



UNIVERSIDAD NACIONAL DE COLOMBIA

Visual Attention Models and Sparse Representations for Morphometrical Image Analysis

Andrea del Pilar Rueda Olarte

Universidad Nacional de Colombia
Facultad de Ingeniería, Departamento de Ingeniería de Sistemas e Industrial
Bogotá D.C., Colombia
2013

Visual Attention Models and Sparse Representations for Morphometrical Image Analysis

Andrea del Pilar Rueda Olarte

In fulfillment of the requirements for the degree of:
Doctor en Ingeniería - Ingeniería de Sistemas y Computación

Advisor:
Eduardo Romero Castro, Ph.D.

Research Field:
Applied Computing - Image Processing
Research Group:
BioIngenium Research Group
Computer Imaging and Medical Applications Laboratory - CIM@LAB

Universidad Nacional de Colombia
Facultad de Ingeniería, Departamento de Ingeniería de Sistemas e Industrial
Bogotá D.C., Colombia
2013

A mis padres Rosalba y Alberto

A mi abuelita Rosalba

A mi hermana, mi prima, mis tíos y tías

A todos los que han apoyado este loco sueño
durante estos años

Acknowledgements

First, I want to express my immense gratitude to my advisor, professor Eduardo Romero, for his unconditional support, guidance, and above all, for his patience. During the time that I have been working with him, he has inspired and encouraged me to continue this adventure, to give my best at all times and to trust in my abilities and skills. His advice has helped me deeply, both in the academical and personal fields.

My special thanks go to professors Oscar Acosta, Norberto Malpica and Fabio González for their continued support and interest in my work, and for the invaluable assistance they constantly gave to me. Their explanations and insights have constantly helped me to improve my understanding of their areas of expertise related to my thesis. Also, I want to thank all members of the Biomedical Imaging group at the Australian e-Health Research Center in Brisbane, Australia, and to the people of the Laboratorio de Análisis de Imagen at the Universidad Rey Juan Carlos and the Fundación Reina Sofía in Madrid, Spain, for their warm welcome and support during my research internships.

Many people from the research group, formerly the BioIngenium Research Group and recently the CIM@LAB and MINDLAB Laboratories, have contributed at various levels to complete this work. My thanks go to Fabio Martínez, Juan Carlos Caicedo, Ángel Cruz, Jorge Camargo and Ricardo Gutiérrez for their support, suggestions, comments, and for their invaluable friendship. Thanks also to Gerardo Tibamoso, Nelson Velasco, Andrea Pulido and Katherine Baquero for allowing me to contribute in their own research problems and share some of my limited experience with them. Thanks to all other members of the research group for making every day a constructive and fun experience.

I also want to thank to the institutions that have sponsored my thesis during these years. This work has been supported by project “Visual Attention Models and Sparse Representations for Morphometrical Image Analysis” (number 12108) funded by Universidad Nacional de Colombia through “Apoyo de la DIB a tesis de investigación en posgrados”; and partially funded by projects “Anotación Automática y Recuperación por Contenido de Imágenes Radiológicas usando Semántica Latente” (number 110152128803) and “Sistema para la Recuperación de Imágenes Médicas utilizando In-

de la "Asociación Multimodal" (number 110152128767) by Convocatoria Colciencias 521 de 2010.

And, last but not least, I want to specially thank my wonderful family and my friends from the "Combo Alpinito" for their love and support through all these years, and for always being there for me, both in good and bad times.

Abstract

Medical diagnosis, treatment, follow-up and research activities are nowadays strongly supported on different types of diagnostic images, whose main goal is to provide an useful exchange of medical knowledge. This multi-modal information needs to be processed in order to extract information exploitable within the context of a particular medical task. In despite of the relevance of these complementary sources of medical knowledge, medical images are rarely further processed in actual clinical practice, so the specialists take decisions only based in the raw data. A new trend in the development of medical image processing and analysis tools follows the idea of biologically-inspired methods, which resemble the performance of the human vision system. Visual attention models and sparse representations are examples of this tendency. Based on this, the aim of this thesis was the development of a set of computational methods for automatic morphometrical analysis, combining the relevant region extraction power of visual attention models with the incorporation of a priori information capabilities of sparse representations. The combination of these biologically inspired tools with common machine learning techniques allowed the identification of visual patterns relevant for pathology discrimination, improving the accuracy and interpretability of morphometric measures and comparisons. After extensive validations with different image data sets, the computational methods proposed in this thesis seems to be promising tools for the definition of anatomical biomarkers, based on visual pattern analysis, and suitable for patient's diagnosis, prognosis and follow-up.

Keywords: Computational neuroanatomy, Sparse representations, Visual attention models, Machine learning techniques, Alzheimer's disease, Semantic-based representations, Visual pattern analysis.

Resumen

Las actividades de diagnóstico, tratamiento, seguimiento e investigación en medicina están actualmente soportadas en diferentes clases de imágenes diagnósticas, cuyo objetivo principal es el de proveer un intercambio efectivo de conocimiento médico. Esta información multimodal necesita ser procesada con el objetivo de extraer información aprovechable en el contexto de una tarea médica particular. A pesar de la relevancia de estas fuentes complementarias de información clínica, las imágenes médicas son raramente procesadas en la práctica clínica actual, de forma que los especialistas sólo toman decisiones basados en los datos crudos. Una nueva tendencia en el desarrollo de herramientas de análisis y procesamiento de imágenes médicas persigue la idea de métodos biológicamente inspirados, que se asemejan al sistema de visión humana. Son ejemplos de esta tendencia los modelos de atención visual y las representaciones escasas (*sparse representations*). Con base en esto, el objetivo de esta tesis fue el desarrollo de un conjunto de métodos computacionales para soportar automáticamente los análisis morfométricos, combinando el poder de extracción de regiones relevantes de los modelos de atención visual junto con la capacidad de incorporación de información a priori de las representaciones escasas. La combinación de estos métodos biológicamente inspirados con técnicas de aprendizaje de máquina facilitó la identificación de patrones visuales relevantes para discriminar patologías cerebrales, mejorando la precisión e interpretabilidad de las medidas y comparaciones morfométricas. Después de extensivas validaciones con diferentes conjuntos de imágenes, los métodos computacionales propuestos en esta tesis se perfilan como herramientas prometedoras para la definición de biomarcadores anatómicos, basados en el análisis visual de patrones, y convenientes para el diagnóstico, pronóstico y seguimiento del paciente.

