Corpus Callosum as a biomarker for Alzheimer and Multiple Sclerosis

Master of Science Thesis

Georgios Sotiropoulos

Supervisors:
Prof. Dr. Ir. Boudewijn Lelieveldt
Dr. Ir. Martijn van de Giessen
Dr. Ir. Dirk Smeets

14 January 2015
Abstract

With the increase of the life expectancy and the deterioration of the quality of life that neurodegenerative diseases can cause to the patients, they have been receiving more and more attention the last few decades in the developed countries. Two of the leading neurodegenerative diseases are Alzheimer’s Disease and Multiple Sclerosis. In diagnosis and assessment progression Magnetic Resonance Imaging (MRI) is playing an important role. MRI is a non-invasive tool and can achieve a great contrast between the different tissues that exist inside the brain. In the search of biomarkers for the faster diagnosis of these diseases, the Corpus Callosum appears to be an interesting brain structure, as it facilitates inter-hemispheric communication and it is not sensitive to brain hydration and dehydration effects.

This thesis investigates if the regional and global shape and area changes of the Corpus Callosum could set the Corpus Callosum as a biomarker for Alzheimer’s Disease and Multiple Sclerosis. In order to achieve that a segmentation pipeline is proposed, so as to extract the Corpus Callosum from Magnetic Resonance brain images. After such a pipeline is constructed and validated, the quantification of the Corpus Callosum area is investigated for scans acquired in scanners of different vendors and in scanners of different magnetic field strengths, in order to gain better insights of the differences that can exist between such scans. Finally, the potential of the Corpus Callosum to be used as a biomarker is investigated, by attempting to correlate the global and regional shape and area changes of the Corpus Callosum with neurodegenerative diseases. During the thesis the segmentation algorithm showed a high segmentation accuracy performance with a mean dice of 93%, a segmentation reproducibility error of 1.93% and the classification accuracy between AD and MS patients with NC was above 95% for both groups.
I would like to thank Prof. Boudewijn Lelieveldt for giving me the opportunity on working on this project. Also, I would like to thank Martijn van de Giessen and Dirk Smeets for their useful and necessary guidance they provided me during the realization of this project. Without their suggestions and comments, this work would have been a lot more difficult. Moreover, I would like to thank Vasilis Terzopoulos and Saurabh Jain for their advices, explanations, support and constant help with all the problems I encountered.

I want also to thank all my friends and fellow master students who stood by me and made these two years in Delft and Belgium really fun. I also want to thank my friends who are in Greece, Belgium, Netherlands and UK or any other place and who with their visits provided me with calmness and relaxing days, which I really needed sometimes.

Closing, I want to deeply thank my parents, Nikos and Ritsa, my cousin Giorgos and the rest of my family for their moral and psychological support that showed me all this time and for being there for me every time I needed them, because without them nothing would have happened.
Contents

Aknowledgements i
Table of Contents ii
List of Figures vi
List of Tables ix

1 Introduction 1

2 Clinical Background 3
  2.1 Alzheimer’s disease . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
  2.2 Multiple Sclerosis . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
  2.3 Other Sources of Variation in CC measurements . . . . . . . . . . . 7
  2.4 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7

3 Previous Work on Corpus Callosum Segmentation 9
  3.1 Model-based segmentation . . . . . . . . . . . . . . . . . . . . . . 9
  3.2 Intensity-based segmentation . . . . . . . . . . . . . . . . . . . . . 11
  3.3 Atlas-based segmentation . . . . . . . . . . . . . . . . . . . . . . . 12
  3.4 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13

4 Statistical Modeling of the Corpus Callosum 14
  4.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14
  4.2 Materials & Methods . . . . . . . . . . . . . . . . . . . . . . . . . . 14
    4.2.1 Active shape models . . . . . . . . . . . . . . . . . . . . . . . 14
    4.2.2 Spatial Normalisation . . . . . . . . . . . . . . . . . . . . . . . 16
    4.2.3 Training Data . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
    4.2.4 Establishing corresponding landmarks between Corpus Callosa 17
    4.2.5 Fitting of the model in unseen images . . . . . . . . . . . . . 18
  4.3 Experiments . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
    4.3.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
    4.3.2 Spatial Normalization . . . . . . . . . . . . . . . . . . . . . . . 20
    4.3.3 Sensitivity of the model to landmark placement . . . . . . . . 21
    4.3.4 Accuracy of the model with healthy and diseased people . . . 22
    4.3.5 Segmentation reproducibility of the model . . . . . . . . . . . 23
    4.3.6 Segmentation Refinement . . . . . . . . . . . . . . . . . . . . 23
  4.4 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23
    4.4.1 Spatial Normalisation . . . . . . . . . . . . . . . . . . . . . . . 23
7 Conclusions & Suggestions for Future Work
   7.1 Conclusions .............................................. 57
   7.2 Future Work .............................................. 58

A Partitioning of the Corpus Callosum 59
List of Figures

1.1 Corpus Callosum in Sagittal Slice [copied from [30]]. 2
2.1 Corpus Callosum Compartments [copied from [2]]. 3
2.2 Corpus Callosum Index [copied from [51]]. 6
3.1 Normalized histogram of a mid sagittal slice in 3T [copied from [9]]. 12
4.1 The first three eigenvalues of the model in ±3 std. 16
4.2 (a): Mid Sagittal Plane Extraction after affine registration of the patient MR volume to the MNI space. (b): Middle Slice of the WM segmentation after it was resampled to the MNI space. (c): CC extraction by automatic masking and using largest connected component. (d): Extraction of CC contour. 17
4.3 Flowchart of the CC segmentation pipeline. 18
4.4 ASM fitting process [copied from [23]]. (a) Initial landmark positions. (b) Long search profiles sampled on contour normals. (c) Search for line piece with smallest Mahalanobis distance to training profiles. (d) Move landmark points to center of line piece with lowest Mahalanobis distance. (e) Map the x,y points to model coordinates, constrain them to be within 3 standard deviations from the training shapes, and transform them back to x,y positions. 19
4.5 (a): 16 Landmarks on the CC. (b): 32 Landmarks on the CC. (c): 44 Landmarks on the CC. (d): 64 Landmarks on the CC. 22
4.6 Failing image in rigid test set 24
4.7 (a) Cosine similarity of the b values for the affine test set on the OASIS NC Test Re-test data (0.934 / 0.971 / 0.109). (b) Cosine similarity of the b values for the rigid test set on the OASIS NC Test Re-test data (0.94 / 0.971 / 0.088). 24
4.8 (a) Normalised b values by the corresponding eigenvalues for the affine test set. (b) Normalised b values by the corresponding eigenvalues for the rigid test set. 25
4.9 (a) Mean shape of the 32-landmarks-set (blue) and 44-landmarks-set(red). (b) Modes of variation of the 32-landmarks-set (blue) and 44-landmarks-set(red). 25
4.10 Dice coefficient for different number of landmarks (16: 84.47/85.03/4.72, 32: 84.95/85.55/4.26, 44: 84.64/84.98/4.48, 64: 84.89/85.1/4.06) . . . 26
4.12 Percentage Area Change between the same patient in close time points (1.92 / 1.56 / 1.61).  

4.13 (a) Minimum thresholding in the segmented intensities OASIS NC Cross-Sectional data (Original: 91.72/92.43/2.72, Improved: 94.36/94.99/2.22).  
(b) Minimum thresholding in the segmented intensities OASIS NC Test Re-test data (Original: 91.87/92.26/2.08, Improved: 93.88/94.64/2.09).  

4.14 (a) Originally segmented image, (b) Minimum thresholded image .  


5.1 (a) Absolute PAC between scans of 3T and 1.5T (Nifty Reg: 6.87 / 6.1 / 4.84, Elastix: 6.92 / 5.99 / 4.96). (b) PAC between scans of 3T and 1.5T (Nifty Reg: 5.72 / 5.43 / 6.16, Elastix: 6.92 / 5.99 / 4.96).  

5.2 (a) Absolute difference of the normalised $b$ values with the use of an affine registration for the spatial normalisation. (b) Absolute difference of the normalised $b$ values with the use of a similarity registration for the spatial normalisation.  

5.3 (a) Histogram of the Corpus Callosum Intensities of 1.5T and 3T. (b) Histogram of the CSF Intensities of 1.5T and 3T in the mid sagittal plane.  

5.4 (a) PAC within each vendor with the use of an affine registration for the spatial normalisation (Philips: 2.46 / 2.43 / 1.57, Siemens: 1.91 / 2.35 / 1.07, GE: 1.3 / 1.55 / 1.0). (b) PAC within each vendor with the use of a similarity registration for the spatial normalisation (Philips: 1.73 / 1.14 / 1.7, Siemens: 2.98 / 2.81 / 1.54, GE: 1.69 / 1.54 / 0.74).  

5.5 (a) PAC between vendors with the use of an affine registration for the spatial normalisation (Philips-Siemens: 5.14 / 4.94 / 2.22, Siemens-GE: 6.71 / 7.76 / 3.7, Philips-GE: 2.79 / 2.74 / 1.5). (b) Absolute difference of the normalised $b$ values with the use of a similarity registration for the spatial normalisation (Philips-Siemens: 8.13 / 6.84 / 2.56, Siemens-GE: 7.52 / 8.54 / 3.22, Philips-GE: 2.57 / 2.74 / 1.12).  

5.6 (a) Absolute difference of the normalised $b$ values with the use of an affine registration for the spatial normalisation. (b) PAC between vendors with the use of a similarity registration for the spatial normalisation.  

5.7 (a) Intensity Profiles of the CC in the three vendors. (b) Intensity Profiles of the CSF in the mid sagittal plane of the three vendors.  

6.1 (a) Population graph of the area of the CC using an affine registration. (b) Population graph of the circularity of the CC using an affine registration.  

6.2 (a) Population graph of the area of the CC using a similarity registration. (b) Population graph of the circularity of the CC using a similarity registration.  

6.3 (a) ROC curve with respect to the LDA dimension for NC, AD patients with the use of an affine registration. (b) ROC curve with respect to the LDA dimension for NC, AD patients with the use of a similarity registration.
6.4 (a) ROC curve with respect to the LDA dimension for NC, MS patients with the use of an affine registration. (b) ROC curve with respect to the LDA dimension for NC, MS patients with the use of a similarity registration.

6.5 (a) Histogram of the total area of the CC in NC and AD scans with the use of an affine registration. (b) Histogram of the total circularity of the CC in NC and AD scans with the use of an affine transformation.

6.6 (a) Histogram of the area of the genu in NC and AD scans with the use of a similarity registration. (b) Histogram of the area of the isthmus in NC and AD scans with the use of a similarity registration.

6.7 (a) Histogram of the area of the genu in NC and AD scans with the use of an affine registration. (b) Histogram of the area of the isthmus in NC and AD scans with the use of an affine registration.

6.8 (a) Histogram of the area of the genu in NC and AD scans with the use of a similarity registration. (b) Histogram of the area of the isthmus in NC and AD scans with the use of similarity registration.

6.9 (a) Histogram of the second $b$ value of the segmented scans of NC and AD scans with the use of an affine registration. (b) Histogram of the second $b$ value of the segmented scans of NC and AD scans with the use of a similarity registration.

6.10 (a) Histogram of the total area of the CC in NC and MS scans with the use of an affine registration. (b) Histogram of the area of the anterior body in NC and MS scans with the use of a similarity registration.

6.11 (a) Histogram of the area of the CC in NC and MS scans with the use of a similarity registration. (b) Histogram of the total area of the CC in NC and MS scans with the use of an affine registration.

6.12 (a) Histogram of the area of the anterior body in NC and MS scans with the use of an affine registration. (b) Histogram of the area of the mid body in NC and MS scans with the use of an affine registration.

6.13 (a) Histogram of the area of the anterior body in NC and MS scans with the use of a similarity registration. (b) Histogram of the area of the mid body in NC and MS scans with the use of a similarity registration.

6.14 (a) Histogram of the first $b$ value of the segmented scans of NC and MS scans with the use of affine registration. (b) Histogram of the third $b$ value of the segmented scans of NC and MS scans with the use of an affine registration.

6.15 (a) Histogram of the second $b$ value of the segmented scans of NC and MS scans with the use of an affine registration. (b) Histogram of the second $b$ value of the segmented scans of NC and MS scans with the use of a similarity registration.

6.16 Modes of variation of the ASM model of the CC.

A.1 Wittelson Partitioning [copied from [50]].
A.2 Radial Partitioning [copied from [38]].
A.3 Medial axis partitioning [copied from [38]].
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Dicom Header protocol values for the different vendors.</td>
<td>31</td>
</tr>
<tr>
<td>5.2</td>
<td>Dicom Header protocol values for the Philips scanner of different magnetic field strength.</td>
<td>32</td>
</tr>
<tr>
<td>6.1</td>
<td>Accuracy, Sensitivity and Specificity of SVM for the separation of NC-MS with the whole feature set</td>
<td>46</td>
</tr>
<tr>
<td>6.2</td>
<td>Accuracy, Sensitivity and Specificity of SVM for the separation of NC-AD in different feature subsets with the use of an affine registration</td>
<td>46</td>
</tr>
<tr>
<td>6.3</td>
<td>Accuracy, Sensitivity and Specificity of SVM for the separation of NC-AD in different feature subsets with the use of a similarity registration</td>
<td>46</td>
</tr>
<tr>
<td>6.4</td>
<td>Accuracy, Sensitivity and Specificity of SVM for the separation of NC-MS with the whole feature set</td>
<td>47</td>
</tr>
<tr>
<td>6.5</td>
<td>Accuracy, Sensitivity and Specificity of SVM for the separation of NC-MS in different feature subsets with the use of an affine registration</td>
<td>48</td>
</tr>
<tr>
<td>6.6</td>
<td>Accuracy, Sensitivity and Specificity of SVM for the separation of NC-MS in different feature subsets with the use of a similarity registration</td>
<td>48</td>
</tr>
<tr>
<td>6.7</td>
<td>Mean and Standard deviation of the CC and its compartments for NC and AD with the use of an affine registration</td>
<td>49</td>
</tr>
<tr>
<td>6.8</td>
<td>Mean and Standard deviation of the CC and its compartments for NC and AD with the use of a similarity registration</td>
<td>49</td>
</tr>
<tr>
<td>6.9</td>
<td>Mean and Standard deviation of the CC and its compartments for NC and MS with the use of an affine registration</td>
<td>52</td>
</tr>
<tr>
<td>6.10</td>
<td>Mean and Standard deviation of the CC and its compartments for NC and MS with the use of a similarity registration</td>
<td>52</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

In the last few years, neurodegenerative diseases have been receiving more and more attention in the developed countries. Each year the amount of people that are affected by them is increasing and in some cases a lot of time is elapsed until the disease is diagnosed. Neurodegeneration is a process that affects either the function or the structure of the neurons, which at the end are not able to operate at the same extent as before, causing degradation of the activities that the patient can perform. In the detection of these diseases magnetic resonance imaging (MRI) plays an important role. This is due to the fact that it is a non-invasive and non-harmful technique, and can provide great contrast between the different soft tissues that exist in the brain, enabling the assessment of different structures and abnormalities in the brain. Two of the leading neurodegenerative diseases are Alzheimer’s Disease (AD) and Multiple Sclerosis (MS), which although they have low incidence rates, they can decrease dramatically the quality of life of the patients [49].

AD mostly affects the elderly people that are aged above 55 and as they are getting older the probability of getting affected by AD grows as well [17]. In addition, due to the growth of the life expectancy more people might get affected by AD. Recent study predicts an increase in the number of AD patients from 4.5 millions in 2000 to 13.2 millions in 2050 in the United States alone [17]. AD progressively affects the function of neurons and slowly degenerates the cognitive processes, such as the memory and thinking skills, and at the end a person loses the ability to perform the simplest tasks [32]. The most common symptoms of AD are memory loss, linguistic and behavioural (confusion, irritability, aggression) problems.

On the other hand, MS is a disease that mainly affects young adults between the age of 20 and 50 years old. MS is believed to be an autoimmune disease of the central nerve system, which affects the myelin of the neurons leading to axonal loss [5, 31]. The breakage of the myelin of the neurons is called lesion. The number of people that are affected throughout the world are between 2 and 2.5 millions, and the costs for the patients in UK alone are estimated at about 1-2 billions per year [5, 27, 31]. Furthermore, MS, in opposition to AD, can affect all the functions of the diseased person e.g. talking, movement, sight, etc.

