pt 2):817–827
7. Frangi A, Niessen W, Vincken K, Viergever M (1998) Multiscale vessel enhancement filtering. In: Wells W, Colchester A, Delp S (eds) *Medical image computing and computer-assisted intervention—MICCAI'98*. Springer, Berlin Heidelberg, pp 130–137
8. Han JS, Kaufman B, El Yousef SJ, Benson JE, Bonstelle CT, Alfidji RJ, Haaga JR, Yeung H, Huss RG (1983) NMR imaging of the spine. *AJR Am J Roentgenol* 141:1137–1145
9. Harrison DE, Cailliet R, Harrison DD, Troyanovich SJ, Harrison SO (1999) A review of biomechanics of the central nervous system—part II: spinal cord strains from postural loads. *J Manipulative Physiol Ther* 22:322–332
10. Hickman SJ, Hadjiprocopis A, Coulon O, Miller DH, Barker GJ (2004) Cervical spinal cord MTR histogram analysis in multiple sclerosis using a 3D acquisition and a B-spline active surface segmentation technique. *Magn Reson Imaging* 22:891–895

11. Horsfield MA, Sala S, Neema M, Absinta M, Bakshi A, Sormani MP, Rocca MA, Bakshi R, Filippi M (2010) Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: application in multiple sclerosis. *NeuroImage* 50:446–455
12. Janke A, Zhao H, Cowin GJ, Galloway GJ, Doddrell DM (2004) Use of spherical harmonic deconvolution methods to compensate for nonlinear gradient effects on MRI images. *Magn Reson Med* 52:115–122
13. Kearney H, Yiannakas MC, Abdel-Aziz K, Wheeler-Kingshott CA, Altmann DR, Ciccarelli O, Miller DH (2014) Improved MRI quantification of spinal cord atrophy in multiple sclerosis. *J Magn Reson Imaging JMRI* 39:617–623
14. Koh J, Kim T, Chaudhary V, Dhillon G (2010) Automatic segmentation of the spinal cord and the dural sac in lumbar MR images using gradient vector flow field. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2010:3117–3120
15. Laule C, Vavasour IM, Zhao Y, Traboulsee AL, Oger J, Vavasour JD, Mackay AL, Li DK (2010) Two-year study of cervical cord volume and myelin water in primary progressive multiple sclerosis. *Mult Scler* 16:670–677
16. Losseff NA, Webb SL, O’Riordan JI, Page R, Wang L, Barker GJ, Tofts PS, McDonald WI, Miller DH, Thompson AJ (1996) Spinal cord atrophy and disability in multiple sclerosis. A new reproducible and sensitive MRI method with potential to monitor disease progression. *Brain: J Neurol* 119(Pt 3):701–708
17. Lukas C, Bellenberg B, Rexilius GW, Hahn HK, Köster O, Schimrigk SK (2006) MR-based measurement of spinal cord atrophy in multiple sclerosis: reproducibility and sensitivity of a new semi-automated procedure. *Eur Radiol* 16:458
18. Lukas C, Knol DL, Sombekke MH, Bellenberg B, Hahn HK, Popescu V, Weier K, Radue EW, Gass A, Kappos L, Naegelin Y, Uitdehaag BM, Geurts JJ, Barkhof F, Vrenken H (2015) Cervical spinal cord volume loss is related to clinical disability progression in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 86:410–418
19. Modic MT, Weinstein MA, Pavlicek W, Starnes DL, Duchesneau PM, Boumprey F, Hardy RJ Jr (1983) Nuclear magnetic resonance imaging of the spine. *Radiology* 148:757–762
20. Mukherjee DP, Cheng I, Ray N, Mushahwar V, Lebel M, Basu A (2010) Automatic segmentation of spinal cord MRI using symmetric boundary tracing. *IEEE Trans Inform Technol Biomed: Publ IEEE Eng Med Biol Soc* 14:1275–1278
21. Norman D, Mills CM, Brant-Zawadzki M, Yeates A, Crooks LE, Kaufman L (1983) Magnetic resonance imaging of the spinal cord and canal: potentials and limitations. *AJR Am J Roentgenol* 141:1147–1152
22. Oh J, Seigo M, Saidha S, Sotirchos E, Zackowski K, Chen M, Prince J, Diener-West M, Calabresi PA, Reich DS (2014) Spinal cord normalization in multiple sclerosis. *J Neuroimaging: Off J Am Soc Neuroimaging* 24:577–584
23. Pezold S, Amann M, Weier K, Fundana K, Radue E, Sprenger T, Cattin P (2014) A semi-automatic method for the quantification of spinal cord atrophy. In: Yao J, Klinder T, Li S (eds) *Computational methods and clinical applications for spine imaging*. Springer International Publishing, pp 143–155
24. Pezold S, Fundana K, Amann M, Andelova M, Pfister A, Sprenger T, Cattin P (2015) Automatic segmentation of the spinal cord using continuous max flow with cross-sectional similarity prior and tubularity features. In: Yao J, Glocker B, Klinder T, Li S (eds) *Recent advances in computational methods and clinical applications for spine imaging*. Springer International Publishing, pp 107–118
25. Reid JD (1960) Effects of flexion-extension movements of the head and spine upon the spinal cord and nerve roots. *J Neurol Neurosurg Psychiatry* 23:214–221
26. Smith SM, De Stefano N, Jenkinson M, Matthews PM (2001) Normalized accurate measurement of longitudinal brain change. *J Comput Assist Tomogr* 25:466–475
27. Smith SM, Zhang Y, Jenkinson M, Chen J, Matthews PM, Federico A, De Stefano N (2002) Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *NeuroImage* 17:479–489
28. Stevenson VL, Leary SM, Losseff NA, Parker GJ, Barker GJ, Husmani Y, Miller DH, Thompson AJ (1998) Spinal cord atrophy and disability in MS: a longitudinal study. *Neurology* 51:234–238
29. Valsasina P, Rocca MA, Horsfield MA, Copetti M, Filippi M (2015) A longitudinal MRI study of cervical cord atrophy in multiple sclerosis. *J Neurol* 262:1622–1628
30. Van Uitert R, Bitter I, Butman JA (2005) Semi-automatic spinal cord segmentation and quantification. *Int Congr Ser* 1281:224–229
31. Yiannakas MC, Mustafa AM, De Leener B, Kearney H, Tur C, Altmann DR, De Angelis F, Plantone D, Ciccarelli O, Miller DH, Cohen-Adad J, Gandini Wheeler-Kingshott CAM (2016) Fully automated segmentation of the cervical cord from T1-weighted MRI using PropSeg: Application to multiple sclerosis. *NeuroImage: Clinical* 10:71–77
32. Yuan J, Bae E, Tai XC (2010) A study on continuous max-flow and min-cut approaches. *Proc Cvpr Ieee*:2217–2224

## 8 Automatic, Robust, and Globally Optimal Segmentation of Tubular Structures

In the approach presented in this chapter, we continue our adaptations of the continuous max flow algorithm. Rather than requiring a fixed axis of orientation for the tubular structure of interest, like we did in Chapters 6 and 7, we now exploit the directional information of the vesselness feature and thus the tubular structure's orientation itself for guiding the segmentation. In order to enable regularization along arbitrary directions, we integrate an anisotropic total variation regularizer into the segmentation algorithm. The presented approach is fully automatic, which reduces necessary user interaction. Furthermore, it makes the previously required segmentation refinement step of our toolchain obsolete, which results in a shorter overall processing time and reduces the number of possible points of failure.

For evaluation, we segment noisy images of a simulated helical phantom both with isotropic and anisotropic regularization, demonstrating the benefits of the latter. We compare the proposed approach to a state-of-the-art spinal cord segmentation framework using images of healthy subjects, where we achieve comparable segmentation quality in terms of mean surface distance, Hausdorff distance, and Dice coefficient with respect to corresponding manual segmentations that were provided by two clinical experts.

Two appendices conclude the chapter: In Appendix 8.A, we derive Eq. (4) of the paper, that is, the minimum cut–maximum flow equality for the nonterminal flow  $p$ . In Appendix 8.B, we derive the first algorithm step for updating  $p$ . We added the derivations here for completeness as we could not add the information to the paper due to a page limit.

**Publication.** The proposed approach was presented at the *19th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, October 2016, Athens, Greece. It was published<sup>1</sup> as part of the conference proceedings [36].

---

<sup>1</sup>[http://dx.doi.org/10.1007/978-3-319-46726-9\\_42](http://dx.doi.org/10.1007/978-3-319-46726-9_42) (last accessed on November 3, 2016)

# Automatic, Robust, and Globally Optimal Segmentation of Tubular Structures

Simon Pezold<sup>1</sup>(✉), Antal Horváth<sup>1</sup>, Ketut Fundana<sup>1</sup>, Charidimos Tsagkas<sup>2</sup>,  
Michaela Andělová<sup>2,3</sup>, Katrin Weier<sup>2</sup>, Michael Amann<sup>2</sup>,  
and Philippe C. Cattin<sup>1</sup>

<sup>1</sup> Department of Biomedical Engineering, University of Basel, Basel, Switzerland  
simon.pezold@unibas.ch

<sup>2</sup> Department of Neurology, University Hospital Basel, Basel, Switzerland

<sup>3</sup> Department of Internal Medicine, University Hospital Motol,  
Prague, Czech Republic

**Abstract.** We present an automatic three-dimensional segmentation approach based on continuous max flow that targets tubular structures in medical images. Our method uses second-order derivative information provided by Frangi et al.'s vesselness feature and exploits it twofold: First, the vesselness response itself is used for localizing the tubular structure of interest. Second, the eigenvectors of the Hessian eigendecomposition guide our anisotropic total variation-regularized segmentation. In a simulation experiment, we demonstrate the superiority of anisotropic as compared to isotropic total variation-regularized segmentation in the presence of noise. In an experiment with magnetic resonance images of the human cervical spinal cord, we compare our automated segmentations to those of two human observers. Finally, a comparison with a dedicated state-of-the-art spinal cord segmentation framework shows that we achieve comparable to superior segmentation quality.

