

# Image processing techniques for quantification and assessment of brain MRI

This thesis was typeset by the author using L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub>

**ISBN:** 978-90-393-6037-8

**Cover design & layout:** Hugo J. Kuijf

**Printed by:** Proefschriftmaken.nl || Uitgeverij BOXPress

# Image processing techniques for quantification and assessment of brain MRI

Beeldverwerkingstechnieken voor de kwantificatie en  
beoordeling van brein MRI  
(met een samenvatting in het Nederlands)

## **Proefschrift**

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op dinsdag 10 december 2013 des middags te 2.30 uur

door

Hugo Jaco Kuijf  
geboren op 28 september 1986 te Woerden

Promotoren: prof. dr. ir. M. A. Viergever  
prof. dr. G. J. Biessels

Copromotor: dr. ir. K. L. Vincken

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van Philips Healthcare, Alzheimer Nederland, Herinneringen Op Linnen BV, de Röntgenstichting en de Internationale Stichting Alzheimer Onderzoek. Financial support by the Dutch Heart Foundation for the publication of this thesis is gratefully acknowledged.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Cerebral microbleeds . . . . .	1
1.2	Cortical cerebral microinfarcts . . . . .	4
1.3	Midsagittal plane . . . . .	5
1.4	Midsagittal surface . . . . .	5
1.5	Manual image processing . . . . .	7
1.6	Thesis outline . . . . .	7
<b>I</b>	<b>Manual image processing</b>	<b>11</b>
<b>2</b>	<b>How to assess the reliability of cerebral microbleed rating?</b>	<b>13</b>
<b>II</b>	<b>Semi-automatic image processing</b>	<b>19</b>
<b>3</b>	<b>Semi-automated detection of cerebral microbleeds on 3.0 T MR images</b>	<b>21</b>
<b>4</b>	<b>Efficient detection of cerebral microbleeds on 7.0 T MR images using the radial symmetry transform</b>	<b>37</b>
<b>5</b>	<b>Detecting cortical cerebral microinfarcts in 7.0 T MR images</b>	<b>57</b>

<b>III Automatic image processing</b>	<b>69</b>
<b>6 Assessment of methods to extract the midsagittal plane from brain MR images</b>	<b>71</b>
<b>7 Automatic extraction of the midsagittal surface from brain MR images using the Kullback–Leibler measure</b>	<b>83</b>
<b>8 Summary and discussion</b>	<b>101</b>
8.1 Summary . . . . .	101
8.2 Discussion . . . . .	103
8.2.1 Manual image processing . . . . .	103
8.2.2 Semi-automatic image processing . . . . .	104
8.2.3 Automatic image processing . . . . .	110
8.3 Conclusion . . . . .	112
<b>Bibliography</b>	<b>115</b>
<b>List of abbreviations</b>	<b>121</b>
<b>Samenvatting</b>	<b>123</b>
<b>Dankwoord</b>	<b>131</b>
<b>List of publications</b>	<b>133</b>
<b>About the author</b>	<b>137</b>

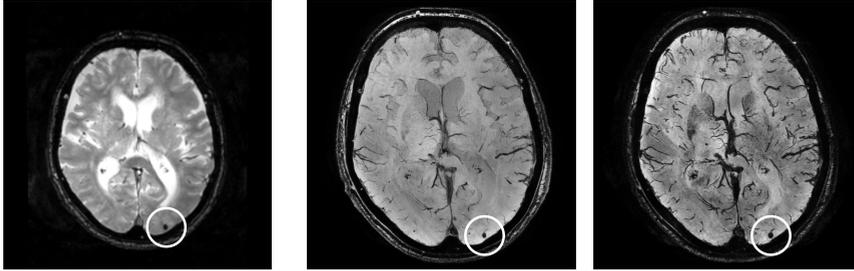
# Introduction

Magnetic resonance imaging (MRI) is a widely used technique to acquire digital images of the human brain. A variety of acquisition protocols is available to generate images in vivo and noninvasively, giving great opportunities to study the anatomy and physiology of the human brain. In this thesis, image processing techniques for quantification and assessment of these images are discussed. These techniques are applied in the context of brain anatomy and pathology, in particular small vessel disease. Different types of image processing techniques can be distinguished: manual, semi-automatic, and automatic techniques.

Advances in MR imaging have led to a tremendous increase in the amount of data and level of detail that can be acquired. This causes manual assessment of these images to become increasingly difficult and time-consuming, thereby threatening the quality of such assessments. Image processing techniques are indispensable to human observers in achieving the best possible results. The main challenge is to translate requirements for quantification and assessment into image processing techniques and apply them in practice.

## 1.1 Cerebral microbleeds

A large part of this thesis focuses on the detection of cerebral microbleeds, either manually or with semi-automated techniques. Microbleeds are small

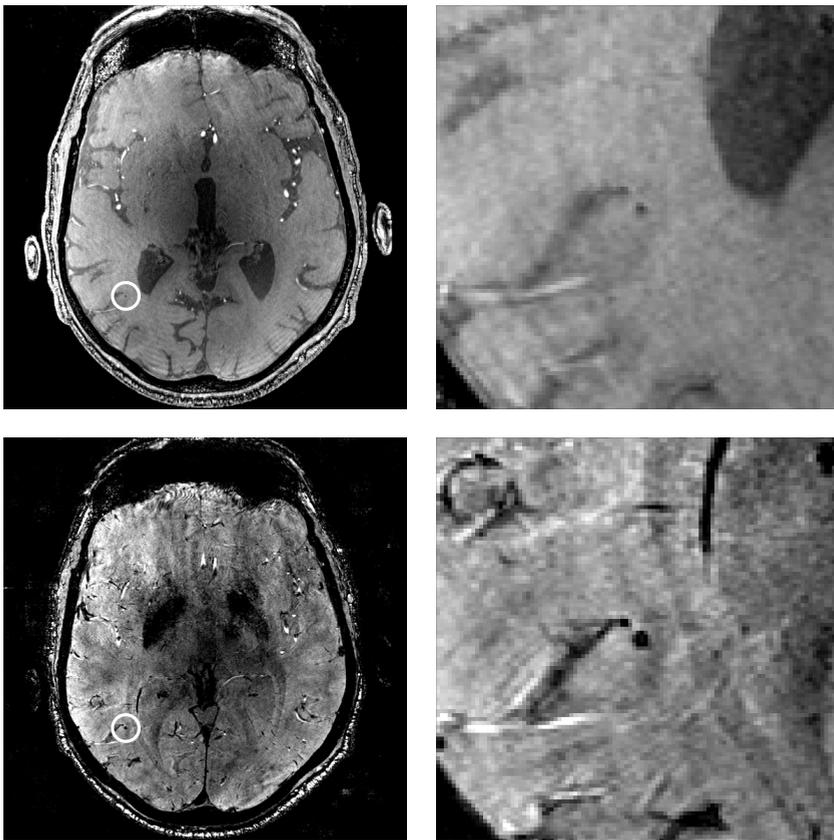


**Figure 1.1:** Subject with a microbleed (encircled) visible on a 3 T T2\*-weighted image (left) and a 7 T dual-echo gradient echo image (right).

hæmorrhages in the human brain that are visualized with gradient echo MRI. They have generated increasing interest among neurologists and researchers, because of their association with various forms of cerebrovascular diseases, Alzheimer’s disease and dementia, and cerebral amyloid angiopathy (Charidimou et al. 2011; Cordonnier et al. 2007; Greenberg et al. 2009; Knudsen et al. 2001; Theysohn et al. 2011; Vernooij et al. 2008; Wardlaw et al. 2006; Wardlaw et al. 2013). Together with this increasing interest, there is an growing demand for robust and reliable microbleed detection techniques.

Microbleeds consist of a paramagnetic hemosiderin deposit that causes a local field inhomogeneity surrounding the microbleed in the magnetic field of an MR scanner (Fazekas et al. 1999). Because of this, microbleeds appear as a spherical signal void on gradient echo MRI. Owing to the so-called blooming effect, the visualized size of the signal void is variable depending on the field strength and echo time (McAuley et al. 2011). This is shown in Figure 1.1, where an example is given of a microbleed visualized on a 3 T T2\*-weighted brain MRI and a 7 T gradient echo dual-echo brain MRI. Another example of a microbleed on a 7 T image is shown in Figure 1.2.

Manual detection of microbleeds is performed using validated visual rating scales, such as the Microbleed Anatomical Rating Scale (MARS, Gregoire et al. (2009)) or the Brain Observer MicroBleed Scale (BOMBS, Cordonnier et al. (2009)). These rating scales define the criteria for a microbleed and give guidelines for reliable visual detection (Wardlaw et al. 2013).



**Figure 1.2:** Subject with a microbleed (encircled) visible on a 7 T dual-echo gradient echo image. Top row: first echo. Bottom row: second echo.

### 1.1.1 Challenges in microbleed detection

Visual rating of microbleeds is a time-consuming task (typically over 30 min for a 7 T image), is rater-dependent (as expressed by a moderate inter-rater agreement), and has limited reproducibility (de Bresser et al. 2013; Cordonnier et al. 2009). These issues will be addressed in this thesis.

Semi-automatic techniques for microbleed detection are likely to aid in reducing or eliminating the aforementioned issues. The technique presented in this thesis is used to detect potential microbleed locations, after which a human observer identifies true microbleeds and censors any false positives. The goal is to achieve a high sensitivity and a moderate number of false positives is allowed. The rationale is that false positives can be quickly censored by human observers. Using this technique, the required rating time and rater-dependency are reduced, and rating reproducibility is improved.

The reliability of microbleed ratings is quantified by determining the inter-rater agreement. If two or more observers perform microbleed detection in a group of subjects, the agreement between observers is expressed with measures as Cohen's kappa coefficient ( $\kappa$ ) or the intraclass correlation coefficient (ICC). These two measures prove to be unreliable in the case of outlier subjects or multiple microbleeds per subject. This is not uncommon with improved microbleed detection, for example using 7 T brain MRI. The use of a more robust measure, the Dice similarity coefficient, is investigated.

## 1.2 Cortical cerebral microinfarcts

Cortical cerebral microinfarcts (in this thesis denoted as microinfarcts) are a hot research topic, since they seem clinically highly relevant (Smith et al. 2012). They are suggested to be the most widespread form of brain infarction and are linked to small vessel disease and dementia. So far, they could only be studied in the setting of neuropathology. In autopsy studies, microinfarcts appear as regions of tissue pallor with neuronal death, gliosis, and sometimes cavitation, having their size varying between 0.05 to 4 mm (Brundel et al. 2012a).

Owing to their small size, microinfarcts are described as "the invisible lesions" and "not visible with the naked eye" (Smith et al. 2012). However, they were recently visualized on 7 T brain MRI (van Veluw et al. 2013b). Van Veluw et al. used FLAIR, T1-, and T2-weighted MRI sequences with a high,

isotropic spatial resolution. Microinfarcts appear hyperintense on FLAIR and T2-weighted images and hypointense on T1-weighted images. An example is shown in Figure 1.3.

Microinfarcts are detected visually by simultaneous examination of all three images in two viewing directions (transversal and sagittal). Assessing this sheer amount of data visually is a difficult and time-consuming task. This hampers a proper and reliable rating and leads to low sensitivities.

In this thesis, a semi-automated detection technique is presented that can help to improve the quality of microinfarct ratings. This technique fully benefits from the 3D high spatial resolution and highlights potential microinfarct locations. The goal is to achieve a high sensitivity, while limiting the number of false positives that need to be manually censored by a human rater.

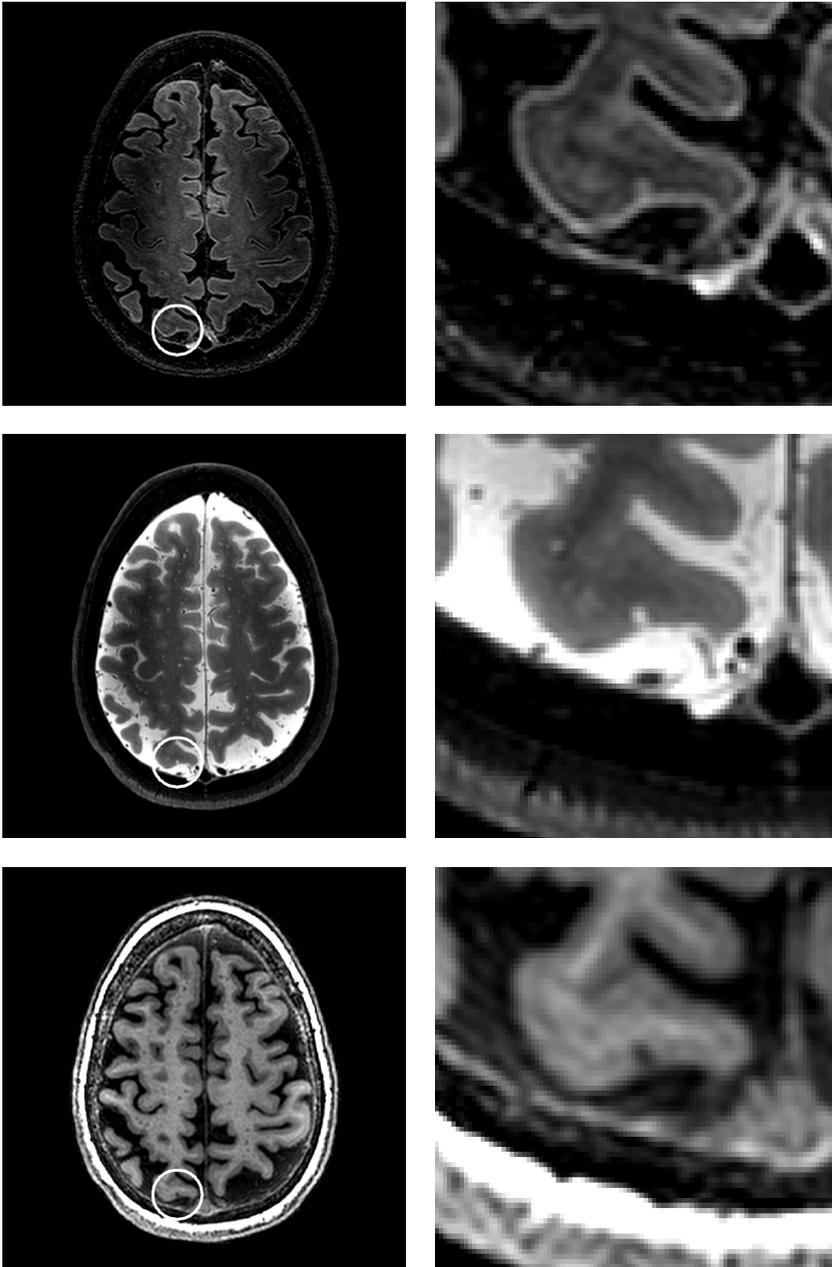
### **1.3 Midsagittal plane**

The two hemispheres of the human brain are separated by the interhemispheric fissure. It is known that each of the two hemispheres has similarities and differences with respect to the contralateral hemisphere. A well-known example is the lateralization of brain function (e.g. Nagel et al. (2013) and Sommer et al. (2001)). Anatomical differences can suggest the presence of pathology (e.g. mass-effect brain tumours, (Joshi et al. 2003)). In order to perform analyses of the hemispheres, accurate extraction of the interhemispheric fissure is a prerequisite. Often, the interhemispheric fissure is modelled as a geometrical plane, the so-called midsagittal plane.

Many methods exist that can extract the midsagittal plane from brain CT or MR images. However, a proper categorization and evaluation of existing methods is lacking. Each individual method was developed on a specific set of images and the used evaluation methods and measures are not comparable. In this thesis, existing methods are categorized in distinct classes and a leading method from each class is implemented. The performance of these methods is evaluated on a representative collection of brain MR images and compared to manual annotations made by multiple human observers.

### **1.4 Midsagittal surface**

Although many techniques use the midsagittal plane to model the interhemispheric fissure, the actual shape of the fissure does not always resemble a



**Figure 1.3:** Left column: microinfarct in a subject, shown on transversal views of 7.0T FLAIR (top), T2 (middle), and T1 (bottom) images. Right column: enlarged visualization of the microinfarct.

true geometrical plane. Owing to variations in anatomy, the shape of the interhemispheric fissure can be curved. This effect is known as brain torque (Toga et al. 2003) and is present in almost every individual, although not to the same extent. The left occipital and right frontal lobes are larger than their counterparts in the other hemisphere.

In order to correctly segment the two hemispheres in 3D, a new approach is needed: the extraction of the midsagittal surface. This surface follows the natural shape of the interhemispheric fissure and thus provides a better segmentation. The performance of the presented technique is compared to manual delineations of the midsagittal surface, in subjects with and without brain torque.

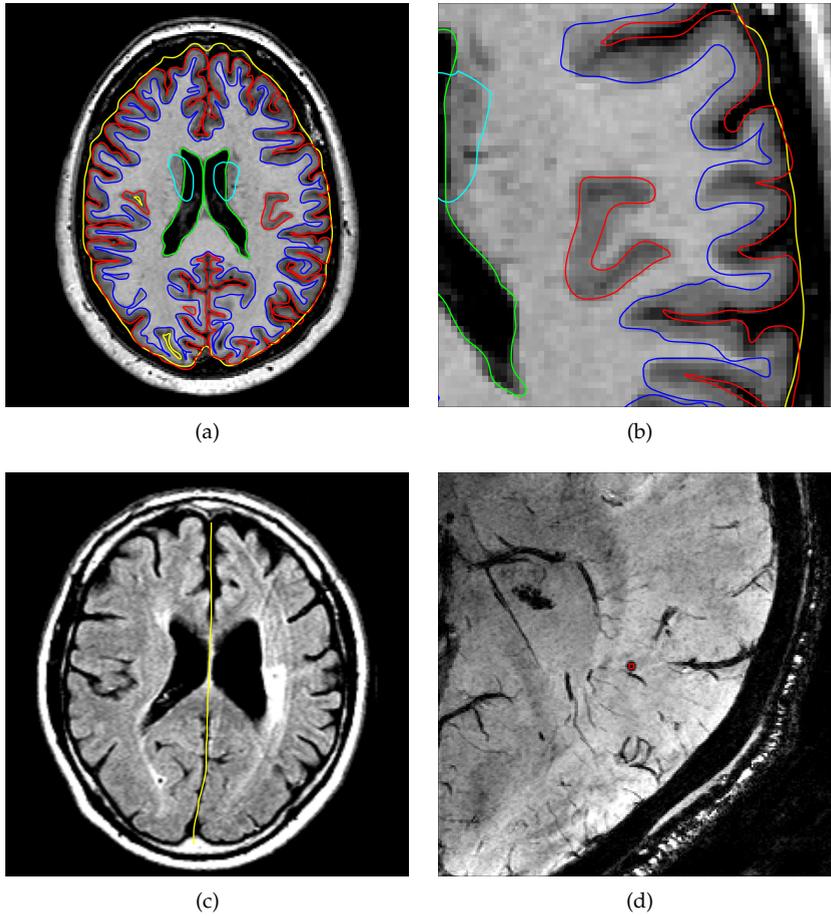
## 1.5 Manual image processing

Every technique that is presented in this thesis uses some form of manual image processing, by rating microbleeds, microinfarcts, or creating delineations of the midsagittal plane or surface. The results of manual ratings or delineations are used to train or test (semi-)automatic image processing techniques.

Dedicated user-friendly computer applications were developed and used by human observers to create these manual ratings, delineations, or other segmentations. Examples are given in Figure 1.4, showing manual segmentations of brain tissue on a 3 T T1-weighted MR image created with a freehand spline drawing technique (Figures 1.4(a) and (b)). This technique was also used to create manual delineations of the midsagittal surface (Figure 1.4(c)) and in works by Wisse et al. (2012), Schlösser et al. (2013), and others. For the detection of microbleeds and microinfarcts, specific point locations in the brain MR images could be annotated (Figure 1.4(d)).

## 1.6 Thesis outline

This thesis describes manual, semi-automatic, and automatic image processing techniques for brain MRI analyses. Manual quantification of cerebral microbleeds is discussed in **Chapter 2**, where measures that determine the reliability of visual detection of microbleeds are addressed. Research on microbleeds is continued in **Chapters 3 and 4**. In these chapters, semi-automatic techniques for the detection of microbleeds are presented. Semi-automatic



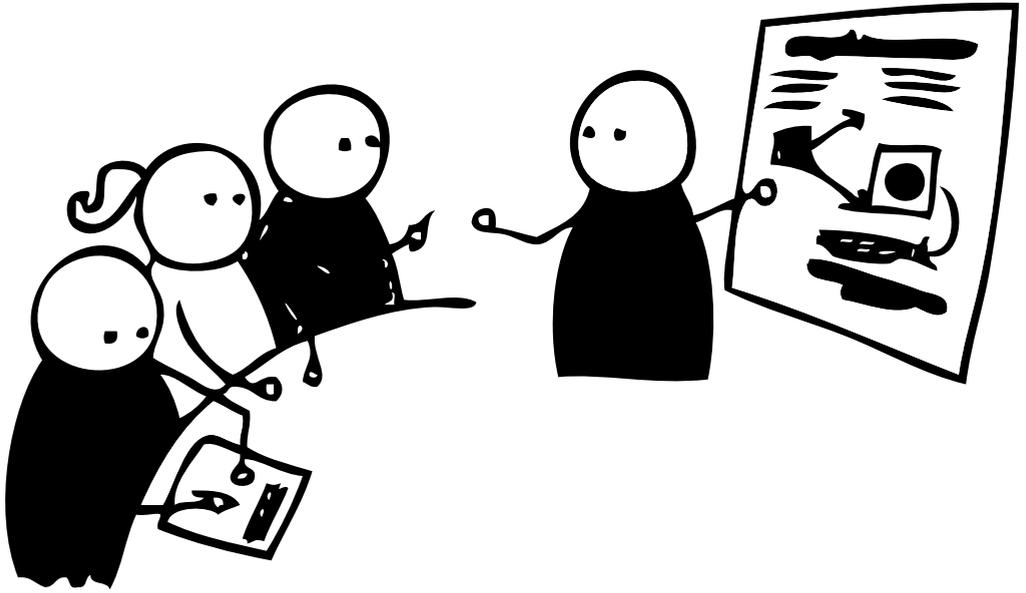
**Figure 1.4:** Examples of manual image processing with dedicated user-friendly computer applications. (a), (b) Manual brain tissue segmentation with freehand spline drawing. (c) Manual delineation of the midsagittal surface. (d) Manual microbleed detection.

techniques are likely to improve the reliability of microbleed ratings. These techniques are used to detect potential microbleed locations in brain MR images and their findings are presented to a human observer for the final quantification and assessment. A similar approach is taken in **Chapter 5** for the semi-automatic detection of cortical cerebral microinfarcts. Microinfarcts are a manifestation of small vessel disease that only recently has been visualized with 7 T MRI. Visual rating of microinfarcts is a difficult and time-consuming task and semi-automatic detection can assist human observers in achieving high quality and reliable ratings. The final part of this thesis focuses on automatic image processing techniques. Extraction of the midsagittal plane is discussed in **Chapter 6**, where the quality of existing methods is evaluated. In **Chapter 7**, a new technique is described for the extraction of the midsagittal surface, which is an improvement over the midsagittal plane.



## **Part I**

# **Manual image processing**



# How to assess the reliability of cerebral microbleed rating?

**Published as:** H.J. Kuijf, S.J. van Veluw, M.A. Viergever, K.L. Vincken, G.J. Biessels, "How to assess the reliability of cerebral microbleed rating?", *Frontiers in Aging Neuroscience*, 2013, nr. 57, vol. 5, pp. 1-2.

Interest in cerebral microbleeds has grown rapidly over the past years. The need for sensitive and reliable detection of microbleeds has spurred the development of new MR sequences and standardized visual rating scales (Cordonnier et al. 2009; Gregoire et al. 2009). The value of these rating scales is currently assessed by measuring the inter-rater agreement, which is commonly determined using Cohen's kappa coefficient ( $\kappa$ ) or the intraclass correlation coefficient (ICC). With the recent increase of MR scanner field strength to 3T and even 7T, the sensitivity of microbleed detection has grown significantly, whence often multiple microbleeds are found in a single subject (de Bresser et al. 2013; Brundel et al. 2012b). As a result of this, researchers no longer solely focus on the absence or presence of microbleeds, but aim at determining their exact count and location as well.

Our concern is that, with this shift of focus, the measures that are in use to validate the reliability of microbleed ratings are no longer up-to-date. If the interest is confined to the presence or absence of microbleeds, the inter-rater agreement can be adequately assessed using  $\kappa$ . However, with multiple microbleeds in an individual subject, determining the inter-rater agreement using a measure that does not consider the number and location of the microbleeds appears inadequate. In other words, raters who agree on the presence or absence of microbleeds in an individual subject might disagree on their count or distribution.

The fact that  $\kappa$  might be an unsuitable measure for studies that are interested in microbleed count has not gone unnoticed. Recently, more studies are reporting the ICC as a measure for inter-rater agreement. This measure partly solves the aforementioned problem, because it takes the number of microbleeds into account. However, microbleed location is not taken into account in determining the ICC. Two raters might agree on the same microbleed count, while having counted different microbleeds. A second important drawback of the ICC is that it is data-dependent. An outlier subject that has many more microbleeds than the other subjects (e.g., a count of  $> 100$  when the median is 2), will highly influence the ICC. This will thwart reliable determination of the inter-rater agreement that does not change because of an outlier subject. This is illustrated in Example A.

**Example A.** In this example, the influence of an outlier subject on the determined inter-rater agreement using ICC is demonstrated. A group of 45 subjects (18 with early Alzheimer's disease and 27 controls) was recruited from a consecutive series of patients referred to our hospital. All subjects un-

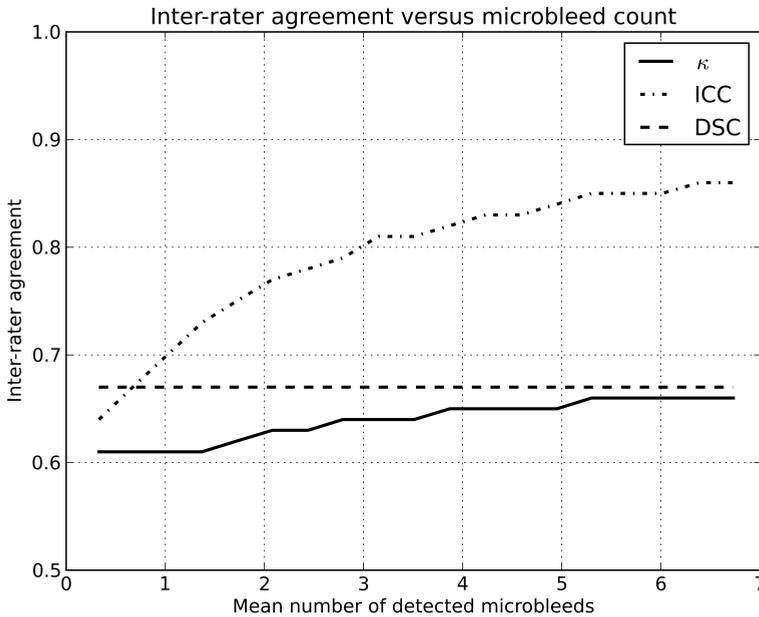
derwent a 7T MRI acquisition with, amongst others, a 3D dual-echo gradient echo weighted sequence (see Brundel et al. (2012b) for details). Written informed consent was given by all subjects and the study was approved by the institutional review board. Presence, count, and exact location of microbleeds was assessed by two independent human raters. Based on these ratings, the average number of microbleeds per subject was four, the median was one, and the ICC = 0.91. Among these subjects, there was an outlier subject who had eighty microbleeds. If this subject was excluded, the average number of microbleeds dropped to two, the median was still one, and the ICC decreased to 0.41. This decrease cannot be attributed to a variation in performance of the raters, but is solely caused by the in- or exclusion of the outlier subject. This is further supported by simulations, available to the reader on: <http://www.isi.uu.nl/People/Hugok/microbleeds/simulator/>.

The data-dependency of the ICC is quite obvious in the given example. Nevertheless, a similar but more subtle effect will occur when newer, highly sensitive MR scanning techniques are used. As the prevalence and number of microbleeds increases (owing to more sensitive detection techniques), the ICC increases as well, even when all ratings are performed by the same raters. This is demonstrated in Example B.

**Example B.** Using the simulator, two (virtual) human raters were simulated. Microbleed detection by these raters was simulated in two groups of subjects with different microbleed prevalence. The prevalence of microbleeds in the first group of subjects ( $L$ ) was low, with on average 0.2 microbleeds/subject and a prevalence of 19%. The second group of subjects ( $H$ ) had on average 3.4 microbleeds/subject with a prevalence of 84%. The two (virtual) human raters ( $X$  and  $Y$ ) had a fixed sensitivity for microbleed detection of 75% and 60%, respectively. The intra-rater agreement between  $X$  and  $Y$  on  $L$  was ICC = 0.64 and on  $H$  was ICC = 0.81. More data are shown in Figure 2.1.

This rise in inter-rater agreement as expressed by the ICC is solely caused by the increase in microbleed count, since all other factors were stable or eliminated in the simulations. This clearly shows the data dependency of the ICC (or  $\kappa$ ).

To overcome these issues, we propose the use of a similarity index, notably the Dice similarity coefficient (DSC), as a more reliable measure of inter-rater agreement (Dice 1945). This measure could be reported in studies rating microbleeds, or other pathologies, to complement existing measures as  $\kappa$  or ICC.



**Figure 2.1:** Inter-rater agreement in simulations where two raters (with fixed sensitivities of 75 % and 60 %) rate microbleeds in different datasets with an increasing number of microbleeds.

The DSC is computed according to the following formula:  $DSC = \frac{2|X \cap Y|}{|X| + |Y|}$ , where  $X$  and  $Y$  are the microbleeds rated by each of the two raters individually. The set  $|X \cap Y|$  contains the microbleeds that are identified by both raters at the exact same location, i.e. the overlap. For example: a subject has five microbleeds, of which three were identified by rater  $X$  and four by rater  $Y$ . Two microbleeds identified by rater  $X$  were also identified by rater  $Y$ . The resulting inter-rater agreement  $DSC = \frac{2 \times 2}{3 + 4} = 0.6$ . If the DSC is computed for Examples A and B, the inter-rater agreement does not change because of outlier subjects or an increase in prevalence (see also Figure 2.1 and the online simulator).

The DSC has added value, because it considers the agreement of detection for every single microbleed, regardless if it occurs in a subject with a single microbleed or many microbleeds. If two raters annotate a microbleed in a subject, they will only reach agreement if this is the exact same micro-

bleed on the exact same location in the brain. This removes the influence of outlier subjects and provides a more direct reflection of the performance of raters in microbleed detection. Computation of the DSC should not require additional time as compared to computing  $\kappa$  or ICC. The verification that two raters identified the same microbleed is already standard procedure during a consensus meeting, thus this should simply be noted to compute the DSC.

The simulated effect demonstrated in Figure 2.1 is also present in real data, albeit more subtle. As MR field strength increases and more microbleeds are detected, thus increasing  $\kappa$ /ICC, the rating of microbleeds becomes more difficult, thus decreasing the  $\kappa$ /ICC. However, this difficulty of rating microbleeds at high resolution images is not expressed to its full extent, as the inter-rater agreement is artificially high owing to the used inter-rater agreement measure.

The DSC has more advantages over existing measures for inter-rater agreement. With the increasing number of studies performing microbleed rating, there is also an increasing request for (semi-) automated microbleed detection techniques (Kuijf et al. 2013d). To compare these techniques with human raters, measures as sensitivity are reported, with a reference ('ground truth') defined by one or multiple human raters. Alternatively, the DSC may be used as a goodness-of-agreement measure. When the DSC is also used for inter-rater agreement, a direct and unbiased comparison can be made between rater performance and the performance of automated techniques. Furthermore, when training novice raters for microbleed detection on scans acquired with high MR field strength, the use of the DSC may indicate the performance of a rater compared with an established ground truth.

In conclusion, we have given arguments why the DSC provides a good measure of inter-rater agreement in studies that aim at determining cerebral microbleed count and locations. The ICC and  $\kappa$  are valid measures of rater agreement for detecting absence or presence of microbleeds in a subject, but become inaccurate in studies where subjects have multiple microbleeds, as is typically true for high-field (3 T, 7 T) brain MRI acquisitions.



## **Part II**

# **Semi-automatic image processing**



# Semi-automated detection of cerebral microbleeds on 3.0 T MR images

**Published as:** H.J. Kuijf, M. Brundel, J. de Bresser, S.J. van Veluw, S.M. Heringa, M.A. Viergever, G.J. Biessels, K.L. Vincken, "Semi-automated detection of cerebral microbleeds on 3.0 T MR images", *PLoS One*, 2013, nr. 6, vol. 8, p. e66610.

### Abstract

Cerebral microbleeds are associated with vascular disease and dementia. They can be detected on MRI and receive increasing attention. Visual rating is the current standard for microbleed detection, but is rater dependent, has limited reproducibility, modest sensitivity, and can be time-consuming. The goal of the current study is to present a tool for semi-automated detection of microbleeds that can assist human raters in the rating procedure.