With the rise of neurodegenerative diseases, a trend has been created towards finding biomarkers that can assist medical experts in assessing neurodegeneration more accurately, as well as quantifying its results in the everyday clinical practice. In this thesis the utilisation of the Corpus Callosum (CC) as a candidate biomarker for diagnosing Alzheimer’s Disease and Multiple Sclerosis is being examined.
The Corpus Callosum can be considered as a potential biomarker as it is the largest white matter structure in the brain, it consists of the majority of fibers that connect the two cortical hemispheres and is responsible for the interconnection and the transfer of cognitive, motor and sensory information between these hemispheres [18, 47]. Furthermore, it is located in the center of the mid-sagittal slice in a T1 weighted MR image as it is shown in the Figure 1.1, where its boundaries are indicated with red. Various studies have attempted to link the size, thickness and shape changes of the corpus callosum with these neurodegenerative diseases [1, 4, 11, 12, 13, 20, 24, 34, 35, 36, 40, 51, 52].

Figure 1.1: Corpus Callosum in Sagittal Slice [copied from [30]].

In order to examine the CC quantitatively, a 2D segmentation approach is proposed, where image segmentation refers to the process of subdividing an image into regions of interest [15]. A 3D segmentation was not followed, due to the fact that the CC protrudes into the two hemispheres, without having clear boundaries. After the construction of a segmentation pipeline, the shape and area changes of the CC were investigated if they could be correlated with the diagnosis of Alzheimer’s Disease and Multiple Sclerosis.

In Chapter 2 the clinical background behind this study is presented along with the research that has already been done, in using the corpus callosum as a biomarker for neurodegenerative diseases. In Chapter 3 the related work that has been performed for the segmentation of the corpus callosum from MR images is presented. Chapter 4 describes the construction and validation of the model that was used to segment the corpus callosum from the MR images. In Chapter 5 the variability on measurements of the corpus callosum that can be found in acquisitions performed in MR scanners of different vendors and of different magnetic field strengths was investigated. Chapter 6 investigated the shape and area changes of the corpus callosum that can be observed in healthy and diseased populations and whether these changes can set the corpus callosum as a biomarker. Finally, our conclusions are summarised in Chapter 7 along with discussion on our work and suggestions for future improvements.
Chapter 2

Clinical Background

Neurodegeneration is a process of a range of conditions that happen to the neurons in the human brain during neurodegenerative diseases (e.g. Alzheimer’s Disease, Multiple Sclerosis, Huntington’s disease etc.). A common characteristic symptom of neurodegeneration in which the clinicians are interested in discovering and quantifying is the atrophy of the human brain. Unfortunately, the human brain is prone to hydration and dehydration effects [44], which can produce deceiving quantification results, when e.g. measuring the volume of the tissues in the brain. Consequently, in the last few years clinicians and researchers have turned their interest towards a more stable structure, the Corpus Callosum. CC also known as collosal commisure, is a banana shaped structure which is located in the center of the brain (Figure 2.1).

![Figure 2.1: Corpus Callosum Compartments [copied from [2]].](image)

The CC is the largest white matter structure that exists in the brain, composed by dense fiber bundles facilitating interhemishperic communication. The fibers that pass though the CC consist of different sizes, generating five main compartments (Figure 2.1). Several studies have been recently performed, trying to correlate the callosal area and shape changes with the progression of AD and MS, which are described in the Sections 2.1 and 2.2. Moreover, some of them try also to conclude if callosal area and callosal morphology can be correlated with sex, handedness and brain volume, as these can be other forms of variation which can affect the callosal measurements.
2.1 Alzheimer’s disease

Recent studies have explored the correlation of the CC area with the progression of AD and most of them have indicated a reduced area primarily in the genu and splenium of the CC [1, 4, 12, 13]. The recent studies that have been performed [4, 12, 13], support that the deformation of the CC in these areas is mainly caused due to the different myelination process in these compartments. In the genu exist small diameter fibers, which myelinate much later in normal development, while in splenium there are large diameter fibers, which myelinate early in development. These two ways of myelination produce different degeneration processes and consequently it is believed that splenium fibers suffer from Wallerian degeneration, while genu fibers degenerate from myelin breakdown.

Bachman et al. [4] investigated the correlation of the CC mid-sagittal cross-sectional area (CCA) and the CC circularity (CIR) in a group of 147 subjects, 75 normal controls (NC), 51 patients with very mild AD (AD-VM) and 21 patients with mild AD (AD-M), taken from the Open Access Series of Imaging Studies (OASIS) database [28]. The CC was partitioned according to the radial partitioning, splitting the Corpus Callosum into five regions (genu, anterior body, mid body, posterior body, splenium). The study concluded that the shrinkage of the genu over time is statistically significant in all three groups and also the rate of atrophy of the splenium is progressive as the disease progresses. Moreover, they found that the rate of change of CIR was significant in all three groups, and consequently can serve as a sensitive metric of brain integrity and disease progression, as it is sensitive to structural brain changes associated with disease progression.

Di Paola et al. [12, 13] focused on correlating the changes of the CCA between NC patients (20), AD-M (20) and severe AD (10) patients by using voxel-based morphometry (VBM). The comparison between severe AD patients to normal controls yielded a significant reduction in anterior and posterior sections, which led to the conclusion that callosal atrophy is present predominantly in the latest stage of AD. Furthermore, when mild AD and severe AD patients were compared, significant differences in the genu, anterior body and splenium were found. Their findings indicated that callosal atrophy is both detectable and fairly generalized in severe AD, while callosal reductions in milder and pre-clinical stages of AD appear to be less pronounced.

Ardekani et al. [1] also explored the correlation of CCA and CIR between 98 NC, 70 AD-VM and 28 AD-M also from OASIS database. The findings of this research were similar to Bachman’s et al. [4] and they confirmed that CIR can be used as a more sensitive marker for monitoring AD progression than CCA. Moreover, the authors found that CCA decreases with age and dementia severity and increases with intracranial capacity confirming the findings of Luders et al. [26]. In addition, they also found that CIR is correlated with age, intracranial capacity and dementia and it has a decreasing behavior with all of them. Finally, the authors found that for subjects above 60 years the shape of the CC becomes less circular with age and also that CC in larger brains tends to be less circular as well. The decrease in circularity can be explained as a decrease in the area, an increase in the perimeter or a combination of both factors.
2.2 Multiple Sclerosis

In MS the clinical score that is most commonly used by clinicians to assess the progression of the disease is the Expanded Disability Status Scale (EDSS). EDSS is a score that is used to quantify the disability of a patient with MS, based on the functionality of its functional systems. Several studies have attempted to correlate the changes of the CC with the EDSS score, but also with the cognitive clinical score of Functional Systems Scale (FSS) or other cognitive and non-cognitive tests (e.g., MS functional composite (MSFC) score, paced serial auditory addition test (PASAT), and 9-hole peg test components). The studies that are discussed below, investigated the correlation of these scores with the shape and area changes of the CC ([20, 34, 35, 36, 40, 51, 52]), after segmenting the CC with a 2D scheme apart from the research of Prinster et al. [36], where the CC was segmented in 3D.

Sigirli et al. [40] investigated the shape differences in the CC and cerebellum, of female relapsing-remitting MS (RRMS) (26) and secondary progressive MS (SPMS) (14) patients compared to normal controls (15). The CC was statistically modelled by manual annotation of 16 landmarks on the CC and the CC deformation was demonstrated with a thin-plate spline. The CC shape in both patient groups was found significantly different from the shape of controls with maximum deformation in the anterior region of the CC and with more pronounced deformations in the SPMS group.

Pelletier et al. [35] conducted a 5-year longitudinal study of 30 RRMS patients with mild disability based on EDSS scores, and compared them with normal controls (53) in terms of callosal atrophy and interhemispheric dysfunction. The authors observed that during the 5 years the total CC atrophy significantly increased and the patients presented impairment in the auditory, motor and sensory interhemispheric transfer. The authors supported that CC atrophy and functional transfer (FT) impairment follow a parallel and progressive evolution during the course of the disease and they found a correlation of these two variables with EDSS score, both in the beginning and at the end of the study. Moreover, a relation was also found between CC atrophy and FT impairment with T2 lesion load, which supports the hypothesis that demyelinating lesions of callosal and pericallosal regions induce CC atrophy and interhemispheric impairment in patients with MS. Finally, the authors found a significant correlation between CC morphology and EDSS score.

Ishaq et al. [20] conducted also a longitudinal study with 51 MS patients (421 images scanned in 300 days) analyzing the global and regional shape of the CC. For more accurate results the authors used a medial profile representation, as it can follow the geometry of the CC in terms of four shape measures (bending, stretch, inner and outer thickness), which were measured globally and regionally. Additionally, in order to apply regional statistics the Wittelson partitioning was used to split up the CC. The authors found that the CC shape tends to straighten out and lose some of its curvature with disease progression. Moreover, the stretch deformation has a more dominant contribution in the anterior three CC subdivisions compared to the four posterior subdivisions. Furthermore, it was observed that the effect of stretch tends to decrease towards the latter half of the CC, while the effects of bending, inner and outer thickness tend to have the opposite behaviour. Finally, it was concluded that stretch and bending effects contribute more to the variation of the CC shape than inner and outer thickness.

In the study of Ozturk et al. [34], the authors attempted to correlate the callosal damage with EDSS and MSFC scores. 69 MS subjects and 29 healthy volunteers participated in the study where T1, Flair and DTI images were used. The authors reported a weak correlation (r=0.21-0.31) of the DTI parameters (Fractional anisotropy
and perpendicular Diffusivity) in different compartments of the CC with PASAT-3 and 9-hole peg-test time, but they didn’t find any correlation with the EDSS score. They argued that such a correlation, of the callosal MRI indices to the EDSS score, was not found as cognitive and non-cognitive impairment do not always develop in parallel during MS and the EDDS is heavily weighted towards motor disability.

Prinster et al. [36] applied a voxel-based morphometry approach on 128 patients and tried to assess the correlation of regional loss of grey and white matter with indices of clinical and radiological severity in RRMS patients, including EDSS and Lesion Load (LL). For more reliable results, the authors included the total intracranial volume, age and sex in the model, so as to account for any possible effects of these variables. From the VBM analysis the authors showed WM loss on the splenium of the CC. Moreover, regions of WM loss, including the anterior part of the CC, were correlated with EDSS, while genu and splenium were also correlated with LL. EDSS score was correlated with motor areas bilaterally, when testing for correlations of GM with disease severity. Loss of WM in the motor brain areas was significantly correlated with EDSS, supporting the hypothesis that motor neurons are more severely damaged as disability progresses.

Yaldizli et al. [51, 52] used the CC index (CCI) as an indicator of MS progression and tried to correlate it initially with long-term disability and afterwards with total and regional atrophy. The CCI was obtained from the mid sagittal plane of a T1 weighted image by drawing a line at the greatest anterior posterior diameter of the CC and a perpendicular line at its midpoint and was computed as \((a'a' + bb' + cc')/ab'\), where the points can be seen in Figure 2.2.

![Figure 2.2: Corpus Callosum Index](copied from [51]).

In their first study [51] the authors analysed the longitudinal data of 169 patients, most of which were treated with disease modifying drugs. In this study the authors observed that higher CCI values corresponded to a higher CC volume and CCI was correlated with cognitive impairment in RRMS and SPMS patients. The main findings of the study were that the brain atrophy as measured by CCI was associated with disability progression and that EDSS at diagnosis was found to be the best predictor for future disability. Moreover, the authors observed that the CCI decreased more in SPMS than RRMS patients and also in patients not treated by disease modifying drugs than in patients that were treated with them.

In the second study of Yazildi et al. [52] the cross sectional data of 113 patients were processed, from which 83 were treated with disease modifying therapy. The authors correlated the CCI with T2, T1 lesion volume and whole brain volume, but
no correlation was found with EDSS score or disease duration. Moreover, the atrophy of the posterior CC segment was significantly associated with cognitive impairment, while the anterior CC segment was significantly associated with fatigue severity. The authors supported that EDSS score depends entirely on the ability of the patient to ambulate and the low mean disease duration and EDSS score of their patients might be a bias towards their analysis.

Besides structural MRI, Diffusion Tensor Imaging (DTI) is used increasingly in MS studies [11, 37, 53], where DTI parameters are correlated with clinical scores. Unfortunately, DTI images are more difficult to acquire for MS patients, as the time needed in an MR scan is much more than that of a structural MR. This additional time can produce severe artifacts in the DTI images, and this is the reason that DTI studies were excluded from our search.

2.3 Other Sources of Variation in CC measurements

In the work of Luders et al. [26], the authors investigated if the sex differences in brain size or the biological sex, play an important role in the callosal morphology. They used a dataset of 96 subjects, which was consisted of equal number of men and women, and was further split into two groups of subjects with similar brain volumes and into subjects with extremely different brain volumes. The study concluded that brain volume is the main cause of differences in callosal morphology and especially in callosal thickness between the two sexes and not the gender. In this way the authors verified the findings of a former study, which argued that “the differences in callosal size and shape between men and women result from size variation, not from sex-related characters” [25]. For this reason, during the master thesis the total brain volume will be taken into account in the measurements as a mean of normalisation of the CC area. Moreover, the age can be an important factor of the disease progression, as with age the chemical changes in the human brain are delayed and maybe the neurodegeneration is delayed as well.

2.4 Conclusions

Considering the studies presented above, in AD the results are more consistent as the atrophy in genu and splenium is mostly correlated with the disease progression, while in MS the uncertainty of the results is bigger, as some of the studies succeeded in correlating the CC atrophy with the EDSS score, while others failed. This can be explained from the fact that cognitive and non-cognitive impairment do not develop in parallel in the MS patients, and the patient groups that were used for the studies might have not been representative for correlating the disease progression with the EDSS score [34]. Moreover, external values like age, sex and intracranial volume, which can have an effect on the results, might not have been taken into account.

The focus of this thesis is to reduce the uncertainty in the MS studies and to furthermore try to correlate global and regional shape and area changes of the CC with MS. Towards this direction, the verification of the results of the AD studies can be used as a sanity check. Overall, the AD patients are expected to present area changes in the anterior and posterior regions of the CC (genu and splenium), which are connecting the prefrontal cortex and the parietal and occipital lobe respectively. Moreover, the body is expected to be spared, as it is responsible for the motor cortex, which is not
severely affected during AD. On the other hand, the CC of the MS patients is expected to be highly atrophied in the part of the body, as the MS patients are frequently facing disability problems.
Chapter 3

Previous Work on Corpus Callosum Segmentation

Image segmentation is the process in which the image is separated into regions of interest, based on prior information about the region to be extracted. In this case the segmentation procedure can be roughly separated into three main categories, model-based segmentation, intensity-based segmentation and atlas-based segmentation. All these segmentation methods have been used in order to segment the CC and are listed below, with the majority of them using model-based segmentation schemes. Due to the fact that the manual delineation of the CC is time consuming and can be prone to inter-subject variability, publications of non full automatic techniques were excluded from the search.

3.1 Model-based segmentation

In the model-based segmentation schemes a high precision model, which captures the variation of the boundaries of the structure to be segmented from a training set, is fitted into the region of interest (ROI) with the use of an optimization method, often by the minimization of an energy function [45]. Some examples of model-based segmentation schemes are deformable templates, level sets, active shape models (ASM) and active appearance models (AAM), which were firstly introduced by Cootes et al. [6, 7, 8].

Szekely et al. performed the first work on automatic CC segmentation, in which they used snakes with the contours of the CC parameterised by Fourier descriptors [45]. An advantage of using Fourier descriptors is that the parameter shape that is defined by the first-degree terms of the Fourier series can be invariant under similarity transformations (translating, rotating and scaling). The authors performed a principal component analysis (PCA) on the covariance matrix of the Fourier coefficients, so as to acquire the major modes of deformation from the training set. Afterwards, a subset of these modes that could explain the variance of the CC was chosen. For the fitting the authors used a modified version of the Fourier snakes, so that they could allow deformations only based on the selected eigenmodes. Finally, the whole procedure was also expanded to 3D as well. Although the authors tried to overcome the main drawback of snakes, which is the stuck to local minima, by using the Hough Transform to achieve a better initialisation near the final solution, at the end the results were unsatisfactory. The authors applied their method to a dataset of 30 mid-sagittal slices, but unfortunately they
did not validate their method quantitatively, but only qualitatively.

