Automated shape modeling and analysis of brain ventricles : findings in the spectrum from normal cognition to Alzheimer disease

Ferrarini, L.

AUTOMATED SHAPE MODELING AND ANALYSIS OF BRAIN VENTRICLES:

FINDINGS IN THE SPECTRUM FROM NORMAL COGNITION TO ALZHEIMER DISEASE

Luca Ferrarini

Colophone

About the cover
Alzheimer Disease slowly and irreversibly changes the perception of reality: the three watches on the front cover represent the decline from normal cognition to a demented state, in which time and space are no longer the same. Moreover, the change in shape of the watches recalls the shape changes observable in the brain ventricles due to atrophy in periventricular structures. The back cover shows axial views of the brain ventricles in an healthy subject.

Automated Shape Modeling and Analysis of Brain Ventricles: findings in the spectrum from normal cognition to Alzheimer Disease

Ferrarini, Luca

Printed by Ponsen & Looijen b.v., The Netherlands

ISBN: 978-90-6464-194-7

AUTOMATED SHAPE MODELING AND ANALYSIS OF BRAIN VENTRICLES:
FINDINGS IN THE SPECTRUM FROM NORMAL COGNITION TO ALZHEIMER DISEASE

GEAUTOMATISEERD VORM MODELLERING EN ANALYSE VAN HERSENVENTRIKELS:
BEVINDINGEN IN HET SPECTRUM VAN NORMALE COGNITIE AAN ALZHEIMER DISEASE

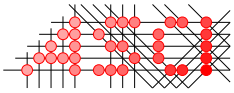
Proefschrift
ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus prof. mr. dr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 19 maart 2008
klokke 13.45 uur

door

Luca Ferrarini
geboren te Carpi (Italië)
in 1978

Promotiecommissie

Promotor:	Prof. dr. ir. J.H.C. Reiber
Co-promotor:	Dr. F. Admiraal-Behloul
Referent:	Prof. dr. D. Rueckert
	Imperial College, London
Overige leden:	Prof. dr. M.A. van Buchem, M.D.
	Prof. dr. W. Niessen
	Erasmus Medical Center, Rotterdam



This work was carried out in the ASCI graduate school.
ASCI dissertation series number 156.

The research reported in this manuscript was financially supported by the Dutch Technology Foundation (STW), under the research grant number 06122.

Financial support for the publication of this thesis was kindly provided by:

- Medis medical imaging systems bv., Leiden
- Stichting Beeldverwerking, Leiden
- Foundation Imago, Oegstgeest
- Internationale Stichting Alzheimer Onderzoek (ISAO)
- Bio-Imaging Technologies bv., Leiden

Contents

Colored Figures from Chapter 4 and Chapter 6

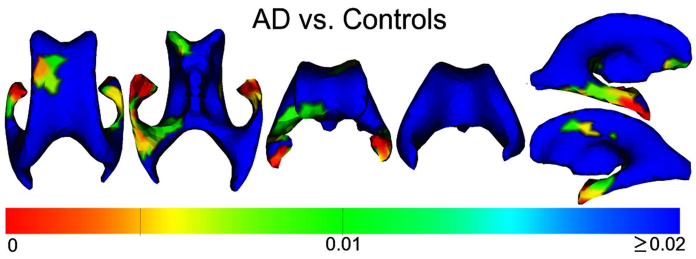


Figure 4.5 (page 71) The color-coded map shows, for each location, the *median* p value evaluated over the N_{iter} runs ($p > 0.01$ are shown in blue).

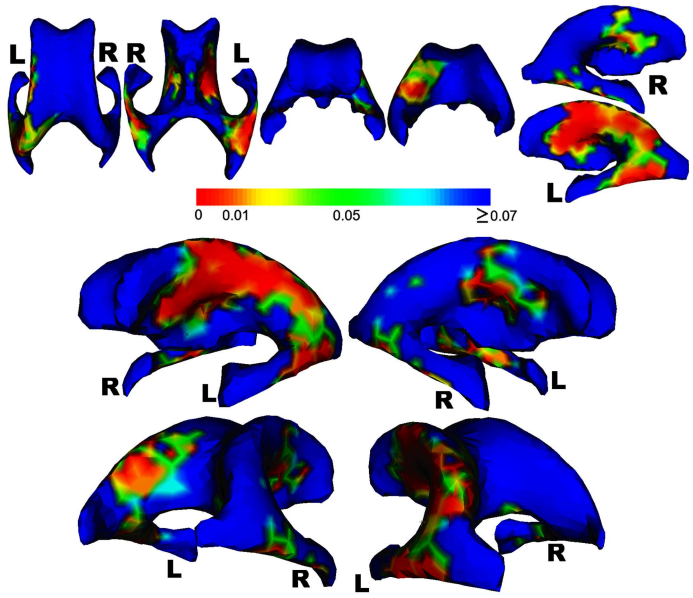


Figure 6.3 (page 103) Local correlation between severity of atrophy and cognitive impairment, across all subjects (controls and patients). Orthogonal (*top*) and prospective (*bottom*) views (p-values higher than 0.05 are plotted in blue).

CHAPTER 1

Introduction

He who cannot draw on three thousand years is living from hand to mouth.

Johann Wolfgang von Goethe, 1749-1832.

1.1 A Stroll along the History of Medical Imaging

The history of modern western medicine[1] goes back in time for more than two millennia, revealing the story of an eternal struggle against diseases. Across the centuries, we have come to realize how essential the observation of the human anatomy is: *medical imaging* as we know it today is the result of such a process.

Before the *Ancient Greek* period, medicine and religion were strongly intertwined: diseases were considered as punishments from the Gods and only curable with religious remedies. It was only with Thales (ca. 624-546 BC) that humans turned their attention away from the Gods and hypothesized the existence of four elements governing the entire Universe: air, water, earth and fire. In the mid-fifth century BC, Alcmaeon of Crotone suggested that diseases were simply a consequence of disequilibrium between such elements, and that among the organs the brain was the most important. His ideas were later turned into a fundamental philosophy by Hippocrates (ca. 460-370 BC), who is commonly recognized as the father of rational medicine. Unfortunately, following the principles of the four elements, philosophers disregarded the observation of the human body as a useless practice: diseases could only be cured by expelling the elements in excess. Nevertheless, the investigation of the human anatomy survived and developed thanks to painters and sculptors, constantly aiming at a perfect representation of the human body.

While the *Ancient Greek* period was turning to an end, it witnessed the life of its greatest philosopher and biologist, Aristotle (384-322 BC). Among his many contributions, he elaborated a physiological system to describe the functionality of the human body. His teachings were passed on to his students, among who was Alexander the Great (356-323 BC), whose conquests in the Mediterranean area facilitated the spreading of Aristotle's philosophy through the *Hellenistic and Roman* era. Alexandria, in Egypt, soon became the center of the epoch's knowledge: medical philosophers trained in Alexandria rediscovered the art of dissection, thanks to the pluri-millenary tradition of mummification. It was within the school of Alexandria that Democrito (ca. 460 BC) developed his atomistic theory, contrasting the widespread philosophy of the four elements. A few centuries later, Alexandria hosted a Greek physician whose theories were bound to rule the medical domain for more than a millennium: Galen of Pergamum (ca. 129-200 AD). After completing his education in Alexandria, Galen moved to Rome where he worked as doctor of gladiators: he performed the first brain and eye surgical operations, acquired a profound knowledge of wound treatment, and promoted first-hand observations of the human anatomy based on dissections and vivisections. Unfortunately, he also contributed to the theory of humours, connected with the elements of Hippocrates, leading to severe limitations in disease therapy. Despite the valuable efforts of many in anatomical studies, knowledge was still passed on through practical experience, and there is no evidence of anatomical drawings in this period.

In the following centuries (ca. 476 AD), the Roman Empire fell, and its knowledge was slowly transferred to the Byzantine Empire: influenced by the Koran, physicians once again took distance from the exploration of the inner human anatomy: any cut inflicted on a human being could lead to his soul departure. The study of the human body would have to undertake

[1] The historical information reported in this section were collected from the lectures on The History of Medicine by Prof. Riva (Faculty of Medicine, University of Cagliari), available online at http://pacs.unica.it/biblio/history.htm.

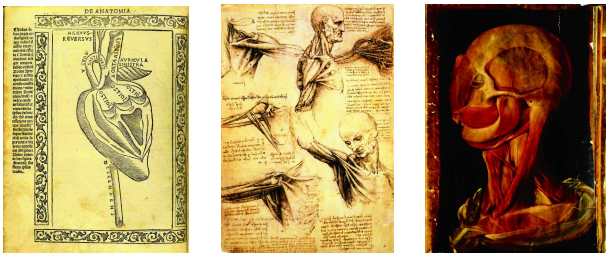


Figure 1.1: From left to right: Anatomical Sketches by Berengario da Carpi (1535), Leonardo Da Vinci (1680), and Fabricius ab Acquapendente (1533-1619).

a long trip back to its source: passing first through the Moorish civilization in Spain, and then finally returning to France and Southern Italy around the year 1000 AD.

The first Universities were founded around 1200 AD, the one in Bologna (Italy) being the oldest in Europe. The first hospitals were built in the same period, and with them the practice of dissection returned. Although medicine was considered an academic field, the link between Universities and Hospitals was not formalized. The first University Clinic was founded only four centuries later (around 1600 AD), in Leiden (The Netherlands). By this time, the importance of reproducing images of the human body was well established: it all happened during the *Renaissance*. The prototype of the *Renaissance man* embraced knowledge from the most diverse fields, such as arts, anatomy, engineering, etc., with Leonardo Da Vinci (1452-1519) being one of the most representative figures of this period. The finest arts had to reproduce nature in its very details, and required a profound knowledge of the human anatomy: Berengario from Carpi (1460-1530, Carpi-Italy) published the first anatomical book with illustrations [1], emphasizing the beauty behind the discovery of the human body; Fabricius ab Acquapendente (1533-1619) introduced colored illustrations (see Fig. 1.1), and Vesalius (1514-1564) published *De Humani Corporis Fabrica* in 1543, a collection of anatomical views in woodcuts [2].

By the beginning of the XVII century, the valuable contribution of anatomical drawings was largely recognized. The description of the human anatomy via illustrations became essential in medical teaching. The idea of a scientific approach to research developed across the

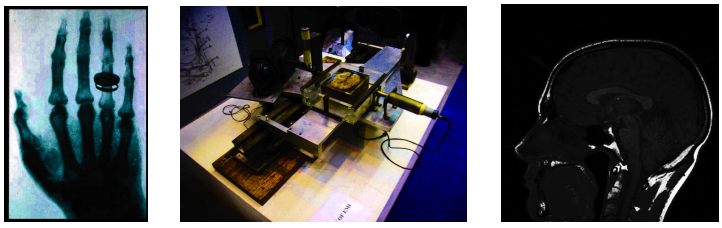
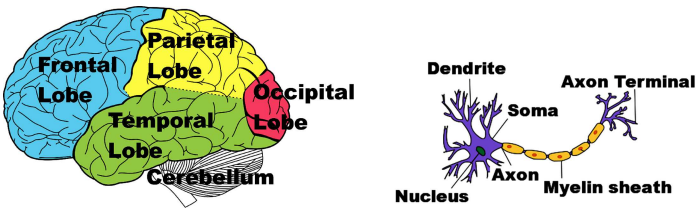


Figure 1.2: From left to right: The first X-ray image, taken by Roentgen in 1896; The first CT scanner prototype designed by Houndsfield; Sagittal MR image of the brain.



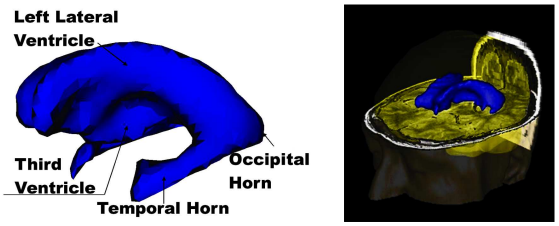
(a) (b)

Figure 1.3: (a) Lobes in the brain. (b) Neural cells: the different nuclei form the gray matter, while the axons bundle together in what is known as the white matter.

XVI-XVII centuries, thanks to Galileo Galilei: reproducibility of experiments and first-hand experiences through senses constituted the basis of such a revolution. The microscope was invented, laying the foundations for the cellular pathology theory which would lead, two centuries later, to the use of colorants for lesion detection. Galileo's background as a physicist and the invention of the microscope originated a new wave of thinking: the human body was considered as a machine, made of several parts (the organs) which needed to be analyzed separately.

Medical imaging as we know it today moved its first steps in the XIX century: the use of X-rays for imaging the human body was discovered in 1895 by Wilhelm Konrad Roentgen, marking the beginning of radiology [3] (see Fig. 1.2). Afterwards, new techniques were developed: ultra-sounds for echography (1940, Dr. George Ludwig [4]), computerized tomography (CT, by Godfrey Hounsfield, 1967 [5]), nuclear medical imaging (first introduced in 1950), and magnetic resonance imaging (MRI, first demonstrated in 1973 by Paul Lauterbur [6]). Nowadays, thanks to increased computational power and advanced technologies, efforts are made to merge existent imaging modalities.

Our stroll started when the human body was nothing but a unique entity at the mercy of Gods; walking through the centuries, we have come to value and increase our knowledge in the anatomy and functionalities of different organs and cells; nowadays, we feel the need to glue pieces together in a better understanding of how they interact within the whole system.



(a) (b)

Figure 1.4: The ventricles (a) are cavities filled with cerebrospinal fluid; located in the middle of the brain (b), they serve as a cushion protecting the brain from concussions.

Medical imaging has developed along this path, and will doubtlessly continue consolidating its role.

1.2 Brain Anatomy in Magnetic Resonance Images

Located in the head and protected by the skull, the brain is probably the most complex organ in the human body: a quote often attributed to E.M. Pugh states that *if the brain were so simple that we could understand it, we would be too simple to understand it.* Whether this paradox corresponds to truth or not is let to science to debate. Certainly, a thorough description of the brain and its functions goes behind the scope of this manuscript: this section provides a general overview of the brain anatomy, and introduces the structures involved in this research.

Functional areas and biological background
The brain can be divided in different functional areas (see Fig. 1.3.a): the *frontal lobe* plays an important role in many cognitive functions like judgment, impulsive control, language, and planning; the *temporal lobe* is predominantly involved in the auditory processing, speech and vision semantic, and memory formation; the *occipital lobe*, receiving information from the retina, is actively involved in visual processing, color discrimination, and motion perception; *parietal lobes* integrate the sensory information perceived from different parts of the human body. The different lobes form what is usually referred to as cerebral cortex. Finally, the *cerebellum* plays an important role in the integration of sensory perception and motor output and is connected to the motor cortex. From the biological point of view, the brain is composed of cells: a distinction can be made between the *gray matter*, consisting of neuron's bodies, and *white matter*, consisting of axons connecting neurons (see Fig. 1.3.b). The complexity of the brain can be appreciated considering that it is made of more than 100 billion neurons, each connected to approximately 10.000 other neurons.

The Brain Ventricles - The brain ventricles are a system of cavities located in the center of the brain and filled with cerebrospinal fluid (CSF) (see Fig. 1.4). The CSF serves as a cushion for the brain, protecting it from concussions. The brain ventricles are surrounded by white and gray matter structures, usually defined as *periventricular structures*.

Corpus Callosum - The corpus callosum is the largest bundle of white matter in the brain; located above the brain ventricles, it allows communication between the left and right hemispheres (Fig. 1.5.a).

Caudate Nuclei - The caudate nuclei (one for each hemisphere) are located within the basal ganglia (a set of nuclei in the brain connected with several cortical and subcortical structures) and play an important role in learning. Their position relative to the brain ventricles is shown in Fig. 1.5.b.

Thalamus - The thalamus is located, symmetrically, in the proximity of the third ventricle (see Fig. 1.5.c): among many other functions, it is responsible to encode and transmit pre-

thalamic inputs to different areas in the cortex.

Hippocampus - The hippocampus is part of the limbic system, and is located in the medial temporal lobe (see Fig. 1.5.d): it has an important role in memory and spatial navigation.

Amygdala - Also part of the limbic system (see Fig. 1.5.b), these almond-shaped groups of neurons are involved in processing and memory of emotional reactions.

1.2.1 MR acquisition

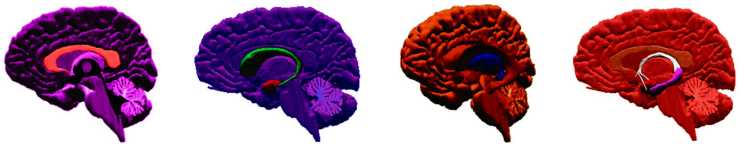
Neuroscientists can rely on several imaging techniques to investigate brain's anatomical and functional properties, both in clinical practice and in research. Imaging modalities differ for the type of information they provide: Multi Slice Computed Tomography (MSCT) and Magnetic Resonance Imaging (MRI) mainly focus on structural properties, while nuclear based techniques such as Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT) are more suitable for imaging the brain's functional properties. The research described in this manuscript is based on clinical MR images, whose main characteristics are discussed in this section.

Magnetic Resonance Imaging was developed in the 1970s. Hydrogen nuclei present in water and lipid are properly excited by radio frequencies waves in a strong static magnetic field: the different relaxation properties of the nuclei can be weighted to generate images with different intensity contrast. Important properties are the *longitudinal relaxation time* T1 and the *transverse relaxation time* T2 (or the $T2^*$ variation); finally, proton-densities (PD) images are acquired without considering the relaxation time. The MR technique presents several advantages:

- images can be acquired at any plane in 3D;
- T1, T2, $T2^*$ properties can be weighted to provide different intensity contrasts;
- the patient is not exposed to radiation;
- contrast agents (like gadolinium) used for image enhancement have proved safer than those used for X-rays and CT;
- aside for possible contrast administration, it is an entirely noninvasive technique.

The main drawback of MR is related to the applied magnetic field: patients with ferromagnetic implants are generally not allowed in MR scanners.

In clinical settings, images can be acquired with an approximate isotropic resolution of $1mm^3$, applying a magnetic field of 1.5 Tesla: higher magnetic field (i.e. 3 Tesla) have also become part of clinical routine, improving the spatial resolution. Depending on the chosen protocol, brain tissues, such as gray and white matter, and CSF appear with different contrasts, as shown in Fig. 1.6.



(a) (b) (c) (d)

Figure 1.5: (a) Corpus Callosum (in white); (b) The caudate nuclei (in green) and the amygdala (almond-shaped red blob); (c) The Thalamus (in blue); (d) The hippocampus (in violet).

1.3 The Alzheimer's Disease

It was in 1901, when doctor Aloysius Alzheimer was presented with a case of a 51 year old woman with cognitive and language deficits, and a loss of short-term memory. Alzheimer followed the case, and when the woman died five years later, he was given her brain for further examinations [7]. By means of staining techniques, Alzheimer and his team could highlight the presence of amyloid plaques and neurofibrillary tangles in the brain. In a book published in 1910, Emil Kraepelin used for the first time the term *Alzheimer's Disease* to refer to the pathology [7]. Since then, the scientific community has spent considerable efforts studying the Alzheimer's Disease (AD), trying to unveil the mechanisms behind its development, introducing diagnostic tests, trying to detect biomarkers, and developing therapies.

The reasons behind such conspicuous efforts are to be found in the syndrome caused by the disease: *dementia*. Defined as a progressive process of cognitive decline, dementia can be caused by diseases or damages affecting the brain. The risk of developing dementia strongly correlates with age. Fratiglioni et al. [8] showed that the percentage of people affected by dementia almost doubles every five years beyond the age of 65: from 1% in the age range 60-64, up to 45% above 95. More recently, Wimo et al. [9] estimated that the societal costs of health care for dementia in 2005 reached 315.4 billion dollars worldwide, of which more than 72% was spent in developed continents (26.5% in North America, 28.1% in Asia, and 38.2% in Europe). Alzheimer Disease is the most common cause of dementia, accounting for about 74% of the cases in North America (where AD was the 7^{th} leading cause of death in

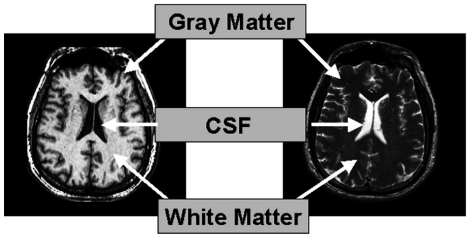
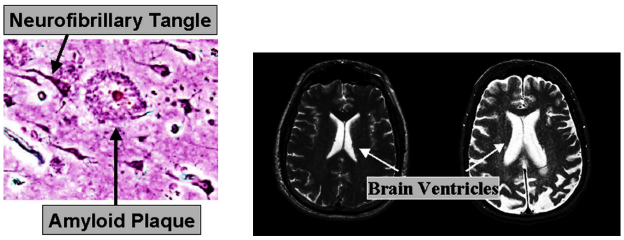


Figure 1.6: MR T1-weighted (left) and T2-weighted (right) axial views of the brain. In T1w images, gray matter is in gray, white matter in white, and CSF in black. In T2w images, the colors are inverted.



(a) (b)

Figure 1.7: (a) Amyloid plaques and neurofibrillary tangles in AD; (b) MR T2w images of brain ventricles in a healthy subject (*left*) and in AD (*right*): the ventricles are visibly enlarged in AD, due to atrophy.

2004), 61% in Europe, and 46% in Asia[2]. Considering the high rate of conversion to AD, the economical and social effects are bound to increase considerably in the future, if no remedy is found. Emotionally, AD affects the life of the patient and his family drastically. The initial loss of short-term memory is followed by a progressive cognitive decline: in its final and most severe stage, the patient is unable to perform any simple task and needs constant supervision. On average, the disease leads to death in about 7-10 years, and up to 20 years after the onset of the symptoms.

Nowadays, no definitive cure exists for AD: available therapies can only slow down the cognitive decline, and have been proved more effective when given in the early stage of the disease [10]. Therefore, it is vital to improve our understanding of how the disease develops, and to identify biomarkers sensitive to brain changes related to AD progression. The most accredited hypotheses seek the causes of the disease in the pathology of amyloid and tau proteins, which accumulate to form amyloid plaques and neurofibrillary tangles (see Fig. 1.7.a). Consequently, neurons' apoptosis and neuronal synapses loss spread in different regions of the brain, causing drastic anatomical changes (i.e. atrophy) which can be highlighted with imaging techniques. Many studies have shown a correlation between the progression of the disease and the decrease of volume in periventricular structures, such as the hippocampus and the amygdala [11–18]: a direct consequence is the enlargement of the brain ventricles [19], clearly visible in MR images (see Fig. 1.7.b).

Alzheimer's disease is at the end of a spectrum ranging from normal cognition to full dementia. An important intermediate phase is the one referred to as Mild Cognitive Impairment (MCI): patients with MCI present mild forms of memory loss, but are otherwise cognitively capable. It has been shown that MCI subjects can eventually degenerate into AD [20]: nevertheless, since not all MCIs turn into ADs [21, 22], defining MCI as an intermediate step towards dementia would be far too reductive. Because of its connections with AD, and because of its cognitive implications, MCI has also been thoroughly investigated.

A definitive diagnosis of AD is possible only post-mortem, via microscopic examination. Expert clinicians can diagnose AD with an accuracy of 85-90%, but the final decision

[2]Information available from the European Commission Report on AD, *http://ec.europa.eu/health/ph_information/dissemination/diseases/alzheimer_en.htm*

is based more on symptoms and on the absence of alternative explanations, rather than on specific biomarkers. Nuclear imaging techniques such as PET and SPECT might be useful to diagnose dementia [23], but are not yet sufficiently specific in differentiating the causes behind the syndrome. Techniques based on structural MR images have been developed to estimate volumetric brain changes in AD: Boundary Shift Integral (BSI, [11]), Structural Image Evaluation using Normalization of Atrophy (SIENA, [24]), and Jacobian Integration [25]. Comparisons between these methods have consolidated the strong correlation between brain volume loss and AD progression [25, 26]. Local analyses, such as Voxel Based Morphometry and Deformation Based Morphometry, have also been used to highlight patterns of gray matter loss through the spectrum from MCI to AD [27]. Several tests have been developed to assess the cognitive functions of an individual. One of the most widely used is the Mini Mental State Examination (MMSE), introduced by Folstein in 1975 [28]: the test consists of several questions and tasks, and a maximum final score of 30 points indicates full cognitive capabilities.

Although imaging techniques and cognitive tests have contributed considerably to the study of AD, the quest for a final diagnostic tool is still ongoing. The goal of this research was to improve our knowledge on how MCI and AD affect periventricular structures in the brain, and how such effects correlate with cognitive decline. Measuring atrophy in periventricular structures is a challenging task: the delineation of such small regions in clinical MR images is time consuming and prone to errors. Moreover, it is not easy to predict beforehand which structures will provide the best biomarkers for the disease. We have tackled these issues by focusing on shape variations of the brain ventricles across different populations: brain ventricles are more easily detectable in clinical MR images, due to the high contrast of the CSF with the rest of the parenchyma; moreover, loss of volume in any periventricular structure will directly be reflected in localized shape changes along the ventricular surface.

1.4 Shape Modeling and Analysis

Variations in biological shapes can reasonably be described by statistical models over a large population of similar instances. The first attempt to formalize such approach was done by Cootes et al. [29]: developed initially for two-dimensional (2D) images, the shape modeling requires, as a first step, the establishing of corresponding locations through a large dataset. The correspondence, particularly in medical images, should be based on similar anatomical locations and/or similar position and shape characteristics: the final set of locations through the entire dataset is referred to as Point Distribution Model (PDM). Manual identification of corresponding landmarks might be possible in 2D, but its complexity increases drastically when 3D shapes are considered: the task is time-consuming, prone to error, and poorly reproducible. Thus, (semi-) automatic methods to generate PDMs are highly desirable. Several methods have been proposed in the literature.

Spherical Harmonics (SPHARM) - First introduced by Gerig et al. [30], SPHARMs model object shapes with spherical harmonics: the method is multi-scale and was successfully applied to the modeling of lateral ventricles. Nevertheless, the objects to be modeled require a

spherical topology. Moreover, the parameters used to model the objects are not intuitively related to the object's shape: thus, interpreting how changes in the parameters relate to changes in shape is not straightforward.

Medial Representation (M-reps) - M-reps were first introduced by Pizer et al. [31]. Complex objects are modeled at multi-scale as compositions of different figures: for each figure, a medial model is created and surface locations are parameterized relative to it. The method was successfully applied to the analysis of kidney and hippocampus (see Styner et al. [32]). Nevertheless, M-reps present some limitations: the definition of medial models can be either manual (time consuming and error prone), or automated [33]: in this case, an initial SPHARM-based parametrization is needed, which leads back to the requirement of a spherical topology. Moreover, to detect medial models in elongated and thin structures (like temporal horns) is a challenging task.

Minimum Description Length (MDL) - Rooted in an information theory framework, MDL (Davies et al. [34]) is based on descriptive functions which are optimized with genetic algorithms: first, a set of functions is used to identify corresponding locations; subsequently, a PDM is built and its performances are tested relative to several error functions; finally, the parameters defining the descriptive functions are tuned in order to maximize an objective function. Although the MDL has been successfully applied to several anatomical structures [35, 36], it still presents some limitations: the objective function is usually highly non-linear and the genetic algorithms used for optimization increase the computational load considerably, without guaranteeing an optimal solution.

Deformation Fields - Rueckert et al. [13] suggested a different approach to the analysis of shape variations. The statistical model is built on the deformation fields obtained by non-rigid registration: thus, the method does not require any pre-segmentation to isolate the objects of interest. Moreover, being based on the entire dataset, it provides information on inter-objects variability (e.g., in a brain dataset, changes of the brain ventricles are related to changes in the surrounding structures). Still, the authors consider shape-based approaches to be preferable when inter-object variability is not important, or when it might confound the analysis.

The ventricular system presents a very challenging shape: it is highly concave and therefore not easily reducible to a spherical topology. Moreover, the temporal and occipital horns of healthy subjects can be very thin, especially when imaged in clinical setups. Part of the research described in this thesis aimed at developing a new technique for shape modeling which could correctly represent shape variations in brain ventricles.

Regardless of the particular modeling technique, once the PDM has been acquired and validated, different shape analyses can be performed. Intra-population variability can be studied with Principal Component Analysis (PCA) [29]: eigen-analysis in the shape-space results in an average shape model and orthogonal directions (eigen-vectors) along which one can move to analyze modes of variations. Inter-populations variability can also be investigated: thanks to the correspondence property of PDMs, shapes in one population can locally be compared with shapes in another population via non-parametric statistical tools, such as

permutation tests [37]. The results presented in this thesis are based on both these kinds of analysis.

1.5 Structure of this thesis

This manuscript is organized as a collection of scientific papers: consequently, a certain degree of overlapping is to be expected through the most general parts of the following sections. The context and novelty of each chapter are described here: Chapter 2 presents a new technique to model the shape of challenging structures with comparable meshes. The new algorithm is thoroughly validated on challenging synthetic shapes, and applied to the modeling of two populations of brain ventricles: healthy subjects and ADs. Comparisons with other approaches show the good performances of our method. In Chapter 3, the new shape modeling algorithm is applied to two populations of AD and age-matched healthy individuals, showing that only well-defined areas of the ventricular surface differ significantly, and highlighting the corresponding periventricular structures. In Chapter 4, permutation tests and support vector machines are used to prove the existence of ventricular shape-based biomarkers for AD in clinical MR images, and to assess their performances. Differences between Memory Complainers, MCIs, and ADs are investigated in Chapter 5, focusing both on the severity and extent of atrophy in periventricular structures. Finally, in Chapter 6 the whole spectrum of cognitive impairment (as measured by the MMSE score) is correlated, via linear regression, with local enlargements in the brain ventricles. Chapters 7 and 8 summarize and conclude the manuscript.

Bibliography

[1] da Carpi B. Anatomia Carpi. 1535;.

[2] Vesalius A. De Humani Corporis Fabrica. 1543;.

[3] Roentgen WK. Über eine neue Art von Strahlen (On a new kind of rays). Sitzungsberichte der Würzburger Physikalischen-Medicinischen Gesellschaft. 1895;p. 132–141.

[4] Ludwig G, Struthers FW. Considerations underlying the use of Ultrasound to detect Gallstones and Foregn Bodies in Tissue. Nav Med Research Inst Proj#004-001. 1949;No. 4.

[5] Hounsfield G. Computerized transverse axial scanning (tomography). Br J Radiol. 1973;46(552):1016–1022.

[6] Lauterbur PC. Image formation by induced local interactions: examples employing nuclear magnetic resonance. Nature. 1973;242:190–191.

[7] Maurer K, Maurer U. Alzheimer: The Life of a Physician and the Career of a Disease, ISBN: 0-231-11896-1. 2003;.

[8] Fratiglioni L, Rocca WA. Epidemiology of dementia. In: Boller F, Cappa SF, eds Handbook of neuropsychology 2nd ed Amsterdam Elsevier. 2001;p. 193–215.

[9] Wimo A, Winblad B, Jönsson L. An estimate of the total worldwide societal costs of dementia in 2005. Alzheimer's & Dementia: The Journal of the Alzheimer's Association. 2007;3(2):81–91.

[10] Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L. Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impariment: a follow-up study. Lancet Neurol. 2006;5:228–234.

[11] Freeborough PA, Fox NC. The boundary shift integral: an accurate measure of cerebral volume changes from registered repeat MRI. IEEE Trans Med Imaging. 1997;16(5):623–629.

[12] Pitiot A, Delingette H, Toga AW, Thompson PM. Learning Object Correspondences with the Observed Transport Shape Measure. Information Processing in Medical Imaging (IPMI). 2003;LNCS 2732:25–37.

[13] Rueckert D, Frangi AF, Schnabel JA. Automatic Construction of 3-D Statistical Deformation Models of the Brain Using Nonrigid Resitration. IEEE Transactions on Medical Imaging. 2003;22(8):1014–1025.

[14] Schott JM, Price SL, Frost C, Whitwell JL, Rossor MN, Fox NC. Measuring atrophy in Alzheimer diseases: a serial MRI study over 6 and 12 months. Neurology. 2005;65:119–124.

[15] Kodama N, Shimada T, Fukumoto I. Image-based Diagnosis of Alzheimer-type Dementia: Measurements of Hippocampal and Ventricular Areas in MR Images. Magn Reson Med Sci. 2002;1(1):14–20.

[16] Wang L, Swank JS, Glick IE, Gado MH, Miller MI, Morris JC, et al. Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. NeuroImage. 2003;20:667–682.

[17] Barnes J, Whitwell JL, Frost C, Josephs KA, Rossor MN, Fox NC. Measurements of the Amygdala and Hippocampus in Pathological Confirmed Alzheimer Disease and Frontotemporal Lobar Denegeneration. Arch Neurol. 2006;63:1434–1439.

[18] Tarroun A, Bonnefoy M, Bouffard-Vercelli J, Gedeon C, Vallee B, Cotton F. Could linear MRI measurements of hippocampus differentiate normal brain aging in elderly persons from Alzheimer disease? Surg Radiol Anat. 2007;29(1):77–81.

[19] Thompson PM, Hayashi KM, de Zubicaray GI, Janke AL, Rose SE, Semple J, et al. Mapping hippocampal and ventricular change in Alzheimer disease. NeuroImage. 2004;22(4):1754–1766.

[20] Apostolova LG, Dutton RA, Dinov ID, Hayashi KM, Toga AW, Cummings JL, et al. Conversion of Mild Cognitive Impairment to Alzheimer disease Predicted by Hippocampal Atrophy Maps. Arch Neurol. 2006;63:693–699.

[21] Carmichael O, Thompson P, Dutton R, Lu A, Lee S, Lee J, et al. Mapping ventricular changes related to dementia and mild cognitive impairment in a large community-based cohort. ISBI Proc. 2006;p. 315–318.

[22] Chételat G, Landeau B, Mezenge F, Viader F, de la Sayette V, Desgranges B, et al. Using voxel-based morphometry to map structural changes associated with rapid conversion in MCI: a longitudinal MRI study. NeuroImage. 2005;27:934–946.

[23] Zipurski RB, Meyer JH, Verhoeff NP. PET and SPECT imaging in psychiatric disorders. Can J Psychiatry. 2007;52(3):146–157.

[24] Smith SM, Zhang Y, Jenkinson M, Chen J, Matthews PM, Federico A, et al. Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis. NeuroImage. 2002;17:479–489.