The radial symmetry transform is originally a technique to highlight circular-shaped objects in two-dimensional images. In the current study, the three-dimensional radial symmetry transform was adapted to detect spherical microbleeds in a series of 72 patients from our hospital, for whom a ground truth visual rating was made by four raters. Potential microbleeds were automatically identified on T2\*-weighted 3.0 T MRI scans and the results were visually checked to identify microbleeds. Final ratings of the radial symmetry transform were compared to human ratings.

After implementing and optimizing the radial symmetry transform, the method achieved a high sensitivity, while maintaining a modest number of false positives. Depending on the settings, sensitivities ranged from 65 to 84 % compared to the ground truth rating. Rating of the processed images required 1 to 2 min per participant, in which 20 to 96 false positive locations per participant were censored. Sensitivities of individual raters ranged from 39 to 86 % compared to the ground truth and required 5 to 10 min per participant per rater.

The sensitivities that were achieved by the radial symmetry transform are similar to those of individual experienced human raters, demonstrating its feasibility and usefulness for semi-automated microbleed detection.

## 3.1 Introduction

Cerebral microbleeds are small spherical lesions that are visible as hypointensities on T2\*-weighted MR images. Recent studies showed that microbleeds are associated with cerebrovascular disease and dementia (Charidimou et al. 2011). Therefore, interest in microbleeds is increasing rapidly and they are being investigated in large studies (e.g. the population-based Rotterdam Scan Study (Vernooij et al. 2008)). The current standard for detection of microbleeds on T2\*-weighted MR scans is rating with validated visual

rating scales, for example the Microbleed Anatomical Rating Scale (MARS) or the Brain Observer MicroBleed Scale (BOMBS) (Cordonnier et al. 2009; Gregoire et al. 2009). However, this process is rater dependent, has limited reproducibility, a modest sensitivity, and can be time-consuming (de Bresser et al. 2013; Cordonnier et al. 2009).

Semi-automated techniques for the detection of cerebral microbleeds are likely to improve the aforementioned issues. These techniques can reduce rater dependence and improve sensitivity, both for experienced and less experienced raters. Other potential advantages are a reduction in required rating time and a higher reproducibility. As such, semi-automated detection may be especially useful for large research studies, in which higher field strengths, higher spatial resolution of scans, and a higher required sensitivity increase the rating difficulty and rating time (de Bresser et al. 2013).

Several methods currently exist for the semi-automated detection of microbleeds. Seghier et al. (2011) described a method that uses a unified segmentation-normalization approach to detect microbleeds. This method detected 50% of the microbleeds present in their participants. While the number of false positives was not reported, manual removal of the false positives required 5 to 10 min on average per participant. Barnes et al. (2011) used a combination of statistical thresholding and a support vector machine supervised learning classifier on susceptibility weighted images. This method had a sensitivity of 82%. On average, over 100 false positives were found per participant, which takes a human rater 5 to 15 min to censor. In a previous study (Kuijff et al. 2012a), we used the 3D radial symmetry transform (RST) to detect microbleeds on non-clinical dual-echo 7.0 T gradient echo data. The 3D RST had a sensitivity of 71%, which is higher than the sensitivity of individual human raters on 7.0 T scans. On average, 17 false positives were found per participant, requiring 2 min to remove, as opposed to 30 min for a full manual rating.

Since clinical use of ultra-high field strength (7.0 T) MRI is still limited, the ability to apply the RST on 3.0 T MR images would be highly useful. Such images are more widely available, but this comes with some challenges, including multi-slice scanning protocols, lower spatial resolution, and an anisotropic voxel size.

In the present study, the use of the RST on 3.0 T T2\*-weighted images was investigated. The approach of the RST, detecting hypointense spherical objects using uncomplicated image processing techniques, combined with its

excellent results on dual echo 7.0 T scans forms an ideal starting base for the detection of microbleeds on 3.0 T images. In order to successfully apply the RST on 3.0 T images, the aforementioned challenges and the lack of additional echoes had to be overcome. The goal was to develop a practical tool for semi-automated detection of microbleeds that can support human raters, offering a high sensitivity while minimizing the number of false positives that have to be censored.

## 3.2 Materials and Methods

### 3.2.1 Participants

For the present study, 72 patients (mean age: 77 years, sd: 8 years, 38 men, 34 women) were included from a consecutive series of patients referred to the memory clinic of the University Medical Center Utrecht, the Netherlands. Inclusion criteria for this study cohort were: cognitive complaints, Mini-Mental State Examination (Folstein et al. 1975) score of  $\geq 20$  and Clinical Dementia Rating  $\leq 1$  (Hughes et al. 1982). Exclusion criteria were: contraindications for MRI (e.g. pacemaker, claustrophobia), a psychiatric or neurological disorder that could influence cognitive functioning, recent non-disabling stroke ( $>2$  years) or any disabling stroke, major depression or a history of alcohol or substance abuse. All 72 patients underwent an MRI-examination between November 2009 and June 2011. Although some scans contained motion artefacts, no patients were excluded for that reason. Diagnoses were established at a multidisciplinary meeting, including: subjective cognitive complaints ( $n=9$ ), mild cognitive impairment (according to Petersen criteria (Petersen et al. 1999),  $n=23$ ), possible ( $n=4$ ) and probable ( $n=33$ ) Alzheimer's Disease (according to the clinical criteria of the National Institute of Neurological Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (McKhann et al. 1984)), vascular dementia (Román et al. 1993) ( $n=1$ ) and semantic dementia (Neary et al. 1998) ( $n=2$ ).

The study was approved by the medical ethics committee of the University Medical Center Utrecht. Written informed consent was given by all participants.

### 3.2.2 MRI acquisition

The participants underwent a standardized MR exam on a 3.0T Philips Achieva MR scanner using an eight-channel head coil, including, among others, a multi-slice T2\*-weighted sequence (TR: 1653 ms, TE: 20 ms), a multi-slice FLAIR sequence (TR: 11000 ms, TE: 125 ms, TI: 2800 ms), and a 3D T1-weighted turbo field echo sequence (TR: 7.9 ms, TE: 4.5 ms). The T2\*-weighted and FLAIR sequences were reconstructed to a voxel size of  $0.96 \times 0.95 \times 3.00 \text{ mm}^3$ . The T1-weighted sequence was reconstructed to a voxel size of 1.0 mm isotropic.

### 3.2.3 Study design

In the current study, we used the RST that was originally designed for 7.0T dual-echo gradient echo images and had a high accuracy (Kuijf et al. 2012a). Because there are still relatively few ultra-high field strength MR scanners available around the world, the ability to apply this method on (single echo) 3.0T T2\*-weighted images leads to higher generalization and opportunities for its use.

First, microbleeds were visually rated on the scans of all participants by four human raters. Based on these ratings, a visual ground truth was created. Second, the RST was implemented for the use on 3.0T MR scans. Five participants with microbleeds were randomly selected and used to determine suitable parameters for the RST and to optimize the work flow. Next, the output of the RST had to be thresholded, to distinguish potential microbleeds from false positives. A series of thresholds was applied and the results were evaluated using a free-response receiver operating characteristic (FROC) curve. Finally, a human rater used a user-friendly interface to inspect the results of the RST after thresholding and censored any remaining false positives. If extra positives were detected by the RST (i.e. microbleeds that were not in the original ground truth), these were added to form the final ground truth if the majority of raters confirmed them as true microbleeds.

### Visual rating of microbleeds

Microbleeds were rated visually on the T2\*-weighted scans of all participants by four raters with different degrees of experience, blinded to all other clinical information. All raters were sufficiently trained prior to the rating sessions.

Rating was performed according to the MARS, as described by Gregoire et al. (2009). In this rating scale, 'definite' microbleeds are defined as hypointense round lesions. If a rater was uncertain about a lesion being a microbleed, it was rated as 'possible'. Mimics like symmetric calcifications in the basal ganglia were disregarded. The detected microbleed locations were compared between raters and in case of discordance, a majority decision was made or an experienced fifth rater was consulted. This consensus-based lesion rating formed the visual ground truth rating.

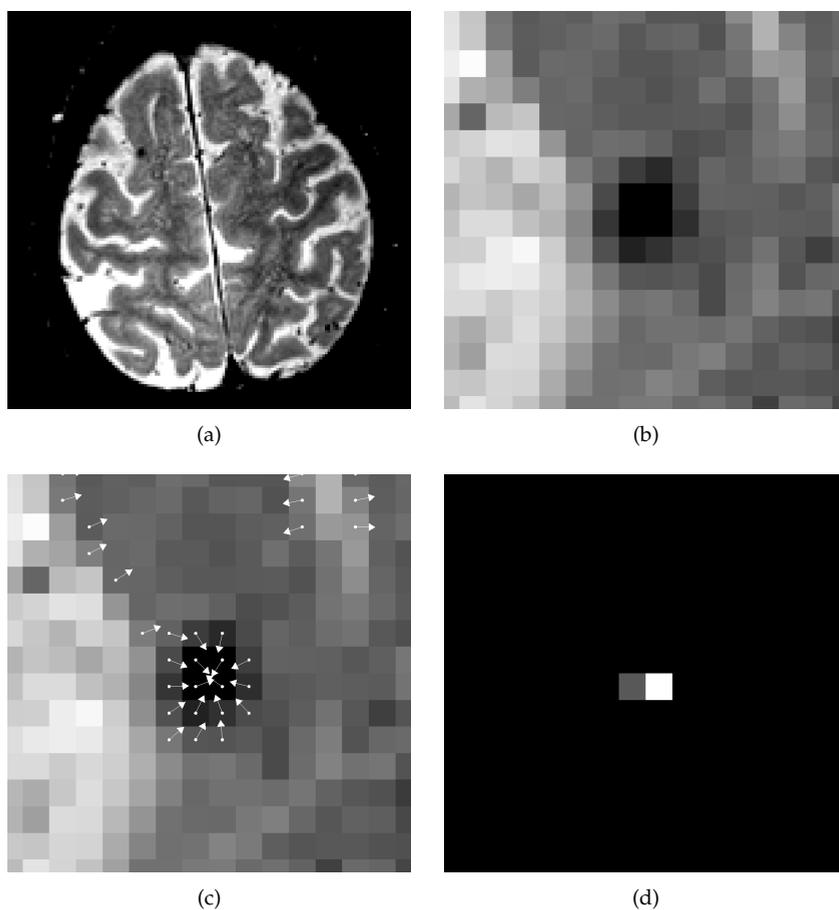
### Radial symmetry transform

The RST is an image processing technique that can be used to highlight spherical-shaped objects in an image. It has numerous applications, both medical and non-medical. For example, the detection of nuclei in H&E stained breast cancer biopsy images and the detection of eyes in pictures of human faces (Loy et al. 2003; Veta et al. 2011). Recently, a 3D version of the RST was used for automatic detection of microbleeds on 7.0 T MR images (Kuijf et al. 2012a).

Before applying the RST, a binary mask of the grey and white matter was obtained using unified segmentation as implemented in SPM8 (Ashburner et al. 2005). For this, the T1-weighted sequence was used and the resulting mask was transformed to the T2\*-weighted sequence using elastix (Klein et al. 2010). Within this mask, the intensities of the T2\*-weighted sequence were normalized using a histogram range-matching procedure. Cocosco et al. 2003 Inspection of the intensities of microbleeds present in the selected five participants showed that all intensity values were below the 6th percentile of the histogram. This was used to normalize the intensity values to a range of [0, 255] using the 6th and 95th percentile.

For each voxel in the image, the 3D RST resulted in a so-called radial symmetry value. This value corresponds to the sphericalness of a local neighbourhood around the voxel. In the case of a microbleed, the centre voxel of the microbleed will receive a high radial symmetry value, because the surrounding hypointense voxels form a spherical spot. This is illustrated in Figure 3.1.

After computation of the 3D RST, a threshold value  $\theta_1$  was set on the radial symmetry value. Locations within the mask with a value above  $\theta_1$  were considered potential microbleed locations, thereby removing locations that have no radial symmetry (e.g. normal brain tissue) or low radial symmetry



**Figure 3.1:** Workflow of the radial symmetry transform. An example slice of a 3.0 T T2\*-weighted image with a microbleed is shown in (a). (b) shows a zoom-in of the microbleed shown in (a). An intermediate step of the radial symmetry transform is shown in (c), where each voxel contributes to some neighbouring voxel. If multiple arrows point towards the same target voxel, this target voxel receives a high radial symmetry value. The output of the radial symmetry transform is shown in (d), displaying a large value at the centre voxel of the microbleed.

(e.g. vessels). This procedure is similar to the method described earlier for detection of microbleeds on 7.0 T dual-echo gradient echo images (Kuijf et al. 2012a). However, the lower spatial resolution of the 3.0 T images and the lack of a second echo posed a challenge: the number of false positives remaining after the previous steps was significantly higher (up to 200 per participant). The availability of the second echo at 7.0 T enabled the removal of false positive locations that were only visible on one echo, but this was not possible for the 3.0 T images. Therefore, an additional step to the method was introduced for 3.0 T images after computation of the 3D RST, involving a 2D RST computed on a transversal minimum intensity projection (minIP) image. The rationale behind this additional step was derived from the visual rating process of microbleeds, in which a rater searches for 2D hypointense circular spots on a minIP representation of the scan. On a minIP, microbleeds show up as dark round spots, while typical false positives (such as vessels) are displayed as elongated structures. This facilitates the distinction between microbleeds and vessels.

A minIP slab of 12 mm thickness was created in a square region of interest ( $1 \text{ cm}^2$ ) at each potential microbleed location after the 3D RST. The 2D RST was computed on this transversal slab with settings identical to the 3D RST. If the new radial symmetry value of the potential microbleed did not exceed the threshold  $\theta_2$ , it was discarded as a potential microbleed.

## Experiments

Parameter settings that influence the quality of the 3D RST were optimized on the selected five participants with microbleeds, similar to a study published earlier (Kuijf et al. 2011).

Threshold values  $\theta_1$  and  $\theta_2$  were used after computation of the 3D and 2D RST, respectively. These threshold values determined the locations that were considered as potential microbleeds. Different values gave different results and therefore a range of threshold values was heuristically established. This was done on the selected five participants with microbleeds. Subsequently, an exhaustive search that computed the results of the RST for all combinations of  $\theta_1$  and  $\theta_2$  was performed on all 72 participants. By plotting a FROC curve with the results of this search, the sensitivity of each combination of  $\theta_1$  and  $\theta_2$  can be assessed together with the number of false positives.

Three combinations of  $\theta_1$  and  $\theta_2$  were selected, representing: A) moder-

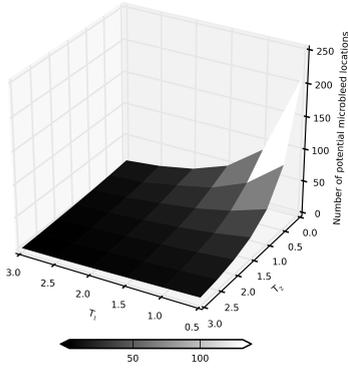
ate, B) good, and C) high sensitivity for detection of microbleeds. For the three combinations, rater 4 censored all potential microbleed locations that were identified. The censoring of the potential microbleed locations was performed several months after the initial visual rating of rater 4, to ensure that previous ratings were not remembered. Additionally, if microbleeds were detected by the RST that were not present in the original visual rating, they were presented to all human raters for inspection. If at least three raters had confirmed them as true microbleeds, they were added as 'extra positives' to the final ground truth rating. Adding the extra microbleeds improved the ground truth rating and gave a fair comparison of the sensitivities of human raters versus the RST.

### 3.3 Results

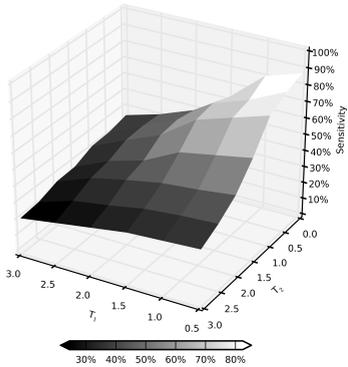
During the visual rating, 244 locations were marked as definite or possible microbleeds by at least one of the four raters. After the consensus meeting, 148 locations in 38 participants (53 %) were confirmed as microbleeds and defined as the visual ground truth: 106 definite and 42 possible microbleeds. Of those microbleeds: 93 (63 %) were detected by rater 1, 130 (88 %) by rater 2, 60 (41 %) by rater 3, and 104 (70 %) by rater 4.

To determine the appropriate range of  $\theta_1$  and  $\theta_2$ , the radial symmetry values at the locations of microbleeds in the five participants were inspected. As a result,  $\theta_1$  was chosen to range from 0.5 to 3.0 and  $\theta_2$  from 0.0 to 3.0, both with a step size of 0.5. Setting  $\theta_2$  to zero means that the 2D RST was not computed at all.

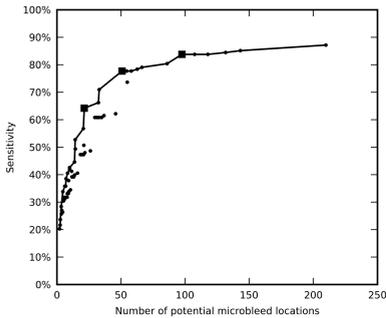
For each combination of  $\theta_1$  and  $\theta_2$ , the number of potential microbleed locations per participant is shown in Figure 3.2(a) and the sensitivity of the method is shown in Figure 3.2(b). These figures clearly show that lower values of  $\theta_1$  and  $\theta_2$  result in more potential microbleed locations (i.e. a higher false positive rate), but also a higher sensitivity. Therefore, when searching for the optimal  $\theta_1$  and  $\theta_2$ , a trade-off has to be made between a high sensitivity and a long censoring time, or a lower sensitivity with a shorter censoring time. This is illustrated in Figure 3.2(c), where all 42 combinations of  $\theta_1$  and  $\theta_2$  are plotted in a FROC curve. As the sensitivity increases, the number of false positives increases accordingly. The solid line in the graph indicates the optimal combinations of  $\theta_1$  and  $\theta_2$ , i.e. the highest sensitivity having the lowest number of false positives.



(a)



(b)



(c)

**Figure 3.2:** Results of the radial symmetry transform for detecting microbleeds. Appropriate values for  $\theta_1$  and  $\theta_2$  were determined on five randomly selected participants with microbleeds. The figures show the results of these values on all 72 participants.

(a) Number of detected potential microbleed locations per participant ( $n=72$ ), visualized as a function of  $\theta_1$  and  $\theta_2$ . Potential microbleeds need to be censored by a human rater to identify true microbleeds (on average 2 per participant) and reject false positives that remained after the 3D and 2D radial symmetry transform.

(b) Sensitivity of the radial symmetry transform on the visual ground truth, visualized as a function of  $\theta_1$  and  $\theta_2$ .

(c) This figure shows the relationship between sensitivity and the number of potential microbleed locations per participant, where each dot is a combinations of  $\theta_1$  and  $\theta_2$ . Depending on the preferred sensitivity, there is an optimal combination of  $\theta_1$  and  $\theta_2$  with the lowest number of locations. These optimal combinations are annotated by the solid line. Three combinations (*A*, *B*, and *C*, annotated from left to right with squares), with moderate, good, and high sensitivity, were selected for inspection by a human rater.

**Table 3.1:** Results of the radial symmetry transform after censoring potential microbleed locations per participant. Total number of detected true positives, extra positives, false positives, and total time required to censor the false positives in all participants. Median rating time required by rater 4, with the interquartile range (Q1 and Q3), is shown. In the visual ground truth rating, 148 microbleeds were found in the 72 participants.

	Number of microbleeds			Required rating time (min)			
	#TP	#EP	#FP (mean $\pm$ sd)	Total	Median	Q1	Q3
A	95	4	1436 (20 $\pm$ 12)	70.8	0.7	0.4	1.1
B	115	4	3538 (49 $\pm$ 25)	121.0	1.4	0.9	1.9
C	124	4	6891 (96 $\pm$ 44)	197.5	2.2	1.9	3.3

**Table 3.2:** Sensitivity of manual and semi-automated rating. Final sensitivities of the individual human raters, all raters combined, and the three chosen combinations of the radial symmetry transform. A total of 152 microbleeds was found (148 during visual rating + 4 extra positives by the radial symmetry transform) in 38 participants (53 %).

	Sensitivity (count)
Rater 1	61 % (93)
Rater 2	86 % (130)
Rater 3	39 % (60)
Rater 4	68 % (104)
All raters	97 % (148)
RST A	65 % (99)
RST B	78 % (119)
RST C	84 % (128)

Based on this graph, three combinations of  $\theta_1$  and  $\theta_2$  (*A*, *B*, and *C*) were selected and presented to rater 4 for censoring. The results of the semi-automatic method after manual censoring are shown in Table 3.1.

In total, 4 extra positives (all possible microbleeds) were found that were not present in the visual consensus rating, but were marked as potential microbleeds by the RST and thereafter confirmed by the majority of raters. These were added to the final ground truth, resulting in a total of 152 microbleeds. The results are shown in Table 3.2.

Of the 38 participants that had at least one microbleed, 30 were identified at setting *A* (i.e. an identification rate of 79 %). All eight participants that were not identified had only one microbleed. If a participant had at least one true microbleed detected at setting *A* and subsequently setting *C* was

used on only these participants, the final sensitivity of the method for the detection of individual microbleeds was 83%. This enables fast rating, as participants without microbleeds can be quickly discarded at setting  $A$ , while still achieving a high sensitivity by subsequently using setting  $C$ .

The time needed for the original visual rating was approximately 5 to 10 min per participant per rater. The required human rater time for visual censoring of the results of the RST is reported in Table 3.1, medians ranged from 1 to 2 min per participant. A dedicated user-friendly interface supported the human rater to censor the results of the RST. Time required for loading the scan into the computer memory, adjusting the contrast settings of the scan prior to rating to normalize the viewing, and breaks were not taken into account. Computation time of the 3D RST was approximately 20 s per participant. The computation time of the 2D RST depended on the number of potential microbleeds, influenced by the values of  $\theta_1$  and  $\theta_2$ , and required another 5 s to 2 min per participant. All computations were performed on a standard workstation.

### 3.4 Discussion

Detection of microbleeds on T2\*-weighted 3.0 T MRI scans can be achieved by the RST, with a high sensitivity and a limited amount of visual rating time required to censor false positive locations.

Most importantly, the sensitivity of the method was similar to the sensitivity of the most experienced human raters. This would allow for the method to be used in practice, for example in research studies to assess the presence of microbleed in large patient cohorts. The use of the method will make the detection of microbleeds less rater dependent and thereby improve the overall quality of the rating.

Comparing the presented method and its results to existing methods, the RST shows some advantages. The resulting sensitivities of 65 to 84% are higher than the reported sensitivities of the methods of Seghier et al. (2011) (50%) and Barnes et al. (2011) (82%). It should be noted that all methods did not include a full FROC curve in their results. Therefore it could be that other sensitivities might be possible and that the presented results are only a single point on the FROC curve, being a trade-off between sensitivity and number of false positives. Furthermore, other factors, such as differences in scan type and quality, might explain differences between various methods.

Recently, Bian et al. (2013) demonstrated the use of the 2D RST for the detection of radiation-induced microbleeds on 3.0 T susceptibility-weighted images. The images underwent a minIP post-processing on which the 2D RST was applied and additional image processing techniques were used to eliminate false positive locations. This resulted in a sensitivity of 87% with on average 45 false positive locations per participant. The work of Bian and the method presented here demonstrate the effectiveness of the RST for microbleed detection. Although there are some differences in the approach, participant population, and cause of microbleeds, these works contribute to the understanding of applying the RST for microbleed detection and complement each other.

Comparing the required human rater time showed lower rating times with the RST, requiring only 1 to 2 min as compared to 5 to 10 min (Seghier et al. 2011) and 5 to 15 min (Barnes et al. 2011). Bian et al. (2013) did not report the required rating time, but only the average numbers of 30 true microbleeds and 45 false positive locations per participant. Seghier et al. (2011) did not report the number of detected potential microbleed locations, so the rating time cannot be explained. Barnes et al. (2011) reported on average 125 potential microbleed locations per participant, with on average 17 true positives and 108 false positives. The longer rating time as compared to the RST might be caused by the higher number of true microbleeds that were present. During the censoring of the results of the RST, rater 4 needed more time to correctly identify a true microbleed location than to reject a false positive location. Assuming this behaviour also holds for the raters in the study of Barnes et al. (2011), this might explain the longer required human rater time. Next to this, false positives of the RST appeared mostly at locations where they could be easily rejected by a trained human rater. Furthermore, individual differences between human raters can also have an effect on the required rating time. Besides this, it is noticeable that censoring the results of the RST on 3.0 T scans required less time than censoring the results of the RST on 7.0 T scans (Kuijff et al. 2012a). This is likely caused by the availability of the second echo on 7.0 T scans, that required additional time for inspection. In addition, microbleeds appear smaller on 7.0 T scans and thus require additional zooming by the human rater during censoring.

The most significant advantage of the RST over existing methods is the easy and uncomplicated implementation. The main unique features of a microbleed (circular shape and hypointense on a T2\*-weighted scan) are directly

translated into the RST. Loy et al. (2003) provide a clear and straightforward description of the implementation of the RST and offer open source implementations online. Furthermore, the method is (largely) independent of scan parameters and field strength of the MR, thus making it generic and suitable for widespread usage. The required amount of training data is minimal, whereas techniques using unified segmentation or classifiers require a substantial amount of training data. Training and optimization of parameters was performed on just a few participants, as demonstrated here and earlier (Kuijf et al. 2011). After the initial training and optimization of the parameters, the entire procedure can be automated and processed offline: extracting images from the picture archiving and communication system, computing the 3D RST, and computing the 2D RST. After these automated steps, limited human rater time is required only for visual censoring of the results using a user-friendly interface.

Since the automatic part of the method always produces the same results, the reproducibility of the method is high but still dependent on the rater that performs the final visual censoring. However, the rater dependence that is usually involved with visual rating of microbleeds can be reduced to a large extent. Potential microbleed locations are annotated by the RST, but identifying true microbleeds and rejecting false positives remains up to the decision and interpretation of a human rater. Less experienced raters now have a tool that will guide them to achieve high sensitivities and experienced raters will benefit from the reduced rating time.

Some microbleeds that were detected by the original visual rating were missed by the RST, even at the setting with a relatively high number of false positives (setting C). Most of those false negatives were too small in size to be detected by the RST.

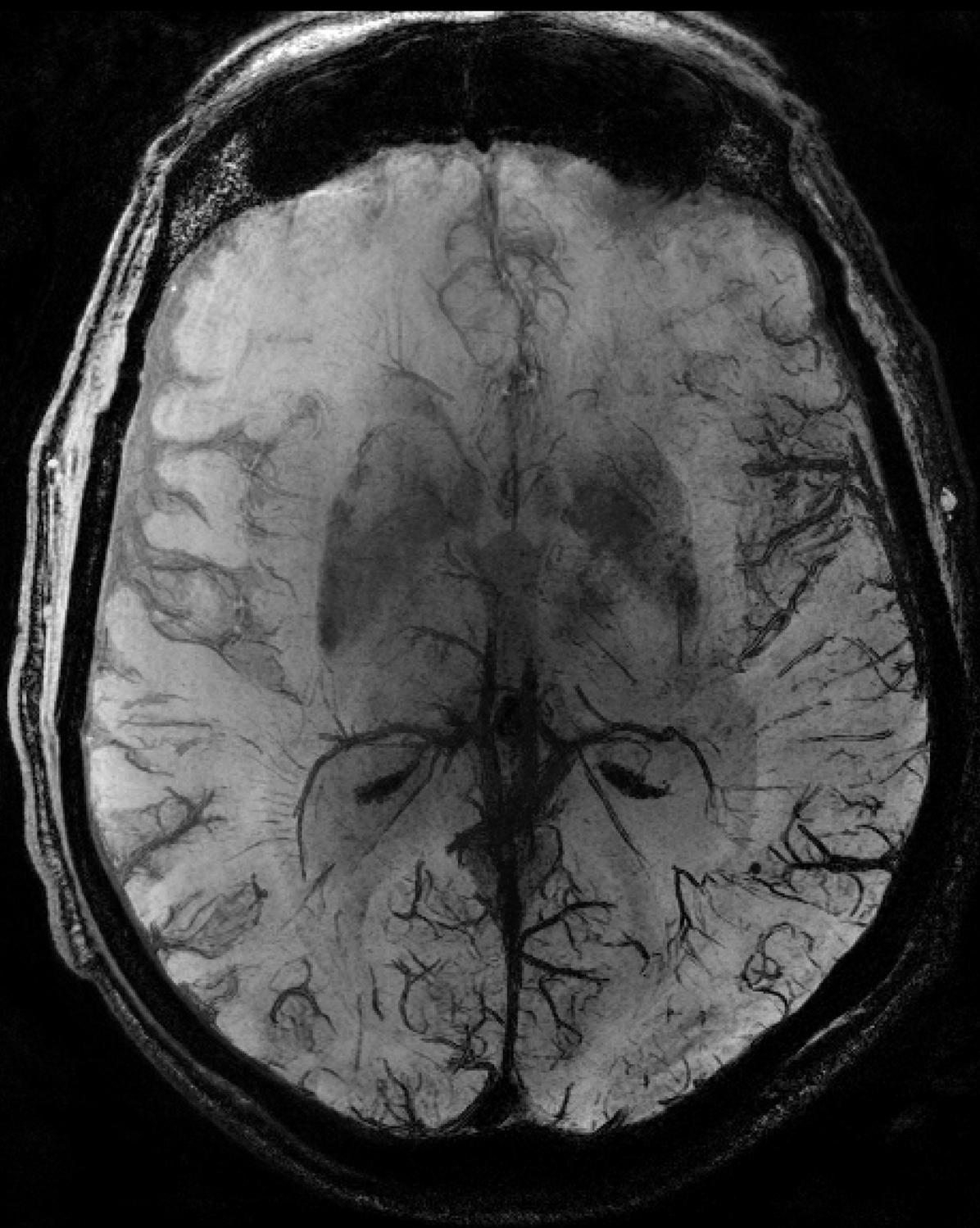
The false positive locations that were detected by the RST needed to be removed by a human rater. These false positive locations were mostly vessels (e.g. periventricular, in sulci, in the cerebellum); calcifications in the basal ganglia and choroid plexus; the interhemispheric fissure and Sylvian fissures; and bone and air artefacts, especially located at the skull base. Scans with motion artefacts had more false positives, but true microbleeds were still reliably detected. Most false positives appeared at locations where they can be easily rejected by a human rater, resulting in a low rating time. Applying the 2D RST removed 50 to 90% of the false positives that were present after the 3D RST, depending on the value of  $\theta_2$ . Higher values of  $\theta_2$  removed more

false positives, but introduced some false negatives as well, and vice versa for lower values of  $\theta_2$  (see Figures 3.2(a) and 3.2(b)). To remove even more false positives, additional scans such as MR angiography could be included to eliminate vessels.

In our opinion, the ability to adapt and fine-tune the RST gives advanced opportunities for its use in a clinical setting. By simply adapting the parameters  $\theta_1$  and  $\theta_2$ , a consideration between sensitivity and rating time can be made. This flexibility of the method can be used to create a time-efficient workflow for microbleed detection in a research setting. With the use of setting *A*, participants without microbleeds can be quickly discarded. Participants with at least one microbleed at setting *A* can be further inspected by using setting *B* or *C*. As such, quick scoring is performed for participants without microbleeds. But for participants with at least one microbleed at setting *A*, a high sensitivity is achieved by using (the more time-consuming) setting *B* or *C*. Another interesting approach would be to apply the RST on quantitative susceptibility mapping (QSM) images (Liu et al. 2012). By acquiring multiple echo times, QSM images will be able to provide a better contrast between microbleeds and surrounding tissue. This was not done in the clinical acquisition protocol used, but would probably lead to a higher sensitivity and a lower number of false positives.

### **3.5 Conclusions**

A flexible system for efficient detection of cerebral microbleeds using the radial symmetry transform was presented. The sensitivity of this method is high, while the number of false positives that need to be censored was minimized.



# Efficient detection of cerebral microbleeds on 7.0 T MR images using the radial symmetry transform

**Published as:** H.J. Kuijf, J. de Bresser, M.I. Geerlings, M. Conijn, M.A. Viergever, G.J. Biessels, K.L. Vincken, "Efficient detection of cerebral microbleeds on 7.0T MR images using the radial symmetry transform", *NeuroImage*, 2012, nr. 3, vol. 59, pp. 2266-2273.

### Abstract

Cerebral microbleeds (CMBs) are commonly detected on MRI and have recently received an increased interest, because they are associated with vascular disease and dementia. Identification and rating of CMBs on MRI images may be facilitated by semi-automatic detection, particularly on high resolution images acquired at high field strength. For these images, visual rating is time-consuming and has limited reproducibility. We present the radial symmetry transform (RST) as an efficient method for semi-automated CMB detection on 7.0 T MR images, with a high sensitivity and a low number of false positives that have to be censored manually.