Vachet et al. [47] build on the work of Szekely et al. by adding several preprocessing steps such as the anterior-posterior commissure (AC-PC) alignment of all the subjects, by rigidly registering them to the Montreal Neurological Institute (MNI) space, and having the tissue segmentation of the different brain tissues (WM, GM) as prior information to perform intensity inhomogeneity corrections. This AC-PC alignment was performed to capture more accurate the modes of variation from the PCA and also to derive the mid-sagittal plane of the image, which is important in the CC segmentation especially when calculations of its area, size and orientation are performed. Moreover, in order to enhance the segmentation procedure the authors successfully detached the fornix from the CC, which is a structure attached to the CC with similar intensities, by averaging the intensities of four adjacent slices to the mid sagittal slice (two in each side) before applying the segmentation algorithm. Additionally, a multi-scale search approach was used to improve the robustness of the algorithm. As a final step to improve the segmentation, the authors used a semi-automatic segmentation my manually placing repulsive points after the first optimisation iteration has been finished. This step helped to overcome the major problems of snakes with incorrect initialisation, the presence of noise and the existence of other areas close to the CC that have similar intensities. In this study the authors used a training set of more than 150 subjects, which included balanced groups of adult controls, schizophrenics, pediatric controls and autistics and applied their method to 366 subjects, between 1 and 2 years old, of which 86 (23.5%) cases needed additional user interaction to correct the segmentations. Unfortunately, also this work didn’t include any accuracy results and its validation was qualitatively.

Stegmann et al. [41, 42, 43] proposed several improvements for the accurate automatic segmentation of the corpus callosum. The authors applied AAMs for the CC segmentation, as they statistically model the shape of the CC and its appearance. For the training set 17 cross-sectional MR images were used, in which experts delineated the CC contour as a closed landmark free curve. Due to the fact that there are not any widely accepted landmarks on the CC, the authors defined only two landmarks on the delineated contours, one in the rostrum and one in the splenium (left and right tip of the CC). The rest of the points were generated automatically with the use of a minimum description length (MDL) parameterization [10], in order to achieve a strong point correspondence with the least number of points. Moreover, due to the fact that AAMs model only appearance of the object, the authors used additional normals outwards from the shape in the contour points, called whiskers, in order to achieve a better estimation of the background, and also their length was incorporated in the calculation of the model [41, 42]. Furthermore, before applying the PCA to capture the correlation between the shape and the intensities they performed a Procrustes analysis to improve the results of the PCA. Additionally, in order to improve both the convergence and robustness of the AAM, the authors fitted their model with a multi-resolution approach. Finally, an additional novel implementation was the incorporation of a Broyden-Fletcher-Goldfarb-Shanno (BFGS) optimizer to improve the fitting of the AAM model [41]. The validation of the methods was performed by using two measures, the average distance between model and ground truth landmarks (pt.pt), and model and ground truth contour (pt.crv). For these measures a leave-one-out validation scheme of the 17 cross-sectional images yielded an average pt.pt error of 0.28 pixels, an average pt.crv error of 0.18 pixels and a mean shape overlap of 0.93 %.

McIntosh et al. [29] proposed a different scenario for the CC segmentation with the use of deformable shape models and genetic algorithms (GAs) for the minimisa-
tion of the energy function. The advantages that lead the authors to use deformable models are their inherent smoothness properties and the fact that they can fit to missing boundaries. The authors investigated the main trade-off of the deformable model based segmentation, which is between the shape model fidelity and its optimisation. In order to mitigate this trade-off, they introduced the use of a GA with localised shape training. Moreover, instead of using a deformable template that was capturing the boundaries, the authors used a medial based shape representation to capture the inner part of the CC, as the medial axis based shape statistics better respect anatomical variability than global statistics, but their inherent non-convexity makes them challenging to optimise. The medial template was split into nodes and from each node certain values were captured, such as the length from each node to the former one, the distance of the node to the boundaries, its base angle and its location. A Hierarchical Principal Component Analysis (HRPCA) processed all these information in order to capture the local variability and not the global, which it is mostly used with deformable templates. The main novelties of this paper were that the classical gradient descent approach was replaced with GA and the convex, implicit, global shape statistics were replaced with non-convex, explicit, localised ones with the use of the medial based representation. The method was applied to 50 subjects, but the quantitative results were though not very good, as the jaccard distance that was reported had a mean of 0.176. The authors concluded that instead of improving the segmentation results, they should look for another energy function, as the variation of error of the energy function was much higher than the variation of error that the random GAs were creating.

3.2 Intensity-based segmentation

Dagderiven et al. [9] were focused on intensity-based methods for the segmentation of the CC. The authors proposed two novel segmentation techniques and a combination of them, by keeping the best features of the two methods. The first method was a valley matching method which utilised the histogram of the image with some preprocessing steps, such as intensity normalisation to [0,255] to standardize the histogram and histogram smoothing to mitigate the local minima. The purpose of the valley matching was to locate an area around point \( w \) (Figure 3.1), in which the intensities of the corpus callosum are located. A gradient descent approach was used to locate the local minimum of the first valley in the histogram, and from that point, point \( w \) was extracted. The authors used an interval of 5 and 30 intensities left and right of point \( w \) respectively, to search for the CC and this interval was chosen based on the intensities of the CC in their datasets. For every search in this interval a median filter was applied to remove smaller objects and a connected component analysis was used to extract the CC. Finally, the CC was extracted only if the object that was found was following certain rules (area bigger than \( 5cm^2 \), orientation between \(-13^\circ, 13^\circ\), etc.).

The second method was called 'Evolutionary CC detection' and focused on localising the CC by using a GA with a fitness function that was using anatomical ratios, instead of fixed prior knowledge. The anatomical ratios were computed from a subset of 30 subjects out of the 252 (2 datasets of 67 and 185 scans) that were used for the study. The last method, which was the combination of the two other methods and yielded the best results, called 'Evolutionary valley matching', used four genetic algorithms, which also replaced the preprocessing steps. The first three GAs were used to perform the valley matching implementation (locate the peaks, the valley between the peaks and the interval of the values in which the CC should be located), and the last one
performed the evolutionary CC detection. The subjects that were used for this study were in total 252 and the algorithm was able to locate the CC in 95.5% of the data (for Valley matching for the first dataset), but no quantitative results on the accuracy of the segmentation were published. The 'Evolutionary valley matching' performed in average the best in the two study groups (91% (61 out of 67 scans) and 81% (150 out of 185 scans) respectively).

Icer et al. [19] used a complete different method to approach the problem of the segmentation of the CC, which involved the use of Gaussian Mixture Modeling (GMM) with the use of an Expectation Maximization Algorithm (EM) to define the optimal parameters of the GMM, and a Fuzzy-C-means (FCM) method. The main idea behind them is that the different tissues inside the brain (White Matter (WM), Gray Matter (GM), Cerebrospinal Fluid (CSF)) represent different classes and each voxel is attempted to be segmented with the least possible error. With the GMM every pixel of the image is assigned to a class according to a Gaussian distribution, while in the FCM each pixel is assigned according to a fuzzy membership function. Both methods were used iteratively to achieve the best results. The only preprocessing step that was used for these methods was a histogram equalisation in order to reveal better differences in the gray scale image. The training set that was used included 28 subjects (12 males and 16 females) and the results that were published had above 92% segmentation accuracy for both methods. The GMM implementation was more time consuming, but it revealed higher segmentation accuracy than the FCM implementation, as the FCM tended to oversegment the CC.

A major problem of the intensity-based methods is that they only consider the intensities of the image, without taking most of the time into account any information about the shape of the object to be segmented. Due to the fact that a high variability in the intensity ranges between the vendors and between the acquisition protocols used in different sites, the intensity-based segmentation schemes can have a serious disadvantage for the CC segmentation.

### 3.3 Atlas-based segmentation

Seixas et al. [39] used an atlas-based segmentation (ABS) scheme for the extraction of the CC. The first step of their approach included the spatial normalisation of the data as a preprocessing step, so that all datasets were in the same stereotaxic space. This was achieved by affinely registering the image to a T1 template. Afterwards, the data were smoothed such that each voxel would represent the average of itself and its neighbours.
and then split into WM, GM, CSF. The CC is a WM structure and a bounding box was placed in the middle slice of the WM segmentation to extract the CC and due to the fact that the spatial information of the image is kept after registration the CC was labeled back in the initial image. This method was applied in 20 subjects of OASIS database (10 men, 10 women) and the results that were published were not referring to the accuracy of the segmentation, but to the area of the CC that was computed and the ratio of this area to the whole brain volume. By evaluating the images that were published a promising thing is that in only two out of the twenty images the fornix could not be detached from the CC, but this study also did not define the segmentation accuracy. The ABS segmentation can be a promising technique for the segmentation of the CC, since it takes into account the intensities and the shapes of the several structures in the brain. Due to the fact that the CC can be highly variable especially in the diseased brains, the ABS approach may not be able to achieve highly segmentation accuracy for a pathological CC and capture fine details.

3.4 Conclusions

The aforementioned methods show promising quantitative results, but the main problem is that in most of them no accuracy results are provided and even if they had they were tested in small datasets. Consequently, the performance of these algorithms is not really well validated and known. Moreover, due to the high variability of the CC, especially in diseased patients, a method which takes into account as prior knowledge the shape of the CC together with its underlying intensities has to be used, so as to be able to yield high accuracy results. In addition, the chosen method should also be invariant to acquisition protocols (different parameters in T1 acquisition between sites) and vendors, so as it will be able to be used for multivendor studies giving high reproducibility results. For the above reasons, the model - based segmentation techniques were chosen to be used in this thesis, as they fulfil all the specifications.
Chapter 4

Statistical Modeling of the Corpus Callosum

4.1 Introduction

The purpose of this project is to try to correlate global and regional shape and area changes of the CC with the progression of Alzheimer Disease and Multiple Sclerosis. The shape of the CC is highly variable even in healthy people and since we are interested in capturing the shape changes of the CC, a technique that models the shape variation of the CC could be used, such as the ASM. The ASM requires a training set of CCs, from which the mean shape and the shape variation of the CC is modelled. By using this prior knowledge the model is intended to generalise the Corpus Callosoa in new images.

Section 4.2 describes how to construct the model, the data that were used during the experiments and how the construction procedure of the model is implemented. In Section 4.3 the experiments that were used for the construction and optimisation of the model are presented. In Section 4.4 the results of the experiments are listed and in Section 4.5 a discussion of the results is performed.

4.2 Materials & Methods

4.2.1 Active shape models

Active shape models are statistical models of the shape of an object and they were firstly introduced by Tim Cootes and Chris Taylor in 1995 [8]. The ASM uses prior information from a set of training images, which represent the boundaries of the objects of interest. In the training images a set of landmarks need to be placed, as the landmarks are the ones to finally produce the shape variation of the model. The landmarks are points on the contours of the object in the training images that represent distinguishable features, such as points of high curvature (e.g. tips of the fingers in a hand). A large number of landmarks establish a dense point correspondence, which are able to explain the shape variation of the object.

After the placement of the landmarks, the contours of the objects from the training images are aligned with the use of Procrustes analysis, which removes rotation, translation and size differences between the training images, so that the distances between
the corresponding landmarks are minimum. The aligned positions of the landmarks of each shape \((x, y)\) in the training set are then grouped to a row vector (Equation 4.1) and all the landmark vectors of all the training contours \((s)\) are then grouped into a matrix (Equation 4.2):

\[
x_1 = (x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n)
\]

(4.1)

\[
X = (x_1, x_2, \ldots, x_s)^T
\]

(4.2)

From the construction of the matrix of the aligned training objects the mean shape is calculated:

\[
\bar{x} = \frac{1}{s} \sum_{i=1}^{s} x_i
\]

(4.3)

Moreover, the covariance matrix of the training data is constructed, from which the eigenvalues \((\lambda)\) and the eigenvectors \((\Phi)\) are calculated:

\[
S = \frac{1}{s-1} \sum_{i=1}^{s} (x_i - \bar{x})(x_i - \bar{x})^T
\]

(4.4)

With these steps the training data can be reconstructed from a shape model based on the following equation:

\[
x = \bar{x} + \Phi b
\]

(4.5)

where \(\bar{x}\) represents the mean shape of the training data, \(\Phi\) represents the eigenvectors and \(b\) the eigenvalues of the model \((b = (\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_s}))\). A final step is the use of a PCA analysis, by keeping less eigenvalues \((b_k = (\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_k}), k < n)\) of the model and consequently capturing a lower variance (e.g. 98 %, 95 %). With this PCA step noise suppression is performed, as noise is present in the lower eigenvalues of the model. Moreover, every new shape in unseen images can be described by equation (4.5), in which the \(b\) values of the model are restricted to an interval of three standard deviations, \(b_{k,i} \in [-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i}]\) (Figure 4.1). Consequently, the object shape in every new image is restricted based on seen images, restricting the model to take implausible shapes. In Figure 4.1 an example of the modelling of the hand can be seen, with the variation of the first three eigenvalues of the model.
4.2.2 Spatial Normalisation

In this step the mid sagittal plane is extracted from the MR volume, in which the extraction of the CC is going to be performed. An efficient and fast way to perform this step, is by registering affinely every patient MR volume to the MNI space. By registering the patient image to the MNI space, a spatial normalisation of the data is performed, as well as alignment of the brain with the axis, since the T1 atlas is perfectly aligned. Consequently, the middle slice of the registered image can be used as the mid sagittal plane of the MR volume, in which the CC segmentation is going to be performed.

4.2.3 Training Data

As it was explained in Section 4.2.1, in order for the ASM to be able to construct a model that describes the shape of the CC, a training set is needed. The data that were used as training were T1-weighted images and were taken from OASIS database (75) [28], ADNI database (40) [21] and from private hospital data of MS patients (35), reaching in total 150 images for the training set. In order to extract the contours of the CC from the mid sagittal slices of the training data a built in brain segmentation tool from the company Icometrix, in Leuven, Belgium was used. The brain segmentation tool is capable of separating the brain in its three substances WM, GM and CSF. After the training data were N4 bias field corrected [3] and spatial normalised to the MNI space, the WM segmentation of each patient was resampled with the same transformation used in the spatial normalisation step, transferring the WM segmentation to the MNI space as well. From the resampled WM volume the middle slice was chosen (Figure 4.2b), as it corresponds to the mid sagittal plane (Figure 4.2a), which was normalised by histogram stretch.
Figure 4.2: (a): Mid Sagittal Plane Extraction after affine registration of the patient MR volume to the MNI space. (b): Middle Slice of the WM segmentation after it was resampled to the MNI space. (c): CC extraction by automatic masking and using largest connected component. (d) Extraction of CC contour.

The two main structures that can be seen in Figure 4.2b, are the CC and the brain stem. In order to remove the latter a mask was automatically applied to the image and afterwards a largest connected component was used for the extraction of the CC and consequently the extraction of its contour. Sometimes, manual intervention was needed, when the fornix was attached to the CC and was also extracted with the largest component. The images that were used for the training set, were taken from healthy and diseased population, so as to make the model more robust in predicting plausible CC shapes of both healthy and diseased people.

4.2.4 Establishing corresponding landmarks between Corpus Callosoa

After the contour extraction of the CC from the training images, corresponding landmarks between the CC training contours should be determined. The large shape variability of the CC, combined with the fact that there are no widely accepted CC landmarks, makes the task of manual annotation on the training images difficult. For these reasons, it was concluded that an automatic scheme had to be used and Minimum Description Length (MDL) principle seemed to fulfil the objectives. The MDL principle has derived from information theory, where it is used to find regularities in code sequences with the ultimate goal of compressing the data [16]. The MDL principle has recently been used also for placing automatic landmarks on the contours of a set of shapes, with the landmarks having a one to one correspondence with each other in these shapes. The MDL can produce a good point correspondence, as you can get a more compact description of the shape variation of the object and for this step the implementation of Thodberg et al. was used [46].