[25] Boyes RG, Rueckert D, Aljabar P, Whitwell J, Schott JM, Hill DLG, et al. Cerebral atrophy measurements using Jacobian integration: comparison with the boundary shift integral. NeuroImage. 2006;32:159–169.

[26] Smith SM, Rao A, de Stefano N, Jenkinson M, Schott JM, Matthews PM, et al. Longitudinal and cross-sectional analysis of atrophy in Alzheimer's Disease: Cross-validation of BSI, SIENA, and SIENAX. NeuroImage. 2007;36:1200–1206.

[27] Whitwell JL, Przybelsky SA, Weigand SD, Knopman DS, Boeve BF, Petersen RC, et al. 3D Maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. Brain. 2007;130:1777–1786.

[28] Folstein MF, Folstein SE, McHugh PR. Mini-Mental State. J Psych Res. 1975;12:189–198.

[29] Cootes TF, Taylor CJ, Cooper DH, Graham J. Active Shape Models - their training and application. Computer Vision and Image Understanding. 1995;61(1):38–59.

[30] Gerig G, Styner M, Jones D, Weinberger D, Lieberman J. Shape analysis of brain ventricles using SPHARM. MMBIA, Proc of IEEE. 2001;p. 171–178.

[31] Pizer SM, Fletcher PT, Joshi S, Thall A, Chen JZ, Fridman Y, et al. Deformable M-Reps for 3D Medical Image Segmentation. IJCV. 2003;55 (2/3):85–106.

[32] Styner M, Lieberman JA, Pantazis D, Gerig G. Boundary and medial shape analysis of the hippocampus in schizophrenia. Meidcal Image Analysis. 2004;8:197–203.

[33] Styner M, Gerig G, Joshi S, Pizer S. Automatic and Robust Computation of 3D Medial Models Incorporating Object Variability. International Journal of Computer Vision. 2003;55(2/3):107–122.

[34] Davies RH, Twining CJ, Cootes TF, Waterton JC, Taylor CJ. A minimum description length approach to statistical shape modelling. Information Processing in Medical Imaging (IPMI). 2001;LNCS 2082:50–63.

[35] Davies RH, Twining CJ, Allen PD, Cootes TF, Taylor CJ. Shape Discrimination in the Hippocampus Using an MDL Model. Information Processing in Medical Imaging (IPMI). 2003;LNCS 2732:38–50.

[36] Styner M, Rajamani KT, Nolte L, Zsemlye G, Szkely G, Taylor CJ, et al. Evaluation of 3D Correspondence Methods for Model Building. IPMI. 2003;LNCS 2732:63–75.

[37] Nichols TE, Holmes AP. Nonparametric Permutation Tetsts For Functional Neuroimaging: A Primer with Examples. Human Brain Mapping. 2001;15:1–25.

Colored Figures from Chapter 5

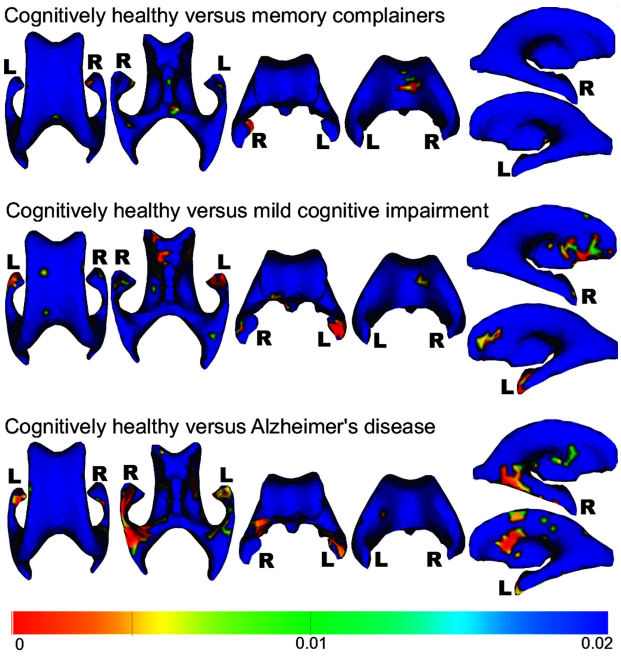


Figure 5.3 (page 86) Local differences in ventricular shape between cognitively healthy and memory complainers, subjects with MCI and subjects with AD. Local shape differences between groups are represented by color-coded p-values (p values > 0.01 are color-coded in blue).

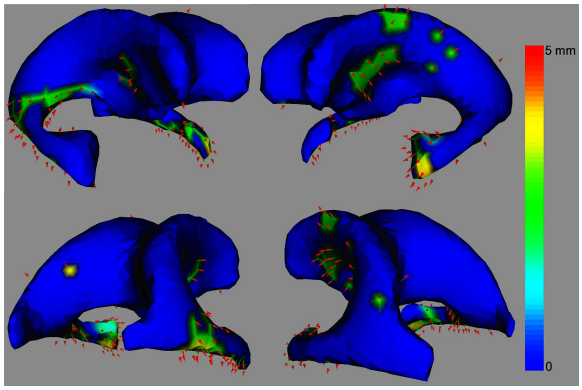
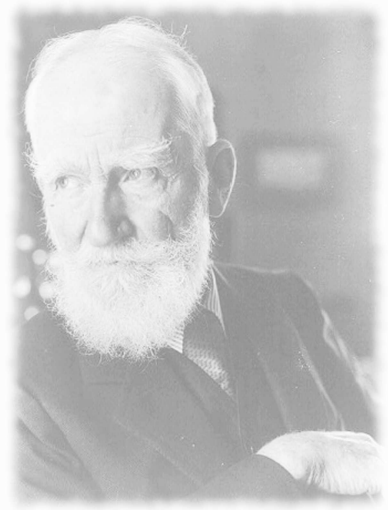


Figure 5.4 (page 87) The direction of the displacement vectors in these images demonstrates that changes between an average cognitively healthy subject and an average AD occur as a result of ventricular enlargement.

CHAPTER 2

GAMEs: Growing and Adapting MEshes for Fully Automatic Shape Modeling and Analysis



Reasonable people adapt themselves to the world. Unreasonable people attempt to adapt the world to themselves. All progress, therefore, depends on unreasonable people.

George Bernard Shaw, 1856-1950.

This chapter is reprinted from:

Medical Image Analysis 2007, Vol. 11(3)

Authors: L. Ferrarini, H. Olofsen, W.M. Palm, M.A. van Buchem, J.H.C. Reiber, and F. Admiraal-Behloul

Title: **GAMEs: Growing and Adapting MEshes for Fully Automatic Shape Modeling and Analysis**, pp. 302-314, Copyright (2007), with permission from Elsevier

Abstract This paper presents a new framework for shape modeling and analysis, rooted in the pattern recognition theory and based on artificial neural networks. Growing and Adaptive MEshes (GAMEs) are introduced: GAMEs combine the Self-Organizing Networks which Grow When Required (SONGWR) algorithm and the Kohonen's Self-Organizing Maps (SOMs) in order to build a mesh representation of a given shape and adapt it to instances of similar shapes. The modeling of a surface is seen as an unsupervised clustering problem, and tackled by using SONGWR (topology-learning phase). The point correspondence between point distribution models is granted by adapting the original model to other instances: the adaptation is seen as a classification task and performed accordingly to SOMs (topology-preserving phase). We thoroughly evaluated our method on challenging synthetic datasets, with different levels of noise and shape variations, and applied it to the analysis of a challenging medical dataset. Our method proved to be reproducible, robust to noise, and capable of capturing real variations within and between groups of shapes.

2.1 Introduction

Statistical shape modeling and analysis have been increasingly used during the last decade as a basis for image segmentation and interpretation. Successful 2D-applications have been described in the literature [1]. Building a statistical model often requires the establishing of correspondence between shape surfaces over a set of training examples. Defining corresponding points on different shapes is not trivial: in some 2D applications, manual landmark definition might be possible but it becomes unpractical when 3D/4D shapes are considered.

Different techniques have been proposed in literature to address this problem. In [2], the authors successfully applied a shape representation based on spherical harmonics (SPHARM) to the analysis of brain ventricles. The SPHARM is a multi-scale approach, which allows for smooth shape representation, even at a very fine scale. The two major limitations of this method are (1) the need of pre-processing steps to generate a spherical-topology object, and (2) the non-intuitive nature of the parameters, which does not allow for an easy interpretation when significant shape differences are found. In [3], the authors introduced deformable medial representations (M-reps) for segmentation of 3D medical structures: they successfully applied their method to the analysis of kidney and hippocampus (see also [4]). The M-reps are multi-scale and can represent objects as composition of multiple figures related with each other. A key step to build up M-reps is the construction of medial models, which can either be created manually (time consuming and error prone), or generated automatically as suggested in [5]. Nevertheless, the first step described in [5] requires a boundary parametrization using SPHARM: thus, the spherical-topology constraint still remains. Moreover, to identify medial models in thin elongated structures is not a trivial task. Another solution, rooted in an information theory framework, is the Minimum Description Length (MDL) approach described in [6]: the authors suggest to (1) use a *descriptive function* to describe corresponding points on different shapes, (2) build up a first model, and (3) evaluate its performances through an objective function. New models are generated with new parameters which are tuned in order to optimize the objective function. The process continues until convergence. The MDL has been successfully applied to the modeling of different structures [7, 8]. Nevertheless, due to the highly non-linear objective function, genetic algorithms are used for the optimization: the computation load is therefore high, and a general global optimum is not granted. A completely different approach is the one suggested in [9]: instead of modeling shapes, the authors suggest to model deformation fields obtained by non-rigid registration. This approach has several advantages when the whole anatomy of a certain organ is studied: it does not require prior segmentation and it provides inter-object relationships. On the other hand, the authors consider shape-based approaches to be a better choice whenever the inter-object relationship is not important or confounds the modeling process. All the modeling approaches described before can be divided into two groups: *groupwise* and *pairwise*. *Groupwise* analyses aim at optimizing an objective function over the whole dataset while creating the statistical model [6], while *pairwise* solutions start with a representative shape of the dataset and build up a model rooted in it. Many advantages of *groupwise* analysis have already been highlighted in the literature [7]; nevertheless, in a recent study on non-rigid registration and segmentation [10], the authors compared *pairwise* and *groupwise* approaches, showing that the simpler framework of *pairwise* methods performed systematically better.

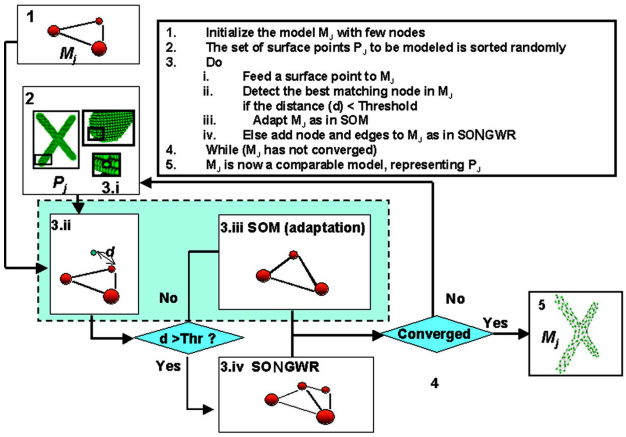


Figure 2.1: *Growing Phase*. During the unsupervised clustering algorithm, a model M_j grows and adapts to learn the topology of the input space P_j. In the *adapting phase*, the input is the M_j model (output of the growing phase), and only the SOM part of the algorithm is used (box with dashed line).

The method presented in this work is a *pairwise* approach to shape modeling. The shape-modeling problem can be summarized as follows: given a set of shapes, one needs to generate Point Distribution Models (PDMs) which describe them. Since shapes need to be compared, a correspondence between points in the PDMs has to be established. Three questions rise: How many nodes are needed? Where should they be located on the surface? How can one define correspondence across shapes? In this work, we introduce a pattern recognition framework to address these questions. In a first phase, a topology-learning unsupervised clustering algorithm is used to select the optimal number of nodes (clusters) and their locations in the input space (3-dimensional space of surface points). In a second phase, the correspondence problem among models is tackled as a classification task: the generalization property of a classifier is used to match unseen cases to similar previously seen points, preserving the topology learned through unsupervised clustering. Growing and Adaptive MEshes (GAMEs) are introduced to implement both the unsupervised clustering and the adapting algorithms. The *growing* phase of GAMEs (unsupervised topology-learning clustering) is based on Self-Organizing Networks which Grow When Required (SONGWR), introduced in [11] (see also [12] and [13] for more details on growing cell structures): SONGWR proved to be (1) more data-driven while growing, and (2) faster in learning input representation, when compared with previous models. The *adaptation* phase of GAMEs is based on Self-Organizing Maps (SOM), which have been proved to be perfectly topology-preserving [14]. To the best of our knowledge, this is the first work showing how a combination of growing structures and self-organizing maps can be used to address the issues of shape modeling and analysis.

In this paper, we provide a thorough evaluation of our method. Working with challenging synthetic shapes, we tested GAMEs' robustness, reproducibility, and accuracy in detecting landmarks. We used the outcome of GAMEs to build statistical models and tested their ability

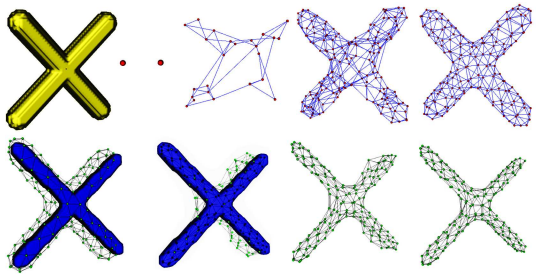


Figure 2.2: *Top* Growing phase: The neural network (model) grows and adapts to the set of surface points representing the shape, detecting the optimal number of clusters (nodes) and their locations. *Bottom* Adapting phase: a model is adapted to a new shape.

of representing real variations within and between groups of shapes. We also successfully applied GAMEs to the analysis of shape variation in populations of brain ventricles, both for healthy elderly individuals and for patients subject to Alzheimer's disease.

The rest of the paper is organized as follows. In section 2.2 we provide an overview of our method. In section 2.3 the methods used to evaluate and compare different PDMs are presented, together with the results for synthetic shapes. Section 2.4 shows the successful modeling of the brain ventricles in Magnetic Resonance Imaging (MRI). We finally give a detailed discussion of the proposed method, highlighting our main contributions, and provide a general conclusion.

2.2 Method

The shape modeling technique presented in this work is based on self-organizing networks which grow when required (SONGWR), introduced in [11]. The general aim of these networks is to provide (learn) an accurate topological representation of a given input space.

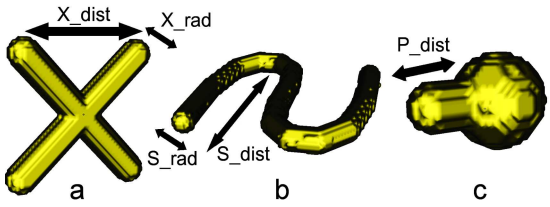


Figure 2.3: Synthetic shapes used for validation: (a) *XShape*, (b) *S-shaped Tube*, and (c) *Sphere*. Some characteristics of these shapes can vary according to the test one wants to perform. *XShape*: distance between tubes (*X_dist*) and radius of the tubes (*X_rad*); *S-shaped Tube*: length of the straight sections (*S_dist*) and radius of the tube (*S_rad*); *Sphere*: length of the protrusion (*P_dist*). Table 2.1 shows the corresponding parameters.

Marsland used his networks for two main applications: novelty detection in a general feature space, and mapping of unknown environments for robots' explorations [11, 15]. Regardless the application, the core of the algorithm is the same (see Fig 2.1): the network (in our case a mesh) is initialized with only few nodes, not connected with each other, and randomly displaced in the input space. Every time a new data point is drawn from the input space, the best matching node is selected among those forming the network at the current state; if the accuracy with which the node represents the data is below a certain threshold a_T, a new node is added, and new connections are created among neighbor nodes (topology-learning behavior); otherwise, the position of the best matching node is adapted. Data points are drawn from the input space until convergence is reached (that is the topology of the network does not change significantly anymore). In Appendix 2.7, a detailed description of the SONGWR algorithm is provided.

The GAMEs method is, to a large extent, similar to the SONGWR method. The main difference is that we adapt the mesh using the Kohonen SOM algorithm: when the mesh is adapting, the motion of the best matching node affects the neighboring nodes as well (see Appendix 2.8 for a recall of the SOM algorithm). The use of the SOM algorithm allows smooth network changes and topology preservation [14]. In the frame of shape modeling and analysis, the input space is three-dimensional and the training data sets are the surface points of the object to be modeled. The GAMEs approach goes through two main steps: (i) the creation of the first model, based on a representative instance of the shapes, using the growing and adapting phase, and (ii) the adaptation of the model to all other shape instances of the data set.

2.2.1 Growing phase: learning the topology

Let us consider a dataset of segmented objects $T = \{O_1, ..., O_n\}$. For each instance in T, the cloud of surface points is detected: $\{P_1, ..., P_n\}$. As a first step, one needs to create a network (or model) of a given cloud P_i (either chosen as a good representative of the training set, or generated as an average based on the training set). Our input space is the set of surface points P_i: iteratively, all the surface points are randomly extracted and given to the network. The whole process can be seen as an unsupervised clustering of the surface points: the new nodes added to the network are new clusters, and the new edges encode neighborhood information. The network grows and adapts until convergence is reached: that is, the network structure does not change significantly when new inputs are given.

One of the key parameters of a general clustering technique is the similarity measure (or distance) used to select the best matching node (cluster). The SONGWR method uses the Euclidian distance. In the GAMEs method, once the network has converged, we further finalize it by repeating the whole process using the Mahalanobis distance when looking for the best matching node. Studying the effect of different distance measures is not in the scope of this paper. In section 2.4 we show how the use of Mahalanobis distance improves the overall performances of the PDM in a medical application. Figure 2.2 illustrates the growing and adapting phase of the GAMEs.

2.2.2 Adaptation phase with topology preservation

After the convergence of the Mahalanobis loop, the final model $M_i = \{A_i, C_i\}$ represents the given cloud of points: A_i and C_i being the set of nodes and edges respectively. The original training set includes several instances of similar shapes $T = \{O_1, ..., O_n\}$: in order to perform shape analysis, one needs comparable models of each instance $\{M_1, ..., M_i, ..., M_n\}$. We define two models (M_i, M_j) *comparable* if each node in model M_i is uniquely associated with a node in M_j, and vice versa: the two models can be *compared* on a node (local) basis. Results of the shape analysis are meaningful only if corresponding nodes are representative of topologically equivalent regions (this is tested by assessing the performances of the final PDM, as illustrated in sections 2.3.3, 2.3.4, and 2.3.5).

Given the initial model M_i and a new cloud of points P_j, a new comparable model M_j is generated through the *adapting phase*: this process can be considered as a classification task, in which the model M_i is the classifier, and the surface points in P_j the objects to classify. As illustrated in Fig. 2.1, the only difference between this phase and the first one is that nodes and edges are not added nor removed. For each point, we identify the best matching cluster and adapt it and its neighbors as in SOM. The whole set of surface points is given iteratively until convergence is reached: the final model M_j is comparable with M_i. Moreover, by adapting the clusters as in SOM, the topology is preserved. The set of models $\{M_1, ..., M_i, ..., M_n\}$ is a PDM: for each node (point) in M_i, the set of corresponding nodes in the other models define its distribution in space.

2.3 Evaluation methods for PDMs: Validation on Synthetic Data

Good performances in shape modeling and analysis rely on a good PDM, whose performances can be assessed by different tests discussed in this section. The synthetic shapes used for the evaluation of our method are shown in Fig. 2.3: *XShape*, *STube*, and *Sphere*. In Table 2.1 we report the geometrical characteristics of each shape, how they vary in the different tests, and how many shapes were generated for each test.

These shapes were chosen because they resemble realistic situations in anatomical structures. Some of them present branches: the *XShape* was chosen to simulate bifurcations with different thickness and angles. The variable protrusion in the sphere simulates structures which can collapse, like temporal horns of brain ventricles. Moreover the highly symmetrical shape of the sphere makes it a challenging shape to model: many false modes of variation can be identified as a result of wrong adaptations along the surface. The *STubes* can present different curvatures, resembling blood vessels (like the cerebral vasculature).

2.3.1 Reproducibility

In both the unsupervised clustering and adapting algorithms, the list of surface points is randomized to assure a more homogenous growth and adaptation for the network. Reproducibility over different runs is an important requirement which needs to be evaluated.

Table 2.1: Dimensions and variability in synthetic datasets (see also Fig 2.3). n is the number of shapes generated for each dataset.

Shape	Surf. Pts. average shape *(Surf. Voxels)*	Shape Characteristics	Dimensions (Voxels) μ	σ	n
XShape	3800	*X_dist*	40	5	40
		X_rad	5	1.25	
STube	6490	*S_dist*	70	3.3	40
		S_rad	5	1.67	
Sphere	1290	*P_dist*	10	3	40

The *reproducibility* test consists in generating n models $\{M_i^1, ..., M_i^n\}$ starting from the same cloud of points P_i, and evaluating both the average accuracy with which the models represent the cloud of points P_i, and the average dissimilarity between models. Given $P_i = (p_i^1, ..., p_i^N)$, with $N = \sharp\{P_i\}$, and a corresponding model $M_i = \{A_i, C_i\}$, the accuracy with which M_i approximates P_i is given by:

$$Acc(M_i, P_i) = \frac{1}{N} \sum_{k=1..N} ||p_i^k - s_i^k||, \quad with \quad s_i^k = \arg\min_{q \in A_i} ||q - p_i^k||, \tag{2.1}$$

where s_i^k is the closest node to p_i^k (according to a pre-defined distance function). The average accuracy is given by averaging all the $Acc(M_i, P_i)$ through the n models of P_i $\{M_i^1, ..., M_i^n\}$.

To define how similar (or dissimilar) two models are is not trivial. Two models can differ both in terms of nodes (number and locations) and in terms of edges (number and lengths). Given a model $M_i^{'}$, its energy is defined as the total squared length of its edges:

$$E_i^{'} = \sum_{q=1..\sharp\{C_i^{'}\}} ||edge_q||^2. \tag{2.2}$$

Given two models $M_i^{'}$ and $M_i^{''}$, the dissimilarity between them is defined as:

$$d(M_i^{'}, M_i^{''}) = Acc(M_i^{'}, M_i^{''}) + Acc(M_i^{''}, M_i^{'}) + (1 - \frac{E_m}{E_M}), \tag{2.3}$$

where $Acc(M_i^{'}, M_i^{''})$ is the accuracy of model $M_i^{'}$ in approximating the nodes of model $M_i^{''}$ (in contrast to the points of P_i, see equation 2.1), and

$$E_m = \min(E_i^{'}, E_i^{''}), \quad E_M = \max(E_i^{'}, E_i^{''}), \tag{2.4}$$

as defined in [11]. The average dissimilarity is given by averaging all the possible $d(M_i^{'}, M_i^{''})$ over $\{M_i^1, ..., M_i^n\}$.

Table 2.2: Reproducibility test: results are given as mean (standard deviation), and averaged over n = 20 cases.

	Dissimilarity ($Voxel$)	Accuracy ($Voxel$)	N. Nodes	N. Edges
X Shape	2.8 (0.11)	1.62 (0.001)	163 (0)	460.85 (4.12)
S Shaped Tube	2.81 (0.08)	1.63 (0.001)	295 (0)	840.35 (5.21)
Sphere	2.83 (0.15)	1.62 (0.002)	64 (0)	177.25 (2.88)

Results are reported in Table 2.2. The average accuracy resulted in 1.62 voxels, and the average dissimilarity between the models was 2.8 voxels. In these experiments the maximum number of nodes in the network was upper-bounded by the number needed to guarantee the required accuracy (standard deviation is zero). The results also show very small standard deviations for the total number of edges (always below 2% of the mean value), which was not constrained. These results proved the reproducibility of our method.

2.3.2 Robustness to noise

The *robustness to noise* test is similar to the *reproducibility* one: only in this case noise is randomly added to the cloud of points before creating a new model. Starting from the original cloud of points P_i, n new clouds are generated $\{P_i^1, ..., P_i^n\}$ by randomly moving $p\%$ of the nodes in space, and removing $q\%$ of them. We generated 20 data sets for each shape with different levels of noise. Results are reported in Table 2.3: both accuracy and dissimilarity did not change significantly compared to the *reproducibility* test. The algorithm proved to be robust to noise.

2.3.3 Accuracy in landmark detection

Correspondence between nodes in different models is useful only when corresponding nodes represent similar landmarks (i.e. similar anatomical locations for medical datasets). When a model is deformed into a new mesh, nodes representing important landmarks on the original shape should move in locations close to equivalent landmarks on the new shape. Given a set of shapes $\{O_1, ..., O_n\}$, one manually identifies t landmarks on each of them $\{l_1^1, ..., l_1^t, ..., l_n^1, ..., l_n^t\}$, where landmarks $l_1^j, ...l_n^j$ ($j = 1..t$) represent similar locations through the set of shapes. Starting from a shape O_i, a model $M_i = \{A_i, C_i\}$ is built up, with $A_i = \{s_i^1, ..., s_i^N\}$ being the set of nodes. For each landmark l_i^j on shape O_i, the best matching node (the closest according to a pre-defined distance function) $s_i^j \in A_i$ is automatically identified. Comparable models are then generated, adapting M_i to the other shapes: $\{M_1, ..., M_n\}$. The nodes previously labeled in M_i ($s_i^j \in A_i$) can be followed in the new model M_r: $s_r^j \in A_r$, $j = 1..t$. The average error for M_r is evaluated averaging the distances between the landmarks on O_r

Table 2.3: Robustness to Noise test (averaged over n = 20 cases). Results are reported (in voxels, as mean(std)) for the three noise conditions (p% - q%), both for Dissimilarity (first column) and Accuracy (second column).

	60%-10%		70%-10%		70%-20%	
X Shape	2.75 (0.1)	1.63 (0.003)	2.74 (0.1)	1.64 (0.003)	2.74 (0.11)	1.64 (0.003)
S Shaped Tube	2.76 (0.09)	1.64 (0.002)	2.74 (0.08)	1.64 (0.002)	2.74 (0.07)	1.64 (0.002)
Sphere	2.78 (0.17)	1.63 (0.005)	2.77 (0.15)	1.64 (0.005)	2.76 (0.17)	1.64 (0.004)

($\{l_r^1, ..., l_r^t\}$) and $s_r^j \in A_r$, $j = 1..t$:

$$AvgDist_r = \frac{1}{t} \sum_{j=1..t} ||s_r^j - l_r^j||, \quad s_r^j \in A_r, \quad l_r^j \in O_r. \tag{2.5}$$

The final error is given by averaging through all the shapes and models[1]. For the *STube* and *Sphere* cases, critical locations were identified and studied (see Fig. 2.4). Table 2.4 reports the average error for each landmark and for the whole set. For the *STube*, the averaging global error was 2.86 voxels ($\sigma = 0.75$ voxels); for the *Sphere* the global error resulted in 3.09 voxels ($\sigma = 0.69$ voxels). We can compare these results with the landmark errors reported in [8]. In their work, the authors compared different methods for model building: SPHARM, DetCov, and MDL. These methods were applied to both lateral ventricles (LatV) (analyzed separately) and femoral head (FH). The SPHARM algorithm resulted in an average error of 7.24 mm ([4.83-14.48] voxels) for FH, and (in average) 4.40 mm ([2.93-4.68] voxels) for LatV; both DetCov and MDL were comparable with manual selection, resulting in an average error of 4.15 mm ([2.77-4.41] voxels) for LatV, and 3.30 mm ([2.20-6.60] voxels) for FH.[2]

[1]While evaluating the error in equation 2.5, it is important to use the nodes s_r^j corresponding to the nodes s_i^j automatically detected on the first model. If one used, for each shape O_r, the best matching nodes in M_r, the final error would be much smaller.

[2]Unfortunately, in [8] no information was given on the image resolutions. Nevertheless, concerning the lateral ventricles, the authors refer to a dataset used for schizophrenia studies. In other works on schizophrenia from the same authors [4, 16], the images acquired were high resolution T1, with a voxel size of 0.94x0.94x1.5 mm^3. The error range in voxel is given considered both the slice thickness and the in-plane resolution as extreme cases. The

Table 2.4: Accuracy in following critical points. Results are given (in voxels) as mean (std) (averaged over n= 40 cases per shape).

	Point 1	Point 2	Point 3	All Points
STube	2.93 (0.89)	2.91 (0.66)	2.74 (0.68)	2.86 (0.75)
Sphere	3.09 (0.69)	-	-	-

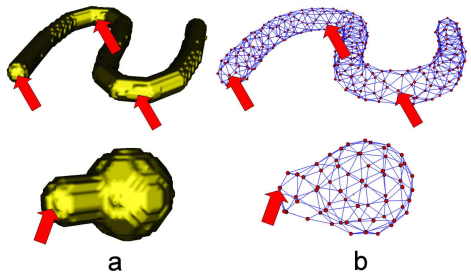


Figure 2.4: Given a shape, we identify important locations (a) and the corresponding best matching nodes in the model (b).

2.3.4 Shape analysis validation

The performances of a statistical model are assessed by evaluating the following functions, as described in [7]:

- *Compactness (C)*: number of parameters needed to describe 90% of the total variation in the dataset;
- *Reconstruction Error (R)*: accuracy with which the model can represent shapes in the training set;
- *Generalization Error (G)*: accuracy with which the model can represent previously unseen shapes;
- *Specificity Error (Sp)*: average dissimilarity between shapes in the training set, and new shapes generated by the model.

Given a set of models, a Principal Component Analysis (PCA) can be used to identify the average model and its main modes of variation, needed to cover up to 90% of the total variation in the training set. The modes are represented by the first k eigenvectors $\{e_1, ...e_k\}$ and eigenvalues $\{\lambda_1, ..., \lambda_k\}$. Starting from the average model, new models can be generated by adding a linear combination of eigenvectors, where the coefficients are constrained by a function of the corresponding eigenvalues. Under these conditions, the functions previously introduced can be formalized as:

- $C(k) = \sum_{m=1}^{k} p_m$, with $p_m = \frac{\lambda_m}{\sum_{i=1..n} \lambda_i}$: the percentage of variance in the dataset which is covered by the first k modes of variation. A better model requires a smaller number of modes to cover the same amount of variance;
- $R(k) = \frac{1}{n} \sum_{i=1}^{n} |S_i^n(k) - S_i|$: the average accuracy in approximating the n shapes of the training set using k modes. S_i is the original shape, while $S_i^n(k)$ is its approximation through a model based on all the n shapes;

femoral images were acquired through computer tomography (CT): a standard voxel size for CT is 0.5x0.5x1.5 mm^3.

Table 2.5: Shape analysis test on synthetic data. For each shape we report the first two modes of variation (% of variation covered, and cumulative %), and the total number of other modes with the cumulative %. We also report the ranges for Reconstruction, Generalization, and Specificity errors.

Shape	Main Modes %			R	G	Sp
	First (cum.)	Second (cum.)	♯ Others (cum.)	(voxels)	(voxels)	(voxels)
XShape	60% (60%)	31% (91%)	0 (91%)	[0.4 - 1.2]	[0.5 - 1.4]	[2.7 - 3.5]
STube	65% (65%)	23% (88%)	1 (90%)	[0.4 - 1.0]	[0.5 - 1.1]	[3.0 - 3.8]
Sphere	70% (70%)	3.5% (73.5%)	13 (90%)	[0.2 - 0.3]	[0.2 - 0.3]	[1.2 - 1.5]

- $G(k) = \frac{1}{n}\sum_{i=1}^{n}|S_i^{n-1}(k) - S_i|$: for a given i, the shape S_i is approximated by a model generated with all the other $n-1$ shapes, and considering the first k modes. The approximation is indicated as $S_i^{n-1}(k)$;
- $Sp(k) = \frac{1}{n_s}\sum_{i=1}^{n_s}[Diss(S_i(k), ObjCls)]$, with $Diss(S_i(k), \{S_1^T, ..., S_n^T\}) = \frac{1}{n}\sum_{i=1}^{n}|S_i^T(k) - S_i|$: is evaluated by creating $n_s = 100$ instances $S_i(k)$ of the object class according to the model (using the first k modes), and averaging the dissimilarity between them and the shapes in the training set S_i^T.

Which function should be regarded as the most important one depends on the application. When the statistical model is used to segment objects in new datasets, the specificity and generalization errors are very important (the model should not generate illegal shapes, nor being overtrained). When we analyze shape variations within a population or between groups, the reconstruction error becomes the most important index: models have to be accurate in order to draw correct conclusions. For each shape, we applied the PCA on 40 instances (see Table 2.1): results are reported in Table 2.5. The right number of main modes was detected in each class, and the identified modes reflected the variations artificially introduced in the datasets. For the *Sphere* case, one important mode of variation was detected (covering almost 70% of the total variation), and other 14 modes all below the 3.5%. The explanation for all these extra "small" modes is the highly symmetrical shape of the *Sphere* object: nodes describing the sphere are more prone to registration errors during the alignment of the shapes, and this accounts for many small modes of variations in which nodes slightly move on the surface.

2.3.5 Local shape analysis

Another criteria to assess the performances of a PDM is the analysis of local shape differences between populations. A good PDM should capture local shape characteristics when compar-

ing different groups: in order to validate this property, one can generate two populations of similar objects and introduce artificial changes in the average shape of one population. Subsequently a PDM is generated to model all the objects in the two populations and the final models are compared statistically. This is done through permutation tests, a statistical tool which has already been successfully applied to the analysis of medical images (see [17] and [18] for a description of the method and applications). The problem can be formulated as following: given two populations G_1 and G_2, are the shapes in G_1 significantly different from those in G_2? If the answer is positive, is it possible to localize the differences on the average surface? In terms of feature space, the first question can be reformulated as "are G_1 and G_2 separable in the feature space?", while the second question seeks for features which mostly differentiate the two populations. Outcome of the permutation test is, in the first place, a *p value* for the omnibus hypothesis "the two groups G_1 and G_2 are drawn from the same population". Moreover, we obtain a *p value* for each node in the model (corrected for multiple tests), telling us whether the distribution of that node in space is the same in G_1 and G_2 or not.