The RST was computed on both echoes of a dual-echo T2\*-weighted gradient echo 7.0 T MR sequence in 18 participants from the Second Manifestations of ARterial disease (SMART) study. Potential CMBs were identified by combining the output of the transform on both echoes. Each potential CMB identified through the RST was visually checked by two raters to identify probable CMBs. The scoring time needed to manually reject false positives was recorded.

The sensitivity of 71.2% is higher than that of individual human raters on 7.0 T scans and the required human rater time is reduced from 30 to 2 min per scan on average. The RST outperforms published semi-automated methods in terms of either a higher sensitivity or less false positives, and requires much less human rater time.

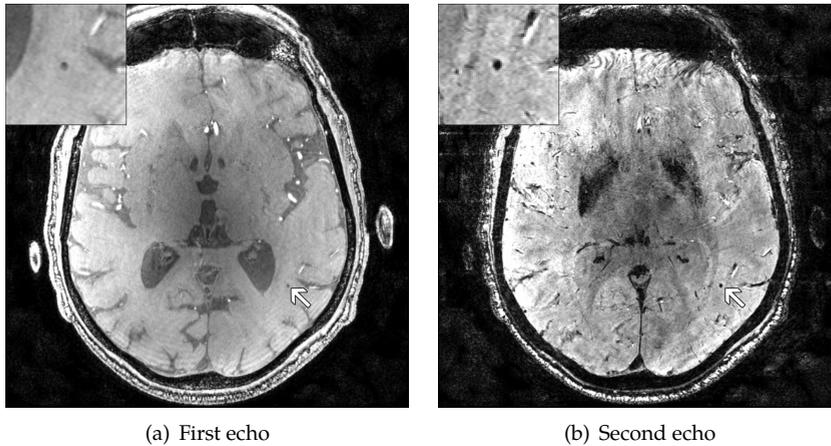
## 4.1 Introduction

Interest in cerebral microbleeds (CMBs) is increasing rapidly since a few years. CMBs are seen as phenomena distinct from larger haemorrhages. CMBs are associated with hypertensive vasculopathy, white matter hyperintensities and lacunar infarcts, and they are a key MRI marker of cerebral amyloid angiopathy (Greenberg et al. 2009; Knudsen et al. 2001; Theysohn et al. 2011; Vernooij et al. 2008; Wardlaw et al. 2006). CMBs consist of hemosiderin deposits (Fazekas et al. 1999) that are paramagnetic and cause a local susceptibility effect inside the magnetic field of the MR scanner. As a result, CMBs can be visualized as round, hypointense spots on a T2\*-weighted gradient echo MR sequence. At regular field strength (1 to 3 T), CMBs are usually defined as having a diameter ranging from 2 to 10 mm (Cordonnier et al. 2007).

The current standard for microbleed detection is visual rating with validated visual rating scales (Greenberg et al. 2009; Gregoire et al. 2009). As visual rating is time-consuming and has limited reproducibility, (semi-)automated detection may improve rating quality and decrease rating time. Recently, two methods on semi-automatic detection of CMBs have been published by Seghier et al. (2011) and Barnes et al. (2011). Seghier et al. described a method using a unified segmentation-normalization approach to detect microbleeds. The method identified 77% of patients with microbleeds; no results were given on detection of the actual individual microbleeds. While numbers of false positives were not reported, manual removal of the false positives required 5 to 10 min on average. Barnes et al. used a combination of statistical thresholding and a support vector machine supervised learning classifier on susceptibility weighted images. This method detected 81.7% of all individual microbleeds present in their data (identifying all patients). On average, over 100 false positives were found per patient, which takes a human rater 5 to 15 min to remove.

With the introduction of high-field 7.0 T MR scanners, detection of much smaller CMBs has become feasible (Conijn et al. 2010; De Reuck et al. 2011). However, while visual rating may be suitable for scans acquired at regular field strength, it takes a single rater about 30 min to assess a high resolution 7.0 T MR scan (typically  $570 \times 570 \times 333$  voxels) for the presence of CMBs. This is due to the higher number of slices, the smaller size of the CMBs, and the fact that structures other than CMBs, in particular blood vessels, also show up at higher field strength. In a proof-of-principle experiment preceding this study, promising results have been published showing a large reduction in rating time when using the radial symmetry transform (RST) for semi-automatic microbleed detection (Kuijf et al. 2011). Moreover, an automated method might improve the accuracy of microbleed rating on 7.0 T MR scans. This is important, because it was observed that the inter-rater accuracy of visual rating of CMBs by experienced raters dropped from an ICC of 0.95 on 1.5 T MR scans to an ICC of 0.46 on 7.0 T MR scans (de Bresser et al. 2013). This is mainly due to reduced sensitivity, because CMBs on 7.0 T scans are easily missed by the human eye.

The goal of the present study was to develop a system with a high sensitivity, while minimizing the number of false positives, that provides automated detection of potential CMBs in a representative group of participants, and enables time-efficient visual censoring of false positive CMBs.



**Figure 4.1:** A typical slice of a 7.0 T T2\*-weighted MR scan, showing the first and second echo. The arrows annotate a microbleed (enlarged in the upper-left corner).

## 4.2 Methods and materials

### 4.2.1 Participants

For this study, 18 participants (mean age: 60 years, sd: 13 years, 15 men and 3 women) with microbleeds were included from the Second Manifestations of ARterial disease (SMART) study (Simons et al. 1999). The objectives of the SMART study are to determine the prevalence of vascular risk factors and concomitant arterial disease and to study the incidence of future cardiovascular events and its predictors in patients newly referred to our hospital with atherosclerotic disease. The SMART study and the 7.0 T imaging were approved by the Medical Ethics Committee. Written informed consent was given by all participants. Of the 18 participants, 6 were included with diabetes mellitus, 4 with cerebrovascular disease, 2 with coronary artery disease, 2 with peripheral artery disease, 1 with hypertension, and 3 with other vascular risk factors or disease.

### 4.2.2 MRI acquisition

MRI acquisition was performed as described previously by Conijn et al. (2010), on a 7.0 T whole-body system (Philips Healthcare, Cleveland, OH).

A dual echo T2\*-weighted sequence was acquired (first echo time: 2.5 ms, second echo time: 15.0 ms, repetition time: 20.0 ms). A flyback gradient was applied between the first and second readout. Sensitivity encoding (SENSE) was applied in the RL direction. The images were reconstructed to  $0.35 \times 0.35 \times 0.3 \text{ mm}^3$  voxels and the built-in phase correction, partial-echo filter and homogeneity correction of the MR system were applied during reconstruction. A typical transversal slice of a high-resolution 7.0 T scan is shown in Figure 4.1.

Within the same scanning session, a 3D T1-weighted turbo field echo MR sequence was acquired, with a repetition time of 7.0 ms and an echo time of 3.0 ms. This scan was reconstructed to  $0.66 \times 0.66 \times 0.7 \text{ mm}^3$  voxels and SENSE was applied in the RL direction. This scan was used for preprocessing (see §4.2.4).

### 4.2.3 Visual rating of microbleeds

All scans were independently and visually scored by two raters, both neuroradiologists with more than 20 years of experience, as part of a previous study by Conijn et al. (2011). Minimal intensity projection post-processing of both echo times were presented simultaneously. The raters were blinded to all other clinical information. Rating of the microbleeds was performed according to the Microbleed Anatomical Rating Scale (MARS), as described by Gregoire et al. (2009). In this rating scale, microbleeds can be defined as 'definite' or 'possible' microbleeds. A slight adjustment was made to the MARS by excluding the size-criterion. Owing to the blooming effect (McAuley et al. 2011), microbleeds appear larger on a T2\*-weighted scan than the actual volume of the hemosiderin deposit. Therefore, a size criterion is ambiguous as microbleed sizes will vary among different scan sequences and parameter settings. For instance, increasing the echo time will increase the apparent size of the microbleeds on the T2\*-weighted scan.

On T2\*-weighted images, definite microbleeds were defined as black, round lesions on the first or second echo image. If such lesions were visible on the first echo, but were not larger on the second echo, they were not scored as microbleeds, for lack of a blooming effect by the longer echo time. If a rater was uncertain about a lesion being a microbleed, it could be rated as a possible microbleed. The lesions on which the two raters disagreed were evaluated in a consensus meeting to obtain a final score, containing definite

and possible microbleeds.

In total, 54 microbleeds were visually detected in the 18 participants: 45 definite and 9 possible microbleeds. Of those microbleeds, 25 were detected by both raters, 9 only by rater 1, 19 only by rater 2, and 1 microbleed was initially not detected by one of the raters, but accidentally found during the consensus meeting. Both raters found additional lesions that were rejected as microbleeds during the consensus meeting: 4 were found by rater 1 and 18 by rater 2.

#### 4.2.4 Preprocessing

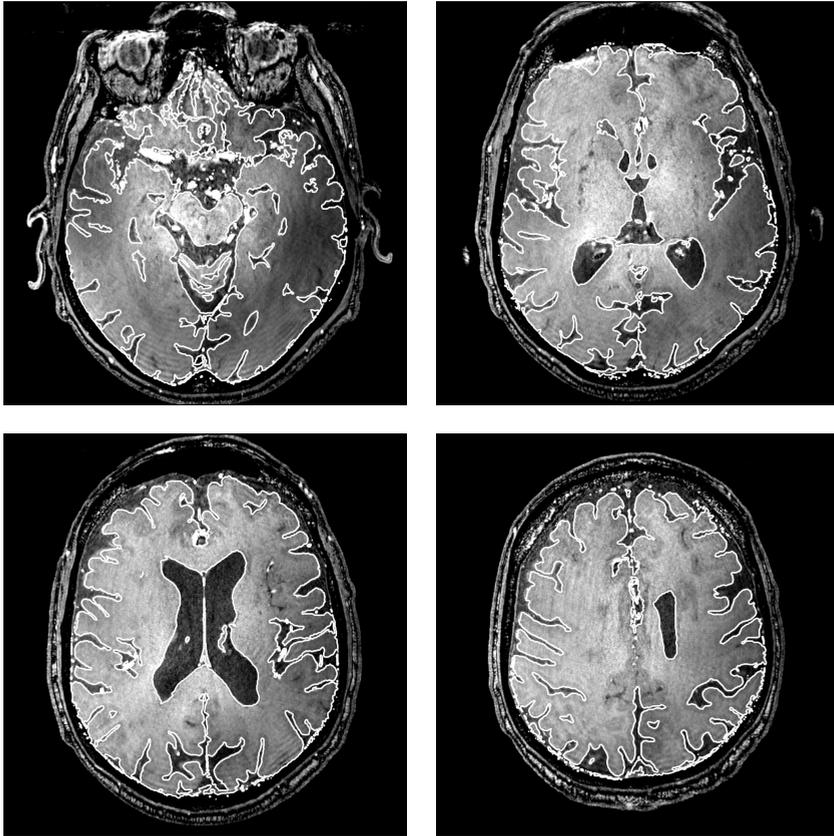
The first two steps in the proposed method are the creation of a binary mask containing the grey and white matter, and intensity normalization. A probabilistic grey and white matter mask was generated from the T1-weighted sequence, using unified segmentation as implemented in SPM8 (Ashburner et al. 2005). The T1-weighted sequence was registered to the first echo of the T2\*-weighted sequence and the resulting transform was applied to the probabilistic mask. A binary mask was created by thresholding the probabilistic mask at a level of 90%. An example is shown in Figure 4.2.

The intensity values of the T2\*-weighted scans were normalized to a range of [0, 255], using the 5th and 95th percentile of the histogram within the mask as lower and upper bound respectively. This standardized the grey values among different scans and will map the centre of visible microbleeds to a grey value of zero.

#### 4.2.5 Radial symmetry transform

For the automatic detection of microbleeds, a 3D version of the original RST described by Loy et al. (2003) was implemented. The RST is a technique that utilizes local radial symmetry to highlight spherical points of interest in an image.

The choice for the RST is derived from the properties and shape of a CMB on the T2\*-weighted MR sequence. The actual hemosiderin deposit is very small and accumulated in one point, as shown in histopathological research of Fazekas et al. (1999) and De Reuck et al. (2011), so it is safe to assume that the hemosiderin deposit is contained within a single voxel of the MR scan. The blooming effect of the microbleed, influenced by the echo time, is of approximately equal strength in all directions, which creates a spherical hypo-



**Figure 4.2:** Example of the binary mask created during preprocessing, where the white contour indicates the boundaries of the binary mask.

intense spot in the scan at the location of the microbleed. This fits perfectly into the concept of the RST.

The transform is computed for all radii  $n \in N$ , where  $N$  is a set of radii of microbleeds that need to be detected. In this study, the radii of CMBs to be detected ranged from 0.3 to 2.0 mm (step size of 0.1 mm), corresponding to the range of CMB radii that were detected by the raters.

Image gradients and orientations are used by the transform to infer the centre of mass of spherical objects in the scan, corresponding to potential microbleeds. The output value at a specific point of the transform with radius  $n$  indicates the contribution to the radial symmetry of the gradients a distance  $n$  from that point. The gradient information is computed using a fast 3D Sobel kernel of size  $3 \times 3 \times 3$  voxels, giving positive values if the gradient direction points from dark to light. Since microbleeds are hypointense, the gradient direction information needs to be negated to have the gradient pointing towards the centre of the microbleeds. This is called ‘negatively-affected’ in the original method by Loy et al. (2003).

At each radius  $n \in N$ , an orientation projection image  $O_n$  and a magnitude projection image  $M_n$  are computed, counting only negatively-affected voxels. The coordinates of a negatively-affected voxel  $p_{-ve}$  are given by

$$p_{-ve}(p) = p - \text{round} \left( \frac{g(p)}{\|g(p)\|} n \right),$$

where  $p$  is a vector with the coordinates of the current voxel,  $g$  is the gradient and ‘round’ rounds each vector element to the nearest integer value. This is illustrated in Figure 4.3.

The orientation and magnitude projection images were initialized at zero. For each voxel  $p$  in the input image, the value at the corresponding point  $p_{-ve}$  in  $O_n$  and  $M_n$  was decremented by 1 and  $\|g(p)\|$ , respectively<sup>1</sup>:

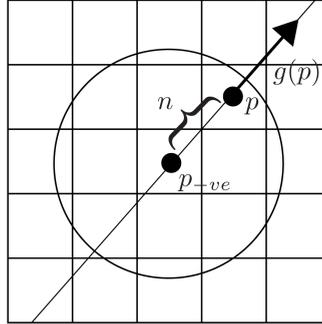
$$O_n(p_{-ve}(p)) = O_n(p_{-ve}(p)) - 1,$$

$$M_n(p_{-ve}(p)) = M_n(p_{-ve}(p)) - \|g(p)\|.$$

For a specific CMB with some radius  $n$ , each gradient on the surface of the CMB will map to the centre of that CMB, resulting in a local extreme in  $O_n$ ,

---

<sup>1</sup>Only negatively-affected voxels were taken into account here, whereas the original method also considers positively-affected voxels. Positively-affected voxels are ignored, as they correspond to hyper-intense spots.



**Figure 4.3:** Given point  $p$ , the negatively-affected voxel  $p_{-ve}$  is found by following the gradient direction  $g(p)$  in the opposite direction for some length  $n \in N$ . This process is repeated for all relevant gradients (controlled by  $\beta$ , see text) in the input image.

indicating the location of the centre of the CMB. As suggested by Loy et al. (2003), small gradients should be ignored during the computation of  $O_n$  and  $M_n$ . To this end, a parameter  $\beta$  was introduced as lower bound and set to 7.5% of the highest gradient magnitude present in the image. This value was found during the preceding proof-of-principle experiment and eliminates noise and small variations in the scan that do not correspond to microbleeds.

The radial symmetry contribution of a single radius  $n \in N$  is given by the convolution

$$S_n = F_n * A_n,$$

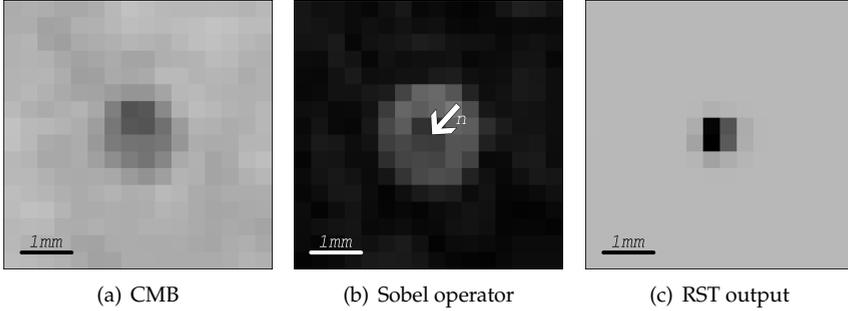
where

$$F_n(p) = \frac{M_n(p)}{k_n} \left( \frac{|\tilde{O}_n(p)|}{k_n} \right)^\alpha$$

and

$$\tilde{O}_n(p) = \begin{cases} O_n(p) & \text{if } -k_n < O_n(p) \\ -k_n & \text{otherwise.} \end{cases}$$

In this equation,  $A_n$  was a 3D Gaussian used to smooth the output of the transform. The properties of this Gaussian were described by Loy et al. (2003) and is adjusted for each  $n$  as suggested. The used  $\alpha$  is the radial strictness parameter, defining how strictly spherical the transform must be to return a high value. Its value was taken from our preceding proof-of-principle experiment (Kuijff et al. 2011), in which the optimal settings were investigated, and



**Figure 4.4:** Output of the Sobel operator and the RST on the first echo of the microbleed, as shown in Figure 4.1. The  $n$  in 4.4(b) shows a possible radius from the set  $N$ . When using smaller or larger  $n$ , the radial symmetry contribution does not sum up in a single voxel.

was set to 3. The parameter  $k_n$  is a scaling factor used to normalize  $O_n$  and  $M_n$  across different radii. When the radius increases, the number of gradient elements that can affect a voxel will also increase, since this can reach the number of voxels on the surface of the microbleed. The value for  $k_n$  was estimated empirically, as suggested by Loy et al. (2003), by inspecting the minimum value of  $O_n$  at microbleed locations. An expected minimum value of  $O_n$  was determined at each radius  $n \in N$ .

The resulting transform was computed as the sum over all radial symmetry contributions:  $S = \sum_{n \in N} S_n$ . The transform was computed separately per echo, resulting in an output for the first echo ( $S_1$ ) and the second echo ( $S_2$ ). Computation of the RST was performed on the whole scan and the result was masked afterwards with the grey and white matter mask obtained during preprocessing. An example output for a typical microbleed is shown in Figure 4.4.

As a final step,  $S_1$  and  $S_2$  had to be combined to generate a list of microbleed locations. Non-minimum suppression within a 2 mm neighbourhood was applied on both  $S_1$  and  $S_2$  to remove multiple minima close to each other. An upper-threshold was determined to extract the minima in  $S$  that might correspond to microbleeds. Separate upper-thresholds  $\theta_1$  and  $\theta_2$  were applied to  $S_1$  and  $S_2$ , which were set at  $-20$  and  $-100$ , respectively. These thresholds were different from the preceding proof-of-principle experiment, partly because we extended the set of input scans on which the thresholds were determined heuristically and partly because the set of radii  $N$  differed.

If within a 26-voxel neighbourhood around a minimum in  $S_1$  a minimum in  $S_2$  exists, this location is considered a potential microbleed. Finally, if the normalized grey value at the second echo of a potential microbleed was not equal to zero, it was removed from the list, since a microbleed should cause a signal void on the  $T_2^*$ -weighted images.

#### 4.2.6 Experiments

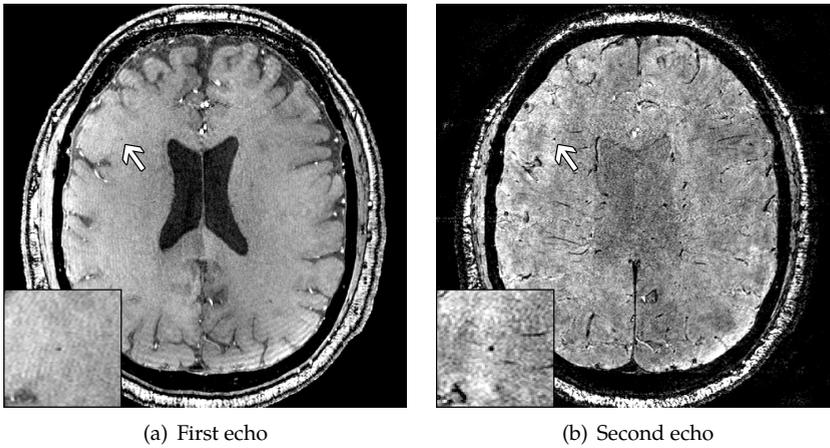
In the experiments that were performed, the number of false positives present in the results was assessed together with the time to visually censor the false positives. A dedicated tool for this was developed using MeVisLab (MeVis Medical Solutions AG, Bremen, Germany). Two raters, rater 3 and 4, who were not involved in the original visual rating to ensure that they were not biased, each censored all potential CMBs identified by the RST. A consensus meeting was held for the cases on which the two raters disagreed.

When CMBs were detected by the RST that were not present in the original visual rating, they were added to the ground truth if raters 3 and 4 had confirmed them as true microbleeds. Adding the CMBs confirmed by raters 3 and 4 improved the ground truth rating and gave a fair comparison of the sensitivities of human raters versus the RST.

### 4.3 Results

The implemented 3D RST algorithm detected in total 353 potential microbleed locations in the 18 participants, on average  $19.6 \pm 12.1$  (mean  $\pm$  sd) microbleeds per participant. Among these locations, 35 true positives (present in the original visual rating) were found, 309 false positives, and 12 extra positives that were not present in the original visual rating, but were marked as potential CMB by the RST and confirmed by raters 3 and 4. Rater 3 required 43 min to censor all detected microbleeds ( $2.5 \pm 1.5$  min per participant;  $7.9 \pm 3.3$  s per detected location), rater 4 28 min ( $1.5 \pm 1$  min per participant;  $5.1 \pm 2.2$  s per detected location). Computation of the RST took about one hour per participant, using one core of a standard workstation.

In total, 66 microbleeds were detected (54 in the original visual rating after consensus, 12 extra positives were found by the RST), of which 34 were detected by rater 1, 44 by rater 2, 53 by both raters, and 47 by the RST (see Table



**Figure 4.5:** A typical microbleed that is detected by the RST.

4.1). The single microbleed found only during the consensus meeting was not taken into account.

Sensitivities of the individual raters and the automatic detection are reported in Table 4.2. The RST had a sensitivity of 71.2%, and the individual observers had a sensitivity of 51.5% (rater 1) and 66.7% (rater 2). The rating after the consensus meeting of raters 1 and 2 had a sensitivity of 80.3%.

Some typical examples of the method, including true positives, false positives, and false negatives are shown in Figures 4.5–4.8. Figure 4.5 shows a typical microbleed present in one of the participants, which is detected by the RST. Another true positive is shown in Figure 4.6, where the CMB is adjacent to a vessel. Since there is still enough radial symmetry contribution from the sides of the CMB that are not touching the vessel, detection is not a problem for the RST. False positives were mainly caused by large blood vessels, as shown in Figure 4.7. The anatomical nature of vessels can sometimes generate significant radial symmetry contribution at highly curved regions to be detected by the RST as a CMB, especially with large vessels that have many contributing gradients on the surface. Figure 4.8 shows a typical false negative caused by the microbleed not being visible on the first echo time image.

**Table 4.1: Upper part:** overview of the microbleeds detected during visual rating. **Middle part:** overview of the microbleeds detected by the RST. **Bottom part:** final overview of all detected microbleeds.

Participant:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Total
Visual <sup>a</sup>	Number of microbleeds detected																		Total
R1&R2	2	0	0	1	0	2	0	0	0	0	10	0	4	2	1	1	1	1	25
R1	1	0	0	0	0	0	0	0	3	0	2	0	0	0	0	0	2	1	9
R2	1	2	1	0	1	0	1	2	0	1	3	1	1	1	4	0	0	0	19
C	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Total (VT)	4	2	1	1	1	2	1	2	4	1	15	1	5	3	5	1	3	2	54
Automatic <sup>b</sup>	Number of locations found																		Total
Total	20	47	13	24	21	30	8	11	27	36	18	7	11	37	18	16	3	6	353
TP	2	2	1	1	1	0	1	2	3	1	9	1	5	3	2	0	1	0	35
EP	1	0	0	0	0	4	0	0	1	0	1	0	1	0	1	0	1	2	12
FP	17	45	12	23	20	26	7	9	23	35	8	6	5	34	15	16	1	4	306
Final <sup>c</sup>	Number of microbleeds																		Total
VT+EP	5	2	1	1	1	6	1	2	5	1	16	1	6	3	6	1	4	4	66
FN	2	0	0	0	0	2	0	0	1	0	6	0	0	0	3	1	2	2	19

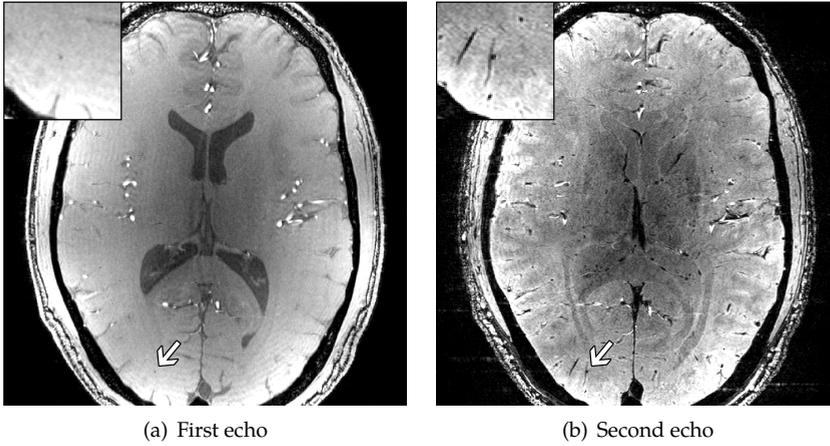
<sup>a</sup> R1&R2: CMBs detected by both rater 1 and 2. R1, R2: CMBs detected exclusively by rater 1 or by rater 2. C: additional CMBs detected at consensus meeting.

<sup>b</sup> Total of detected locations; divided in true positives (TP), extra positives (EP), and false positives (FP). Extra positives were CMBs detected by the RST, but not during the initial visual rating.

<sup>c</sup> Final total combining visual total (VT) and extra positives (EP), together with the false negatives (FN).

**Table 4.2:** Sensitivity of the individual raters, both raters combined, and the automatic detection. A total of 66 microbleeds was present in the images.

	#TP	#FP	#FN	Sensitivity
Rater 1	34	4	32	51.5 %
Rater 2	44	18	22	66.7 %
Rater 1+2	53	0	13	80.3 %
Automatic	47	309	19	71.2 %

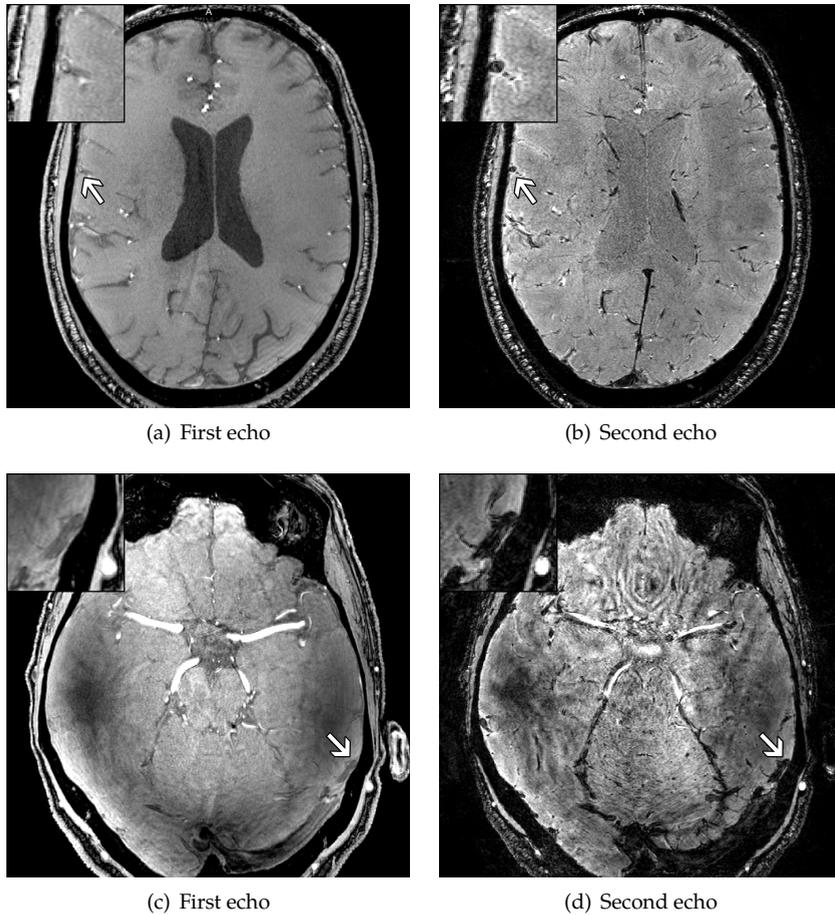


**Figure 4.6:** A microbleed detected by the RST, where the microbleed is adjacent to a vessel.

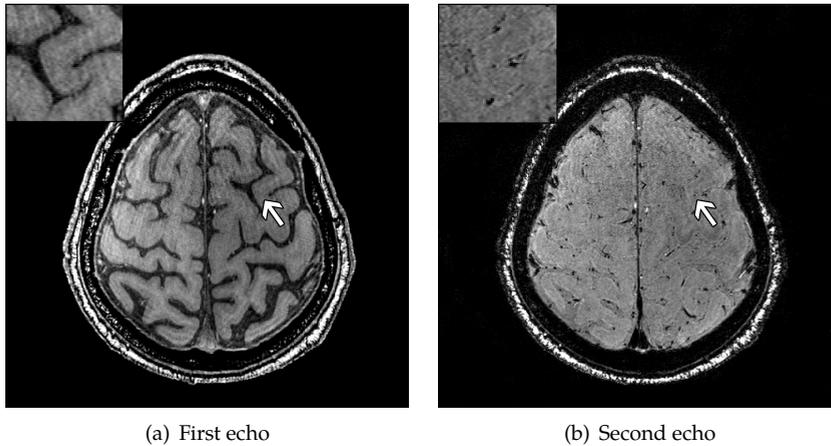
## 4.4 Discussion

Detection of cerebral microbleeds can be achieved by the radial symmetry transform on 7.0 T MRI scans, with a high sensitivity and a limited number of false positives. The semi-automatic method showed a large reduction in rating time: from 30 to 2 min per participant. Most importantly, the method outperformed individual raters in terms of sensitivity and even detected microbleeds that were missed in the visual rating.

We tried to optimize human rating so as to arrive at a reliable reference standard. However, this was only successful to a limited extent, because of the fundamental problems that arise during microbleed rating on high-resolution 7.0 T MR images. 7.0 T MR imaging is known to have a higher sens-



**Figure 4.7:** Two typical false positives, in two different participants, that were detected by the RST as microbleeds. Both involved large blood vessels, where false positives typically occur at curved sites within the vessel.



**Figure 4.8:** A typical false negative, where the microbleed is not visible on the first echo.

itivity for microbleed detection than scans acquired at lower field strength (Conijn et al. 2011). However, the limited size of some of the lesions and susceptibility effects of adjacent other structures complicates detection, which leads to a modest sensitivity of the visual rating on 7.0 T MRI scans (de Bresser et al. 2013). The described sensitivities of raters 1 and 2 are in line with sensitivities reported by De Bresser et al., who found a mean sensitivity of 60% (range: 44 to 83%). De Bresser et al. also reports higher sensitivities of the same raters when performing visual rating on 1.5 T MR scans, which indicates that the reduced sensitivity is caused by the increased difficulty in rating at 7.0 T MR scans.

Visual rating of CMBs on 7.0 T MR images will not result in a ground truth rating that can be considered as a golden standard, but rather as a “silver” standard. The original visual rating was combined with the extra positives detected by the RST to construct a ground truth rating with maximum sensitivity, even though this will introduce a bias in the sensitivities in favour of the RST. The extra positives were confirmed as true microbleeds by two raters independently. Visual rating by human raters can not guarantee that 100% of all present CMBs have been detected, thus adding CMBs found by other means to the ground truth rating will improve the ground truth.

Nineteen microbleeds that were detected by the initial human raters were missed by the RST. Of these false negatives, 7 were not visible on the first echo

and 4 were not visible on the second echo, which was a prerequisite for our variant of the RST to detect CMBs. The microbleeds that were not visible on the first echo were either completely absent or too small in size to have any response to the used Sobel kernel. Their mean diameter on the second echo, in which CMBs appear larger than on the first echo owing to the blooming effect, was  $0.78 \pm 0.18$  mm. This makes it difficult for human raters to detect those microbleeds and in most other studies such small microbleeds were even ignored (as indicated in the review study by Cordonnier et al. (2007)). The microbleeds that were invisible on the second echo disappeared in the blooming effect, caused by vessels nearby, and had a radial symmetry value far above the threshold  $\theta_2$ . Work by Conijn et al. (2010) showed that 15.4% of microbleeds were not visible on the first echo and 1.9% were not visible on the second echo. As the missing microbleeds on the second echo were relatively large and thus had a radial symmetry value far below the threshold  $\theta_1$  on the first echo, there should be a possibility to detect them. This will be investigated in future work.