The contours from the training images are initially aligned using Procrustes analysis and the arc length along each contour is normalised to run from 0 to 1 [46]. The MDL attempts to fit $N = 2^L$ marks on each CC contour in a hierarchical manner, where $L$ represents the number of levels to be used. For example, let's assume that we want to place 64 landmarks (6 levels) on the contour of the CC. Since the CC has a closed contour landmarks 0 and 64 are the same. During the first level only landmarks 0 and 32 are placed based on their absolute arc length position. In level 2, marks 16 (between landmarks 0 and 32) and 48 (between landmarks 32 and 64) are placed and so on. After initial placement of the landmarks an iterative optimisation procedure is followed to compute the mean shape and the landmark positions.
procedure on the landmarks attempts to adjust the node parameters to optimise the correspondence of all the marks over the training contours. The mean shape is computed by representing the landmark positions as complex numbers \((x - iy)\) and diagnosing the hermitian \(N \times N\) covariance matrix of the training set [14]. In addition, a PCA analysis is performed on the covariance matrix, yielding the eigenvalue spectrum. The description length is computed as:

\[
\text{DescriptionLength} = \sum L_m \tag{4.6}
\]

\[
L_m = 1 + \log \left( \frac{\lambda_m}{\lambda_{cut}} \right) \quad \text{for} \quad \lambda_m \geq \lambda_{cut} \tag{4.7}
\]

\[
L_m = \frac{\lambda_m}{\lambda_{cut}} \quad \text{for} \quad \lambda_m < \lambda_{cut} \tag{4.8}
\]

where \(\lambda_m\) is the value of the eigenvalue of mode \(m\) and the \(\lambda_{cut}\) is arranged to the pixel resolution we want to achieve.

4.2.5 Fitting of the model in unseen images

The whole procedure that was followed for the extraction of the CC from a T1 MR image can be seen in Figure 4.3. Initially, the T1 patient scan is N4 bias field corrected [3] and then spatially normalised by affinely registering the patient MR scan to a T1 atlas in the MNI space, so as to extract the mid sagittal plane, in which the CC is best represented. Additionally, an intensity normalisation with the use of histogram stretch is performed to the mid sagittal plane. Afterwards, the ASM model is fitted in the mid sagittal plane of the new image segmenting the CC and later on the CC is partitioned into meaningful compartments.

![Flowchart of the CC segmentation pipeline.](image)

In Figure 4.4 the fitting process of the mean shape in new images is described. Initially, the mean shape is initialised as close as possible to the final solution (Figure 4.4a), so as not to get stuck in local minima, which is the most common problem of segmentation using ASM’s [8]. In our case, a good initialisation position was obtained by affinely registering the new MR volume to a T1 MNI atlas. After the initialisation of the mean shape, the intensities across the normals in the landmark positions of the fitted image in each direction (k pixels) are sampled, yielding an intensity gray profile of \(2k+1\) pixels for each landmark \((s_i)\) (Figure 4.4b).

In order to adjust the landmarks in the new image, Cootes et al. [8] proposed the use of the Mahalanobis distance between the gradients of the intensity profiles in the fitted image and their corresponding gradient intensity profiles in the training set through an inverse correlation matrix:
\[
f(g_l) = (g_l - \bar{g})^T (S_g)^{-1} (g_l - \bar{g}) \tag{4.9}
\]

\[
g_l = s(i + 2) - s(i), \quad i = 1, \ldots, 2k - 1 \tag{4.10}
\]

where \(g_l\) is the gradient of the intensity profile in each landmark, \(\bar{g}\) is the mean intensity derivative of the respective landmark \(l\) in the training set and \(S_g\) is the covariance matrix of the intensity derivatives of all the landmarks in the training set. Moreover, by finding the line piece with the smallest Mahalanobis distance to the training profiles (Figure 4.4c) the landmark is moved to the center of this line piece (Figure 4.4d) and then the \(b\) values of the model are restricted to \(\pm 3\text{std's}\) (Figure 4.4e).

Figure 4.4: ASM fitting process (copied from [23]). (a) Initial landmark positions. (b) Long search profiles sampled on contour normals. (c) Search for line piece with smallest Mahalanobis distance to training profiles. (d) Move landmark points to center of line piece with lowest Mahalanobis distance. (e) Map the \(x, y\) points to model coordinates, constrain them to be within 3 standard deviations from the training shapes, and transform them back to \(x, y\) positions.

In our case, instead of using the intensity derivatives, the intensities were used because they were yielding better results. In more detail, a PCA was performed across each of the normal on the landmarks on the training images, from which the resulting eigenvectors, normalised by their variance, were used to calculate the model parameter values for new profiles in the test data [23].

\[
f(s_i) = \frac{(s_i - \bar{s}_x) \cdot \Phi_{PCA}}{b_{PCA}} \tag{4.11}
\]

where \(s_i\) are the intensities in the normals of the landmarks in the fitted image, \(\bar{s}_x\) are the mean intensities on the normals of the landmarks in the training images, \(\Phi_{PCA}\) and \(b_{PCA}\) are...
are the eigenvectors and $b_{PCA}$ are the eigenvalues of the PCA on the intensity profiles in the training images. So, in every new image the distance between the intensity profiles in the fitted image and the normalised PCA mean values of the corresponding training profiles was computed, from which the pixel movement of each landmark was derived. This is an iterative process and a number of 40 iterations was chosen until all the landmarks were adjusted to their final position.

Finally, in order to improve the accuracy of the fitting process, a coarse to fine resolution approach was followed. At the end, from the ASM segmentation the positions of the CC contour were recovered, but they did not correspond to actual pixels. In order to retrieve a pixel based representation, the coordinates of the contour points were rounded.

4.3 Experiments

4.3.1 Introduction

In this section the experiments that were performed in order to construct and validate the model are listed. The experiments that were performed were the following:

Spatial Normalisation
In this experiment the type of registration to perform the spatial normalisation along with the mid sagittal plane extraction was evaluated. The main question that was answered, was if the size of the Corpus Callosum should be included or not in the ASM model.

Sensitivity of the model to landmark placement
In this experiment the optimal number of landmarks to be placed on the CC was investigated.

Accuracy of the model
In this experiment the accuracy of the model was tested.

Segmentation reproducibility of the model
This experiment investigated the segmentation reproducibility of the model.

Segmentation Refinement
In this section an approach to improve the segmentation procedure was presented and tested.

4.3.2 Spatial Normalization

The approach that was used for the spatial normalisation of the data along with the mid sagittal plane extraction, was the registration of every image to a T1 MNI atlas as a pre-processing step. This step enables the spatial normalisation of the data and the alignment of the brain with the axis of the coordinate system performing in this way the mid sagittal plane extraction. The question that was investigated in this step was if rigid or affine registration should be performed in order to achieve higher segmentation accuracy results and consequently if the size of the CC should be included or not in the ASM modelling. The main difference between rigid and affine registration is the scaling factor, as the rigid registration compensates only for rotation and translation, while the affine registration performs translation, rotation, shearing and scaling. With
the rigid registration the original shape and size of the CC are maintained, while with the affine transformation the size information is kept in the affine transformation matrix. The number of landmarks that were used in this experiment on the training images was 32.

The measures that were used to decide which one of the registration techniques is better, were the dice coefficient and the consistency of the $b$ values arising from the ASM segmentation. The dice coefficient is defined between the surface of the ASM segmentation and the ground truth annotation. 50 images of healthy subjects were chosen from the OASIS Longitudinal database as a training set. A rigid training set and an affine training set were constructed, after registering rigidly and affinely the training images to the MNI space, following the procedure in section 4.2.3. As a test set, the 160 images from the OASIS Test Retest dataset were used (20 subjects with 2 time points each and 4 scans in each time point), after they were also rigidly and affinely registered to the MNI space. The idea behind the consistency of the $b$ values is that in a test retest dataset the $b$ values of each segmentation and the area of the CCs should be identical, since it is the same patient, but small differences can be expected. The measures that were used for the consistency of the $b$ values, where the cosine similarity and the standard deviation of the normalised $b$ values. The cosine similarity was computed by calculating the angle of the $b$ values between the scans of the same subject in different time points and then computing the mean angle for each subject (Equation: 4.12).

$$\theta_k = \frac{\sum_{i=1}^{4} b_{t1}^i \cdot b_{t2}^i}{4 \| b_{t1}^i \| \| b_{t2}^i \|}, \quad k = 1, ..., 20 \quad (4.12)$$

where $\theta_k$ represents the mean angle computed for every subject.

Finally, each one of the $b$ values of the model was normalised by a standard deviation of the eigenvalue of the model ($\sqrt{\lambda}$) and the standard deviation in each time point for each patient was computed (Equation: 4.13).

$$b_{ij} = \text{std}\left( \frac{b_{t1}^i}{\sqrt{\lambda_i}}, \ldots, \frac{b_{t4}^i}{\sqrt{\lambda_i}} \right), \quad i = 1, \ldots, n, \quad j = 1, 2 \quad (4.13)$$

where $b_{ij}$ represents the standard deviation of each $b$ value $i$ in time point $j$ for the four scans of each patient and $b_{t1}^i$ represents the $i^{th}$ $b$ value of scan 1 in time point $j$.

Finally, for this experiment a subset of the $b$ values was used (5 out of 9 for the rigid set and 7 out of 14 for the affine set, where 9 and 14 are the $b$ values capturing 98% of the variance of the model), since the lower $b$ values can be corrupted by noise.

### 4.3.3 Sensitivity of the model to landmark placement

Using the optimal registration that was decided in Section 4.3.2, the number of landmarks that would be placed on the training images needed to be defined, so as to achieve a dense point correspondence for the model. As it was explained before, the MDL principle was used for the automatic generation of landmarks on the CC and the implementation of Thodberg et al. [46] was used. In the MDL implementation a maximum of 256 landmarks can be produced for each CC contour and a minimum of 8 can be used for establishing a dense point correspondence. For this experiment the number of landmarks that were used, were a sub-selection of the maximum landmarks. Due to the
fact that the tips of the CC (genu and splenium), are highly variable especially in diseased patients, the experiment that was performed, investigated if additional landmarks on both tips of the CC could yield a higher accuracy segmentation of its boundaries. The 50 images that were used for spatial normalisation in Section 4.3.2 were used as well as a training set for this experiment after they were affinely registered to the MNI space. With the use of the MDL approach, 16, 32, 44 (6 additional in genu and 6 in splenium) and 64 landmarks were placed on them creating four training datasets (Figure 4.5). As a test set the 160 images of the OASIS Test Retest dataset were chosen. The 4 training datasets were used to train the ASM model and each one of the ASM models was used to segment the 160 images of the test set. The measure that was used for the evaluation of the experiment was the dice coefficient of the segmented area for each one of the training sets to the ground truth segmentations.

Figure 4.5: (a) : 16 Landmarks on the CC. (b) : 32 Landmarks on the CC. (c) : 44 Landmarks on the CC. (d) : 64 Landmarks on the CC.

4.3.4 Accuracy of the model with healthy and diseased people

For the evaluation of the segmentation accuracy of the model visual inspection and computation of the dice coefficient between manual annotations and model based segmentations were performed. The visual inspection was used for verification that the model can capture well the boundaries of the CC and identify cases in which the model based segmentation gives unsatisfactory results. The manual annotation involved 30 scans, 10 from each category (Healthy Controls, Alzheimer patients and Multiple Sclerosis patients) and the dice coefficient was used as it can be an efficient and representative measure of the trade off between false positives and false negatives pixels in the segmentation procedure. As a training set of the model, the whole training set (150 images) as it described in Section 4.2.3 was used.
4.3.5 Segmentation reproducibility of the model

An important factor for the ASM model was its segmentation reproducibility, as it is going to be used for the segmentation of healthy and diseased people, where the CC shape is highly variable. For this reason, the segmentation reproducibility using the model was tested with the OASIS test retest dataset, by investigating if the segmentation of sequential scans of the same patient were yielding similar areas. As a training set for the model, the whole training set (150 images) was used. The measure that was used for the evaluation of the experiment was the Percentage Area Change (PAC) and it was computed as follows:

\[
PAC = \frac{|A_{1i}^{t_1} * S^{t_1}_i - A_{2i}^{t_2} * S^{t_2}_i|}{(A_{1i}^{t_1} * S^{t_1}_i + A_{2i}^{t_2} * S^{t_2}_i)/2} \quad j = 1, 2, 3, 4
\]  

(4.14)

where \(A_{1i}^{t_1}\) represents the segmented CC area of scan \(i\) in time point \(1\) and \(S^{t_1}_i\) represents the scaling factor of scan \(i\) in time point \(1\) computed from the determinant of the affine transformation matrix, which was performed in the spatial resolution step.

4.3.6 Segmentation Refinement

The CC is a white matter structure, which has clear boundaries in the mid sagittal plane. This originates from the fact that, in the mid sagittal plane the CC, which has high intensities in T1 images, is surrounded by CSF, which has very low intensities in T1 images. It was observed that during the segmentation procedure, the ASM model tended to oversegment the CC, yielding a high number of false positives, while the false negatives were relatively low. Due to this reason along with the fact that the CC has homogeneous intensities, a minimum threshold was set to the intensities of the original segmentation (minimum threshold = the median of the segmented intensities minus their standard deviation). By thresholding the segmented intensities, the segmentation accuracy was improved, as the misclassified pixels in the boundaries of the CC were excluded to a large degree. In order to verify and quantify the improvement of the thresholding measure, the ASM segmented the OASIS Test Retest database (160 scans) and the OASIS Normal Controls (NC) Cross - Sectional database (1344 scans).

4.4 Results

4.4.1 Spatial Normalisation

For the spatial normalisation of the data, it was investigated whether the rigid or affine registration can yield higher segmentation accuracy results. The experiment yielded similar mean dice coefficients, 84.91 % ± 3.99 % for the rigid registration and 84.58 % ± 4.41 % for the affine registration. The dice coefficients are really close between the two registration types, but from the 160 segmented images 4 of them failed in the rigid set, while none of them failed in the affine set. By failing, it is meant that the ASM model was unable to capture the boundaries of the CC in these cases and it was because the CC was too big to be captured.

23
Apart from the accuracy measurements, the consistency of the $b$ values was also investigated. The consistency of the $b$ values was investigated with respect to two measures, the cosine similarity and the normalisation towards the standard deviation of the eigenvalues of the model. The caption at the bottom of each image represents the mean / median / standard deviation of the measured data. The cosine similarity yielded 0.934 ± 0.109 for the affine set and 0.94 ± 0.088 for the rigid set (Figure 4.7).

In Figure 4.8 the normalised $b$ values (5 out of 9 for the rigid and 7 out of 14 for the affine set) can be seen.
4.4.2 Sensitivity of the model to landmarks placement

In Figure 4.9 the difference in the mean shape and the first six modes of variation are shown for the 32-landmarks-set (blue) and the 44-landmarks-set (red). As it can be seen, the first two modes of variation in the 44-landmarks-set are more elongated in the tips of the CC, while the mean shape is a bit shifted with respect to the mean shape of the 32-landmarks-set, both of which can be explained by the procrustes analysis that is performed in the beginning of the MDL approach.

Moreover, in 4.10 box plots of the dice coefficients for the different number of landmarks are shown.
4.4.3 Accuracy of the model with healthy and diseased people

The dice coefficients of the segmentations of the 30 images can be seen in Figure 4.11:

Figure 4.11: Dice coefficients of manual delineated Corpus Callosum healthy and diseased patients (NC:92.8/93.2/1.608, AD:88.978/88.53/3.317, MS:83.135/86.77/6.241).

The caption in the bottom of the Figure represents the mean/median/standard deviation values of each set.

4.4.4 Segmentation reproducibility of the model

The Percentage Area Change on the OASIS Test Retest dataset can be seen in Figure 4.12. The values below the Figure correspond to the mean/median/standard deviation or the measurements.
Figure 4.12: Percentage Area Change between the same patient in close time points (1.92 / 1.56/ 1.61).

4.4.5 Segmentation Refinement

The results on the dice coefficients that were achieved by thresholding the data can be seen in Figures 4.13a and 4.13b for the OASIS Cross-sectional dataset and the OASIS Test Retest dataset. Moreover, in Figure 4.14 it can be seen that the lesion on the CC is excluded from the segmentation. The red dots represent the false positives pixels and the green dots the false negatives. Finally, the improvement on the dice coefficients that was achieved in manual delineated images from Section 4.4.3 can be seen in Figure 4.15.