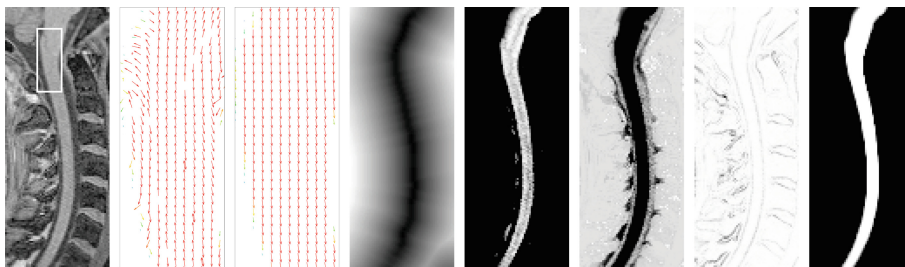
**Keywords:** Convex optimization · Anisotropic total variation · Vesselness

## 1 Introduction

Segmenting tubular structures is an important task in medical image analysis; for example, for assessing vascular diseases or tracking the progress of neurological disorders that manifest in spinal cord atrophy. Especially when used in large-scale clinical trials, largely automated segmentation is desirable to reduce the workload on clinical staff. Such automated segmentation approaches, in turn, should be robust with respect to specific choices of parameterization.

In this paper, we propose a segmentation method that fulfills both criteria: it is completely automated, and it creates segmentations of similar quality over a wide range of parameter choices. Our method adapts Yuan et al.'s continuous max flow approach [10] and combines it with an anisotropic total variation (ATV)

© Springer International Publishing AG 2016  
S. Ourselin et al. (Eds.): MICCAI 2016, Part III, LNCS 9902, pp. 362–370, 2016.  
DOI: 10.1007/978-3-319-46726-9\_42



**Fig. 1.** *Left to right:* T1 image  $I$  of the cervical spinal cord; closeups of the vessel directions  $v_1$  before GVF and  $\tilde{v}_1$  after GVF (vectors scaled by the segmentation  $u^*$  for visualization); distance map  $D$  of the vessel ridge  $\mathcal{R}$  (normalized for visualization); source capacities  $C_s$ ; sink capacities  $C_t$ ; nonterminal capacities  $C$ ; segmentation  $u^*$ .

regularization term. ATV keeps changes of the segmentation's boundary small along the course of the tubular structure. We use Frangi et al.'s well-established vesselness feature [3] as our measure of tubularity, which we exploit twofold: both for finding the location and the orientation of the structures of interest.

The directional information of vesselness, which is usually neglected, has previously been used: Manniesing et al. [6] construct an anisotropic tensor from it for accentuating vascular structures in angiography images for image enhancement. Gooya et al. [4] use this tensor in an active contour framework for blood vessel segmentation. ATV-regularized segmentation has been generally described by Olsson et al. [7] and has been used, for example, by Reinbacher et al. [8] who segment thin structures of known volume based on first-order derivatives and a volume constraint. A review of vessel segmentation is given by Lesage et al. [5]. De Leener et al. [2] review the more specific topic of spinal cord segmentation.

Our contributions lie in incorporating Hessian-based vesselness into ATV-regularized segmentation and in integrating ATV into the continuous max flow framework [10]. To the best of our knowledge, both has not been tried, so far.

## 2 Methods

In Sect. 2.1, we motivate our choice of ATV regularization. In Sect. 2.2, we state the ATV-regularized segmentation problem in the continuous max flow framework and propose an algorithm for solving it. In Sect. 2.3, we describe how we incorporate the vesselness feature. In Sect. 2.4, we present our choice of flow capacities. For a good general introduction to continuous max flow, see [10].

### 2.1 Isotropic and Anisotropic Total Variation

Segmentation on the  $d$ -dimensional image domain  $\Omega \subset \mathbb{R}^d$  can be formulated as the problem of finding a binary labeling  $u : \Omega \rightarrow \{0, 1\}$  for the given image  $I : \Omega \rightarrow \mathcal{I}$  (e.g. with  $\mathcal{I} = [0, 1]$  for a normalized single-channel image). In practice,

364 S. Pezold et al.

the problem is often relaxed such that  $u : \Omega \rightarrow [0, 1]$ , and the final labeling is determined by applying a threshold to the result of the relaxed problem [7, 10].

A common regularization term in segmentation is the *total variation* TV, which minimizes the surface area of the segmented region, penalizing jumps between segmentation foreground ( $u = 1$ ) and background ( $u = 0$ ) and thus allowing for smooth segmentations even if  $I$  is noisy (here,  $|\cdot|$  denotes the  $l_2$  norm):

$$\text{TV}[u] = \int_{\Omega} |\nabla u| \, dx. \tag{1}$$

While TV is a good regularizer for many applications, it seems not optimal in the context of tubular structure segmentation. This is because TV is isotropic; that is, changes of  $u$  are penalized regardless of orientation. If we want to segment a tube, however, we would like to employ the prior knowledge that its shape ideally does not change along its course; thus we would like to penalize changes along the tube’s direction more strongly than changes perpendicular to it. In other words, we would prefer an anisotropic regularization term.

In the proposed method, we thus use *anisotropic total variation* ATV [7, 8]:

$$\text{ATV}[u; A] = \int_{\Omega} (\nabla u^T A \nabla u)^{1/2} \, dx = \int_{\Omega} |S^T \nabla u| \, dx \quad \text{with} \quad A = S S^T, \tag{2}$$

where  $A : \Omega \rightarrow \mathbb{R}^{d \times d}$  is strongly positive definite in the sense of Olsson et al. [7] and  $S$  is a decomposition of  $A$ . For our particular choice of  $A$  and  $S$ , see Sect. 2.3.

If we assume, as a simple three-dimensional example, that  $A = \text{diag}(1, a, a)$  with  $0 < a < 1$ , we see that changes along the  $x_1$  axis will be more strongly penalized than changes along  $x_2$  and  $x_3$ . This would in fact be a meaningful choice if the tubular structure of interest was oriented along  $x_1$ . From the example we can also see that ATV is a generalization of TV, as ATV becomes TV for  $a = 1$ .

## 2.2 ATV in Continuous Max Flow

The dual formulation of the max flow problem as stated in [10], with TV replaced by ATV regularization as we propose, is given by the min cut problem

$$\min_{u \in [0, 1]} \int_{\Omega} (1 - u) C_s + u C_t + |S^T \nabla u| C \, dx, \tag{3}$$

with the source capacities  $C_s$ , sink capacities  $C_t$ , and nonterminal capacities  $C$  ( $C : \Omega \rightarrow \mathbb{R}_{\geq 0}$ ). Using integration by parts and the geometric definition of the scalar product, we can show for the nonterminal flow  $p : \Omega \rightarrow \mathbb{R}^d$  that

$$\int_{\Omega} |S^T \nabla u| C \, dx = \max_{|p| \leq C} \int_{\Omega} u \, \text{div}(Sp) \, dx. \tag{4}$$



---

**Algorithm 1.** Augmented Lagrangian-based max flow algorithm with ATV.

---

- Set bound  $\hat{\epsilon}$ , steps  $\gamma$ ,  $c$ ; calculate  $C$ ,  $C_s$ ,  $C_t$ ,  $S$ ; arbitrarily initialize  $p^0$ ,  $p_s^0$ ,  $p_t^0$ ,  $u^0$ .
- Starting from  $n = 0$ , iterate until  $\frac{1}{|\Omega|} \int_{\Omega} |\epsilon^{n+1}(x)| dx < \hat{\epsilon}$ :

$$\begin{cases} \tilde{p}^{n+1} = p^n + \gamma S^T \nabla \left( \operatorname{div}(Sp^n) - p_s^n + p_t^n - \frac{u^n}{c} \right) \\ p^{n+1} = \frac{\tilde{p}^{n+1}}{|\tilde{p}^{n+1}|} \min \{ |\tilde{p}^{n+1}|, C \} & \text{if } \tilde{p}^{n+1} \neq 0 \\ & \text{else } 0 \\ p_s^{n+1} = \min \left\{ \left( \frac{1-u^n}{c} + \operatorname{div}(Sp^{n+1}) + p_t^n \right), C_s \right\} \\ p_t^{n+1} = \min \left\{ \left( \frac{u^n}{c} - \operatorname{div}(Sp^{n+1}) + p_s^{n+1} \right), C_t \right\} \\ \epsilon^{n+1} = c \left( \operatorname{div}(Sp^{n+1}) - p_s^{n+1} + p_t^{n+1} \right) \\ u^{n+1} = u^n - \epsilon^{n+1}. \end{cases}$$


---

Together with the respective equalities for  $C_s, C_t$  and the source and sink flows  $p_s, p_t : \Omega \rightarrow \mathbb{R}$  (see Eqs. (18) and (19) in [10]), we derive the primal-dual formulation as  $\max_{p_s, p_t, p} \min_{u \in [0,1]} E[p_s, p_t, p, u]$  with

$$E = \int_{\Omega} (1-u) p_s + u p_t + u \operatorname{div}(Sp) dx = \int_{\Omega} p_s + u (\operatorname{div}(Sp) - p_s + p_t) dx, \quad (5)$$

subject to the flow capacity constraints  $p_s \leq C_s$ ,  $p_t \leq C_t$ , and  $|p| \leq C$ .