Palabras clave: Neuroanatomía computacional, Representaciones escasas, Modelos de atención visual, Técnicas de aprendizaje de máquina, Enfermedad de Alzheimer, Representaciones basadas en semántica, Análisis de patrones visuales.

Contents

Acknowledgements	iv
Abstract	vi
1 Introduction	1
1.1 Research Problem	1
1.1.1 Research Question	2
1.2 Computational Brain Morphometry using MRI	3
1.3 Contributions and Academic Products	6
1.3.1 Segmentation of Medical Images	6
1.3.2 Super-Resolution of Medical Images	7
1.3.3 Semantic Representation of Medical Information	9
1.4 Thesis Outline	10
2 Improving partial volume classification with topological operators	12
2.1 Introduction	12
2.2 Methods	14
2.2.1 Pure tissue segmentation	15
2.2.2 Initial partial volume labelling	15
2.2.3 A topology preserving segmentation	16
2.2.4 Partial volume relabelling and fractional content	17
2.3 Experiments	21
2.3.1 Simulated MR data	22
2.3.2 Manually segmented real MR data	22
2.3.3 Cross sectional series of real MR scans	22
2.3.4 Error and similarity measures	23
2.4 Results and discussion	23
2.4.1 BrainWeb	23
2.4.2 Real MR Data	26
2.4.3 Computational performance	28
2.4.4 Deep sulci cutting and cortical thickness estimation on real data	29

2.5	Conclusion	33
3	Single-image Super-Resolution of Brain MR Images	34
3.1	Introduction	34
3.2	Sparse Representations in Image Processing	37
3.3	Super-Resolution of Brain MR Images	38
3.3.1	Dictionary Considerations	39
3.4	Proposed Methodology	39
3.4.1	Dictionary Construction	39
3.4.2	Local reconstruction by sparsity	42
3.4.3	Global regularization by back-projection	43
3.5	Experimental Results	44
3.5.1	Implementation Details	44
3.5.2	Brain MR Data Sets	46
3.5.3	Volume Processing Tools	47
3.5.4	Quantitative Performance Measures	47
3.5.5	Tests on Dictionary Construction	48
3.5.6	Tests on the Super-Resolution Approach	52
3.5.7	Comparison with Other Approaches	60
3.5.8	Impact on Post-processing Tasks	65
3.6	Discussion	68
4	Extracting Brain Patterns using a Visual Saliency Model	74
4.1	Introduction	75
4.2	Representation of Relevant Structural Regions by their Saliency	78
4.3	Adaptive Learning of Salient Regions	80
4.4	Materials and Methods	82
4.4.1	OASIS public data set	82
4.4.2	Proposed Approach	83
4.4.3	Performance Evaluation	90
4.4.4	Anatomical Interpretation	91
4.5	Experimental Results	92
4.5.1	Classification Results	92
4.5.2	Anatomical Patterns and Saliency	98
4.6	Discussion	106
5	Conclusions and Perspectives	109
5.1	Perspectives	110

A	Topology Preservation and Homotopic Transformations	112
B	The Super-Resolution Generative Model	114
C	Influence of Dimensionality Reduction in Super-Resolution	116
D	Cortical and Subcortical Regions in Harvard-Oxford Atlases	118
	References	120

List of Figures

1-1	Schematic description of the principal morphometric analyses used in computational neuroanatomy.	4
2-1	Overall process for topology-corrected PV estimation in MR images. . .	15
2-2	Illustration of the homotopic operations applied for correction of the grey and white matter segmentations.	17
2-3	Illustration of topological corrections in a synthetic phantom.	18
2-4	2D and 3D views of brain tissue segmentations before and after the topological correction.	19
2-5	Schematic view of the local tissue averages for a given mixed voxel. . .	20
2-6	Averaged intensity within the connected components of the pure gray matter.	20
2-7	3D reconstruction showing the regional differences in contrast between white and gray matters.	21
2-8	Partial volume segmentation of a simulated BrainWeb volume (3% noise, 20% bias field).	24
2-9	Example of PV estimation of a simulated BrainWeb volume (3% noise, 20% bias field).	25
2-10	RMS error per AAL region, in white and grey matters, for different noise levels.	26
2-11	PV estimation errors for white and grey matters on BrainWeb.	27
2-12	Tissue segmentation in IBSR images.	28
2-13	Jaccard similarity results for WM (a) and GM (b).	29
2-14	Example of cortical thickness estimation from MR.	30
2-15	Surface representation of cortical thickness, computed at different steps for two scans of a single subject (OASIS).	31
2-16	Illustration of the improvements in cortical thickness estimation after topological correction.	32
3-1	Illustration of low- and high-resolution dictionary construction.	40
3-2	Illustration of patch-based local super-resolution reconstruction.	43

3-3	Illustration of global regularization.	43
3-4	Comparison of local super-resolution of a low-resolution patch using several dictionaries.	50
3-5	Results of the reconstruction with $1mm$ (<i>top</i>) and $3mm$ (<i>bottom</i>) slice thicknesses.	55
3-6	Comparison of the local high-resolution reconstruction using different edge extraction operators.	56
3-7	Comparison of super-resolution reconstruction with different amounts of PCA coefficients.	57
3-8	Comparison of super-resolution of a low-resolution image using a scaling factor of 4.	58
3-9	Comparison of super-resolution using low- and high-resolution images acquired from the same subject in the same scan session.	59
3-10	Comparison of super-resolution of a previously downsampled low-resolution image with the corresponding high-resolution image.	60
3-11	Comparison of super-resolution results obtained after applying our proposal and the method in [133].	62
3-12	Reconstruction with 4200 dictionary patches from multiple MR brain images.	63
3-13	Comparison of super-resolution results obtained after applying our proposal and the method in [88].	64
3-14	Comparison of super-resolution results obtained on denoised images with our proposal and the method in [88].	65
3-15	Improvement of gray matter segmentation using the proposed method.	66
3-16	Comparison of improvement in gray matter segmentation.	66
3-17	Statistical map ($p < 0.01$) of comparison between control subjects and MCI patients.	68
3-18	Reconstruction of images from different centers (different scanners and acquisition parameters).	69
4-1	Graphical overview of the proposed method.	84
4-2	Construction of the saliency maps.	85
4-3	Overview of the graph-based approach for computing saliency.	86
4-4	Schematic description of the relevance weight map construction.	92
4-5	Illustration of Harvard-Oxford cortical (left) and subcortical (right) atlases.	92
4-6	Relevance maps for AD classification of the different OASIS subject groups.	100