Figure 4.13: (a) Minimum thresholding in the segmented intensities OASIS NC Cross-sectional data (Original: 91.72/92.43/2.72, Improved: 94.36/94.99/2.22). (b) Minimum thresholding in the segmented intensities OASIS NC Test Retest data (Original: 91.87/92.26/2.08, Improved: 93.88/94.64/2.09).
4.5 Conclusions and Discussion

In this Chapter a pipeline for the 2D segmentation of the CC was proposed, described and validated. For the spatial normalisation of the data as a preprocessing step, the affine registration was chosen, as no images failed to be segmented. The failed images in the rigid set were due to bad initialisation, which is a problem of the size of the mean shape of the training set and the size of the CC in the image to be segmented. This issue arose from the fact that the length of the normal that is used in the landmarks is set to 8 pixels, which makes it impossible to capture the boundaries of a CC, when the CC is much bigger than the mean shape of the training set. This problem is raised in images which are rigidly registered, as during rigid registration only rotation and translation is performed, leaving the CC in its original size (Figure 4.6).

In addition, due to the fact that in the rigid training set the CC is in its original shape, the highest eigenvalues of the ASM model also include information about the size of the CC on the training set. On the other hand, with the use of affine registration the size of the CC is not included in the model and it is relatively the same in the training and test images. The scaling that was performed during the affine registration can be retrieved from the determinant of the affine transformation matrix. Having the scale separately from the model, can increase the reproducibility of the model, as it is more difficult for images to fail.
As far as the consistency of the $b$ values is concerned for the spatial normalisation, the rigid-set yields lower mean values in both measures (Figure 4.7, 4.8). Although, it seems that the rigid-set has more stability than the affine-set, the reproducibility in the affine-set is higher than in the rigid-set and the accuracy almost the same.

Concerning the sensitivity of the model to landmark placement, the dice coefficients of the four landmark-training-sets (Figure 4.10) yielded more or less the same results, with the 32-landmark-set having the highest mean dice coefficient and the second higher being the 64-landmark-set. Since, the 32-landmark-set yielded the highest dice coefficient results, without having fundamental differences from the rest of the landmark-sets, the 32 landmarks were chosen to be used in the further data processing.

Moreover, the accuracy of the ASM model is high enough ($\sim 88\%$) and can vary depending on the status of the person, diseased or healthy. Especially for MS patients, where lesions can exist on the CC and cannot be excluded from the model the dice coefficient can reach lower values. This problem was outreached with the use of the thresholding measure, which improved the segmentation accuracy from a mean dice of 83% to a mean of 93% for the MS patients, as it was seen in Figure 4.15. On the other hand, with the use of the affine registration as a preprocessing measure, the reproducibility of the model is extremely high ($> 99\%$), as none of the scans failed to be segmented. This arises from the fact that the CC in every new image has similar size as the mean shape of the ASM model and exists in the same position, which solves the classical issue of the ASM with initialisation. Finally, the current implementation yielded high segmentation reproducibility results (mean error of 1.93%), which can consist that the current ASM implementation can be effectively used to segment the CC from healthy and diseased patients images, with high segmentation accuracy.
Chapter 5

Variability in Corpus Callosum measurements

5.1 Introduction

In Chapter 4 the construction and fitting of the ASM model was extensively explained and validated. In this Chapter the pipeline is going to be used to investigate potential differences in measuring the area of the Corpus Callosum in acquisitions of the same patient in scanners of different vendors (Philips, Siemens, GE) and for different Magnetic Fields (Philips 1.5T and 3T).

Section 5.2 describes the data of the patients that were used, their acquisition protocols and some preprocessing steps that were applied. Section 5.3 presents the experiments that were performed in order to explore potential differences between acquisitions in different vendors and between different magnetic fields, while Section 5.4 illustrates the results of the experiments. Finally, the results are discussed in Section 5.5.

5.2 Materials & Methods

5.2.1 Test Retest dataset in Philips, Siemens and GE scanners

This dataset included 3D T1 weighted images of 9 MS patients, which were scanned at the same day in all three vendor MR scanners (Philips-Achieva3T, Siemens-Skyra, GE-MR750w). The procedure included the acquisition of two scans in each vendor, by putting each patient in and out in each scanner. The acquisition protocols that were followed for the three vendors were the following ones:

As it can be seen in Table 5.1, the voxel size of each scanner is different, which can introduce differences in the computations later. For this reason, the scans of Philips and GE were resampled without interpolation, so as to be as close to isotropic voxel size (1.0, 1.0, 1.0). This step was performed by packing together two pixels in every direction, producing one new with the double size. Consequently, a voxel after resampling was equal to 8 voxels of the old resolution. The resampling step was performed due to the fact that the MNI atlas that is used for the spatial normalisation of the data has isotropic voxel size of 1.0 and consequently the resampling of the initial datasets close to isotropic voxel of 1.0 can improve the results of the registration. The T1 MNI atlas is the standard NiftyReg atlas [33] with isotropic voxel size (1.0, 1.0, 1.0) and dimension.
Table 5.1: Dicom Header protocol values for the different vendors.

<table>
<thead>
<tr>
<th></th>
<th>Philips</th>
<th>Siemens</th>
<th>General Electric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Echo Time (ms)</td>
<td>2.303</td>
<td>2.29</td>
<td>3.144</td>
</tr>
<tr>
<td>Repetition Time (ms)</td>
<td>4.963</td>
<td>2300.0</td>
<td>7.328</td>
</tr>
<tr>
<td>Inversion Time (ms)</td>
<td>-</td>
<td>900.0</td>
<td>450.0</td>
</tr>
<tr>
<td>Flip Angle (degrees)</td>
<td>8</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Acquisition Matrix</td>
<td>432x432</td>
<td>256x256</td>
<td>512x512</td>
</tr>
<tr>
<td>Field Strength (Tesla)</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Number of Frames</td>
<td>310</td>
<td>176</td>
<td>328</td>
</tr>
<tr>
<td>Voxel Size (mm)</td>
<td>0.5324x0.5324x0.5</td>
<td>0.9375x0.9375x0.94</td>
<td>0.4297x0.4297x0.5</td>
</tr>
</tbody>
</table>

Furthermore, resampling without interpolation was used, so as not to induce additional error in our measurements due to interpolation. The procedure that was followed to measure the area of the CC was the one described in Section 4.2.5, with two small changes described below.

As it was seen in Section 4.2.5 an affine registration was used for the spatial normalisation of the data. Due to the fact that the affine registration includes shearing effects, which can induce deformations on the contours of the CC, a similarity registration was also used, so as to capture additional shape variations with the segmentation procedure. The similarity registration is similar to the affine registration, since it includes rotation, translation and isotropic scaling without the shearing effects of the affine. Both types of registration were used in the following experiments, with the use of the registration package NiftyReg [33] for the affine registration and Elastix [22] for the similarity registration, both of which use mutual information as a similarity measure. Consequently, two individual test sets were created after registering the data affinely and with a similarity registration to the MNI space. In order to compute the area of the CC in the patient space, the scaling factor of each registration procedure needed to be computed. For the similarity registration with Elastix, the scaling factor which performed the isotropic scaling was at once retrieved from a Transformation parameters file, while for the affine registration with Nifty Reg the scaling factor was computed from the determinant of the affine transformation matrix. By computing the determinant of the affine transformation matrix, the scaling factor of the affine registration included also the deformation of the CC due to shearing effects.

The second change that was performed, included the accurate measurement of the area of the CC. As it was referred in Section 4.2.5, when the ASM was fitted in new images the points of the segmented contour were retrieved and rounded, in order to represent actual pixels, so that the computation of the dice coefficient was possible. In this chapter, since the dice coefficient is not computed a procedure to measure the area of the CC more accurately is introduced. Since the contour points of the segmented CC are available after fitting the ASM model and are quite dense as well, the area integral of the points is computed with the help of the Green-Stokes theorem:

\[
\int \int_{CC} dA = \frac{1}{2} \oint_{C} (-ydx + xdy) = \oint_{C} xdy = -\oint_{C} ydx = -\sum_{i} \oint_{x_{i}}^{x_{i+1}} ydx = (5.1)
\]

\[
\oint_{x_{i}}^{x_{i+1}} ydx = \frac{1}{2}(y_{i+1} - y_{i})(x_{i+1} - x_{i})
\]

where \(i\) represents the number of points that describe the contour of the CC.
5.2.2 Test Retest dataset in 1.5T and 3T Philips scanner

This dataset included 3D T1 weighted images of 20 MS patients that were scanned in 1.5T Philips Intera scanner and 3T Philips Achieva scanner. The acquisition procedure that was followed was a single scan in each one of the two scanners, so in total 40 MR volumes were acquired. The acquisition protocol that was followed can be seen in Table 5.2:

Table 5.2: Dicom Header protocol values for the Philips scanner of different magnetic field strength.

<table>
<thead>
<tr>
<th></th>
<th>1.5T</th>
<th>3T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Echo Time (ms)</td>
<td>4.217</td>
<td>4.6</td>
</tr>
<tr>
<td>Repetition Time (ms)</td>
<td>8.818</td>
<td>9.8618</td>
</tr>
<tr>
<td>Flip Angle( degrees)</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Acquisition Matrix</td>
<td>288x288</td>
<td>288x288</td>
</tr>
<tr>
<td>Number of Frames</td>
<td>130</td>
<td>160</td>
</tr>
<tr>
<td>Voxel Size (mm)</td>
<td>0.8194x0.8194x1.2</td>
<td>0.8333x0.8333x1</td>
</tr>
</tbody>
</table>

The procedure that was followed in order to measure the area of the CC, was the same as the one described in Section 5.2.1, and also both Nifty Reg and Elastix were used. Moreover, during the similarity registration performed with Elastix, 6 out of the 20 subjects failed to be registered. For this reason, only 14 subjects were used in the results of the experiments with Elastix (similarity registration).

5.3 Experiments

5.3.1 Introduction

Percentage Area Change between scanners of different magnetic field strength
In this experiment the differences in the area of the Corpus Callosum between scans of the same subject, that were acquired in scanners of different magnetic field strength were investigated.

Consistency of $b$ values between between scanners of different magnetic field strength
In this experiment the consistency of the $b$ values of the segmented CC of the same subject acquired in 1.5T and 3T Philips scanner was tested. The $b$ are the model shape parameters retrieved from equation 4.5.

Intensity profiles of the Corpus Callosum between scanners of different magnetic field strength
In this experiment the intensity profiles of the Corpus Callosum between scans of the same subject, that were acquired in scanners of different magnetic fields were investigated if they can induce problems in the fitting process.

Consistency of the Corpus Callosum Area within vendors
In this experiment the Percentage Area Change of the CC for scans that were acquired in the same vendor was investigated.
Consistency of the Corpus Callosum Area between vendors
In this experiment the Percentage Area Change of the CC for scans of the same subject acquired in different vendors was computed.

Consistency of the $b$ values within vendors
In this experiment the consistency of the $b$ values within vendors was tested. The $b$ are the model shape parameters retrieved from equation 4.5.

Intensity profiles of the Corpus Callosum in the different vendors
In this experiment the intensity profiles of the Corpus Callosum in the different vendors were investigated if they can induce problems in the fitting process.

5.3.2 Corpus Callosum differences in scanners of different magnetic field strength

Percentage Area Change between between scanners of different magnetic fields
In this experiment the Percentage Area Change of the Corpus Callosum between the scans of the same subject in 1.5T and 3T was computed based on the following formula:

$$PAC_i = \frac{|A_{1.5Ti}^3 \cdot S_{1.5Ti}^3 - A_{3Ti}^1 \cdot S_{3Ti}^1|}{(A_{1.5Ti}^3 \cdot S_{1.5Ti}^3 + A_{1.5Ti}^3 \cdot S_{1.5Ti}^3)/2}, \quad i = 1, \ldots, 20 \quad (5.3)$$

where $A_{1.5Ti}^3$, $A_{3Ti}^3$ represent the area of the CC in the scan of 1.5T and 3T respectively, $S_{1.5Ti}^3$, $S_{3Ti}^3$ represent the scaling factor in 1.5T and 3T respectively and $i$ represents the 20 subjects.

Moreover, the same equation without the absolute value was also used, in order to define if a systematic overestimation or underestimation in the area of the CC is happening between the different magnetic fields.

Consistency of $b$ values between between scanners of different magnetic field strength
In this experiment the consistency of the $b$ values between the scans of the same subject acquired in scanners of different magnetic field strength was investigated. The first 10 out of the 25 $b$ values of the model were used, as the bigger $b$ values of the model define the main shape of the segmented object and are also more invariant to noise.

The measure that was used was the absolute difference of the normalised $b$ values:

$$db_{ij} = \frac{|b_{1.5Tj}^i - b_{3Tj}^i|}{\sqrt{\lambda_j}}, \quad i = 1, \ldots, 20, \quad j = 1, \ldots, 10 \quad (5.4)$$

where $j$ represents the 10 $b$ values that were investigated and $i$ the different subjects.

Intensity profiles of the Corpus Callosum between scanners of different magnetic field strength
In this experiment the histogram of the Corpus Callosum intensities was plotted, as well as the histogram of the intensities of the CSF in the mid sagittal slices. In this way, better insights for the capability of the ASM to fit in images acquired in scanners of different magnetic field strength can be concluded.
5.3.3 Differences in the Corpus Callosum between scanners of different vendors

Consistency of the Corpus Callosum Area within vendors

In this experiment the consistency of the calculated area of the CC for each vendor was computed. Consequently, the Percentage Area Change (PAC) was computed for each subject for each vendor based on the following formula:

\[
PAC_{i,l} = \frac{|A_{l,1}^i \cdot S_{l,1}^i - A_{l,2}^i \cdot S_{l,2}^i|}{(A_{l,1}^i \cdot S_{l,1}^i + A_{l,2}^i \cdot S_{l,2}^i)/2}, \quad i = 1, \ldots, 9, \quad l = 1, 2, 3 \quad (5.5)
\]

where \(A_{l,1}^i\) represents the area of subject \(i\) for the first scan in vendor scanner \(l\) and \(S_{l,1}^i\) the scaling factor of subject \(i\) for the first scan in vendor scanner \(l\). The scaling factor is taken from the transformation parameters for the similarity transformation and by computing the determinant of the affine transformation matrix for the affine registration.

Consistency of the Corpus Callosum Area between vendors

This experiment investigated if the same area of the CC can be found in scans of the same subject in the three vendors. In order to achieve that, the PAC was computed again, but now between respective scans of the different vendors.

\[
PAC_{PH-SI,i} = \frac{2}{\sqrt{\lambda_i}} \sum_{k=1}^{2} \frac{|A_k^{l_1} \cdot S_k^{l_1} - A_k^{l_2} \cdot S_k^{l_2}|}{(A_k^{l_1} \cdot S_k^{l_1} + A_k^{l_2} \cdot S_k^{l_2})}, \quad i = 1, \ldots, 9 \quad (5.6)
\]

\[
PAC_{PH-GE,i} = \frac{2}{\sqrt{\lambda_i}} \sum_{k=1}^{2} \frac{|A_k^{l_1} \cdot S_k^{l_1} - A_k^{l_3} \cdot S_k^{l_3}|}{(A_k^{l_1} \cdot S_k^{l_1} + A_k^{l_3} \cdot S_k^{l_3})}, \quad i = 1, \ldots, 9 \quad (5.7)
\]

\[
PAC_{SI-GE,i} = \frac{2}{\sqrt{\lambda_i}} \sum_{k=1}^{2} \frac{|A_k^{l_2} \cdot S_k^{l_2} - A_k^{l_3} \cdot S_k^{l_3}|}{(A_k^{l_2} \cdot S_k^{l_2} + A_k^{l_3} \cdot S_k^{l_3})}, \quad i = 1, \ldots, 9 \quad (5.8)
\]

where \(k\) represents the two scans in each vendor, \(i\) represents each one of the 9 subjects, PH stands for Philips scanner, SI for Siemens scanner and GE for General Electric scanner.

Consistency of the \(b\) values

The consistency of the \(b\) values in the segmentation of the 9 subjects was also investigated. The ASM model that was used for the segmentation of these images consists of 25 \(b\) values, so every segmented CC is produced from these 25 \(b\) values. From these 25 values the first 10 were chosen to be used for the investigation of their consistency, as the highest \(b\) values describe the main segmented shape and are more invariant to noise.