The remaining eight false negatives were caused by various problems. Artefacts visible in one T2\*-weighted image (acquired in 2008) caused some false negatives. Advances in the quality of 7.0 T MR imaging might improve this in the future. In the remaining cases, there was not enough contribution of the microbleeds to the radial symmetry in order to be detectable. A total of 7 false negatives were scored as possible microbleeds by raters 1 and 2, indicating that they were not sure about the lesion being a definite microbleed or something else. The lack of radial symmetry at these locations might possibly indicate that these lesions were not microbleeds and should have been ignored.

During preprocessing, a binary mask of the grey and white matter was created using SPM8, but another segmentation method can be used as well. Completely removing all cerebrospinal fluid by setting a lower-threshold of 90% on the probability-mask produced by SPM8 is required to reduce the number of false positives. The sulci present in the brain can have a large radial symmetry value, especially in the far end of a sulcus, resulting in false positives.

Computation time of the RST was about one hour per participant, using one core of a standard workstation. However, the computation of the radial symmetry contribution of multiple radii could be parallelized, which e.g. would reduce the computation time to 15 min on a typical quad-core

workstation.

The RST performs excellently in terms of sensitivity and false positives, in comparison with the other two semi-automated CMB detection methods published so far, by Seghier et al. (2011) and Barnes et al. (2011). Our study and these previous studies clearly differ in MR acquisition protocols and field strengths. Nevertheless, the method described by Seghier et al. identified 77% of all patients, while the RST identified 94% of the participants. The participant that was not identified by the RST had a single microbleed that was rated 'possible' and not 'definite' by the human raters. As numbers on false positives were not provided, only the time required to remove the false positives could be compared. The method of Seghier et al. required on average 5 to 10 min time to remove the false positives, compared to 1.5 to 2.5 min for the RST, which is likely due to a high number of false positives in the method of Seghier et al. The method described by Barnes et al. had a sensitivity of 81.7%, which is higher compared to the sensitivity of 71.2% for the RST. However, the reported number of false positives in the method of Barnes et al. is much higher: on average 107.5 per patient compared to 17.2 per participant for the RST. This explains the longer rating time required after processing for the method of Barnes et al. The higher sensitivity might be explained by the differences in size of the microbleeds found in the patients of Barnes and the participants in this study. Although not precisely specified, Barnes et al. reported microbleed sizes of <5 to 10 mm. In this study, the microbleeds visible on the first echo of the used T2\*-weighted MR scan range from 0.3 to 2.0 mm and seven microbleeds were not visible at all. Since the RST does not produce many false negatives on larger microbleeds, this method might perform better when searching only for larger microbleeds. The latter was commonly done in most studies (Cordonnier et al. 2007).

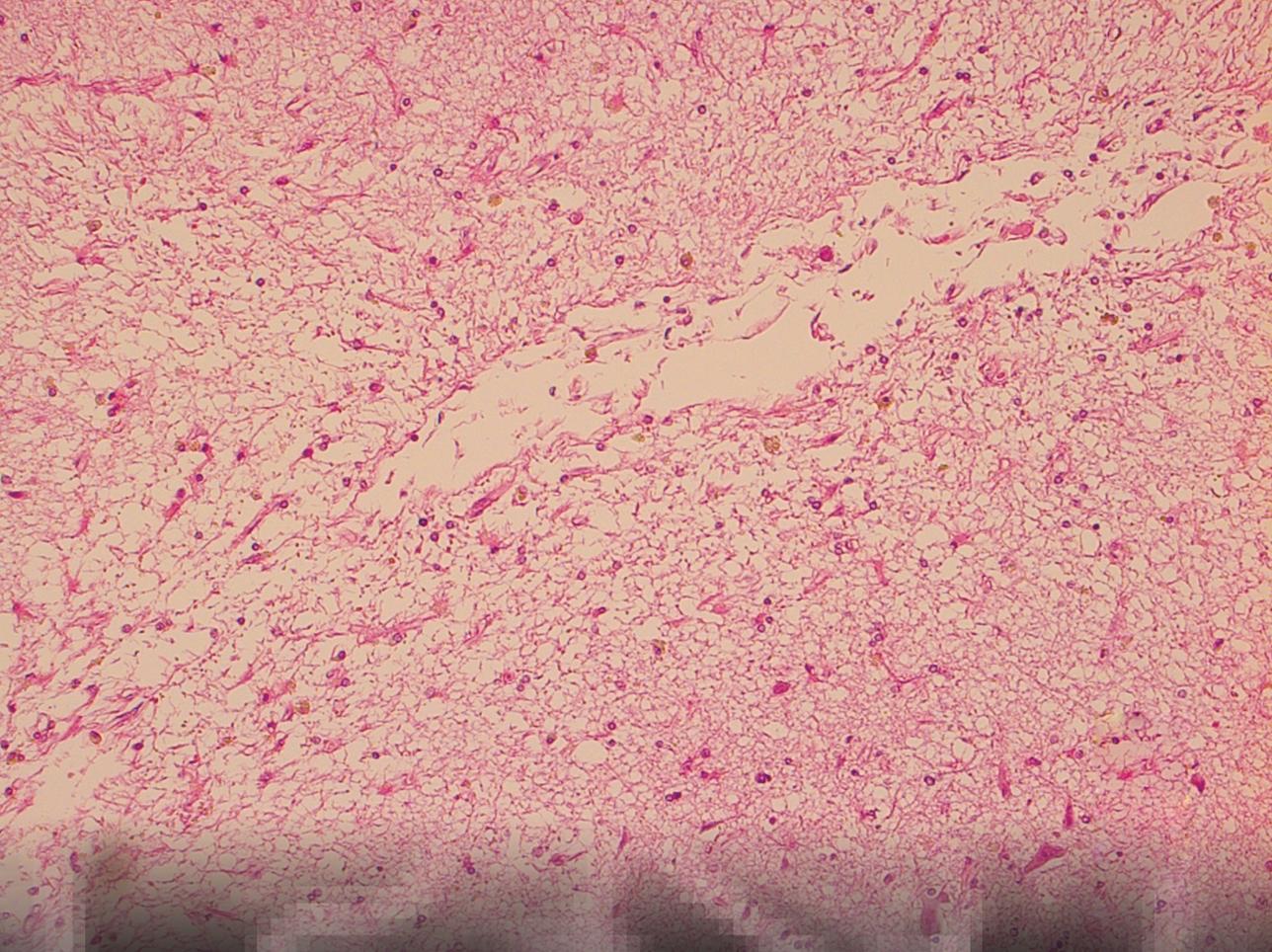
Applying the RST to an anisotropic scan acquired at lower field strength is also possible. The individual components of the gradient direction are scaled with the voxel size, to account for anisotropic voxels. However, the voxel size gives rise to a lower limit on the detection of CMBs, since it is not possible to detect CMBs of only one or a few voxels in size. There would hardly be any radial symmetry contribution for microbleeds that consist of so few voxels. This could be counteracted by adapting the thresholds, which can result in additional false positives and more censor-time afterwards, or increasing the echo time, resulting in a larger apparent size of the CMBs. With an increase in apparent size, CMBs could consist of more voxels and might be detectable

by the RST.

The RST outperforms individual human raters in terms of sensitivity and required human rater time. Combining the rating of two individual raters in a consensus meeting will result in the highest sensitivity, but still not 100%, and takes even more time. This indicates the difficulty in the manual rating of cerebral microbleeds in high-resolution MR scans. In daily practice, a scan is usually rated by a single rater, since time constraints do not permit multiple raters and a consensus meeting. Semi-automatic detection of microbleeds with the RST will decrease rating time and increase the sensitivity significantly.

## **4.5 Conclusion**

The radial symmetry transform can be used to detect cerebral microbleeds on high-resolution, nearly isotropic brain MR scans with a high sensitivity. Removing the false positives requires 2 min by an experienced rater, as compared with 30 min for full visual rating of the cerebral microbleeds.



# Detecting cortical cerebral microinfarcts in 7.0 T MR images

**Published as:** H.J. Kuijf, F. Zijlstra, S.J. van Veluw, M.A. Viergever, K.L. Vincken, G.J. Biessels, "Detecting Cortical Cerebral Microinfarcts in 7.0 T MR Images", in: *IEEE International Symposium on Biomedical Imaging*, 2013, pp. 982-985.

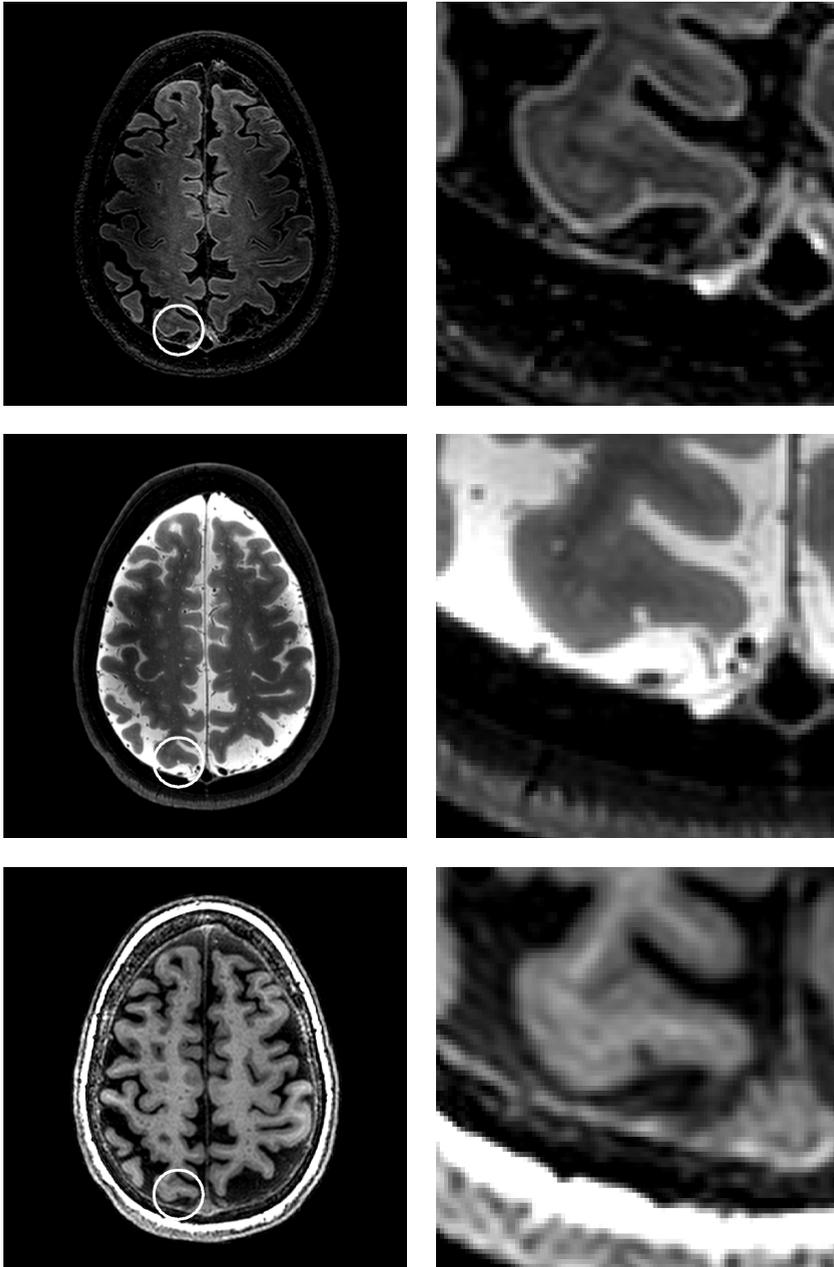
### Abstract

Known as the “invisible lesions”, cerebral microinfarcts have been attracting increased attention because of their key role in cognitive decline and dementia. Recently, cerebral microinfarcts have been visualized for the first time in vivo on high resolution 7.0 T MR images. The detection and scoring of microinfarcts requires extensive manual evaluation, is very time-consuming, and highly observer-dependent. The goal of this proof of principle study is to present a technique for semi-automated detection of microinfarcts with a high sensitivity and a low number of false positives, since the removal of these requires additional efforts by a human observer.

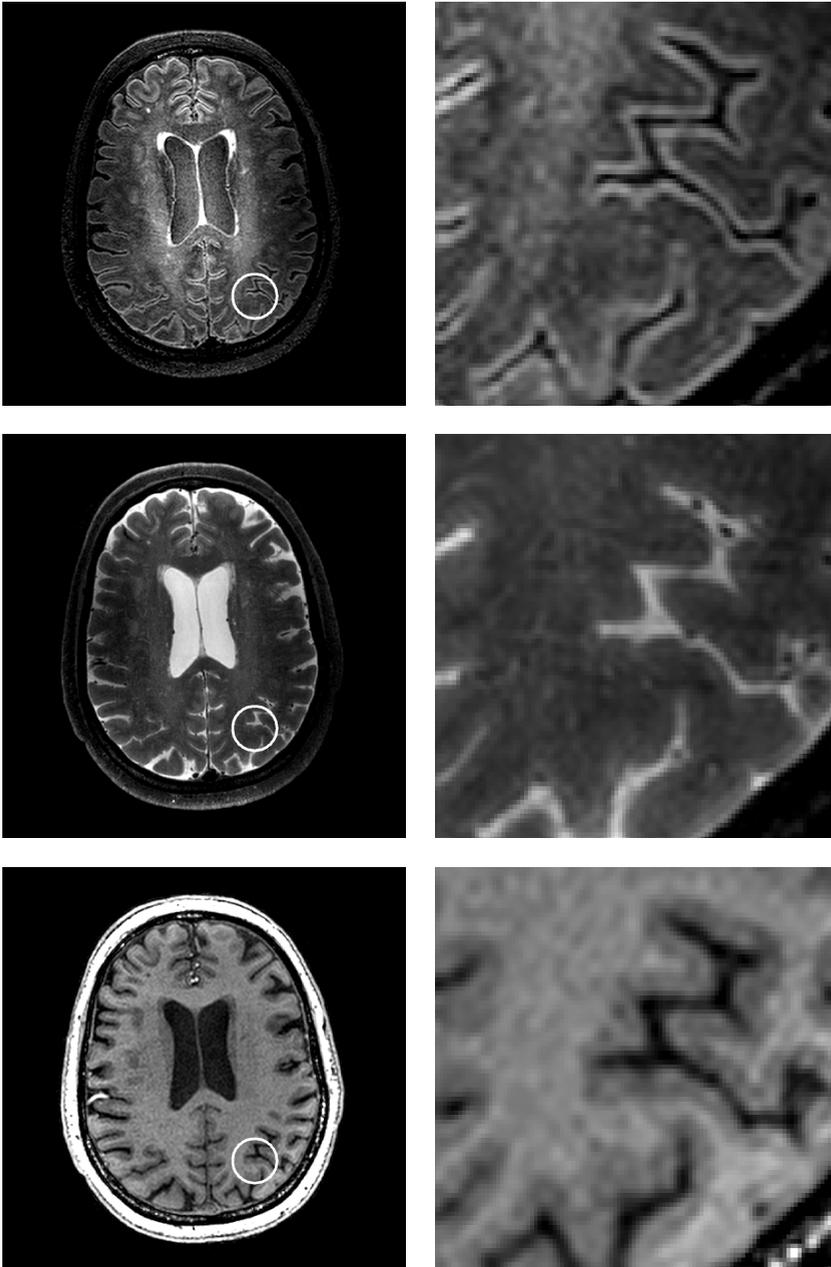
## 5.1 Introduction

Recently, cerebral microinfarcts (CMIs) have received increasing attention, because they may play a key role in ageing and dementia. CMIs are the most widespread form of brain infarction and have been identified as hallmarks of small vessel disease (Smith et al. 2012). CMIs have been studied extensively in autopsy studies. In such studies they are commonly defined as ‘not visible to the naked eye’. Upon microscopic examination they appear as delineated regions of tissue loss, with gliosis and sometimes cavitation. Estimated sizes vary between 50  $\mu\text{m}$  and 4 mm (Brundel et al. 2012a). Owing to their small sizes, they escape detection on conventional MRI. Recently, cortical CMIs have for the first time been visualized on high resolution 7.0 T MR images (van Veluw et al. 2013b). In that study CMIs were found to be hyperintense on both FLAIR and T2 weighted images, as can be seen in Figures 5.1 and 5.2. The in vivo findings were validated with high resolution ex vivo MRI on post-mortem brain tissue, providing histopathological confirmation that the lesions were indeed CMIs. The detection and scoring of cortical CMIs requires extensive manual evaluation, is very time-consuming, and highly observer-dependent.

In this study, we have performed a proof-of-principle experiment to investigate the possibility of detecting all potential CMI locations from an MR brain scan. The ultimate goal is to develop a system that has a high sensitivity, while minimizing the number of false positives that have to be censored manually.



**Figure 5.1:** Left column: CMI detected in subject 1, shown on transversal views of 7.0T FLAIR (top), T2 (middle), and T1 (bottom) images. Right column: enlarged visualization of the CMI.



**Figure 5.2:** Left column: CMI detected in subject 2, shown on transversal views of 7.0 T FLAIR (top), T2 (middle), and T1 (bottom) images. Right column: enlarged visualization of the CMI.

## **5.2 Materials and Methods**

### **5.2.1 Subjects**

For this study, two female subjects (subjects 1 and 2) with CMIs were included. Both subjects were 66 years old, with a Mini-Mental State Examination score of 30 (Folstein et al. 1975). Subject 2 had type 2 diabetes mellitus. The study was approved by the medical ethics committee of the UMC Utrecht and written informed consent was given by the subjects.

### **5.2.2 MRI**

MRI scans were acquired on a whole-body 7.0 T MR system (Philips Healthcare, Cleveland, OH, USA). The protocol included, amongst others, a 3D FLAIR (0.8 mm<sup>3</sup>, TR=8000 ms, TE=300 ms, TI=2325 ms, scan duration 12:48 min) (Visser et al. 2010), a 3D T2-weighted (0.7 mm<sup>3</sup>, TR=3158 ms, TE=301 ms, scan duration 10:15 min) (Wisse et al. 2012), and a 3D T1-weighted (1.0 mm<sup>3</sup>, TR=4.8 ms, TE=2.2 ms, TI=1240 ms, scan duration 1:57 min) sequence. On the same day, the subjects also underwent a 3.0 T MRI (Intera, Philips, Best, the Netherlands), which included, amongst others, a 3D T1-weighted (1.0 mm<sup>3</sup>, TR=7.9 ms, TE=4.5 ms, TI=955 ms) sequence. Examples of the 7.0 T scans are shown in the top rows of Figures 5.1 and 5.2.

### **5.2.3 Visual rating of microinfarcts**

Cortical CMIs were defined as small circumscribed lesions in cortical grey matter that are hyperintense on FLAIR and T2 images, and hypointense on T1. This can be seen in Figures 5.1 and 5.2. Sagittal views of the FLAIR and T2 scans were thoroughly examined by an experienced human observer to identify possible cortical CMIs and the findings were compared to the T1-weighted scans. Potential cortical CMIs had to be smaller than 4 mm, and not form a direct connection with perivascular spaces (PVS) in the white matter. A second observer evaluated all identified lesions. In the case of disagreement, the decision was made by consensus.

### **5.2.4 Detection of microinfarcts**

The method consists of three parts: segmentation of cortical grey matter, detection of potential CMI locations, and rejection of false positives. After these

steps, a human observer assessed the results.

### Segmentation

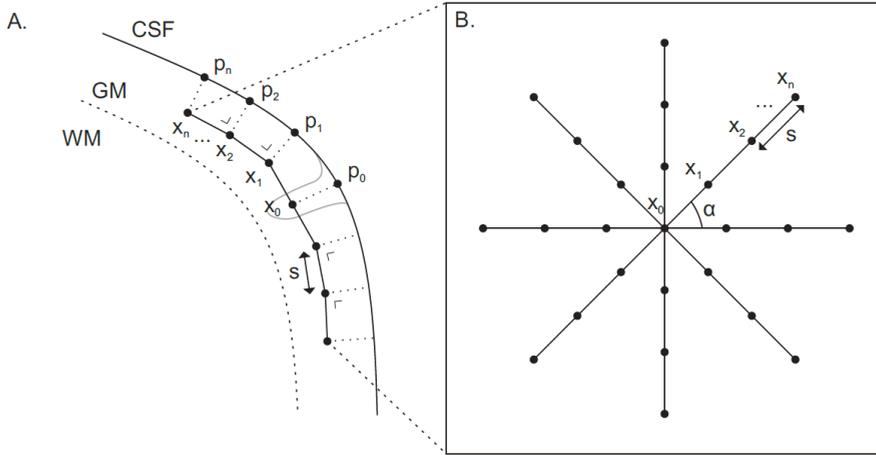
Segmentation of cortical grey matter was performed with FreeSurfer on the 3.0 T T1-weighted scan, resulting in two meshes: of the pial and white matter surfaces (Dale et al. 1999). From this segmentation, the temporal lobes were ignored for lack of signal on the 7.0 T images. Using elastix (Klein et al. 2010), the 3.0 T T1-weighted scan was registered to the 7.0 T T1-weighted scan. Next, the 7.0 T T1-weighted scan was registered to the 7.0 T FLAIR scan. The results were used to transform the surface meshes of the grey matter to the 7.0 T FLAIR scan. The 7.0 T T2-weighted scan was also registered to the 7.0 T FLAIR scan.

To improve the segmentation of cortical grey matter on the 7.0 T FLAIR scan, the mesh of the pial surface was locally deformed to align with the hyperintense rim (corresponding to the upper layer(s) of the cortex) that is visible on the 7.0 T FLAIR scan (Zwanenburg et al. 2009).

### Detection of potential CMI locations

First, intensity values of the 7.0 T FLAIR and T2 scans within the segmented grey matter regions were normalized to correct for the RF field inhomogeneity of the 7.0 T scanner. This was done by dividing the original images by a smoothed median filtered version of the image (2.5 mm<sup>3</sup> kernel for median filtering and 5.0 mm<sup>3</sup> kernel for smoothing the median image), resulting in a unit mean in the grey matter.

Next, a method was developed to detect potential CMI locations. CMIs are hyperintense structures in grey matter on FLAIR and T2 scans with an approximate cylindrical shape orthogonal to and attached to the pial surface of the brain (see Figures 5.1 and 5.2). Accordingly, a (curved) cross-section parallel to the pial surface should contain a detectable local maximum at the location of a CMI and not at non-CMI locations. This cross-section was sampled as a 2D curved circular plane. This procedure is illustrated in Figure 5.3. For each sampled line, the median intensity of all sampling points (locations  $x_{1..n}$  in Figure 5.3) should be low compared to the intensity at the centre (location  $x_0$  in Figure 5.3). To this end, the highest median of all sampling lines was determined (that should be low since the tissue surrounding a CMI should have low intensity values) and subtracted from the input image (having a high



**Figure 5.3:** A) Schematic drawing of the cortical grey matter, with a CMI at location  $x_0$ . B) The complete sampling pattern of the curved cross-section shown in A. The angle  $\alpha = 11.25^\circ$ , resulting in 32 lines. The step size  $s = 0.5$  mm and the number of sampling points  $n = 5$ .

value at CMI locations). This was stored in a result image  $Y$  that contained high values at the location of CMIs and low values (near zero) elsewhere.

The previous step was computed on the FLAIR and T2 images separately. The T1 image was ignored, since CMIs are not always clearly visible on this sequence, as can be seen in Figure 5.2. Both results were combined into a single image  $Z = Y_{FLAIR} + w * Y_{T2}$ . The weighting factor  $w$  was determined experimentally. All values in  $Z$  above an experimentally determined threshold  $z$  were considered potential CMI locations.

### Rejection of false positives

Among the potential CMI locations, there will be a number of false positive locations that should be removed before presenting them to a human observer.

First of all, PVS visible on the T2 image that originate from white matter and continue into the cortex should be removed. This was achieved by performing an iterated greyscale geodesic dilation from deep white matter on the T2 images. If a PVS grew into the cortex, that cortex location was no longer considered a possible CMI location.

Next, locations that were above the threshold  $z$  in  $Z$ , but only exist in either  $Y_{FLAIR}$  or  $Y_{T2}$  were removed, because a CMI should be visible in both.

Hysteresis reconstruction was used to reconstruct the full CMI around detected locations and each CMI less than four voxels was removed.

Finally, at each remaining potential CMI location, an intensity profile orthogonal to the pial surface was inspected. If the potential CMI location was too close to the CSF ( $< 0.5$  mm) or disconnected from the pial surface, it was rejected.

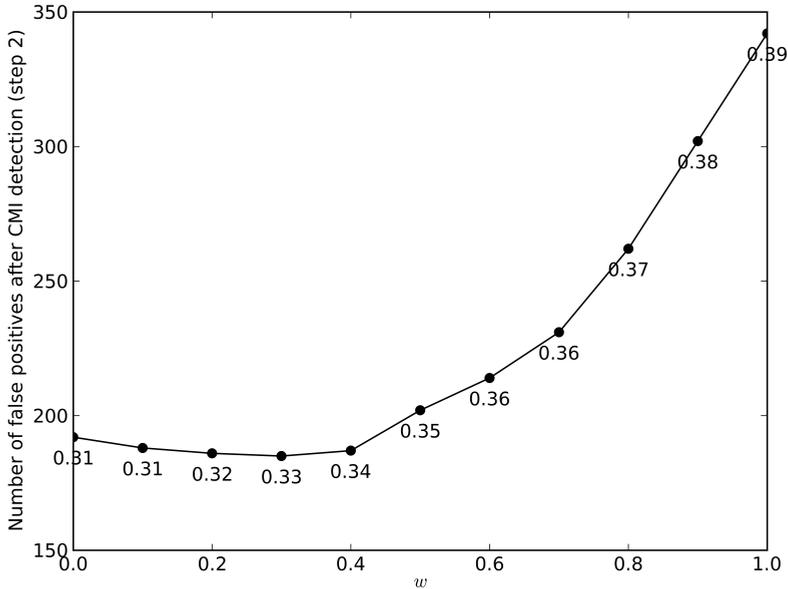
### 5.2.5 Experiments

Two parameters needed to be optimized: the weighting factor  $w$  and the threshold  $z$ . Possible values for  $w$  were taken from the range  $[0.0, 1.0]$  with a step size of 0.1. For each  $w$ , the optimal threshold  $z$  was determined automatically to extract all locations that might correspond to CMI locations. As the location of true CMIs was known, it was possible to ensure that the applied threshold preserved the true CMIs in  $Z$  for both subjects.

The optimal setting of  $w$  and  $z$  (having the least number of false positives) was selected for inspection by a human observer. For each detected location, the observer visually annotated whether it was a definite CMI. The observer was blinded for the original rating of the scans and the time it took to censor the false positives was recorded.

## 5.3 Results

During the visual rating of CMIs, one CMI was found in each subject. These CMIs are shown in Figures 5.1 and 5.2. Visual rating required about one hour of observer time per subject. The quality of segmenting grey matter by FreeSurfer was compared to manual segmentations of the full brain of two other subjects, not included in this study, but who received the exact same 3.0 T MR examination. Dice's similarity coefficient (DSC) on grey matter was 0.84 for these two subjects. For the two subjects included in this study, manual segmentations of grey matter were made on one slice of the 7.0 T T1-weighted sequences. DSC of the segmentation of the grey matter on these slices after registration was on average 0.77 for the two subjects. The parameters  $w$  and  $z$  were optimized after the second step of the proposed method (the part described in §5.2.4 *Detection of potential CMI locations*). The result of this optim-



**Figure 5.4:** Graph showing the optimal threshold values of  $z$  (the data labels) that ensure maximal sensitivity for a given  $w$ , together with the number of false positive locations.

ization is shown in Figure 5.4. Based on these results,  $w = 0.3$  and  $z = 0.33$ , resulting in a total of 185 false positives. After the third step (§5.2.4 *Rejection of false positives*), a total number of 13 potential CMI locations were detected in subject 1 and 49 in subject 2.

Visual inspection of the detected potential CMI locations required 5 min and 20 min for subjects 1 and 2, respectively. In subject 1, the true CMI was identified by the observer and all other locations were censored. In subject 2, the true CMI was identified by the observer, 5 additional CMIs were identified as true CMIs, and the other locations were censored.

## 5.4 Discussion

This study presents a semi-automated method for detecting CMIs. To our knowledge, this is the first such method aimed at CMI detection. Given the

possible role of CMIs in age-related cerebral diseases, the large variety of locations in the brain at which a CMI can occur, and the high difficulty of visually rating CMIs, the method has potential for clinical usage. Not only has a large reduction in rating time as compared with full visual detection been shown, but the method also appears capable of identifying potential CMIs that were missed by the human observer, but were recognized as true CMIs upon evaluation.

The main limitation of this proof-of-principle study is the low number of subjects that was included to develop and test the proposed method. In light of the highly promising results, we will continue to develop this technique in much larger study samples.

The current standard for detection of CMIs is visual rating. Although visual rating should result in a so-called “gold standard”, this is not the case for detecting CMIs on high resolution 7.0 T scans. Owing to the large difficulty of visual rating, a large number of CMIs was missed by the observers. Visual rating was performed by inspecting grey matter on sagittal slices. CMIs that were oriented parallel to the viewing direction were often missed and very subtle CMIs were only visible after zooming in on a specific location.

The five additional CMIs that were found by the automated method were verified as true CMIs by the observer who served as evaluator. This might cause a bias in the results in favour of the method, but more importantly it demonstrates the high difficulty of visual rating. Future work should include visual ratings by more than one observer and repeated ratings by the same observer(s) to overcome this bias. When more subjects are included, a larger test set will become available and this will allow a separation of training and testing data. It should be noted that visualization of CMIs is still in its infancy and that the criteria for detection of CMIs are work in progress. The process of establishing robust criteria for the detection of CMIs will benefit from all detected potential CMI locations identified by the proposed method.

Besides having detected extra CMIs, the method returned a number of false positive locations. Most false positives resemble CMIs, but were censored based on expert knowledge of the human observer. The occurrence of such look-alikes is common for similar small pathologies, such as microbleeds, which are seen on 7.0 T scans. These occurrences increase the number of false positives identified by automated methods, since they can hardly be distinguished from true pathology based on image information only (Kuijff et al. 2012a). Some reasons for censoring include: flow artefacts of nearby

vessels that resemble a CMI, the presence of other imaging artefacts or noise, segmentation errors in small sulci, or vagueness of the potential CMI location on the T2 image. The last issue might be resolved by including more training data and adapting  $w$  accordingly.

The technique described in §5.2.4 *Detection of potential CMI locations* showed strong responses at the location of CMIs. The use of a local (curved) coordinated system to detect the hyperintense CMIs is beneficial, but is dependent on the availability of a high quality surface mesh. While FreeSurfer produced good results on 3.0 T images, the low resolution of this scan posed some difficulties in transforming the surfaces to 7.0 T. The applied non-rigid deformation to align the pial surface with the hyperintense rim visible on the FLAIR image improved the segmentation to a great extent.

Currently the method works locally on a 2D circular plane, but more 3D information should be included to improve the results. Inspecting intensity profiles orthogonal to the pial surface during the rejection of false positives (§5.2.4 *Rejection of false positives*) was a first step in this direction.

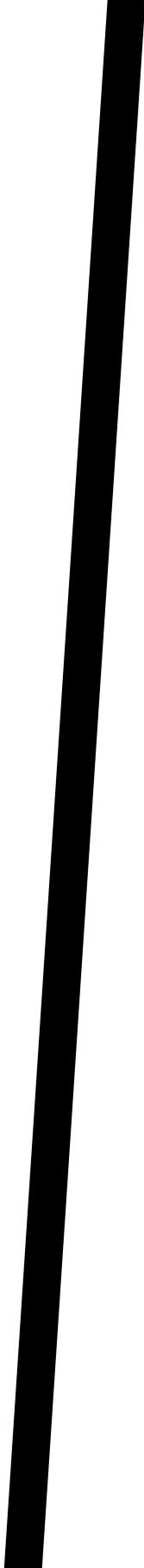
## 5.5 Conclusion

In this proof-of-principle study, we have shown that it is feasible to semi-automatically detect CMIs on high-resolution brain MR images. The reduction in rating time off our method (12 min) compared with visual rating (60 min) is considerable. Maybe even more important is that our method was able to detect CMIs that were missed during visual rating, but were confirmed as true CMIs by a human observer.



## **Part III**

# **Automatic image processing**



# Assessment of methods to extract the midsagittal plane from brain MR images

**Published as:** H.J. Kuijf, A. Leemans, M.A. Viergever, K.L. Vincken, “Assessment of methods to extract the mid-sagittal plane from brain MR images”, in: *SPIE Medical Imaging*, 2013, vol. 8673, p. 86731K.