Making use of the anisotropic coarea formula in [7], it can be shown that any  $u^\ell$  for a threshold  $\ell \in (0, 1)$ , given by

$$u^\ell(x) = \begin{cases} 1, & u^*(x) > \ell \\ 0, & u^*(x) \leq \ell \end{cases} \quad \text{with } u^* = \arg \min_{u \in [0,1]} E, \quad (6)$$

is a globally optimal solution for the binary problem corresponding to Eq. (3).

Following [10], we add an augmented Lagrangian term to Eq. (5), gaining

$$\max_{p_s, p_t, p} \min_{u \in [0,1]} \int_{\Omega} p_s + u (\operatorname{div}(Sp) - p_s + p_t) - \frac{c}{2} (\operatorname{div}(Sp) - p_s + p_t)^2 dx \quad (7)$$

as the final problem, which we propose to solve with Algorithm 1.

### 2.3 ATV Regularization with Vesselness

Frangi et al. [3] examine the Hessian matrices; that is, the second-order derivatives, in the scale space of the volumetric image  $I$  to calculate what they call vesselness. The idea is to determine from the ratios of the Hessians' eigenvalues how closely the local structure in  $I$  resembles a tube.

In particular, let  $\mathcal{S}$  be a predefined set of scales that roughly match the expected tube radii. For each scale  $s \in \mathcal{S}$ , let  $H_s(x)$  denote its Hessian approximation in  $x$ , calculated by convolving  $I$  with Gaussian derivatives of standard deviation  $s$ . Let  $\lambda_{i,s}(x)$  ( $i=1, 2, 3$ ) denote the sorted eigenvalues

366 S. Pezold et al.

( $|\lambda_{1,s}| \leq |\lambda_{2,s}| \leq |\lambda_{3,s}|$ ) and  $v_{i,s}(x)$  corresponding eigenvectors of  $H_s$ , such that

$$H_s = V_s \Lambda_s V_s^T \quad \text{with} \quad V_s = [v_{1,s}|v_{2,s}|v_{3,s}], \quad \Lambda_s = \text{diag}(\lambda_{1,s}, \lambda_{2,s}, \lambda_{3,s}). \quad (8)$$

Note that  $V_s^T = V_s^{-1}$ , as  $H_s$  is symmetric. Assuming bright tubular structures on dark background, the vesselness response is  $\nu(x) = \max_{s \in \mathcal{S}} \nu_s(x)$ , where

$$\nu_s = \begin{cases} 0, & \lambda_{2,s} \geq 0 \text{ or } \lambda_{3,s} \geq 0 \\ \left(1 - \exp\left(\frac{-1}{2w_1^2} \frac{\lambda_{2,s}^2}{\lambda_{3,s}^2}\right)\right) \exp\left(\frac{-1}{2w_2^2} \frac{\lambda_{1,s}^2}{\lambda_{2,s}\lambda_{3,s}}\right) \left(1 - \exp\left(\frac{-\sum_i \lambda_{i,s}^2}{2w_3^2}\right)\right), & \text{else,} \end{cases} \quad (9)$$

with the weighting factors  $w_i \in \mathbb{R}_{>0}$ . The eigenvectors for  $\nu$  are  $V = [v_1|v_2|v_3]$ , with  $v_i = v_{i,s^*}$  and  $s^* = \arg \max_{s \in \mathcal{S}} \nu_s$ . In the original description of [3], no use of  $V$  is made. In our approach, we use the eigenvectors to steer the ATV regularizer. We observe that in points where  $\nu$  is high,  $v_1$  points along the local vessel orientation [6]. Recall that we want to regularize strongly along the direction of the vessel. Unfortunately, we cannot use  $v_1$  directly for this purpose, as it reliably gives the vessel's direction in the vessel center only, where  $\nu$  is the highest. Therefore, we use the concept of gradient vector flow (GVF) [9] to first propagate the directions from places where  $\nu$  is high to regions where  $\nu$  is low, creating a smoothly varying vector field. The necessary steps are as follows.

Let  $\mathcal{R}$  be the set of vesselness ridge points; that is, the local maxima of  $\nu$ , down to a noise threshold. As both  $-v_1$  and  $v_1$  are valid eigenvectors, we have to make sure that the vectors of neighboring points approximately point in the same rather than the opposite direction, so that they don't cancel each other out when diffusing them via GVF. Thus, we fix their signs beforehand, gaining  $\bar{v}_1$ : We calculate the minimum spanning tree over the ridge points  $\mathcal{R}$ , select a root point, keep its sign, and traverse the tree. For each child point, we choose  $\bar{v}_1$  as either  $-v_1$  or  $v_1$ , depending on which one maximizes the dot product (i.e. minimizes the angle) with its parent's  $\bar{v}_1$ . After traversal, the signs of the  $v_1$  for all remaining domain points  $x \in \Omega \setminus \mathcal{R}$  are fixed w.r.t. their closest point in  $\mathcal{R}$ , following the same rule. We scale all  $\bar{v}_1$  with  $\nu$ , apply GVF, and scale the resulting vectors back to unit length, gaining  $\tilde{v}_1$ . A comparison of the vector field before and after sign adjustment and GVF is shown in Fig. 1.

Finally, we recomplete  $\tilde{v}_1$  to an orthonormal basis  $\tilde{V} = [\tilde{v}_1|\tilde{v}_2|\tilde{v}_3]$ . The particular choice of  $\tilde{v}_2, \tilde{v}_3$  does not matter, as we will treat all directions perpendicular to  $\tilde{v}_1$  the same when regularizing. From  $\tilde{V}$ , we construct  $A, S$  for Eq. (2) as

$$A = \tilde{V} \tilde{A} \tilde{V}^T \quad \text{and} \quad S = \tilde{V} \tilde{A}^{1/2} \quad \text{with} \quad \tilde{A} = \text{diag}(1, a, a) \quad \text{and} \quad 0 < a \leq 1. \quad (10)$$

Notice the similarity to  $A$  in the example at the end of Sect. 2.1: The idea of regularizing one direction stronger than the others remains the same; however, as we now scale with  $\tilde{A} = \text{diag}(1, a, a)$  in the new basis  $\tilde{V}$ , we target the actual local vessel direction  $\tilde{v}_1$  rather than a fixed axis.

## 2.4 Flow Capacities

For the source and sink capacities  $C_s, C_t$  of Eqs. (3) and (5), we use a combination of the normalized image intensities and the distances to the vessel ridge points  $\mathcal{R}$ . Intuitively, using the intensities enables the distinction of foreground (i.e. the vessel) and background (i.e. everything else), while the distances w.r.t.  $\mathcal{R}$  isolate the vessel surrounding. This helps avoiding oversegmentations in case other structures have intensities similar to those of the vessel of interest. More formally, let  $D : \Omega \rightarrow \mathcal{D} = \mathbb{R}_{\geq 0}$  be a Euclidean distance map of  $\mathcal{R}$ , and let  $I : \Omega \rightarrow \mathcal{I} = [0, 1]$  be the normalized image. Let  $p_b^D, p_f^D, p_b^I, p_f^I$  be predefined estimates of the background and foreground probability densities for  $\mathcal{D}$  and  $\mathcal{I}$ , with  $p_b^D : \mathcal{D} \rightarrow \mathbb{R}_{\geq 0}$  and  $p_f^I : \mathcal{I} \rightarrow \mathbb{R}_{\geq 0}$ . We calculate  $C_s, C_t$  as

$$C_s(x) = \frac{1}{q} \max\{r(x), 0\}, \quad \text{with } r(x) = \ln \left( \frac{p_f^D(D(x)) \cdot p_f^I(I(x)) + \varepsilon}{p_b^D(D(x)) \cdot p_b^I(I(x)) + \varepsilon} \right), \quad (11)$$

$$C_t(x) = \frac{1}{q} \max\{-r(x), 0\}, \quad q = \ln \left( \frac{\max\{\hat{p}_b^D \cdot \hat{p}_b^I, \hat{p}_f^D \cdot \hat{p}_f^I\} + \varepsilon}{\varepsilon} \right), \quad (12)$$

where  $\hat{p} := \max p$ . The small positive constant  $\varepsilon$  avoids zero logarithms and zero divisions in  $r$ . Normalization with  $q$  ensures that  $C_s, C_t \in [0, 1]$ , which eases their balancing with the nonterminal capacities  $C$  of Eqs. (3) and (5).

Using  $C$ , we try to move the segmentation boundary to image edges by making  $C$  small where the intensity gradient magnitude is high and vice versa:

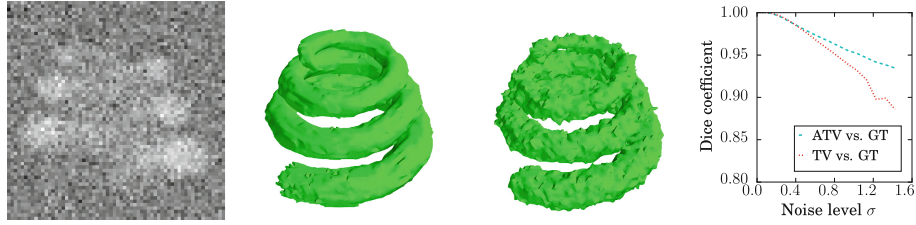
$$C(x) = w \exp \left( -1/\varsigma^2 |\nabla I(x)|^2 \right) \quad \text{with } w, \varsigma \in \mathbb{R}_{>0}, \quad (13)$$

where  $\varsigma$  controls  $C$ 's sensitivity regarding the size of  $|\nabla I|$  and  $w$  balances  $C$  and  $C_s, C_t$ . For an example of the capacities, see Fig. 1.