Two populations of *STubes* were created:

- *Population 1*: starting from the average shape, 20 instances were created with different radius (normally distributed, with $\mu = 5 \quad voxels$, $\sigma = 1 \quad voxel$);
- *Population 2*: same as population 1, but in two locations the average radius and standard deviations are different (location 1: $\mu = 10 \quad voxels$, $\sigma = 2 \quad voxels$; location 2: $\mu = 2 \quad voxels$, $\sigma = 0.4 \quad voxels$). See Fig. 2.5.a and Fig. 2.5.b.

The omnibus hypothesis was rejected with $p < 0.01$, indicating that the two populations are globally significantly different. Local p values were also evaluated and mapped on the average shape (see Fig. 2.5.c). The map shows the lowest p values exactly at the locations where we artificially introduced different mean values for the radius. We would like to point out that almost every location on the shape was found to be significantly different. The tubes were generated through splines: changing radius in two locations slightly and systematically affects the entire shape. Because the difference in the second group is systematic, permutation tests correctly highlighted it as significant. Further analysis could include the amount of displacement each node is subject to, while moving from the average shape of population 1 to the average shape of population 2: this would facilitate further investigations in the areas highlighted by the permutation tests.

2.4 Application to brain ventricles

The brain ventricles are in the center of the brain, surrounded by gray and white matter structures which undergo atrophy when affected by Alzheimer Disease (AD) (see Fig. 2.6). Any change of volume or shape occurring in these structures must affect the shape and volume of the ventricles. Shape analysis can highlight local variations which can directly be associated with adjacent areas, improving our knowledge on how AD affects the peri-ventricular brain structures.

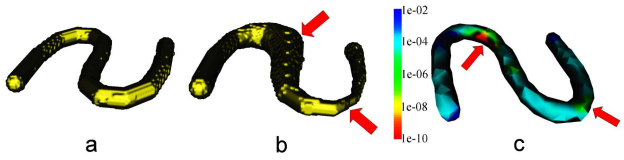


Figure 2.5: Average shapes for population 1 (a) and population 2 (b). The second population is similar to the first one except in two locations, where the radius (in average) changes. (c) Statistical map for shape comparison: the local p values are color-coded on the surface, indicating the areas of most significant difference between the two populations.

We applied our method to study shape variations of brain ventricles in healthy elderly and AD subjects. Twenty-nine patients with probable AD (12 men, mean age 73 years, age range 60-83 years) and 25 volunteers with normal cognitive function (11 men, mean age 74 years, age range 64-89 years) were included. MRI was performed on a 1.5 Tesla MR-system (Philips Medical Systems, Best, The Netherlands) using the following pulse sequences: Dual fast spin-echo (proton density and T2 weighted), time to echo (TE) 27/120 ms, repetition time (TR) 3000 ms, 48 contiguous 3 mm slices without an interslice gap, matrix 256x256, field of view (FOV) 220 mm. FLAIR (fluid attenuated inversion recovery): TE 100 ms, TR 8000, 48 contiguous 3mm slices without an interslice gap, matrix 256x256, FOV 220 mm. For shape modeling purposes, all images were corrected for brain-size and orientation by linear mapping into a stereotaxic space using an automatic affine 12-parameters registration [19]. The LUMC T2-weighted brain template for geriatrics was used as target image [20]. A template-based fuzzy classifier [21] was used to fully automatically extract the cerebrospinal fluid (CSF): interactive 2D tools were then used to re-label the lateral and third ventricles as *ventricles*.

The most suitable statistical model was chosen experimenting with different thresholds a_T (0.01, 0.025, and 0.05), and different initial shapes (a single case visually selected out

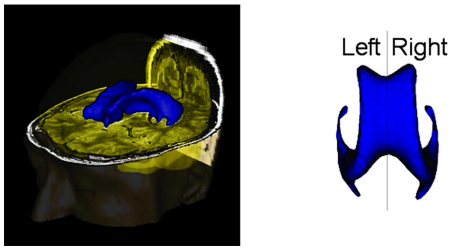


Figure 2.6: The brain ventricles are located in the center of the brain and surrounded by white matter and gray matter structures generally affected by dementia. We refer to *right* and *left* from the patient's point of view.

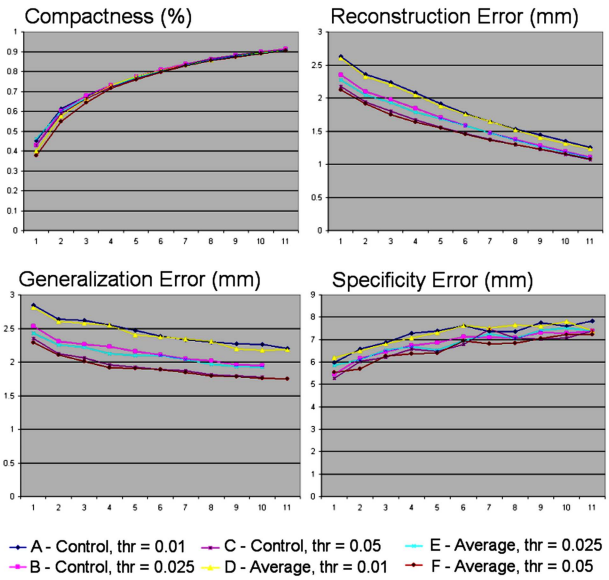


Figure 2.7: PCA performances for statistical modeling of the ventricular shapes in the control group, given as functions of the number of modes. Increasing the accuracy from 0.025 to 0.05 does not improve the model significantly. Using an *Average* shape is generally better than using a good representative from the *Control*.

of the *Control* population, and the average ventricle of the *control* population[3]). Results for *Compactness*, *Reconstruction Error*, *Generalization Error*, and *Specificity Error* are shown in Fig. 2.7 for the control population. Table 3.1 shows the number of nodes and an indication of the adaptation time for different thresholds (average of controls used as initial shape). The use of an average initial shape gives slightly better results than an individual initial shape, when using the same threshold value a_T. Moreover, increasing the accuracy improves the final performances. The difference between average- and individual- initial shape was not pronounced, indicating that, in this application, the algorithm is not significantly sensitive to this initial choice. With respect to the accuracy threshold, a_T = 0.025 and a_T = 0.05 gave significantly better results than a_T = 0.01, while the difference between 0.025 and 0.05 was not relevant. In order to keep a low computational load, we used a_T =0.025, in combination with an average initial shape, for the rest of the experiments. Results under these conditions are shown in Table 2.7, for both controls and AD subjects.

In section 2.2 we stated that the Mahalanobis distance plays an important role while building up the statistical model. In Fig. 2.8 we compare the performances obtained for the *Controls* with (WM) and without (WoM) the use of Mahalanobis distance ($a_T = 0.025$, and

[3]In order to obtain an average ventricle, one first needs to build a voxel-count map of all the *controls*: this map is the sum of all the *control* binary images. A typical ventricle is then obtained by setting a threshold (to 20) on the voxel-count map.

Table 2.6: Computation load for different values of the a_T threshold. The table shows the total number of nodes needed to model the shape, and the average time needed to adapt the model to a new instance (in min., mean $\pm$ std.).

	Num. Nodes	Adaptation Time (min.)
$a_T = 0.01$	337	2 ± 1
$a_T = 0.025$	844	6 ± 2
$a_T = 0.05$	1689	16 ± 4

average control shape as a starting point): the WoM model results more compact, requiring only 7 modes to cover up to 90% of the total variation in the dataset (the WM model requires 11 modes, which are enough for the WoM to reach 100% of covered variation). Nevertheless, when the reconstruction error is considered, the WoM model performs worse than the WM model. Only when all the 11 modes of the WoM model are used, the error is comparable with the WM one: but statistical model usually only considers the first k modes of variation needed to cover the 90% of total variation. Thus, the WoM error, using up to 7 modes of variation, would present a higher error. For what concerns the generalization error, the WoM model performs worse than the WM even when all the 11 modes are considered. Finally, the specificity error does not differ between the two models. We have performed the same analysis on the *AD* group and got similar results.

The performances of our method can be compared with those reported in [8], for both SPHARM and MDL approaches. When applied to lateral ventricles (left and right ventricles considered separately), the MDL resulted in 20 modes of variation, with a generalization error between 3 and 1 mm, and a specificity error between 1.55 and 1.8 mm (SPHARM provided slightly worse results). These results are slightly better than those obtained with our method, but in our case the whole ventricular system (with the exception of the 4^{th} ventricle) was considered as one object, with temporal and occipital horns included in the analysis. This implies a much more challenging shape to model. MDL and SPHARM gave worse results when applied to the femoral head: the generalization error was between 7 and 3 mm for MDL (8 and 4 mm for SPHARM), and the specificity error was between 5.2 and 4.6 mm for MDL (slightly higher for SPHARM). Since we did not apply our method to the same datasets, it is not our goal to claim better performances in respect of other approaches: the quantitative comparison is useful to show that our method provides satisfactory results when applied to medical structures characterized by challenging shapes. Moreover, our method reaches comparable results on clinical T2 weighted images with a slice thickness of 3 mm. Most of the reported work in the literature on brain structures was performed on high resolution T1-weighted images (1mm isotropic).

Table 2.7: Performances of the PDM for brain ventricles ($a_T = 0.025$, average initial shape). The table shows ranges (in mm) for reconstruction error, generalization error, and specificity error.

	Controls	Alzheimer
Compactness	11 modes	15 modes
Reconstruction Error (mm)	[1.1 - 2.3]	[1.1 - 2.4]
Generalization Error (mm)	[1.9 - 2.4]	[2.1 - 2.5]
Specificity Error (mm)	[5.8 - 7.5]	[5 - 7]

2.5 Discussion

2.5.1 Method overview

GAMEs present some advantages when compared to previously published approaches for shape modeling and analysis. Once the cloud of surface point is acquired, the unsupervised clustering automatically detects areas in which more clusters (nodes) are needed: thus, we do not need to iteratively reduce the set of original surface points to a subset with high (sparse) dense point distribution in areas of high (low) curvature, as required in [22]. Differently from methods based on spherical harmonics, GAMEs does not require a spherical topology for the shapes to be parameterized; moreover, being based on node's coordinates, changes in the feature space are directly related to shape changes: thus, the interpretation of the results is straightforward. SPHARM presents a multi-scale feature which we did not investigate in our approach: nevertheless, it would be achievable by a progressive reduction in the accuracy for the growing process (a_T). GAMEs can easily be extended to consider multiple objects representation (as in M-reps). Each node in the mesh can be labeled differently depending on which organ it represents (or which part of an organ, i.e. left- right-ventricle): the growing neural network can then grow independently, and be considered as one in the PCA phase. In [6], the authors highlight how critical some aspects are for the final model: the number of nodes, their locations, and the correspondence across subjects strongly influence the final performances. GAMEs are rooted in a *pairwise* framework: thus, the final model might be influenced by the initial shape one starts with. Nevertheless, not having to deal with genetic optimization, *pairwise* approaches require less computational time and are therefore more suitable for *learning systems*: how do we improve our model when new shapes are available? *Groupwise* approaches have to undertake the whole optimization process again.

2.5.2 Quantitative analysis on synthetic data

The experiments performed on synthetic data showed the reproducibility and robustness to noise of our method. Both these aspects are important when dealing with real datasets, since errors in image acquisition and occlusions are common in medical images. Landmarks detection is a desirable feature for PDMs: shape analysis would be meaningless if corresponding nodes through models did not represent homologous structures. The good performances of

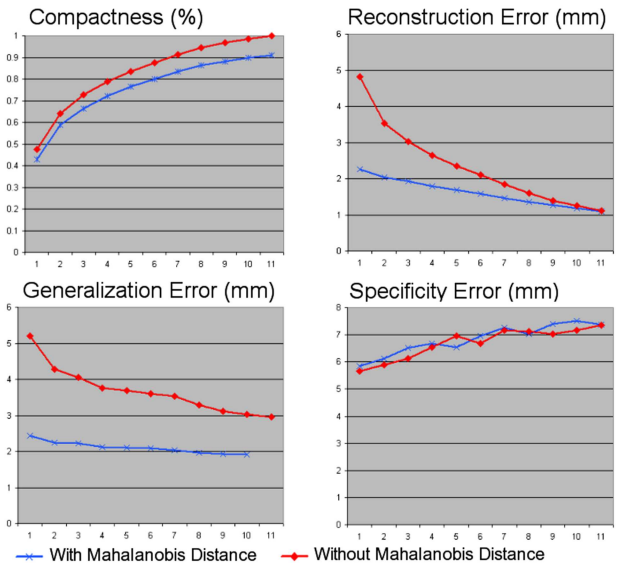


Figure 2.8: Comparison between statistical model obtained using the Mahalanobis distance, and one obtained without. Although not using the Mahalanobis distance leads to a more compact model (7 modes instead of 11 to cover more than 90% of the total variation), both the reconstruction and generalization errors are better when the Mahalanobis distance is used.

GAMEs proved to be very satisfactory for medical image analysis. The PDMs are used to build up statistical models of shape variations. When tested on populations of synthetic shapes, our method could highlight the right modes of variations with a high degree of compactness: the error functions were below the acceptable error for medical image analysis reported in other studies (see [8]). The capability of detecting local shape differences between populations was also tested. On synthetic shapes we could correctly pinpoint the locations of major differences between two populations of *S Shaped* tubes. Statistical maps with color-coded p values were used to visually investigate local deformations. Our method performed successfully in all the steps which characterize shape modeling and analysis processes.

2.5.3 Brain Ventricle Analysis

The analysis of brain ventricles went through the modeling of two populations: patients with Alzheimer's Disease, and age-matched controls. After having chosen the most suitable accuracy level, and having generated an initial average shape, we proceeded with the modeling of all subjects. The PDM performances were assessed through the evaluation of *compactness*, *reconstruction error*, *generalization error*, and *specificity error*. We compared our results with other studies in the literature, and showed that GAMEs provide satisfactory results when applied to challenging medical structures.

Table 2.8: Formulae and values for variables in the unsupervised clustering and adapting algorithms.

acc_T	a_T	f_T	α	β	e_T
0.001	0.025	$1 - \frac{1}{\alpha} * (1 - e^{-\alpha * 5/\beta})$	1.05	3.33 for winner node 14.3 for neighbour node	50

2.6 Conclusions

We presented a new framework for shape modeling and analysis. Growing and Adaptive MEshes (GAMEs) are used to model different instances of similar objects, addressing the problem of point correspondence in a straightforward way. We have thoroughly evaluated our algorithm on challenging synthetic shapes, and successfully applied it to the analysis of brain ventricles in healthy elderly controls and patients subject to Alzheimer disease. Because of its generality, the GAMEs approach can easily be applied to the modeling of other organs. To the best of our knowledge, GAME is the first attempt to merge two sound techniques (SONGWR and SOM) to address the problem of shape modeling and analysis. In future works we will investigate the modeling of multiple objects by growing several meshes simultaneously. Moreover, intensity information could easily be integrated in the modeling as fourth dimension, and different similarity measurements investigated.

2.7 Appendix 1

In this section we present the *growing* phase of GAMEs. The threshold values used in our implementation are reported in Table 2.8: these values were also used in [11], except for the threshold a_T. Given a cloud of surface points $P_i = \{p_i^1, ..., p_i^N\}$, the *unsupervised clustering algorithm is given*:

1. Initialization
2. Do
   (a) Randomize order in P_i
   (b) Cluster P_i
   (c) Remove unused nodes
   (d) Evaluate new accuracy

   while accuracy is above a certain threshold acc_T
3. Evaluate covariance matrixes
4. Repeat loop 2 using Mahalanobis distance
5. Update covariance matrixes

Following the terminology introduced in [11]:

1 Initialization: create the set of nodes $A = \{n_1, n_2\}$, with n_1 and n_2 randomly located in the 3D space; create the set of connections among nodes (edges) $C = \emptyset$
2a Randomize: The sequence of surface points P_i is randomly re-sorted. This ensures a homogeneous growth of the network at each step of the loop.
2b Cluster P_i

- This is the core of SONGWR ([11]), reported for clarity
- For each point $p_k^i \in P_i$:
  1. $\forall \quad n_q \in A$ evaluates $||p_k^i - n_q||$, where $||\cdot||$ is the Euclidean distance in 3D
  2. select the two best matching nodes $s, t \in A$ as
  $$s = \arg\min_{n \in A} ||n - p_k^i||, \quad t = \arg\min_{n \in A - s} ||n - p_k^i|| \tag{2.6}$$
  3. create the connection between s and t $C = C \cup \{(s,t)\}$ (if not present in C), and set the age of the connection $Age_{s,t} = 0$
  4. evaluate the activation for s, $a = e^{-||p_k^i - s||}$
  5. if $a < a_T$ and $f_s < f_T$
     - add a new node r, $A = A \cup \{r\}$, with $r = (s + p_k^i)/2$
     - add edges $C = C \cup \{(r,s),(r,t)\}$, with $Age_{r,s} = 0$, $Age_{r,t} = 0$
     - remove edge $C = C - (s,t)$,

     where a_T and f_T are thresholds, and f_s is the firing value for node s (defined in equation 2.8)
  6. else
     - adapt the positions of s and its neighbours
     $$s = s + e_w * f_s * (p_k^i - s), \quad n^s = n^s + e_n * f_k * (p_k^i - n^s) \tag{2.7}$$
     where
     $0 < e_n < e_w < 1$, $n^s \in A_s$, $A_s \subseteq A$ is s'neighborhood,
     $e_n = 0.1 * e^{-\frac{d^2}{2*\sigma^2}}$, $\sigma = 0.42$, $d =$ distance from s
     $e_w = 0.1, (= e_n, \quad \text{with } d = 0)$
  7. increase age of edges ending in s, $Age_{(s,v)} = Age_{(s,v)} + 1, v \in A_s$
  8. decrease the firing value associated with s and its neighbors
  $$f_s = 1 - \frac{1}{\alpha_w} * (1 - e^{-\alpha_w t/\beta_w}), \quad f_n = 1 - \frac{1}{\alpha_n} * (1 - e^{-\alpha_n t/\beta_n}), \tag{2.8}$$
  where t is the number of times the given node has already been selected as s, or as neighbour of s

9. remove edges with $Age_{(v,z)} > e_T$
10. remove nodes without edges

2c Remove nodes which have never been selected as best matching in loop 2b
2d Evaluate Accuracy with which the model represents the surface points.
3 Evaluate Covariance: for each node in the model, find the subset of surface points which has the node as best-matching node, and evaluate the covariance matrix associated with them.
4 Repeat Loop block 2 using everywhere the Mahalanobis distance for $||\cdot||$.

2.8 Appendix 2

In this section we give a brief overview on how nodes are adapted in Self-Organizing Maps (SOM, [14]).

Each output node has an associated vector of dimensionality D, $w_i = (w_i^1, .., w_i^D)$, $i = 1..N$. When an input vector x is given to the network, the output best matching node is identified:

$$o_i(x) = \arg\min_{q \in O} ||q - x||. \tag{2.9}$$

Each nodes is then modified as follows:

$$w_k^{new} = w_k^{old} + \mu * f(i,k) * (x - w_k^{old}), \quad k = 1..N, \tag{2.10}$$

where $f(i,k)$ is the neighborhood function, and has value 1 when $i = k$ and decreases with the distance $d(i,k)$ between node i and k in the array. An example of neighborhood function is given by:

$$f(i,k) = exp^{-\frac{d(i,k)^2}{2*\sigma^2}}, \tag{2.11}$$

where σ is the width parameter which can be decreased in time to assure convergence. Nodes far from the best matching one are only slightly adapted; nearby nodes responds to similar input and get similar updates, preserving topology.

Bibliography

[1] Cootes TF, Taylor CJ, Cooper DH, Graham J. Active Shape Models - their training and application. Computer Vision and Image Understanding. 1995;61(1):38–59.

[2] Gerig G, Styner M, Jones D, Weinberger D, Lieberman J. Shape analysis of brain ventricles using SPHARM. MMBIA, Proc of IEEE. 2001;p. 171–178.

[3] Pizer SM, Fletcher PT, Joshi S, Thall A, Chen JZ, Fridman Y, et al. Deformable M-Reps for 3D Medical Image Segmentation. IJCV. 2003;55 (2/3):85–106.

[4] Styner M, Lieberman JA, Pantazis D, Gerig G. Boundary and medial shape analysis of the hippocampus in schizophrenia. Meidcal Image Analysis. 2004;8:197–203.

[5] Styner M, Gerig G, Joshi S, Pizer S. Automatic and Robust Computation of 3D Medial Models Incorporating Object Variability. International Journal of Computer Vision. 2003;55(2/3):107–122.

[6] Davies RH, Twining CJ, Cootes TF, Waterton JC, Taylor CJ. A minimum description length approach to statistical shape modelling. Information Processing in Medical Imaging (IPMI). 2001;LNCS 2082:50–63.

[7] Davies RH, Twining CJ, Allen PD, Cootes TF, Taylor CJ. Shape Discrimination in the Hippocampus Using an MDL Model. Information Processing in Medical Imaging (IPMI). 2003;LNCS 2732:38–50.

[8] Styner M, Rajamani KT, Nolte L, Zsemlye G, Szkely G, Taylor CJ, et al. Evaluation of 3D Correspondence Methods for Model Building. IPMI. 2003;LNCS 2732:63–75.

[9] Rueckert D, Frangi AF, Schnabel JA. Automatic Construction of 3-D Statistical Deformation Models of the Brain Using Nonrigid Resitration. IEEE Transactions on Medical Imaging. 2003;22(8):1014–1025.

[10] Crum WR, Camara O, Rueckert D, Bhatia KK, Jenkinson M, Hill DLG. Generalised Overlap Measures for Assesment of Pairwise and Groupwise Image Registration and Segmentation. MICCAI. 2005;LNCS 3749:99–106.

[11] Marsland S, Shapiro J, Nehmzow U. A self-organizing network that grows when required. Neural Networks. 2002;15:1041–1058.

[12] Fritzke B. Growing Cell Structures - A self-organizing Network for Unsupervised and Supervised Learning. Neural Networks. 1994;7(9):1441–1460.

[13] Fritzke B. Kohonen Feature Maps and Growing Cell Structures - a Performance Comaprison. NIPS Proc. 1992;p. 123–130.

[14] Kohonen T. The Self-Organizing Map. Proceedings of the IEEE. 1990;78(9):1464–1480.

[15] Duckett T, Marsland S, Shapiro J. Fast, On-Line Learning of Globally Consistent Maps. Autonomous Robots. 2002;12:287–300.

[16] Styner M, Lieberman JA, McClure RK, Weinberger DR, Jones DW, Gerig G. Morphometric analysis of lateral ventricles in schizophrenia and healthy controls regarding genetic and disease-specific factors. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(13):4872–4877.

[17] Nichols TE, Holmes AP. Nonparametric Permutation Tetsts For Functional Neuroimaging: A Primer with Examples. Human Brain Mapping. 2001;15:1–25.

[18] Terryberry TB, Joshi SC, Gerig G. Hypothesis Testing with Nonlinear Shape Models. Information Processing in Medical Imaging (IPMI). 2005;LNCS 3565:15–26.

[19] Woods RP, Grafton ST, Holmes CJ, Chery SR, Mazziotta JC. Automated Image Registration I. General methods and intrasubject, intramodality validation. J Comput Assist Tomogr. 1998;22:139–152.

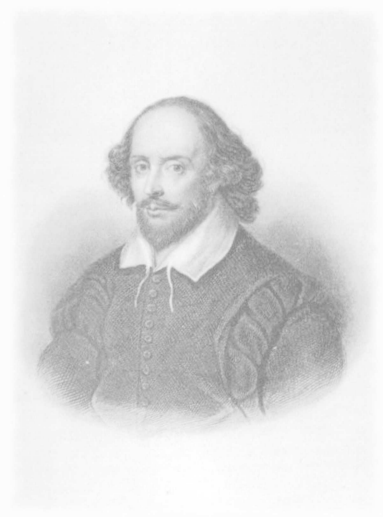
[20] Admiraal-Behloul F, van den Heuvel D, Olofsen H, Schmitz N, van Buchem MA. Brain templates for the elderly. In: International Society for Magnetic Resonance in Medicine (ISMRM) - Toronto; 2003. .

[21] Admiraal-Behloul F, van den Heuvel DMJ, Olofsen H, van Osch MJP. Fully automatic segmentation of white matter hyperintensities in MR images of the elderly. Neuroimage. 2005;28(3):607–617.

[22] Kaus MR, Pekar V, Lorenz C, Truyen R, Lobregt S, Richolt J, et al. Automated 3D PDM Construction Using Deformable Models. Proceedings of the Eigth IEEE International Conference on Computer Vision (ICCV'01). 2001;1:566–572.

CHAPTER 3

Shape Differences of The Brain Ventricles in Alzheimer's Disease



The Devil hath power to assume a pleasing shape.

William Shakespeare (Hamlet), 1564-1616.

This chapter is reprinted from:

NeuroImage 2006, Vol. 32 (3)

Authors: L. Ferrarini, W.M. Palm, H. Olofsen, M.A. van Buchem, J.H.C. Reiber, and F. Admiraal-Behloul

Title: **Shape Differences of The Brain Ventricles in Alzheimer's Disease**, pp. 1060-1069,

Abstract The brain ventricles are surrounded by gray and white matter structures that are often affected in dementia in general, and Alzheimer's disease (AD) in particular. Any change of volume or shape occurring in these structures must affect the volume and shape of the ventricles. It is well known that ventricular volume is significantly higher in AD patients compared to age-matched healthy subjects. However, the large overlap between the two volume distributions makes the measurement unsuitable as a biomarker of the disease. The purpose of this work was to assess whether local shape differences of the ventricles can be detected when comparing AD patients and controls. In this work, we captured the ventricle's shape and shape variations of 29 AD subjects and 25 age-matched controls, using a fully automatic shape modeling technique. By applying permutation tests on every single node of a mesh representation of the shapes, we identified local areas with significant differences. About 22% of the ventricular surface was found to be significantly different ($P < 0.05$) ($\sim$14% of the left, and $\sim$7% of the right side). We found that in patients with Alzheimer disease, not only the lateral horns were significantly affected, but also the areas adjacent to the anterior corpus callosum, the splenium of the corpus callosum, the amygdala, the thalamus, the tale of the caudate nuclei (especially the left one), and the head of the left caudate nucleus.

3.1 Introduction

Alzheimer's disease (AD) is the most common cause of dementia in middle-aged and older adults. It is a progressive, irreversible, and degenerative brain disorder that causes impairment in memory, thinking, and behavior. The vast emotional and economic costs of this disease are continuously growing, as the population gets older. Although the risk of developing AD increases with age, AD is by no means a part of the normal aging process. In the absence of this disease, the human brain can often operate properly to the age of 100 and beyond.

The imaging-based measurements of disease progression in AD that have received the most attention are MRI measures of the rate of change in whole-brain volume and hippocampus volume [1–7]. Early studies showed that the brain volume decreases and cerebrospinal fluid volume increases with advancing age. It has been shown in [8, 9] that different brain regions lose volume at different rates in a non-linear region-dependent manner, with prefrontal volume declining more rapidly than other brain regions [10, 11]. Different diseases with different phenotypic presentations may be associated with specific patterns of regional atrophy. Studies in AD have shown prominent early involvement of medial temporal lobe structures, especially the entorhinal cortex and the hippocampus [12–15].

In AD, there is increasing evidence of marked damage and dysfunction not only in the gray matter but also in the white matter [16–18]. Neuropathological and biochemical studies on white matter from AD subjects have demonstrated dramatic loss of myelin and axons [5, 17]. Several studies have demonstrated that structural white matter changes, such as atrophy of the corpus callosum, are significantly greater in AD than in healthy subjects [19–21].

The brain ventricles (see Fig. 3.1) are in the center of the brain, surrounded by gray and white matter structures that are often affected by dementia diseases: deep white matter, corpus callosum, hippocampus, amygdala, caudate nucleus, and thalamus. Any change of volume or shape occurring in these structures must affect the shape and volume of the ventricles.

It is well known that brain ventricle volume is significantly higher in AD patients compared to age-matched healthy subjects. However, the distributions of volume measurements of both groups present a large overlap, which makes the volume measurement not suitable as a biomarker.

The purpose of this work was first to study the shape of the brain ventricles and their

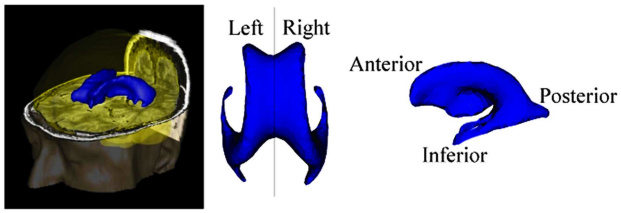


Figure 3.1: The brain ventricles discussed in this paper follow the indications shown in this image: left, right, anterior, and posterior are considered from the patient's point of view.

variations in healthy elderly and AD subjects, and second to investigate the presence of significant shape differences between the two groups. The shape of brain ventricles is highly concave and therefore challenging to model and analyze: it requires therefore sophisticated computing techniques.

3.2 Methods

3.2.1 Subjects

Twenty-nine patients with probable AD (12 men, 17 women, mean age 73 years, age range 60.83 years) and 25 volunteers with normal cognitive function (11 men, 14 women, mean age 74 years, age range 64.89 years) were included. The patients with probable AD had been consecutively referred to our outpatient memory clinic. The volunteers had been recruited through an advertisement in a local newspaper. All subjects were evaluated for memory loss using a standardized dementia screening that included a detailed medical history, a general internal and neurological exam, laboratory tests, neuropsychological testing including the Mini Mental State Examination (MMSE), and magnetic resonance images (MRI) of the brain. Of the 29 patients with probable AD, 2 patients were severely demented (MMSE < 10), 20 patients were moderately demented ($9 <$ MMSE < 21), and 7 patients were mildly demented (MMSE > 20). Diagnoses were made in a multidisciplinary consensus meeting according to the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria for probable AD. Patients and controls were included if they (1) were older than 60 years, (2) had no other neurologic or psychiatric illness, and (3) had no abnormalities on MRI other than white matter hyperintensities or an incidental small lacunar lesion ($>$ 5 mm diameter). The study was approved by the local Medical Ethical Committee. Written informed consent was obtained from all subjects or from a close relative if a patient was demented.

3.2.2 MRI acquisition and pre-processing

MRI was performed on a 1.5-T MR-system (Philips Medical Systems, Best, The Netherlands) using the following pulse sequences: dual fast spin-echo (proton density and T2 weighted): time to echo (TE) 27/120 ms, repetition time (TR) 3000 ms, 48 contiguous 3-mm slices without an interslice gap, matrix 256 x 256, field of view (FOV) 220 mm. FLAIR (fluid attenuated inversion recovery): TE 100 ms, TR 8000, 48 contiguous 3-mm slices without an interslice gap, matrix 256 x 256, FOV 220 mm.

In-house developed automated segmentation software (SNIPER, Software for NeuroImage Processing in Experimental Research) that combines template-based fuzzy clustering, fuzzy inference, and region-growing techniques was used to pre-process the images. Using the method described in Admiraal-Behloul et al. [22], the software extracts fully automatically the intracranial cavity, the cerebrospinal fluids (CSF), and the white matter hyperintensities. The lateral and third ventricles were semiautomatically extracted by re-labeling the wrongly segmented ventricular CSF to ventricles; this was done using interactive editing

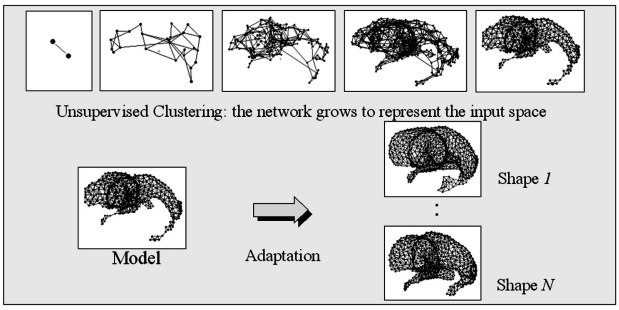


Figure 3.2: The shape modeling algorithm: unsupervised surface point clustering and adaptation phase.

tools: the user, by a simple click, re-labels pre-segmented CSF. The re-labeling has been done slice-by-slice, using 2D region growing.

For shape modeling and statistical comparison purposes, all the images were corrected for brain size and orientation by linear mapping into a stereotaxic space using an automatic affine 12-parameters registration [23]. The LUMC T2-weighted brain template for geriatrics was used as target image [24].

3.2.3 Statistical shape Modeling

The notion of biological shape seems reasonably well explained by a statistical description over a large population of instances [25]. The model building requires, as a key step, the establishing of correspondence between shape surfaces over a large set of training examples. Finding correspondence between two surfaces implies finding corresponding elements that share position and/or shape similarities. Finding the right way to define correspondent points on different shapes is difficult. In some 2D application, manual landmark definition might be possible, but it becomes unpractical when 3D shapes are considered due to the larger amount of data. In 3D, the manual landmark definition is time consuming, poorly reproducible, and error prone. The problem can be summarized in 3 main questions: How many landmarks or nodes (in mesh representation) are needed?, Where should they be located?, and How can one define node correspondence across several instances of shapes?

In a previous work [26], we showed how to use growing neural networks to answer all three questions. Most of the growing neural networks are variations of the growing cell structure (GCS) introduced by Fritzke [27]. Marsland presented a Self- Organizing Network That Grows When Required (SONGWR) [28]. SONGWRs proved to be (1) more data driven while growing and (2) faster in learning input representation when compared with previous models.

In a SONGWR, each node is associated with a subspace of the input space. The network is initialized with few nodes randomly located in the input space and not connected together. At each iteration, a new input is given to the network; in order to maintain a good representation of the input space, the existing nodes are moved or removed (adaptation), or new nodes and

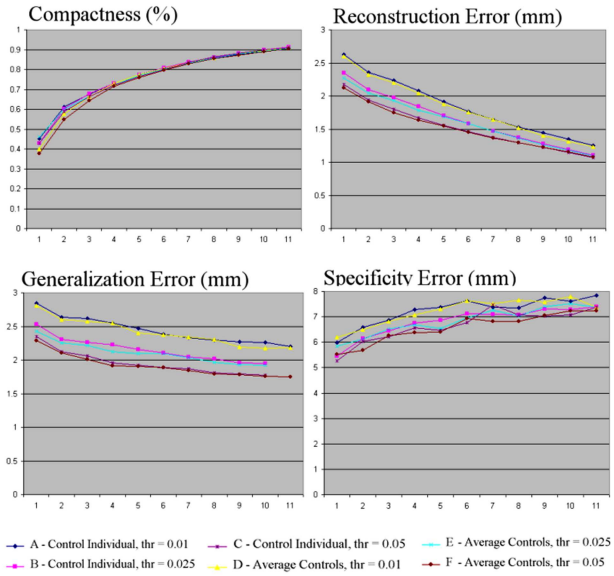


Figure 3.3: PCA performances for statistical modeling of the ventricular shapes in the control group. The plots show how the performances change depending on the total number of modes of variation being considered (horizontal axis). Increasing the accuracy from 0.025 to 0.05 does not improve the model significantly. Using an average shape is generally better than using a good representative from the data set.

connections are added to the network (growing) (see Fig. 3.2). The process is repeated until convergence: a final model of the input space (at a desired accuracy level) is reached. The SONGWR principle is based on self-organizing maps (SOM) [29], which are known to be perfectly topology-preserving: the network preserves neighborhood relations in the data by mapping neighboring inputs onto neighboring nodes in the map.