### Abstract

Automatic detection of the midsagittal plane, separating both hemispheres of the brain, is useful in various applications. Several methods have been developed in the past years, applying different techniques to estimate the position of the midsagittal plane.

These methods can be classified into three distinct classes: feature-based, global symmetry based, and local symmetry based methods. Feature-based methods use the shape or intensity of the interhemispheric fissure to extract the midsagittal plane. Global symmetry based methods reflect the entire image with respect to the sagittal axes and perform a rigid registration. Local symmetry based methods try to optimize a symmetry-measure in a small band covering the interhemispheric fissure.

From each class, one leading method has been implemented. The methods have been evaluated on the same datasets to allow a fair comparison. Manual delineations were made by two experienced human observers.

The results show that the examined methods perform similar to human observers. No significant differences were found between errors (defined as the angle and volume between planes) made by the methods and the inter-observer differences. Feature-based and local symmetry based methods have a low computation time of 1.8 s and 0.5 s, respectively. The global symmetry based method has a higher computation time of 33.6 s, caused by the full 3D rigid registration.

The largest errors, both by the methods and observers, are made in participants with cerebral atrophy. These participants have a widened interhemispheric fissure, allowing many plane orientations and positions to result in a valid division of the hemispheres.

## 6.1 Introduction

The human brain exhibits an approximately bilateral symmetry, having the cerebrum divided into two hemispheres. The hemispheres are separated by the interhemispheric fissure (IF) that is often modelled as a geometric plane, the so-called midsagittal plane. Automatic detection of the midsagittal plane can be useful for a variety of purposes, for example comparing the hemispheres on anatomical and physiological asymmetries, correct the image for

misalignment and miscentering, and as an initial step for image or atlas registration.

In the past years, numerous methods have been published for the automatic extraction of the midsagittal plane. Many of these methods report good results, but often lack quantitative validation and comparison with existing methods.

The present study will classify existing methods and implement one method of each class. These methods will be evaluated by running them on datasets and comparing them to manual delineations of the midsagittal plane. Furthermore, the performance of human observers who manually delineate the midsagittal plane will be assessed as well.

## **6.2 Methods and Materials**

Existing methods can be classified into three general approaches: feature based, global symmetry based, and local symmetry based methods. Feature based methods utilize known features of the midsagittal plane, such as its linear shape or low intensity values of the cerebrospinal fluid (CSF) in the IF. Symmetry based methods attempt to find a plane that optimizes some symmetry-measure, under the assumption that the hemispheres have bilateral symmetry with respect to the midsagittal plane.

### **6.2.1 Feature based methods**

The most well-known feature based method is by Brummer (1991), who uses the Hough transform to detect the linear shape of the IF. However, most feature based methods rely on the low intensity of the CSF (with appropriate MR sequences such as T1-weighted) or the intensity differences between the CSF in the IF and the surrounding tissue (Bergo et al. 2009; Ekin 2006; Nowinski et al. 2006; Puspitasari et al. 2009; Volkau et al. 2006).

Methods based on the image intensity initialize a sagittal plane in the centre of the image and try to optimize the position of the plane by applying a number of translations and rotations. During optimization, the intensity values along the plane have to be minimized or the differences in intensity values with the surrounding tissue have to be maximized. The method by Ekin (2006) differs from this approach: multiple lines are inspected in the

left-right direction, a local minimum in intensity caused by the CSF is detected, and a plane is fitted through all detected points.

### 6.2.2 Global symmetry based methods

Symmetry based methods use the fact that the brain has bilateral symmetry with respect to the hemispheres of the cerebrum. Global symmetry based methods reflect the input image over the sagittal axis and perform a rigid registration to align the reflected image with the original. The resulting transformation can be used to derive the midsagittal plane by taking half of the transformation. Many variations exist with modifications to this generic approach (Ardekani et al. 1997; Joshi et al. 2003; Junck et al. 1990; Liu et al. 1998; Liu et al. 2001; Minoshima et al. 1992; Mykkänen et al. 2005; Prima et al. 2002; Prima et al. 2000; Teverovskiy et al. 2006; Tuzikov et al. 2002; Tuzikov et al. 2003).

Modifications to the generic approach includes the use of different similarity metrics (e.g. cross correlation, mutual information), using only part of the image to speed up the registration, multi-resolution approaches, and the use of derivative images. Some methods perform the rigid registration on (a number of) 2D transversal slices and fit a midsagittal plane through the resulting midsagittal lines (Liu et al. 1998; Liu et al. 2001; Minoshima et al. 1992; Mykkänen et al. 2005).

### 6.2.3 Local symmetry based methods

Whereas the previous methods consider the entire image, local symmetry based methods try to first approximately locate the centre of the brain and then locally optimize a symmetry-measure (Hu et al. 2003; Smith et al. 1999; Stegmann et al. 2005). The rationale behind this is that the global symmetry of the head not necessarily represents the symmetry of the hemispheres. These methods are generally faster than the global symmetry based methods, because they only consider part of the brain.

The midsagittal plane is usually initialized in the centre of the image and a number of line-segments perpendicular to this plane is inspected for symmetry. Translations and rotations are applied to the plane's position to optimize the symmetry of the line-segments based on intensity.

### 6.2.4 Implementation

Three methods, one from each class, were implemented for evaluation on a series of datasets and compared to manual delineations. These methods are most commonly used and cited in existing literature. From the feature-based methods, the method designed by Volkau et al. (2006) (later improved by Nowinski et al. (2006)) was implemented. From the local symmetry based methods, the method described by Hu et al. (2003) was implemented. For the global symmetry based methods, elastix was used to compute a rigid registration between the input image and a left-right reflected image, using advanced Mattes mutual information as a symmetry-metric (Klein et al. 2010).

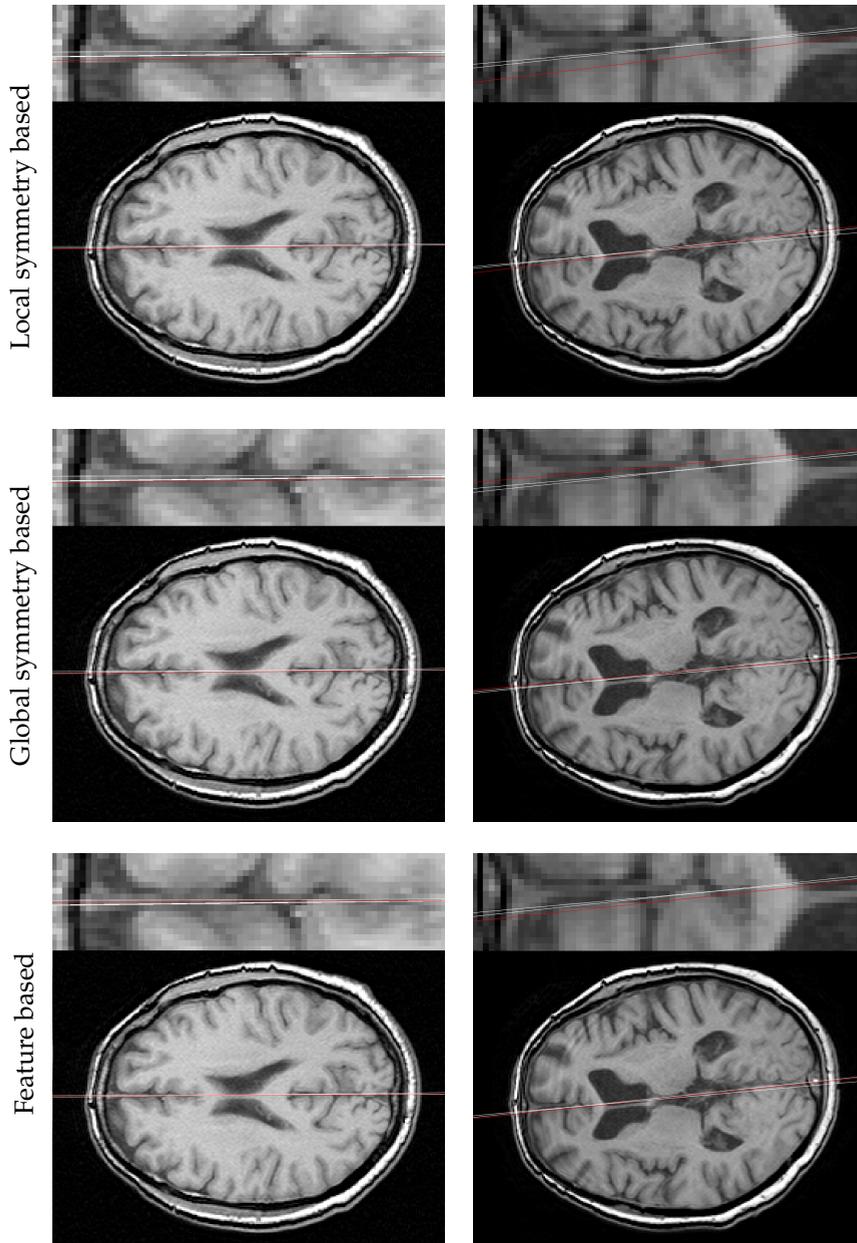
### 6.2.5 Experiments

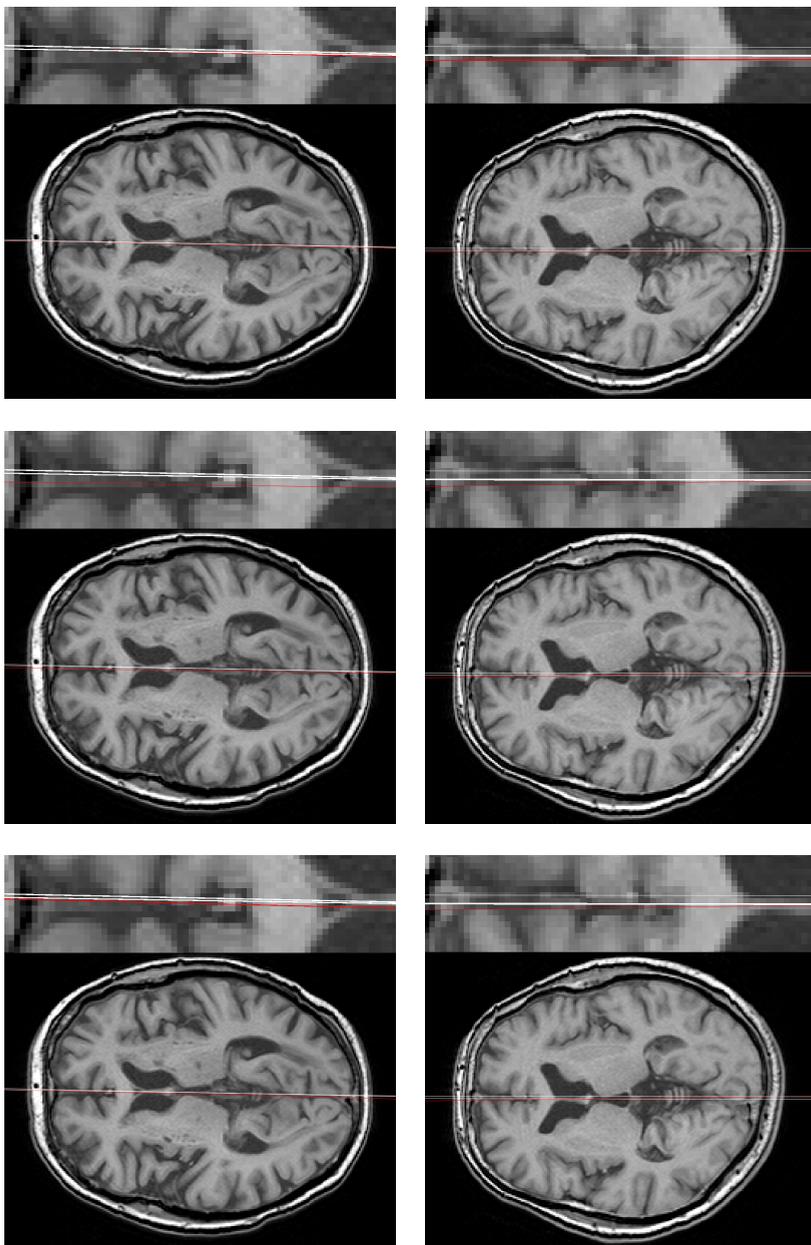
Fifty random participants (mean age: 60 years, sd: 9 years) were included from the Second Manifestations of ARterial disease (SMART) study (Simons et al. 1999). The objectives of the SMART study are to determine the prevalence of vascular risk factors and concomitant arterial disease and to study the incidence of future cardiovascular events and its predictors in patients newly referred to our hospital with atherosclerotic disease. The SMART study is approved by the Medical Ethics Committee. Written informed consent was given by all participants.

MRI acquisition was performed on a Philips Gyroscan ACS-NT 1.5 T whole body system (Philips Medical Systems, Best, The Netherlands). The protocol consisted of, among others, a transversal T1-weighted gradient echo sequence (TR/TE: 235/2 ms) with a reconstructed voxel-size of  $0.9 \times 0.9 \times 4.0$  mm. The brain extraction tool (BET) was used to extract a brain mask from the images that will be used for evaluation of the results (Smith 2002).

Manual delineation of the MSP was performed by two human observers: A and B. The plane was constructed by computing a least-squared fit through observer-annotated points with a custom application developed in MeVisLab (MeVis Medical Solutions AG, Bremen, Germany (Ritter et al. 2011)). Adjustments to the computed plane could be made in an interactive viewer showing the three main orthogonal views of the image. Delineations were performed twice by observer A to compute the intra-observer variability.

To evaluate the results, the angular error (AE) and volume error (VE) were computed. The AE is the 3D angle between the normal of two planes. The VE is the amount of brain volume (i.e. the volume within the brain mask ob-





**Figure 6.1:** Results of the automatic methods (in red) and the manual delineations (in white). The result with on average the smallest volume errors is shown in the top row, average volume errors in the middle rows, and on average the largest volume error in the bottom row.

**Table 6.1:** Results of the evaluation, showing the angular error (AE), volume error (VE), and computation time. For each participant, the difference (both AE and VE) of a computed plane with the first delineation of observer A and the delineation of observer B was determined and averaged. For each method and each error measure, the resulting differences per participant were averaged. For the inter-observer differences, the first delineation of observer A was compared with the delineation of observer B. For the intra-observer differences, the first and second delineations of observer A were used.

Measure, mean (sd):	AE (°)	VE (ml)	Time (s)
Feature based	1.25 (0.53)	1.58 (0.59)	0.5
Global symmetry based	1.54 (0.61)	1.67 (0.54)	33.6
Local symmetry based	1.46 (0.60)	1.68 (0.62)	1.8
Inter-observer	1.39 (0.95)	1.60 (0.87)	
Intra-observer	1.02 (0.66)	1.08 (0.55)	

tained with the BET) in-between the two planes. For each error measure, the differences between the computed plane, and the first delineation of observer A and the delineation of observer B were determined and averaged. In a similar way, the inter-observer and intra-observer differences were computed.

### 6.3 Results

Results of the evaluation are shown in Table 6.1. Example results are shown in Figure 6.1.

No significant differences were found (using a paired-samples Students' t-test, all  $p > 0.05$ ) between any of the errors that the methods made and the inter-observer difference. All errors made by the methods and the inter-observer difference were significantly larger than the intra-observer difference ( $p < 0.05$ ). Human observers required on average one minute to manually delineate the midsagittal plane.

### 6.4 Discussion

This study introduced a classification of various methods to extract the midsagittal plane from neurological images, distinguishing feature-based, global symmetry based, and local symmetry based methods. Furthermore, this

study is the first to compare the performances of the leading methods in each class by evaluation on MRI brain datasets.

The three methods selected for implementation are the most used and cited methods in existing literature. Following the description of each method provided in the literature, the implementation of the methods was done as accurately as possible. However, it cannot be excluded that differences with the originally published methods exist. The method that is actually most cited, by Brummer (1991) using the Hough transform, was not used in the comparison. A common problem with this technique is that the Hough transform sometimes returns two results, for the left and right hemisphere separately, in cases where the IF is widened or in high resolution images where the width of the IF consists of multiple voxels. This cannot be detected reliably and more recent feature based methods do not suffer from this problem. For the global symmetry based methods, a generic implementation was made based on the common concept that all existing methods share: reflecting the image and performing a rigid registration. Based on the specific task, improvements to this generic approach might be possible and thus reducing the errors for this task.

All errors made by the methods are comparable to the inter-observer differences, indicating that the methods perform similar to what human observers can achieve. The feature-based method by Volkau et al. (2006) and Nowinski et al. (2006) has an error that is slightly less than the inter-observer difference, whilst the symmetry based methods have an error that is slightly higher. However, none of the differences are significant, indicating that all methods perform sufficiently.

The higher error of the symmetry-based methods could be caused by the fact that the brain and head are not perfectly symmetric. Especially global symmetry based methods could suffer from the influence of all non-brain tissue that is present in the scan.

The main reason why some (or all) methods have a large error in some of the participants is caused by cerebral atrophy, resulting in a widening of the IF. In that case, many plane orientations and positions result in a valid midsagittal plane separating both hemispheres. If the volume error would be evaluated on a grey and white matter segmentation, instead of the brain mask, the volume error is expected to be smaller. However, if many plane orientations and positions result in a valid midsagittal plane, this could influence the results of subsequent (a)symmetry analysis. If a younger cohort

of participants was used, the errors might have been smaller, as the IF is generally narrower in younger persons.

Computation time of the global symmetry based method is large, compared to the other two methods. This could be improved by computing the registration on a number of 2D slices and fitting a plane through the resulting lines, although it is not expected that this will be profitable. However, this is a technique commonly applied by less recently published methods to reduce the computational power required and to be able to run these methods on older workstations.

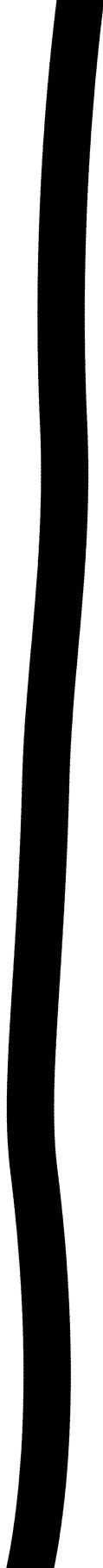
All methods are likely to be able to detect the midsagittal plane on other types of MR sequences (e.g. FLAIR, IR, T2), at higher field strength, or other imaging modalities (e.g. CT). The main requirement for feature based methods is that the IF should be visible, something that almost always holds for medical brain images. Symmetry based methods should have enough contrast in the images to reliably detect the symmetry of the hemispheres. When using images with a higher spatial resolution, the computation times are expected to increase.

The midsagittal plane is not always the best solution to separate the hemispheres of the cerebrum, as can be seen in the bottom row of Figure 6.1. Owing to so-called brain torque, where the left occipital and right frontal lobes are larger than their counterparts in the other hemisphere, the IF is not a perfectly geometrical plane. Computing a curved midsagittal surface results in a better separation of both hemispheres in cases with noticeable brain torque (Kuijf et al. 2013a).

## 6.5 Conclusion

Methods to extract the midsagittal plane can be classified in three classes: feature based methods, global symmetry based methods, and local symmetry based methods. For the first time, leading methods of each class have been compared and evaluated on the same dataset. Results showed that each method performs similar to what human observers can achieve when delineating the midsagittal plane manually.





# Automatic extraction of the midsagittal surface from brain MR images using the Kullback–Leibler measure

**Based on:** H.J. Kuijf, S.J. van Veluw, M.I. Geerlings, M.A. Viergever, G.J. Biesels, K.L. Vincken, “Automatic extraction of the midsagittal surface from brain MR images using the Kullback–Leibler measure”, (submitted)

### Abstract

The midsagittal surface separates the two hemispheres of the cerebrum. This surface is often typified as a geometrical plane: the midsagittal plane. However, in subjects with a considerable amount of naturally occurring brain torque, the midsagittal surface deviates to a large extent from a plane. In the present study, an automated method to extract the midsagittal surface is proposed, evaluated on a large dataset, and compared to a conventional midsagittal plane representation.

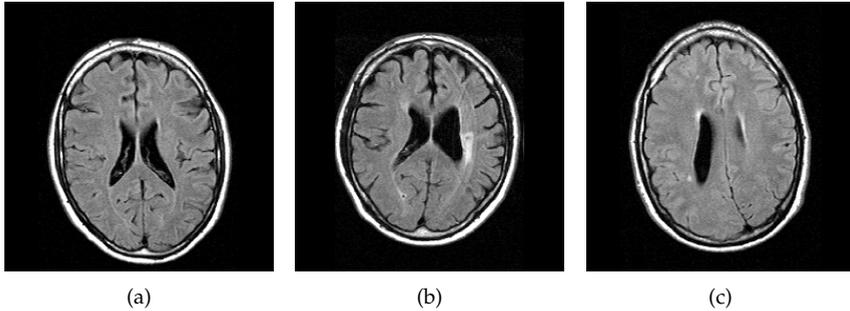
The midsagittal plane was extracted from MR images with a technique based on the Kullback–Leibler measure. This plane was used to initialize a surface, that was deformed to represent the midsagittal surface. One hundred subjects were selected from the SMART-MR study: fifty subjects with brain torque and fifty random subjects. Manual delineations of the midsagittal surface were used for evaluation.

The extracted midsagittal planes and surfaces were compared to the manual delineations by assessing the absolute volume of misclassified cerebrum tissue. The midsagittal surface resulted in significantly better separations of the hemispheres. In the randomly selected subjects, the error reduced from  $2.71 \pm 1.05$  ml to  $2.20 \pm 0.66$  ml and in subjects with brain torque from  $4.85 \pm 2.79$  ml to  $2.23 \pm 0.77$  ml, with improvements up to 16.6 ml in individual subjects with marked brain torque.

## 7.1 Introduction

Bilateral symmetry is an essential concept in biology and many animal species, including humans. Our appearance exhibits bilateral symmetry and some organs in our body come in symmetrical pairs. The cerebrum of the human brain is divided into two hemispheres, separated by the interhemispheric fissure (IF). Both hemispheres are connected by the corpus callosum and smaller commissures, including the anterior, posterior, and hippocampal commissure. Comparison of the two hemispheres and detection of differences has been a topic of interest for many years (Rentería 2012). Besides lateralization of function (e.g. Nagel et al. (2013) and Sommer et al. (2001)), anatomical differences can suggest the presence of pathology (e.g. mass-effect brain tumours, (Joshi et al. 2003)).

The midsagittal plane (also called median plane) is a geometric plane that separates the two hemispheres of the cerebrum and thus represents the IF. Identifying this plane is a prerequisite for many (automated) analyses that

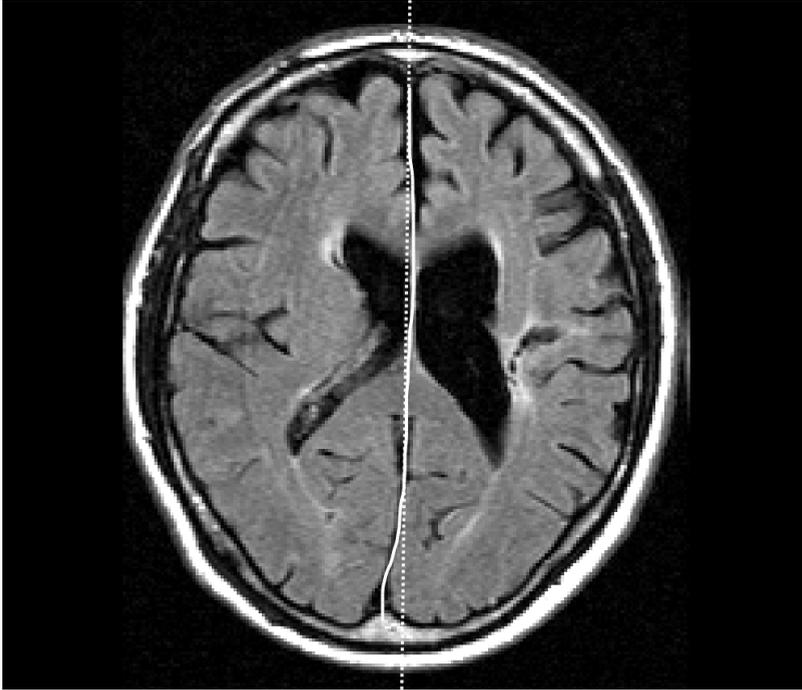


**Figure 7.1:** Example MRI FLAIR images showing the natural variation in brain torque. (a) No visible brain torque. (b) Slightly visible brain torque, showing an enlarged left occipital lobe. (c) Extreme brain torque, showing an enlarged right occipital lobe whereas usually the left occipital lobe is larger.

compare the hemispheres with one another. Therefore, the extraction of the midsagittal plane from brain images is a topic that has been studied for decades (Brummer 1991; Junck et al. 1990; Phillips 1964) and is still highly relevant, since new methods are still actively developed (Jayasuriya et al. 2013; Puspitasari et al. 2009). Most methods assume bilateral symmetry of the brain and extract the midsagittal plane by optimizing a symmetry measure between the original brain scan and a reflected version of itself (e.g. Hu et al. (2003), Liu et al. (2001), Prima et al. (2002) and Tuzikov et al. (2003)). Other methods use features—such as the intensity of the CSF in the IF (Nowinski et al. 2006; Puspitasari et al. 2009; Volkau et al. 2006) or the linear shape of the IF (Brummer 1991; Zhang et al. 2008)—to extract the midsagittal plane.

However, the human brain has no perfect bilateral symmetry. The left occipital lobe and right frontal lobe are often larger than their counterparts in the other hemisphere. This effect is known as brain torque (also known as petalia and Yakovlevian torque, (Toga et al. 2003)). Natural variation causes this effect to be larger in some humans than in others, as can be seen in Figure 7.1. Next to this, the presence of mass-effect tumours or other pathology could induce asymmetries in the cerebrum. Because of this, extracting the midsagittal plane might not be the optimal solution to separate the two hemispheres.

The midsagittal surface is a (curved) surface that corresponds with the IF and separates the two hemispheres. Contrary to the midsagittal plane,



**Figure 7.2:** Example scan (also shown in Figure 7.1(b)) showing the difference between the midsagittal plane (dotted line) and the manually delineated midsagittal surface (solid line). It is clear that the midsagittal plane makes a segmentation error where it intersects the tissue of the left occipital lobe.

the midsagittal surface follows the natural shape of the IF. In the presence of asymmetries, either caused by natural variation or pathology, a midsagittal surface is able to correctly separate the two hemispheres, whereas a midsagittal plane would intersect or misclassify some brain tissue. This is demonstrated in Figure 7.2.

In the present study, an automated method for the extraction of the midsagittal surface will be presented. An existing technique by Volkau et al. (2006) and Nowinski et al. (2006) will be used to extract the midsagittal plane, which then is used as an initialization of the midsagittal surface. Optimizing the configuration of the midsagittal surface is based on the same principle as that for computation of the midsagittal plane.

We hypothesize that the midsagittal surface will give a better separation

of the left and right hemispheres. In a random selection of subjects, manual left-right delineations of the cerebrum will be compared to the automatically extracted midsagittal plane or surface. Since considerable brain torque is not commonly found in the average (randomly selected) subject, no significant improvement of the midsagittal surface over the midsagittal plane is expected. Therefore, another set of subjects with considerable brain torque will be selected, to demonstrate the added value of the midsagittal surface in these subjects. In these subjects, the error made by the midsagittal plane, as compared to manual delineations, should be significantly reduced when the midsagittal surface is used.

The midsagittal surface can be used for more purposes than for a better segmentation of the two hemispheres. The amount of brain torque can be determined by quantifying the deviation of the midsagittal surface from a plane, i.e. the surface tortuosity (Zhang et al. 2004). In subjects with considerable brain torque or midsagittal surface tortuosity, automated analyses that rely on brain symmetry should not be performed. An example is brain perfusion analysis, where the perfusion in one hemisphere is compared to the contralateral hemisphere by using the midsagittal plane as a “mirror”. Furthermore, brain asymmetry is speculated to be linked to handedness and other cognitive functions (Balzeau et al. 2012; Galaburda et al. 1978). Determining the shape of the midsagittal surface or its tortuosity can aid in analysis of this topic. In addition, the shape of the midsagittal surface is a parameter that is ignored in large voxel-based morphometry studies (Ashburner et al. 2000), since this parameter is lost in the spatial normalization step. Brain torque or midsagittal surface tortuosity is therefore a parameter that is still relatively unexplored in large studies.

The aims of this study are to present a method that can extract the midsagittal surface and to show that this is an improvement over the midsagittal plane, by reducing the error that is made in left-right segmentation. It is hypothesized that in the average subject the midsagittal surface will not improve (nor degrade) the quality of the left-right segmentation as compared to the midsagittal plane, but that in subjects with considerable brain torque the accuracy will improve significantly.

## 7.2 Methods and Materials

### 7.2.1 Midsagittal plane

A variety of methods exist to extract the midsagittal plane from neuroimages (Kuijf et al. 2013a). These methods use different approaches and can be classified into symmetry-based methods and feature-based methods. In the present study, a method developed by Volkau et al. (2006) and Nowinski et al. (2006) was used. This method employs the difference in image intensity between the CSF in the IF and the neighbouring tissue (grey and white matter) that is visible with most imaging modalities. This approach can be extended to extract the midsagittal surface and thus forms an ideal starting point for the present study.

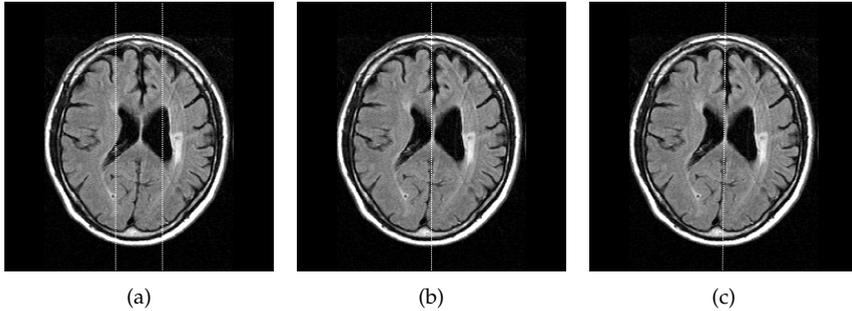
The method assumes that the brain is approximately located in the centre of the image and initializes two reference planes 2 cm apart from the central sagittal slice of the image (see Figure 7.3(a)). These reference planes will generally contain mostly grey and white matter. A single probability distribution  $p$  of the intensity values in the two reference planes was created.

Next, all sagittal slices in between the two reference planes were inspected. For each slice, a probability distribution  $q$  of the intensity values was created. The Kullback–Leibler (KL) divergence was used to compute the difference  $d$  between  $p$  and  $q$ :

$$d(p/q) = \sum_i p_i \log(p_i/q_i),$$

where  $p_i$  and  $q_i$  are the probabilities of the intensity having the value  $i$ . Since the midsagittal plane generally will contain mostly CSF, whose intensity is very different from that of grey and white matter, it is expected that the difference  $d$  of the midsagittal plane with the reference slices will be large. Accordingly, the sagittal slice in between the two reference slices with the largest difference  $d$  was chosen as the central midsagittal plane (see Figure 7.3(b)).

The midsagittal plane of the brain does not necessarily coincide exactly with a sagittal slice of the image, but may have been rotated owing to an angulated position in the scanner. Therefore several small translations and rotations were applied to the central midsagittal plane, using an iterative optimization strategy described by Volkau et al. (2006) to maximize  $d$ . The plane with the largest difference  $d$  was taken as the final midsagittal plane (see Figure 7.3(c)).



**Figure 7.3:** Extraction of the midsagittal plane as described by Volkau et al. (2006) (a) The lines indicating the two reference planes. (b) The sagittal slice with the largest difference as compared with the reference slices. (c) The final midsagittal plane.

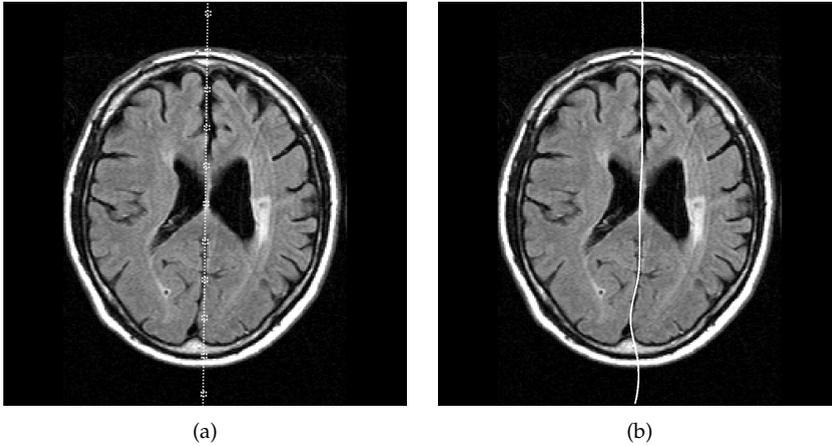
### 7.2.2 Midsagittal surface

The midsagittal plane computed in the previous section was used to initialize the computation of the midsagittal surface. The midsagittal surface was represented as a bicubic spline, as implemented in ALGLIB (Bochkanov et al. 2012). Control points for the spline were placed in a regular grid on the computed midsagittal plane, with distance  $m$  between the control points. An example is shown in Figure 7.4(a).