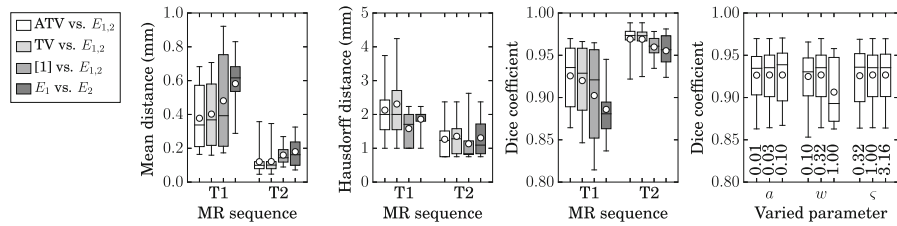
## 3 Experiments and Results

*Implementation:* Most of the method's steps can be described as *embarrassingly parallel*, which means they can be calculated independently for different voxels. This is true, for example, for the vesselness  $\nu$ , GVF, the capacities  $C_s, C_t, C$ , and large parts of Algorithm 1. For this reason, we ported code to the GPU wherever possible. To reduce memory consumption, which is still a limiting factor for GPU programming, we represent the pointwise basis matrices  $\tilde{V}$  in Eq. (10) as equivalent unit quaternions. Derivatives are approximated using forward ( $\nabla$ , Algorithm 1), backward (div, Algorithm 1), and central ( $\nabla$ , Eq. 13) differences. *Parameterization:* For all experiments, the following parameters were chosen. Algorithm 1:  $\hat{\varepsilon} = 10^{-6}$ ,  $\gamma = 0.11\Delta x^2$ ,  $c = 0.2\Delta x^2$  ( $\Delta x$ : minimum voxel edge length in mm); GVF: 316 iterations with a regularization parameter of  $\mu = 3.16$  and step size determined as defined and described in [9]; Eq. (6):  $\ell = 0.5$ ; Eq. (9):  $w_1 = w_2 = 0.5$  as suggested in [3],  $w_3$  determined following [3]; Eq. (12):  $\varepsilon = 10^{-9}$ .

368 S. Pezold et al.



**Fig. 2.** Phantom experiment. *Left to right:* Mean intensity projection at  $\sigma = 1.5$ ; segmentation with ATV; segmentation with TV; Dice coefficients w.r.t. ground truth (GT) for all noise levels  $\sigma$ .



**Fig. 3.** Spinal cord experiment. *Left to right:* Mean surface distances for ATV/TV/PropSeg [1] vs.  $E_1$  and  $E_2$ ; corresponding Hausdorff distances; corresponding Dice coefficients; T1 Dice coefficients for ATV vs.  $E_1$  and  $E_2$  with varying parameter values (central bar: median, circle: mean, box limits: 25/75th percentile, whiskers: extrema).

**Helical Phantom.** In this simulation experiment, we rendered images of a synthetic helical phantom with values in  $[0, 1]$  to which we added Gaussian noise of standard deviation  $\sigma$  (Fig. 2). The phantom's tube radius varied between 3 mm and 6 mm, so we set  $\mathcal{S} = [2.5 \text{ mm}, 7.2 \text{ mm}]$  (16 scales) in Eq. (9). We segmented the images with TV and ATV regularization, setting  $a = 1$  and  $a = 0.03$  in Eq. (10), respectively. We modeled  $p^f$  as normal distributions, using the true background, foreground, and noise level values. For the sake of simplicity, we set  $p^D = 1$  always. For  $w$  and  $\zeta$  in Eq. (13), we made grid searches for each noise level, using the Dice coefficients w.r.t. the ground truth as optimization criterion.

Figure 2 shows the Dice coefficients for the best  $w, \zeta$  combinations. The advantage of ATV becomes apparent as soon as the noise level increases.

**Spinal Cord.** In this experiment with real data, two clinical experts ( $E_1, E_2$ ) manually segmented 10 MR scans (5 T1, 5 T2) of the healthy human cervical spinal cord over the C1–C3 region. For each image, we then used the remaining four images of the same sequence (T1/T2) to estimate  $p_i$  and to find an optimal parameter combination for  $a, w, \zeta$ . The distributions  $p_i$  were estimated from the manual labelings of the remaining four images, modeling  $p_b$  as mixtures of four Gaussians and  $p_f$  as normal distributions. We set  $\mathcal{S} = [2 \text{ mm}, 4 \text{ mm}]$  (16 scales) in Eq. (9). The parameters  $a, w, \zeta$  were optimized via grid search, using the mean

Dice coefficients of the remaining four images w.r.t. their manual segmentations as optimization criterion. These distributions and the determined optimum parameterization were then used to segment the left-out image for evaluation of the method. For comparison with the state of the art, we also segmented all images with *PropSeg* [1].

Figure 3 shows the averaged mean surface distances, Hausdorff distances, and Dice coefficients w.r.t. their manual segmentations. Especially the T1 images profit from ATV, as they have both a lower resolution (T1:  $1 \times 1 \times 1 \text{ mm}^3$ , T2:  $0.75 \times 0.38 \times 0.38 \text{ mm}^3$ ) and a lower contrast-to-noise ratio (about one eighth) than the T2 images. On the right, Fig. 3 shows the Dice coefficients for applying a wide range of parameter values to the T1 images, demonstrating the robustness of our method w.r.t. parameterization. For each varied parameter, the others were kept constant ( $a = 0.03$ ,  $w = 0.32$ ,  $\zeta = 1.00$ ).

#### 4 Discussion and Conclusion

We presented a fully automated method for the segmentation of tubular structures. For a single image of  $256^3$  voxels, the complete process of calculating the vesselness, GVF, capacities, and segmentation takes about 1 to 1.5 min (GPU: Nvidia GeForce GTX 770). Although image segmentation in general and tubular structure segmentation in particular have often been addressed, the results of comparison with a state-of-the-art approach lead us to believe that our method may be of value to the scientific community. Future experiments will have to show in more detail how strong is the dependence of the segmentation quality on the outcome of the Frangi vesselness response and what is the influence of a particular GVF parameterization in this context. Furthermore, the use of alternative vesselness indicators will have to be considered. We provide our reference implementation at <https://github.com/spezold/miccai2016>.

#### References

1. De Leener, B., Kadoury, S., Cohen-Adad, J.: Robust, accurate and fast automatic segmentation of the spinal cord. *NeuroImage* **98**, 528–536 (2014)
2. De Leener, B., Taso, M., Cohen-Adad, J., Callot, V.: Segmentation of the human spinal cord. *Magn. Reson. Mater. Phys., Biol. Med.* **29**(2), 125–153 (2016)
3. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: Wells, W.M., Colchester, A.C.F., Delp, S.L. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 130–137. Springer, Heidelberg (1998)
4. Gooya, A., Liao, H., Sakuma, I.: Generalization of geometrical flux maximizing flow on Riemannian manifolds for improved volumetric blood vessel segmentation. *Comput. Med. Imaging Graph.* **36**(6), 474–483 (2012)
5. Lesage, D., Angelini, E.D., Bloch, I., Funka-Lea, G.: A review of 3D vessel lumen segmentation techniques: models, features and extraction schemes. *Med. Image Anal.* **13**(6), 819–845 (2009)
6. Manniesing, R., Viergever, M.A., Niessen, W.J.: Vessel enhancing diffusion: a scale space representation of vessel structures. *Med. Image Anal.* **10**(6), 815–825 (2006)

370 S. Pezold et al.

7. Olsson, C., Byrod, M., Overgaard, N., Kahl, F.: Extending continuous cuts: anisotropic metrics and expansion moves. In: 2009 IEEE 12th International Conference on Computer Vision, pp. 405–412 (2009)
8. Reinbacher, C., Pock, T., Bauer, C., Bischof, H.: Variational segmentation of elongated volumetric structures. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3177–3184 (2010)
9. Xu, C., Prince, J.L.: Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Process.* **7**(3), 359–369 (1998)
10. Yuan, J., Bae, E., Tai, X.C.: A study on continuous max-flow and min-cut approaches. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2217–2224 (2010)

## 8.A Derivation of the Nonterminal Flow Equality

Here we show for the nonterminal flow variable  $p : \Omega \rightarrow \mathbb{R}^d$  that the following equality holds:

$$\int_{\Omega} |S^T \nabla u| C \, dx = \max_{|p(x)| \leq C(x)} \int_{\Omega} u \operatorname{div}(S p) \, dx, \quad (8.1)$$

where  $S$  is defined as above.