4-7 Anatomical relevance maps for AD classification of the different OASIS subject groups.	101
--	-----

List of Tables

2-1	Fractional content RMS error on BrainWeb.	25
2-2	Mean (\pm standard deviation) of Jaccard similarity index for each method.	28
2-3	Pearson correlation coefficient and differences between scans for the OASIS dataset, grouped by brain lobules.	31
3-1	Accuracy super-resolution values under influence of noise.	53
3-2	Accuracy super-resolution values under influence of INU.	54
4-1	Comparison of classification performance of the proposed approach with respect to the FBM approach [121].	93
4-2	Classification performance measures for KG 1 for the different OASIS subject groups.	94
4-3	Kernel weights learned for KG 1 on the different OASIS subject groups.	94
4-4	Classification performance measures for KG 2 for the different OASIS subject groups.	95
4-5	Kernel weights learned for KG 2 on the different OASIS subject groups.	96
4-6	Classification performance measures for KG 3 for the different OASIS subject groups.	97
4-7	Kernel weights learned for KG 3 on the different OASIS subject groups.	98
4-8	List of relevant cortical and subcortical regions highlighted by the classification process using a single kernel.	102
4-9	List of relevant cortical and subcortical regions highlighted by the classification process using the KG 1.	103
4-10	List of relevant cortical and subcortical regions highlighted by the classification process using the KG 2.	104
4-11	List of relevant cortical and subcortical regions highlighted by the classification process using the KG 3.	105
D-1	Names of relevant cortical regions listed in Tables 4-8 to 4-11	118
D-2	Names of relevant subcortical regions listed in Tables 4-8 to 4-11	119

1 Introduction

Current medical and clinical practices have been improved by new technical and technological developments in medical imaging devices. The increasing production and compilation of medical images in digital form, coming from diverse modalities such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT) or Ultrasound, provide a useful exchange of medical knowledge. In fact, thanks to these advances, physicians and specialists have now more useful structural and functional information to facilitate and increase accuracy of medical diagnosis, treatment and follow-up, as well as to support medical and biological research and training.

Medical image processing refers to those tools and methods that can assist a medical expert (radiologist, pathologist) to identify, interpret and analyze all the useful information available in a medical image. Most of these methods are based on known facts about how image interpreters work on different kinds of medical images. Image interpretation can be seen then as a process to generate a content-based representation of image data [15]. Thus, image processing, interpretation and/or analysis deal, in general, with the process of extracting information from images. Information extraction tasks range from simple processes such as finding regions with different color or texture on an image, to complex procedures that could determine the absence or presence of a specific pathology in a diagnostic image. As the complexity level of the medical task increases, a high-level interpretation of the information must be attained, by relating the visual information with semantic concepts associated to a particular knowledge domain. The process of inferring and associating the visual information present in an image with semantic concepts is currently known as image understanding [108]. Some medical applications that can be benefited with the integration of automatic (or semi-automatic) image understanding capabilities, one of the most important among them is the automatic quantification or morphometry on medical images.

1.1 Research Problem

As already mentioned, medical diagnosis, treatment, follow-up and research activities are nowadays strongly supported on the different types of images produced in clinics and hospitals. This multi-modal information, which combines visual structural and

functional information with textual descriptions and other types of medical information, needs to be processed in order to extract information useful within the context of a particular medical task. In spite of the relevance of these complementary sources of medical knowledge, medical images are rarely further processed in actual clinical practice, so the specialists take decisions only based in the raw data.

Recognition and interpretation with semantic meaning of medical images offer a wide range of possibilities such as serve as a second opinion in diagnosis, or to complement the training of new specialists. These processes are nowadays considered as challenging tasks, because they should be performed by hand, implying large amounts of working time of human experts. A lot of image processing tools, such as different segmentation, recognition and classification methods, have been introduced into the medical domain in order to semi- or fully-automatize this medical information-extraction process. However, most of them are designed to work under specific assumptions, in controlled environments and for a particular pathology and/or organ. For this reasons, only a few number of image processing applications have reached the status of stable tools for computational medical analysis.

Inference and interpretation of medical information for a single patient are valuable tools for diagnosis and treatment. In addition, the comparison of this information with the one from other patients that are in similar conditions or present the same pathology, helps to understand the behavior of pathologies in particular environments and human groups, and also provides important information for development of public health policies. In most of the cases, the comparison of groups of patients is not an immediate process, because the anatomical variability between subjects induces errors that bias the statistical conclusions that could be extracted from the population analysis. Effective modeling and/or quantification of this inter- and intra-subject anatomical variability helps to understand which image structures are common or uncommon within a given population, which image or anatomic features could characterize the subject variability and in which ways a subject could be defined as similar to or different from the population.

1.1.1 Research Question

The referred situations unveil the lack of computational tools that could exploit all the specific knowledge contained in medical images and infer as much as possible accurate and pertinent information that could be helpful in actual clinical practice. This leads us to the main research question in this work: *how can we infer or interpret medical useful knowledge from diagnostic images?*

1.2 Computational Brain Morphometry using MRI

The target clinical application selected for the work developed in this thesis corresponds to the area of *brain morphometry* (also known as computational neuroanatomy or neuromorphometry [92]), which involves the analysis of features derived from the form (size and shape) of subject brains. Given the particular and complex anatomy of the human brain, and the importance of this organ as the central control system of the whole human body, this problem results of high interest in many different domains. The techniques proposed in this thesis aim to provide both accurate and clinically meaningful morphometric tools for brain analysis, based only on structural information coming from brain Magnetic Resonance (MR) images. The effectiveness of MRI as a valuable diagnostic technique in neurological diseases has been widely proved, usually on the T1-weighted imaging data [74]. Such popularity comes from the obtained good contrast between soft tissues, giving the possibility of identifying the distribution of changes in neuroanatomical structures, such as estimates of damaged tissue or atrophy rates. The study of complex neurological diseases, namely the Alzheimer's disease (AD), the Parkinson disease, the schizophrenia or the multiple sclerosis, have benefited by using MR images in the actual clinical analysis and diagnosis, allowing to improve the disease physiopathological knowledge.