The control points could be moved in the left-right direction in order to optimize the configuration. The KL divergence was used as the cost function for optimization, by maximizing  $d$ . For each configuration of control points, the intensity probability distribution  $q_i$  was computed for the spline.

A limited-memory Broyden-Fletcher-Goldfarb-Shanno quasi-newton method (L-BFGS, (Liu et al. 1989)), as implemented in the dlib C++ library (King 2012), was used to determine the direction of the search. The gradient information of the cost function was numerically approximated. The step size of each control point was scaled with the gradient, allowing sub-voxel positioning of the control points. The optimization method was terminated when two consecutive optimization steps had a difference  $\delta d < 1 \times 10^{-5}$ . An example result can be seen in Figure 7.4(b).

Visual inspection showed that in some extreme cases, the optimizer converged to a local maximum and did not reach the optimal position of the midsagittal surface. Such an example is given in Figure 7.5(a). This was solved by

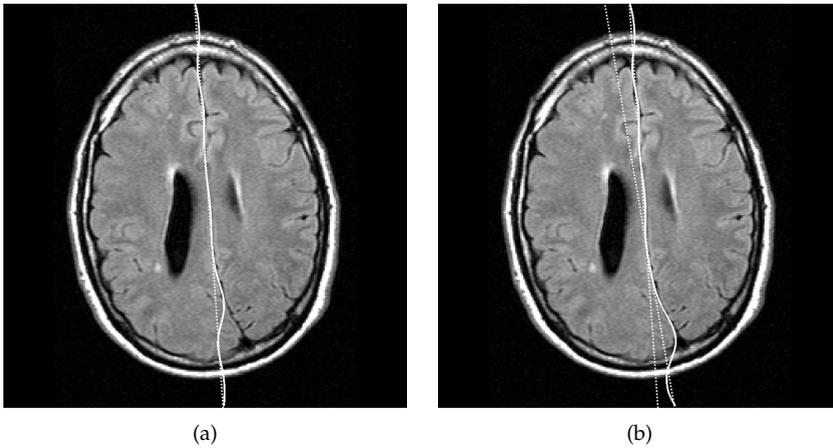


**Figure 7.4:** Computation of the midsagittal surface. (a) Initialization of the control points in a regular grid along the computed midsagittal plane shown in Figure 7.3(c). Control points can move in the left-right direction only. (b) Final midsagittal surface.

computing the midsagittal plane twice: once on the anterior half of the image and once on the posterior half of the image. Both planes were used to initialize the midsagittal surface: the intersection-line of the two planes was determined and the control points of the surface anterior to this line were placed on the anterior midsagittal plane and vice versa for the posterior part. By doing this, the procedure reached the optimal position as shown in Figure 7.5(b). This dual-plane approach was used if the in-slice angle of the two planes was larger than  $3.5^\circ$  and the computation of both planes succeeded (see the parameter justification in the Results and Discussion sections for more details). Otherwise, the single midsagittal plane computed on the entire image was used, as described above.

### 7.2.3 Experiments and Validation

To assess the quality of the automated midsagittal plane/surface extraction, the results were compared to manual delineations of the midsagittal surface. The absolute amount of cerebrum tissue volume  $V$  between two delineations/methods (either manual, plane or surface) was used as the evaluation metric. An example of this is given in Figure 7.6, where  $V$  is determined for



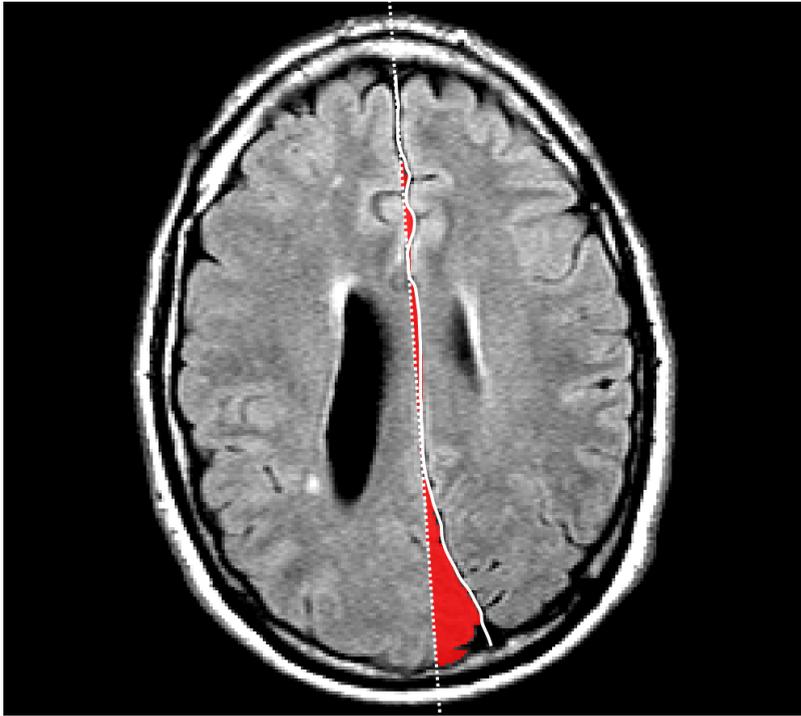
**Figure 7.5:** (a) Unsuccessful computation of the midsagittal surface, because the optimization procedure halted in a local optimum. (b) With the dual-plane approach, the optimizer reached the optimal position.

a manual delineation and the computed midsagittal surface. The value of  $V$  was also used to evaluate the intra- and inter-observer agreement. The cerebellum was ignored, as is commonly done in segmentation algorithms (Liang et al. 2007), because a left-right segmentation of the cerebellum is ambiguous and ill-defined.

## Subjects

Subjects used for evaluation were included from the Second Manifestations of ARterial disease - Magnetic Resonance (SMART-MR) study, a cohort study of brain ageing in 1309 patients with manifest arterial disease (Geerlings et al. 2010). The SMART-MR study was approved by the Medical Ethics Committee of the University Medical Center Utrecht. Written informed consent was given by all subjects.

From the SMART-MR study, fifty consecutive subjects (set  $A$ , mean age: 59 years, sd: 11 years) were included. Because the average subject does not have a considerable brain torque, the improvement in these fifty subjects was neither expected to be large nor significant. Therefore another fifty subjects (selected from the remaining subjects not included in  $A$ ) having considerable brain torque were automatically selected (set  $B$ ). This was achieved through



**Figure 7.6:** The error made by the midsagittal plane (dotted line) as compared with the manual delineation of the midsagittal surface (solid line) is annotated in red.

an automatic left-right segmentation, performed by deformable registration of the MNI152 atlas to each subject in the SMART-MR study with the elastix toolbox (Fonov et al. 2009; Fonov et al. 2011; Klein et al. 2010; Kuijf et al. 2013a). The fifty subjects whose midsagittal plane had the largest error  $V$  with the generated left-right segmentation were selected (set  $B$ , mean age: 60 years, sd: 12 years).

## MRI

MR imaging was performed on a 1.5 T whole-body system (Gyrosan ACS-NT, Philips Medical Systems, Best, the Netherlands). The protocol included, among other sequences, a transversal T1-weighted gradient-echo sequence (repetition time (TR)/echo time (TE): 235/2 ms); a transversal T2-weighted

fluid-attenuated inversion recovery (FLAIR) sequence (TR/TE/inversion time (TI): 6000/100/2000 ms), and a transversal inversion recovery (IR) sequence (TR/TE/TI: 2900/22/410 ms), all with a reconstructed voxel size of  $0.9 \times 0.9 \times 4.0$  mm. For extraction of the midsagittal plane and surface, the FLAIR sequence was used.

Segmentation of brain tissue into grey and white matter was performed with a probabilistic k-Nearest Neighbour classification segmentation method. This method used the T1-weighted, IR, and FLAIR sequences, as described by Anbeek et al. (2005). The registration with the MNI152 atlas was used to exclude the cerebellum from the evaluation.

### Manual delineation

The midsagittal surface was manually delineated on the FLAIR scan of the selected subjects by two experienced human observers. The first observer delineated set  $A$  twice and set  $B$  once. The second observer delineated set  $A$  once. Delineations were made with a dedicated tool developed in MeVisLab (MeVis Medical Solutions AG, Bremen, Germany, (Ritter et al. 2011)) using a freehand spline drawing technique. The instructions were to separate both hemispheres of the cerebrum, when visible, on each slice of the scan. The first delineation of observer 1 was used in all evaluations.

### Experiments

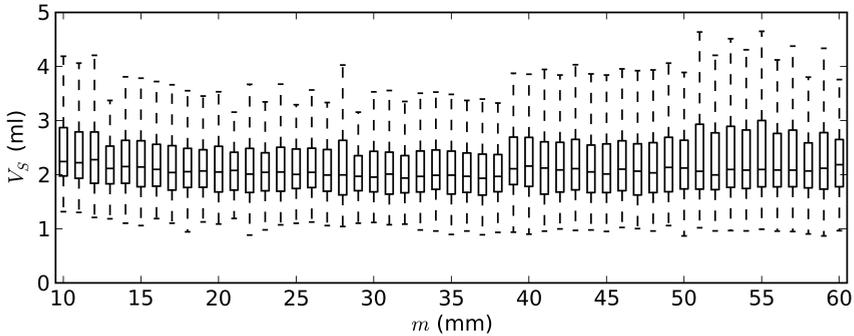
The most important parameter for the midsagittal surface is the grid size  $m$ . To investigate the effect of  $m$ , the results were computed for multiple values of  $m$  ranging from 10 to 60 mm with a step size of 1 mm.

An in-slice angle of  $3.5^\circ$  was empirically chosen at the cutoff value for using the dual-plane approach. To test the robustness of this choice, the results were computed for  $3.0^\circ$  and  $4.0^\circ$  as well.

The quality of the automatic left-right segmentation by registration of the MNI152 atlas was evaluated by comparison to the manual delineations.

## 7.3 Results

The intra-observer error  $V_{\text{intra}}$  on set  $A$  was (mean  $\pm$  sd)  $1.24 \pm 0.36$  ml and the inter-observer error  $V_{\text{inter}}$  on set  $A$  was  $1.40 \pm 0.40$  ml. Using a paired



**Figure 7.7:** The error  $V_S$  on all subjects from sets  $A$  and  $B$  is shown for different values of  $m$ . For all reported results,  $m = 30$  mm will be used.

samples Student's t-test, the difference between the intra- and inter-observer errors proved to be significant ( $p < 0.001$ ).

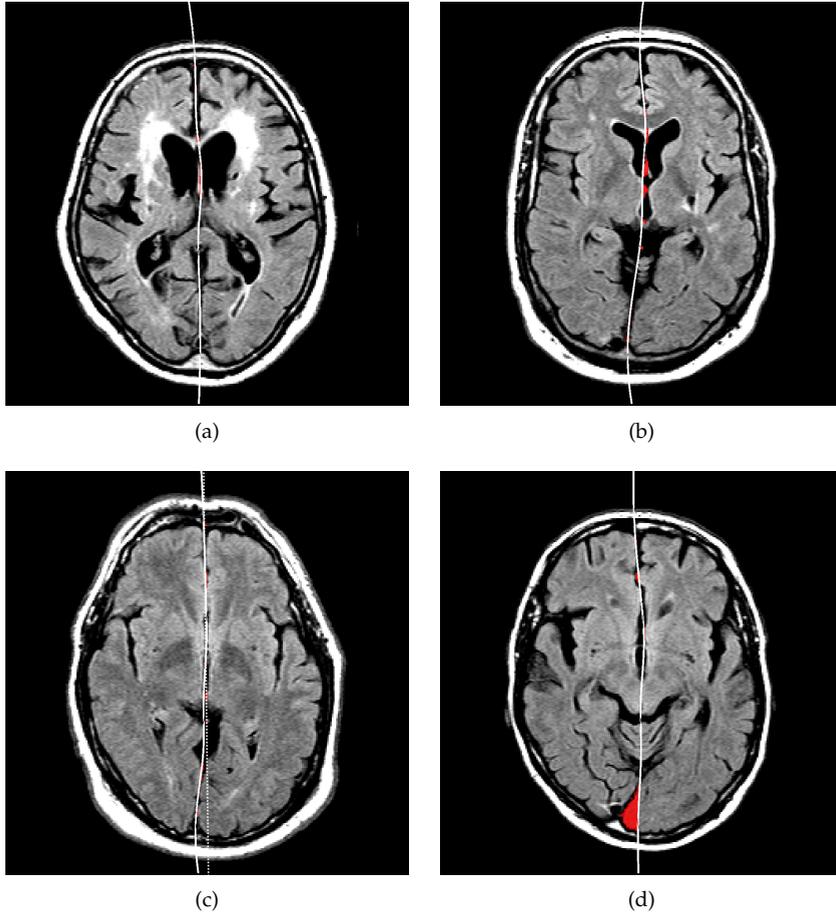
The error  $V_P$  of the midsagittal plane with the first delineation of observer 1 was;  $A$ :  $2.71 \pm 1.05$  ml,  $B$ :  $4.85 \pm 2.79$  ml,  $A \cup B$ :  $3.78 \pm 2.36$  ml. All errors were significantly larger than the intra- and inter-observer errors for manual delineation ( $p < 0.001$ ). Moreover, the error  $V_P$  on  $B$  was significantly larger than the error on  $A$  ( $p < 0.001$ ).

The error  $V_S$  of the resulting midsagittal surface with the first delineation of observer 1 is dependent on parameter  $m$ , as shown in Figure 7.7. All results reported in this section are for  $m = 30$  mm.

The error  $V_S$  of the midsagittal surface with the first delineation of observer 1 was;  $A$ :  $2.02 \pm 0.66$  ml,  $B$ :  $2.23 \pm 0.77$  ml,  $A \cup B$ :  $2.12 \pm 0.72$  ml. All errors were significantly larger than the intra- and inter-observer errors for manual delineation ( $p < 0.001$ ). The error  $V_S$  on  $B$  is not significantly larger than the error on  $A$  ( $p = 0.14$ ).

All errors  $V_S$  (for  $A$ ,  $B$ , and  $A \cup B$ ) of the midsagittal surface are significantly smaller ( $p < 0.001$ ) than the errors  $V_P$  of the midsagittal plane. The average improvement of the midsagittal surface over the midsagittal plane is;  $A$ :  $0.69 \pm 0.84$  ml,  $B$ :  $2.63 \pm 2.45$  ml,  $A \cup B$ :  $1.66 \pm 2.07$  ml. In addition, the standard deviations of the results of the midsagittal surface are smaller, which signifies that the surface-based results are more consistent.

Some example results of the computation of the midsagittal surface are



**Figure 7.8:** Example results of the computed midsagittal surface; error  $V_S$  is shown in red. (a) Result with a small error  $V_S$ . The midsagittal surface is comparable to the midsagittal plane, since this subject has no brain torque. The midsagittal surface is incorrect around the lateral ventricles, where the surface intersects with the left ventricle. (b) Subject where the midsagittal surface showed a large improvement over the midsagittal plane, because of the present brain torque. Some errors still occurred around the lateral ventricles. (c) Another subject where the midsagittal surface improved over the midsagittal plane (dotted line). (d) The midsagittal surface computed in this subject was unable to account for the present brain torque.

shown in Figure 7.8.

The computation time of the midsagittal plane was approximately 2 s. De-

pending on the required number of iterations of the optimizer, the computation of the midsagittal surface required an additional 2 to 10 s.

The value of  $3.5^\circ$  as a cutoff for the use of the dual-plane approach is fairly robust; setting the angle at  $3.0^\circ$  or  $4.0^\circ$  did not result in any significant differences in the results ( $p > 0.35$ ).

A registration with the MNI152 atlas was used to select subjects with brain torque. The left-right segmentation of this atlas was compared to the manual delineations. Because the MNI152 atlas contains a gap between the left and right hemispheres, different errors are reported for both hemispheres. The error  $V_{\text{MNI}}$  of the left-right segmentation by atlas registration with the first delineation was;  $A \cup B$  for the left hemisphere:  $7.46 \pm 4.56$  ml and  $A \cup B$  for the right hemisphere:  $10.01 \pm 5.70$  ml.

## 7.4 Discussion

The results of the midsagittal surface showed a significant improvement over those of the midsagittal plane, not only for subjects with considerable brain torque, but also for subjects without brain torque. Overall, the error  $V$  was reduced by 1.66 ml. This is relatively small compared to the whole brain volume, but is still a considerable amount of brain tissue in the vicinity of the IF. Improvements in subjects with considerable brain torque were larger, with reductions in  $V$  up to 16.6 ml for the subject shown in Figures 7.5 and 7.6.

The midsagittal surface as computed by the presented method is sometimes incorrectly positioned at the lateral ventricles. During optimization, the cost function will avoid the septum pellucidum, the membrane separating the lateral ventricles, and fit the spline through one of the CSF-filled ventricles. However, this had no influence on the left-right segmentation of the tissue in the cerebrum. An example is shown in Figures 7.8(a) and (b). If a correct left-right segmentation of the ventricles is needed, an (approximate) segmentation of the ventricles would be required. Using this segmentation, the cost function can then ignore this region in the optimization. Another solution would be to invert the contrast within the segmented ventricles (causing the CSF to have tissue-like intensities and the septum pellucidum to have CSF-like intensities), because then the spline-fitting will be correct. This was not done in the present study to keep the proposed method as efficient as possible, while the results still show useful improvements.

Many methods are available for the extraction of the midsagittal plane (Kuijf et al. 2013b). Although all methods have comparable results, it would be unintuitive to choose a symmetry-based method in the case of (asymmetric) brain torque. The method published by Volkau et al. (2006) and later improved by Nowinski et al. (2006) was considered most suitable. Furthermore, the used KL measure could be easily transformed into a cost function that could be optimized during the computation of the midsagittal surface.

Subjects with considerable brain torque were automatically selected. Obtaining manual delineations of all MRI scans in the SMART-MR study was not achievable and therefore an automatic left-right segmentation of all scans was achieved by atlas registration with the MNI152 template. The midsagittal plane was computed for all subjects. Subjects where the error  $V$  between the midsagittal plane and the left-right segmentation was largest were selected as subjects with considerable brain torque. Following this, manual delineations were made for the fifty subjects with the highest error  $V$ . Because the left-right segmentation by registration of the MNI152 template itself had a large error  $V$  with the manual delineations, owing to misregistration, it is likely that not all subjects with considerable brain torque were selected. The results reported are thus an underestimation of the true improvement of the presented method. Next to this, the number of 50 subjects included in  $B$  was an arbitrary choice and does not reflect the (unknown) prevalence of considerable brain torque in the study subjects. The probable inclusion of subjects without considerable brain torque in set  $B$  leads to a reduction of the reported results. It can therefore be assumed that the true improvement of the midsagittal surface over the midsagittal plane is higher than reported.

The errors made by the computed midsagittal surface are larger than the intra- and inter-observer errors. This is mainly because of small protrusions that cannot be captured by a smooth surface. This is shown in Figure 7.6, where small protrusions of the right hemisphere at the anterior part are delineated by the observers but not captured by the smooth surface, and in Figure 7.8(d).

The proposed method computed the midsagittal plane twice for subjects with extreme brain torque, such as in Figure 7.5. Both planes, computed separately on the anterior and posterior halves of the image, were used to initialize the midsagittal surface if the in-slice angle between the two planes was larger than  $3.5^\circ$ . This value was chosen empirically after inspecting a few subjects with no and extreme brain torque. In case the dual-plane approach

was used, the line resulting from the intersection of both planes was determined. Control points for the spline anterior to the line were placed on the midsagittal plane computed on the anterior half of the scan and vice versa for the posterior part of the scan. If the in-slice angle between both planes is set too small, the intersection line can appear at unwanted locations or even outside the scan. If the in-slice angle is too large, cases in which the dual-plane approach is superior are missed. The results showed that the selected value of  $3.5^\circ$  is fairly robust.

Alternative methods to compute the midsagittal surface have been proposed by Stegmann et al. (2005). The first method of Stegmann et al. (2005) is symmetry-based: on each left-right line in the image, the point with the highest bilateral symmetry is found. A thin-plate spline is fitted through all detected points. This method assumed that the image is aligned to achieve optimal symmetry along the left-right lines, something that can be done if the midsagittal plane is known. The computation of symmetry points is performed in the cerebrum only, so a full brain segmentation is required. The second suggestion of Stegmann et al. (2005) is similar to the method presented here. Assuming that the CSF-filled IF is hypointense on MRI scans, Stegmann et al. (2005) try to minimize the sum of intensities along the midsagittal surface in the cerebrum. Both methods have been evaluated qualitatively and the first method was considered more robust, consistent, and conservative (Stegmann et al. 2005). No quantitative validation with an established ground truth is given. The subjects used by Stegmann et al. (2005) were randomly selected elderly, similar to set *A* presented here. It is therefore unknown how the proposed method performs in the presence of considerable brain torque, especially because the method is symmetry-based whereas brain torque is an asymmetric phenomenon.

The use of the KL-measure as a cost function to optimize the midsagittal surface has the main advantage that it is not directly dependent on the (absolute) image intensities. The only prerequisite is that the CSF in the IF has a distinct contrast with its surroundings, not necessarily hypointense as required by Stegmann et al. (2005). This makes the method applicable to other MRI contrasts, such as T2-weighted scans where the CSF is hyperintense. Furthermore, Puspitasari et al. (2009) and Vos et al. (2013) showed that the KL-measure can be used for CT images as well. This makes the proposed method widely usable, e.g. for the computation of midline shift on brain CT images (Liao et al. 2010).

Extracting the midsagittal plane is commonly done to perform symmetry analyses of the cerebrum. Given a midsagittal plane, this is a simple and straightforward procedure. Having a midsagittal surface rather than a plane makes it more complicated (or even impossible) to perform symmetry analyses, since the midsagittal surface can be (highly) curved as shown in Figure 7.5(b). On the other hand, performing symmetry analyses on brains with considerable brain torque is questionable and possibly even incorrect. It is advisable to compute the midsagittal surface and evaluate the deviation from a plane before performing symmetry analyses.

## **7.5 Conclusion**

Computation of the midsagittal surface results in a significant better left-right segmentation of the cerebrum than the midsagittal plane, especially for subjects with considerable brain torque.



# Summary and discussion

## 8.1 Summary

In this thesis, a number of image processing techniques for quantification and assessment of brain MRI were presented, distinguishing manual, semi-automatic, and automatic techniques. Image processing techniques are needed to ensure high quality assessments in situations where manual assessment is difficult or too time-consuming. This included the detection of cerebral microbleeds and cortical microinfarcts on 7 T brain MRI. It is infeasible to manually assess the large amount of data and level of detail that is acquired with 7 T MR scanning. In order to achieve robust and reliable ratings, semi-automatic image processing techniques were introduced.

The current standard for microbleed detection is visual rating with validated visual rating scales. The reliability of ratings is quantified by determining the inter-rater agreement. This is commonly done by computing Cohen's kappa coefficient ( $\kappa$ ) or the intraclass correlation coefficient (ICC). In **Chapter 2** it was shown that these measures itself are not reliable in the case of outlier subjects or multiple microbleeds per subject. The use of the Dice similarity coefficient (DSC) is proposed as an alternative that is robust in these cases.

In **Chapters 3 and 4**, techniques were presented for the semi-automated detection of microbleeds on 3 T and 7 T brain MRI, respectively. Based on the

radial symmetry transform (RST), it was possible to detect microbleeds with a high sensitivity and a modest number of false positives that needed to be censored manually.

For the detection on 3 T T2\*-weighted brain MRI, a combination of 3D and 2D RST computations was applied to select potential microbleed locations. The thresholds that influenced the outcome of the detection were variable, so that a trade-off could be made between sensitivity and the time required to censor the false positives. Depending on the chosen settings, the sensitivity ranged from 65 to 84 % compared with the ground truth rating. Censoring false positives required 1 to 2 min per subject. Sensitivities of individual observers during visual rating ranged from 39 to 86 % and required 5 to 10 min per subject per observer.

The 3D RST was used to detect microbleeds on 7 T gradient echo dual-echo brain MRI. Both echo time images were used to detect microbleeds. Potential microbleed locations that were visible on each of the two echo images were presented to a human observer who identified true microbleeds and censored false positives. One of the most important results of the RST was the detection of extra microbleeds that were not identified earlier during visual rating, but were confirmed as true microbleeds by observers. The resulting sensitivity of the RST was 71.2 %, which was higher than the sensitivity of individual observers. The time required to censor the false positives was on average 2 min per subject. Full visual rating required on average 30 min per subject.

The ability to visualize cortical cerebral microinfarcts in vivo on 7 T brain MRI required a semi-automatic detection technique to complement manual detection. Manual detection is time-consuming and highly observer dependent. A proof of principle study for such a technique has been presented in **Chapter 5**. The developed technique was constructed such that it detected all manually annotated microinfarcts, whilst limiting the number of false positives. Visual inspection of the results required 5 to 20 min. A number of additional microinfarcts—not present in the original visual rating—were identified and confirmed as true microinfarcts. This stresses the importance of semi-automatic techniques to support human observers during detection.

**Chapter 6** started with an overview of methods to extract the midsagittal plane from brain MR images. The midsagittal plane separates the two hemispheres of the brain and many competing methods exist to extract it from brain MR or CT images. The methods can be classified into three distinct classes: feature-based, global symmetry based, and local symmetry based

methods. A leading method of each class has been implemented and evaluated. Evaluations consisted of comparing the detected midsagittal planes to manually annotated planes in a dataset of fifty subjects. Manual annotations were made by two human observers. The errors made by each of the methods were not significantly different from the inter-observer variability, i.e. all methods performed similar to humans.

In **Chapter 7**, one of the midsagittal plane methods was extended to compute the midsagittal surface. The computed midsagittal plane was used to initialize a bicubic spline with control points placed in a regular grid. The configuration of the control points was optimized to extract the midsagittal surface. The results of the midsagittal plane and surface extractions were evaluated on two datasets of fifty subjects each: one with randomly selected subjects and one with subjects having considerable brain torque. Manual annotations of the midsagittal surface were made by two human observers. The midsagittal plane showed errors significantly larger than the inter-observer variation. The midsagittal surface performed significantly better than the midsagittal plane, resulting in smaller errors. However, these errors were still larger than the inter-observer variation.

## 8.2 Discussion

### 8.2.1 Manual image processing

Manual image processing, i.e. performing ratings or creating segmentations of images manually, is an important step preceding any (semi-)automatic image processing. Every chapter in this thesis used some form of manual image processing, by rating microbleeds, microinfarcts, or creating delineations of the midsagittal plane or surface. The results of manual ratings or delineations were used to train or test automatic image processing techniques (including the automated part of semi-automatic techniques).

Dedicated tools are used to create manual ratings or segmentations. For example, rating of microbleeds in this thesis was done according to the Microbleed Anatomical Rating Scale (MARS, Gregoire et al. (2009)). Initially, the paper MARS Rating Form was completed for each subject. In order to use these ratings to automatically evaluate the results of the radial symmetry transform, the ratings were digitized. Later ratings and manual evaluation of the results of the RST were performed digitally, with a dedicated user-friendly

computer application. A similar application was used for the manual rating of microinfarcts and the evaluation of the automatic results. Manual delineations of the midsagittal plane and surface were also created with dedicated computer applications.

There are multiple advantages in using computer applications for manual image processing and the creation of digital ratings and segmentations. The availability of digitized ratings and segmentations can be effectively used to create new (semi-)automated image processing techniques. Both training and testing of such techniques rely heavily on digital manual ratings. Although not every manual rating or segmentation task is translated into a (semi-)automated technique, having digital ratings available can still be useful to quickly assess results or to compute quantitative measures.

Determining the quality of manual ratings or segmentations is a difficult task. Experts on a particular topic can review the results generated by others, but this results in a subjective description of the quality. Computing the inter-rater agreement is an attempt to quantify such reviews and to measure quality. However, the fact that multiple observers (or experts) produce the same results does not necessarily imply that these results are of high quality. Despite this drawback, determining the inter-rater agreement and organizing consensus meetings about ratings or segmentations should assure a high quality.

Ironically, measures for inter-rater agreement that are used to determine the quality of microbleed ratings are itself not always of high quality, as demonstrated in Chapter 2. Inter-rater agreement measures should only measure differences in the performance of observers and should not be influenced by other factors, such as the data that is being rated. The use of the DSC, which is a de-facto standard for evaluating (semi-)automated segmentation techniques, proved to be a good solution. To compute the DSC for ratings made by two observers, the number of microbleeds detected by each observer individually is compared to the number of microbleeds that is rated by both observers (the overlap).

### **8.2.2 Semi-automatic image processing**

The semi-automatic image processing techniques that were presented for the detection of microbleeds and microinfarcts share a number of properties. Each technique included important features such as intensity, shape, and loc-

ation. Both microbleeds and microinfarcts have distinct features in each of these categories that make them stand out in the large amount of data that needed to be processed. These features were translated from their definitions for manual rating into semi-automatic image processing techniques. Next to this, each technique employed a similar work-flow, in which potential microbleed or microinfarct locations were detected automatically and the results were assessed by a human observer. A trade-off could be made between sensitivity and required human observer time (related to the number of false positives), by choosing parameter settings.

### **From definition to detection**

Microbleeds are defined as round, hypointense lesions; criteria that were elegantly captured by the radial symmetry transform. Especially the high spatial resolution of 7 T images appreciated the spherical shape of microbleeds. This shape was less pronounced at 3 T brain MRI, owing to the large slice thickness that resulted in anisotropic voxels. Overall, the direct translation of the problem definition (microbleed detection) into an image processing technique resulted in high quality results. To overcome issues caused by the anisotropic voxel size and absence of additional echo times of 3 T brain MR images, an additional step using minimum intensity projection images was added. This step was directly derived from the manual rating procedure, a typical example of translating expert knowledge into a semi-automatic technique.

The definition of microinfarcts is still subject of discussion among human observers, which hampers detection results that are widely accepted. An observer may decide that detected locations are not actual microinfarcts, but microinfarct resemblances or locations on which a decision cannot be made for various reasons. Consequently, this resulted in a relatively high number of false positives, but it is important to highlight such locations and leave the final decision to the observer. An important advantage of the presented detection technique is the detection of extra microinfarcts. Some of these locations were absent in the ground truth, but turned out to be true microinfarcts. Presenting microinfarct resemblances to an observer might help in developing a better definition of true microinfarcts and sharpens the criteria for detection.

A common problem in the detection of microbleeds and microinfarcts on

7T brain MRI was the segmentation of relevant tissue classes in these images. To avoid false positive detections at locations in the images where microbleeds or microinfarcts cannot occur, a proper segmentation was needed. However, there are no dedicated solutions yet to segment a 7T brain MR image and other (suboptimal) solutions were used. For example, SPM8 (Ashburner et al. 2005) was used to directly segment brain tissue on T1-weighted 7T images and subsequent image processing was applied to improve segmentation results. Although SPM8 was not designed to process 7T images, the results were sufficient to be used for microbleed detection. Segmentation of grey matter for the purpose of microinfarct detection was performed on the 3T images that were available for each subject. FreeSurfer (Fischl 2012) was used to segment the 3T images and the elastix toolbox for rigid and nonrigid registration of images (Klein et al. 2010) was used to transform the segmentation results to the corresponding 7T image of each subject. The quality of these segmentations was verified by assessing the DSC for a number of slices, which was around 0.8 and that is considered good. However, having a dedicated and standardized way to segment 7T brain MR images directly would be preferable.

### **Extra positives**

The detection of extra microbleeds and microinfarcts was a key finding of the presented semi-automated techniques. These extra findings were not seen before by human observers. This demonstrates the high difficulty involved in visual rating and the added value of semi-automated techniques. Improvements of semi-automatic microbleed detection at 7T made it possible to push the sensitivity to 90% (Kuijf et al. 2013c), an achievement that human observers are unlikely to match. The semi-automatic detection of microinfarcts was a proof-of-principle study and is likely to be expanded with more subjects and additional features to improve the detection. The technique presented in this thesis has already been applied to a larger group of subjects ( $n = 56$ ) to complement visual rating of microinfarcts (van Veluw et al. 2013a). Visual rating by two observers identified 40 microinfarcts. Semi-automatic detection identified 28 (70%) of those, but identified an additional 37 true microinfarct locations. Of the twelve false negative locations, four were later rejected as true microinfarcts in a consensus meeting. Further improvements to the semi-automatic detection should focus on reducing the number of false

positives and increasing the sensitivity. As human observers become more experienced in visual detection of microinfarcts, which is still a relatively new finding on 7 T, the definitions that can be used in semi-automatic detection will become more clear. Additional features and constraints can be added that reduce the number of false positive locations.