*Derivation:* In the first step, we show that

$$\int_{\Omega} |S^T \nabla v| C \, dx = \max_{|q(x)| \leq C(x)} \int_{\Omega} S^T \nabla v \cdot q \, dx, \quad (8.2)$$

with  $\nabla v, q \in \mathbb{R}^d$ . The cases in which  $|S(x)^T \nabla v(x)| = 0$  or  $C(x) = 0$  are similar to the case described for the first summand in Section 3.5. Equality for  $|S(x)^T \nabla v(x)| > 0$  and  $C(x) > 0$  can be shown by a geometric argument: We let  $w = S^T \nabla v$  and define  $\mathcal{Q}_c = \{q \in \mathbb{R}^d \mid |q| = c\}$ , that is, the set of  $d$ -dimensional vectors of length  $c \in \mathbb{R}_{>0}$ . The product  $w \cdot q$  with  $q \in \mathcal{Q}_c$  is maximized for a given  $w$  by  $q = c w/|w|$ , that is, by the feasible vector that points in the same direction as  $w$ . This immediately follows from the geometric definition of the dot product,

$$w \cdot q := |w| |q| \cos \alpha, \quad (8.3)$$

where  $\alpha$  is the angle between  $w$  and  $q$ : The product  $|w| |q|$  remains constant for all feasible  $q$  by definition of  $\mathcal{Q}_c$ , while  $\cos \alpha$  is maximized for  $\alpha = 0$  with  $\cos 0 = 1$ . We now define the set  $\mathcal{Q}_{\hat{c}, w} = \{c w/|w| \in \mathbb{R}^d \mid 0 < c \leq \hat{c}\}$ , that is, the set of vectors pointing in the same direction as  $w$ , with the length bounded by  $\hat{c} \in \mathbb{R}_{>0}$ . The product  $w \cdot q$  with  $q \in \mathcal{Q}_{\hat{c}, w}$ , is maximized by  $q = \hat{c} w/|w|$ , that is, by the longest feasible vector. This can again be seen from the definition of the dot product: Now the angle remains constant with  $\alpha = 0$ , therefore  $\cos \alpha = 1$ , and the product  $|w| |c w/|w|| = c |w| |w/|w||$  is maximized by  $c = \hat{c}$ . In other words, now we employ the linearity of the dot product with respect to the length of  $q$ . As a consequence,

$$\max_{|q(x)| \leq C(x)} \int_{\Omega} w \cdot q \, dx = \int_{\Omega} w \cdot \left( C \frac{w}{|w|} \right) dx = \int_{\Omega} \frac{w \cdot w}{|w|} C \, dx = \int_{\Omega} |w| C \, dx \quad (8.4)$$

$$\Leftrightarrow \max_{|q(x)| \leq C(x)} \int_{\Omega} S^T \nabla v \cdot q \, dx = \int_{\Omega} |S^T \nabla v| C \, dx, \quad (8.5)$$

where we leverage the fact that  $w \cdot w = |w|^2$  in the last term of Eq. (8.4).

In the second step, we rearrange:

$$\int_{\Omega} S^T \nabla v \cdot q \, dx = \int_{\Omega} (S^T \nabla v)^T q \, dx = \int_{\Omega} \nabla v^T S q \, dx = \int_{\Omega} \nabla v \cdot S q \, dx. \quad (8.6)$$

In the third step, we employ the equality

$$\int_{\Omega} \nabla v \cdot S q \, dx = - \int_{\Omega} v \operatorname{div}(S q) \, dx, \quad (8.7)$$

which holds if either  $v = 0$  or  $q = 0$  on the boundary of  $\Omega$ , and which can be shown using integration by parts.

Combining the steps, we can thus write

$$\int_{\Omega} |S^T \nabla v| C \, dx = \max_p \int_{\Omega} S^T \nabla v \cdot p \, dx = \max_p \int_{\Omega} \nabla v \cdot S p \, dx = \max_p \int_{\Omega} -v \operatorname{div}(S p) \, dx, \quad (8.8)$$

subject to  $|p(x)| \leq C(x)$ .

Finally, if we substitute  $u = -v$ , constrain  $u \in [0, 1] \subset \mathbb{R}$ , and realize that  $|\nabla(-u)| = |\nabla u|$ , we get the equality for the nonterminal flow, as stated in Eq. (8.1).

## 8.B Derivation of an Update Rule for the Flow $p$

**Derivation Using Euler–Lagrange Equation.** Here we derive an update rule based on gradient ascent for the nonterminal flow variable  $p$  with anisotropic total variation regularization in two dimensions. A generalization to more dimensions is straightforward and leads to the same final result in matrix/vector notation. For the sake of notational clarity, we use the shorthand  $\partial_{\beta} \alpha$  to denote the partial derivative  $\partial \alpha / \partial \beta$  throughout this section. The three-dimensional version of the result is used in the algorithm proposed above.

*Derivation:* The augmented Lagrangian-based segmentation energy term  $E$  that is to be maximized in  $p$  is given by

$$E = \int_{\Omega} L \, dx, \quad (8.9)$$

with

$$L = p_s + u (\operatorname{div}(S p) - p_s + p_t) - \frac{c}{2} (\operatorname{div}(S p) - p_s + p_t)^2, \quad (8.10)$$

where

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad S = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}, \quad \text{and} \quad p = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}.$$

We can write out  $\operatorname{div}(S p)$  in Eq. (8.10) as

$$\operatorname{div}(S p) = \partial_{x_1} (s_{11} p_1 + s_{12} p_2) + \partial_{x_2} (s_{21} p_1 + s_{22} p_2) \quad (8.11)$$

$$= \partial_{x_1} (s_{11} p_1) + \partial_{x_1} (s_{12} p_2) + \partial_{x_2} (s_{21} p_1) + \partial_{x_2} (s_{22} p_2) \quad (8.12)$$

$$= \partial_{x_1} s_{11} p_1 + s_{11} \partial_{x_1} p_1 + \partial_{x_1} s_{12} p_2 + s_{12} \partial_{x_1} p_2 + \quad (8.13)$$

$$\partial_{x_2} s_{21} p_1 + s_{21} \partial_{x_2} p_1 + \partial_{x_2} s_{22} p_2 + s_{22} \partial_{x_2} p_2$$

$$= (\partial_{x_1} s_{11} + \partial_{x_2} s_{21}) p_1 + (\partial_{x_1} s_{12} + \partial_{x_2} s_{22}) p_2 + \quad (8.14)$$

$$s_{11} \partial_{x_1} p_1 + s_{12} \partial_{x_1} p_2 + s_{21} \partial_{x_2} p_1 + s_{22} \partial_{x_2} p_2.$$

The Euler–Lagrange equations for Eq. (8.10) are given by

$$\partial_{p_i} L - \sum_j \partial_{x_j} (\partial_{\partial_{x_j} p_i} L) \stackrel{!}{=} 0 \quad \text{with} \quad i, j = 1, 2. \quad (8.15)$$



It can be seen that each component  $\partial_{\psi} L$  in Eq. (8.15) is of the form

$$\partial_{\psi} L = \partial_{\psi} \operatorname{div}(S p) m \quad \text{with} \quad m = u - c (\operatorname{div}(S p) - p_s + p_t). \quad (8.16)$$

We can directly get  $\partial_{\psi} \operatorname{div}(S p)$  from Eq. (8.14), for example:

$$\partial_{\partial_{x_1} p_1} \operatorname{div}(S p) = s_{11} \quad \Leftrightarrow \quad \partial_{\partial_{x_1} p_1} L = s_{11} m = s_{11} (u - c (\operatorname{div}(S p) - p_s + p_t)). \quad (8.17)$$

As a result, we can write out the two Euler-Lagrange equations ( $i = 1, 2$ ) as

$$(\partial_{x_1} s_{11} + \partial_{x_2} s_{21}) m - (\partial_{x_1} (s_{11} m) + \partial_{x_2} (s_{21} m)) = 0, \quad (8.18)$$

$$(\partial_{x_1} s_{12} + \partial_{x_2} s_{22}) m - (\partial_{x_1} (s_{12} m) + \partial_{x_2} (s_{22} m)) = 0. \quad (8.19)$$

Applying the product rule to the  $\partial_{x_j} (s_i m)$  terms, we can simplify to

$$-(s_{11} \partial_{x_1} m + s_{21} \partial_{x_2} m) = 0, \quad (8.20)$$

$$-(s_{12} \partial_{x_1} m + s_{22} \partial_{x_2} m) = 0, \quad (8.21)$$

which we can combine in matrix/vector notation to

$$-\begin{bmatrix} s_{11} & s_{21} \\ s_{12} & s_{22} \end{bmatrix} \begin{pmatrix} \partial_{x_1} m \\ \partial_{x_2} m \end{pmatrix} = 0 \quad \Leftrightarrow \quad S^T \nabla_x (c (\operatorname{div}(S p) - p_s + p_t) - u) = 0. \quad (8.22)$$

In order to maximize  $p$ , we introduce a time variable  $\tau$  and its time step  $\Delta\tau$ , such that

$$\frac{p^{\tau+\Delta\tau} - p^{\tau}}{\Delta\tau} \approx c S^T \nabla_x \left( \operatorname{div}(S p) - p_s + p_t - \frac{u}{c} \right), \quad (8.23)$$

and we end up with the update rule

$$p^{n+1} = p^n + \gamma S^T \nabla_x \left( \operatorname{div}(S p^n) - p_s + p_t - \frac{u}{c} \right), \quad (8.24)$$

where  $n$  is the iteration count and  $\gamma = c \Delta\tau$  is the step size.

**Alternative Derivation.** A more elegant derivation of the update rule for  $p$  in  $d$  dimensions is achieved by following a common way to derive the Euler-Lagrange equation. We thank Antal Horváth for providing this alternative.

*Derivation:* Let  $E$  and  $L$  be defined as above, that is

$$E = \int_{\Omega} L \, dx, \quad (8.25)$$

$$L = L(x, p(x), \operatorname{div}(S p)(x)), \quad (8.26)$$

with

$$L = p_s + u (\operatorname{div}(S p) - p_s + p_t) - \frac{c}{2} (\operatorname{div}(S p) - p_s + p_t)^2. \quad (8.27)$$

We now define  $p_\varepsilon : \Omega \rightarrow \mathbb{R}^d$  as the result of a perturbation of  $p$ :

$$p_\varepsilon(x) := p(x) + \varepsilon \eta(x), \quad (8.28)$$

where  $\varepsilon \in \mathbb{R}$  is a weighting factor and the perturbation  $\eta : \Omega \rightarrow \mathbb{R}^d$  with  $\eta \in C_0^\infty(\Omega)$  is a differentiable function that is zero on the boundary of the domain  $\Omega$ .