In case of shape modeling, the input space of the neural network is the whole set of surface points extracted from the object to be modeled (i.e., extracted from segmented brain ventricles). A representative instance of the shape is used as input for the growing and adapting iterative algorithm. Once the SONGWR algorithm has converged, the optimal number of nodes and their spatial locations are identified. The next step is to find corresponding nodes across different instances of the shape to be modeled. At this stage, only the adaptation part of the SONGWR algorithm is used. Adapting the network to a new instance is equivalent to using the model as a classifier: for each given point in the new shape instance, the best-matching node is selected (as the closest to the given input according to a predefined distance function) and adapted accordingly. The process has similarity with SOMs because there are no added/removed nodes anymore (note that the shape instances must be normalized before modeling). After adapting the initial mesh to all other instances of the data set individually, every node of the initial mesh will have a corresponding cloud of matching nodes across the

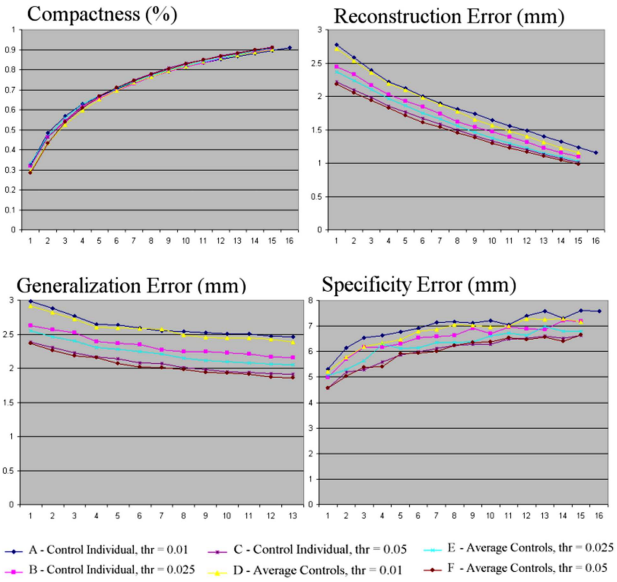


Figure 3.4: PCA performances for statistical modeling of the ventricular shapes in the Alzheimer's disease group. The plots show how the performances change depending on the total number of modes of variation being considered (horizontal axis). The results are comparable to the modeling of the control group shown in Fig. 3.3.

data set.

The aim of this work is to investigate local differences between two populations of shapes. It is therefore important that the models (meshes) used to describe the objects are sensitive to shape differences across individuals. In this respect, we followed the approach described by Davies et al. [30] to evaluate the obtained mesh representation.

A mesh is composed of N nodes, where each node is characterized by three coordinates in the 3-dimensional space. Thus, each mesh is a 3 x N -dimensional object x in the feature space. Given a set of meshes, one can carry out a principal component analysis (PCA) to get a mean mesh and main eigen variations.

The performances of the resulting statistical model can be evaluated using the following

functions:

$$C(M) = \sum_{m=1}^{M} p_m \tag{3.1}$$

$$E_r(M) = \frac{1}{N_s} \sum_{i=1}^{N_s} |S_i^{N_s}(M) - S_i| \tag{3.2}$$

$$G(M) = \frac{1}{N_s} \sum_{i=1}^{N_s} |S_i^{N_s-1}(M) - S_i| \tag{3.3}$$

$$S_p(M) = \frac{1}{N} \sum_{i=1}^{N} Diss(S_i(M), \{S_1^T, ..., S_{N_s}^T\}) \tag{3.4}$$

$$Diss(S_i(M), \{S_1^T, ..., S_{N_s}^T\}) = \frac{1}{N_s} \sum_{i=1}^{N_s} |S_i^T(M) - S_i(M)|, \tag{3.5}$$

where:

- Compactness $C(M)$: the percentage of variance in the data set that is covered by the first M modes of variation (where P_m is the percentage of total variation covered by the m^{th} mode). A better model requires a smaller number of modes to cover the same amount of variance.
- Reconstruction error $E_r(M)$: the average accuracy in approximating the N_s shapes of the training set using M modes. S_i is the original shape whereas $S_i^{N_S}(M)$ is its approximation through a model based on all the N_s shapes. This function shows how good the model is in reconstructing the shapes seen in the training set.
- Generalization error $G(M)$: for a given i, the shape S_i is approximated by a model generated with all the other shapes. The approximation is indicated as $S_i^{N_S-1}(M)$; this function shows how good the model is in generalizing what it has learnt from a training set, working on an unseen shape.
- Specificity error $S_p(M)$ is evaluated by uniformly sampling $N = 100$ instances $S_i(M)$ of the object class according to the model, and averaging the dissimilarity between them and the shapes in the training set S_i^T. This function measures the similarity between the training set and the shapes one can create through the model.

There are mainly two factors that can strongly influence the performances of the final statistical model: the choice of the first shape-instance to be used to build the first mesh representation [31], and the accuracy threshold a_T [28] that defines how accurate the first model should be in approximating the corresponding object. In this study, we have experimented with these two factors and chosen the best statistical model according to Eqs. 3.1-3.5.

The PCA approach described in this section is used mainly to identify global variations within each population and to estimate the performances with which the growing and adapting phases model the given shapes. An independent component analysis (ICA) approach

could be used to highlight local variations within a population [32], but this lies outside the scope of this work.

3.2.4 Surface-based statistical comparison

The highlighting of local shape differences between populations is possible because each node in the first model is uniquely labeled; it is thus possible to follow the node position to find corresponding points within each population and across populations.

When comparing two populations, several statistical tests can be used. Permutation tests have been successfully applied to the analysis of brain images [2, 33]; they require limited assumptions and provide correct results for multi-test analysis when local information is needed (i.e., local surface comparison of two or more groups).

Given two populations G_1 and G_2, the question one tries to answer is: are the shapes in G_1 significantly different from those in G_2? Moreover, if such a global difference exists, can one localize the most different areas on the surface? In the feature space, these questions translate as: "are G_1 and G_2 separable?", and "which feature set leads to the best separability?".

The permutation tests provide us with a P value for the omnibus hypothesis "the two groups G_1 and G_2 are drawn from the same population". Moreover, we obtain the P values for each node in the model, showing whether the distribution of that node in space is the same in G_1 and G_2 or not. Permutation tests can be summarized as follows:

1. Considering two groups G_1 and G_2:
   (a) Each node in the average model has two corresponding clouds of points C_1 and C_2, obtained considering the positions the corresponding nodes assume through all the shapes in G_1 and G_2.
   (b) C_1 and C_2 are compared via a Hotelling's T2 statistical test; the outcome of the test is the t value for the node comparison (are the node positions distributed in C_1 and C_2 in significantly different ways?).
2. For 10,000 times, two groups of shapes A and B are built up by randomly mixing G_1 and G_2, and point 1 is performed on them. Only the highest t value is stored for each iteration.
3. A critical t value t_c is evaluated as the k^{th} highest value of all the 10,000 t values previously stored (plus the t_{max} for the original subdivision in G_1 G_2), where

$$k = \lfloor \alpha * 10,000 \rfloor + 1, \alpha = 0.05. \tag{3.6}$$

4. The P value for the omnibus hypothesis "G_1 and G_2 are the same" is evaluated as

$$P_{value} = \frac{N}{N_{test}}, where \tag{3.7}$$

$$N = \{stored \quad t_{values} | t_{value} > t_c\}, \tag{3.8}$$

$$N_{tests} = 10,000 + 1. \tag{3.9}$$

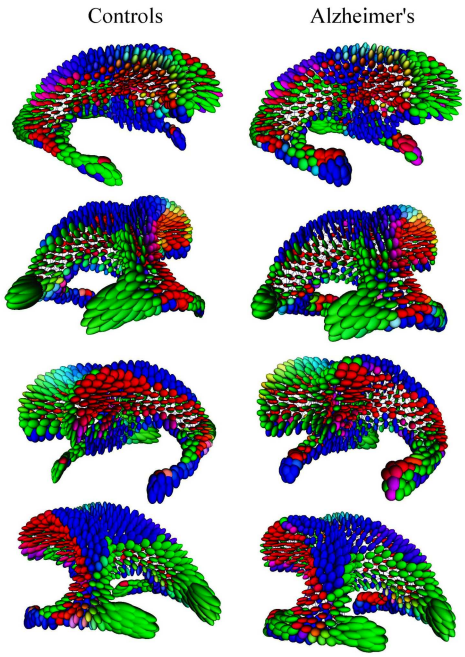


Figure 3.5: Tensor-based representation of the nodes within controls (left column) and AD patients (right column). Each tensor is color-coded according to the direction of the main eigenvector: green for the anterior-posterior direction, blue for the inferior-superior direction, and red for the left right direction.

5. Finally, point 4 is applied to each single node, counting how many t values are higher than the t value associated with a particular node in the original G_1 G_2 grouping of shapes, and dividing the number for $10000 + 1$; this leads to a P value (corrected for multi-tests) for each node in the model.

3.3 Results

3.3.1 Optimal shape modeling parameters

In order to find the most suitable statistical model, we experimented with different threshold a_T and different initial shapes. As threshold level, we compared 0.01, 0.025, and 0.05. Concerning the initial shape, we considered a single case randomly selected out of the control population, and the average ventricle of the control population.

Results for compactness, reconstruction error, generalization error, and specificity error, obtained for the modeling of the ventricular shapes of the control group, are shown in Fig. 3.3. Those obtained for the AD group are given in Fig. 3.4. Table 3.1 shows the number of

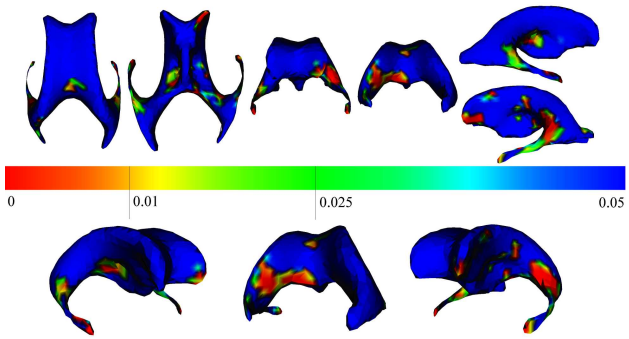


Figure 3.6: Color-coded maps showing the P value associated with each node while comparing controls and AD patients. The P values were evaluated at $\alpha = 0.05$.

nodes and an indication of the adaptation time for different thresholds (average of controls used as initial shape).

When an average initial shape was used we achieved slightly better results than with an individual initial shape (at the same a_T threshold). Moreover, the performances improved while increasing the accuracy: a_T = 0.025 and a_T = 0.05 gave significantly better results than a_T = 0.01, whereas the difference between 0.025 and 0.05 was not relevant. Finally, the choice between an average- and individual initial shape did not seem to influence the final results relevantly. In order to keep a low computational load (see Table 3.1), we used a_T = 0.025, in combination with an average initial shape, for the rest of the experiments.

3.3.2 Tensor maps for groups representation of global and local variations

Given a node in the model, we considered the cloud of locations the node assumed moving through matched nodes in the population: computing the covariance matrix associated with the cloud and its eigenvectors allows us to generate a tensor representation of the node (see Fig. 3.5). Each ellipsoid represents the distribution in space of a particular node while moving across individuals belonging to the same group. The size of the ellipsoids is proportional to

Table 3.1: Computation load for different thresholds: we report the number of nodes of the final model and the time needed to adapt the model to a new shape.

	No. of ndoes	Adaptation time (in minutes, per shape)
$a_T = 0.01$	337	2 ± 1
$a_T = 0.025$	844	6 ± 2
$a_T = 0.05$	1689	16 ± 4

Table 3.2: Number of nodes in the model and percentage of surface area found to be significantly different at 95% and 99% of confidence interval (LV = left ventricle, RV = right ventricle).

	No. of nodes	% Total Area	No. nodes LV	% LV	No. nodes RV	% RV
$\alpha = 0.05$	183	22	121	14	62	7
$\alpha = 0.01$	80	9	55	7	25	3

the eigenvalues of the tensor, whereas the color-coding maps the direction in space of the main eigenvector.

Most of the tensors present a first (main) mode of variation perpendicular to the surface rather than along the surface. Variations along the surface are generally a consequence of misplacement in adaptation (i.e., corresponding nodes that do not represent similar anatomical locations) or registration errors (i.e., small errors due to incorrect transformation matrix). Perpendicular variations are more likely to reflect real variations in shape because the nodes are always forced to be on the surface. Intuitively, the more perpendicular to the surface the tensor is, the better the node can capture the variations within a population and across populations. One can notice that the main modes of variation in both groups are the global dilation/shrinking of the ventricles and the elongation/contraction of the temporal horns.

The most striking results are the asymmetrical variations (left vs. right) between the inferior temporal horns within healthy controls as well as in the AD group. Note that in the control group, the left inferior horn presents mainly inferior-superior thickening variations (blue) whereas the right inferior horn shows elongation variations in the anterior-posterior direction (green).

Both left and right horns are showing more pronounced variations (according to the sizes of the ellipsoids) within the AD group compared to the controls. In AD, the variations are also asymmetrical but different from those seen in controls: the right inferior horn is showing inferior-superior widening (blue) whereas the left horn is showing mainly an enlargement in all directions; the tensors are almost isotropic (spherical). The color code, based on the value of the largest eigen value, shows a slight elongation tendency in the posterior-anterior direction (green) and a widening in left-right direction (red).

3.3.3 Statistical shape comparison of AD vs. controls using statistical maps

For each node, the P values derived from the permutation tests have been color-coded and mapped on an average model of the brain ventricles (see Fig. 3.6). Table 3.2 gives the percentage of area found to be significantly different at α = 0.05 and α = 0.01. Figure 3.6 shows significant differences on the inferior lateral horns (left and right) and the areas adjacent to the left side of the splenium of the corpus callosum, a midsagittal area of the corpus callosum, the amygdala (left and right), the tale of the caudate nuclei (especially the left one), the head

of the left caudate nucleus, and some areas adjacent to the thalamus (on the third ventricle).

A total area covering 22% of brain ventricular surface was significantly different with $P < 0.05$ and 9% of total brain ventricles presented significance at $P < 0.01$. The left lateral ventricle presented more significant differences (14% of the total surface at $P < 0.05$) as compared to the right ventricle (7%, $P < 0.05$).

3.3.4 Local shape comparison using displacement maps

Once the local area presenting the most significant local shape differences were identified, we computed the displacement vectors that would move a node from an average control shape to an average AD shape. Figure 3.7 shows the directions of the displacement vectors for the significantly different nodes ($P < 0.05$). The length of the vectors is color-coded and mapped on the surface. In Fig. 3.8, we show both the main direction of normal variations (derived from the control group) and the directions of the control-to- Alzheimer transformation of Fig. 3.7. For most nodes, the direction control to Alzheimer is clearly not along (parallel) the normal variations (in control group); this confirms the significance of the local shape differences.

3.4 Discussion

The hippocampus and the corpus callosum are frequently studied structures in AD [6, 21, 34–43]. Some other structures such as the lateral ventricular horns have also attracted the attention of researchers [33, 35, 44]. Thompson et al. [33] presented a longitudinal study to analyze the hippocampal and ventricular changes in AD. An anatomical surface modeling approach was combined with surface-based statistics to visualize the regions of significant shape differences between the two groups at both base line and follow-up scans as well as rate of atrophy in the serial MRI scans. They concluded that although both temporal horns and hippocampus maps correlated with clinical deterioration, temporal horn expansion maps were more sensitive to AD progression. In this paper, we extended the cross-sectional part of the analysis to the whole cerebral ventricular system (with the exception of the 4^{th} ventricle) motivated by the fact that many adjacent brain structures are affected by dementia in general and AD in particular. Any change of shape and/or volume of the surrounding structure is directly reflected in the shape and volume of the ventricles.

Segmenting brain structures, such as the hippocampus, is tedious, time consuming, and often imprecise because most of these structures are relatively small and often do not present sufficient contrast or sharp borders. Furthermore, the white/gray matter image contrast in the elderly is considerably reduced as compared to young individuals, making any delineation (automatic or manual) a quite difficult and unreliable task. However, the contrast between CSF and the parenchyma remains sharp, thanks to clear CSF image intensity, irrespective of age. Therefore, the brain ventricle delineation is easier and more reliable.

Although the brain ventricles are relatively easy to delineate, their complex 3D shape is challenging to model and analyze. In this paper, we applied a fully automatic technique that is based on growing neural networks [26]. A first mesh model is generated on the average

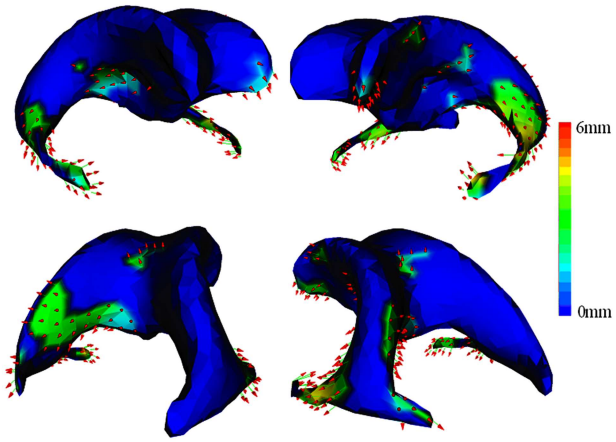


Figure 3.7: Local changes required to transform an average control shape into an average AD shape. The direction of movement is indicated by the arrows; the amplitude of movement is color coded.

brain ventricles of the control group using the unsupervised growing phase of the algorithm. After convergence, the shape is adapted to every instance of the data set (healthy and AD subjects). The point correspondence is automatically established during the adaptation phase (automatic deformation) of the mesh.

We performed permutation tests for every node of the mesh and mapped the corresponding P values with a color code to highlight the local areas with significant differences (Fig. 3.6). For these areas, we computed the displacement vectors that would move a node from an average control shape to an average AD shape (Figs. 3.7 and 3.8).

Our results show that in patients with AD, not only the lateral horns were significantly affected by the disease, but also the areas adjacent to the anterior corpus callosum, the splenium of the corpus callosum, the amygdala, the thalamus, the tale of the caudate nuclei, and the head of the left caudate nucleus.

Although most of the analysis on brain morphology in AD are performed on high-resolution T1-weighted images, the MR images used in this work are quite standard in clinical settings (T2- weighted and FLAIR, 3 mm slice thickness). When scan time is not an issue, it is obviously preferable to acquire high-resolution images for morphology analysis. It would be therefore beneficial to run our experiments on higher resolution images and compare the extent of the areas showing significant differences. On the other hand, our work shows that even on relatively low-resolution images, shape differences can be picked up when studying the brain ventricles.

To our knowledge, this is the first study to report on local shape differences of the whole cerebral ventricular system in AD patients. The most widely used biomarker for AD is the volume of the hippocampus. However, the hippocampal atrophy occurs also in other dementias [43]. In our preliminary study, we observed that apart from shape changes in the tem-

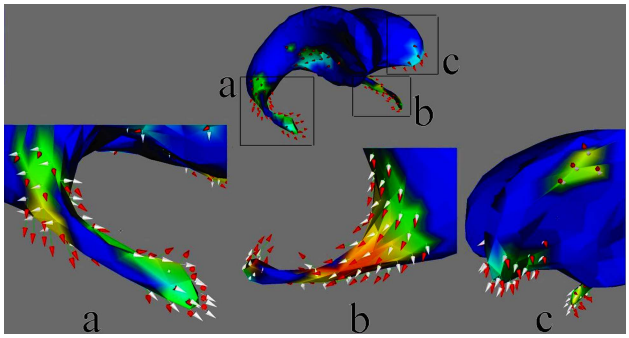


Figure 3.8: Local changes required to transform an average control shape into an average AD shape (red arrows) compared with the main direction of variation within controls (white arrows). The images show the right and left temporal horns (a)-(b), and the frontal part of the left lateral ventricle (c). Note that the arrows are not parallel, indicating different shape changes between the two groups.

poral horn, reflecting hippocampal atrophy, other brain parts of the ventricles also showed differences in shape when comparing AD patients and controls. Whether these additional ventricular shape changes permit superior differentiation of AD patients and patients with other neuro-degenerative conditions should be subject to further research.

Bibliography

[1] Freeborough PA, Fox NC. The boundary shift integral: an accurate measure of cerebral volume changes from registered repeat MRI. IEEE Trans Med Imaging. 1997;16(5):623–629.

[2] Nichols TE, Holmes AP. Nonparametric Permutation Tetsts For Functional Neuroimaging: A Primer with Examples. Human Brain Mapping. 2001;15:1–25.

[3] Pitiot A, Delingette H, Toga AW, Thompson PM. Learning Object Correspondences with the Observed Transport Shape Measure. Information Processing in Medical Imaging (IPMI). 2003;LNCS 2732:25–37.

[4] Rueckert D, Frangi AF, Schnabel JA. Automatic Construction of 3-D Statistical Deformation Models of the Brain Using Nonrigid Resitration. IEEE Transactions on Medical Imaging. 2003;22(8):1014–1025.

[5] Scheltens P, Barkhof F, Leys D, Wolters EC, Ravid R, Kamphorst W. Histopathologic correlates of white matter changes on MRI in Alzheimer's disease and normal aging. Neurology. 1995;45:883–888.

[6] Schott JM, Price SL, Frost C, Whitwell JL, Rossor MN, Fox NC. Measuring atrophy in Alzheimer diseases: a serial MRI study over 6 and 12 months. Neurology. 2005;65:119–124.

[7] Thodberg HH. Minimum Description Length Shape and Apperance Models. Information Processing in Medical Imaging (IPMI). 2003;LNCS 2732:51–62.

[8] Showell ER, Peterson BS, Thompson PM. Mapping cortical changes across the human life span. Nat Neuroscience. 2003;6:309–315.

[9] Thompson PM, Hayashi KM, Zubicary G. Dynamics of gray matter loss in Alzheimer's disease. J Neuroscience. 2003;23:994–1005.

[10] Coffey CE, Wilkinson WE, Parashos IA. Quantitative cerebral anatomy of aging human brain: a cross-sectional study using magnetic resonance imaging. Neurology. 1992;42:527–536.

[11] Raz N, Gunning FM, Head D. Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. Cereb cortex. 1997;7:268–282.

[12] Barns J, Scahill RI, Boyes RG, Frost C, Lewis EB, Rossor CL, et al. Differentiating AD from Aging using semiautomated measurement of hippocamapal atrophy rates. Neuroimage. 2004;23:574–581.

[13] Dekaban AS. Changes in brain weights during the span of human life: relation of brain weights to body heights and body weights. Ann Neurol. 1978;4:345–356.

[14] Lerch JP, Priessmer JC, Zijdenbos A, Hamperl H, Teipel SJ, Evans AC. Focal Decline of Cortical Thickness in Alzheimer's Disease Identified by Computational Neuroanatomy. Cerebral Cortex. 2004;15:995–1001.

[15] Pruessner JC, Collins DL, Pruessner M, Evans AC. Age and gender poredict volume decline in the anterior and posterior hippocampus in early adulthood. The journal of neuroscience. 2001;21(1):194–200.

[16] Armstrong CL, Traipe E, Hunter JV, Haselgrove JC, Ledakis GE, Tallent EM, et al. Age-related, regional, hemispheric, and medial-lateral differences in myelin integrity in vivo in the normal adult brain. AJNR Am J Neuroradiol. 2004;25(6):977–984.

[17] Bronge L, Bogdanovic N, Wahlumd LO. Postmortem MRI and histopathology of white matter changes in Alzheimer brains. A quantitative, comparative study. Dement Geriatr Cogn Disord. 2002;13:205–212.

[18] Spilt A, Geeraedts T, de Craen AJ, Westendorp RG, Blauw GJ, van Buchem MA. Age-related changes in normal-appearing brain tissue and white matter hyperintensities: more of the same or something else? AJNR Am J Neuroradiol. 2005;26(4):725–729.

[19] Bozzao A, Floris R, Baviera ME, Apruzzese A, Simonetti G. Diffusion and perfusion MR imaging in cases of Alzheimer's disease: correlations with cortical atrophy and lesion load. AJNR AM J Neuro-radiol. 2001;22:1030–1036.

[20] Hanyu H, Asano T, Sakurai H, Imon Y, Iwamoto T, Takazaki M, et al. Diffusion-weighted and magnetization transfer imaging of the corpus callosum in Alzheimer's disease: a quantitative study. J Neurol Sci. 1999;167:37–44.

[21] Teipel SJ, Bayer W, Alexander GE, Zebhuhr Y, Teichberg D, Kulic L, et al. Progression of corpus callosum atrophy in Alzheimer disease. Arch Neurol. 2002;59:243–248.

[22] Admiraal-Behloul F, van den Heuvel DMJ, Olofsen H, van Osch MJP. Fully automatic segmentation of white matter hyperintensities in MR images of the elderly. Neuroimage. 2005;28(3):607–617.

[23] Woods RP, Grafton ST, Holmes CJ, Chery SR, Mazziotta JC. Automated Image Registration I. General methods and intrasubject, intramodality validation. J Comput Assist Tomogr. 1998;22:139–152.

[24] Admiraal-Behloul F, van den Heuvel D, Olofsen H, Schmitz N, van Buchem MA. Brain templates for the elderly. In: International Society for Magnetic Resonance in Medicine (ISMRM) - Toronto; 2003. .

[25] Cootes TF, Taylor CJ, Cooper DH, Graham J. Active Shape Models - their training and application. Computer Vision and Image Understanding. 1995;61(1):38–59.

[26] Ferrarini L, Olofsen H, van Buchem MA, Reiber JHC, Admiraal-Behloul F. Fully automatic shape modelling using growing cell neural networks. MICCAI. 2005;LNCS 3750:451–458.

[27] Fritzke B. Growing Cell Structures - A self-organizing Network for Unsupervised and Supervised Learning. Neural Networks. 1994;7(9):1441–1460.

[28] Marsland S, Shapiro J, Nehmzow U. A self-organizing network that grows when required. Neural Networks. 2002;15:1041–1058.

[29] Fritzke B. Kohonen Feature Maps and Growing Cell Structures - a Performance Comaprison. NIPS Proc. 1992;p. 123–130.

[30] Davies RH, Twining CJ, Allen PD, Cootes TF, Taylor CJ. Shape Discrimination in the Hippocampus Using an MDL Model. Information Processing in Medical Imaging (IPMI). 2003;LNCS 2732:38–50.

[31] Davies RH, Twining CJ, Cootes TF, Waterton JC, Taylor CJ. A minimum description length approach to statistical shape modelling. Information Processing in Medical Imaging (IPMI). 2001;LNCS 2082:50–63.

[32] Üzümcü M, Frangi AF, Sonka M, Reiber JHC, Lelieveldt BPF. ICA vs. PCA Active Appearance Models: Application to Cardiac MR Segmentation. MICCAI. 2003;LNCS 2878:451–458.

[33] Thompson PM, Hayashi KM, de Zubicaray GI, Janke AL, Rose SE, Semple J, et al. Mapping hippocampal and ventricular change in Alzheimer disease. NeuroImage. 2004;22(4):1754–1766.

[34] Adachi M, Kawakatsu S, Hosoya T, Otani K, Honma T, Shibata A, et al. Morphology of the inner structure of the hippocampal formation in Alzheimer disease. Am J Neuroradiology. 2003;24(8):1575–1581.

[35] Bracco L, Piccini C, Manfredi G, Fonda C, Falcini M, Amaducci L. Magnetic resonance measures in Alzheimer disease: their utility in early diagnosis and evaluating disease progression. Alzheimer Dis Assoc Disord. 1999;13(3):157–164.

[36] Gootjes L, Bouma A, van Strien JW, van Schijndel R, Barkhof F, Scheltens P. Corpus Callosum size correlates with asymmetric performance on a dichotic listening task in healthy aging but not in Alzheimer's disease. NeuroPhycologia;44(2):208–217.

[37] Gosche KM, Mortimer JA, Smith CD, Markesbery WR, Snowdon DA. An automated technique for measuring hippocampal volumes from MR imaging studies. Am J Neuroradiology. 2001;22(9):1686–1689.

[38] Kaneko T, Kodama N, Kaeriyama T, Kawase Y, Shibukawa M, Fukumoto I. Automatic extraction of corpus callosum from midsagittal head MR image and examination of Alzheimer-type dementia objective diagnostic system in feature analysis. Nippon Hoshasen Gijutsu Gakkai Zasshi. 2004;60(2):293–298.

[39] Kantarci K, Jack CR. Neuroimaging in Alzheimer disease: an evidence-based review. Neuroimaging Clin N Am. 2003;13(2):197–209.

[40] Kodama N, Shimada T, Fukumoto I. Image-based Diagnosis of Alzheimer-type Dementia: Measurements of Hippocampal and Ventricular Areas in MR Images. Magn Reson Med Sci. 2002;1(1):14–20.

[41] Morys J, Boek-Billewicz B, Dziewiatkowaski J, Bidzan L, Ussorowska D, Narklewicz O. Changes in the volume of temporal lobe structures related to Alzheimer's type dementia. Folia Neuropathol. 2002;40(2):47–56.

[42] Petrella JR, Coleman RE, Doraiswamy PM. Neuroimaging and early diagnosis of Alzheimer disease: a look to the future. Radiology. 2003;226(2):315–336.

[43] van de Pol LA, Hensel A, van der Flier WM, Visser PJ, Pijnenburg YA, Barkhof F, et al. Hippocampal atrophy on MRI in frontotemporal lobar degeneration and Alzheimer's disease. J Neurol Neurosurg Psychiatry. 2006;77:439–442.

[44] Frisoni GB, Geroldi C, Beltramello A, Bianchetti A, Binetti G, Bordiga G, et al. Radial width of the temporal horn: a sensitive measure in Alzheimer disease. Am J Neuroradiol. 2002;23(1):35–47.

CHAPTER 4

Ventricular Shape Biomarkers for Alzheimer's Disease in Clinical MR Images

Better to add life to our days, than days to our lives
(Meglio aggiungere vita ai giorni, che giorni alla vita).

Rita Levi Montalcini, 1909- .

This chapter is reprinted from:

Magnetic Resonance in Medicine (MRM)

Authors: L. Ferrarini, W.M. Palm, H. Olofsen, R. van der Landen, M.A. van Buchem, J.H.C. Reiber, and F. Admiraal-Behloul

Title: **Ventricular Shape Biomarkers for Alzheimer's Disease in Clinical MR Images**, Accepted for publication (October 2007)

Abstract Aim of our work was to identify ventricular shape-based biomarkers in Magnetic Resonance (MR) images to discriminate between patients with Alzheimer Disease (AD) and healthy elderly. Clinical MR images were collected for 58 patients and 28 age-matched healthy controls. After normalizing all the images, the ventricular cerebrospinal fluid was semi-automatically extracted for each subject and an innovative technique for fully automatic shape modeling was applied to generate comparable meshes of all ventricles. The search of potential biomarkers was carried out with repeated permutation tests. Results highlighted well-defined areas of the ventricular surface as being discriminating features for AD: the left inferior medial temporal horn, the right medial temporal horn (superior and inferior), and the areas close to the left anterior part of the corpus callosum and the head of the right caudate nucleus. The biomarkers were then used as features to build an intelligent machine for AD detection: a Support Vector Machine was trained on AD and healthy subjects, and subsequently tested with leave-1-out experiments and validation tests on previously unseen cases. Results showed a sensitivity of 76% for AD, with an overall accuracy of 84%, proving that suitable biomarkers for AD can be detected in clinical MR images.

4.1 Introduction

Alzheimer's Disease (AD) is commonly recognized as the most common form of dementia in elderly. Due to an increased life expectancy and a high conversion rate in elderly [1], the emotional and economic costs of AD are continuously growing. Treatment strategies for AD have proved more successful if administered during the early stages of the disease [1]: thus, biomarkers for early diagnosis of AD are highly desirable.

Attempts to detect biomarkers have been previously presented in the literature. Hansson et al. [1] investigated different proteins, and proved how their concentrations in CereboSpinal Fluid (CSF) can be correlated with the progression of AD: the CSF samples were acquired with lumbar punctures, a highly invasive procedure. Leon et al. [2] stated that volumetric changes of hippocampal atrophy and medial temporal atrophy, although associated with AD progression, are not specific for AD: therefore, they suggested to combine neuroimaging techniques (Magnetic Resonance Imaging (MRI) and Fluorodeoxyglocose-Positron Emission Tomography (FDG-PET)) with CSF biomarkers for a more sensitive early detection of AD in both cognitively normal subjects and in patients with Mild Cognitive Impairment (MCI). Even in this case, the need for CSF biomarkers implies an invasive procedure. Finally, Kantarci et al. [3] tackled the problem of noninvasive techniques, providing an overview of quantitative MR measurements which can help assessing the progression from a normal elderly to AD, passing through the MCI state. Particularly interesting is their conclusion on how different subcortical structures undergo atrophy during the progression of the disease: hippocampal atrophy develops rapidly in the early stage, and tends to slow down even before AD is diagnosed; temporal lobe atrophy is more characteristic of the MCI to AD conversion, while frontal lobe atrophy is characteristic of the last AD phase.