There will be a point in the future at which automated techniques have a superior sensitivity in detecting abnormalities on high resolution brain MR images, especially on images acquired at 7 T. Here, the enormous amount of data and high level of detail cannot be assessed solely by a human observer. However, precautions must be taken to ensure the quality of semi-automated approaches. The deterministic nature of image processing technique guarantees that everything that adheres to the definition will be detected and nothing else. This approach relies heavily on a good, sound definition of what to detect. If the criteria for detection are not good or too strict, false negative results will occur. This is already demonstrated by the detection of extra positives during microbleed or microinfarct detection. If the method was tuned such that these additional findings were not detected, experts would still be satisfied with the results. However, now we know that there were still undetected lesions after manual detection. These undetected lesions were detected. Interestingly, this leads to a new question: are there even more undetected lesions, considering that the semi-automatic detection is also not perfect? The answer is probably yes. Further improvements to the detection techniques might reveal them, but this situation will repeat itself, *ad infinitum*. However, future iterations of a detection technique should not be based upon detections by a previous iteration of itself. This situation must be avoided to prevent a bias in the results. High quality manual ratings by multiple observers are likely to reduce this effect to a minimum, but probably not eliminate fully.

### **Multiple images**

An advantage of 7 T brain MRI over 3 T for the detection of microbleeds was the availability of two echo times. This enabled a reduction in the number of false positives that were caused by noise or microbleed resemblances that were only present on one of the two echo times. A microbleed has to be visible on both echo times, so this was added as a constraint to the technique. In some cases this theory did not match the actual situation, because the bloom-

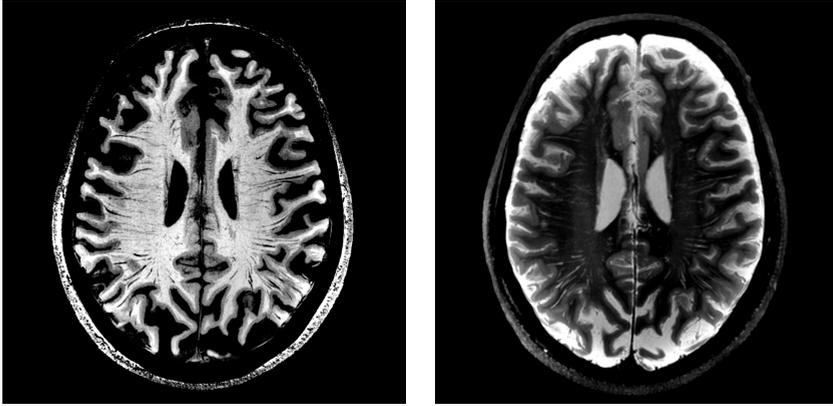
ing effect of nearby vessels could overshadow microbleeds on images with a longer echo time. The characteristic spherical shape of a microbleed would then disappear. The clinical scan protocol that was used for 3 T brain MRI did not include additional echo time images or other sequences that visualized microbleeds. Additional image processing was introduced to constrain the number of false positives to an acceptable amount. Since many false positives occur at the location of vessels, a potential improvement is to include a vessel mask to suppress these locations. This mask might be based on the same image or on other scan sequences (for example a proton density weighted MR image (Ghafaryasl et al. 2012)).

The availability of different MR sequences acquired at the same time is highly advantageous for (semi-)automatic techniques. The positive effect of having two echo times available on microbleed detection was already described. Detection of microinfarcts was performed on 7 T FLAIR and T2-weighted images, on which microinfarcts are best visible. Combining these images resulted in a better detection. However, the T2-weighted sequence is not always available. There is an open challenge in microinfarct detection based on only the FLAIR sequence. In general, combining information from multiple sources results in a better detection. Note that a lesion or structure of interest does not have to be visible on all sequences, because its absence on one and presence on another might be characteristic as well.

Besides the magnitude information of various MR sequences, it is also interesting to look at the phase information of MR images. By using a phase gradient magnitude image, the number of false positives in microbleed detection could be reduced at no additional cost (Bakker et al. 2008; Kuijf et al. 2012b; de Leeuw et al. 2009). Phase information is inherent in MR acquisition, but it is not always stored to disk. Developments in quantitative susceptibility mapping (QSM) are likely to improve the quality of images and microbleed detection techniques (Kuijf et al. 2013d; Liu et al. 2012). QSM is a technique that highlights the iron deposit of which a microbleed consists and thereby improves the visibility of microbleeds on brain MRI.

## Prospectives

As demonstrated in this thesis, the translation of definitions and guidelines for visual rating into image processing techniques for semi-automatic detection turned out to be highly valuable. Although semi-automatic detection



**Figure 8.1:** Enlarged perivascular spaces visualized with 7T brain MRI, showing as elongated tubes in the white matter. Left: 5.8 mm thick minimum-intensity projection slab of a 7T magnetization-prepared rapid acquisition with gradient echo (MP-RAGE) image, showing hypointense perivascular spaces. Right: 5.8 mm thick maximum-intensity projection slab of a 7T 3D turbo spin echo (TSE) image, showing hyperintense perivascular spaces.

in this thesis was limited to microbleeds and microinfarcts, general concepts and techniques can be applied in other settings as well. Features as intensity, shape, and location can be derived for almost everything that is visible and needs to be detected on brain MRI. Future semi-automatic techniques might focus on the detection of perivascular spaces, defined as elongated tubes visible on high-resolution 7T images. An example is shown in Figure 8.1. Enlarged perivascular spaces are an MRI marker of small vessel disease and associated with dementia (Bouvy et al. 2013). They have distinct intensities on different MR sequences and are located in the subcortical white matter. With the advances in acquisition, the need for semi-automatic image processing techniques will increase further as it will become less feasible to process these images manually. Individual cases or small scale studies might still employ manual assessment, but larger studies—or studies including longitudinal data that require reproducible assessments—will need robust, reliable, and validated image processing techniques.

A variant on semi-automatic image processing is interactive image processing. Hereby the observer interacts with an automated detection system by applying corrections. These corrections are returned to the automated system and applied to similar situations and in subsequent steps. The perform-

ance of the automated part of the system should improve gradually as the automated system adapts to the observer.

Another possibility is rapid application development or rapid prototyping of semi-automated techniques. Semi-automatic image processing techniques are deployed quickly to clinical researchers, whilst features and functionality are continuously added. Immediate and early feedback from clinical researchers can lead to improvements in the technique and this routine can be repeated. Win-win potential exists because clinical researchers have quick access to novel techniques (that can improve performance or save time) and image processing researchers receive suggestions for improvement. With every iteration, the benefits for both parties increase. Adoption of new techniques by clinical researchers might go faster and the semi-automated approach ensures supervised quality control.

This feedback-loop might be extended to the acquisition of brain MR images. Requirements imposed by clinical researchers for image processing techniques might demand additional or different brain MRI acquisitions in order to fulfil them. Currently, image processing techniques are mostly applied to brain MR images that are already acquired, but these might be sub-optimal for the (semi-)automated image processing task for a particular problem. In a rapid prototyping situation with continuous feedback, this can be quickly detected and acted upon. Such feedback-loops are already employed in the setting of image guided treatments, such as radiotherapy (e.g. motion correction caused by breathing) or high intensity focused ultrasound (HIFU) treatments (to monitor temperature).

### 8.2.3 Automatic image processing

Automatic image processing techniques yield results without the intervention of a human observer. This puts high requirements on the robustness and reliability of an automatic technique, since erroneous outcomes will not be noticed. For example in cases where automated techniques are used as an initial step prior to further processing. Next to this, the performance of automated methods should be high enough to be beneficial over manual or semi-automatic image processing.

Three automated techniques for the extraction of the midsagittal plane were implemented and described in Chapter 6. These techniques form a typical example of automatic image processing techniques, requiring robustness,

reliability, and a good performance. These aspects have been evaluated by Hu et al.; Liu et al.; Volkau et al. However, each technique was evaluated on a specific type of images and different evaluation measures were used, making it hard to compare results. By implementing various techniques and evaluating the performance on the same dataset, the techniques could be compared. The results showed that the performance of automated techniques did not differ from the inter-observer variability, i.e. the automated techniques are as good as human observers. This is a respectable achievement for automated techniques, since further improvement is not measurable.

It also highlights the importance of having manual ratings, segmentations, or delineations from multiple human observers and a good measurement of the inter-observer variability. It is sensible to compare the performance of automated techniques to the inter-observer variability. An automated technique can only be as good as the 'gold standard' itself, which is not necessarily perfect. This aspect is often overlooked.

The performance of the technique to extract the midsagittal surface is high, but the remaining errors are larger than the inter-observer variation. While this technique performs much better than the midsagittal plane, there is still room for improvement. The robustness and reliability of the technique were high and subjects with extreme brain torque were processed correctly. There was no significant difference in performance between subjects with or without considerable brain torque.

Although the automatically extracted midsagittal surface is not similar to manual delineations (given the remaining error), it is questionable to make further adaptations to improve this. The famous 80/20 rule (also known as the Pareto principle) dictates that this final 20% improvement might require 80% of the total resources invested, thereby potentially threatening the robustness and reliability of the technique. Future work should focus on translating this technique to other imaging modalities such as CT or CT perfusion imaging. Currently, manual hemisphere separation (or non-robust automated separations) is used in practice to study brain perfusion after stroke on CT. Notably CT perfusion analyses will highly benefit from a decent extraction of the midsagittal surface. Another possible application of this technique is the analysis of the midsagittal surface tortuosity, since brain asymmetry is speculated to be linked to handedness and other cognitive functions (Balzeau et al. 2012; Galaburda et al. 1978).

### **Prospectives**

Since erroneous outcomes of automated techniques are not detected, they are expected to be robust and reliable. However, achieving robustness and reliability can be a difficult task, because all possible exceptions to the standard procedures have to be captured in the technique. Even situations that might not be represented in the given training set. Humans are incredibly good at detecting new and unexpected situations, but a computer program will handle every situation according to the rules given.

Instead of focussing on a high robustness and reliability, which are nevertheless important, future developments might look into the detection of failure or erroneous outcomes. If the mere failure of an automated technique can be detected automatically, the consequences of continuing further computations with wrong values can be avoided. An automated technique becomes very powerful if a mechanism for self-detection of failure is functioning properly. In such situations, automated processing can be halted and a human observer should look into the situation at hand.

### **8.3 Conclusion**

This thesis described a number of image processing techniques for quantification and assessment of brain MRI, distinguishing manual, semi-automatic, and automatic techniques. These techniques have been developed as to assist human observers in their existing workflow.

In developing new image processing techniques, the evermore famous question is: what is the goal? The benefit of techniques should be clear, but can be versatile: ranging from a reduction in required human observer time to improvements in sensitivity, performance, or usability. With the specified goals and requirements in mind, an image processing technique can be designed.

Image processing techniques developed in the forefront of clinical and scientific research show a gradual transition from manual, to semi-automatic, and eventually to automatic techniques. In this “sliding window”, new findings are often detected manually at first. Upon increasing interest and the generation of additional data, semi-automatic techniques may be introduced to support the demand for improved detection. Once the definition of a new finding and its criteria for detection are crystallized, automatic tech-

niques may be employed to facilitate large-scale or fully automated image processing.

This “sliding window”-effect will have a continuous impact on the focus of researchers in both the clinical and technical environment. Solid cooperation and proper feedback between the academic researchers involved will lead to solutions that eventually find their way, through all stages, into daily clinical and research use.



# Bibliography

- Anbeek, P. et al. (2005). 'Probabilistic segmentation of brain tissue in MR imaging'. In: *NeuroImage* 27.4, pp. 795–804.
- Ardekani, B. A. et al. (Dec. 1997). 'Automatic Detection of the Mid-Sagittal Plane in 3-D Brain Images'. In: *IEEE Transactions on Medical Imaging* 16.6, pp. 947–952.
- Ashburner, J. and K. J. Friston (2000). 'Voxel-based morphometry - The methods'. In: *NeuroImage* 11.6 I. Cited By (since 1996): 2013, pp. 805–821.
- (2005). 'Unified segmentation'. In: *NeuroImage* 26, pp. 839–851.
- Bakker, C. J. G. et al. (2008). 'Phase gradient mapping as an aid in the analysis of object-induced and system-related phase perturbations in MRI'. In: *Physics in Medicine and Biology* 53.18, N349.
- Balzeau, A., E. Gilissen and D. Grimaud-Hervé (Jan. 2012). 'Shared Pattern of Endocranial Shape Asymmetries among Great Apes, Anatomically Modern Humans, and Fossil Hominins'. In: *PLoS ONE* 7.1, e29581.
- Barnes, S. R. S. et al. (2011). 'Semiautomated detection of cerebral microbleeds in magnetic resonance images'. In: *Magnetic Resonance Imaging* 29.6, pp. 844–852.
- Bergo, F. P. G. et al. (2009). 'Fast, Accurate and Precise Mid-Sagittal Plane Location in 3D MR Images of the Brain'. In: *Biomedical Engineering Systems and Technologies*. Ed. by A. Fred, J. Filipe and H. Gamboa. Vol. 25. Communications in Computer and Information Science. Springer Berlin Heidelberg, pp. 278–290.
- Bian, W. et al. (2013). 'Computer-aided detection of radiation-induced cerebral microbleeds on susceptibility-weighted MR images'. In: *NeuroImage: Clinical* 2, pp. 282–290.
- Bochkanov, S. and V. Bystritsky (2012). *ALGLIB* ([www.alglib.net](http://www.alglib.net)).
- Bouvy, W. H. et al. (2013). 'Visualization of Perivascular Spaces and perforating arteries with 7 Tesla MRI'. In: *submitted*.
- Bresser, J. de et al. (2013). 'Visual Cerebral Microbleed Detection on 7T MR Imaging: Reliability and Effects of Image Processing'. In: *American Journal of Neuroradiology* 34.6, E61–E64.
- Brummer, M. E. (Mar. 1991). 'Hough transform detection of the longitudinal fissure in tomographic head images'. In: *IEEE Transactions on Medical Imaging* 10.1, pp. 74–81.
- Brundel, M. et al. (2012a). 'Cerebral microinfarcts: a systematic review of neuropathological studies'. In: *Journal of Cerebral Blood Flow and Metabolism* 32.3, pp. 425–436.
- Brundel, M. et al. (2012b). 'High Prevalence of Cerebral Microbleeds at 7Tesla MRI in Patients with Early Alzheimer's Disease'. In: *Journal of Alzheimer's Disease* 31, pp. 259–263.
- Charidimou, A. and D. J. Werring (2011). 'Cerebral microbleeds: detection, mechanisms and clinical challenges'. In: *Future Neurology* 6, pp. 587–611.

- Cocosco, C. A., A. P. Zijdenbos and A. Evans (2003). 'A fully automatic and robust brain MRI tissue classification method'. In: *Medical Image Analysis 7.4*. Medical Image Computing and Computer Assisted Intervention, pp. 513–527.
- Conijn, M. M. A. et al. (2010). 'Visualization of cerebral microbleeds with dual-echo T2\*-weighted magnetic resonance imaging at 7.0 T'. In: *Journal of Magnetic Resonance Imaging* 32.1, pp. 52–59.
- Conijn, M. M. A. et al. (2011). 'Cerebral Microbleeds on MR Imaging: Comparison between 1.5 and 7T'. In: *AJNR Am J Neuroradiol* 32.6, pp. 1043–1049.
- Cordonnier, C., R. A.-S. Salman and J. M. Wardlaw (2007). 'Spontaneous brain microbleeds: systematic review, subgroup analyses and standards for study design and reporting'. In: *Brain* 130.8, pp. 1988–2003.
- Cordonnier, C. et al. (2009). 'Improving Interrater Agreement About Brain Microbleeds: development of the Brain Observer MicroBleed Scale (BOMBS)'. In: *Stroke* 40.1, pp. 94–99.
- Dale, A. M., B. Fischl and M. I. Sereno (Feb. 1999). 'Cortical surface-based analysis: I. Segmentation and surface reconstruction'. In: *Neuroimage* 9.2, pp. 179–194.
- De Reuck, J. et al. (2011). 'Comparison of 7.0 T T2\*-Magnetic Resonance Imaging of Cerebral Bleeds in Post-Mortem Brain Sections of Alzheimer Patients with Their Neuropathological Correlates'. In: *Cerebrovasc Dis* 31.5, pp. 511–517.
- Dice, L. R. (1945). 'Measures of the Amount of Ecological Association Between Species'. In: *Ecology* 26.3, pp. 297–302.
- Ekin, A. (2006). 'Feature-based brain mid-sagittal plane detection by RANSAC'. In: *14th European Signal Processing Conference*.
- Fazekas, F. et al. (Apr. 1999). 'Histopathologic Analysis of Foci of Signal Loss on Gradient-Echo T2\*-Weighted MR Images in Patients with Spontaneous Intracerebral Hemorrhage: Evidence of Microangiopathy-Related Microbleeds'. In: *American Journal of Neuroradiology* 20, pp. 637–642.
- Fischl, B. (2012). 'FreeSurfer'. In: *NeuroImage* 62.2, pp. 774–781.
- Folstein, M. F., S. E. Folstein and P. R. McHugh (1975). "'Mini-mental state": A practical method for grading the cognitive state of patients for the clinician'. In: *Journal of Psychiatric Research* 12.3, pp. 189–198.
- Fonov, V. S. et al. (2009). 'Unbiased nonlinear average age-appropriate brain templates from birth to adulthood'. In: *NeuroImage* 47, Supplement 1. Organization for Human Brain Mapping 2009 Annual Meeting, S102.
- Fonov, V. S. et al. (2011). 'Unbiased average age-appropriate atlases for pediatric studies'. In: *NeuroImage* 54.1, pp. 313–327.
- Galaburda, A. M. et al. (1978). 'Right-left asymmetries in the brain'. In: *Science* 199.4331, pp. 852–856.
- Geerlings, M. I. et al. (2010). 'Brain volumes and cerebrovascular lesions on {MRI} in patients with atherosclerotic disease. The SMART-MR study'. In: *Atherosclerosis* 210.1, pp. 130–136.
- Ghafaryasl, B. et al. (2012). 'A computer aided detection system for cerebral microbleeds in brain MRI'. In: *9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 138–141.
- Greenberg, S. M. et al. (2009). 'Cerebral microbleeds: a guide to detection and interpretation'. In: *The Lancet Neurology* 8.2, pp. 165–174.
- Gregoire, S. M. et al. (2009). 'The Microbleed Anatomical Rating Scale (MARS): Reliability of a tool to map brain microbleeds'. In: *Neurology* 73.21, pp. 1759–1766.

- Hu, Q. and W. L. Nowinski (2003). 'A rapid algorithm for robust and automatic extraction of the midsagittal plane of the human cerebrum from neuroimages based on local symmetry and outlier removal'. In: *NeuroImage* 20.4, pp. 2153–2165.
- Hughes, C. P. et al. (1982). 'A new clinical scale for the staging of dementia.' In: *The British Journal of Psychiatry* 140.6, pp. 566–72.
- Jayasuriya, S., A. Liew and N. Law (2013). 'Brain symmetry plane detection based on fractal analysis'. In: *Computerized Medical Imaging and Graphics*, in press.
- Joshi, S. et al. (2003). 'Structural and radiometric asymmetry in brain images'. In: *Medical Image Analysis* 7.2, pp. 155–170.
- Junck, L. et al. (1990). 'Correlation Methods for the Centering, Rotation, and Alignment of Functional Brain Images'. In: *Journal of Nuclear Medicine* 31.7, pp. 1220–1226.
- King, D. E. (2012). *Dlib C++ library* ([www.dlib.net](http://www.dlib.net)).
- Klein, S. et al. (Jan. 2010). 'elastix: A Toolbox for Intensity-Based Medical Image Registration'. In: *IEEE Transactions on Medical Imaging* 29.1, pp. 196–205.
- Knudsen, K. A. et al. (2001). 'Clinical diagnosis of cerebral amyloid angiopathy: Validation of the Boston Criteria'. In: *Neurology* 56.4, pp. 537–539.
- Kuijf, H. J., M. A. Viergever and K. L. Vincken (2013a). 'Automatic extraction of the curved midsagittal brain surface on MR images'. In: *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*. Vol. 7766. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 225–232.
- Kuijf, H. J. et al. (2011). 'Detecting cerebral microbleeds in 7.0 T MR images using the radial symmetry transform'. In: *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 758–761.
- Kuijf, H. J. et al. (2012a). 'Efficient detection of cerebral microbleeds on 7.0T MR images using the radial symmetry transform'. In: *NeuroImage* 59.3, pp. 2266–2273.
- Kuijf, H. J. et al. (2012b). 'Powerful detection of cerebral microbleeds on 7.0T MR phase gradient magnitude images using the radial symmetry transform'. In: *Proc. Intl. Soc. Mag. Reson. Med.* Vol. 20, p. 1066.
- Kuijf, H. J. et al. (2013b). 'Assessment of methods to extract the mid-sagittal plane from brain MR images'. In: *Proc. SPIE*. Vol. 8673, 86731K.
- Kuijf, H. J. et al. (2013c). 'Observer performance in semi-automated microbleed detection'. In: *Proc. SPIE*. Vol. 8673, p. 867313.
- Kuijf, H. J. et al. (2013d). 'Semi-automated detection of cerebral microbleeds on 3.0 T MR images'. In: *PLoS ONE* 8.6, e66610.
- Leeuw, H. de et al. (2009). 'High resolution phase gradient mapping as a tool for the detection and analysis of local field disturbances'. In: *Proc. Intl. Soc. Mag. Reson. Med.* Vol. 17, p. 261.
- Liang, L. et al. (2007). 'Automatic segmentation of left and right cerebral hemispheres from MRI brain volumes using the graph cuts algorithm'. In: *NeuroImage* 34.3, pp. 1160–1170.
- Liao, C.-C. et al. (2010). 'Automatic recognition of midline shift on brain CT images'. In: *Computers in Biology and Medicine* 40.3, pp. 331–339.
- Liu, D. C. and J. Nocedal (1989). 'On the limited memory BFGS method for large scale optimization'. English. In: *Mathematical Programming* 45.1-3, pp. 503–528.
- Liu, T. et al. (2012). 'Cerebral Microbleeds: Burden Assessment by Using Quantitative Susceptibility Mapping'. In: *Radiology* 262.1, pp. 269–278.

- Liu, Y., R. T. Collins and W. E. Rothfus (1998). 'Automatic bilateral symmetry (midsagittal) plane extraction from pathological 3D neuroradiological images'. In: *Proc. SPIE* 3338, pp. 1528–1539.
- Liu, Y., R. Collins and W. E. Rothfus (Mar. 2001). 'Robust Midsagittal Plane Extraction from Normal and Pathological 3D Neuroradiology Images'. In: *IEEE Transactions on Medical Imaging* 20.1, pp. 175–192.
- Loy, G. and A. Zelinsky (2003). 'Fast radial symmetry for detecting points of interest'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25.8, pp. 959–973.
- McAuley, G. et al. (2011). 'Iron quantification of microbleeds in postmortem brain'. In: *Magnetic Resonance in Medicine* 65.6, pp. 1592–1601.
- McKhann, G. et al. (1984). 'Clinical diagnosis of Alzheimer's disease'. In: *Neurology* 34.7, p. 939.
- Minoshima, S. et al. (1992). 'An Automated Method for Rotational Correction and Centering of Three-Dimensional Functional Brain Images'. In: *Journal of Nuclear Medicine* 33.8, pp. 1579–1585.
- Mykkänen, J. et al. (2005). 'Automatic extraction of brain surface and mid-sagittal plane from PET images applying deformable models'. In: *Computer Methods and Programs in Biomedicine* 79.1, pp. 1–17.
- Nagel, B. J. et al. (2013). 'Hemispheric lateralization of verbal and spatial working memory during adolescence'. In: *Brain and Cognition* 82.1, pp. 58–68.
- Neary, D. et al. (1998). 'Frontotemporal lobar degeneration'. In: *Neurology* 51.6, pp. 1546–1554.
- Nowinski, W. L. et al. (2006). 'Rapid and Automatic Calculation of the Midsagittal Plane in Magnetic Resonance Diffusion and Perfusion Images'. In: *Academic Radiology* 13.5, pp. 652–663.
- Petersen, R. C. et al. (1999). 'Mild Cognitive Impairment: Clinical Characterization and Outcome'. In: *Arch Neurol* 56.3, pp. 303–308.
- Prima, S., S. Ourselin and N. Ayache (Feb. 2002). 'Computation of the mid-sagittal plane in 3-D brain images'. In: *IEEE Transactions on Medical Imaging* 21.2, pp. 122–138.
- Prima, S., S. Ourselin and N. Ayache (2000). 'Computation of the Mid-Sagittal Plane in 3D Images of the Brain'. In: *Computer Vision - ECCV 2000*. Ed. by D. Vernon. Vol. 1843. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, pp. 685–701.
- Puspitasari, F. et al. (2009). 'Robust calculation of the midsagittal plane in CT scans using the Kullback–Leibler's measure'. In: *International Journal of Computer Assisted Radiology and Surgery* 4 (6), pp. 535–547.
- Rentería, M. E. (May 2012). 'Cerebral Asymmetry: A Quantitative, Multifactorial, and Plastic Brain Phenotype'. In: *Twin Research and Human Genetics* 15 (Special Issue 03), pp. 401–413.
- Ritter, F. et al. (Nov. 2011). 'Medical Image Analysis: A Visual Approach'. In: *IEEE Pulse* 2.6, pp. 60–70.
- Román, G. C. et al. (1993). 'Vascular dementia'. In: *Neurology* 43.2, p. 250.
- Schlösser, T. P. C. et al. (2013). 'Quantitative Analysis of the Closure Pattern of the Neurocentral Junction as related to Pre-existent Rotation in the Normal Immature Spine'. In: *The Spine Journal* 13.7, pp. 756–763.
- Seghier, M. L. et al. (Mar. 2011). 'Microbleed Detection Using Automated Segmentation (MIDAS): A New Method Applicable to Standard Clinical MR Images'. In: *PLoS ONE* 6.3, e17547.
- Simons, P. et al. (1999). 'Second Manifestations of ARterial disease (SMART) study: Rationale and design'. In: *European Journal of Epidemiology* 15 (9), pp. 773–781.
- Smith, E. E. et al. (2012). 'Cerebral microinfarcts: the invisible lesions'. In: *The Lancet Neurology* 11.3, pp. 272–282.

- Smith, S. M. (Nov. 2002). 'Fast robust automated brain extraction'. In: *Human Brain Mapping* 17.3, pp. 143–155.
- Smith, S. M. and M. Jenkinson (1999). 'Accurate Robust Symmetry Estimation'. In: *Medical Image Computing and Computer-Assisted Intervention - MICCAI'99*. Ed. by C. Taylor and A. Colchester. Vol. 1679. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, pp. 308–317.
- Sommer, I. et al. (2001). 'Handedness, language lateralisation and anatomical asymmetry in schizophrenia: Meta-analysis'. In: *The British Journal of Psychiatry* 178.4, pp. 344–351.
- Stegmann, M. B., K. Skoglund and C. Ryberg (2005). 'Mid-sagittal plane and mid-sagittal surface optimization in brain MRI using a local symmetry measure'. In: *Proc. SPIE 5747*. Ed. by J. M. Fitzpatrick and J. M. Reinhardt. San Diego, CA, USA, pp. 568–579.
- Phillips, H. J. (1964). 'A method of determining the midsagittal plane of the cranium'. In: *American Journal of Orthodontics* 50.10, p. 788.
- Teverovskiy, L. and Y. Liu (Apr. 2006). 'Truly 3D Midsagittal Plane Extraction for Robust Neuroimage Registration'. In: *3rd IEEE International Symposium on Biomedical Imaging: Macro to Nano, 2006*, pp. 860–863.
- Theysohn, J. M. et al. (2011). '7 tesla MRI of microbleeds and white matter lesions as seen in vascular dementia'. In: *Journal of Magnetic Resonance Imaging* 33.4, pp. 782–791.
- Toga, A. W. and P. M. Thompson (2003). 'Mapping brain asymmetry'. In: *Nature Reviews Neuroscience* 4, pp. 37–48.
- Tuzikov, A. V., O. Colliot and I. Bloch (2002). 'Brain symmetry plane computation in MR images using inertia axes and optimization'. In: *Pattern Recognition, 2002. Proceedings. 16th International Conference on*. Vol. 1, 516–519 vol.1.
- (2003). 'Evaluation of the symmetry plane in 3D MR brain images'. In: *Pattern Recognition Letters* 24.14, pp. 2219–2233.
- Veluw, S. J. van et al. (2013a). 'Cerebral Cortical Microinfarcts at 7Tesla MRI in Patients with early Alzheimer's Disease'. In: *Journal of Alzheimer's Disease*, (in press).
- Veluw, S. J. van et al. (2013b). 'In vivo detection of cerebral cortical microinfarcts with high-resolution 7T MRI'. In: *Journal of Cerebral Blood Flow and Metabolism* 33, pp. 322–329.
- Vernooij, M. W. et al. (2008). 'Prevalence and risk factors of cerebral microbleeds: The Rotterdam Scan Study'. In: *Neurology* 70.14, pp. 1208–1214.
- Veta, M. et al. (Apr. 2011). 'Marker-controlled watershed segmentation of nuclei in H&E stained breast cancer biopsy images'. In: *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pp. 618–621.
- Visser, F. et al. (July 2010). 'High-resolution magnetization-prepared 3D-FLAIR imaging at 7.0 Tesla'. In: *Magnetic Resonance in Medicine* 64.1, pp. 194–202.
- Volkau, I. et al. (2006). 'Extraction of the midsagittal plane from morphological neuroimages using the Kullback-Leibler's measure'. In: *Medical Image Analysis* 10.6, pp. 863–874.
- Vos, P. C. et al. (2013). 'Automatic detection and segmentation of ischemic lesions in computed tomography images of stroke patients'. In: *Proc. SPIE 8670, Medical Imaging 2013: Computer-Aided Diagnosis*, p. 867013.
- Wardlaw, J. M. et al. (2006). 'Cerebral Microbleeds Are Associated With Lacunar Stroke Defined Clinically and Radiologically, Independently of White Matter Lesions'. In: *Stroke* 37.10, pp. 2633–2636.
- Wardlaw, J. M. et al. (2013). 'Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration'. In: *The Lancet Neurology* 12.8, pp. 822–838.

- Wisse, L. E. M. et al. (2012). 'Subfields of the hippocampal formation at 7 T MRI: In vivo volumetric assessment'. In: *NeuroImage* 64.4, pp. 1043–1049.
- Zhang, T. and G. Nagy (2004). 'Surface tortuosity and its application to analyzing cracks in concrete'. In: *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004*. Vol. 2, pp. 851–854.
- Zhang, Y. and Q. Hu (2008). 'A PCA-based approach to the representation and recognition of MR brain midsagittal plane images'. In: *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3916–3919.
- Zwanenburg, J. J. M. et al. (2009). '3d flair at 7 tesla highlights peripheral layers of the cortex'. In: *Proc. Intl. Soc. Mag. Reson. Med.* P. 276.

# List of abbreviations

<b>AE</b>	Angular error
<b>BET</b>	Brain extraction tool
<b>BOMBS</b>	Brain observer microbleed scale
<b>CMB</b>	Cerebral microbleed
<b>CMI</b>	Cortical cerebral microinfarct
<b>CSF</b>	Cerebrospinal fluid
<b>DSC</b>	Dice similarity coefficient
<b>FROC</b>	Free-response receiver operating characteristic
<b>ICC</b>	Intraclass correlation coefficient
<b>IF</b>	Interhemispheric fissure
$\kappa$	Cohen's kappa coefficient
<b>KL</b>	Kullback–Leibler divergence
<b>MARS</b>	Microbleed anatomical rating scale
<b>minIP</b>	Minimum intensity projection
<b>MRI</b>	Magnetic resonance imaging
<b>MSP</b>	Midsagittal plane
<b>RST</b>	Radial symmetry transform
<b>SMART</b>	Second manifestations of arterial disease
<b>VE</b>	Volume error



# Samenvatting

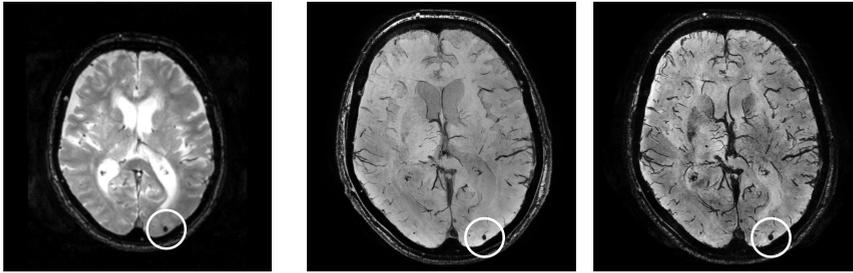
## Introductie

Door middel van *magnetic resonance imaging* (MRI) kunnen er 3-dimensionale afbeeldingen worden gemaakt van het menselijk brein. Dit gebeurt zonder het lichaam te openen, binnen te dringen of anderszins te beschadigen. Hierdoor is het mogelijk om de anatomie en fysiologie van het brein te bestuderen in levende mensen.