Making use of  $p_\varepsilon$ , we define  $E_\varepsilon$  and  $L_\varepsilon$  as

$$E_\varepsilon := \int_{\Omega} L_\varepsilon \, dx, \quad (8.29)$$

$$L_\varepsilon := L(x, p_\varepsilon(x), \operatorname{div}(S p_\varepsilon)(x)). \quad (8.30)$$

Calculating  $dE_\varepsilon/d\varepsilon$ , that is,  $E_\varepsilon$ 's total derivative with respect to  $\varepsilon$ , we yield

$$\frac{dE_\varepsilon}{d\varepsilon} = \frac{d}{d\varepsilon} \left( \int_{\Omega} L_\varepsilon \, dx \right) = \int_{\Omega} \frac{dL_\varepsilon}{d\varepsilon} \, dx, \quad (8.31)$$

with

$$\frac{dL_\varepsilon}{d\varepsilon} = \frac{\partial L_\varepsilon}{\partial x} \cdot \frac{dx}{d\varepsilon} + \frac{\partial L_\varepsilon}{\partial p_\varepsilon} \cdot \frac{dp_\varepsilon}{d\varepsilon} + \frac{\partial L_\varepsilon}{\partial \operatorname{div}(S p_\varepsilon)} \frac{d \operatorname{div}(S p_\varepsilon)}{d\varepsilon} \quad (8.32)$$

$$= \frac{\partial L_\varepsilon}{\partial p_\varepsilon} \cdot \eta + \frac{\partial L_\varepsilon}{\partial \operatorname{div}(S p_\varepsilon)} \operatorname{div}(S \eta). \quad (8.33)$$

If the current  $p$  maximizes  $L$ , then any perturbation with  $\varepsilon \eta$  will be non-increasing in  $E_\varepsilon$ . If in this case  $\varepsilon = 0$ , then  $p_\varepsilon = p$ ,  $L_\varepsilon = L$ , and  $E_\varepsilon$  has a maximum value. In combination we thus have

$$\left. \frac{dE_\varepsilon}{d\varepsilon} \right|_{\varepsilon=0} = \int_{\Omega} \left[ \frac{\partial L}{\partial p} \cdot \eta + \frac{\partial L}{\partial \operatorname{div}(S p)} \operatorname{div}(S \eta) \right] dx \stackrel{!}{=} 0. \quad (8.34)$$

Using integration by parts on the second summand, we get

$$\int_{\Omega} \frac{\partial L}{\partial \operatorname{div}(S p)} \operatorname{div}(S \eta) \, dx = - \int_{\Omega} S \eta \cdot \nabla_x \left( \frac{\partial L}{\partial \operatorname{div}(S p)} \right) dx, \quad (8.35)$$

where we recall that  $\eta$  is zero on  $\partial\Omega$ . Equation (8.34) thus becomes

$$\int_{\Omega} \left[ \frac{\partial L}{\partial p} \cdot \eta - S \eta \cdot \nabla_x \left( \frac{\partial L}{\partial \operatorname{div}(S p)} \right) \right] dx = \int_{\Omega} \left( \frac{\partial L}{\partial p} - S^\top \nabla_x \left( \frac{\partial L}{\partial \operatorname{div}(S p)} \right) \right) \cdot \eta \, dx \stackrel{!}{=} 0. \quad (8.36)$$

From the fundamental lemma of the calculus of variations, it follows that

$$\frac{\partial L}{\partial p} - S^\top \nabla_x \left( \frac{\partial L}{\partial \operatorname{div}(S p)} \right) \stackrel{!}{=} 0, \quad (8.37)$$

where  $\partial L / \partial p = 0$ , as  $p$  maximizes  $L$ .

## 8.B Derivation of an Update Rule for the Flow $p$

After inserting Eq. (8.27) into Eq. (8.37), we get

$$-S^T \nabla_x (u - c (\operatorname{div}(Sp) - p_s + p_t)) = 0 \quad (8.38)$$

$$\Leftrightarrow c S^T \nabla_x \left( \operatorname{div}(Sp) - p_s + p_t - \frac{u}{c} \right) = 0, \quad (8.39)$$

which again, compare Eqs. (8.22)–(8.24), motivates the update rule

$$p^{n+1} = p^n + \gamma S^T \nabla_x \left( \operatorname{div}(Sp^n) - p_s + p_t - \frac{u}{c} \right), \quad (8.40)$$

where  $n$  is the iteration count and  $\gamma$  is the step size.



## 9 Discussion and Conclusion

It was our goal to develop software tools for our clinical partners for reliably measuring the spinal cord volume and quantifying neural atrophy. To achieve this goal, we developed and adjusted a toolchain to segment the spinal cord, reconstruct its surface, handle image distortions, and acquire measurements relative to an anatomical landmark. In addition to volume measurements, the toolchain enables cross-sectional area measurements. The user interaction is limited to providing the landmark location as well as a small number of initial hints that label voxels of the spinal cord and its surroundings.

In our first version (Chapter 5), we relied on graph cut-based segmentation. For calculating the segmentation cost functions, we relied on the image intensities only. An evaluation on healthy subjects and MS patients indicated that our measurement quality was on par with previously published approaches. Nevertheless, this initial version of our toolchain had some limitations: Due to evaluating solely the image intensities in the segmentation, it worked reliably on T1-weighted MR images only. Furthermore, as graph cut segmentation is by design bound to the image resolution, a second step of segmentation refinement was necessary to get a surface reconstruction that was fine enough for exact measurements.

To overcome the first issue, we integrated two Hessian-based features into our segmentation approach: a vesselness feature that responded to the spinal cord itself, and a customized csfness feature that responded to the CSF, which immediately surrounds the spinal cord in the vertebral canal. Furthermore, we switched from graph cut-based segmentation [14] to a continuous max flow-based algorithm [47]. In this new segmentation algorithm, we integrated a cross-sectional similarity prior that allowed only gradual changes along a predefined image axis and thus helped to guide the segmentation in noisy or low-contrast image regions. As a result from these adjustments, the second version of our toolchain (Chapters 6 and 7) could now handle both T1- and T2-weighted MR images, could work on larger images, and provided an overall more robust segmentation outcome. As a further result of replacing the segmentation algorithm, we were able to port large parts of the toolchain to the GPU, which gained us noticeable improvements in execution time.

In the third version (Chapter 8), rather than predefining an axis for the cross-sectional similarity prior, we leveraged the directional information from the vesselness feature to regularize along the spinal cord's direction itself, thus evading the need for predefining an axis and allowing for the handling of more curvy tubular structures, as we demonstrated on a helical phantom. Furthermore, we could now omit the previously required surface refinement step, reducing the number of possible points of failure in the toolchain. Again we could show that the quality of our segmentation results was competitive with other published approaches.

We deployed our software toolchain with our clinical partners, and it has already been used for larger-scale MS patient trials, analyzing more than 1000 follow-up scans of more than 200 patients acquired over a seven-year range. The results, which have yet to be published,

are well in line with the MS literature, showing different baseline spinal cord volumes and different degrees of atrophy that are characteristic for the respective subtypes of MS. Thus, in addition to reaching comparable performance to the state of the art, our approach proved its practical usability for large-scale analyses.

**Future Research.** Despite demonstrating that at present our approach is already well-suited for application in clinical research, we still see numerous opportunities for advancing it.

Regarding the degree of automation, we are confident that it is possible to make our toolchain fully automatic. As other groups have already shown, for example De Leener et al. [12], and as we have demonstrated by proof of concept (Chapters 6 and 8), completely automated spinal cord segmentation with reliable results is possible. We note that by now we have numerous segmentations and anatomical landmark annotations at our disposal that were acquired with our toolchain. These data could be used as training samples for integrating machine learning techniques for automating the landmark localization and for extending or perhaps even replacing the segmentation algorithm. Nevertheless, we think that a certain degree of interactivity should always be possible, enabling an expert to adjust or overrule the software tool's results.

As to the acquired measurements, our approach currently segments the whole spinal cord, while research interest also lies in segmenting its inner gray and white matter structures. This part of future research is already being approached in another project of our group, which can be seen as a continuation of the presented work. Potentially, even further localization of atrophy measurements down to pointwise locations is possible, as has been indicated in recent work of Rocca et al. [41] and Taso et al. [44].

**Conclusion.** In conclusion, we are confident that we provided our clinical partners with the necessary tools for quantifying atrophy in the spinal cord comfortably and reliably. In this way, we added our share to the ongoing research in MS, which may ultimately help to understand the disease better. To make our suggested approaches easier accessible for the scientific community, we are currently preparing *cordial*, the *cord image analyzer*, to be released as open source software<sup>1</sup>.