The measure that was used, was the absolute difference of each \(b\) value, normalised by a standard deviation of the eigenvalue of the model \(\sqrt{\lambda}\), between the two scans that were acquired in each vendor:

\[
b_{lj} = \frac{|b_{l_1,j}^i - b_{l_2,j}^i|}{\sqrt{\lambda_j}}, \quad l = 1, 2, 3 \quad j = 1, \ldots, 10 \quad (5.9)
\]

where \(b_{l_1,j}^i\) represents the \(b\) value \(j\) of vendor \(l\) in the first scan.
Intensity profiles of the Corpus Callosum in the different vendors

During this experiment, the histogram of the CC intensities in the different vendors was plotted, as well as the histogram of the CSF in the mid sagittal slices.

5.4 Results

5.4.1 Corpus Callosum differences in scanners of different Magnetic Field Strength

Percentage Area Change between between scanners of different magnetic fields

In Figure 5.1a the absolute PAC between the scans of 3T and 1.5T with the registration performed with an affine and a similarity registration is plotted, while in Figure 5.1b the PAC is again plotted but without getting its absolute value. As it can be seen a mean of 7% difference between the Corpus Callosum Area of the 3T and 1.5T is found, and from Figure 5.1b it can be concluded that a systematic overestimation of mean around 6-7% of the area of the CC in the 3T scanner, both for the similarity and the affine registration, is taking place. It is worth mentioning, that the mean / median / standard deviation has not changed when performing the spatial normalisation with a similarity registration, while it changes when using an affine one. This means, that by using a similarity registration a constant overestimation of the CC area in the 3T is taking place, while with the affine registration, in some subjects, the opposite results is occurring (overestimation of 1.5T CC area).

![Figure 5.1: (a) Absolute PAC between scans of 3T and 1.5T (Nifty Reg: 6.87 / 6.1 / 4.84, Elastix: 6.92 / 5.99 / 4.96). (b) PAC between scans of 3T and 1.5T (Nifty Reg: 5.72 / 5.43 / 6.16, Elastix: 6.92 / 5.99 / 4.96).](image-url)
**Consistency of $b$ values between between scanners of different magnetic fields**

In Figure 5.2 the absolute difference of the first 10 $b$ values of the model is plotted, both for the affine and similarity registration. As it can be seen, the first $b$ value is stable for both registration methods, which means that the accuracy segmentation is quite good. On the other hand, most of the rest of the $b$ values have higher variance in the similarity registration than in the affine one and especially the second and the third, which implies that quite different shapes are acquired during segmentation.

![Figure 5.2](image)

**Intensity profiles of the Corpus Callosum between scanners of different magnetic fields**

In Figure 5.3a the intensity distribution of the CC between 1.5T and 3T can be seen, while in Figure 5.3b the intensity distribution of the CSF between 1.5T and 3T in the mid sagittal plane is shown. It can be observed, that the intensities of the CC in the 1.5T scanner are lower that the ones in the 3T scanner, while the CSF intensities are not different.

![Figure 5.3](image)
5.4.2 Differences in the Corpus Callosum between scanners of different vendors

Consistency of the Corpus Callosum Area within vendors

In Figure 5.4a the PAC between scans of the same vendors is presented for the affine (5.4a) and the similarity registration (5.4b). The caption in the bottom of the figure represents the mean / median / standard deviation of the PAC.

![Figure 5.4a](image1)

(a) PAC within each vendor with the use of an affine registration for the spatial normalisation (Philips : 2.46 / 2.43 / 1.57, Siemens : 1.91 / 2.35 / 1.07, GE : 1.3 / 1.55 / 1.0).

![Figure 5.4b](image2)

(b) PAC within each vendor with the use of a similarity registration for the spatial normalisation (Philips : 1.73 / 1.14 / 1.7, Siemens : 2.98 / 2.81 / 1.54, GE : 1.69 / 1.54 / 0.74).

Consistency of the Corpus Callosum Area between vendors

In Figure 5.5 the PAC between scans of the same subject in different vendors was computed for the affine (5.5a) and the similarity registration (5.5b).

Consistency of the $b$ values

In Figure 5.6 the absolute difference of the first 10 $b$ values of the model is plotted, both for the affine and similarity registration. It can be observed, that the first $b$ value is consistently estimated for both registration types, while this quickly decreases for the higher $b$ values, apart from the 6th $b$ value. Moreover, the variance of the $b$ values with the similarity registration is higher in almost all the modes, than in the affine one. This means that the ASM model can capture more accurately the CC, by performing an affine that a similarity transformation.

Intensity profiles of the Corpus Callosum in the different vendors

In Figure 5.7 the difference in the intensity distribution of the Corpus Callosa between the vendors is presented (Figure 5.7a), as well as the intensity distribution of the CSF in the mid sagittal planes between the vendors (Figure 5.7b). It can be observed, that the CC intensities in the Siemens scanner are lower than the CC intensities in the Philips and GE scanners, where they are quite similar. On the other hand, the CSF intensities
in the Siemens scanner are also lower than the ones in Philips and GE scanners, but less far away than in the CC intensities.
5.5 Discussion & Conclusions

In this chapter we compared the size and shape of the Corpus Callosum in scans at different field strengths and acquired in machines from different vendors. The assumption that could be made for the data that were used in this chapter is that the CC area of the same subject either scanned in different vendors or scanned in scanners of different magnetic field strengths should be the same. The difference in area that was found between scans of 1.5T and 3T, both with Nifty Reg and Elastix, was around a median of 7% (Figure 5.1), with a constant overestimation of the CC area for the 3T scanners. This is a rather high percentage if we take into account that the atrophy of the CC area of 2% in the first year of disease, was proposed as a predictor for MS progression [48].

After applying the ASM model to segment the CC from the mid sagittal planes of both scans a visual inspection was performed, which confirmed that the accuracy segmentation of the CC was satisfactory. A reason for the overestimation of the CC area in the 3T scanners can be due to the higher magnetic field strength. In the 3T scanners, due to the higher magnetic field strength the resolution of the image is higher, producing stronger edges between the different tissues in the brain. Consequently, in the pixel level the ASM might tend to over segment the CC area. Nevertheless, from Figures 5.2 and 5.3 it can be observed that the ASM model was capable of producing similar segmented shapes of the CC. The first three $b$ values that are the most reliable, as they are more invariant to noise, and define the main shape of the segmented object are stable (Figure 5.2), while a higher variability is observed in the other $b$ values. This variance can be explained due to registration mismatch, as during the acquisitions in the two scanners, the subject is positioned differently. The different position especially of the neck, can induce small differences in the mid sagittal planes of the two acquired scans, which is later on shown in the $b$ values of the model.

As far as the analysis of the scans acquired in different vendors is concerned, the PAC within the vendors (Figure 5.4b) with the use of an affine registration was found to have a median of 2.43%, 2.35%, 1.55% for Philips, Siemens, GE, while with the similarity one a median of 1.14%, 2.81%, 1.54% was reported respectively. These values are relatively low and are similar to the values reported in Chapter 4 (Figure 4.12) of a median of 1.56% for healthy subjects. Furthermore, in Figure 5.5 the PAC between vendors is shown and it can be observed that for Philips and GE similar areas are found (median PAC of 2.74%), while for Siemens compared to the other vendors much larger differences of a median of about 6-7% are computed. In addition, the first $b$ value is
consistently computed, while the rest of the $b$ values have a higher variance. The $b$ values from the affine registration have in general lower variance, than the ones with the similarity transformation.

A fact that is worth mentioning, is that in Figure 5.4b, the similarity registration performs better than the affine registration for Philips scanner and in Figure 5.5, the similarity registration shows higher compliance for Philips-GE than the affine one. On the first sight this is strange, as the affine transformation has more degrees of freedom than the similarity transformation. An explanation for that is that during the affine registration, shearing effects are also induced which cause deformation on the shape of the CC and consequently squeezing the CC. This means, that a segmentation error of the ASM model in images registered with the affine registration accounts for a bigger area change than in images registered with a similarity one.

A possible reason for the high differences between the Siemens scanner and the other ones, can be the differences in the intensity profiles of the CC in the different scanners. As it can be seen in Figure 5.7, the intensities of the CC in the Siemens scanners are lower, than the ones of GE and Philips. Since, the CC is a WM structure and we examine it in a T1 weighted image it has high values, while the CSF that exist in its surroundings has much lower values. The higher values in the GE and Philips scanners, can produce stronger edges in the boundaries of the CC and as a result the ASM segmentation can work better, since it is also based on the background intensities of the CC and not only in its shape.

To sum up, it can be concluded that the PAC within the vendors is reasonably low and can be due to the segmentation procedure used, while the PAC between vendors and between scans of different magnetic field strengths is relatively high. These differences can be due to the fitting of the ASM model, as its use is based on the intensities of the object relative to the background or it can be due to the differences in the acquisition and reconstruction algorithms that are used from the different vendors. Finally, a clinical conclusion, that could be derived from this analysis is that for the construction of longitudinal data, the follow up scans of the patients should be acquired in the same scanner type, as the differences in acquisition protocols may yield unwanted results.
Chapter 6

Corpus Callosum shape and area changes in the ageing & diseased population

6.1  Introduction

The purpose of this thesis is the investigation of the correlation of global and regional shape and area changes of the CC with Alzheimer’s Disease (AD) and Multiple Sclerosis (MS). Using our pipeline for accurate segmentation of the CC from T1 MR images of healthy and diseased populations, we investigate if the CC could be used as a biomarker for Alzheimer’s Disease and Multiple Sclerosis.

Section 6.2 describes the data and the methods that were used to define the clinical use of the CC pipeline. Section 6.3 lists the experiments that were performed, while section 6.4 shows the results of the experiments. Finally, in Section 6.5 the results are discussed.

6.2  Materials & Methods

6.2.1  Data

The data that were used for these experiments were 1209 scans of healthy controls taken from 416 (160 male and 256 female) subjects from the OASIS Cross sectional database aged between 18 and 92 years old, 210 (97 male and 113 female) Alzheimer patients taken from ADNI database with ages between 56 and 88 and another 134 (70 male and 64 female) Alzheimer patients with unknown age, and 66 MS patients taken from private hospital data with unknown inclusion criteria. The Alzheimer scans that were chosen from the ADNI database were the first scans after diagnosis for the Alzheimer patients.

The procedure that was followed to segment the CC is the one described in Section 4.2.5. Both Nifty Reg and Elastix softwares were used for the spatial normalisation of the data, by performing an affine and a similarity registration respectively, as in Chapter 5. Moreover, the area and the shape changes of the CC were computed for the ageing population. In order to calculate shape changes, apart from the shape parame-
ters ($b$ values) of the ASM model, the Circularity (CIR) of the CC was also used and it was defined as:

$$CIR = \frac{4 \piCCA}{CCP^2}$$ (6.1)

where CCA represents the area of the CC and CCP represents its arc length.

Furthermore, in order to evaluate regional area and shape changes of the CC, the CC was partitioned in five compartments according to Wittelson partitioning, which is described in Appendix A. By splitting the CC, a more thorough investigation of its regional changes was performed. Finally, the possibility of the use of the CC as a biomarker for Alzheimer’s Disease and Multiple sclerosis was investigated.

6.3 Experiments

6.3.1 Introduction

Healthy population graphs based on age
In this experiment population graphs of the healthy subjects based on the area of the CC and its circularity were constructed, in order to get a better understanding of how the CC changes with ageing.

Corpus Callosum as a Biomarker for Alzheimer’s Disease
In this experiment the separability of the Healthy Controls from the Alzheimer patients based on the age of the subject, the area and shape of the CC was investigated.

Corpus Callosum as a Biomarker for Multiple Sclerosis
In this experiment the separability of the Healthy Controls from the Multiple Sclerosis patients based on the area and shape of the CC was investigated.

Global & Regional shape and area changes during Alzheimer’s Disease
In this experiment statistical tests were performed for the CC as a whole and in each one of its compartments, in order to investigate if the CC changes can be correlated with Alzheimer’s Disease.

Global & Regional shape and area changes during Multiple Sclerosis
In this experiment statistical tests were performed for the CC as a whole and in each one of its compartments, in order to investigate if the CC changes can be correlated with Multiple Sclerosis.

6.3.2 Healthy population graphs

Initially, the changes of the CC area and the circularity were investigated in the ageing population. For this reason, 1209 healthy scans were used from the OASIS Cross Sectional database, and population graphs were constructed. The area and the circularity of the CC in each scan were computed from the segmentations of the CC in the MNI space, since in the MNI space the area of the CC is normalised. This means that in the MNI space the CC is normalised with respect to the total brain volume. This is a compulsory normalisation as in a former study [26] it has been shown, that people with bigger brains tend to have a bigger CC as well. Finally, for the registration of the subjects to the MNI space both a similarity (Elastix) and an affine (Nifty Reg) registration were used.
6.3.3 Corpus Callosum as a Biomarker for Alzheimer’s Disease

In this experiment the capability of using the shape and area changes of the CC as a biomarker for Alzheimer’s Disease was evaluated. 210 Alzheimer scans were used and they were the first scans after diagnosis of each patient. The reasoning behind that was to investigate if we could differentiate healthy people from Alzheimer patients in their initial stage. From the healthy patients 393 were chosen, from subjects above 55 years of age, which is the onset age of Alzheimer’s Disease [32]. For the purpose of the experiment, a vector was constructed for each scan, which included the 38 values listed below:

1. age of the subject
2. total circularity of the CC
3. total area of the CC of the subject
4. area of each one of the five CC compartments
5. circularity of each one of the five CC compartment
6. 25 \( b \) values of the ASM model

In this way a matrix of 603 scans was constructed. A PCA was performed on the data, which reduced the dimensions of the data to 22, keeping 99% of the total variance. Moreover, an LDA was performed on the data yielding one dimension of separability, since the classes that we have are two (healthy and Alzheimer people). The 22 dimensions from the PCA with the 1 dimension from the LDA constructed the new matrix of the data, which was fed to an SVM. In order to validate the efficiency of the SVM a 5-fold cross validation was performed to the new matrix of the data. Moreover, the ROC curve of the data with respect to the LDA dimension was plotted. The whole procedure was performed twice, by spatially normalising the scans with a similarity and an affine registration.

6.3.4 Corpus Callosum as a Biomarker for Multiple Sclerosis

In this experiment, the ability of the CC to be used as a biomarker for Multiple Sclerosis was investigated. From the healthy subjects 816 were chosen, aged below 55, and all the Multiple Sclerosis scans (66). As in Section 6.3.3 a similar vector was also constructed for each scan, but due to the fact that the ages of the MS patients were unknown, the vector included the following features (37):

1. total area of the CC of the subject
2. total circularity of the CC
3. area of each one of the five CC compartments
4. circularity of each one of the five CC compartment
5. 25 \( b \) values of the ASM model
A matrix of 882 scans was constructed, in which a PCA and an LDA was performed. After the PCA, 22 dimensions were kept explaining 99% of the variance and one dimension was yielded from the LDA, constructing in this way a matrix of 882 by 23, which included the projected data. The projected data, were fed to an SVM and a 5-fold cross validation was performed. In order to balance the large difference in the sample size of the data, a different class prior was used in the parameters of the SVM. Finally, the ROC curve of the data with respect to the LDA dimension was constructed. The whole procedure was performed twice, by spatially normalising the scans with a similarity and an affine registration.

6.3.5 Global & Regional shape and area changes during Alzheimer’s Disease

In order to define if global or regional area and shape changes of the CC could be correlated with Alzheimer’s Disease, statistical tests with a significance level of 5% were performed for the area and circularity of the CC as a whole and for each one of its compartments. More specifically, if the groups that were tested followed a normal distribution and had similar variances a student’s T-test was performed, else if the variances were different a Welch test was performed. Moreover, if one of the groups didn’t follow a normal distribution a Mann Whitney U-test was performed.

The Corpus Callosum was split in 5 regions based on the Wittelson partitioning and 7 statistical tests were performed, for the area of its one of the compartments of the CC, the total CC area and total CC circularity. Finally, the first three \( b \) values were chosen from the segmented NC and AD scans and statistical tests were performed in each one of them. The first three \( b \) values were chosen as they define the main segmented shape of the CC, are more invariant to noise and as it was observed in Chapter 5 are stable at least for the affine registration.

The subjects that were used were the 393 normal control scans that were aged above 55 years old and the Alzheimer group consisted of 344 Alzheimer scans (210 scans in which the age was known and 134 scans in which their age was unknown). The whole procedure was performed twice, by spatially normalising the scans with a similarity and an affine registration.