As pointed out by Mietchen and Gaser [92], a brain morphometric study is composed of two important elements: a common spatial representation of the brain and the corresponding morphometrical measures and statistical analyses. A common reference frame is needed given the intrinsic anatomical variability and the image acquisition conditions, so that all images are warped or registered together to a specific template or an atlas. The whole set of working images can be extracted from different brains (a cross-sectional study) or from one specific brain along the time (a longitudinal study). The particular morphometrical measures come from different sources of information, such as locations (landmarks), voxel intensities, template deformations or surface representations; each of them springing up a specific morphometric technique: landmark-based morphometry [34], voxel-based morphometry [5], deformation-based or tensor-based morphometry [6] and surface-based morphometry [102]. The most commonly used approaches are detailed hereafter, and their relationships are illustrated in Figure 1-1. Voxel-Based Morphometry (VBM), proposed by Ashburner and Friston in 2000 [5], is by far the morphometric approach most commonly used by the neuroscience research community (the proposing manuscript has, up to date, about 3600 cites in Google Scholar¹). It is based on using the brain tissue segmentations for comparing their volume across a subject population, in a voxel-by-voxel basis. In this approach, brain

¹<http://scholar.google.com/scholar?q=Voxel-based+morphometry-the+methods>

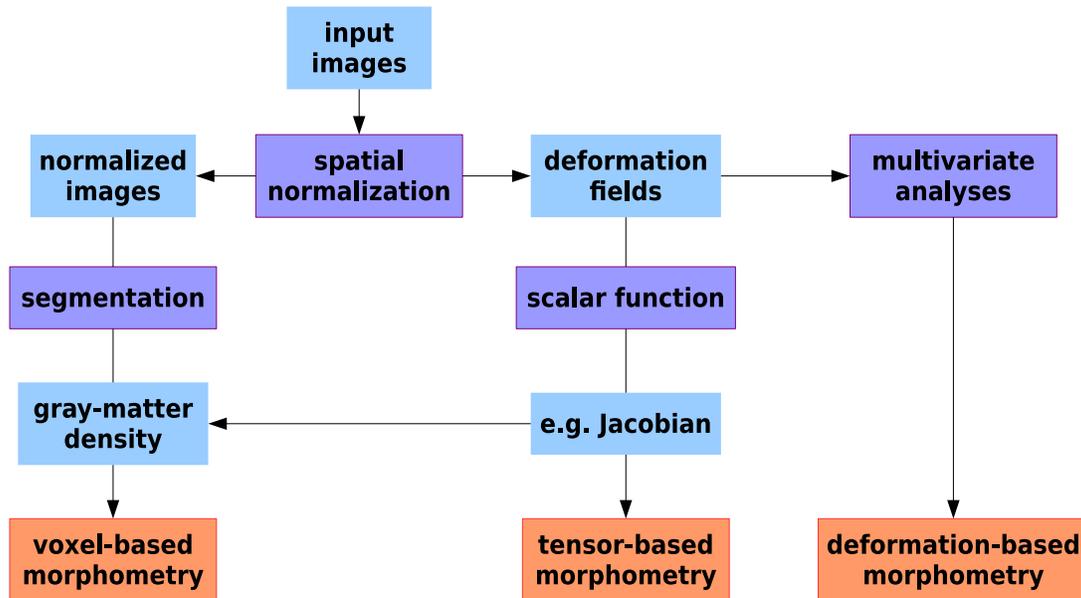


Figure 1-1: Schematic description of the principal morphometric analyses used in computational neuroanatomy, indicating the different information sources. (Illustration adapted from <http://www.fil.ion.ucl.ac.uk/spm/doc/intro/Figure3.htm>)

tissue membership of each voxel is identified using a mixture of gaussians, generating up to three tissue probability maps. Then, tissue maps are warped together to a common anatomical space, using some intersubject registration model², and then they are spatially blurred (with an isotropic gaussian kernel) to obtain a weighted sum of the tissue around each voxel. Finally, voxel-by-voxel statistical analyses of the whole data are performed by fitting a general linear model at each voxel. This morphometric strategy has been incorporated into currently used software packages for brain analysis, such as SPM (Statistical Parametric Mapping) [51] and FSL (FMRIB Software Tools) [69].

VBM approach is considered to be oriented to mesoscopic differences, that is to say, to local differences. In contrast, identification of macroscopic anatomical differences among brains can be modeled using the deformation fields that warp individual brains to a common reference space. This is known as Deformation-Based Morphometry (DBM), and was proposed in 1998 by Ashburner et al. [6]. The deformation fields thus encode the shapes of individual brains, providing additional information to the analysis about lengths, areas and angles, among others. Because the deformation

²DARTEL [4] is the most accurate and the most used so far

fields are multivariate, the comparisons among them require the use of multivariate statistical techniques, allowing to establish the nature of the differences and to make inferences from them. This approach has also been included into the very well known SPM software package [51].

A main drawback of these morphometrical analyses is the requirement of an inter-subject registration as accurate as possible, in order to guarantee that the statistical analysis compares homologous structures across all subject brains and to avoid unreliable interpretation of the results. However, this kind of one-to-one correspondence between subjects need not be achieved for every case, mainly because of the inherent intersubject anatomical variability and the effects of a brain pathology. In fact, the same anatomical structure may not be present in all subjects, or may exhibit multiple morphologies across the population. On the other hand, some pathologies may affect not only a single anatomical structure or interconnected regions, but specific structures localized far away from each other. This kind of patterns are difficult to find and analyze with the standard morphometrical techniques. To cope with this issue, Toews et al. have proposed in 2010 a new approach, called Feature-Based Morphometry (FBM) [121], focused in modeling such natural image patterns that might not occur in all subjects. Four stages are involved in this approach: first, an affine alignment of the subject brains is performed, assuming an approximate arrangement of similar image structures. Then, a set of SIFT (Scale Invariant Feature Transform) features is extracted per volume. These features are characterized by their robustness to geometrical and intensity variations and by their image appearance distinctiveness. With these features, each subject volume is modeled as a collection of brain features, and a probabilistic framework is formulated to estimate the relationship between a feature and the brain. These estimations allow to cluster features across subjects, taking into account the geometry, appearance and group class similarities. Finally, group analyses are performed using cooccurrence statistics between features and class groups.