In dit proefschrift worden beeldverwerkingstechnieken voor de kwantificatie en beoordeling van MRI beelden van de hersenen beschreven. Deze technieken worden gebruikt voor onderzoek naar de anatomie van het brein en eventuele ziektebeelden, in het bijzonder aandoeningen van de bloedvaten in de hersenen. Er wordt onderscheid gemaakt tussen verschillende varianten van beeldverwerkingstechnieken: handmatige, semiautomatische en automatische technieken.

**Handmatige beeldverwerking** Dit is de “traditionele” vorm van beeldverwerking. Er wordt een afbeelding van de hersenen gemaakt en een arts of onderzoeker beoordeelt deze door er naar te kijken. Hierbij kan er worden gezocht naar ziektebeelden of kunnen er handmatige metingen worden gedaan.

**Semiautomatische beeldverwerking** Bij semiautomatische beeldverwerking wordt de computer ingezet om bepaalde taken van een arts of onderzoeker over te nemen. De computer gaat op zoek naar afwijkingen in de afbeeldingen of verricht automatisch metingen. De resultaten hiervan worden aan de arts of onderzoeker getoond en deze kan vervolgens eventuele foutjes corrigeren of andere aanpassingen maken.



**Figuur 1:** Persoon met een microbloeding (omcirkeld), zichtbaar gemaakt met 3 T T2\*-gewogen MRI (links) en 7 T dual-echo gradiënt echo MRI (rechts).

**Automatische beeldverwerking** Hierbij verricht de computer automatisch alle handelingen, metingen en onderzoeken. Er komt geen arts of onderzoeker meer aan te pas om tot resultaten te komen.

Technische ontwikkelingen op het gebied van MRI-scanners en bijbehorende scanprotocollen hebben er voor gezorgd dat er steeds meer en betere afbeeldingen worden gemaakt. Het handmatig verwerken van al deze afbeeldingen wordt echter steeds moeilijker en neemt meer tijd in beslag. De kwaliteit van beoordelingen kan hiermee in gevaar komen. Goede beeldverwerkingstechnieken kunnen artsen en onderzoekers ondersteunen en zijn daarmee onmisbaar voor het waarborgen van de kwaliteit. Het overkoepelende doel in dit proefschrift is om vereisten voor de kwantificatie en beoordeling van brein MRI te vertalen naar beeldverwerkingstechnieken die in de praktijk kunnen worden toegepast.

## Microbloedingen

Een groot deel van dit proefschrift richt zich op de detectie van microbloedingen in het brein, zowel met handmatige als semiautomatische technieken. Microbloedingen zijn zeer kleine bloedingen die kunnen worden afgebeeld met gradiënt echo MRI. Op deze afbeeldingen zijn microbloedingen zichtbaar als kleine, zwarte, ronde puntjes, zoals te zien is in Figuur 1.

Onder artsen en onderzoekers is er een toenemende interesse voor deze microbloedingen, omdat ze geassocieerd zijn met vaatschade in het brein, de ziekte van Alzheimer, dementie en amyloïde angiopathie (zoals de Katwijkse

ziekte). Samen met deze toenemende interesse is er ook een groeiende vraag naar goede en betrouwbare technieken om microbloedingen op te sporen. Deze uitdaging wordt in dit proefschrift besproken.

## Handmatige opsporing

Voor het handmatig opsporen van microbloedingen wordt gebruik gemaakt van de Microbleed Anatomical Rating Scale (MARS); een papieren formulier dat wordt ingevuld tijdens het bekijken van een afbeelding van het brein van een patiënt. Deze procedure is echter zeer tijdsintensief (duurt vaak meer dan 30 min voor een 7 T afbeelding), is sterk afhankelijk van wie het doet en heeft een matige reproduceerbaarheid.

Om de kwaliteit van beoordelingen te kunnen garanderen wordt deze handmatige procedure door meerdere artsen of onderzoekers (beoordelaars) herhaalt. Als twee of meer beoordelaars dezelfde microbloedingen vinden, dan wordt aangenomen dat de beoordeling van hoge kwaliteit is. Als ze van mening verschillen en niet dezelfde microbloedingen vinden, dan is de beoordeling van een lagere kwaliteit. Deze kwaliteit kan in een getal worden uitgedrukt: de interbeoordelaarsbetrouwbaarheid.

In Hoofdstuk 2 wordt er gekeken naar bestaande methoden voor het berekenen van de interbeoordelaarsbetrouwbaarheid. Twee veelgebruikte methoden in microbloedingenonderzoek zijn Cohen's kappa coëfficiënt ( $\kappa$ ) en de intraclass correlatiecoëfficiënt (ICC). Echter is gebleken dat deze twee methoden niet nauwkeurig zijn als er meerdere microbloedingen per patiënt zijn of als er een *outlier* is (één patiënt met veel meer microbloedingen dan de rest). In deze situaties berekenen  $\kappa$  en ICC een hogere interbeoordelaarsbetrouwbaarheid dan daadwerkelijk het geval is. Door deze overschatting is het niet mogelijk om te concluderen dat een beoordeling van hoge kwaliteit is, omdat dit beïnvloed kan zijn door de gebruikte methode.

Een mogelijk alternatief is het gebruik van de Dice coëfficiënt. Deze methode kijkt niet alleen naar het totale aantal gevonden microbloedingen, maar ook of dit exact dezelfde microbloedingen op exact dezelfde locaties zijn. Daardoor treedt er geen overschatting op en kan de correcte interbeoordelaarsbetrouwbaarheid worden bepaald.

## Semiautomatische opsporing

Het handmatig opsporen van microbloedingen is een tijdsintensieve procedure, is afhankelijk van de beoordelaar en heeft vaak maar een matige kwaliteit. Dit geldt in het bijzonder voor het opsporen van microbloedingen op 7 T MRI afbeeldingen, die door de enorm hoge resolutie en hoeveelheid details lastig te beoordelen zijn. Een semiautomatische aanpak kan helpen om de opsporing te vereenvoudigen en de kwaliteit te verhogen.

In dit proefschrift wordt een dergelijke semiautomatische techniek gepresenteerd. Mogelijke microbloedingen worden automatisch opgespoord in een MRI afbeelding van de hersenen en deze locaties worden aan een arts of onderzoeker getoond. Deze beoordelaar bekijkt alle gevonden locaties en selecteert hieruit de echte microbloedingen en verwijdert eventuele foutieve locaties.

De techniek die hiervoor wordt gebruikt heet de *radial symmetry transform* (RST). De computer gaat hierbij pixel-voor-pixel op zoek naar kleine, zwarte, ronde puntjes in de afbeelding. Deze eigenschappen (klein, zwart, rond) worden uitgedrukt in één getal: de RST-waarde. Aan het einde wordt er een afkapwaarde gekozen en alle locaties met een RST-waarde die daar bovenuit komt zijn mogelijke microbloedingen.

In Hoofdstuk 3 wordt deze techniek toegepast op afbeeldingen die zijn gemaakt met een 3 T MRI scanner. Dit is een scanner die in de dagelijkse, klinische praktijk wordt gebruikt om patiënten mee te onderzoeken. Om de RST te testen zijn 72 patiënten gescand en de afbeeldingen zijn door 4 beoordelaars onderzocht op microbloedingen. In totaal werden er 152 microbloedingen gevonden in 38 patiënten. Elke individuele beoordelaar vond tussen de 39 en 86 % van deze microbloedingen. Dit duurde ongeveer 5 tot 10 min per patiënt per beoordelaar. De RST detecteerde, afhankelijk van de gekozen afkapwaarde, tussen de 65 en 84 % van alle microbloedingen. Het bekijken van alle gedetecteerde locaties en het verwijderen van de foutieve locaties duurde 1 tot 2 min.

In Hoofdstuk 4 wordt de RST toegepast op afbeeldingen gemaakt met een 7 T MRI scanner. Deze relatief nieuwe scanner wordt nog niet in de dagelijkse praktijk gebruikt, maar al wel veel voor wetenschappelijk onderzoek. De afbeeldingen die hiermee worden gemaakt zijn van een veel hogere resolutie en kwaliteit dan die van een 3 T scanner. Er is veel meer detail zichtbaar, maar dit zorgt er ook voor dat het moeilijker wordt om afbeeldingen handmatig te

beoordelen. Om de RST te testen zijn 18 patiënten gescand en de afbeeldingen zijn door 2 beoordelaars onderzocht op microbloedingen. Deze beoordelaars vonden in totaal 54 microbloedingen. De RST detecteerde er hier 35 van, maar vond ook nog eens 12 extra microbloedingen die eerder niet door de beoordelaars waren gezien. Dit brengt het totaal op 66 microbloedingen, waarvan 71 % werd gedetecteerd door de RST en 52 % en 67 % door de individuele beoordelaars. Het bekijken van de resultaten van de RST en het verwijderen van foutieve locaties duurde slechts 2 min per patiënt, tegenover 30 min voor een handmatige beoordeling. Verdere verbeteringen aan de RST (niet in dit proefschrift) hebben het percentage gedetecteerde microbloedingen omhoog gebracht naar 90 %.

Het meest bijzondere resultaat is de detectie van extra microbloedingen die nog niet eerder door beoordelaars waren gezien. Dit geeft aan hoe moeilijk het is om microbloedingen betrouwbaar handmatig op te sporen. Goede beeldverwerkingstechnieken zullen in de toekomst onmisbaar worden om betrouwbaar MRI beelden te kunnen beoordelen.

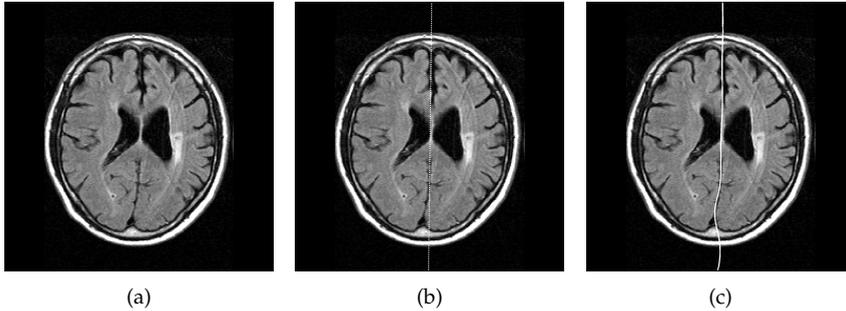
## Corticale microinfarcten

Microinfarcten worden ook wel omschreven als “de onzichtbare laesie” en “niet zichtbaar met het blote oog”. Deze zeer kleine infarcten komen waarschijnlijk veel voor bij oudere mensen en zijn gerelateerd aan vaatschade en dementie. Tot dusver konden ze alleen worden onderzocht met een microscoop, waarbij kleine gebiedjes afgestorven weefsel zichtbaar zijn.

Echter, microinfarcten blijken ook zichtbaar te zijn op afbeeldingen gemaakt met een 7 T MRI scanner. Door verschillende afbeeldingen van één patiënt te combineren kunnen microinfarcten worden gevonden, zoals te zien is in Figuur 5.1 op Pagina 59.

Het handmatig opsporen van microinfarcten gebeurt door alle drie de afbeeldingen gelijktijdig te bekijken. Omdat MRI afbeeldingen 3D zijn wordt dit vanuit verschillende richtingen gedaan, om niets te missen. Het beoordelen van deze enorme hoeveelheid afbeeldingen is erg moeilijk en zeer tijdsintensief. Dit staat een goede en betrouwbare beoordeling in de weg. Een semiautomatische aanpak zou hierbij kunnen helpen: de computer detecteert mogelijke microinfarcten en een beoordelaar bekijkt de resultaten.

Een eerste opzet tot zo'n methode is gepresenteerd in Hoofdstuk 5. Waar het handmatig opsporen van microinfarcten meer dan één uur per patiënt



**Figuur 2:** (a) Beide hersenhelften worden gescheiden door de interhemisferische spleet. (b) Het midsagittale vlak is een rechte lijn die beide hersenhelften van elkaar scheidt. (c) De interhemisferische spleet is niet per se een recht vlak, maar kan gekromd lopen.

duurde, kon het beoordelen van de resultaten van de semiautomatische methode in 5 tot 20 min worden gedaan. De methode vond ook extra microinfarcten die nog niet eerder gezien waren tijdens de handmatige beoordeling.

## Midsagittale vlak

De grote hersenen van het menselijk brein bestaan uit twee hersenhelften die van elkaar worden gescheiden door de interhemisferische spleet. Dit is te zien in Figuur 2(a). De twee hersenhelften lijken elkaars spiegelbeeld, maar er zijn ook verschillen tussen beide. Zo worden bepaalde functies door de ene of de andere hersenhelft uitgevoerd. Anatomische verschillen tussen de hersenhelften kunnen wijzen op de aanwezigheid van een pathologie, bijvoorbeeld een infarct of bloeding. Om de twee hersenhelften afzonderlijk te kunnen onderzoeken of met elkaar te kunnen vergelijken, is het nodig om ze automatisch te kunnen scheiden. Er wordt vaak een rechte lijn (in 3D een vlak) getrokken door de interhemisferische spleet: het midsagittale vlak. Dit is te zien in Figuur 2(b).

Er bestaan meerdere methoden voor het automatisch bepalen van het midsagittale vlak en in Hoofdstuk 6 zijn deze met elkaar vergeleken. Voor een grote groep patiënten is het midsagittale vlak met de hand aangegeven door twee verschillende onderzoekers. Omdat niet iedere onderzoeker het midsagittale vlak op dezelfde plek zet, zijn de verschillen tussen beide onderzoekers bepaald. Vervolgens hebben drie automatische technieken het midsagittale

vlak berekend voor elke patiënt. Voor iedere afzonderlijke methode was het verschil tussen de midsagittale vlakken van de automatische methode en de onderzoeker niet groter dan het verschil tussen de onderzoekers onderling. Oftewel: de automatische methoden waren net zo goed als de onderzoekers zelf.

### **Niet recht, maar krom**

Zoals in Figuur 2 te zien is, is het midsagittale vlak niet per se kaarsrecht. In veel mensen zijn de rechter-frontale en linker-occipitale hersenkwabben net iets groter dan die aan de tegenovergestelde zijde. In Figuur 2(b) is te zien dat de linker occipitaalkwab (dat is de onderkant van de rechterhersenhelft in de afbeelding) het midsagittale vlak doorsnijdt. Dit effect is bij sommige mensen sterker dan bij anderen en meer voorbeelden worden getoond in Hoofdstuk 7.

Hoewel het midsagittale vlak veel wordt gebruikt om de hersenhelften met elkaar te vergelijken, is het dus niet per se een “correcte” scheiding van beide hersenhelften. Hierdoor kunnen fouten ontstaan in onderzoeken die dit soort vergelijkingen uitvoeren. Daarnaast kan de mate van kromming of asymmetrie tussen beide hersenhelften meer betekenissen hebben. Er wordt gedacht dat de mate van kromming samenhangt met links- of rechtshandigheid en andere hersenfuncties. Ook zijn er aanwijzingen dat kromming of asymmetrie samenhangt met schizofrenie.

In Hoofdstuk 7 wordt een techniek gepresenteerd die een gekromd midsagittaal vlak kan bepalen dat zorgt voor een betere scheiding tussen beide hersenhelften. Het “normale” midsagittale vlak (Figuur 2(b)) wordt in kleine stukjes vervormt zodat het beter “past” tussen beide hersenhelften. Er wordt geprobeert het gekromde midsagittale vlak zoveel mogelijk in de interhemisferische spleet te laten lopen.

In een groep van 100 patiënten (50 met een “recht” brein en 50 met een “krom” brein) zijn de resultaten van deze automatische methode vergeleken met handmatig getekende, gekromde midsagittale vlakken. De links/rechts scheiding van de hersenen wordt veel beter door gebruik te maken van het gekromde midsagittale vlak in plaats van het “normale” midsagittale vlak.

## Conclusie

In dit proefschrift zijn een aantal beeldverwerkingstechnieken voor de kwantificatie en beoordeling van brein MRI gepresenteerd. Deze technieken zijn ontwikkeld om artsen en onderzoekers te ondersteunen in hun dagelijkse werkzaamheden.

Bij het ontwikkelen van nieuwe beeldverwerkingstechnieken moet altijd de vraag “wat is het doel?” in het achterhoofd worden gehouden. Technieken moeten een duidelijke toegevoegde waarde hebben, maar deze kan zeer divers zijn: van het versnellen van bepaalde procedures tot het verhogen van de kwaliteit, betrouwbaarheid of gebruiksvriendelijkheid. Aan de hand van deze eisen kan een nieuwe techniek worden ontwikkeld.

Het onderscheid tussen handmatige, semiautomatische en automatische beeldverwerkingstechnieken is niet zomaar gekozen, maar volgt vanuit de manier waarop beeldverwerking voor brein MRI verloopt. Nieuwe ontdekkingen worden vaak met de hand gedaan, waarbij een arts of onderzoeker iets ziet op een afbeelding dat vragen opwekt. Als de interesse in een nieuwe bevinding toeneemt en er meer afbeeldingen van worden gemaakt, kunnen semiautomatische beeldverwerkingstechnieken helpen in een goede detectie en kwantificatie. Als uiteindelijk bekend wordt wat er is gevonden, dan kunnen er duidelijke definities en criteria worden opgesteld voor de detectie. Vervolgens kunnen automatische beeldverwerkingstechnieken worden ontwikkeld om grootschalig onderzoek mogelijk te maken.

Dit verloop heeft een continue impact op het aandachtsgebied van klinische en technische wetenschappers. Een goede samenwerking en terugkoppeling tussen alle betrokken wetenschappers zal er voor zorgen dat beeldverwerkingstechnieken uiteindelijk terecht zullen komen in dagelijkse klinische en wetenschappelijke praktijk.

# Dankwoord

Ik wil iedereen die op enige wijze heeft bijgedragen aan dit proefschrift hartelijk bedanken. Ook al is het mijn naam die op de voorkant van dit proefschrift staat, de inhoud ervan had niet tot stand kunnen komen zonder de talloze enthousiaste mensen die mij hierbij hebben geholpen. Dank daarvoor!

Koen, jouw aanstekelijke enthousiasme voor medische beeldverwerking bracht je al op me over tijdens mijn Master-studie, toen ik bij jou voor een kort project aanklopte. De keuze om bij het ISI, en bij jou, te gaan promoveren was daarna snel gemaakt. Je bent een geweldige bron van inspiratie, kennis en gezelligheid... ik had me geen betere copromotor kunnen wensen.

Max, heel erg bedankt dat ik bij het ISI heb kunnen promoveren met jou als promotor. Jouw waardevolle kennis, ervaring en inzicht in de wetenschap en beeldverwerking hebben mij enorm geholpen. Inspirerend vond ik je voorraad aan obscure Engelse woorden die mijn artikelen hebben verrijkt.

Geert Jan, ik voelde mij vereerd om deel uit te mogen maken van jouw VCI-groep. Jouw volharding in de vraag "wat is je doel?" dreef mij in het begin tot wanhoop, maar zorgde uiteindelijk wel voor een beter inzicht.

Geachte leden van de beoordelingscommissie: prof.dr. Van der Graaf, prof.dr. Kappelle, prof.dr. Mali, prof.dr. Breeuwer, prof.dr. Niessen. Dank voor het beoordelen van mijn proefschrift en ik kijk er naar uit om met u in discussie te gaan.

Mijn beide paranimfen, Sjoerd en Susanne. Sjoerd, vanaf dag één zijn wij, samen met Nynke, kamergenoten geweest. Om met mij dag-in-dag-uit in een kantoor te zitten, dat vraagt nogal wat... We liepen ongeveer gelijktijdig door ons promotietraject heen en daardoor had ik veel steun aan jou. Zowel op het werk als daarbuiten hadden we het ook nog eens heel gezellig met onze discussies, grapjes, verhalen en uitjes. Susanne, ik ben heel blij dat ik jou

heb leren kennen en dat we zo veel hebben kunnen samenwerken. Je bent de perfecte wetenschappelijke *partner in crime*.

Alle collega's van het ISI en met name mijn mede-promovendi uit de *OIO-steeg*. Ik vond het heel fijn om jullie als collega's te hebben, zowel op wetenschappelijk gebied, maar wellicht nog meer daarbuiten. Jullie collegialiteit en gezelligheid was erg prettig en de vele sociale activiteiten op congressen, borrels, Ardennen-weekend, Sinterklaas, Kerst, of wanneer dan ook maakten mijn promovendus-tijd tot een hele fijne periode.

Naast mijn directe collega's bij het ISI heb ik veel samengewerkt met de meisjes (en jongens) van de neuro. Het combineren van jullie kennis van de neurologie en mijn kennis van beeldverwerking was een goede zet. Het was geweldig om te werken aan de praktische toepassing van mijn technieken en deze samenwerking heeft mij veel geleerd.

Ook wil ik graag alle andere collega's, zowel binnen als buiten het UMC, die nog niet genoemd zijn erg graag bedanken. Er is te weinig plek om iedereen te kunnen noemen, maar de vele samenwerkingen en interessante discussies met iedereen vormden voor mij een belangrijk deel van mijn werk.

Buiten de wetenschappelijke wereld wil ik natuurlijk mijn vrienden, familie en kennissen bedanken voor hun steun en interesse in mijn werk. Het was fijn om voor de nodige ontspanning bij jullie terecht te kunnen.

Lieve papa en mama, wat jullie altijd voor mij hebben gedaan is niet in woorden op te schrijven. Jullie steun en vertrouwen in mijn keuzes, in studie, werk, promotie en op persoonlijk vlak, betekende heel veel voor mij. Dankzij jullie heb ik zonder zorgen de dingen kunnen doen die ik leuk vond en ik hoop dit altijd te kunnen blijven doen. Simon, ook jij maakte hier een groot deel van uit. Het is heel fijn om jou als broertje te hebben en te kunnen praten over van alles en nog wat, van hobby's en elektronica tot sport en werk.

Lieve Jessica, op een moment dat het bijna niet beter met mij kon gaan—ik was klaar met studeren en ging promoveren—gebeurde er toch nog iets mooiers: jij kwam in mijn leven. Ook als er hobbels op mijn promotie-weg kwamen stond jij voor me klaar met goede adviezen waar ik versteld van kon staan. Jouw niet-wetenschappelijke kijk op de wetenschap dwong mij om na te denken op manieren die nog nooit iemand had gedaan. Bedankt voor deze positieve invloed op mijn promotie.



# List of publications

## Papers in international journals

**H.J. Kuijf**, S.J. van Veluw, M.I. Geerlings, M.A. Viergever, G.J. Biessels, K.L. Vincken, "Automatic extraction of the midsagittal surface from brain MR images using the Kullback–Leibler measure", (submitted).

S.J. van Veluw, S.M. Heringa, **H.J. Kuijf**, H.L. Koek, P.R. Luijten and G.J. Biessels, "Cerebral Cortical Microinfarcts at 7Tesla MRI in Patients with early Alzheimer's Disease", *Journal of Alzheimer's Disease*, (in press).

**H.J. Kuijf**, S.J. van Veluw, M.A. Viergever, K.L. Vincken, G.J. Biessels, "How to assess the reliability of cerebral microbleed rating?", *Frontiers in Aging Neuroscience*, 2013, nr. 57, vol. 5, pp. 1-2.

S.J. van Veluw, L.E.M. Wisse, **H.J. Kuijf**, W.G.M. Spliet, J. Hendrikse, P.R. Luijten, M.I. Geerlings, G.J. Biessels, "Hippocampal T2 hyperintensities on 7 Tesla MRI", *NeuroImage: Clinical*, 2013, vol. 3, pp. 196-201.

T.P.C. Schlösser, K.L. Vincken, H. Attrach, **H.J. Kuijf**, M.A. Viergever, M.M.A. Janssen, R.M. Castelein, "Quantitative Analysis of the Closure Pattern of the Neurocentral Junction as related to Pre-existent Rotation in the Normal Immature Spine", *The Spine Journal*, 2013, nr. 7, vol. 13, pp. 756-763.

J.M. Biesbroek, **H.J. Kuijf**, Y. van der Graaf, K.L. Vincken, A. Postma, W.P.T.M. Mali, G.J. Biessels, M.I. Geerlings, "Association between Subcortical Vascular Lesion Location and Cognition: A Voxel-Based and Tract-Based Lesion-Symptom Mapping Study. The SMART-MR Study", *PLoS One*, 2013, nr. 4, vol. 8, p. e60541.

**H.J. Kuijf**, M. Brundel, J. de Bresser, S.J. van Veluw, S.M. Heringa, M.A. Viergever, G.J. Biessels, K.L. Vincken, "Semi-automated detection of cerebral microbleeds on 3.0 T MR images", *PLoS One*, 2013, nr. 6, vol. 8, p. e66610.

**H.J. Kuijf**, J. de Bresser, M.I. Geerlings, M. Conijn, M.A. Viergever, G.J. Biessels, K.L. Vincken, "Efficient detection of cerebral microbleeds on 7.0T MR images using the radial symmetry transform", *NeuroImage*, 2012, nr. 3, vol. 59, pp. 2266-2273.

L.E.M. Wisse, L. Gerritsen, J.J.M. Zwanenburg, **H.J. Kuijf**, P.R. Luijten, G.J. Biessels, M.I. Geerlings, "Subfields of the hippocampal formation at 7T MRI: In vivo volumetric assessment", *NeuroImage*, 2012, nr. 4, vol. 61, pp. 1043-1049.

## Papers in conference proceedings

**H.J. Kuijf**, J.M. Biesbroek, M.A. Viergever, G.J. Biessels, K.L. Vincken, "Registration of brain CT images to an MRI template for the purpose of lesion-symptom mapping", in: *Multimodal Brain Image Analysis*, Springer International Publishing, 2013, vol. 8159, Lecture Notes in Computer Science, pp. 119-128.

**H.J. Kuijf**, F. Zijlstra, S.J. van Veluw, M.A. Viergever, K.L. Vincken, G.J. Biessels, "Detecting Cortical Cerebral Microinfarcts in 7.0 T MR Images", in: *IEEE International Symposium on Biomedical Imaging*, 2013, pp. 982-985.

**H.J. Kuijf**, A. Leemans, M.A. Viergever, K.L. Vincken, "Assessment of methods to extract the mid-sagittal plane from brain MR images", in: *SPIE Medical Imaging*, 2013, vol. 8673, p. 86731K.

**H.J. Kuijf**, M. Brundel, J. de Bresser, M.A. Viergever, G.J. Biessels, M.I. Geerlings, K.L. Vincken, "Observer performance in semi-automated microbleed detection", in: *SPIE Medical Imaging*, 2013, vol. 8673, p. 867313.

**H.J. Kuijf**, M.A. Viergever, K.L. Vincken, "Automatic Extraction of the Curved Midsagittal Brain Surface on MR Images", in: *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*, Springer Berlin Heidelberg, 2013, vol. 7766, Lecture Notes in Computer Science, pp. 225-232.

**H.J. Kuijf**, J. de Bresser, G.J. Biessels, M.A. Viergever, K.L. Vincken, "Detecting cerebral microbleeds in 7.0 T MR images using the radial symmetry trans-

form", in: *IEEE International Symposium on Biomedical Imaging*, 2011, pp. 758 - 761.

S. Rit, **H.J. Kuijf**, S. van Kranen, M. van Herk, J.-J. Sonke, "Computer Assisted Analysis of Lung Tumor Regression During Radiotherapy", in: *International Conference on the Use of Computers in Radiation Therapy*, 2010.

## Abstracts

J.M. Biesbroek, **H.J. Kuijf**, N.A. Weaver, L.J. Kappelle, M. van Zandvoort, A. Postma, G.J. Biessels, "The relation between ischemic stroke location and visuoconstructive and visuospatial deficits: a voxel-based lesion-symptom mapping study", *Vas-Cog*, 2013.

S.J. van Veluw, L.E.M. Wisse, **H.J. Kuijf**, W.G.M. Spliet, J. Hendrikse, P.R. Luijten, M.I. Geerlings, G.J. Biessels, "Hippocampal T2 hyperintensities on 7 Tesla MRI", *Vas-Cog*, 2013.

L.E.M. Wisse, G.J. Biessels, S.M. Heringa, **H.J. Kuijf**, H.L. Koek, P.R. Luijten, M.I. Geerlings, "Alzheimer disease differentially affects subfield volumes of the hippocampal formation: a 7T MRI study", *Vas-Cog*, 2013.

S.J. van Veluw, S.M. Heringa, **H.J. Kuijf**, H.L. Koek, P.R. Luijten, G.J. Biessels, "Cerebral cortical microinfarcts at 7 Tesla MRI in patients with early Alzheimer's disease", *Vas-Cog*, 2013.

W.H. Bouvy, J.J.M. Zwanenburg, **H.J. Kuijf**, F. Visser, L.J. Kappelle, P.R. Luijten, G.J. Biessels, "Visualization of perivascular spaces and perforating arteries in the brain with 7 Tesla MRI", *Vas-Cog*, 2013.

M. Derieppe, B. Denis de Senneville, C. Bos, **H.J. Kuijf**, C.T.W. Moonen, "Tracking of Cell Nuclei for Assessment of In Vitro Uptake Kinetics in Ultrasound-Mediated Drug Delivery using Fibered Confocal Fluorescence Microscopy", *European Molecular Imaging Meeting*, 2013.

R. Kleinloog, J. Zwanenburg, **H. Kuijf**, E. Korkmaz, J. Post, P. Luijten, Y. Ruigrok, G. Rinkel, L. Regli, B. Verweij, "Validation of In-Vivo Imaging of Aneurysmal Wall on 7 Tesla MRI", *Society of NeuroInterventional Surgery*, 2013

**H.J. Kuijf**, M. Brundel, M.A. Viergever, G.J. Biessels, K.L. Vincken, "Semi-automatic detection of cerebral microbleeds on clinical 3.0T T2\*-weighted im-

ages using the radial symmetry transform", *International Society Magnetic Resonance in Medicine*, 2012.

**H.J. Kuijf**, H. de Leeuw, C.J.G. Bakker, K.L. Vincken, "Powerful detection of cerebral microbleeds on 7.0T MR phase gradient magnitude images using the radial symmetry transform", *International Society Magnetic Resonance in Medicine*, 2012.

L.E.M. Wisse, G.J. Biessels, S. Heringa, **H.J. Kuijf**, H. Koek, P.R. Luijten, M. Geerlings, "Alzheimer's disease-related changes in subfields of the hippocampal formation using 7T MRI", *Alzheimer's Association International Conference 2012*, 2012.

**H.J. Kuijf**, G.J. Biessels, M. Brundel, J. de Bresser, M.A. Viergever, K.L. Vincken, "Efficient detection of cerebral microbleeds", *Cerebral Amyloid Angiopathy Meeting*, 2012.

L.E.M. Wisse, L. Gerritsen, **H.J. Kuijf**, J.J.M. Zwanenburg, H.L. Koek, P.R. Luijten, G.J. Biessels, M.I. Geerlings, "Hippocampal subfields at 7T MRI: in vivo volumetric assessment", *International Conference on Alzheimer's Disease*, 2011.

L.E.M. Wisse, L. Gerritsen, **H.J. Kuijf**, J.J.M. Zwanenburg, H.L. Koek, P.R. Luijten, G.J. Biessels, M.I. Geerlings, "Hippocampal subfields at 7T MRI: in vivo volumetric assessment", *Vas-Cog*, 2011.

# About the author



Hugo Jaco Kuijf werd geboren op 28 september 1986 te Woerden en groeide op in Linschoten. In 2004 behaalde hij zijn vwo-diploma en ging Informatica studeren aan de Universiteit Utrecht. Zijn Bachelor behaalde hij in 2007, met minors in Softwaretechnologie en Game- en Mediatechnologie. Tijdens zijn daaropvolgende Master in Game and Media Technology werd zijn interesse gewekt voor medische beeldverwerking. Gedurende zijn Master volgde hij extra vakken bij het Image Sciences Institute (ISI), UMC Utrecht en voltooide zijn Master met een project bij het Nederlands Kanker Instituut. Eind 2009 begon Hugo aan zijn promotietraject bij het ISI en richtte zich hier op beeldverwerkingstechnieken, in het bijzonder toegepast op MRI afbeeldingen van het brein. Na het afronden van zijn promotieonderzoek in 2013 is Hugo verder gegaan als postdoctoraal onderzoeker bij het ISI.

Hugo Jaco Kuijf was born on 28 September 1986 in Woerden and grew up in Linschoten. He finished high school in 2004, graduating from the Kalsbeek College in Woerden, and started a Computer Science study at the Utrecht University. He completed his Computer Science Bachelor in 2007, with academic minors in Software Technology and Game- and Mediatechnology. During his Master Game and Media Technology, he developed an interest in medical imaging. He took extracurricular courses offered by the Image Sciences Institute (ISI), UMC Utrecht and finished his Masters' with an internship at the Netherlands Cancer Institute, Amsterdam. End of 2009, Hugo started as a PhD-candidate at the ISI and focussed on image processing techniques for brain MR images. After completing his PhD in 2013, Hugo continued as a postdoctoral researcher at the ISI.