6.3.6 Global & Regional shape and area changes during Multiple Sclerosis

The same experiments as in Section 6.3.5 were performed, but now between 816 normal control scans, aged below 55 years old, and 66 Multiple Sclerosis scans.

6.4 Results

6.4.1 Healthy population graphs based on age

In Figure 6.1 the population graphs of the area and circularity of the CC using Nifty Reg are shown, while in Figure 6.2 the same graphs with the use of Elastix are displayed. The dashed red lines represent the 5th and 95th percentile of the data and their in between red line represents the 50th percentile of the data. Finally, the green crosses represent the data points.
In both Figures it can be seen that the area fluctuates with age and it does not have a standardised pattern in the whole age range. A constant decrease in the area of the CC can be seen only for the age range above 65 years old. On the other hand the circularity has a constant decrease in the whole age span. From this experiment it can be concluded, that circularity is more consistent and it could be more reliably used as a measure of change of the CC, than the CC area.

Figure 6.1: (a) Population graph of the area of the CC using an affine registration. (b) Population graph of the circularity of the CC using an affine registration.

Figure 6.2: (a) Population graph of the area of the CC using a similarity registration. (b) Population graph of the circularity of the CC using a similarity registration.
6.4.2 Corpus Callosum as a Biomarker for Alzheimer’s Disease

In Table 6.1 the results of the 5-fold cross validation of the projected data (after the PCA and LDA was performed) fed to the SVM are depicted. As it can be observed the affine transformation yields higher results for all the values, which means that the data registered with an affine transformation are more separable. This can also be observed in the ROC curves of the data with respect to the LDA dimension in Figure 6.3. In this experiment a true positive is defined as a correct classification of a healthy person and a true negative as the correct classification of an Alzheimer patient.

Moreover, in order to define which one of the features contributes more to the classification, subsets of the features were used. A PCA keeping 99% of the variance and an LDA were used and then the data were fed to the SVM performing a 5-fold cross validation. From the results in Tables 6.2 and 6.3, is can be observed that most of the contribution to the classification emerges from the $b$ values of the model and especially from the first 10 both for the affine and the similarity registration.

Table 6.1: Accuracy, Sensitivity and Specificity of SVM for the separation of NC-MS with the whole feature set

<table>
<thead>
<tr>
<th>NC-AD</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affine Registration</td>
<td>96.01 ± 1.70</td>
<td>94.83 ± 2.78</td>
<td>98.01 ± 1.81</td>
</tr>
<tr>
<td>Similarity Registration</td>
<td>93.51 ± 2.05</td>
<td>92.56 ± 2.66</td>
<td>95.43 ± 2.72</td>
</tr>
</tbody>
</table>

Table 6.2: Accuracy, Sensitivity and Specificity of SVM for the separation of NC-AD in different feature subsets with the use of an affine registration

<table>
<thead>
<tr>
<th>Affine Registration NC-AD</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age-Total Area</td>
<td>68.95</td>
<td>70.89</td>
<td>64.13</td>
</tr>
<tr>
<td>Age-Total Area-Total CIR</td>
<td>69.21</td>
<td>71.19</td>
<td>64.94</td>
</tr>
<tr>
<td>Area of Compartments</td>
<td>67.06</td>
<td>69.20</td>
<td>65.85</td>
</tr>
<tr>
<td>CIR of Compartments</td>
<td>55.32</td>
<td>55.44</td>
<td>54.12</td>
</tr>
<tr>
<td>$b$ values</td>
<td>96.16</td>
<td>94.98</td>
<td>97.50</td>
</tr>
<tr>
<td>10 $b$ values</td>
<td>94.25</td>
<td>92.80</td>
<td>95.59</td>
</tr>
</tbody>
</table>

Table 6.3: Accuracy, Sensitivity and Specificity of SVM for the separation of NC-AD in different feature subsets with the use of a similarity registration

<table>
<thead>
<tr>
<th>Similarity Registration NC-AD</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age-Total Area</td>
<td>68.56</td>
<td>69.70</td>
<td>62.94</td>
</tr>
<tr>
<td>Age-Total Area-Total CIR</td>
<td>68.87</td>
<td>70.25</td>
<td>65.06</td>
</tr>
<tr>
<td>Area of Compartments</td>
<td>66.23</td>
<td>69.62</td>
<td>63.36</td>
</tr>
<tr>
<td>CIR of Compartments</td>
<td>54.95</td>
<td>54.94</td>
<td>52.87</td>
</tr>
<tr>
<td>$b$ values</td>
<td>92.77</td>
<td>92.99</td>
<td>92.60</td>
</tr>
<tr>
<td>10 $b$ values</td>
<td>90.16</td>
<td>90.74</td>
<td>89.54</td>
</tr>
</tbody>
</table>
6.4.3 Corpus Callosum as a Biomarker for Multiple Sclerosis

For the validation of the results of the SVM, a 5-fold cross validation was performed to the projected data and the results can be seen in Table 6.4. From the results it can be observed that the scans registered with the similarity transformation yield a higher Specificity close to 100%, which means that when a scan is classified as MS, then it is definitely an MS patient. This can be seen in the ROC curves in Figure 6.4, where scans registered with a similarity transformation yield a mean area under the curve of 0.96. In this experiment a true positive is defined as a correct classification of a healthy person and a true negative as the correct classification of an MS patient.

Moreover, in order to define which one of the features contributes more to the classification, subsets of the features were used. A PCA keeping 99% of the variance and an LDA were used and then the data were fed to the SVM performing a 5-fold cross validation. From the results in Tables 6.5 and 6.6, it can be concluded that most of the contribution to the classification stems from the b-values of the model. It is interesting to observe that more that 10% of the specificity results from the 15 last b-values both for the affine and similarity registration. This means that the small deviations in the shape of the CC represented by the smaller b-values contribute quite a lot to the separation of the NC and MS scans. Moreover, the high sensitivity can be deceiving, as it means that most of the normal controls are correctly classified and due to the large group size difference the sensitivity of the classifier is always high.

Table 6.4: Accuracy, Sensitivity and Specificity of SVM for the separation of NC-MS with the whole feature set

<table>
<thead>
<tr>
<th>NC-MS</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affine Registration Reg</td>
<td>94.94 ± 2.23</td>
<td>95.04 ± 2.58</td>
<td>94.18 ± 7.91</td>
</tr>
<tr>
<td>Similarity Registration</td>
<td>95.46 ± 1.61</td>
<td>95.42 ± 1.63</td>
<td>97.78 ± 4.44</td>
</tr>
</tbody>
</table>

Figure 6.3: (a) ROC curve with respect to the LDA dimension for NC,AD patients with the use of an affine registration. (b) ROC curve with respect to the LDA dimension for NC,AD patients with the use of a similarity registration.
Table 6.5: Accuracy, Sensitivity and Specificity of SVM for the separation of NC-MS in different feature subsets with the use of an affine registration

<table>
<thead>
<tr>
<th>Affine Registration NC-MS</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Area-Total CIR</td>
<td>90.58</td>
<td>95.90</td>
<td>48.88</td>
</tr>
<tr>
<td>Area of Compartments</td>
<td>85.14</td>
<td>95.83</td>
<td>32.42</td>
</tr>
<tr>
<td>CIR of Compartments</td>
<td>78.28</td>
<td>96.26</td>
<td>23.77</td>
</tr>
<tr>
<td>b values</td>
<td>94.96</td>
<td>94.95</td>
<td>95.85</td>
</tr>
<tr>
<td>10 b values</td>
<td>93.89</td>
<td>95.04</td>
<td>74.78</td>
</tr>
</tbody>
</table>

Table 6.6: Accuracy, Sensitivity and Specificity of SVM for the separation of NC-MS in different feature subsets with the use of a similarity registration

<table>
<thead>
<tr>
<th>Similarity Registration NC-MS</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Area-Total CIR</td>
<td>90.22</td>
<td>95.26</td>
<td>39.76</td>
</tr>
<tr>
<td>Area of Compartments</td>
<td>84.81</td>
<td>96.55</td>
<td>27.26</td>
</tr>
<tr>
<td>CIR of Compartments</td>
<td>82.47</td>
<td>96.39</td>
<td>23.70</td>
</tr>
<tr>
<td>b values</td>
<td>96.11</td>
<td>96.06</td>
<td>97.69</td>
</tr>
<tr>
<td>10 b values</td>
<td>95.73</td>
<td>96.11</td>
<td>87.55</td>
</tr>
</tbody>
</table>

Figure 6.4: (a) ROC curve with respect to the LDA dimension for NC,MS patients with the use of an affine registration (b) ROC curve with respect to the LDA dimension for NC,MS patients with the use of a similarity registration

6.4.4 Global & Regional shape and area changes during Alzheimer’s Disease

From the results of the statistical tests that are in Tables 6.7 and 6.8 it can be observed that statistical differences were found for the total area and circularity of the CC (p ≪ 0.01) both for the affine (Figure 6.5) and the similarity registration (Figure 6.6).

Moreover, for the area of the compartments of the CC, the tests showed statistical differences both for the affine and similarity registration in all the compartments apart from the splenium. In that compartment, the affine registration showed statistical differences (p = 0.013), while the similarity registration didn’t yield any differences (p ≫ 0.05). The more pronounced differences in both the affine and the similarity registration, were found in genu and isthmus and the histograms of the areas found in these two compartments can be seen in Figures 6.7 and 6.8. As it can be seen, the differences
with the use of a similarity registration are more pronounced, and in both cases the area of the AD patients is lower than the area of the NC.

Finally, both for an affine and a similarity registration statistical differences were found in the second $b$ value and their histograms can be seen in Figure 6.9.

Table 6.7: Mean and Standard deviation of the CC and its compartments for NC and AD with the use of an affine registration

<table>
<thead>
<tr>
<th>Affine Reg</th>
<th>Mean NC / AD</th>
<th>Std NC / AD</th>
<th>Norm test NC/AD</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total CC Area</td>
<td>672.58 / 612.72</td>
<td>86.57 / 83.09</td>
<td>0.003 / 0.08</td>
<td>1.44 · 10⁻¹⁸</td>
</tr>
<tr>
<td>Total CIR</td>
<td>0.158 / 0.146</td>
<td>0.023 / 0.024</td>
<td>0.0007 / 0.04</td>
<td>6.53 · 10⁻¹¹</td>
</tr>
<tr>
<td>Genu</td>
<td>276.51 / 244.84</td>
<td>42.46 / 35.99</td>
<td>0.0007 / 4.49 · 10⁻⁵</td>
<td>1.94 · 10⁻²⁴</td>
</tr>
<tr>
<td>Anterior Body</td>
<td>76.04 / 69.68</td>
<td>11.43 / 10.75</td>
<td>0.08 / 6.63 · 10⁻²⁰</td>
<td>1.52 · 10⁻¹⁵</td>
</tr>
<tr>
<td>Mid Body</td>
<td>63.87 / 60.19</td>
<td>8.85 / 9.07</td>
<td>3.20 · 10⁻⁷ / 2.47 · 10⁻¹⁶</td>
<td>8.78 · 10⁻⁸</td>
</tr>
<tr>
<td>Isthmus</td>
<td>65.64 / 54.49</td>
<td>22.73 / 18.57</td>
<td>9.53 · 10⁻⁴³ / 1.12 · 10⁻⁴⁰</td>
<td>3.45 · 10⁻²³</td>
</tr>
<tr>
<td>Splenium</td>
<td>190.52 / 183.52</td>
<td>38.10 / 36.02</td>
<td>1.6410⁻¹⁶ / 3.25 · 10⁻⁵</td>
<td>0.013</td>
</tr>
<tr>
<td>$b_1$</td>
<td>-6.64 / -1.59</td>
<td>136.20 / 126.47</td>
<td>0.534 / 0.53</td>
<td>0.31</td>
</tr>
<tr>
<td>$b_2$</td>
<td>-44.31 / -25.36</td>
<td>83.81 / 75.00</td>
<td>7.193 · 10⁻⁹ / 0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>$b_3$</td>
<td>69.81 / 73.61</td>
<td>79.58 / 75.95</td>
<td>0.01 / 0.055</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Table 6.8: Mean and Standard deviation of the CC and its compartments for NC and AD with the use of a similarity registration

<table>
<thead>
<tr>
<th>Similarity Reg</th>
<th>Mean NC / AD</th>
<th>Std NC / AD</th>
<th>Norm test NC/AD</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total CC Area</td>
<td>618.58 / 564.52</td>
<td>80.87 / 85.71</td>
<td>0.001 / 0.06</td>
<td>7.31 · 10⁻¹⁷</td>
</tr>
<tr>
<td>Total CIR</td>
<td>0.156 / 0.142</td>
<td>0.022 / 0.023</td>
<td>0.001 / 0.025</td>
<td>6.25 · 10⁻¹⁶</td>
</tr>
<tr>
<td>Genu</td>
<td>259.95 / 227.58</td>
<td>41.11 / 37.92</td>
<td>0.104 / 0.01</td>
<td>4.69 · 10⁻²⁴</td>
</tr>
<tr>
<td>Anterior Body</td>
<td>72.09 / 66.52</td>
<td>10.11 / 9.70</td>
<td>0.121 / 0.0003</td>
<td>6.22 · 10⁻¹⁴</td>
</tr>
<tr>
<td>Mid Body</td>
<td>59.56 / 57.10</td>
<td>8.09 / 8.66</td>
<td>0.614 / 2.59 · 10⁻¹⁵</td>
<td>5.14 · 10⁻⁶</td>
</tr>
<tr>
<td>Isthmus</td>
<td>74.14 / 57.69</td>
<td>31.00 / 24.50</td>
<td>5.97 · 10⁻²⁵ / 6.13 · 10⁻³⁰</td>
<td>1.14 · 10⁻¹⁴</td>
</tr>
<tr>
<td>Splenium</td>
<td>152.84 / 155.63</td>
<td>46.05 / 39.39</td>
<td>1.76 · 10⁻⁹ / 2.32 · 10⁻⁰⁶</td>
<td>0.39</td>
</tr>
<tr>
<td>$b_1$</td>
<td>-20.76 / -2.88</td>
<td>134.00 / 151.20</td>
<td>0.77 / 0.10</td>
<td>0.063</td>
</tr>
<tr>
<td>$b_2$</td>
<td>-60.72 / -80.23</td>
<td>78.19 / 97.55</td>
<td>1.46 · 10⁻⁵ / 0.00049</td>
<td>0.004</td>
</tr>
<tr>
<td>$b_3$</td>
<td>41.79 / 51.75</td>
<td>72.81 / 90.83</td>
<td>0.29 / 0.08</td>
<td>0.13</td>
</tr>
</tbody>
</table>
Figure 6.5: (a) Histogram of the total area of the CC in NC and AD scans with the use of an affine registration. (b) Histogram of the total circularity of the CC in NC and AD scans with the use of an affine transformation.

Figure 6.6: (a) Histogram of the area of the genu in NC and AD scans with the use of a similarity registration. (b) Histogram of the area of the isthmus in NC and AD scans with the use of a similarity registration.

Figure 6.7: (a) Histogram of the area of the genu in NC and AD scans with the use of an affine registration. (b) Histogram of the area of the isthmus in NC and AD scans with the use of an affine registration.
6.4.5 Global & Regional shape and area changes during Multiple Sclerosis

As it can be observed in Tables 6.9 and 6.10 the tests that were performed in the total area and circularity of the CC yielded statistically significant differences ($p \ll 0.01$), both for the affine (Figure 6.10) and the similarity registration (Figure 6.11). Moreover, for the area of the compartments of the CC, the tests yielded also statistical differences in all the compartments, both for the affine and the similarity registration, with the more pronounced differences found in the anterior body and the mid-body of the CC. The histograms of the areas found in these two compartments can be seen in Figures 6.12 and 6.13. As it can be seen, the differences with the use of a similarity registration are more pronounced and in both cases the mean area of the MS patients is lower than the mean area of the NC. Due to the large sample difference between MS patients and healthy controls, in order to illustrate more accurately the distribution of the MS patients in the histograms, the sample of the MS patients was multiplied by 10.

Finally, the similarity registration yielded statistical differences only for the second $b$ value (Figure 6.15), while in the affine registration statistical differences were found in all the $b$ values (Figure 6.14). In order to explain these differences in the $b$ values,
the first three modes of variation were plotted (Figure 6.16).