Clinical MR images of the brain can be easily obtained nowadays, and provide detailed insights on brain structures in a noninvasive way. Challenging questions arise from these insights. Are there markers in clinical images which characterize AD? Where are they located? How sensitive are they in discriminating AD from healthy subjects? Alzheimer Disease is known to cause atrophy in periventricular structures; a rational choice would therefore be to analyze shape and volume changes in these areas, with shape changes being generally more accurate than volume changes. Shape changes of hippocampus and ventricles have been associated with AD progression [4, 5]. Unfortunately, the contrast between white and gray matter tends to decrease with age, making it difficult to correctly delineate periventricular gray matter structures from their white matter surroundings in elderly. On the other hand, the contrast between CSF and brain parenchyma remains mostly unaffected: thus, the analysis of ventricular shape changes is more reliable. As an example, in [6] the authors studied ventricular shape differences to successfully discriminate healthy subjects and patients with schizophrenia.

In this work, we focus on ventricular shape-based biomarkers for AD detection. Volume and shape changes in periventricular structures are reflected on the volume and shape of the brain ventricles: thus, by localizing differences in the ventricles, one can draw conclusions about which periventricular structures are mostly affected by the disease. In previous works, we introduced a new method for shape modeling and analysis, Growing and Adaptive MEshes (GAMEs) [7], and applied it to the analysis of brain ventricular shapes in populations of

healthy elderly and patients with AD [8]. In this study, GAMEs and repeated permutation tests are used to prove the existence of well-defined ventricular areas which consistently discriminate AD subjects from healthy elderly, showing the suitability of clinical MR for AD detection. After detecting such biomarkers, these features will be used to train a Support Vector Machine (SVM): by assessing the accuracy, generalization, and sensitivity properties of the SVM, the feasibility of an intelligent machine able to detect AD from clinical MR images will be demonstrated.

4.2 Material and Method

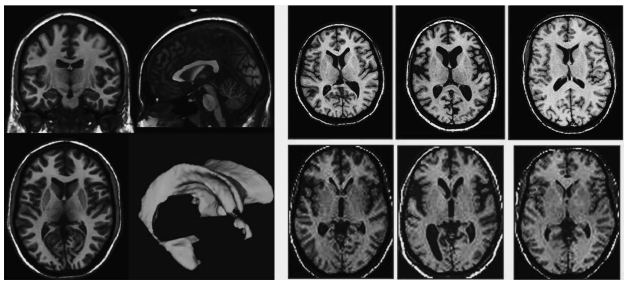
4.2.1 Subjects and MRI acquisition

Fifty-eight patients with probable AD (27 men, mean age 74 years, age range 60- 95 years), and 28 volunteers with normal cognitive functions (12 men, mean age 74 years, age range 64-89 years) were included in the study. The patients with probable AD had been consecutively referred to our outpatient memory clinic. Volunteers were recruited through advertisements in local newspapers. All subjects were evaluated for memory loss using a standardized dementia screening including a detailed medical history, a general internal and neurological exam, laboratory tests, neuropsychological testing including the Mini Mental State Examination (MMSE), and MRIs of the brain. Diagnoses were made in a multidisciplinary consensus meeting based on the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria for probable AD. Patients and controls were included if they: were older than 60 years, had no other neurologic or psychiatric illness, and had no abnormalities on MRI other than white matter hyperintensities or an incidental small lacunar lesion ($\leq$5 mm diameter). The study was approved by the local Medical Ethical Committee. Written informed consent was obtained from all subjects or from a close relative if a patient was demented.

MRI was performed on a 1.5-T MR-system (Philips Medical Systems, Best, The Netherlands) using the following pulse sequences: dual fast spin-echo (proton density and T2 weighted): time to echo (TE) 27/120 ms, repetition time (TR) 3000 ms, 48 contiguous 3-mm slices without an interslice gap, matrix 256 x 256, field of view (FOV) 220 mm. FLAIR (fluid attenuated inversion recovery): TE 100 ms, TR 8000, 48 contiguous 3-mm slices without interslice gap, matrix 256 x 256, FOV 220 mm.

4.2.2 Pre-processing steps

In-house developed software for automatic-segmentation (SNIPER, Software for Neuro-Image Processing in Experimental Research [9]) was used to pre-process the images: the software extracted fully automatically the intracranial cavity, the CSF, and the white matter hyperintensities. Interactive 3D tools based on region growing allowed the re-labeling of ventricular CSF as *ventricles*. From each ventricular volume, the surface was extracted (see Fig. 4.1.a). Before analyzing shape differences between populations, normalization is needed to correct for brain-size and orientation. All images were automatically registered into a stereotaxic



(a) (b)

Figure 4.1: (a) The ventricular CSF is extracted semi-automatically, and the ventricle surface is reconstructed from the volume. (b) An affine 12-parameters registration brings all the original images (first row) into a common stereotaxic space, removing differences due to brain-size and orientation (second row, example given on T1 images).

space, the LUMC T2-weighted brain template for geriatrics [10], using automatic affine 12-parameters image registration [11] (see Fig. 4.1.b). The same procedure was applied to the segmented ventricular CSF.

4.2.3 Shape Modeling and Analysis

The local analysis of shape differences between two populations requires, as a first step, the modeling of all the instances' shapes. For this purpose, we used the GAMEs method described in [7], and summarized in Fig. 4.2. An average shape is created from the control population: a voxel-count map is built up by adding all controls' binary volumes; subsequently, a threshold is applied to the voxel-count map, obtaining an average control ventricle. During the first step of GAMEs, a mesh grows (adding nodes and edges) until it converges to a good representation of the average control shape. Once the first mesh is obtained, the number of nodes and edges is frozen. In a second phase, the mesh is adapted to all the instances in the two populations using the Kohonen self-organizing map algorithm [12] (see Fig. 4.3). Since nodes are not added nor removed, all the final meshes are *comparable*: each node in a given mesh is uniquely associated with another node in another mesh, allowing comparison between meshes. Comparable meshes are necessary for shape analysis, but not sufficient: corresponding nodes in comparable meshes have to be representative of similar anatomical locations. A thorough analysis of our method was performed in [7, 8], following the suggestions given by Davies et al. [13, 14]: GAMEs proved successful in terms of reconstruction error, generalization error, compactness, specificity error, as well as in detecting similar anatomical landmarks and highlighting local shape differences between populations.

Local shape differences between controls and AD were investigated in [8]: for each node in the meshes, permutation tests were applied, resulting in a p-value indicating the probability that the local shape in the two populations was drawn from the same distribution in space. A p-value lower than 0.05 indicated that, at a confidence level of 95%, the local shape was significantly different between controls and ADs. Since permutation tests were applied to

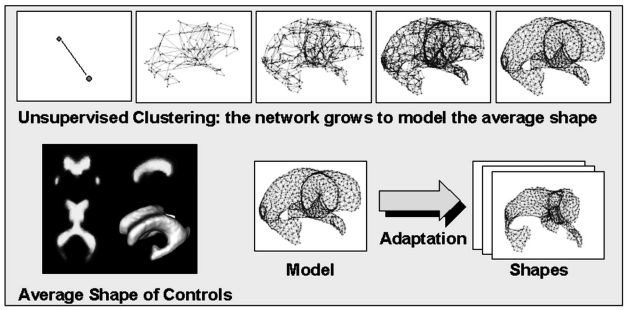


Figure 4.2: GAMEs approach: (1) an average shape is built up from the Control population; (2) through an unsupervised clustering algorithm, a mesh grows and adapts to model the average shape; (3) the mesh is adapted to all the controls and AD shapes, without adding nor removing nodes.

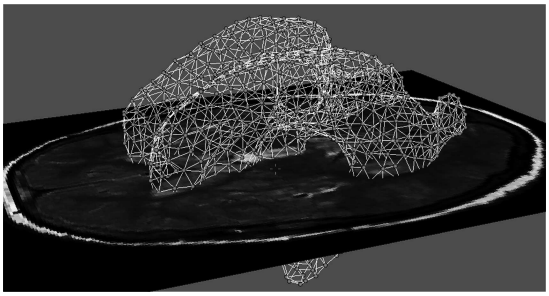
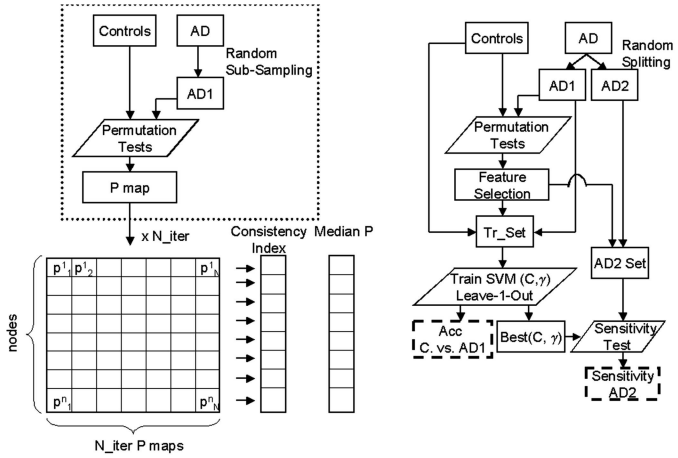


Figure 4.3: A mesh is adapted to a new instance in the population to represent the ventricular surface.

all the subjects in the two groups, the ventricular regions highlighted in [8] were considered characteristics for AD in those particular datasets. In this work, we expand the AD dataset, and perform *repeated* permutation tests on randomly sub-sampled AD groups: thus, regions consistently highlighted through the iterations can be safely considered as representative of the pathology (AD).

When applied to the subjects of this study, our method provided us with 86 meshes (28 Controls and 58 AD) of 949 nodes: thus, each ventricle is a point in a feature space of dimension 3 x 949 (three coordinates for each node). After detecting proper biomarkers, the dimensionality can be reduced, simplifying the training of a classifier.



(a) (b)

Figure 4.4: (a) Biomarker Selection: Permutation Tests (PT) can be used to determine a feature set discriminating between Controls and AD. In order to assess the consistency of the results, PT were run several times and a consistency index was evaluated for each node. (b) Design of the SVM: training and testing. After randomly splitting the AD group in two sub-groups, the SVM is trained on Controls/AD1 and tested on other dataset combinations.

4.2.4 Biomarker Selection

Our first goal was to detect consistent biomarkers for AD. An elegant guess is to start with those nodes (i.e. areas) highlighted as significantly different by permutation tests. A proper biomarker set must reflect differences due to the disease, rather than a casual pattern due to the particular subjects included in the study. In order to guarantee the consistency of the nodes highlighted by the permutation tests, we set up the following experiment (schematically represented in Fig. 4.4.a):

1. For $N_{iter} = 10$ times, randomly select 50% of ADs to generate an AD1 group. Run permutation tests on Controls-AD1, generating N_{iter} P-value maps;

2. In the final matrix, each row is associated with a node in the mesh, while each column reports the corresponding p-value for a particular iteration;

3. In each iteration (column), a node is considered significantly different between Controls and AD if:

   - its $p \leq P_{thr}$ or
   - one of its neighbors (up to 2 links away in the mesh) had $p \leq P_{thr}$;

4. For each node (row), evaluate the percentage of iterations in which it was significantly different (*Consistency Index*);

5. Finally, for each node (row) assess the *median* of all the p-values through the N_{iter} iterations.

If nodes highlighted as significantly different by permutation tests are actually discriminating subjects depending on their healthy/AD condition, their p-values should not be affected by the random sub-sampling through the N_{iter} iterations. Therefore, a given node should be either significantly different or not significantly different in most of the iterations: we define such a node (characterized by a consistency index either lower than 20% or higher than 80%) as a *consistent node*. Within the consistent nodes, we define as *biomarkers* those which are significantly different between Controls and ADs. Depending on the P_{thr} threshold, the amount of consistent nodes and biomarkers will change: ideally, one aims at having a high percentage of consistent nodes, with a low percentage of biomarker nodes (fewer biomarker nodes will lead to a smaller feature space in which training the classifier).

Different P_{thr} thresholds were used (0.05, 0.01, and 0.001) to assess the best set of shape-based biomarkers. Results (reported in Table 4.1 and presented in section 4.3.1) led us to two consistent sets. It is important to notice that the consistency of almost the entire ventricular surface through the iterations (obtained with P_{thr} set to 0.01 and 0.001) proves the independence of the biomarker sets (i.e. feature sets) from the particular sub-groups used to generate these: regardless how AD is sub-sampled, similar locations are identified as biomarkers.

4.2.5 AD Detection with SVM

The second goal of this study was to use the biomarker sets to build an intelligent machine (i.e. classifier) for AD detection in clinical MR images. During a supervised training phase, the classifier learns how to discriminate between healthy subjects and ADs; subsequently, new subjects can be classified by the machine. Important properties for a classifier are:

- *Accuracy (Separability in the feature space)*: successful classification rate while classifying the objects of the training set;

Table 4.1: Consistent nodes and Biomarker nodes. For three different P_{thr} thresholds, we evaluated the amount of consistent nodes and biomarker nodes (consistent nodes which are significantly different).

	% of Consistent Nodes	% of Biomarker Nodes
$p \leq 0.05$	70%	42%
$p \leq 0.01$	80%	14%
$p \leq 0.001$	90%	4%

- *Generalization*: successful classification rate while classifying subjects previously not used in the training phase (assessable with leave-1-out tests);

- *Sensitivity and Specificity*: successful classification rate while classifying previously unseen ADs and Controls respectively.

SVM's Description

A Support Vector Machine (SVM) is a well-known pattern recognition tool for data classification [15, 16]. Given a training set of l data-labels pairs $\{(\mathbf{x_1}, y_1), ..., (\mathbf{x_l}, y_l)\}$, with $\mathbf{x_i} \in R^n$ (e.g., a subject in our populations), and $y_i \in \{-1, 1\}$ (e.g., the healthy/AD status), the SVM seeks for a hyperplane which separates the datasets, minimizing the classification error. When a perfect separation of the training set can not be found, the SVM seeks for a sub-optimal solution which allows misclassifications of some subjects $\mathbf{x_i}$ in the training set: the misclassifications are weighted by a positive constant term $C > 0$. After training, a SVM can be considered as a decision-box which accepts a given subject $\mathbf{x_i}$, and associates it to the corresponding status.

Linearly separation of the training set in the original input space might not provide the desired results: thus, non-linear solutions might be necessary. Non-linear classifiers use a kernel function of the data $K(\mathbf{x_i}, \mathbf{x_j})$ to transform the input space into a high-dimensional space in which the desired hyperplane can be found: the resulting hyperplane in the transformed space will translate into a non-linear classifier in the original space. We implemented the Radial Basis Functions (RBFs) kernel, defined as:

$$K(\mathbf{x_i}, \mathbf{x_j}) = \exp(-\gamma||\mathbf{x_i} - \mathbf{x_j}||^2), \quad \gamma > 0. \tag{4.1}$$

Thus, the SVM is fully characterized by the two parameters C and γ, tuned during the training phase (see Appendix 4.5 for more details). A non-linear classifier might be more prone to over-fitting, especially when trained on a limited number of examples: therefore, we experimented also with a linear SVM. Since the linear classifier did not present any significant improvement (see Appendix 4.5), in the remainder of the paper we focus only on the RBFs-based classifier. A more complete overview of the SVM's mathematical background can be found in [17].

SVM's Training and Testing

Given a certain feature set (i.e. biomarker set), each mesh in the dataset can be represented by a subset of its nodes in a high-dimensional feature space. The next step is to train a SVM to discriminate between Controls and ADs: better feature sets will lead to a more discriminative classifier. The procedure for training and testing is schematically reported in Fig. 4.4.b, and follows the suggestions of Hsu et al.: *A Practical Guide to Support Vector Classification*

(*citeseer.ist.psu.edu/689242.html*)[1]. After properly scaling the data[2], we:

1. randomly split the AD group into two equally-sized groups: AD1 and AD2;
2. apply permutation tests on Controls vs. AD1 to detect the features associated with a given P_{trh};
3. build up a training set Tr_Set with Controls and AD1 (using the feature sets of point 2);
4. train and tune the SVM on Tr_Set;
5. assess the SVM's accuracy on Controls-AD1;
6. assess the SVM's generalization property (leave-1-out experiment on Controls-AD1);
7. assess the SVM's sensitivity on AD2.

The training and testing was performed for $P_{thr} = 0.01$ and $P_{thr} = 0.001$. Results are reported in Table 4.2 and presented in section 4.3.2. It is worth commenting on point 2 of the training and testing procedure. After splitting the AD group in AD1 and AD2, we apply permutation tests on Controls and AD1 to identify the appropriate feature set. Two critics could be made to this step. One could argue that using the Controls and AD1 both for feature selection and for the assessment of accuracy and generalization properties might bias the results: nevertheless, following the method presented in section 4.2.4 and the results of section 4.3.1, one can conclude that for $P_{thr} = 0.01$ and $P_{thr} = 0.001$ the biomarker sets are consistent regardless of the sub-sampled AD group. Thus, the bias on the results, if any, would be minimal. Considering this, one could argue that it is not necessary to perform permutation tests again and that we could use the results of the repeated permutation tests of section 4.2.4. Although this is true, we still prefer to avoid any potential minimal bias while

[1] All the experiments were performed using the libsvm library (Chang et al., *LIBSVM: a Library for Support Vector Machines: http://www.csie.ntu.edu.tw/ cjlin/papers/libsvm.pdf*) in MatLab 7.0.1.

[2] Scaling data is highly recommended. Features ranging over large numeric domains might dominate over those presenting smaller range: linearly scaling all the features in a prefixed interval, e.g. $[-1, 1]$, might help preventing this problem.

Table 4.2: Performances of the SVM-RBFs: results show the separability in the future space (i.e. Accuracy, evaluated on Controls-AD1), the generalization capability (leave-1-out experiment on Controls-AD1) reported as Success-Rate (Specificity/Sensitivity), and the sensitivity on unseen pathological cases (AD2).

	Separability in the Feature Space	Generalization (Specificity/Sensitivity)	Sensitivity (unseen dataset)
$p \leq 0.01$	100%	88% (89% / 86%)	66%
$p \leq 0.001$	93%	84% (82% / 86%)	76%

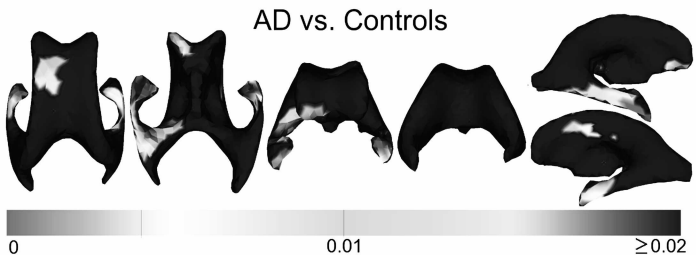


Figure 4.5: The color-coded map shows, for each location, the *median* p value evaluated over the N_{iter} runs ($p > 0.01$ are shown in blue). The corresponding colored Figure can be found at page viii, before the Introduction.

evaluating the sensitivity on AD2, since the detection of AD (sensitivity) has the highest priority within this work (more details can be found in the Appendix 4.5).

4.3 Results

4.3.1 AD Biomarkers

Results for biomarker selection with repeated permutation tests are reported in Table 4.1. The case with $p \leq 0.05$ does not provide a good feature set: only 70% of the nodes are consistent, and the high percentage of biomarker nodes would lead to a high-dimensional feature space. The other two cases, $p \leq 0.01$ and $p \leq 0.001$, provide valid feature sets: they both present high percentages of consistent nodes, and a sufficiently small number of biomarker nodes (about 130 and 30 nodes respectively).

The areas which better discriminate Controls and AD can be visualized: in Fig. 4.5, each location is color-coded with the *median* p-value evaluated over the N_{iter} runs. From the map, it is clear that the most discriminating areas of the ventricles are the left inferior medial temporal horn, the right (superior and inferior) medial temporal horn, and the areas close to the genu of the left side of the corpus callosum and the head of the right caudate nucleus. Thus, we have detected two potentially good biomarker sets which can be used as features during the training and testing of the SVM: the best biomarker set will lead to the best SVM performances.

4.3.2 Biomarker's Performances

As reported in Table 4.2, both scenarios ($p \leq 0.01$ and $p \leq 0.001$) present high performances in terms of separability in the feature space (i.e. accuracy, 100% and 93% for $p \leq 0.01$ and $p \leq 0.001$ respectively) and generalization (specificity 89% and 82%, sensitivity 86% and 86%, for $p \leq 0.01$ and $p \leq 0.001$ respectively); eventually, one is mostly interested in detecting AD: the second test run on AD2 (sensitivity on unseen pathological cases) showed

that the best performances are given by $p \leq 0.001$, where a sensitivity of 76% is reached: the corresponding areas are the inferior (left and right) medial temporal horns. A final note on the results for the separability in the feature space: the good performances on this test strengthen the hypothesis that permutation tests correctly highlight features on the ventricular surface which strongly characterize the two groups.

4.4 Discussion

In this work, we have introduced a new framework to detect ventricular shape biomarkers for AD, in clinical MR images. Shape modeling techniques, statistical nonparametric tests, and support vector machines were merged together to prove the existence of such biomarkers and to assess their performances. The idea of complementing shape analysis and pattern recognition in one solution had already been explored in the literature: Golland et al. [18] showed how, given two populations of shapes, it is possible to extract features, build a classifier-function, and move back from the function to the shape differences in the image domain. In our work, we proceeded in a different way: after having modeled all the ventricular shapes with comparable meshes, repeated permutation tests were run to prove the consistency of shape-based biomarkers differentiating ADs from healthy subjects. Once the consistency had been proved, we assessed the biomarker's performances, using them to build a classifier which was trained to recognize Controls and ADs. Different sub-sets of features were tested, and the most accurate one was selected as a biomarker set for AD.

One of the most common problems in pattern recognition is the *curse of dimensionality*. When the feature space has a too high dimension compared to the number of available samples, the training of the classifier becomes a very difficult task: the risk of over-fitting increases and the generalization capabilities are potentially reduced. In our work, we had a relatively small dataset (86 instances in total, 28 controls and 58 AD). Nevertheless, the SVMs could provide good accuracies for all the experiments (leave-1-out experiment, and sensitivity-test): as clearly explained in [17], SVMs are less subject to the *curse of dimensionality* compared to other classifiers. Moreover, by setting the threshold on the p-values at 0.001, the feature space was fully described by only 30 nodes: under these conditions, our

Table 4.3: Performances of the Linear-SVM: results show the separability in the future space (i.e. Accuracy, evaluated on Controls-AD1), the generalization capability (leave-1-out experiment on Controls-AD1) reported as Success-Rate (Specificity/Sensitivity), and the sensitivity on unseen pathological cases (AD2).

	Separability in the Feature Space	Generalization (Specificity/Sensitivity)	Sensitivity (unseen dataset)
$p \leq 0.01$	96%	86% (82% / 90%)	66%
$p \leq 0.001$	93%	84% (82% / 86%)	76%

dataset was sufficiently large to train the classifier.

The results of our analysis were very promising, and allowed us to some further considerations. The repeated permutation tests showed a pattern of shape-based features which was consistent regardless of the chosen training set. This strongly supports the idea that stable biomarkers for AD can be found in the shape of the brain ventricles: particularly important areas are the left inferior medial temporal horn, the right inferior and superior medial temporal horn, and the areas close to the left anterior part of the corpus callosum and the head of the right caudate nucleus.

The second important contribution of our work was the assessment of features' reliability, by means of a SVM. Different sub-sets of the pattern highlighted by the permutation tests were analyzed to determine the most accurate biomarker set. Both for $p \leq 0.01$ and $p \leq 0.001$, the analysis showed a common trend: the leave-1-out experiments gave excellent results (88% and 84% successful rate), and the sensitivity on unseen AD cases resulted in 66% and 76% for $p \leq 0.01$ and $p \leq 0.001$ respectively. Despite the limited number of datasets in the populations, shape-based biomarkers could successfully discriminate healthy subjects and ADs, particularly when the inferior (left and right) medial temporal horns were used ($p \leq 0.001$).

4.4.1 Comparison with other methods

When comparing our results with alternative methods, there are two aspects to consider: which biomarkers we have found, and what is the diagnostic accuracy of such biomarkers.

The shape-based biomarkers identified by our method can be grouped in four main areas: (1) the left inferior medial temporal horn, (2) the right inferior and superior medial temporal horn, (3) the area close to the left anterior part of the corpus callosum, and (4) the head of the right caudate nucleus. These results are in agreement with the recent publications on AD progression. In [19], Kodama et al. showed that the hippocampal area index was significantly different between controls and AD (the same was true for the ventricular area index). In Wang et al. [20], it was shown that different patterns of hippocampal shape change, and different rate of hippocampal volume loss can distinguish mild AD from healthy subjects. In [21], Barnes et al. investigated subjects with AD and Frontotemporal Lobe Degeneration (FTLD): they showed that volume loss rate of both hippocampus and cingulate gyrus are useful to discriminate between healthy subjects and AD, and healthy subjects and FTLD. Similar conclusions about the hippocampus were drawn in [22] and [23], where the authors also reported an important role for the (left) amygdala volume loss rate in differentiating between healthy subjects, AD, and FTLD. Finally, Thompson et al. [24] have investigated cortical changes in longitudinal studies, showing that cortical deficits on the right hemisphere spread from the temporal-parietal and entorhinal areas to the cingulate gyrus (while a more global cortical loss is reported on the left side). Our method, indirectly, pinpointed to the same regions: the left inferior medial temporal horn reflects atrophy in the left hippocampus and left amygdala; the right medial temporal horn reflects atrophy in the right hippocampus; finally, the areas close to the left genu of the corpus callosum and head of caudate nuclei might reflect atrophy in the cingulate gyrus. The advantage of our method is that it is based on ventricular CSF: thanks to the good image contrast, the ventricle segmentation can be

accurately performed semi-automatically even in clinical MR images (as reported in this study), and is more reliable and less time-consuming than the delineation of small gray/white matter structures. Moreover, the analysis of one single structure (the brain ventricles) directly highlights local atrophy in all the periventricular structures.

The second important aspect to consider is the accuracy with which our biomarkers differentiate healthy subjects from AD. In most of the papers previously cited, the authors reported the sensitivity (accuracy in detecting AD) and specificity (accuracy in detecting healthy subjects) associated with their biomarkers. Kodama et al. [19] reported a sensitivity/specificity of 90.2%/81.8% (overall accuracy of 87.7%) for their combined Volume - Area - Index and Hippocampus - Area - Index; although the results look very promising, the discriminant analysis, based on Mahalanobis distance from the mean values, was performed on the same dataset (51 AD and 22 controls) used to obtain the mean values and the standard deviations. In [20], Wang et al. reported a sensitivity of 83.3% and a specificity of 84.6% (overall accuracy of about 84%) for their longitudinal study based on 18 AD and 26 controls: in this case, the results were based on a leave-1-out experiment. Finally, Barnes et al. reported a specificity of 80% and sensitivities of 65% for atrophy rate of right amygdala volume, 80% for atrophy rate of left amygdala volume, 90% for atrophy rate of left and right cingulate gyrus, and about 93% for atrophy rate of left and right hippocampus [21, 23]; these values were obtained by using regression models on datasets of 10 AD and 11 controls. The accuracies obtained with our method are certainly in line with the previously reported results, and present some advantages. Considering the cases $p < 0.01$ and $p < 0.001$, the accuracy for the leave-1-out experiment (Contr. vs AD1) is about 86% (in average), slightly better than the overall accuracy reported by Wang et al. [20] (the only study in which we found a clear reference to the leave-1-out test). One of the main contributions of our work is that the sensitivity of 76% was evaluated on a previously unseen dataset of ADs, in contrast to what reported in all the other studies.

In conclusion, our results suggest that biomarkers for AD can be detected in clinical MR images: although our datasets are larger than those reported recently in the literature, future work should aim at larger populations to improve the detection of the biomarkers and to build stronger machines for the detection of AD. Moreover, different biomarker sets and classifiers might be investigated for different pathological groups, like MCI converters versus MCI non converters: such a machine would be extremely helpful in clinical routine to decide which patients should be treated for AD.

4.5 Appendix

The SVM presented in this work is a non-linear classifier based on RBFs. A non-linear classifier is more prone to over-fitting, especially when trained on a relatively small dataset. Linear classifiers, on the other hand, are less prone to over-fitting, but might provide worse performances.

Thus, we implemented a SVM linear classifier and evaluated its performances. Results are shown in Table 4.3. As expected, the separability in the feature space gets slightly worse. The generalization capability is the same for $p \leq 0.001$, and gets slightly worse for $p \leq 0.01$

Table 4.4: Performances of the RBFs- and Linear-SVM with feature selection based on Controls-AD1 and testing based on Controls-AD2.

	RBF		Linear	
	Separability in the Feature Space	Generalization (Spec./Sens.)	Separability in the Feature Space	Generalization (Spec./Sens.)
$p \leq 0.01$	100%	79% (82%/76%)	86%	77% (79%/76%)
$p \leq 0.001$	100%	79% (71%/86%)	95%	79% (75%/83%)

(from 88% to 86%, with sensitivity of the leave-1-out increasing from 86% to 90%); finally, the sensitivity on the unseen dataset does not change. The results are comparable, suggesting that the SVM based on RBFs does not suffer for over-fitting.

4.5.1 Tuning of the SVM

Two parameters need to be tuned for a SVM based on RBFs: the constant $C > 0$, and the γ parameter of the kernel function. The procedure and values used for tuning are the same as suggested by Hsu et al.: *A Practical Guide to Support Vector Classification* (*citeseer.ist.psu.edu/689242.html*).

Given a training set Tr_Set, a grid search is performed over the C x γ space:

1. $\forall$ (C,γ) $\in [2^{-5}, 2^{-3}, 2^{-1}, .., 2^{15}]x[2^{-15}, 2^{-13}, 2^{-11}, .., 2^{3}]$
   (a) assess the generalization capability of the SVM(C,γ) with leave-1-out over Tr_Set;
   (b) store the success rate for (C, γ): $succRate_{(C,\gamma)}$;
2. Select the best $(\overline{C}, \overline{\gamma}) = argmax_{(C,\gamma)}$ $(succRate_{(C,\gamma)})$;
3. Perform a refined grid search (as in point 1) in the surrounding of $(\overline{C}, \overline{\gamma})$: e.g., if $\overline{C} = 2^{-3}$ and $\overline{\gamma} = 2^{-13}$, search within $[2^{-5}, 2^{-1}]$x$[2^{-15}, 2^{-11}]$ with steps equal to one tenth of the intervals;
4. Select the final best $(\overline{C}, \overline{\gamma})$;
5. Train the SVM($\overline{C}, \overline{\gamma}$) on the entire training set.

In case of a linear SVM, only C needs to be tuned. The results for the RBFs SVM are $(\overline{C}, \overline{\gamma}) = (2, 0.3125)$ and $(\overline{C}, \overline{\gamma}) = (128, 4.8828 * 10^{-4})$ for $p \leq 0.01$ and $p \leq 0.001$ respectively.

4.5.2 Training and Testing

While training and testing a classifier, results should not be biased, nor based on too small datasets. Ideally, one would want to perform feature selection and tuning/training on one training set, and validate the results on an independent testing set. Both training set and testing set should contain enough objects per category, to make sure the classifier does not over-fit on one class. When not enough data is available, leave-1-out tests can be used.

In our work, we had 28 controls and 58 ADs. Because of the limited number of controls, we had to compromise on the experiments, still trying to maintain unbiased results. Priority was given to the sensitivity test on the dataset of unseen AD. Therefore, we first split the AD group in two (AD1 and AD2), and kept the AD2 for testing. Since we have shown in this work that biomarker sets are rather independent by the sub-sampled AD group, we are quite confident that feature selection on Controls-AD1 does not bias the results of the validation test on Controls-AD1. Moreover, this gives us the opportunity of testing the sensitivity on an entire dataset of unseen objects (AD2).

A different choice could have been made: selecting the features on Controls-AD1, and validate the SVM on Controls-AD2 (removing the bias for the AD subjects). Nevertheless, the sensitivity could then be tested only with leave-1-out tests on Controls-AD2. We performed these experiments for both the RBFs- and linear-SVM. Results are shown in Table 4.4. The results are quite similar to those obtained with the original framework, showing that selecting features on Controls-AD1 and testing on Controls-AD1 does not lead to biased results.

Bibliography

[1] Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L. Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impariment: a follow-up study. Lancet Neurol. 2006;5:228–234.

[2] Leon MJ, DeSanti S, Zinkowski R, Mehta PD, Pratico D, Segal S, et al. MRI and CSF studies in the early diagnosis of Alzheimer's disease. Journal of Internal Medicine. 2004;256:205–223.

[3] Kantarci K, Jack CR. Quantitative Magnetic Resonance Technique as Surrogate Markers of Alzheimer's Disease. NeuroRX: The Journal of the American Society for Experimental NeuroTherapeutics. 2004;1:196–205.

[4] Thompson PM, Hayashi KM, de Zubicaray GI, Janke AL, Rose SE, Semple J, et al. Mapping hippocampal and ventricular change in Alzheimer disease. NeuroImage. 2004;22(4):1754–1766.

[5] Apostolova LG, Dutton RA, Dinov ID, Hayashi KM, Toga AW, Cummings JL, et al. Conversion of Mild Cognitive Impairment to Alzheimer disease Predicted by Hippocampal Atrophy Maps. Arch Neurol. 2006;63:693–699.

[6] Styner M, Lieberman JA, McClure RK, Weinberger DR, Jones DW, Gerig G. Morphometric analysis of lateral ventricles in schizophrenia and healthy controls regarding genetic and disease-specific factors. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(13):4872–4877.

[7] Ferrarini L, Olofsen H, Palm WM, van Buchem MA, Reiber JHC, Admiraal-Behloul F. GAMEs: Growing and Adaptive Meshes for Fully Automatic Shape Modeling and Analysis. Medical Image Analysis. 2007;11(3):302–314.

[8] Ferrarini L, Palm WM, Olofsen H, van Buchem MA, Reiber JHC, Admiraal-Behloul F. Shape differences of the brain ventricles in Alzheimer's disease. NeuroImage. 2006;32(2):1060–1069.

[9] Admiraal-Behloul F, van den Heuvel DMJ, Olofsen H, van Osch MJP. Fully automatic segmentation of white matter hyperintensities in MR images of the elderly. Neuroimage. 2005;28(3):607–617.

[10] Admiraal-Behloul F, van den Heuvel D, Olofsen H, Schmitz N, van Buchem MA. Brain templates for the elderly. In: International Society for Magnetic Resonance in Medicine (ISMRM) - Toronto; 2003. .

[11] Woods RP, Grafton ST, Holmes CJ, Chery SR, Mazziotta JC. Automated Image Registration I. General methods and intrasubject, intramodality validation. J Comput Assist Tomogr. 1998;22:139–152.