In spite of the increasing research interest in this kind of morphometric approaches, these are still far from emulating the process that a radiologist follows when examining a particular case and, moreover, from the required medical interpretability for any method to be used in actual clinical scenarios. In the computational attempt of emulating the radiologist visual perception, the use of biologically inspired tools for image analysis is a new perspective that naturally leads to useful interpretation. Therefore, the techniques proposed in this thesis incorporate some of these approaches, such as the visual attention models and the sparse representations, exploring their contributions in the development of efficient support systems in medicine.

1.3 Contributions and Academic Products

This work presents several contributions to the morphometric analyses of medical images, with an important focus in the study of the Alzheimer's disease, using measurements obtained only from structural brain MR images. These contributions are mainly focused to the extraction of relevant image regions with medical and diagnostic meaning and also to the effective introduction of prior medical knowledge. These goals have been achieved by the use of two important biologically-inspired tools for image analysis: sparse representations and visual attention models. The following are the main applications covered by the methods proposed in this thesis, together with the corresponding references to published works.

1.3.1 Segmentation of Medical Images

As the accuracy of any morphometrical analysis is subject to the quality of the input information, early efforts were devoted to improve tissue classification, mainly in brain MR images. As a complementary work, some contributions were also made for a liver segmentation approach.

Brain Tissue Segmentation

Refinement of brain tissue segmentations, improving at the same time the Partial Volume (PV) classification and the fractional content computation, was achieved by imposing topological constraints to an initial binary segmentation. This procedure allowed to identify a set of mixed voxels, for which the fractional content is estimated using a local averaging of pure tissue voxels. This work was published in:

- **Andrea Rueda**, Oscar Acosta, Michel Couprie, Pierrick Bourgeat, Jurgen Fripp, Nicholas Dowson, Eduardo Romero and Olivier Salvado. *Topology-corrected segmentation and local intensity estimates for improved partial volume classification of brain cortex in MRI*. Journal of Neuroscience Methods, 188 (2), pag. 305-315, 2010.
- **Andrea Rueda**, Oscar Acosta, Pierrick Bourgeat, Jurgen Fripp, Erik Bonner, Nicholas Dowson, Michel Couprie, Eduardo Romero and Olivier Salvado. *Partial volume estimation of brain cortex from MRI using topology-corrected segmentation*. Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009. June 28 - July 1, 2009. Boston, Massachusetts, United States of America.

Liver Segmentation

Additional contributions have made to a semi-automatic approach for segmentation, reconstruction and estimation of the liver volume in Computed Tomography (CT) images. The proposal includes an intensity-based deformation method, which acts on an averaged liver shape, placed by hand over the region of interest. This work was published (in Spanish) in:

- Gerardo Tibamoso, **Andrea Rueda** and Eduardo Romero. *Segmentación Semi-automática del Volumen del Hígado en Imágenes de Tomografía Computarizada*. Acta Biológica Colombiana, 15 (3), 2010.
- Gerardo Tibamoso, **Andrea Rueda** and Eduardo Romero. *Segmentación Semi-automática del Volumen del Hígado en Imágenes de TAC*. Proceedings of the Seminario de Ingeniería Biomédica y Seminario Internacional de Procesamiento y Análisis de Imágenes Médicas, SIB-SIPAIM 2009. November 26-27, 2009. Bogotá, Colombia.

1.3.2 Super-Resolution of Medical Images

Another limitation for accurate morphometric measurements comes from the intrinsic resolution of the acquired images. Hence, improving image resolution (directly in the acquisition process or as a post-processing step) is one of the main challenges in medical image processing. A powerful technique was proposed for solving the super-resolution of brain MR images problem, while also some contributions in super-resolution of 4D cardiac MR images were made.

Brain MR Images

An automatic technique for resolution enhancement of complete 3D brain MR volumes was proposed, comprising a multi-scale feature analysis, semantic-based dictionaries and a dimensionality reduction scheme. Coupled low- and high-resolution were constructed from training images, and then used to locally describe a new low-resolution image as a sparse combination of the dictionary patches. Finally, a global correction is performed to ensure consistency and remove discontinuity effects. This work was published in:

- **Andrea Rueda**, Norberto Malpica and Eduardo Romero. *Single-image Super-Resolution of Brain MR Images using Overcomplete Dictionaries*. Medical Image

Analysis³, 17 (1), pag. 113-132. 2013.

- **Andrea Rueda**, Gloria Díaz and Eduardo Romero. *Learning compact dictionaries for brain MR image super-resolution*. Proceedings of the Seminario Internacional de Procesamiento y Análisis de Imágenes Médicas, SIPAIM 2010, December 1-4, 2010. Bogotá, Colombia.
- **Andrea Rueda** and Eduardo Romero. *Super-Resolution of Brain MR Images based on Sparse Representations*. Proceedings of the Interdisciplinary Workshop on Sparsity and Modern Mathematical Methods for High Dimensional Data. April 6-10, 2010. Brussels, Belgium.

Cardiac MR Images

A different strategy for super-resolution of cardiac MR image series was proposed, given the highly anisotropic voxels and the number of non-orthogonal series available per subject. A Bayesian approach is then used to model the contribution of each low-resolution voxel in the final intensities of the high-resolution image. This work was published in:

- Nelson Velasco, **Andrea Rueda**, Cristina Santa Marta and Eduardo Romero. *Super-resolution in cardiac MRI using a Bayesian approach*. SPIE Medical Imaging 2013. February 9-14, 2013. Lake Buena Vista, Florida, USA.
- Nelson Velasco, **Andrea Rueda**, Cristina Santa Marta and Eduardo Romero. *Estimación por Máxima Verosimilitud para Super-Resolución en Imágenes de Resonancia Magnética Cardíaca*. Proceedings of the International Seminar on Medical Information Processing and Analysis, SIPAIM 2012. November 12-15, 2012. San Cristóbal, Venezuela.
- Nelson Velasco, **Andrea Rueda** and Eduardo Romero. *Combinación de proyecciones no ortogonales de imágenes de RM cardíaca*. Proceedings of the Seminario Internacional de Procesamiento y Análisis de Información Médica, SIPAIM 2011. December 5-7, 2011. Bucaramanga, Colombia.