---

<sup>1</sup><https://github.com/spezold/cordial> (last accessed on November 3, 2016)

## Bibliography

- [1] Amann, M., Pezold, S., Naegelin, Y., Fundana, K., Andělová, M., Weier, K., Stippich, C., Kappos, L., Radue, E.W., Cattin, P., Sprenger, T.: Reliable volumetry of the cervical spinal cord in MS patient follow-up data with cord image analyzer (Cordial). *Journal of Neurology* 263(7), 1364–1374 (May 2016) [55](#)
- [2] Ascherio, A., Munger, K.L.: Epstein–Barr Virus Infection and Multiple Sclerosis: A Review. *Journal of Neuroimmune Pharmacology* 5(3), 271–277 (Apr 2010) [5](#)
- [3] Barkhof, F.: The clinico-radiological paradox in multiple sclerosis revisited. *Current Opinion in Neurology* 15(3), 239–245 (Jun 2002) [1](#), [6](#)
- [4] Boykov, Y., Kolmogorov, V.: Computing geodesics and minimal surfaces via graph cuts. In: Ninth IEEE International Conference on Computer Vision, 2003. Proceedings. vol. 1, pp. 26–33 (Oct 2003) [9](#), [18](#)
- [5] Boykov, Y., Veksler, O.: Graph Cuts in Vision and Graphics: Theories and Applications. In: Paragios, N., Chen, Y., Faugeras, O. (eds.) *Handbook of Mathematical Models in Computer Vision*, pp. 79–96. Springer US (2006) [9](#), [15](#)
- [6] Boykov, Y., Veksler, O., Zabih, R.: Markov random fields with efficient approximations. In: 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1998. Proceedings. pp. 648–655 (Jun 1998) [13](#)
- [7] Boykov, Y.Y., Jolly, M.P.: Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In: Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings. vol. 1, pp. 105–112 (2001) [9](#), [16](#)
- [8] Bresson, X., Esedoğlu, S., Vandergheynst, P., Thiran, J.P., Osher, S.: Fast Global Minimization of the Active Contour/Snake Model. *Journal of Mathematical Imaging and Vision* 28(2), 151–167 (Jul 2007) [9](#), [13](#), [14](#)
- [9] Chan, T., Esedoğlu, S., Nikolova, M.: Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM Journal on Applied Mathematics* 66(5), 1632–1648 (2006) [9](#), [13](#), [20](#)
- [10] Chan, T., Vese, L.: Active contours without edges. *IEEE Transactions on Image Processing* 10(2), 266–277 (2001) [9](#), [11](#), [13](#)
- [11] Compston, A., Coles, A.: Multiple sclerosis. *The Lancet* 372(9648), 1502–1517 (2008) [4](#), [5](#)

## Bibliography

- [12] De Leener, B., Cohen-Adad, J., Kadoury, S.: Automatic Segmentation of the Spinal Cord and Spinal Canal Coupled With Vertebral Labeling. *IEEE Transactions on Medical Imaging* 34(8), 1705–1718 (Aug 2015) [84](#)
- [13] De Leener, B., Taso, M., Cohen-Adad, J., Callot, V.: Segmentation of the human spinal cord. *Magnetic Resonance Materials in Physics, Biology and Medicine* 29(2), 125–153 (Jan 2016) [24](#)
- [14] Delong, A., Boykov, Y.: A Scalable graph-cut algorithm for N-D grids. In: *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008*, pp. 1–8 (Jun 2008) [2, 9, 18, 83](#)
- [15] Filippi, M., Mastronardo, G., Rocca, M.A., Pereira, C., Comi, G.: Quantitative volumetric analysis of brain magnetic resonance imaging from patients with multiple sclerosis. *Journal of the Neurological Sciences* 158(2), 148–153 (Jun 1998) [6](#)
- [16] Ford Jr., L.R., Fulkerson, D.R.: Maximal flow through a network. *Canadian Journal of Mathematics* 8, 399–404 (1956) [17, 18](#)
- [17] Geman, S., Geman, D.: Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6(6), 721–741 (Nov 1984) [17](#)
- [18] Goldberg, A.V., Tarjan, R.E.: A New Approach to the Maximum-flow Problem. *J. ACM* 35(4), 921–940 (Oct 1988) [18](#)
- [19] Grassiot, B., Desgranges, B., Eustache, F., Defer, G.: Quantification and clinical relevance of brain atrophy in multiple sclerosis: a review. *Journal of Neurology* 256(9), 1397–1412 (Apr 2009) [6](#)
- [20] Gray, H.: *Anatomy of the Human Body*. Lea & Febiger, Philadelphia, 20th edn. (1918) [3, 4](#)
- [21] Greig, D.M., Porteous, B.T., Seheult, A.H.: Exact Maximum A Posteriori Estimation for Binary Images. *Journal of the Royal Statistical Society. Series B (Methodological)* 51(2), 271–279 (Jan 1989) [9, 11, 14, 15, 17](#)
- [22] Ising, E.: Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik* 31(1), 253–258 (Feb 1925) [9, 12, 17](#)
- [23] Kameyama, T., Hashizume, Y., Sobue, G.: Morphologic Features of the Normal Human Cadaveric Spinal Cord. *Spine* 21(11), 1285–1290 (1996) [3](#)
- [24] Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. *International Journal of Computer Vision* 1(4), 321–331 (1988) [9](#)
- [25] Kurtzke, J.F.: Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 33(11), 1444–1452 (Nov 1983) [5](#)



- [26] Long, J., Shelhamer, E., Darrell, T.: Fully Convolutional Networks for Semantic Segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3431–3440 (2015) [8](#)
- [27] Losseff, N.A., Webb, S.L., O’Riordan, J.I., Page, R., Wang, L., Barker, G.J., Tofts, P.S., McDonald, W.I., Miller, D.H., Thompson, A.J.: Spinal cord atrophy and disability in multiple sclerosis. *Brain* 119(3), 701–708 (Jun 1996) [24](#)
- [28] Ludwin, P.S.K., Antel, D.J., Arnold, D.D.L.: Multiple Sclerosis. In: Pfaff, D.W. (ed.) *Neuroscience in the 21st Century*, pp. 2623–2665. Springer New York (2013) [4, 5](#)
- [29] Lukas, C., Knol, D.L., Sombekke, M.H., Bellenberg, B., Hahn, H.K., Popescu, V., Weier, K., Radue, E.W., Gass, A., Kappos, L., Naegelin, Y., Uitdehaag, B.M.J., Geurts, J.J.G., Barkhof, F., Vrenken, H.: Cervical spinal cord volume loss is related to clinical disability progression in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry* 86(4), 410–418 (Apr 2015) [1, 6, 23](#)
- [30] MacKay, D.J.C.: *Information Theory, Inference and Learning Algorithms*. Cambridge University Press (Sep 2003) [9](#)
- [31] Miller, D.H., Barkhof, F., Frank, J.A., Parker, G.J.M., Thompson, A.J.: Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. *Brain* 125(8), 1676–1695 (2002) [6](#)
- [32] Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics* 42(5), 577–685 (Jul 1989) [9, 11, 13, 17](#)
- [33] Olsson, C., Byrod, M., Overgaard, N., Kahl, F.: Extending continuous cuts: Anisotropic metrics and expansion moves. In: 2009 IEEE 12th International Conference on Computer Vision. pp. 405–412 (Sep 2009) [9, 13, 14](#)
- [34] Pezold, S., Amann, M., Weier, K., Fundana, K., Radue, E., Sprenger, T., Cattin, P.: A Semi-automatic Method for the Quantification of Spinal Cord Atrophy. In: Yao, J., Klinder, T., Li, S. (eds.) *Computational Methods and Clinical Applications for Spine Imaging, Lecture Notes in Computational Vision and Biomechanics*, vol. 17, pp. 143–155. Springer International Publishing (2014) [27](#)
- [35] Pezold, S., Fundana, K., Amann, M., Andelova, M., Pfister, A., Sprenger, T., Cattin, P.: Automatic Segmentation of the Spinal Cord Using Continuous Max Flow with Cross-sectional Similarity Prior and Tubularity Features. In: Yao, J., Glocker, B., Klinder, T., Li, S. (eds.) *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging, Lecture Notes in Computational Vision and Biomechanics*, vol. 20, pp. 107–118. Springer International Publishing (2015) [41](#)
- [36] Pezold, S., Horváth, A., Fundana, K., Tsagkas, C., Andělová, M., Weier, K., Amann, M., Cattin, P.C.: Automatic, Robust, and Globally Optimal Segmentation of Tubular Structures. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) *Medical*

## Bibliography

- Image Computing and Computer-Assisted Intervention – MICCAI 2016, pp. 362–370. No. 9902 in Lecture Notes in Computer Science, Springer International Publishing (Oct 2016) [67](#)
- [37] Polman, C.H., Reingold, S.C., Banwell, B., Clanet, M., Cohen, J.A., Filippi, M., Fujihara, K., Havrdova, E., Hutchinson, M., Kappos, L., Lublin, F.D., Montalban, X., O’Connor, P., Sandberg-Wollheim, M., Thompson, A.J., Waubant, E., Weinshenker, B., Wolinsky, J.S.: Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. *Annals of Neurology* 69(2), 292–302 (Feb 2011) [6](#)
- [38] Potts, R.B.: Some generalized order-disorder transformations. *Mathematical Proceedings of the Cambridge Philosophical Society* 48(01), 106–109 (Jan 1952) [13](#)
- [39] Pratt, W.K.: *Digital image processing: PIKS Scientific inside*. Wiley-Interscience, Hoboken, N.J, 4th edn. (2007) [8](#)
- [40] Ray, S.S.: Cut Sets and Cut Vertices. In: *Graph Theory with Algorithms and its Applications*, pp. 115–124. Springer India (2013) [15](#)
- [41] Rocca, M.A., Valsasina, P., Damjanovic, D., Horsfield, M.A., Mesaros, S., Stosic-Opincal, T., Drulovic, J., Filippi, M.: Voxel-wise mapping of cervical cord damage in multiple sclerosis patients with different clinical phenotypes. *Journal of Neurology, Neurosurgery & Psychiatry* 84(1), 35–41 (Jan 2013) [84](#)
- [42] Sanfilippo, M.P., Benedict, R.H., Sharma, J., Weinstock-Guttman, B., Bakshi, R.: The relationship between whole brain volume and disability in multiple sclerosis: A comparison of normalized gray vs. white matter with misclassification correction. *NeuroImage* 26(4), 1068–1077 (Jul 2005) [6](#)
- [43] Strang, G.: Maximal flow through a domain. *Mathematical Programming* 26(2), 123–143 (Jun 1983) [18](#)
- [44] Taso, M., Le Troter, A., Sdika, M., Cohen-Adad, J., Arnoux, P.J., Guye, M., Ranjeva, J.P., Callot, V.: A reliable spatially normalized template of the human spinal cord — Applications to automated white matter/gray matter segmentation and tensor-based morphometry (TBM) mapping of gray matter alterations occurring with age. *NeuroImage* 117, 20–28 (Aug 2015) [84](#)
- [45] Thompson, A.J., Baneke, P.: Atlas of MS 2013. Report, Multiple Sclerosis International Federation (2013), <http://www.msif.org/wp-content/uploads/2014/09/Atlas-of-MS.pdf> [1](#), [4](#), [5](#)
- [46] Trapp, B.D., Nave, K.A.: Multiple Sclerosis: An Immune or Neurodegenerative Disorder? *Annual Review of Neuroscience* 31(1), 247–269 (2008) [4](#), [6](#)
- [47] Yuan, J., Bae, E., Tai, X.C.: A study on continuous max-flow and min-cut approaches. In: *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 2217–2224 (Jun 2010) [2](#), [9](#), [18](#), [19](#), [22](#), [83](#)