[12] Kohonen T. The Self-Organizing Map. Proceedings of the IEEE. 1990;78(9):1464–1480.

[13] Davies RH, Twining CJ, Cootes TF, Waterton JC, Taylor CJ. A minimum description length approach to statistical shape modelling. Information Processing in Medical Imaging (IPMI). 2001;LNCS 2082:50–63.

[14] Davies RH, Twining CJ, Allen PD, Cootes TF, Taylor CJ. Shape Discrimination in the Hippocampus Using an MDL Model. Information Processing in Medical Imaging (IPMI). 2003;LNCS 2732:38–50.

[15] Boser B, Guyon I, Vapnik V. A training algorithm for optimal margin classifiers. In Fifth Annual Workshop on Computational Learning Theory - Pittsburgh, ACM. 1992;p. 144–152.

[16] Vapnik V. The Nature of Statistical Learning Theory. Springer-Verlag, New York. 1995;.

[17] Burges CJC. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery. 1998;2:121–167.

[18] Golland P, Grimson WEL, Shenton ME, Kikinis R. Deformation Analysis for Shape Based Classification. Information Processing in Medical Imaging (IPMI). 2001;LNCS 2082:517–530.

[19] Kodama N, Shimada T, Fukumoto I. Image-based Diagnosis of Alzheimer-type Dementia: Measurements of Hippocampal and Ventricular Areas in MR Images. Magn Reson Med Sci. 2002;1(1):14–20.

[20] Wang L, Swank JS, Glick IE, Gado MH, Miller MI, Morris JC, et al. Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. NeuroImage. 2003;20:667–682.

[21] Barnes J, Godbolt AK, Frost C, Boyes RG, Jones BF, Scahill RI, et al. Atrophy rates of the cingulate gyrus and hippocampus in AD and FTLD. Neurobiology of Aging. 2007;28:20–28.

[22] Tarroun A, Bonnefoy M, Bouffard-Vercelli J, Gedeon C, Vallee B, Cotton F. Could linear MRI measurements of hippocampus differentiate normal brain aging in elderly persons from Alzheimer disease? Surg Radiol Anat. 2007;29(1):77–81.

[23] Barnes J, Whitwell JL, Frost C, Josephs KA, Rossor MN, Fox NC. Measurements of the Amygdala and Hippocampus in Pathological Confirmed Alzheimer Disease and Frontotemporal Lobar Denegeneration. Arch Neurol. 2006;63:1434–1439.

[24] Thompson PM, Hayashi KM, Dutton RA, Chiang MC, Leow AD, Sowell ER, et al. Tracking Alzheimer's Disease. Annals of the New York Academy of Sciences: Special Issue on Imaging and the Aging Brain. 2006;.

CHAPTER 5

Variation in Ventricular Shape between Cognitively Normal, Memory Complainers, Mild Cognitive Impairment and Alzheimer's Disease

Of all the things I have lost, I miss my mind the most.

Mark Twain, 1835-1910.

This chapter is adapted from:

Variation in Ventricular Shape between Cognitively Normal, Memory Complainers, Mild Cognitive Impairment and Alzheimer's Disease

Authors: W.M. Palm, L. Ferrarini, H. Olofsen, G.J. Blauw, R.G.J. Westendorp, E.L.E.M. Bollen, H.A.M. Middelkoop, J.H.C. Reiber, J. van der Grond, F. Admiraal-Behloul, and M.A. van Buchem

Manuscript in preparation

Abstract Our objective was to assess whether differences in cerebral ventricular shape exist between groups of cognitively healthy subjects, individuals with subjective memory complaints (MC), patients with mild cognitive impairment (MCI) and patients with probable Alzheimer's Disease (AD). We included 28 healthy subjects, 21 MC subjects, 28 MCI patients, and 63 probable AD patients from our outpatient memory clinic. FLAIR, T2 and PD-weighted brain MRI scans were acquired at 1.5 T. Volumetric assessment was performed through semi-automated segmentation of the lateral and third ventricles, followed by statistical shape modeling based on GAMEs (Growing and Adaptive MEshes). We generated mesh models for all the ventricle shapes and used them to highlight local areas with significant differences between the healthy subjects and the remaining groups: permutation tests with a predefined threshold ($p = 0.01$) were used for the statistical analysis. Results were analyzed both within a group, to highlight asymmetric behaviour in left and right structures due to a pathological condition, and between groups to highlight trends in ventricular shape changes across the clinical spectrum of cognitive impairment: MC and MCI presented significantly different areas mainly on the right side of the ventricular system, AD mainly on the left side. The analysis across the spectrum showed that some ventricular areas mostly characterize AD (left side of the corona radiata, right inferior medial temporal horn, etc.), others MCI (right side of the thalamus, right side of the corona radiata, right caudate nuclei), and others present a clear trend through the spectrum.

5.1 Introduction

Magnetic Resonance Imaging (MRI) has been a valuable tool for the detection of structural brain changes in subjects with Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD) [1–3]. Many authors have studied atrophy by segmenting brain parenchyma into gray matter (GM) and white matter (WM), using measurements such as the total gray matter volume [2, 4], hippocampal volume [1, 5, 6], and cortical thickness [3]: all these methods depend on the distinction between cerebral gray matter and white matter in MRI. The decline of contrast between white matter and gray matter is inherent in aging, and has a histological basis that can not be resolved with MRI. Thus, the age-associated decline of contrast may lead to skewed conclusions as a result of the image processing.

A different approach to detect and quantify atrophy is to measure the volume of the intracranial CSF compartments: the loss of parenchyma leads to a compensatory dilation of the CSF within the fixed intracranial volume. The contrast between CSF and brain parenchyma is much more pronounced than the contrast between GM and WM, and does not decrease significantly with aging: therefore, this method is not likely to suffer from the same pitfalls as GM and WM segmentation.

The quantification of CSF volumes to study atrophy has previously been applied to MCI and AD research; many studies used the ventricle-to-brain ratio, a measure reflecting global parenchyma loss [1, 7]. Others described the CSF volume of a specific part of the ventricular system, such as the temporal horns, reflecting hippocampal atrophy [8, 9]. The use of the total volume of the ventricular system to determine the amount of atrophy is a simple, but coarse approach, since atrophy of different parts of the brain are blended into a single measurement: cognitive impairment associated with atrophy may involve multiple brain areas, which can not be distinguished using one global measurement of atrophy.

Our aim is to visualize and quantify localized cerebral ventricle shape differences between cognitively healthy subjects, subjects with memory complaints in the absence of cognitive impairment (memory complainers, MC), MCI patients and probable AD patients. For this purpose, we used automatic shape modeling to generate mesh models of all the ventricle shapes, and used them to highlight local areas with significant differences between the four groups of subjects.

The automatic shape modelling technique has previously been applied to the study of cognitively healthy and AD [10]. In the present study, we expand on these results with an increased number of subjects and the addition of both MC and MCI patients, in an attempt to study a more detailed spectrum from cognitively healthy to AD.

5.2 Materials and Methods

5.2.1 Subjects

Elderly subjects that were consecutively referred to our outpatient memory clinic participated in this study. We included 63 patients with probable AD, 28 patients with MCI, 21 patients who were evaluated for memory complaints (MC patients) but who had no detectable

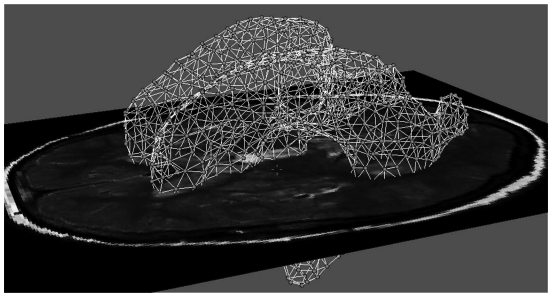


Figure 5.1: Mesh model of a brain ventricle, projected on a PD MRI image.

cognitive deficits on neuropsychological testing, and 28 volunteers without memory complaints who were recruited through an advertisement in a local newspaper. All subjects were evaluated using a standardized dementia screening analysis that included a detailed medical history, a general internal and neurological exam, laboratory tests, neuropsychological testing (including the Mini Mental State Examination (MMSE)) and MRI of the brain. Diagnoses were made in a multidisciplinary consensus meeting, according to the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria for probable AD, and the Petersen criteria for MCI: (1) memory complaint; (2) impaired memory function on cognitive testing; (3) preserved general cognitive functioning; (4) intact activities of daily living; (5) not demented. Patients and controls were included if they: (1) were older than 60 years; (2) had no other neurological or psychiatric illness; and (3) had no abnormalities on MRI other than white matter hyperintensities or an incidental small lacunar lesion ($\leq$ 5mm diameter). The study was approved by the local Medical Ethical Committee. Written informed consent was obtained from all subjects or from a close relative if a patient was demented.

5.2.2 MR image acquisition and pre-processing

MRI was performed on a 1.5 T MR system (Philips Medical Systems, Best, The Netherlands). Conventional dual fast spin-echo sequences [48 axial slices; slice thickness = 3mm; no gap; TR/TE (ms) 3000/27/120; flip angle = $90\,^{\circ}$; field of view = 220 mm; matrix 256x256] were obtained for all patients. The line through the inferior border of the genu and splenium of the corpus callosum defined the direction of scanning. Using in-house developed software, we automatically extracted the intra-cranial cavity, the CSF, and white matter hyperintensities: semi-automatic region growing was used to re-label the ventricular CSF as brain ventricles. All the images were then corrected for brain-size and orientation using automatic affine 12-parameters registration to the LUMC T2-weighted brain template for geriatrics [11–13].

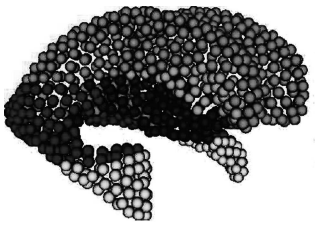


Figure 5.2: Clustering of nodes in areas corresponding coarsely to the corona radiata (light brown), corpus callosum (green), caudate nucleus (light blue), thalamus (brown), upper medial temporal horn (red) and inferior medial temporal horn (yellow).

5.2.3 Ventricular shape modeling

Local shape analysis requires, as a first step, the modeling of the brain ventricles with comparable meshes: following a previously introduced procedure for automatic shape modeling technique [14], healthy subjects were used to generate an average brain ventricle volume, whose surface was modeled by a growing and adapting mesh; subsequently, the topology of the mesh (i.e. nodes and edges) was frozen and its shape was adapted to all the ventricles in the study (see Fig. 5.1). This resulted in a set of comparable meshes: each node in a given mesh is uniquely associated with corresponding nodes in other meshes.

5.2.4 Ventricular surface parcellation

In order to quantify the extent of local shape differences in the different groups and relate it to particular periventricular structures, we manually parcellated the total ventricular surface into regions, according to the adjacent anatomical structures: the left and right corona radiata, the left and right side of the corpus callosum, the left and right caudate nuclei, the left and right thalamus, the left and right upper medial temporal horns and the left and right inferior medial temporal horns. This manual delineation was performed only once, using a regular anatomical atlas, on the template used to normalize all the brain images [15]. The mesh's nodes were associated with the closest overlapping region, and clustered together in corresponding surface areas (see Fig. 5.2). The purpose of our manual delineation was to generate a meaningful regional parcellation of the ventricular surface, rather than to obtain an accurate delineation of the corresponding periventricular structures: such a parcellation allows a more detailed quantitative analysis of the differences between the groups.

5.2.5 Statistical shape comparison

Given two populations of brain ventricle meshes, permutation tests[1] were applied at each node location to evaluate the significance of local shape differences: each location (i.e. node in the mesh) of the ventricles was associated with a p value, rendered with a color code on the

[1] Permutation tests [16] are sound non-parametric statistical tests: they require minimum hypothesis (i.e. same kind of distribution for the populations) and correct for multiple comparisons.

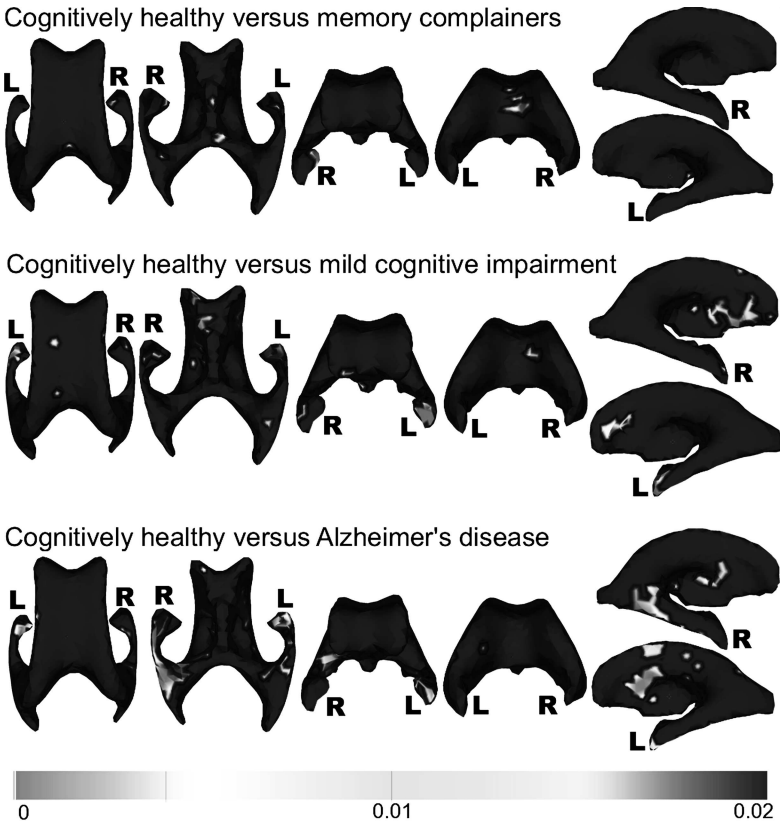


Figure 5.3: Local differences in ventricular shape between cognitively healthy and memory complainers, subjects with MCI and subjects with AD. Local shape differences between groups are represented by color-coded p-values (p values > 0.01 are color-coded in blue). The corresponding colored Figure can be found at page 16, before Chapter 2

ventricular surface, as shown in Fig. 5.3. Locations with a p value below 0.01 were chosen as characteristic features for the brain ventricle's shapes. Displacement vectors were also evaluated at each location found to be significantly different: considering two groups, i.e. healthy subjects and ADs, one could generate two average meshes and evaluate the displacement in space between corresponding nodes. Visual inspection of displacement vectors between cognitively healthy and the remaining groups showed that changes occurred as a result of enlargement of the ventricles in the group with the lower cognitive status. This is illustrated in Fig. 5.4, for the case of cognitively healthy vs. AD patients. Ventricular shape differences between populations were expressed in *extent of change*, reflecting the *extent of local atrophy* and represented by the percentage of nodes significantly different in each area, and *severity of change*, reflecting the *severity of atrophy* and represented by the average displacement per area.

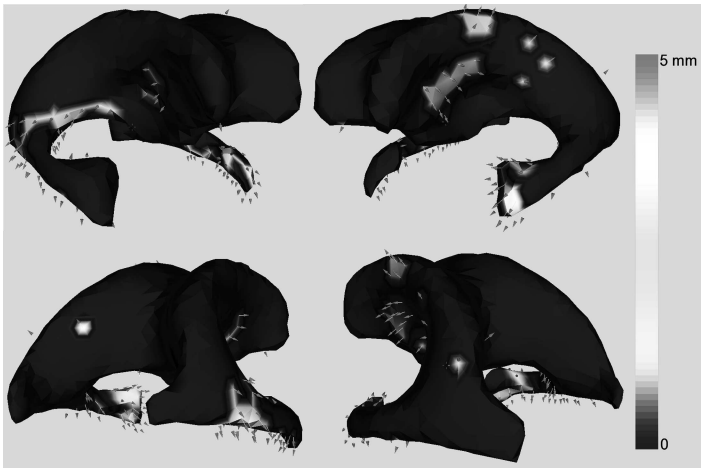


Figure 5.4: The direction of the displacement vectors in these images demonstrates that changes between an average cognitively healthy subject and an average AD occur as a result of ventricular enlargement. The corresponding colored Figure can be found at page 16, before Chapter 2.

5.3 Results

Statistical shape modeling and analysis resulted in the assessment and visualization of localized ventricular changes along the clinical cognitive spectrum. These changes reflect atrophy in adjacent periventricular gray and white matter structures, whose extent and severity are described in Table 5.1. While analyzing these results, it is important to account for small image registration errors: therefore, changes in extent of atrophy among group comparisons (i.e. from controls to MCI and to AD) are considered relevant only if they differ for at least 5%, while changes in severity of atrophy are considered relevant when around 1 mm or higher.

The results reported in Table 5.1 lead to two different kinds of analysis. Focusing on a given structure, e.g. the corona radiata, and on a given column (e.g. Controls vs. AD), one can assess the asymmetry between left and right sides due to the particular condition (e.g. AD), both for the extent and the severity of atrophy: asymmetries in the extent of atrophy (e.g., left/right inferior medial temporal horns in AD) were considered relevant when (1) the extents differ for more than 5% (e.g., 27% in the right and 47% in the left), and (2) involved more than 5% of the nodes (27% and 47% $>$ 5%); asymmetries in the severity of atrophy were considered relevant when (1) the severities differ for more than 1 mm and (2) involved more than 5% of the nodes. When focusing on a row (e.g., left corona radiata), on the other hand, one can highlight trends across the cognitive spectrum.

Table 5.1: Six periventricular areas were chosen to study localized ventricular shape differences. The left and right hemispheres were studied separately. Percentage of significantly different nodes (in %, per area) and average displacement (in mm, per area) reflect the *extent* and *severity* of atrophy respectively.

		Total Number of Nodes	Controls vs. MC		MCI		AD	
			%/mm		%/mm		%/mm	
Corona Radiata	R	88	0	0	1	+3.4	0	0
	L	91	0	0	2	+2.7	11	+2.8
Corpus Callosum	R	140	5	+1.8	5	+2.6	1	+1.7
	L	153	2	+0.7	1	+2.6	1	+3.6
Caudate Nuclei	R	32	0	0	41	+2.2	6	+2.5
	L	30	0	0	17	+2.6	27	+2.9
Thalamus	R	50	2	+0.9	16	+1.4	8	+2.1
	L	38	3	+0.4	0	0	3	+2.6
Sup. Med. Temp. Horn	R	99	2	+1.3	1	+3.2	30	+2.8
	L	122	0	0	3	+1.2	11	+2.1
Inf. Med. Temp. Horn	R	55	13	+1.1	15	+2.0	27	+2.9
	L	51	2	+1.3	49	+2.5	47	+3.1

5.3.1 Left and right asymmetry

The results for left and right asymmetry are highlighted in Table 5.2: for each structure and for each group comparison, the table reports relevant asymmetrical behaviour, both for the extent and the severity of atrophy. The same insight can be achieved visually, comparing the statistical maps in Fig. 5.3. Memory complainers present a higher extent of atrophy in the right inferior medial temporal horn. Patients with MCI present more severe atrophy in the right side of the thalamus, as well as a higher extent to atrophy in the right caudate nucleus, right side of the thalamus, and left inferior medial temporal horn. Finally, subjects with AD present more severe atrophy in the left corona radiata, as well as a higher extent of atrophy in the left corona radiata, left caudate nucleus, left inferior medial temporal horn, right side of the thalamus, and right superior medial temporal horn. In general, it is notable that MC and MCI are mostly affected on the right side of well-defined structures, while ADs are apparently more affected on the left side.

5.3.2 Trends across the cognitive spectrum

Corona Radiata - The left side of the corona radiata presents a relevant change in extent of atrophy in AD, as well as a relevant severity of atrophy both in MCI and AD. The right side, on the other hand, does not show any relevant change in the extent of atrophy: it is instead notable that MCIs present a relevant severity of atrophy when compared to both MC and AD.

Table 5.2: Six periventricular areas were chosen to investigate asymmetry of localized ventricular shape differences across the cognitive spectrum. Considering the results in Table 5.1, this table shows, for each structure, whether a relevant asymmetry is present or not (both in terms of extent and severity of atrophy). An *R(L)* indicates that the right (left) side is more affected.

	MC		MCI		AD	
	Ext./Sev.		Ext./Sev.		Ext./Sev.	
Corona Radiata	-	-	-	-	L	L
Corpus Callosum	-	-	-	-	-	-
Caudate Nuclei	-	-	R	-	L	-
Thalamus	-	-	R	R	R	-
Sup. Med. Temp. Horn	-	-	-	-	R	-
Inf. Med. Temp. Horn	R	-	L	-	L	-

Corpus Callosum - The extent of atrophy does not change relevantly across the cognitive spectrum, neither on the left nor on the right side. The severity of atrophy, on the other hand, presents a relevant trend on the left side, increasing while moving from MC to AD.

Caudate Nuclei - The right caudate nucleus presents a relevant change in extent of atrophy both in MCI and in AD, when compared to MC; even more striking is the much higher extent in MCI, compared to both MC and AD. Concerning the severity of atrophy, relevant changes are highlighted both in MCI and AD, compared to MC, while no relevant difference is found between MCI and AD. The same severity patter is present in the left caudate nucleus, while the extent of atrophy presents a clear trend through the cognitive spectrum.

Thalamus - Concerning the extent of atrophy, the right side of the thalamus presents the same behaviour previously observed for the right caudate nucleus: MCIs show a relevant change when compared to both MC and AD, although even ADs present a higher extent of atrophy than MC. The change in severity of atrophy, both in the right and left sides of thalamus, is relevant only in AD.

Superior medial temporal horn - The extent of atrophy presents a relevant change in AD, when compared to MC and MCI, both in the left and right horns. Changes in the severity of atrophy are relevant both in MCI and AD (both in the right and left horns), when compared to MC, while no relevant difference appears between MCI and AD.

Inferior medial temporal horn - The right inferior medial temporal horn presents a relevant change in extent of atrophy in AD, while the left inferior medial temporal horn presents relevant changes both in MCI and AD. The severity of atrophy follows a similar pattern: in the right inferior horn, only AD shows a relevant change, while in the left one both MCI and AD presents relevant changes when compared to MC.

5.4 Discussion

In this study, we have investigated local ventricular shape differences among different groups, across the spectrum of cognitive function: healthy subjects where compared to memory complainers, mild cognitive impairments, and subjects with Alzheimer Disease. The cross-sectional analysis of group comparisons (e.g. Controls vs. AD) highlighted relevant asymmetrical shape differences between left and right structures due to the particular pathology, while the analysis across the cognitive spectrum highlighted particular trends in the extent and severity of atrophy.

The most striking findings, concerning left and right asymmetry, were predominantly worse conditions (both in terms of extent and severity of atrophy) on the right side of well-defined structures (such as caudate nucleus and thalamus) in MC and MCI. Subjects with AD, on the other hand, presented worse conditions mostly on the left side of periventricular structures: corona radiata, caudate nucleus, and inferior medial temporal horn. In AD, a predominant involvement of the right side was shown only in the thalamus and superior medial temporal horn.

The interpretation of our results across the cognitive spectrum leads to further considerations. The left caudate nucleus and the left side of the corpus callosum presented a relevant trend while moving from MC to AD: the former presented increasing extent of atrophy, while the latter presented increasing severity of atrophy. The left inferior medial temporal horn presented a relevant change in both extent and severity of atrophy while moving from MC to MCI/AD: the same pattern was found for the severity of atrophy in the caudate nuclei (left and right), and the superior medial temporal horns (left and right). Alzheimer Disease differed from both MC and MCI for the extent of atrophy in the right and left superior medial temporal horns, and right inferior medial temporal horn; similar patterns were found for the severity of atrophy in the thalamus and the right inferior medial temporal horn. Finally, well specific regions seemed to characterize MCI: relevant changes when compared to both MC and AD were found for the extent of atrophy in the right caudate nucleus and right thalamus, and for the severity of atrophy in the right corona radiata and right side of the corpus callosum.

Few publications describe a direct relationship between damage to the corona radiata and cognitive decline. Subtle cognitive impairment and emotional disturbances were found in patients with a single lacunar infarct in the corona radiata [17]. Similar findings are described in a number of reports on lacunar infarcts in supratentorial structures [18–20]. In the literature, no consistent pattern was found between the involved supratentorial structures and the neuropsychological impairments. To the best of our knowledge, the corona radiata has never been previously described as a possible biomarker for the detection and follow up of neurodegenerative diseases involving memory loss. Our results indicate that local changes in the surface of this structure may differentiate between cognitively healthy subjects and subjects with AD and MCI.

The caudate nucleus forms part of a cholinergic network, consisting of central cholinergic neurons, which are thought to be involved in a number of neurodegenerative diseases, including AD [21]. Disturbance of the central cholinergic system does not appear to be closely related to the etiology, but to the development of clinical symptoms [21]. A relationship be-

tween caudate nucleus volume and a number of dementing illnesses such as AD, Dementia with Lewy Bodies and Vascular Dementia has been described before [22]. Volumetric analysis of the caudate nucleus on MRI was not found to discriminate between diagnostic groups. However, our results indicate that changes in the ventricular area adjacent to the caudate nucleus occur between cognitively healthy subjects, subjects with MCI and AD, which may point to the caudate nucleus as a biomarker for MCI and AD.

The role of the thalamus in the development of cognitive decline in AD may be of vascular origin, superimposed on AD pathology [23]. Subcortical small lesions involving the thalamus disrupt cortico-subcortical circuits, resulting in cognitive dysfunction. Our results show that the extent of atrophy in the right side of the thalamus might be a good indication for MCI, while the severity of atrophy on the left side seems to be characteristic of AD. It is widely accepted that structural changes of the medial temporal lobe, where the amygdala and hippocampus are located, are associated with an increased risk of MCI and AD [1, 8, 24, 25]. Though our association between medial temporal lobe atrophy, MCI and AD is not new, it emphasizes the ability of our technique to detect clinically relevant changes to periventricular gray matter structures.

Finally, significant cerebral shape differences were found between cognitively healthy subjects and subjects with subjective memory complaints. Although the latter group of subjects has non-objectifiable memory impairment, our results indicate that anatomically these subjects may be classified in a distinct group. These results are in accordance with those from numerous other studies [26, 27].

The strengths of the present study include the clear classification of subjects into four cognitive groups through an extensive neuropsychological test battery, as well as the use of a validated automated set-up of our statistical shape modeling method. To our knowledge, this is the first study that not only describes the association of structural changes of the surface of the medial temporal lobe with cognitive decline in MCI and AD, but also includes the entire surface of the ventricular system, in search of potential biomarkers of neurodegenerative disease involving memory loss (MC, MCI, and AD). The results might be strengthened in the future by means of larger datasets; moreover, a further distinction between MCI converters[2] and non-converters might lead to more consistent patterns across the spectrum of cognitive function.

5.5 Conclusions

In this work, we have shown that localized changes in ventricular shape may differentiate between cognitively healthy subjects, subjects with subjective memory complaints, subjects with MCI and subjects with AD. The corona radiata, caudate nucleus and thalamus are added to the well-established medial temporal lobe as possible structural biomarkers of cognitive decline.

[2] With MCI-converters we refer to MCI patients who eventually convert to an AD type of dementia.

Bibliography

[1] Jack CR, Shiung MM, Weigand SD, O'Brien PC, Gunter JL, Boeve BF, et al. Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. Neurology. 2005;65(8):1227–1231.

[2] Karas GB, Scheltens P, Rombouts SA, Visser PJ, van Schijndel RA, Fox NC, et al. Global and local gray matter loss in mild cognitive impairment and Alzheimer's disease. Neuroimage. 2004;23(2):708–716.

[3] Singh V, Chertkow H, Lerch JP, Evans AC, Dorr AE, Kabani NJ. Spatial patterns of cortical thinning in mild cognitive impairment and Alzheimer's disease. Brain. 2006;129(11):2885–2893.

[4] Allen JS, Bruss J, Brown CK, Damasio H. Normal neuroanatomical variation due to age: the major lobes and a parcellation of the temporal region. Neurobiol Aging. 2005;26(9):1245–1260.

[5] Du AT, Schuff N, Laakso MP, Zhu XP, Jagust WJ, Yaffe K, et al. Effects of subcortical ischemic vascular dementia and AD on entorhinal cortex and hippocampus. Neurology. 2002;58(11):1635–1641.

[6] Mungas D, Harvey D, Reed BR, Jagust WJ, DeCarli C, Beckett L, et al. Longitudinal volumetric MRI change and rate of cognitive decline. Neurology. 2005;65(4):565–571.

[7] Bigler ED, Kerr B, Victoroff J, Tate DF, Breitner JC. White matter lesions, quantitative magnetic resonance imaging, and dementia. Alzheimer DisAssocDisord. 2002;16(3):161–170.

[8] Apostolova LG, Dinov ID, Dutton RA, Hayashi KM, Toga AW, J L Cummings P M Thompson PM. 3D comparison of hippocampal atrophy in amnestic mild cognitive impairment and Alzheimer's disease. Brain. 2006;129(11):2867–2873.

[9] Giesel FL, Hahn HK, Thomann PA, Widjaja E, Wignall E, Tengg-Kobligk H, et al. Temporal horn index and volume of medial temporal lobe atrophy using a new semiautomated method for rapid and precise assessment. AJNR AmJNeuroradiol. 2006;27(7):1454–1458.

[10] Ferrarini L, Palm WM, Olofsen H, van Buchem MA, Reiber JHC, Admiraal-Behloul F. Shape differences of the brain ventricles in Alzheimer's disease. NeuroImage. 2006;32(2):1060–1069.

[11] Admiraal-Behloul F, van den Heuvel D, Olofsen H, Schmitz N, van Buchem MA. Brain templates for the elderly. In: International Society for Magnetic Resonance in Medicine (ISMRM) - Toronto; 2003. .

[12] Admiraal-Behloul F, van den Heuvel DMJ, Olofsen H, van Osch MJP. Fully automatic segmentation of white matter hyperintensities in MR images of the elderly. Neuroimage. 2005;28(3):607–617.

[13] Woods RP, Grafton ST, Holmes CJ, Chery SR, Mazziotta JC. Automated Image Registration I. General methods and intrasubject, intramodality validation. J Comput Assist Tomogr. 1998;22:139–152.

[14] Ferrarini L, Olofsen H, Palm WM, van Buchem MA, Reiber JHC, Admiraal-Behloul F. GAMEs: Growing and Adaptive Meshes for Fully Automatic Shape Modeling and Analysis. Medical Image Analysis. 2007;11(3):302–314.

[15] Weir J, Abrahams PH. Imaging Atlas of Human Anatomy. 2003;.

[16] Nichols TE, Holmes AP. Nonparametric Permutation Tetsts For Functional Neuroimaging: A Primer with Examples. Human Brain Mapping. 2001;15:1–25.

[17] Zandvoort MJV, Kappelle LJ, Algra A, Haan EHD. Decreased capacity for mental effort after single supratentorial lacunar infarct may affect performance in everyday life. JNeurolNeurosurgPsychiatry. 1998;65(5):697–702.

[18] Pullicino P, Lichter D, Benedict R. Micrographia with cognitive dysfunction: minimal sequelae of a putaminal infarct. Mov Disord. 1994;9(3):371–373.

[19] Schnider A, Gutbrod K, Hess CW, Schroth G. Memory without context: amnesia with confabulations after infarction of the right capsular genu. JNeurolNeurosurgPsychiatry. 1996;61(2):186–193.

[20] Yamanaka K, Fukuyama H, Kimura J. Abulia from unilateral capsular genu infarction: report of two cases. JNeurolSci. 1996;143(1-2):181–184.

[21] Oda Y. Choline acetyltransferase: the structure, distribution and pathologic changes in the central nervous system. PatholInt. 1999;49(11):921–937.

[22] Barber R, McKeith I, Ballard C, O'Brien J. Volumetric MRI study of the caudate nucleus in patients with dementia with Lewy bodies, Alzheimer's disease, and vascular dementia. JNeurolNeurosurgPsychiatry. 2002;72(3):406–407.

[23] Mori E. Impact of subcortical ischemic lesions on behavior and cognition. AnnNYAcadSci. 2002;977:141–148.

[24] Jack CR, Petersen RC, Xu YC, O'Brien PC, Smith GE, Ivnik RJ, et al. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. Neurology. 1999;52(7):1397–1403.

[25] Mungas D, Reed BR, Jagust WJ, DeCarli C, Mack WJ, Kramer JH, et al. Volumetric MRI predicts rate of cognitive decline related to AD and cerebrovascular disease. Neurology. 2002;59(6):867–873.

[26] Saykin AJ, Wishart HA, Rabin LA, Santulli RB, West LAFJD, McHugh TL, et al. Older adults with cognitive complaints show brain atrophy similar to that of amnestic MCI. Neurology. 2006;67(5):834–842.

[27] van der Flier WM, van Buchem BA, Weverling-Rijnsburger AW, Mutsaers ER, Bollen EL, Admiraal-Behloul F, et al. Memory complaints in patients with normal cognition are associated with smaller hippocampal volumes. JNeurol. 2004;251(6):671–675.

CHAPTER 6

MMSE Scores Correlate with Local Ventricular Enlargement in the Spectrum from Cognitively Normal to Alzheimer Disease

All change is not growth, as all movement is not forward.

Ellen Glasgow, 1873-1945.

This chapter is reprinted from:

NeuroImage

Authors: L. Ferrarini, W.M. Palm, H. Olofsen, R. van der Landen, G.J. Blauw, R.G.J. Westendorp, E.L.E.M. Bollen, H.A.M. Middelkoop, J.H.C. Reiber, M.A. van Buchem, and F. Admiraal-Behloul

Title: **MMSE Scores Correlate with Local Ventricular Enlargement in the Spectrum from Cognitively Normal to Alzheimer Disease**, Accepted for publication (November 2007)

Abstract In this work, we aimed at correlating focal atrophy in periventricular structures with cognitive function, in the spectrum from healthy subjects to severe Alzheimer Disease: 28 subjects with normal cognition and 84 patients presenting various degrees of cognitive impairment were included in the study. The cognitive level of each subject was assessed with the Mini Mental State Examination (MMSE). Atrophy in periventricular structures was inferred by modeling and analyzing local shape variations of brain ventricles: for a given subject, we distinguished between the *severity of atrophy*, estimated as local enlargement (in mm) of the ventricular surface relative to an average normal subject, and the *extent of atrophy*, defined as the percentage of the ventricular surface (global or per anatomical region) significantly different from an average control. Linear regression across subjects was performed to evaluate the correlation between atrophy and MMSE score. The severity of atrophy showed good correlation with MMSE score in the left thalamus, the left temporal horn, the left corona radiata, and the right caudate nuclei. The extent of atrophy showed no significant correlations. In conclusion, the MMSE scores correlate with localized depth of atrophy in well-defined periventricular structures.