³This journal is one of the Top-10 journals in Computer Science, according to SCImago Journal & Country Rank (http://www.scimagojr.com/journalrank.php?area=1700&category=1701&country=all&year=2011&order=sjr&min=0&min_type=cd)

1.3.3 Semantic Representation of Medical Information

In the search of image-based computational tools useful for support training and diagnosing processes, research efforts have been focused in the identification of distinctive and reproducible patterns as well as in their relations with a particular diagnosis, thereby accomplishing medical interpretability. Relationships between visual features and pathological conditions can be established in different ways, as proposed in the approaches described hereafter.

Classification of Alzheimer's Disease

Automatic identification and quantification of systematic differences with clinical meaning among normal subjects and Alzheimer's disease (AD) patients was achieved by a fusion strategy that mixes information coming from a multi-scale analysis of visual features together with machine learning and fusion techniques. The proposed approach allows an objective graduation and understanding of the different AD stages. This work was published in:

- **Andrea Rueda**, Fabio A. González and Eduardo Romero. *Learning visual salient patterns for structural MR morphometry of the Alzheimer's disease*. Submitted to Human Brain Mapping. 2013.
- **Andrea Rueda**, Fabio A. González and Eduardo Romero. *Saliency-based characterization of group differences for Magnetic Resonance disease classification*. Revista DYNA, in press, 2013.
- Andrea Pulido, **Andrea Rueda** and Eduardo Romero. *Classification of Alzheimer's disease using regional saliency maps from brain MR volumes*. SPIE Medical Imaging 2013. February 9-14, 2013. Lake Buena Vista, Florida, USA.
- Andrea Pulido, **Andrea Rueda** and Eduardo Romero. *Classification of Alzheimer's Disease using Regional Saliency Maps from Brain MR Images*. Proceedings of the International Seminar on Medical Information Processing and Analysis, SIPAIM 2012. November 12-15, 2012. San Cristóbal, Venezuela.
- **Andrea Rueda**, John Arévalo, Angel Cruz, Eduardo Romero and Fabio A. González. *Bag of Features for Automatic Classification of Alzheimer's Disease in Magnetic Resonance Images*. Proceedings of the Iberoamerican Congress in Pattern Recognition, CIARP 2012. September 3-6, 2012. Buenos Aires, Argentina.

- **Andrea Rueda**, Katherine Baquero and Eduardo Romero. *Saliency-based Characterization of Group Differences for MR Disease Classification*. Proceedings of the Seminario Internacional de Procesamiento y Análisis de Información Médica, SIPAIM 2011. December 5-7, 2011. Bucaramanga, Colombia.

Breast Masses Classification

An application of sparse representations in a classification task was explored in the context of severity (benign or malign) identification for breast masses. A set of Regions of Interest (RoIs) were characterized by their projection onto learned malign and benign dictionaries, and the region class was identified using a decision rule algorithm. This work was published in:

- Fabián Narváez, **Andrea Rueda** and Eduardo Romero. *Breast masses classification using a sparse representation*. Proceedings of the Workshop in Medical Image Analysis and Description for Diagnosis Systems, MIAD 2011. January 28-29, 2011. Rome, Italy.

Classification of Basal Cell Carcinoma

Diagnosis of a histopathology glass slide is a complex process that involves accurate recognition of several structures, their function in the tissue and their relation with other structures. A graph-based semantic representation was then proposed, allowing to describe histopathological concepts suitable for classification. Using learned dictionaries, the proposal models their spatial relations as the co-occurrence of dictionary atoms in the image. This work was published in:

- Ricardo Gutiérrez, **Andrea Rueda** and Eduardo Romero. *Learning semantic histopathological representation for basal cell carcinoma classification*. SPIE Medical Imaging 2013. February 9-14, 2013. Lake Buena Vista, Florida, USA.

1.4 Thesis Outline

The remaining chapters of the thesis are organized as follows:

- **Chapter 2: *Topology-corrected segmentation and local intensity estimates for improved partial volume classification of brain cortex in MRI***. This chapter introduces a new method to label voxels and compute brain tissue fractional content, integrating a mechanism for detecting sulci with topology preserving operators. The proposal allows to improve the computation of the

fractional content of mixed voxels using local estimations of pure tissue intensity means, increasing the precision of cortical thickness estimation in brain regions where this measure is particularly difficult, such as deep sulci. Accuracy and precision of the proposed technique is assessed using simulated and real MR data and comparison with other existing approaches demonstrates its benefits.

- **Chapter 3: *Single-image Super-Resolution of Brain MR Images using Overcomplete Dictionaries.*** This chapter presents a sparse-based super-resolution method, adapted for easily including prior knowledge, which couples up high- and low-frequency information so that a high-resolution version of a low-resolution brain MR image is generated. The proposed approach includes a whole-image multi-scale edge analysis and a dimensionality reduction scheme, which results in a remarkable improvement of the computational speed and accuracy. The method is validated by comparing interpolated and reconstructed versions of various data sets of brain volumes with the original images, and also compared with a recent state-of-the-art algorithm, suggesting a substantial impact in voxel-based morphometry studies.
- **Chapter 4: *Extracting Brain Patterns using Visual Saliency for Imaging-Based Classification of Neurodegenerative Diseases.*** This chapter presents a new fully automatic classification method that finds discriminative brain patterns associated to the presence of Alzheimer’s disease, mining systematic differences and therefore grading objectively any neurodegenerative disorder. This is accomplished by a fusion strategy that mixes together bottom-up and top-down information flows. Bottom-up information comes from a multiscale analysis of different image features, while the top-down stage includes learning and fusion strategies formulated as a max-margin multiple-kernel optimization problem. The proposal shows to outperform a state-of-the-art method in classification performance over different subject groups of a public brain MR data set. In terms of the anatomical analysis, relevant regions found by the proposed approach highly correlates to what has been reported in clinical studies of Alzheimer’s disease.
- **Chapter 5: *Conclusions and Perspectives.*** This final chapter presents the main conclusions of the proposed work, highlighting the main contributions achieved and its impact in the research area. In addition, it depicts some of the future research directions and perspectives promoted by this thesis.