# Curriculum Vitae

## Personal Data

---

Name **Simon Pezold**  
Date of Birth January 15, 1985  
Place of Birth Ludwigsburg, Germany

*For a portrait, see Fig. 2.1 (left) on page 4.*

## Education

---

2011–2016 **Ph.D. in Biomedical Engineering**  
Department of Biomedical Engineering, University of Basel  
Supervisor: Prof. Dr. Philippe C. Cattin

2005–2010 **Diplom in Medical Informatics**  
(Combined BS/MS equivalent)  
Studies at University of Heidelberg and Heilbronn University  
Thesis at University of Utah, Salt Lake City

## References

---

Prof. Dr. Philippe C. Cattin  
Department of Biomedical Engineering, Universität Basel  
Gewerbestrasse 14, 4123 Allschwil (Switzerland)  
philippe.cattin@unibas.ch | +41 61 207 54 00

Prof. Dr. Rolf Bendl  
Fakultät für Informatik, Hochschule Heilbronn  
Max-Planck-Str. 39, 74081 Heilbronn (Germany)  
rolf.bendl@hs-heilbronn.de | +49 7131 504 498

Dennis L. Parker, Ph.D.  
Mark H. Huntsman Professor of Radiology  
Utah Center for Advanced Imaging Research, University of Utah  
729 Arapeen Drive, Salt Lake City, UT 84108–1217 (USA)  
parker@uclair.med.utah.edu | +1 801 581 8654

# Publications

*Asterisks (\*) denote equally contributing first authors.*

## Accepted Peer-reviewed Publications

Amann, M. \*, **Pezold, S. \***, Naegelin, Y., Fundana, K., Andělová, M., Weier, K., Stippich, C., Kappos, L., Radue, E.W., Cattin, P., Sprenger, T.: Reliable volumetry of the cervical spinal cord in MS patient follow-up data with cord image analyzer (Cordial). *Journal of Neurology* 263(7), 1364–1374 (May 2016)

Andermatt, S., **Pezold, S.**, Cattin, P.: Multi-dimensional Gated Recurrent Units for the Segmentation of Biomedical 3d-Data. In: Carneiro, G., Mateus, D., Peter, L., Bradley, A., Tavares, J.M.R.S., Belagiannis, V., Papa, J.P., Nascimento, J.C., Loog, M., Lu, Z., Cardoso, J.S., Cornebise, J. (eds.) *Deep Learning and Data Labeling for Medical Applications*, pp. 142–151. No. 10008 in *Lecture Notes in Computer Science*, Springer International Publishing (Oct 2016)

Horváth, A., **Pezold, S.**, Weigel, M., Weier, K., Bieri, O., Cattin, P.: Variational Segmentation of the White and Gray Matter in the Spinal Cord using a Shape Prior. In: *Fourth MICCAI Workshop and Challenge: Computational Methods and Clinical Applications for Spine Imaging (MICCAI CSI 2016)*. Athens, Greece (Oct 2016), (*in press*)

**Pezold, S.**, Amann, M., Weier, K., Fundana, K., Radue, E., Sprenger, T., Cattin, P.: A Semi-automatic Method for the Quantification of Spinal Cord Atrophy. In: Yao, J., Klinder, T., Li, S. (eds.) *Computational Methods and Clinical Applications for Spine Imaging*, *Lecture Notes in Computational Vision and Biomechanics*, vol. 17, pp. 143–155. Springer International Publishing (2014)

**Pezold, S.**, Fundana, K., Amann, M., Andelova, M., Pfister, A., Sprenger, T., Cattin, P.: Automatic Segmentation of the Spinal Cord Using Continuous Max Flow with Cross-sectional Similarity Prior and Tubularity Features. In: Yao, J., Glocker, B., Klinder, T., Li, S. (eds.) *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, *Lecture Notes in Computational Vision and Biomechanics*, vol. 20, pp. 107–118. Springer International Publishing (2015)

**Pezold, S.**, Horváth, A., Fundana, K., Tsagkas, C., Andělová, M., Weier, K., Amann, M., Cattin, P.C.: Automatic, Robust, and Globally Optimal Segmentation of Tubular Structures. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, pp. 362–370. No. 9902 in *Lecture Notes in Computer Science*, Springer International Publishing (Oct 2016)

Schneider, A., Baumberger, C., Griessen, M., **Pezold, S.**, Beinemann, J., Jürgens, P., Cattin, P.C.: Landmark-Based Surgical Navigation. In: Erdt, M., Linguraru, M.G., Laura, C.O., Shekhar, R., Wesarg, S., Ballester, M.A.G., Drechsler, K. (eds.) *Clinical Image-Based Procedures. Translational Research in Medical Imaging*, pp. 57–64. No. 8361 in *Lecture Notes in Computer Science*, Springer International Publishing (Sep 2013)

Schneider, A., **Pezold, S.**, Baek, K.w., Marinov, D., Cattin, P.C.: Direct Calibration of a Laser Ablation System in the Projective Voltage Space. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *Medical Image Computing and Computer-Assisted Intervention MICCAI 2015*, pp. 274–281. No. 9349 in *Lecture Notes in Computer Science*, Springer International Publishing (Oct 2015)

Schneider, A., **Pezold, S.**, Baek, K.w., Marinov, D., Cattin, P.C.: Simultaneous intrinsic and extrinsic calibration of a laser deflecting tilting mirror in the projective voltage space. *International Journal of Computer Assisted Radiology and Surgery* 11(9), 1611–1621 (Jun 2016)

Schneider, A., **Pezold, S.**, Sauer, A., Ebbing, J., Wyler, S., Rosenthal, R., Cattin, P.C.: Augmented Reality Assisted Laparoscopic Partial Nephrectomy. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) *Medical Image Computing and Computer-Assisted Intervention MICCAI 2014*, pp. 357–364. No. 8674 in *Lecture Notes in Computer Science*, Springer International Publishing (Sep 2014)

Schneider, A., Thalmann, P., **Pezold, S.**, Hieber, S.E., Cattin, P.C.: Augmented Reality Assisted Brain Tumor Extraction in Mice. In: Paolis, L.T.D., Mongelli, A. (eds.) *Augmented and Virtual Reality*, pp. 255–264. No. 9254 in *Lecture Notes in Computer Science*, Springer International Publishing (Aug 2015)

Wyder, S., Hennings, F., **Pezold, S.**, Hrbacek, J., Cattin, P.C.: With Gaze Tracking Toward Noninvasive Eye Cancer Treatment. *IEEE Transactions on Biomedical Engineering* 63(9), 1914–1924 (Sep 2016)

## Posters

Amann, M.\*, **Pezold, S.\***, Naegelin, Y., Fundana, K., Andelova, M., Weier, K., Stippich, C., Kappos, L., Cattin, P., Sprenger, T.: Rapid segmentation of the cervical spinal cord on 3d MRI data with cord image analyzer (cordial): application to three-year follow-up data of MS patients with a progressive disease course. Poster, ISMRM 23rd Annual Meeting and Exhibition, Toronto, ON, Canada (2015)

Tsagkas, C., Magon, S., Gaetano, L., **Pezold, S.**, Naegelin, Y., Amann, M., Cattin, P., Wuerfel, J., Bieri, O., Sprenger, T., Kappos, L., Weier, K.: Spinal cord volume loss in multiple sclerosis patients: a 7-year longitudinal study. Poster, 2nd Congress of the European Academy of Neurology (EAN), Copenhagen, Denmark (2016)

Tsagkas, C., Magon, S., Gaetano, L., **Pezold, S.**, Naegelin, Y., Amann, M., Cattin, P., Wuerfel, J., Bieri, O., Sprenger, T., Kappos, L., Weier, K.: Spinal cord volume loss predicts disease ac-

## *Publications*

tivity and disability over time in multiple sclerosis. Poster, 32nd Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS), London, United Kingdom (2016)

Weier, K., **Pezold, S.**, Andelova, M., Amann, M., Magon, S., Naegelin, Y., Radue, E.W., Stippich, C., Gass, A., Kappos, L., Cattin, P., Sprenger, T.: Both spinal cord volume and spinal cord lesions impact physical disability in multiple sclerosis. Poster, 29th Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS), Copenhagen, Denmark (2013)