6.1 Introduction

Dementia is usually defined as a progressive decline of cognitive function. Among the several possible causes of dementia, Alzheimer Disease (AD) is usually regarded as the most common one. Considering the spectrum of cognitive function, healthy subjects are at one side, presenting normal cognitive level, while patients with severe Alzheimer Disease (AD) are at the very opposite end. An important intermediate position is represented by the Mild Cognitive Impairment (MCI): patients suffering of MCI experience isolated memory deficits (amnesia), but are otherwise cognitively capable and not limited in their daily life. Growing evidence of how AD affects brain structures has previously been presented: hippocampal atrophy [1–5]; rapid degeneration of prefrontal lobe [6, 7]; atrophy of entorhinal cortex [8–11]; pronounced atrophy of the corpus callosum [12–14]. Similar studies have analyzed the effects of MCI and its relationship with AD [15–17].

Several cognitive tests exist to assess the cognitive level of an individual. One of them, the Mini Mental State Examination (MMSE) [18], has largely been used in clinical practice for the diagnostic work up of AD and MCI. Conspicuous attempts to correlate MMSE scores with brain structural changes, as captured by Magnetic Resonance Imaging (MRI), have recently been reported in the literature: in [19], a high correlation in AD between MMSE scores and changes in temporal horns was reported; in [20], a high correlation between cognitive impairment and selective white matter damage in AD, measured as a reduction of fractional anisotropy, was shown (particularly in the splenium of the corpus callosum); in [15], the correlation between MMSE scores and hippocampal volume changes was investigated, finding no significant correlation in MCI-converters[1]; in [21], the authors showed high correlation in AD between MMSE scores and decrease of gray matter in the left temporal lobe; finally, a strong correlation between MMSE scores and gray matter loss in several cortical regions was observed in clinical and pre-clinical AD [22].

The aim of our study was to investigate to what extent the MMSE score correlates with atrophy in periventricular structures. Any change in volume or shape in these structures must be reflected on the shape and volume of the brain ventricles: thus, by analyzing shape variations in the entire ventricular system we could indirectly estimate the degree of atrophy in all the periventricular structures. The delineation of small gray and white matter structures is a challenging task, particularly in elderly subjects, due to the loss of contrast between different tissues [23]: (semi-) automatic segmentation is error prone and not always reproducible, while manual delineation is highly time consuming. Cerebrospinal fluid (CSF), on the other hand, can be automatically and reliably segmented, thanks to its sharp contrast with respect to the rest of the parenchyma. While analyzing atrophy, we introduced a distinction between the *severity of atrophy* and the *extent of atrophy*. The former was estimated as local enlargement (in mm) of the ventricular surface with respect to a normal average case; the latter was given by the percentage of ventricular surface significantly different from an average normal case. Finally, rather than focusing on only one specific pathological condition, such as AD or MCI, we included subjects from both groups, covering a larger area in the spectrum of cognitive impairment.

[1] Among the subjects affected by MCI, some convert in time into AD and are referred to as MCI-converters.

Table 6.1: Demographic Data: Total number of subjects per group (and total number of male), Age (in years), and MMSE score.

		Controls	Patients
Tot (N. Male)		28 (12)	84 (38)
Age $\mu(\sigma)$		74 (6.9)	74 (7.3)
MMSE	$\mu(\sigma)$	27.6 (1.7)	20.8 (5.6)
	[min-max]	[24-30]	[4-30]

6.2 Material and Methods

6.2.1 Subjects

The subjects included in the study were chosen so that a larger area in the spectrum of cognitive function could be represented: 28 volunteers with normal cognitive function, and 84 patients presenting different degrees of cognitive impairment (26 subjects with MCI and 58 patients affected by AD, ranging from mild to severe cases). The volunteers were recruited through advertisements in local newspapers, while patients with cognitive impairment were consecutively referred to our outpatient memory clinic. The cognitive status was evaluated for all subjects using a standardized dementia screening, including a detailed medical history, a general internal and neurological exam, laboratory tests, neuropsychological testing (i.e., MMSE), and magnetic resonance images of the brain. Diagnoses were made in a multidisciplinary consensus meeting, according to the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria for probable AD, and the Petersen criteria for MCI. Table 6.1 shows the demographic data: *t* tests performed on age and sex showed no significant difference between groups ($p \leq 0.01$). Subjects were included if they were older than 60 years, had no other neurologic or psychiatric illness, and had no abnormalities on MRI other than white matter hyperintensities or an incidental small lacunar lesion ($\leq$ 5 mm diameter). Approval by the local Medical Ethical Committee was granted to the study. Written informed consent was obtained from all subjects or from a close relative if a patient was demented.

6.2.2 MRI acquisition and Image pre-processing

Magnetic resonance images were acquired on a 1.5-T MR-system (Philips Medical Systems, Best, The Netherlands) using the following pulse sequences: dual fast spin-echo (proton density and T2 weighted): time to echo (TE) 27/120 ms, repetition time (TR) 3000 ms, 48 contiguous 3-mm slices without an interslice gap, matrix 256 x 256, field of view (FOV) 220 mm. FLAIR (fluid attenuated inversion recovery): TE 100 ms, TR 8000, 48 contiguous 3-mm slices without an interslice gap, matrix 256 x 256, FOV 220 mm.

Images were pre-processed with our in-house developed automatic-segmentation software [24], which automatically extracted the intracranial cavity, the CSF, and the white matter hyperintensities. Three-dimensional region growing was applied to re-label the ventricu-

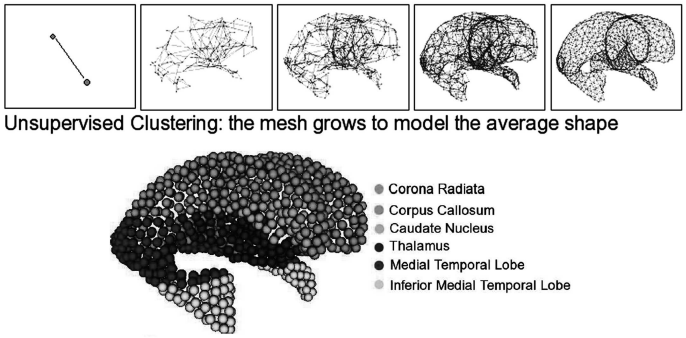


Figure 6.1: (*Top*) Shape modeling of an average ventricle: a mesh grows and adapts to model an average control surface. (*Bottom*) The ventricle volume was manually divided in regions, according to the adjacent periventricular structures, and each node was labeled accordingly.

lar CSF as *ventricles*. Automatic affine 12-parameters image registration [25] was applied to normalize all the images and the ventricular segmentations to the LUMC T2-weighted brain template for geriatrics [26], in order to correct for brain-size and orientation.

6.2.3 Shape Modeling

The measure of focal atrophy in periventricular structures is based on the corresponding deformations occurring in the shape of the brain ventricles. Thus, as a first step one needs to model all the brain ventricles with comparable meshes. This was done with an automatic modeling technique previously introduced in [27]: the healthy subjects are used to generate an average brain ventricle image, whose surface is modeled by a growing and adapting mesh (see Fig. 6.1 (*top*)); subsequently, the mesh's structure is frozen (nodes are not added nor removed), and the mesh's shape is automatically adapted to all the ventricles in the study, ending up with comparable meshes (i.e. each node in a given mesh is uniquely associated with corresponding nodes in the other meshes). Local changes in ventricular shapes should be associated with focal atrophy in periventricular structures: thus, an expert was asked to manually delineate several regions on the initial average ventricular volume, each corresponding to an adjacent periventricular structure. Nodes in the first mesh were then labeled accordingly to the closest region (see Fig. 6.1 (*bottom*)), and labels were propagated to all the other meshes. We visually inspected the results across all the meshes, to guarantee correct correspondence between labels and adjacent periventricular structures. In the remaining of this paper, the terms *region left (right) corona radiata*, *region left (right) caudate nuclei*, etc., will refer to the sets of nodes associated with the corresponding periventricular structures.

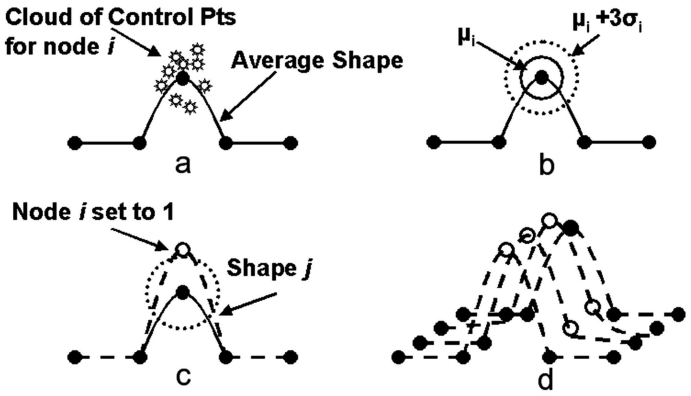


Figure 6.2: Extent of Atrophy measurement. (a) Spatial Distribution of the control population around node i of the average shape. (b) Average distance μ_i and threshold $\mu_i + 3\sigma_i$. (c) In a shape j, the node i is labeled as 1 if its distance from the average node is higher than the threshold. (d) The extent of atrophy is the percentage of nodes set to 1.

6.2.4 Pre-Statistical Shape Analysis

Does MMSE score correlate with severity of focal atrophy ?

We have assessed the significance of the correlation between the severity of focal atrophy and MMSE score for each location (i.e. node) on the ventricular surface. Considering a particular node in a given subject, one can evaluate its Euclidian distance in space to the corresponding node in the average mesh, obtained from the control group. The underlying idea is that focal atrophy in periventricular structures should lead to a more pronounced local ventricular enlargement, and subsequently to larger distances between a given node and the corresponding one in the average-control. We visually checked that all the non-control ventricles presented larger structures than the average-control. For each node, the local enlargements across all subjects were correlated with the corresponding MMSE scores: the statistical analysis and the results are reported in section 6.3.1.

Does MMSE score correlate with the extent of atrophy ?

A similar approach was taken to evaluate the correlation between the MMSE score and the extent of atrophy. Starting with the average mesh previously introduced, for each node i we modeled the local distribution in space of the control population, evaluating the average distance μ_i and the standard deviation σ_i (with $i = 1, 2, .., N_{Nodes}$) (see Fig. 6.2). Subsequently, considering a subject j, for each node i we evaluated its Euclidian distance d_i^j to the corresponding node in the average mesh. The location i in subject j was then labeled as:

$$l_i^j = \begin{cases} 1 & \text{if } d_i^j > \mu_i + 3 * \sigma_i, \\ 0 & \text{otherwise.} \end{cases}$$

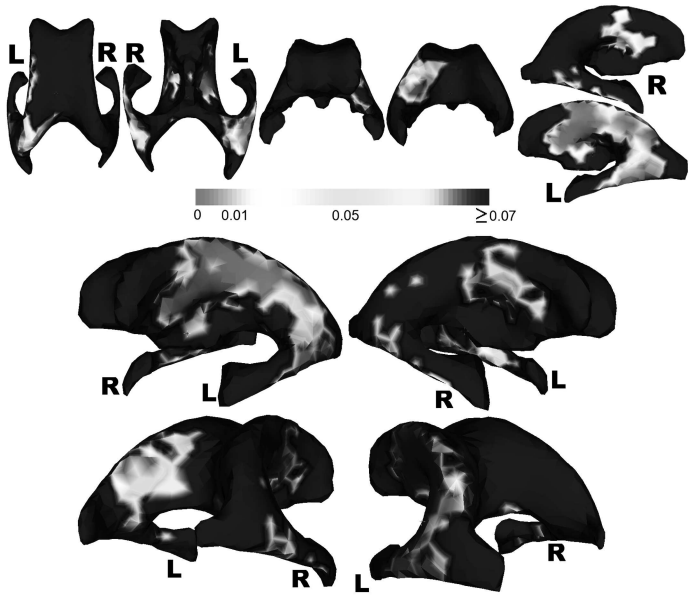


Figure 6.3: Local correlation between severity of atrophy and cognitive impairment, across all subjects (controls and patients). Orthogonal (*top*) and prospective (*bottom*) views (p-values higher than 0.05 are plotted in blue). The corresponding colored Figure can be found at page viii, before the Introduction.

For each subject and for each region, we could therefore evaluate the percentage of locations set to 1, corresponding to the local extent of atrophy. Such values were finally correlated, across all subjects, with the MMSE scores: permutation tests were used to assess the significance level (p-value) of the correlation. Results are reported in section 6.3.2.

6.3 Results

6.3.1 MMSE Correlation with Severity of Atrophy

For each given node, we performed a linear regression correlating local distances of all subjects from the control average node, to the MMSE scores associated with the subjects. The statistical significance of the correlation was assessed via permutation tests, correcting for multiple comparisons. Thus, the results provided us, for each node, with a p-value indicating the significance of the correlation in that particular location. Mapping the p-values on an average control, we could visually highlight areas with higher correlation (see Fig. 6.3). Results were then clustered into the manually delineated regions, and for each region we evaluated the percentage of surface in which the severity of atrophy significantly correlated with the MMSE score: this made it possible to identify the regions (and, consequently, the

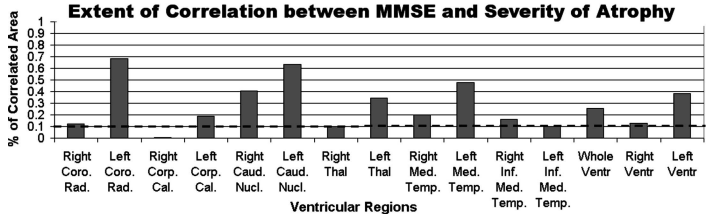


Figure 6.4: Extent of correlation for the severity of atrophy. For each region, we plot the percentage of ventricular surface whose severity of atrophy (see Fig. 6.3) significantly correlates with the MMSE score.

periventricular structures) with higher correlation. Figure 6.4 shows the percentage of area in which the severity of atrophy correlates ($p < 0.05$) with the MMSE score, for each manually delineated region (the whole, left, and right ventricular systems are also considered). Some regions show a high correlation: left corona radiata, right caudate nuclei, left caudate nuclei, left thalamus, and left medial temporal horn. Other regions, on the other hand, are almost completely uncorrelated: right side of the corpus callosum, right side of the thalamus, left inferior medial temporal horn. It is notable that the severity of atrophy in left periventricular structures (reflected in the left ventricular system) presents a higher correlation with the MMSE score than those on the right.

6.3.2 MMSE Correlation with Extent of Atrophy

We performed a linear regression analysis to correlate the MMSE score and the extent of atrophy for the whole, left, and right ventricular systems, as well as for each of the manually delineated regions. All subjects were included in the analysis, and both the R^2 statistic and the corresponding p value were evaluated. Results are reported in Table 6.2. No significance correlation was found in any of the analyzed areas. Only in the left inferior medial temporal horn we could detect a trend, showing that higher MMSE scores reflect smaller extents of atrophy: nevertheless, the correlation was not significant (p value = 0.06).

6.4 Discussion and Conclusions

In this work, we have investigated the correlation between atrophy in periventricular structures and cognitive function, in the spectrum from normal cognition to severe AD. The analysis of structural changes in periventricular structures always presents some serious challenges. First, one should choose between volumetric or shape-based analyses. In [28], the authors analyzed a subset of the population used in this study, showing that MR volumetric measurements of the brain correlate with cognitive decline in non-demented elderly. The volume changes, being inherently global, are less subject to small segmentation errors; shape-based methods, on the other hand, aiming at a localized analysis, are more sensitive to the accuracy of the segmentation. Conversely, the volumetric distributions of periventricular structures in

Table 6.2: Linear regression between extent of atrophy and MMSE score (all 112 subjects included): for each area, we report the R^2 value and the corresponding p value. No significance correlation was found (p threshold set at 0.05).

		R^2	p-value
Whole Ventricle		0.01	0.67
Right Ventricle		10^{-6}	1.00
Left Ventricle		0.03	0.40
Corona Radiata	R	0.01	0.66
	L	0.01	0.75
Corpus Callosum	R	0.03	0.39
	L	0.01	0.70
Caudate Nuclei	R	0.03	0.39
	L	0.03	0.43
Thalamus	R	0.001	0.88
	L	10^{-4}	0.94
Sup. Med. Temp. Horn	R	10^{-4}	0.92
	L	0.05	0.31
Inf. Med. Temp. Horn	R	0.003	0.78
	L	0.14	0.06

normal and non-normal populations might overlap and fail to provide discriminative results; shape analysis, highlighting more localized differences, is less keen to this problem. Thus, a shape-analysis of periventricular structures is to be preferred. Unfortunately, the intensity contrast between white and gray matter decreases with age [23], making it a more difficult and less accurate task to delineate periventricular structures in MR. On the other hand, the contrast between CSF and the remaining parenchyma stays sharp in elderly, allowing automatic tools to reliably segment ventricular CSF. In our study, we have analyzed shape changes of the whole ventricular system as an indirect tool to investigate atrophy in the surrounding periventricular structures: this differs from previous methods which focused more on singular gray/white matter structures or specific ventricular areas [15, 19, 20]. Moreover, we introduced a clear separation between severity of atrophy and extent of atrophy, showing that the former correlates with cognitive impairment, while the second does not: while doing this, we explored a larger area in the spectrum of cognitive function, including both subjects with MCI and AD, while previous methods focused more on AD [19, 20], on AD and a limited number of converting MCI [22], or on converting and non-converting MCI [15].

The assessment of a subject's cognitive function is not a trivial task. In this work, we have chosen to use the MMSE score as an estimation of an individual's cognitive level. This choice was justified by the fact that, despite its limitations, the MMSE score is largely used in the diagnostic work up of AD and MCI in clinical routine. Thus, a better understanding of how the MMSE score relates to periventricular atrophy is important. Global measurements of brain atrophy have already been proved to correlate with MMSE scores [29, 30]. The local

analysis presented in this work aimed at further improving our knowledge on how damage in specific periventricular structures contributes to the assessment of an individual's cognitive level.

The results lead to some further considerations. The severity of cognitive impairment seems to correlate significantly with increased severity of atrophy in well-defined periventricular structures. More specifically, areas like the left corona radiata, right and left caudate nuclei, left side of the thalamus, left medial temporal horn, and the left side of the splenium in the corpus callosum showed high correlation between severity of atrophy and cognitive impairment. These findings are in agreement with those previously reported in the literature: a correlation between dementia and infarcts in left corona radiata was highlighted in [31]; the left temporal horn and its correlation with dementia is well-established [15, 19], and shows how atrophy of the temporal lobe and hippocampus are inherently related with the disease; involvement of the left side of the splenium in the corpus callosum was highlighted by [20], by using global diffusion tensor images for healthy subjects and ADs. In [19], the authors suggested that a lack of correlation between cognitive decline and hippocampal atrophy might be due to the MMSE test itself, since it is mostly related with cortical functions. Our findings on the thalamus might lead to some further considerations: the thalamus is responsible to code and transmit information to the cortex, particularly for what concerns the sensory system (auditory system, visual system, etc.): thus, impairments of cortical functions might also occur as a consequence of thalamus deficits. Finally, looking at the whole, left, and right ventricular systems, it is clear that the severity of atrophy in the left hemisphere presents a higher correlation: this result is consistent with previous studies showing how changes in left periventricular structures relate to dementia. When we tried to correlate the extent of atrophy with cognitive function, we could not find any significant result. The only exception was the left inferior medial temporal horn, which presented an almost significant correlation ($p = 0.06$).

The assumption that MMSE scores can confidently reflect cognitive impairment is open to discussion. Neuropsychological tests usually provide a large set of indicators to estimate the cognitive level of an individual. The MMSE test is a global composite measurement of different cognitive functions, spatially located in different parts of the brain. Some criticisms have been moved to the use of MMSE test, especially when assessing the cognitive level of healthy subjects or patients with MCI. Nevertheless, the MMSE test keeps being intensively used in the diagnosis of AD and MCI: thus, efforts to improve our understanding on its relation with brain structural changes are largely justified. To the best of our knowledge, this is the first study which clearly distinguishes between different kinds of atrophy in periventricular structures, and correlates them with cognitive impairment in MCI and AD, as estimated by the MMSE score.

In conclusion, our results suggest that worse MMSE scores correlate with an increased severity of atrophy in well-defined periventricular structures. Future research might focus on other global tests, such as the Camcog, and more specific cognitive tests: a comparison between tests could strength even further the link between cognitive impairment and atrophy of well-defined periventricular structures.

Bibliography

[1] Freeborough PA, Fox NC. The boundary shift integral: an accurate measure of cerebral volume changes from registered repeat MRI. IEEE Trans Med Imaging. 1997;16(5):623–629.

[2] Pitiot A, Delingette H, Toga AW, Thompson PM. Learning Object Correspondences with the Observed Transport Shape Measure. Information Processing in Medical Imaging (IPMI). 2003;LNCS 2732:25–37.

[3] Rueckert D, Frangi AF, Schnabel JA. Automatic Construction of 3-D Statistical Deformation Models of the Brain Using Nonrigid Resitration. IEEE Transactions on Medical Imaging. 2003;22(8):1014–1025.

[4] Schott JM, Price SL, Frost C, Whitwell JL, Rossor MN, Fox NC. Measuring atrophy in Alzheimer diseases: a serial MRI study over 6 and 12 months. Neurology. 2005;65:119–124.

[5] Thodberg HH. Minimum Description Length Shape and Apperance Models. Information Processing in Medical Imaging (IPMI). 2003;LNCS 2732:51–62.

[6] Coffey CE, Wilkinson WE, Parashos IA. Quantitative cerebral anatomy of aging human brain: a cross-sectional study using magnetic resonance imaging. Neurology. 1992;42:527–536.

[7] Raz N, Gunning FM, Head D. Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. Cereb cortex. 1997;7:268–282.

[8] Barns J, Scahill RI, Boyes RG, Frost C, Lewis EB, Rossor CL, et al. Differentiating AD from Aging using semiautomated measurement of hippocamapal atrophy rates. Neuroimage. 2004;23:574–581.

[9] Dekaban AS. Changes in brain weights during the span of human life: relation of brain weights to body heights and body weights. Ann Neurol. 1978;4:345–356.

[10] Lerch JP, Priessmer JC, Zijdenbos A, Hamperl H, Teipel SJ, Evans AC. Focal Decline of Cortical Thickness in Alzheimer's Disease Identified by Computational Neuroanatomy. Cerebral Cortex. 2004;15:995–1001.

[11] Pruessner JC, Collins DL, Pruessner M, Evans AC. Age and gender poredict volume decline in the anterior and posterior hippocampus in early adulthood. The journal of neuroscience. 2001;21(1):194–200.

[12] Bozzao A, Floris R, Baviera ME, Apruzzese A, Simonetti G. Diffusion and perfusion MR imaging in cases of Alzheimer's disease: correlations with cortical atrophy and lesion load. AJNR AM J Neuro-radiol. 2001;22:1030–1036.

[13] Hanyu H, Asano T, Sakurai H, Imon Y, Iwamoto T, Takazaki M, et al. Diffusion-weighted and magnetization transfer imaging of the corpus callosum in Alzheimer's disease: a quantitative study. J Neurol Sci. 1999;167:37–44.

[14] Teipel SJ, Bayer W, Alexander GE, Zebhuhr Y, Teichberg D, Kulic L, et al. Progression of corpus callosum atrophy in Alzheimer disease. Arch Neurol. 2002;59:243–248.

[15] Apostolova LG, Dutton RA, Dinov ID, Hayashi KM, Toga AW, Cummings JL, et al. Conversion of Mild Cognitive Impairment to Alzheimer disease Predicted by Hippocampal Atrophy Maps. Arch Neurol. 2006;63:693–699.

[16] Carmichael O, Thompson P, Dutton R, Lu A, Lee S, Lee J, et al. Mapping ventricular changes related to dementia and mild cognitive impairment in a large community-based cohort. ISBI Proc. 2006;p. 315–318.

[17] Chételat G, Landeau B, Mezenge F, Viader F, de la Sayette V, Desgranges B, et al. Using voxel-based morphometry to map structural changes associated with rapid conversion in MCI: a longitudinal MRI study. NeuroImage. 2005;27:934–946.

[18] Folstein MF, Folstein SE, McHugh PR. Mini-Mental State. J Psych Res. 1975;12:189–198.

[19] Thompson PM, Hayashi KM, de Zubicaray GI, Janke AL, Rose SE, Semple J, et al. Mapping hippocampal and ventricular change in Alzheimer disease. NeuroImage. 2004;22(4):1754–1766.

[20] Duan JH, Wang HQ, Xu J, Lin X, Chen SQ, Kang Z, et al. White matter damage of patients with Alzheimer's disease correlated with the disease cognitive function. Surg Radiol Anat. 2006;28:150–156.

[21] Baxter LC, Sparks DL, Johnson SC, Lenoski B, Lopez JE, Connor DJ, et al. Relationship of cognitive measures and gray and white matter in Alzheimer's disease. Journal of Alzheimer's disease. 2006;9(3):253–260.

[22] Apostolova LG, Lu PH, Dutton RA, Hayashi KM, Toga AW, Cummings JL, et al. 3D mapping of mini-mental state examination performances in clinical and preclinical Alzheimer disease. Alzheimer Dis Assoc Disord. 2006;20(4):224–231.

[23] Magnaldi S, Ulkmar M, Vasciaveo A, Longo R, Pozzi-Mucelli RS. Contrast between white and grey matter: MRI apperance with aging. European Radiology. 1993;3(6):513–519.

[24] Admiraal-Behloul F, van den Heuvel DMJ, Olofsen H, van Osch MJP. Fully automatic segmentation of white matter hyperintensities in MR images of the elderly. Neuroimage. 2005;28(3):607–617.

[25] Woods RP, Grafton ST, Holmes CJ, Chery SR, Mazziotta JC. Automated Image Registration I. General methods and intrasubject, intramodality validation. J Comput Assist Tomogr. 1998;22:139–152.

[26] Admiraal-Behloul F, van den Heuvel D, Olofsen H, Schmitz N, van Buchem MA. Brain templates for the elderly. In: International Society for Magnetic Resonance in Medicine (ISMRM) - Toronto; 2003. .

[27] Ferrarini L, Olofsen H, Palm WM, van Buchem MA, Reiber JHC, Admiraal-Behloul F. GAMEs: Growing and Adaptive Meshes for Fully Automatic Shape Modeling and Analysis. Medical Image Analysis. 2007;11(3):302–314.

[28] van der Flier WM, van der Vlies AE, Wervelings AWE, de Boer NL, Admiraal-Behloul F, Bollen ELEM, et al. MRI measures and progression of cognitive decline in nondemented elderly attending memory clinic. Int J Geriatric Psychiatry. 2005;20(11):1060–1066.

[29] Fox NC, Scahill RI, Crum WR, Rossor MN. Correlation between rates of brain atrophy and cognitive decline in AD. Neurology. 1999;52(8):1687–1688.

[30] Duarte A, Hayasaka S, Du A, Schuff N, Jahng G, Kramer J, et al. Volumetric correlates of memory and executive function in normal elderly, mild cognitive impairment and Alzheimer's Disease. Neuroscience Letter. 2006;406(1-2):60–65.

[31] Pohjasvaara T, Mantyla R, Salonen O, Aronen HJ, Ylikoski R, Hietanem M, et al. MRI correlates of dementia after first clinical ischemic stroke. J neurol sci. 2000;181(1-2):111–117.

CHAPTER 7

Summary and Conclusions

Concern for man and his fate must always form the chief interest of all technical endeavors. Never forget this in the midst of your diagrams and equations -

Try not to become a man of success, but rather a man of value.

Albert Einstein, 1879-1955.

7.1 Summary and Conclusions

Since the first case of Alzheimer Disease was reported, many efforts have been spent to investigate the causes and mechanisms of this pathology. Therapies developed to slow down the cognitive decline have been proved more effective if administered during the early stage of the disease: consequently, biomarkers for early diagnosis are highly desirable. Growing evidence of the correlation between the progression of AD and the loss of neurons in specific areas of the brain has been reported in the literature. Clinical MR images are a valuable tool for the estimation of volume loss and shape changes in well-defined brain structures: tissue degeneration due to AD has been shown both in gray matter structures, such as hippocampus and amygdala, and in white matter structures, such as the corpus callosum. Thus, investigating changes in these areas might increase our knowledge on the progression of AD. Preliminary to any kind of analysis is the segmentation of the structures of interest. The delineation of small white and gray matter structures remains a challenging task, particularly in elderly subjects: the contrast between white and gray matter decreases with age, yielding to error-prone and less reproducible (semi-) automatic segmentation algorithms. Manual delineation, on the other hand, is more subjective and time consuming. The analysis of brain ventricles could lead to more reproducible results since the contrast between CSF and the rest of the parenchyma remains sharp with aging. Moreover, any change in volume and shape due to atrophy in periventricular structures must be reflected on the shape and volume of the ventricles. The focus of this research was to investigate changes in the ventricular shape within different populations (i.e. healthy subjects, MCIs, and ADs), as well as across the whole spectrum of cognitive impairment. The most relevant results are summarized in this section.

Chapter 1 introduces the general issues investigated in this manuscript: starting with a brief overview on the history of medical imaging, the chapter continues presenting the anatomical structures of interest for this research, and the Magnetic Resonance technique used to acquire brain images. Subsequently, an overview of Alzheimer Disease is given, covering hypotheses on its causes, societal and economical effects, cognitive and anatomical consequences, and relevant research challenges. The chapter concludes with the issue of shape modeling and analysis: the advantages of shape-based statistical models are presented, together with some of the most relevant methods published in the literature.

Chapter 2 presents a newly developed method for shape modeling and analysis: Growing and Adaptive MEshes (GAMEs). The method is based on Self-Organizing Networks which Grow When Required and was introduced by Marsland et al. [1]. In a first phase, a mesh is initialized with only two nodes randomly located in space: surface points from an average instance are iteratively given to the network, which grows and adapts to model the instance's surface. After convergence, the mesh's topology is frozen, and the mesh's shape is adapted to other similar instances: the set of comparable meshes form a Point Distribution Model. Several aspects of the PDM were tested on challenging synthetic shapes: reproducibility, robustness to noise, and accuracy in detecting corresponding landmarks. Given the PDM, a statistical model based on Principal Component Analysis was built and tested for compactness, reconstruction error, generalization error, and specificity error. All the results proved to

be satisfactory, when compared to similar approaches for medical image analysis. Moreover, we tested the method's performances in highlighting local shape differences between populations of similar instances: local comparisons with permutation tests highlighted the exact locations where shape differences had been artificially introduced. Finally, we applied the method to the modeling of two populations of brain ventricles, controls and AD: the statistical models were used to prove the good performances of our method in a challenging medical application.

In Chapter 3, the method previously introduced was applied to the analysis of two populations of brain ventricles: patients with Alzheimer Disease, and age-matched healthy subjects. The controls were used to create an average ventricle which was used to generate a first mesh. Comparable meshes were then created to build up a PDM. A first analysis focused on intra-population variability, showing the main modes of variation within each population by means of tensor maps. Both in controls and AD, we could highlight asymmetrical variations between the left and right inferior temporal horns: AD were characterized by inferior-superior widening in the right inferior medial temporal horn, and a more isotropic enlargement in the left one; controls showed anterior-posterior elongation in the right inferior temporal horn, and an inferior-superior thickening in the left one. Subsequently, permutation tests were applied to detect local shape differences between populations: results showed that about 22% of the total surface was significantly different (at a confidence level of 95%), of which 14% was on the left side. By visual inspection, we could infer significant changes in areas close to temporal horns, splenium of the corpus callosum, amygdala, thalamus, head of the left caudate nuclei, and tail of caudate nuclei. Finally, for each of the significantly different regions we evaluated displacement vectors showing the local deformations from an average control to an average AD ventricle.

Chapter 4 investigates the existence of ventricular shape-based biomarkers for AD in clinical MR images. Considering two populations of AD and age-matched healthy individuals, repeated permutation tests were run on randomly sub-sampled AD groups and controls to highlight locations on the ventricle's surface which consistently discriminated the two groups: by applying different thresholds on the p values, more specific regions could be identified, such as the tips of the inferior temporal horns, the right superior temporal horn, the head of the right caudate nuclei, and the genu of the corpus callosum. These areas were then used to train a support vector machine to discriminate between healthy subjects and ADs. The most successful classification was achieved when only the tips of the inferior temporal horns were used: leave 1 out experiments showed an overall accuracy of 84%, while the sensitivity on a previously unseen AD dataset reached 76%. Our results were compared with previous methods [2–6], proving the good performances of our approach.

Chapter 5 extends the results presented in Chapter 3 to other populations, covering a larger part of the cognitive spectrum. Healthy elderly are compared to Memory Complainers, patients with Mild Cognitive Impairment, and subjects affected by AD. Permutation tests highlighted different patterns of shape differences for the different groups, supporting the classification based on neuropsychological tests. Shape changes (both in terms of extent of affected

surface and severity of changes) reflected atrophy in the surrounding structures: different regions were delineated along the ventricular surface, and results were clustered according to the adjacent periventricular structures. When comparing controls versus a particular group (e.g., AD), we could indicate relevant asymmetries in the way the pathology affects periventricular structures (e.g. major atrophy on the right side for MC and MCI, and on the left side for AD); when focusing on a given region (e.g., close to the caudate nuclei), we could investigate trends across the cognitive spectrum.

In Chapter 6, we correlate local ventricular enlargements in brain ventricles with the degree of cognitive impairment. Subjects were included to cover the entire spectrum of cognitive function, ranging from healthy controls to severe cases of AD. The degree of cognitive impairment was estimated, for each subject, with the Mini Mental State Examination (MMSE). Atrophy in periventricular structures was inferred by shape analysis of the ventricular surface. The amount of local enlargement in the ventricle, with respect to an average normal control, was used to estimate the severity of focal atrophy in periventricular structures; the amount of surface significantly different between a given subject and the average control was used to estimate the extent of atrophy in periventricular structures. The surface of the ventricles was clustered in different regions, each corresponding to adjacent anatomical structures. Linear regression analysis was used to correlate both the severity of atrophy and the extent of atrophy with the MMSE score. Results showed that the severity of atrophy correlates with cognitive impairment in well-defined periventricular structures: left side of the thalamus, left temporal horn, left corona radiata, and right caudate nuclei. The extent of atrophy, on the other hand, does not correlate in any of the analyzed regions. Results suggest that worse cognitive conditions are a consequence of an increased severity of atrophy in well-defined periventricular structures.