2 Topology-corrected segmentation and local intensity estimates for improved partial volume classification of brain cortex in MRI

*In morphometric analyses based on magnetic resonance imaging (MRI), is of paramount importance the accuracy and precision with which brain structures can be quantified. One of the most common artifacts that hampers this measurements is the partial volume (PV) effect, due to the limited spatial resolution of MRI compared to the size of the anatomical structures. Accurate classification of mixed voxels and correct estimation of the proportion of each pure tissue (fractional content) may help to increase the precision of morphometrical tasks, such as the cortical thickness estimation, in regions where these measures are particularly difficult, such as deep sulci. This chapter presents a new PV classification-estimation method which integrates a mechanism for correcting brain tissue delineation using topological operators. The contribution of this work is twofold: on the one hand, we propose a new method to label voxels and compute tissue fractional content, integrating a mechanism for detecting sulci with topology preserving operators. On the other hand, we improve the computation of the fractional content of mixed voxels using local estimation of pure tissue intensity means. The complete content of this chapter has been published as a research article in the **Journal of Neuroscience Methods** (see [111]).*

2.1 Introduction

Accurate segmentation of Magnetic Resonance (MR) images into different brain tissues, namely gray matter (GM), white matter (WM), and cerebro-spinal fluid (CSF), can allow *in-vivo* quantification of structural modifications appearing during neurodegenerative diseases. However, MR-related artifacts, such as intensity inhomogeneity, noise and partial volume (PV) effects, can hamper the precision of this task. Inhomogeneities can be characterized by a low frequency multiplicative bias field and are mostly due

to the sensitivity of the receiver coils and, in some cases, to non-homogeneous tissue MR properties. The noise is Rician distributed and it has been shown to strongly affect the tissue classification [125]. Finally, PV effects appear when the size of anatomical features being imaged is comparable to the voxel size, causing blurring at the interfaces between tissues. In some cases, e.g. with opposed banks of GM in deep sulci, misclassification problems appear, affecting further processings such as cortical thickness estimation.

Topological operators and constraints have been widely used to correct and achieve accurate cortical tissue segmentations [9, 60, 75, 116]. It has been assumed that the cerebral cortex is a folded sheet of GM built upon the WM, which would have the topology of a hollow sphere if the midline hemispheric connections were artificially removed. Due to MR artifacts, the segmentation process cannot guarantee this assumption, generating deviations from the true anatomy of the structures of interest. Proposed approaches that address this issue can be classified in two categories: methods that include topological constraints directly into the segmentation process, based on active contours [116], topology adaptive snakes [91], digital topology models [9, 10] or segmentation by registration to an atlas [75]; and retrospective techniques that correct the topology after the segmentation process [60]. Those approaches are focused on ameliorating the topology of the segmented tissues, working directly on a voxel or on a mesh (surface) space. Voxel-based methods operate directly on the volumetric tissue segmentations, by removing or adding voxels according to topological constraints. However, removal or addition of a whole voxel in thin structures such as the GM may considerably modify the measure of thickness (ranging between ± 1 voxel) if any mechanism such as partial volume is not used to compensate for the structural modifications. In contrast, mesh-based techniques require an initial 3D reconstruction (triangular mesh) of the volumetric segmentations. The approaches for segmentation and cortical thickness estimation operating directly with the surfaces, such as CLASP [71], BrainVISA [86] or Freesurfer [32, 45, 46], incorporate mechanisms to prevent self-intersection of surfaces or topology correction, imposing also some smoothness constraints. Mesh-based approaches are however computationally more expensive, because of the needed additional reconstruction step. Overall, after or during the mesh generation, most of the methods tackle the elimination of tunnels and handles [46, 47, 68, 139].

On the other hand, PV estimation has received considerable attention in the last few years and different approaches have been proposed for classification and computation of fractional content [24, 77, 93, 115, 117, 122, 125]. Most techniques model voxel intensity as a linear combination of the intensity distributions of the possible tissue types within each voxel [25, 93]. Computing the fractional content of voxels therefore requires both pure and mixed voxels to have been previously classified. Shattuck et

al. [117] implemented a maximum *a posteriori* (MAP) classifier, which combined a tissue measurement model with a prior model of the local spatial interactions to obtain six tissue types: three pure and three mixed. The fractional content for the mixed voxels was calculated based on the global intensity mean of pure tissue types. Tohka et al. [122] proposed an algorithm which used statistical estimators, based on the MAP estimation [117]. Recently, Chiverton et al. [24] presented a local adaptive Gradient-controlled spatial regularizer (GSR) using a Markov Random Field to model the class membership and a Markov chain Monte Carlo (MCMC) simulation to adapt the model to the observed data. The labelling error may remain high because the intensity inhomogeneities (not explicitly modelled) and the noise may lead to misdetection of mixed voxels mainly in tight sulci, representing a portion of GM/CSF/GM within the same voxel.

The approaches previously presented have been focused on solving either the PV estimation or the topology correction. Our contribution consists in demonstrating that better results and performance are obtained if both strategies are combined together with a spatial intensity variation modeling. In this paper, we propose a new method aimed at improving both PV classification and fractional content computation, working at a voxel level in order to be accurate and computationally efficient. The improved classification is achieved by imposing topological constraints to the binary segmentation and thus detecting hidden mixed voxels in zones of tight sulci. The accurate fractional content estimation is attained by computing the fractional content as a linear relation between robust local intensity averages of pure tissue voxels. The spatially dependent averaging helps to overcome the problems of intensity inhomogeneity for a given tissue across the image.

In the next section we describe our methods, followed by experiments using simulated and real data. We also compare the results with other previously proposed methods. We demonstrated the utility of our approach by integrating the whole process to our voxel-based cortical thickness estimation pipeline.

2.2 Methods

The proposed strategy follows the steps depicted in Figure 2-1: Firstly, an initial classification of voxels into pure tissues WM, GM and CSF and mixed tissues WM/GM and GM/CSF is performed. Secondly, topology-constraints are introduced in the classification assuming that the GM is a continuous layer covering the WM. A topology preserving dilation of the WM over GM adds robustness to the delineation of mixed voxels GM/CSF in deep sulci. Finally, the estimation of fractional content for mixed voxels is adaptively performed based on a local averaging of the pure tissue voxels.