Table 6.9: Mean and Standard deviation of the CC and its compartments for NC and MS with the use of an affine registration

<table>
<thead>
<tr>
<th>Affine Reg</th>
<th>Mean NC / AD</th>
<th>Std NC / AD</th>
<th>Norm test NC/AD</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total CC Area</td>
<td>737.69 / 622.02</td>
<td>74.94 / 103.74</td>
<td>0.011 / 0.59</td>
<td>5.59 · 10⁻¹⁶</td>
</tr>
<tr>
<td>Total CIR</td>
<td>0.188 / 0.160</td>
<td>0.0154 / 0.028</td>
<td>3.5 · 10⁻¹² / 0.71</td>
<td>1.76 · 10⁻¹⁴</td>
</tr>
<tr>
<td>Genu</td>
<td>308.29 / 263.11</td>
<td>39.37 / 46.52</td>
<td>1.53 · 10⁻⁶ / 0.47</td>
<td>1.1 · 10⁻¹²</td>
</tr>
<tr>
<td>Anterior Body</td>
<td>87.18 / 71.75</td>
<td>9.91 / 10.93</td>
<td>6.61 · 10⁻⁶ / 0.18</td>
<td>6.53 · 10⁻²⁰</td>
</tr>
<tr>
<td>Mid Body</td>
<td>70.94 / 59.23</td>
<td>8.04 / 9.24</td>
<td>3.57 · 10⁻⁷ / 1.52 · 10⁻¹³</td>
<td>5.17 · 10⁻¹⁸</td>
</tr>
<tr>
<td>Isthmus</td>
<td>69.66 / 55.12</td>
<td>17.69 / 19.33</td>
<td>7.35 · 10⁻⁸ / 0.45</td>
<td>1.51 · 10⁻¹⁰</td>
</tr>
<tr>
<td>Splenium</td>
<td>201.63 / 172.81</td>
<td>33.17 / 37.76</td>
<td>0.001 / 1.15 · 10⁻⁶</td>
<td>1.77 · 10⁻⁵</td>
</tr>
<tr>
<td>$b_1$</td>
<td>-110.91 / -57.76</td>
<td>102.35 / 132.46</td>
<td>7.57 · 10⁻⁸ / 0.036</td>
<td>0.0039</td>
</tr>
<tr>
<td>$b_2$</td>
<td>-84.55 / -44.55</td>
<td>82.79 / 89.36</td>
<td>2.38 · 10⁻¹⁶ / 0.17</td>
<td>0.001</td>
</tr>
<tr>
<td>$b_3$</td>
<td>46.19 / 80.57</td>
<td>96.29 / 80.47</td>
<td>4.72 · 10⁻⁸ / 1.75 · 10⁻⁹</td>
<td>9.57 · 10⁻¹⁵</td>
</tr>
</tbody>
</table>

Table 6.10: Mean and Standard deviation of the CC and its compartments for NC and MS with the use of a similarity registration

<table>
<thead>
<tr>
<th>Similarity Reg</th>
<th>Mean NC / MS</th>
<th>Std NC / MS</th>
<th>Norm test NC/MS</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total CC Area</td>
<td>688.62 / 611.54</td>
<td>78.68 / 108.79</td>
<td>8.33 · 10⁻⁸ / 0.29</td>
<td>5.32 · 10⁻⁹</td>
</tr>
<tr>
<td>Total CIR</td>
<td>0.179 / 0.157</td>
<td>0.021 / 0.026</td>
<td>3.1 · 10⁻⁶ / 0.65</td>
<td>6.67 · 10⁻¹²</td>
</tr>
<tr>
<td>Genu</td>
<td>293.77 / 269.06</td>
<td>38.71 / 53.97</td>
<td>6.97 · 10⁻⁶ / 0.40</td>
<td>9.61 · 10⁻⁶</td>
</tr>
<tr>
<td>Anterior Body</td>
<td>83.20 / 69.63</td>
<td>11.37 / 11.16</td>
<td>5.71 · 10⁻¹¹ / 0.036</td>
<td>1.61 · 10⁻¹⁷</td>
</tr>
<tr>
<td>Mid Body</td>
<td>66.95 / 58.53</td>
<td>8.96 / 10.59</td>
<td>2.38 · 10⁻¹⁶ / 0.17</td>
<td>6.32 · 10⁻¹¹</td>
</tr>
<tr>
<td>Isthmus</td>
<td>77.40 / 65.39</td>
<td>26.66 / 33.78</td>
<td>4.72 · 10⁻³⁷ / 1.75 · 10⁻⁹</td>
<td>0.003</td>
</tr>
<tr>
<td>Splenium</td>
<td>167.29 / 148.93</td>
<td>48.05 / 47.19</td>
<td>1.51 · 10⁻⁰⁷ / 0.079</td>
<td>0.001</td>
</tr>
<tr>
<td>$b_1$</td>
<td>-82.57 / -62.10</td>
<td>138.51 / 131.89</td>
<td>2.2 · 10⁻²⁵ / 0.02</td>
<td>0.19</td>
</tr>
<tr>
<td>$b_2$</td>
<td>-94.73 / -8.72</td>
<td>114.38 / 109.72</td>
<td>2.2 · 10⁻³⁶ / 0.39</td>
<td>9.57 · 10⁻¹⁵</td>
</tr>
<tr>
<td>$b_3$</td>
<td>20.27 / -4.76</td>
<td>93.47 / 89.56</td>
<td>0.005 / 0.21</td>
<td>0.067</td>
</tr>
</tbody>
</table>

Figure 6.10: (a) Histogram of the total area of the CC in NC and MS scans with the use of an affine registration. (b) Histogram of the total area of the CC in NC and MS scans with the use of an affine registration.
Figure 6.11: (a) Histogram of the area of the CC in NC and MS scans with the use of a similarity registration. (b) Histogram of the total area of the CC in NC and MS scans with the use of a similarity registration.

Figure 6.12: (a) Histogram of the area of the anterior body in NC and MS scans with the use of an affine registration. (b) Histogram of the area of the mid body in NC and MS scans with the use of an affine registration.

Figure 6.13: (a) Histogram of the area of the anterior body in NC and MS scans with the use of a similarity registration. (b) Histogram of the area of the mid body in NC and MS scans with the use of a similarity registration.
Figure 6.14: (a) Histogram of the first $b$ value of the segmented scans of NC and MS scans with the use of affine registration. (b) Histogram of the third $b$ value of the segmented scans of NC and MS scans with the use of an affine registration.

Figure 6.15: (a) Histogram of the second $b$ value of the segmented scans of NC and MS scans with the use of an affine registration. (b) Histogram of the second $b$ value of the segmented scans of NC and MS scans with the use of a similarity registration.

Figure 6.16: Modes of variation of the ASM model of the CC
6.5 Conclusions

In this Chapter, the potential of the CC to be used as a biomarker for Alzheimer’s Disease and Multiple Sclerosis was investigated. From Figures 6.1 and 6.2, it can be concluded that the Circularity can be used as a stable measure, to observe changes on the CC with ageing, as it consistently decreases with age. Another observation is that the circularity values of the scans registered with the affine registration are higher than the ones registered with the similarity registration. This can be explained due to the affine registration performed with Nifty Reg, which includes shearing effects that ‘squeeze’ the CC. The arc length of the squeezed CC becomes lower, while its area decreases much more, yielding a lower circularity.

Moreover, as far as the separation of NC and AD patients is concerned, the use of the proposed features in Section 6.4.2 with an SVM yielded high accuracy results (∼96%) especially with the use of an affine registration for the spatial normalisation of the data (Table 6.1). Due to the fact that the performance of an SVM can change depending on its parameterisation and the data, the ROC curve of the LDA dimension was plotted as well, which is more constant (Figure 6.3). If we define an acceptable clinical error (false positive rate) of the order of 5% (Figure 6.3a) for the affine registration, around 50% of the scans would be classified correctly, which is rather low. On the other hand, for the separability of the MS patients, the accuracy of the SVM was similar for the affine and similarity registrations (∼95%), but the similarity registration yielded a higher specificity (Table 6.4). The higher specificity is important due to the fact that if a scan is estimated as MS, then it is extremely probable to be MS. In addition, from the ROC curve of the similarity registration (Figure 6.4b), if we define a clinical error of 5%, then around 80% of the patients would be correctly classified. Furthermore, the highest contribution for the separation between normal controls and diseased patients stemmed from the \( b \) values of the model. In addition, the rest of the \( b \) values, which describe small variations contribute a lot for the separation of the MS patients and not so much for the AD patients. From this analysis, it can be seen that the area and shape changes of the CC can be potentially used for the separation between healthy and diseased patients. An analysis between the area and shape differences of AD and MS was not performed due to the fact that MS is a disease, which affects people of young age, while AD is appeared in people aged above 55 years old.

Furthermore, as far as the AD patients are concerned, statistical differences were found in all compartments, apart from the splenium for the similarity registration, with the more pronounced ones in the first and fourth compartments (genu and isthmus) of the CC. Based on recent studies, a reduced area was found in the genu and splenium [4, 12, 13], while in our case the genu was found as well, but more pronounced area changes were found in the isthmus than the splenium. Moreover, statistical differences were found in the total area and circularity of the CC and they were higher, when the scans were spatially normalised with a similarity registration. In addition, statistical significant differences were also observed in the 2nd \( b \) value of the model and as it can be seen in Figure 6.9b the normal controls tend to have a mean 2nd \( b \) value around zero, while the AD patients around -100. This is translated from Figure 6.16, as that the older normal controls tend to have a shape close to the mean shape of our model which means very small variations, while the AD patients tend to have an elongation in the CC shape towards the horizontal direction.

For the MS patients, statistical differences were found for the total area and circularity of the CC (Figures 6.10 and 6.11), as well as for the area in all the compartments of the CC. The more pronounced differences were observed in the 2nd and 3rd com-
partment of the CC (anterior and mid body), which are responsible for the primary motor and somatosensory cortex. This is a reasonable outcome, as the MS patients are more probable to have movement problems. Moreover, statistical differences were found only for the second \( b \) value for the similarity registration and for all the \( b \) values for the affine registration. From Figure 6.15b, it can be observed that the \( b \) values of the young healthy controls tend to have values below zero, while the MS patients tend to have positive values. This can be translated as that in younger ages the CC of the normal controls tend to be elongated in the horizontal direction, while for MS patients the CC tends to elongate along the transversal direction and the CC tips become less distinguishable (Figure 6.16).

To sum up, the changes of the area and the shape of the CC, could be potentially used as a biomarker for Alzheimer’s Disease and Multiple Sclerosis, as it can yield high classification accuracy results. Further investigation towards that direction is needed and studies which include a large number of MS patients need to be performed.
Chapter 7

Conclusions & Suggestions for Future Work

7.1 Conclusions

This thesis has investigated the possibility of using the Corpus Callosum as a biomarker for Alzheimer’s Disease and Multiple Sclerosis. As a first step, a pipeline was proposed and validated for the extraction of the Corpus Callosum from T1 weighed MR brain images. The implementation as it was seen in Chapter 4 yielded high accuracy segmentation results with a mean dice of 93% (Figure 4.15) and high segmentation reproducibility results that yielded a mean error of 1.93% (Figure 4.12). Moreover, with the use of thresholding the lesions that can be found on the CC in MS patients were excluded. Furthermore, with the use of the ASM model the shape variation of the CC was captured and from the b values of the model it was observed that the ASM can accurately segment scans of the same patient at different time points (Figure 4.8).

Using this accurate model we tried to investigate potential differences between scans of a subject, that were acquired in scanners of different vendors and in scanners of different magnetic field strengths. While the segmentation reproducibility within vendors was quite low (1.5% - 2%), the segmentation reproducibility in scans between vendors and between different magnetic fields (7%) was rather high. These differences can be induced due to the fact that the algorithm might not be robust to changes in contrast between vendors, or that there exist real acquisition differences between the vendors that can be mainly attributed to the differences in the reconstruction algorithms used from every vendor to construct the image.

As far as the changes of the CC with ageing are concerned, the circularity can be used as a more stable measure to observe changes of the CC with ageing. Moreover, the mid part of the CC turns to elongate vertically with ageing, as the 2nd b value of the model from negative in young ages (Figure 6.15b) turns to concentrate around zero (Figure 6.9b) in the older ages. Furthermore, it was shown that the area and shape changes of the CC can be used in order to separate healthy from diseased patients with an accuracy of above 95%. In addition, the genu and isthmus yielded the most pronounced statistical differences between NC and AD people, while the anterior body and the mid-body had the most pronounced differences for MS patients. The results for the AD group verified the results of other studies for the compartment of the genu and the results for the MS group were logical as the anterior and mid body are responsi-
ble for the motor cortex, which is mostly affected by MS, verifying this MS symptom. Additionally, the CC of the AD patients tends to be longer towards the horizontal direction than NC volunteers, while for MS patients the CC tends to stretch towards the vertical direction. A separation between AD and MS patients was not attempted, as the age interval in which these diseases affect people is different.

Overall, the area and shape changes on the CC could be used as a biomarker for the prediction of Alzheimer’s Disease and Multiple Sclerosis. Although, our current implementation yielded high segmentation accuracy and reproducibility results, the area of the CC might not be able yet to be used for clinical practice, as the percentage changes in the area of the Corpus Callosum that we look for in order to be used as a biomarker, are of the magnitude of 2% ([48]). On the other hand, the area in addition with the shape changes of the CC might be able to be used, as the classification accuracy based on them was high enough. Finally, in order to draw clinical conclusions about the changes of the CC during the progression of a neurodegenerative disease, the data that are used should have been acquired in the same scanner, as the acquisition in different scanners can yield misleading results.

7.2 Suggestions for Future Work

Some improvements can be suggested for the better construction of the model. First of all, a more accurate way of extracting the mid sagittal plane needs to be used than the registration of each scan in an MNI atlas. Moreover, the 2D model can be used for segmenting the CC in adjacent slices as well, maximum 3 slices in each direction of the mid sagittal plane (7 slices in total) and in this way the computation of the area of the CC can be extrapolated to 3D. These improvements can further improve the segmentation accuracy and reproducibility of the algorithm, which might be later consist it usable for clinical practice. Finally, a further investigation is needed for if the shape changes of the CC can be used for discovering the progression of a neurodegenerative disease based on the $b$ values of the ASM model. For this reason, longitudinal studies need to take place.
Appendix A

Partitioning of the Corpus Callosum

After the 2D segmentation of the CC from the mid sagittal plane, the CC is partitioned into meaningful compartments. Due to the fact that different kinds of fibers pass through the CC, which can undergo different changes during neurodegeneration (e.g. genu and splenium in AD), it is interesting to investigate the changes in the different compartments of the CC and not only the CC as a whole. The main schemes for the segmentation of the corpus callosum into meaningful compartments are the following ones:

1. Witelson Partitioning
2. Radial Partitioning
3. Medial axis Partitioning

Witelson et al. [50] were the first ones to try to partition the CC into meaningful compartments. They tried to achieve a lobar parcellation of the corpus callosum, so they separated it into five main compartments along its first principal axis and the four points that were needed were in distances 1/3, 1/2, 2/3, 4/5 from the most anterior point of the structure (Figure A.1). The first compartment was further subdivided into three compartments and the names of the compartments are the following ones:

1. Rostrum
2. Genu
3. Rostral Body
4. Anterior midbody
5. Posterior midbody
6. Isthmus
7. Splenium

The next partitioning scheme is the radial partitioning, in which the CC is split into equiradial areas [38]. In order to find the center of gravity for the split up of the CC, the first principal axis is translated till it splits the CC into fractions containing 95% of the superior and 5% of the inferior of the total area of the CC.
The last partitioning theme is the medial axis partitioning, in which the chordal axis transform is used to find the medial axis of the CC. After the medial axis is found the CC is partitioned based on its arc length.

The medial axis partitioning is a method, which is unbiased to the shape differences of the CC. Nevertheless, Wittelson partitioning is the scheme that is most commonly used in the studies, as it is easy to implement and gives similar results to the medial axis partitioning.
Bibliography


[38] Charlotte Ryberg. Corpus callosum partitioning schemes and their effect on callosal morphometry. In 14th Scientific Meeting and Exhibition of International Society for Magnetic Resonance in Medicine, 2006.