A better understanding on how Alzheimer Disease affects the brain is an important step towards the development of new therapies and the discovery of potential biomarkers. Quantitative measurements of localized atrophy in MR images are extremely important to this goal. In this thesis, we have shown that quantitative analysis of ventricular shape changes is a powerful tool to estimate atrophy in periventricular structures. A new method for shape modeling and analysis of brain ventricles was introduced and thoroughly validated, both on challenging synthetic shapes and medical images. Combining our method with non-parametric statistical tests and pattern recognition tools, we achieved new insights on how AD and MCI affect periventricular structures. Further analyses based on linear regression highlighted the correlation between cognitive impairment (assessed by the MMSE score) and focal atrophy. In conclusion, the goals of this research stated in Section 1.3 have been achieved.

7.2 Future Work

The shape modeling algorithm introduced in this thesis presents some advantages when compared to other methods: it can model objects with non-spherical topology, it is highly automatic, and its surface parametrization is straightforwardly linked to the shape properties of

the objects. Nevertheless, some characteristics could be improved in the future. First of all, the Euclidian and Mahalanobis distance functions used in the growing and adaptive phase are based on spatial information; different distance functions considering surface connectivity should be implemented to improve the results in particular applications: for instance, distance measurements between locations on the brain cortex should take into account whether the nodes are on different gyri. Moreover, additional information could be introduced: nodes in a mesh are currently described by their spatial coordinates; the method could easily be extended to include a fourth dimension, such as intensity or a given label, to facilitate the modeling of non-binary structures or to allow for the simultaneously modeling of different structures. Finally, because of its generality, the method could be used for different applications.

The results achieved in the analysis of brain ventricles suggest several new research directions. The discrimination between MCI and AD could be further investigated in a longitudinal study: future analyses could incorporate knowledge on the MCI state (converters or non-converters). A longitudinal comparison between healthy individuals and subjects developing AD would also provide more insights on the developing of the disease. Moreover, the correlation between cognitive decline and atrophy in periventricular structures could be strengthened by using different neuropsychological tests, such as the Camcog.

The combination of different image modalities and more invasive techniques could also improve the diagnosis of dementia in general, and AD in particular. Recent studies have shown that CSF biomarkers (i.e. concentration of beta amyloid and total tau) could improve the detection of AD [7], especially if combined with MR-based quantitative analysis [8]. Other studies have investigated the use of nuclear imaging (PET) to highlight the increase of amyloid plaques in AD and MCI [9–11]. Finally, studies on functional MRI have shown that differences in resting-state activation networks could be used to discriminate between healthy subjects, MCIs and ADs [12–14].

Bibliography

[1] Marsland S, Shapiro J, Nehmzow U. A self-organizing network that grows when required. Neural Networks. 2002;15:1041–1058.

[2] Barnes J, Godbolt AK, Frost C, Boyes RG, Jones BF, Scahill RI, et al. Atrophy rates of the cingulate gyrus and hippocampus in AD and FTLD. Neurobiology of Aging. 2007;28:20–28.

[3] Barnes J, Whitwell JL, Frost C, Josephs KA, Rossor MN, Fox NC. Measurements of the Amygdala and Hippocampus in Pathological Confirmed Alzheimer Disease and Frontotemporal Lobar Denegeneration. Arch Neurol. 2006;63:1434–1439.

[4] Kodama N, Shimada T, Fukumoto I. Image-based Diagnosis of Alzheimer-type Dementia: Measurements of Hippocampal and Ventricular Areas in MR Images. Magn Reson Med Sci. 2002;1(1):14–20.

[5] Tarroun A, Bonnefoy M, Bouffard-Vercelli J, Gedeon C, Vallee B, Cotton F. Could linear MRI measurements of hippocampus differentiate normal brain aging in elderly persons from Alzheimer disease? Surg Radiol Anat. 2007;29(1):77–81.

[6] Wang L, Swank JS, Glick IE, Gado MH, Miller MI, Morris JC, et al. Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. NeuroImage. 2003;20:667–682.

[7] Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L. Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impariment: a follow-up study. Lancet Neurol. 2006;5:228–234.

[8] Leon MJ, DeSanti S, Zinkowski R, Mehta PD, Pratico D, Segal S, et al. MRI and CSF studies in the early diagnosis of Alzheimer's disease. Journal of Internal Medicine. 2004;256:205–223.

[9] Choo IH, Lee DY, Youn JC, Jhoo JH, Kim KW, Lee DS, et al. Topographic patterns of brain functional impairment progression according to clinical severity staging in 116 Alzheimer Disease Patients: FDG-PET Study. Alzheimer Disease Assoc Disord. 2007;21(2):77–84.

[10] Kemppainen NM, Aalto S, Wilson IA, Nagren K, Helin S, Brück A, et al. PET amyloid ligand [11C]PIB uptake is increased in mild cognitive impairment. Neurology. 2007;68(19):1603–1606.

[11] Rowe CC, Nq S, Ackermann U, Gong SJ, Pike K, Savage G, et al. Imaging beta-amyloid burden in aging and dementia. Neurology. 2007;68(20):1718–1725.

[12] Rombouts SARB, Barkhof F, Goekoop R, Stam CJ, Scheltens P. Altered Resting State Networks in Mild Cognitive Impairment and Mild Alzheimer's Disease: An fMRI Study. Human Brain Mapping. 2005;26:231–239.

[13] Wang K, Jiang T, Liang M, Wang L, Tian L, Zhang X, et al. Discriminative Analysis of Early Alzheimer's Disease Based on Two Intrinsically Anti-correlated Networks with Resting-State fMRI. MICCAI. 2006;LNCS 4191:340–347.

[14] Wang L, Zang Y, He Y, Liang M, Zhang X, Tian L, et al. Changes in hippocampal connectivity in the early stages of Alzheimer's disease: Evidence from resting state fMRI. NeuroImage. 2006;31:496–504.

CHAPTER 8

Samenvatting en Conclusies

Niels Bohr and Albert Einstein at Ehrenfest's house in Leiden (December 1925)

Prediction is always difficult, especially if it's about the future -

An expert is a man who has made all the mistakes which can be made in a very narrow field.

Niels Bohr, 1885-1962.

8.1 Samenvatting en Conclusies

Sinds het eerste geval van Alzheimer Disease werd gerapporteerd, zijn er vele pogingen gedaan om de oorzaken en de uitwerkingen van deze ziekte te onderzoeken. Therapieën die ontwikkeld zijn om de cognitieve daling te vertragen, hebben bewezen efficiënter te zijn als deze al een vroeg stadium van de ziekte toegepast worden: daarom zijn biomarkers voor een vroege diagnose zeer wenselijk. In de literatuur is wezenlijk bewijs van de correlatie tussen de vooruitgang van AD en het verlies van neuronen in specifieke gebieden van de hersenen beschreven. Klinische MRI beelden zijn een waardevol hulpmiddel om volumeverlies en vormveranderingen in specifieke hersenenstructuren te schatten: weefseldegeneratie als gevolg van AD is zowel in de grijze stof (zoals de hippocampus en de amygdala) als in de witte stof (zoals in het corpus callosum) aangetoond. Dus zou het analyseren van veranderingen in deze gebieden onze kennis over de progressie van AD kunnen vergoten. Voordat het mogelijk is om een analyse te doen is, moeten de structuren die van belang zijn worden gesegmenteerd. Een onderscheid maken tussen kleine structuren in de witte en grijze stof is een moeilijke taak, vooral in de hersenen van ouderen: het contrast tussen witte en grijze stof vermindert met leeftijd en daarom kunnen de (half-)automatische segmentatie algoritmen falen en zijn minder reproduceerbaar. Handmatige segmentatie, anderzijds, is subjectiever en tijdrovender. De analyse van hersenventrikels kan tot reproduceerbaardere resultaten leiden, omdat het onderscheid tussen het CSF (hersenvocht) en de rest van het parenchym (hersenweefsel) met het verouderen duidelijk zichtbaar blijft. Bovendien spiegelen verandering in volume en vorm van periventriculaire structuren als gevolg van AD zich af op het volume en vorm van de hersenventrikels. In dit onderzoek wilden we de veranderingen van hersenventrikels analyseren, zowel binnen verschillende populaties (d.w.z. gezonde individuen, MCI, en AD) als in de breedte van de cognitieve spectrum. De meest relevante resultaten worden in deze sectie samengevat.

Hoofdstuk 1 introduceert de algemene onderzochte kwesties in dit manuscript: het begint met een kort overzicht van de geschiedenis van medische beeldvorming, waarna het hoofdstuk anatomische structuren introduceert die belangrijk zijn voor dit onderzoek, samen met de Magnetische Resonantie techniek die wordt gebruikt om hersenenbeelden te verwerven. Vervolgens wordt een overzicht van de Ziekte van Alzheimer (AD) gegeven, waarin hypothesen over de oorzaken, de sociale en economische gevolgen en de cognitieve en anatomische uitwerkingen behandeld worden, alsmede de openstaande onderzoeksvragen voor relevant onderzoek. Het hoofdstuk besluit met het modelleren en analyseren van vormen: de voordelen van op vorm gebaseerde statistische modellen worden gegeven samen met een aantal van de meest relevante methodes die in de literatuur zijn gepubliceerd.

Hoofdstuk 2 introduceert een nieuwe methode van vorm modellering en analyse: Growing and Adaptive MEshes (GAMEs). De methode is gebaseerd op de Self-Organizing Networks which Grow When Required, geïntroduceerd door Marsland et al. [1]. In de eerste fase wordt een netwerk geïnitialiseerd met slechts twee knooppunten die willekeurig in ruimte worden geplaatst: punten aan het oppervlak van een gemiddelde instantiatie worden achtereenvolgens aan het netwerk gegeven, zodat het groeit en zich aanpast om zo de oppervlakte van de

instantiatie te modelleren. Na convergentie wordt de topologie van het netwerk bevroren en wordt de vorm van het netwerk aangepast aan andere, gelijksoortige instantiaties: deze set van vergelijkbare netwerken vormt een Point Distribution Model (PDM). Verscheidene aspecten van het PDM werden getest met complexe synthetische vormen: reproduceerbaarheid, robuustheid voor ruis en nauwkeurigheid in het ontdekken van overeenkomstige oriëntatiepunten. Aan de hand van het PDM werd een statistisch model, gebaseerd op de Principal Component Analysis, geconstrueerd en getest voor compactheid, reconstructiefout, generalisatiefout, en specificiteitfout. Alle resultaten bleken voldoende vergeleken met gelijksoortige methodes voor medische beeldanalyse. Bovendien hebben wij de prestaties van de methode getest op het accentueren van lokale vormverschillen tussen populaties van gelijksoortige instantiaties: lokale vergelijkingen met permutatietests benadrukten de exacte plaatsen waar de vormverschillen kunstmatig waren geïntroduceerd. Ten slotte pasten wij de methode toe op het modelleren van twee groepen hersenventrikels; gezonde individuen en AD. De statistische modellen werden gebruikt om de goede prestaties van onze methode in een moeilijke medische toepassing te bewijzen.

In Hoofdstuk 3, werd de eerder geïntroduceerde methode toegepast op de analyse van twee populaties van hersenventrikels: patiënten met de Ziekte van Alzheimer en gezonde individuen van vergelijkbare leeftijd. Vanuit de controle groep werd een gemiddeld ventrikel geconstrueerd, dat werd gebruikt om een eerste netwerk te produceren. Daarna werden soortgelijke netwerken gecreëerd om een PDM te construeren. Een eerste analyse concentreerde zich op intra-groep variatie die de belangrijkste vormen van variatie binnen elke groep aantoont door middel van tensormaps. Zowel in de controle groep als in de AD populatie konden wij asymmetrische variaties tussen de linker en rechter inferieure temporale hoornen laten zien: de AD werd gekenmerkt door inferieure-superieure verwijdingen in de rechter inferieure mediale temporale hoorn en een meer isotrope vergroting in linker; de controles toonden een anterieure-posterieure verlenging in de rechter inferieure temporale hoorn en een inferieure-superieure verdikking in linker. Daarna werden de permutatietests toegepast om lokale vormverschillen tussen populaties te ontdekken: de resultaten lieten zien dat ongeveer 22% van de totale oppervlakte significant verschillend was met een confidence interval van 95%, waarvan 14% aan de linkerkant zat. Door visuele inspectie konden wij significante veranderingen aanwijzen in gebieden dicht bij temporale hoornen, splenium van het corpus callosum, amygdala, thalamus, kop van de linker nucleus caudatus en de staart van de nucleus caudatus. Ten slotte werden voor elk van de significant verschillende gebieden de verplaatsingsvectoren geëvalueerd die de lokale vervormingen aangaven tussen een gemiddelde controle ventrikel en een gemiddeld ventrikel van de AD.

Hoofdstuk 4 onderzoekt het bestaan van ventriculaire vormgebaseerde biomarkers voor AD in klinische MR beelden. Twee populaties werden samengesteld (AD en gezonde individuen van vergelijkbare leeftijd), waarbij herhaalde permutatietests werden uitgevoerd op de willekeurige sub-groepen van de AD en de gezonde individuen om vaste plaatsen op de oppervlakte van het ventrikel te vinden die de twee groepen onderscheidde: door verschillende thresholds op de p-waarden konden we specifiekere gebieden identificeren, zoals de uiteinden van de inferieure temporale hoornen, de rechter superieure temporale hoorn, de

kop van de rechter nucleus caudatus, en de genu van het corpus callosum. Deze gebieden werden gebruikt om een support vector machine te trainen om onderscheid te maken tussen gezonde individuen en AD. De meest succesvolle classificatie werd bereikt toen slechts de uiteinden van de inferieure temporale hoornen werden gebruikt: ”leave 1 out” experimenten lieten een algemene nauwkeurigheid van 84% zien, terwijl de sensitiviteit op een onbekende dataset van de AD 76% haalde. Onze resultaten werden vergeleken met eerdere methodes [2–6] waardoor de goede prestaties van onze benadering worden bewezen.

Hoofdstuk 5 breidt de resultaten die in Hoofdstuk 3 gepresenteerd zijn, uit naar andere populaties die een groter deel beslaan van het cognitieve spectrum. Gezonde ouderen worden vergeleken met ouderen met geheugenklachten (Memory Complainers), milde cognitieve beschadigingen (Mild Cognitive Impairment), en de ziekte van alzheimer (AD). Permutatie tests lieten verschillende voorbeelden van vormverschillen voor de verschillende groepen zien die overeen kwamen met de classificatie op basis van neuropsychologische tests. Vormveranderingen (zowel in omvang van aangetast ventrikeloppervlakte als de mate van de veranderingen) wezen op atrofie in de aangrenzende structuren: de verschillende gebieden werden ingetekend langs de ventriculaire oppervlakte en de resultaten werden geclusterd volgens de aangrenzende periventriculaire structuren. Door controles te vergelijke met een bepaalde groep (b.v., AD), konden we relevante asymmetrieën als gevolg van de invloed van de pathologie op de periventriculaire structuren aanwijzen (b.v. belangrijke atrofie aan de rechterkant voor MC en MCI, en aan de linkerkant voor AD); door te concentreren op een bepaald gebied (b.v. dicht bij de nucleus caudatus), konden we trends over het cognitieve spectrum onderzoeken.

In Hoofdstuk 6, correleren we lokale ventriculaire vergrotingen van de hersenenventrikels aan de mate van de cognitieve beschadiging. Patiënten werden toegevoegd om het volledige spectrum van cognitieve functie te dekken, dat zich uitstrekt van gezonde ouderen tot aan ernstige gevallen van AD. De mate van de cognitieve beschadiging werd voor elk individu geschat met het Mini Mental State Examination (MMSE). Atrofie in periventriculaire structuren werd aangeduid door vorm analyse van de ventriculaire oppervlakte. De mate van lokale verwijding van het ventrikel, vergeleken met een gemiddelde normale ventrikel, werd gebruikt om de graad van atrofie in periventriculaire structuren te schatten; het aantal significant verschillende oppervlaktes tussen een bepaalde patiënt en de gemiddelde controle groep werd gebruikt om de omvang van atrofie in periventriculaire structuren te schatten. De ventrikeloppervlakte werd verdeeld in verschillende gebieden corresponderend aan de aangrenzende anatomische structuren. De lineaire regressie analyse werd gebruikt om de mate en de omvang van atrofie met de MMSE score te correleren. De resultaten bewezen dat de mate van atrofie correleert met de cognitieve beschadigingen in specifieke periventriculaire structuren: de linkerkant van de thalamus, de linker temporale hoorn, de linker corona radiata, en de rechter nucleus caudatus. Echter, de atrofie omvang correleert met geen enkele van de geanalyseerde gebieden. De resultaten suggereren dat de slechtere cognitieve condities een gevolg zijn van een hogere mate van atrofie in specifieke periventriculaire structuren.

Een beter begrip over hoe de Ziekte van Alzheimer de hersenen beïnvloedt is een belan-

grijke stap richting de ontwikkeling van nieuwe therapieën en de ontdekking van potentiële biomarkers. De kwantitatieve metingen van gelokaliseerde atrofie in MR beelden zijn enorm belangrijk voor dit doel. In deze thesis, hebben wij aangetoond dat de kwantitatieve analyse van ventriculaire vormveranderingen een krachtig hulpmiddel is om atrofie in periventriculaire structuren te schatten. Een nieuwe methode voor modellering van vorm en analyse van hersenenventrikels werd geïntroduceerd en grondig gevalideerd, zowel op complexe synthetische vormen als medische beelden. Door onze methode te combineren met niet-parametrische statistische tests en patroonerkennings methoden, hebben wij nieuw inzicht bereikt in hoe de AD en MCI periventriculaire structuren beïnvloeden. De analyses, gebaseerd op lineaire regressie, benadrukten de correlatie tussen cognitieve stoornissen (die door de MMSE score wordt beoordeeld) en gelokaliseerde atrofie. Samenvattend, de doelstellingen van dit onderzoek, die in Sectie 1.3 worden genoemd, zijn gehaald.

8.2 Aanbevelingen

Het Shape Modeling algoritme dat in deze thesis wordt geïntroduceerde heeft een aantal voordelen vergeleken met andere methodes: het kan voorwerpen met niet-sferische topologie modelleren, het is hoog automatisch, en zijn oppervlakte- parameterisatie is rechtstreeks verbonden met de vormeigenschappen van de objecten. Niettemin zouden sommige eigenschappen in de toekomst kunnen worden verbeterd. Allereerst zijn de Euclidische en Mahalanobis afstandsfuncties die in de groei- en de aanpassingsfasen worden gebruikt, gebaseerd op spatiële informatie; andere afstandsfuncties die oppervlakte connectiviteit berekenen, moeten worden geïmplementeerd om de resultaten van specifieke toepassingen te verbeteren: bijvoorbeeld op het hersenschors zouden afstandsfuncties rekening moeten houden met knopen die op verschillende gyri liggen. Verder zou extra informatie kunnen worden toegevoegd: de knopen in een netwerk worden nu beschreven door hun spatiële coördinaten; de methode kan gemakkelijk worden uitgebreid met een vierde dimensie, zoals intensiteit of een bepaald label, om ook niet binaire structuren of verschillende structuren tegelijkertijd te modelleren. Ten slotte zou de methode voor verschillende toepassingen kunnen worden gebruikt, omdat hij zeer generiek is.

De bereikte resultaten in de analyse van hersenenventrikels impliceren verscheidene nieuwe onderzoeksrichtingen. Het onderscheid tussen MCI en de AD zou in een longitudinale studie verder onderzocht kunnen worden: toekomstige analyses zouden kennis kunnen opnemen omtrent de MCI status (converters of non-converters). Een longitudinale vergelijking tussen gezonde individuen en individuen die AD ontwikkelen zou ook meer inzicht geven in het zich ontwikkelen van de ziekte. Bovendien zou de correlatie tussen cognitieve daling en atrofie in periventriculaire structuren kunnen worden versterkt door deze te vergelijken met verschillende neuropsychologische tests, zoals de Camcog.

Ook kan de combinatie van verschillende beeldmodaliteiten en meer invasieve technieken de diagnose van dementie in het algemeen en AD in het bijzonder verbeteren. Recente studies hebben aangetoond dat CSF biomarkers, d.w.z. concentratie van beta-amyloid en totale tau, de opsporing van AD konden verbeteren [7], vooral wanneer ze gecombineerd worden met op MR-gebaseerde kwantitatieve analyse [8]. Andere studies hebben het gebruik van nucleare

beeldvormende technieken (PET) onderzocht om de verhoging van amyloid plaques bij AD en MCI te meten [9–11]. Ten slotte hebben de studies met functionele MRI aangetoond dat in de resting state de verschillen in netwerken gebruikt zouden kunnen worden om onderscheid te maken tussen gezonde patiënten, MCIs en ADs [12–14].

Bibliography

[1] Marsland S, Shapiro J, Nehmzow U. A self-organizing network that grows when required. Neural Networks. 2002;15:1041–1058.

[2] Wang L, Swank JS, Glick IE, Gado MH, Miller MI, Morris JC, et al. Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. NeuroImage. 2003;20:667–682.

[3] Kodama N, Shimada T, Fukumoto I. Image-based Diagnosis of Alzheimer-type Dementia: Measurements of Hippocampal and Ventricular Areas in MR Images. Magn Reson Med Sci. 2002;1(1):14–20.

[4] Barnes J, Godbolt AK, Frost C, Boyes RG, Jones BF, Scahill RI, et al. Atrophy rates of the cingulate gyrus and hippocampus in AD and FTLD. Neurobiology of Aging. 2007;28:20–28.

[5] Barnes J, Whitwell JL, Frost C, Josephs KA, Rossor MN, Fox NC. Measurements of the Amygdala and Hippocampus in Pathological Confirmed Alzheimer Disease and Frontotemporal Lobar Denegeneration. Arch Neurol. 2006;63:1434–1439.

[6] Tarroun A, Bonnefoy M, Bouffard-Vercelli J, Gedeon C, Vallee B, Cotton F. Could linear MRI measurements of hippocampus differentiate normal brain aging in elderly persons from Alzheimer disease? Surg Radiol Anat. 2007;29(1):77–81.

[7] Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L. Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impariment: a follow-up study. Lancet Neurol. 2006;5:228–234.

[8] Leon MJ, DeSanti S, Zinkowski R, Mehta PD, Pratico D, Segal S, et al. MRI and CSF studies in the early diagnosis of Alzheimer's disease. Journal of Internal Medicine. 2004;256:205–223.

[9] Choo IH, Lee DY, Youn JC, Jhoo JH, Kim KW, Lee DS, et al. Topographic patterns of brain functional impairment progression according to clinical severity staging in 116 Alzheimer Disease Patients: FDG-PET Study. Alzheimer Disease Assoc Disord. 2007;21(2):77–84.

[10] Rowe CC, Nq S, Ackermann U, Gong SJ, Pike K, Savage G, et al. Imaging beta-amyloid burden in aging and dementia. Neurology. 2007;68(20):1718–1725.

[11] Kemppainen NM, Aalto S, Wilson IA, Nagren K, Helin S, Brück A, et al. PET amyloid ligand [11C]PIB uptake is increased in mild cognitive impairment. Neurology. 2007;68(19):1603–1606.

[12] Rombouts SARB, Barkhof F, Goekoop R, Stam CJ, Scheltens P. Altered Resting State Networks in Mild Cognitive Impairment and Mild Alzheimer's Disease: An fMRI Study. Human Brain Mapping. 2005;26:231–239.

[13] Wang L, Zang Y, He Y, Liang M, Zhang X, Tian L, et al. Changes in hippocampal connectivity in the early stages of Alzheimer's disease: Evidence from resting state fMRI. NeuroImage. 2006;31:496–504.

[14] Wang K, Jiang T, Liang M, Wang L, Tian L, Zhang X, et al. Discriminative Analysis of Early Alzheimer's Disease Based on Two Intrinsically Anti-correlated Networks with Resting-State fMRI. MICCAI. 2006;LNCS 4191:340–347.

Publications

Publications in International Journals

Luca Ferrarini, Walter M. Palm, Hans Olofsen, Roald van der Landen, Gerard J. Blauw, Rudi G.J. Westendorp, Eduard L.E.M. Bollen, Huub A.M. Middelkoop, Johan H.C. Reiber, Mark A. van Buchem, Faiza Admiraal-Behloul, *"MMSE Scores Correlate with Local Ventricular Enlargement in the Spectrum from Cognitively Normal to Alzheimer Disease"*, NeuroImage, Accepted for publication in November 2007

Luca Ferrarini, Walter M. Palm, Hans Olofsen, Roald van der Landen, Mark A. van Buchem, Johan H.C. Reiber, Faiza Admiraal-Behloul, *"Ventricular Shape Biomarkers for Alzheimer's Disease in Clinical MR Images"*, Magnetic Resonance in Medicine (MRM), Accepted for publication in October 2007

Luca Ferrarini, Hans Olofsen, Walter M. Palm, Mark A. van Buchem, Johan H.C. Reiber, Faiza Admiraal-Behloul, *"GAMEs: Growing and Adaptive Meshes for Fully Automatic Shape Modeling and Analysis"*, Medical Image Analysis, 11(3), 302-314, 2007

L. Ferrarini, W.M. Palm, H. Olofsen, M.A. van Buchem, J.H.C. Reiber, F. Admiraal-Behloul, *"Shape differences of the brain ventricles in Alzheimer's disease"*, NeuroImage 32 (3), pp. 1060-1069, 2006

L. Ferrarini, L. Bertelli, J. Feala, A.D. McCulloch, and G. Paternostro, *"A more efficient search strategy for aging genes based on connectivity"*, Bioinformatics 21 (3), pp. 338-348, 2005

Publications in Conference Proceedings

L. Ferrarini, H. Olofsen, W.M. Palm, M.A. van Buchem, J.H.C. Reiber, F. Admiraal-Behloul, *"Growing Cell Neural Networks for Fully Automatic Shape Modeling"*, Medical Image Understanding and Analysis (MIUA), 2006

L. Ferrarini and H. Olofsen and M.A. van Buchem and J.H.C. Reiber and F. Admiraal-Behloul, *"Growing Cell Neural Networks: A Pattern Recognition Framework for Fully Automatic Shape Modeling"*, ASCI Conference 2006, Lommel - Belgium

Luca Ferrarini, Hans Olofsen, Johan H.C. Reiber, and Faiza Admiraal-Behloul, *"A NeuroFuzzy Controller for 3D Virtual Centered Navigation in Medical Images of Tubular Structures"*, International Conference on Artificial Neural Networks - ICANN 2005, LNCS 3697, pp 371-376, Warsaw - Poland

Luca Ferrarini, Hans Olofsen, Mark A. van Buchem, Johan H.C. Reiber, and Faiza Admiraal-Behloul, *"Fully automatic shape modeling using growing cell neural networks"*, International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI 2005, LNCS 3750, pp. 451-458, Palm Springs - U.S.A.

Ferrarini Luca, Hans Olofsen, Johan H.C. Reiber, Faiza Admiraal-Behloul, *"Exploration of 3D Medical Images of Tubular Structures: A NeuroFuzzy Controller for Virtual Centered Navigation"*, ASCI Conference 2005, Heijen - The Netherlands

Faiza Admiraal-Behloul, Boudewijn .P.F. Lelieveldt, **Luca Ferrarini**, Hans Olofsen, Rob van der Geest and J.H.C Reiber, *"A Virtual Exploring Mobile Robot for Left Ventricle Contour Tracking"*, International Joint Conference on Neural Networks 2004 (IJCNN 2004), Budapest - Hungary

Book Chapters

L. Ferrarini, H. Olofsen, J.H.C. Reiber, F. Admiraal-Behloul, *"Autonomous Virtual Mobile Robot for the Exploration of 3D Medical Images"*, In *Medical Robots* - Advanced Robotic Systems (ISBN 978-3-902613-18-9)

Published Abstracts

L. Ferrarini, H. Olofsen, W.M. Palm, M.A. van Buchem, J.H.C. Reiber, F. Admiraal-Behloul, *"Ventricular Shape Biomarkers for Alzheimer's disease in Clinical MR images"*, International Society for Magnetic Resonance in Medicine - ISMRM 2007. Berlin - Germany

H. Olofsen, **L. Ferrarini**, W.M. Palm, M.A. van Buchem, J.H.C. Reiber, F. Admiraal-Behloul, *"Local Volumetric Analysis of the Brain Ventricles in Alzheimer's Disease using MRI"*, International Society for Magnetic Resonance in Medicine - ISMRM 2007. Berlin - Germany

L. Ferrarini, W.M. Palm, H. Olofsen, M.A. van Buchem, J.H.C. Reiber, F. Admiraal-Behloul, *"Significant local shape differences of the brain ventricles in Alzheimer's patients compared to healthy elderly"*, International Society for Magnetic Resonance in Medicine - ISMRM 2006. Seattle - U.S.A.

Awards

Honorable mention as second best paper at Medical Image Understanding and Analysis (MIUA), 2006. **L. Ferrarini**, H. Olofsen, W.M. Palm, M.A. van Buchem, J.H.C. Reiber, F. Admiraal-Behloul, *"Growing Cell Neural Networks for Fully Automatic Shape Modeling"* Steering committee and organizers: Dr E. Claridge, Dr J. Graham, Prof. D. Hill, Dr M. Mirmehdi, Prof. A. Noble, Dr D. Rueckert, Dr N. Thacker, Dr S.M. Astley, Dr T.F. Cootes, Dr J. Graham, Dr N. Thacker

Papers Submitted to International Journals

Luca Ferrarini, Berit M. Verbist, Hans Olofsen, Filiep Vanpoucke, Johan H.M. Frijns, Johan H.C. Reiber, Faiza Admiraal-Behloul, *"Autonomous Virtual Mobile Robot for 3-Dimensional Medical Image Exploration: Application to Micro-CT Cochlear Images"*, Artificial Intelligence In Medicine, 2007

Acknowledgement

Since November 2003, I have been working as a researcher within the Neuro-imaging section of the Laboratory for Clinical and Experimental Image Processing (LKEB), under the supervision of Prof. J.H.C. Reiber and Dr. F. Admiraal-Behloul, at the Leiden University Medical Center (LUMC), The Netherlands. This manuscript presents the results of my research. Along the way, many colleagues and friends have shared part of their lives with me, making these last four years a wonderful experience. My deepest gratitude to all of them.

To the colleagues of the Neuro-imaging section: working with you has always been stimulating and pleasurable. You have supported me professionally, and created the best working-atmosphere I could hope for.

To Alize, Roald, Dennis, Ronald, and Bobika: because of the daily question ”how was your evening?”, and the many laughs and jokes shared together (inscrutable to non-TZs). Thanks for knowing how to help and how to ask for it.

To Hans and Alize, for helping me with the Dutch translation.

To Emmanuelle, for being the most persistent person I have ever met. The friendship we have built up in these years is among the most precious things I have.

To all the friends in Leiden who has filled my life with happiness.

To Prof. Goldoni, Prof. Capone, and Prof. Zanetti. The passion you have always put in your job has shown me what studying and teaching really mean.

To Giusy, who would have deserved a chance to achieve the same goals and never got it, and to her parents for the many hand-written letters: always a pleasurable reading. (*Per Giusy,*

che avrebbe meritato l'opportunitá di raggiungere lo stesso obiettivo e non l'ha avuta, e per i suoi genitori, per le molte e piacevolissime lettere!).

To Sanne, for being close to me during these last months, supporting me, and helping me putting things back in the right perspective every time they started looking a bit too complicated.

To Federico C., Federico M., Nico and Paolo. Although I have been living far from you guys, every time I was back in Carpi I felt like I had never left. Living abroad would have not been possible without friends like you. (*Per Toto, Ponch, Nico e Paolo. Pur vivendo lontano da casa, ad ogni rientro mi avete fatto sentire come se non fossi mai partito. Questa esperienza non sarebbe stata possibile senza amici come voi. Grazie.*).

To my parents, Valeria Barletta and Claudio Ferrarini, to my brother Marco Ferrarini, and to the rest of my family, for being the solid ground I can always rely on. For wanting me there, and supporting me in my staying here. (*Ai miei genitori, Valeria Barletta e Claudio Ferrarini, a mio fratello Marco Ferrarini, e al resto della mia famiglia, per essere quelle solide fondamenta sulle quali posso sempre contare. Per volermi lí con voi, e comunque incoraggiarmi nella mia avventura qui in Olanda. Grazie di cuore!*).

Curriculum Vitae

Luca Ferrarini was born in Carpi, Italy, in 1978. He graduated at the High School I.T.I.S. Leonardo Da Vinci (Carpi, Italy) in 1997. In the same year, he started his academic education at the faculty of Computer Engineering, University of Modena and Reggio Emilia, from which he graduated *cum laude* in 2003, receiving his M.Sc. degree. His graduation project was carried on at *The Burnham Institute Labs, La Jolla - San Diego - CA* at the John Reed lab.: the project involved the processing of high-speed camera videos for cardiac function evaluation in genetically modified Drosophila Melanogaster.

During the University period, he worked as a tutor for students in secondary school, covering subjects as Mathematics, Statistics, Computer Science, Physics, and Chemistry. Immediately after the graduation (March, 2003), he worked as project developer for a system aimed at modeling and simulating complex networks, at *The Burnham Institute Labs, La Jolla - San Diego - CA* at the G. Paternostro lab: the project was developed in collaboration with Prof. Andrew McCulloch of the Cardiac Mechanics Research Center of UCSD (University of California, San Diego), and carried on on several clusters at the San Diego Supercomputer Center (SDSC). In the period from June 2003 until November 2003, he worked as external consultant for the same project.

From November 2003 until December 2007, he worked at the Laboratory for Clinical and Experimental Image Processing (LKEB) at Leiden University Medical Center. The results of his research are presented in this thesis.

His research interests include shape modeling and analysis, pattern recognition, analysis of MR images of the brain (particularly related to Alzheimer Disease), which constitute the basis for his Ph.D. thesis. Other research related interests include complex network theory, parallel programming for clusters, intelligent systems based on neural networks and genetic algorithms.

List of Figures

List of Tables