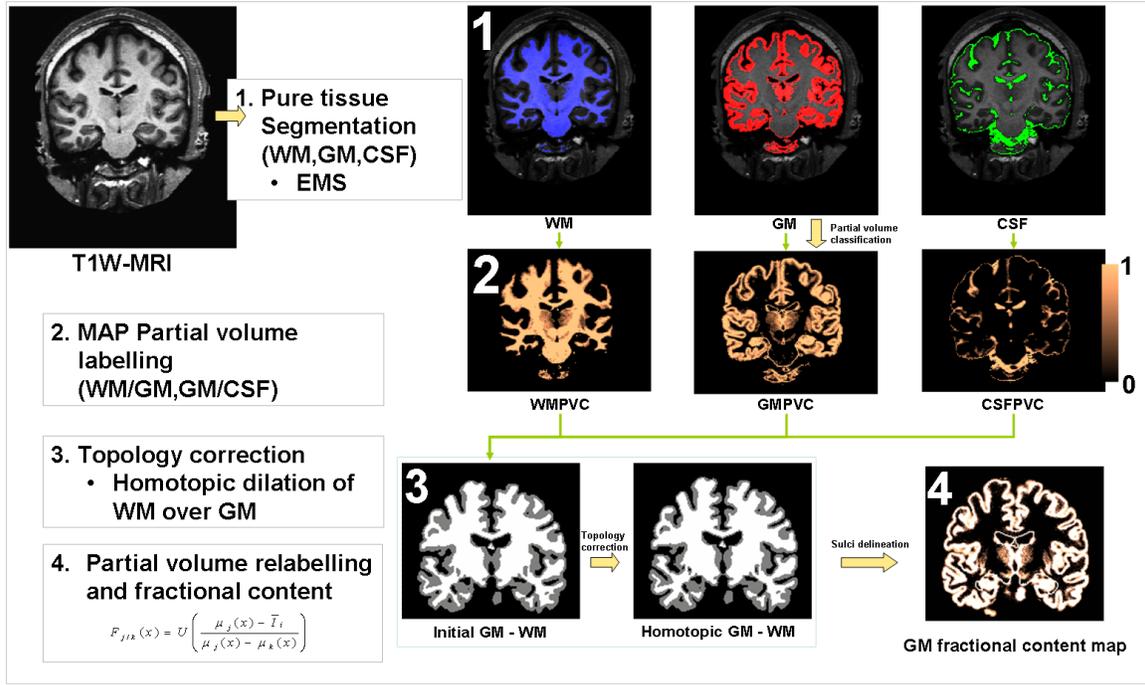


Figure 2-1: Overall process for topology-corrected PV estimation in MR images.

2.2.1 Pure tissue segmentation

A first segmentation of pure brain tissues into GM, WM and CSF is performed based on an implementation of the expectation-maximisation (EM) segmentation method as in [124]. Here, the Colin atlas and associated priors are first affinely registered to the data using a robust block matching approach [98], followed by a diffeomorphic Demons non-rigid registration [127]. Probabilistic tissue maps associated with the atlas were used to initialize the EM segmentation and enforce spatial consistency throughout the segmentation. The probability density functions of the tissues are modelled with 6 Gaussians (WM, GM, CSF and 3 for non brain tissues, skull and background). Finally, hard segmentations are obtained after the EM segmentation by labelling each voxel with the most probable tissue.

2.2.2 Initial partial volume labelling

Using the hard segmentations, a first labelling of partial volume voxels are identified within the hard segmentations and along the interfaces of pure tissues. Three pure tissue classes and two mixture classes are considered $\Gamma = \{GM, CSF, WM, CSF/GM, GM/WM\}$. A maximum *a posteriori* classification (MAP) is made and labels the

voxels as belonging to the set Γ . This procedure, relying on both intensity and spatial information, extends the method proposed by [117], but we assume that each voxel contains at most two tissues [115], and PV classification is restricted to the region formed by a dilated GM region (radius 2) because only the cortical thickness is sought. To take into account dependency on the neighbouring tissue types, a Markov prior that models local spatial interactions was implemented using a Potts model in order to perform the labelling. As in [71,117,122], we use the Iterated Conditional Modes (ICM) algorithm as explained in [13] to search for the optimal labelled image. According to this, every voxel is updated once per iteration until no label changes occur between iterations. This model favors classification of contiguous regions of GM, WM and CSF and encourages configurations of voxels that make physical sense such as GM/CSF or GM/WM voxels adjacent to GM.

2.2.3 A topology preserving segmentation

After the MAP labelling, some of the sulci may be misdetected, as the intensity of buried PV GM/CSF voxels is close to that of the GM. In order to refine the segmentation and identify such buried GM/CSF voxels, we used a homotopic dilation of the consolidated $\mathbf{WM} = \{\text{WM}, \text{WM/GM}\}$ constrained by the GM, leading to a better delineation of deep sulci. To preserve this folds during dilation, the set \mathbf{WM} is corrected first to assure that shares the topology of a filled sphere.

The homotopic transformations that we used are topology-preserving procedures that consist of sequentially deleting or adding single points (voxels) as described in [12]. The algorithms used are detailed in Appendix A. Our topology preserving segmentation of the WM consists in performing a homotopic dilation of a seed set of voxels, called \mathbf{S} , constrained to only add voxels from the set \mathbf{WM} , knowing that \mathbf{S} is topologically equivalent to a filled sphere. The result of this operation is denoted by \mathbf{SWM} . For example, \mathbf{S} could be made of single voxels chosen in the white matter, but we describe below a way to obtain a seed that is closer to the expected result, and thus leads to a more robust segmentation.

To obtain the seed \mathbf{S} , we first compute a surface skeleton \mathbf{SK} of \mathbf{WM} , by dilating using Algorithm 3 as described in Appendix A. Then, we perform an homotopic erosion, constrained by \mathbf{SK} , of a full cuboid that includes \mathbf{SK} . Finally, we perform an homotopic dilation of the same seed set \mathbf{S} , constrained by the set $\mathbf{SWM} \cup \mathbf{GM}$ to only add GM and WM voxels, and we subtract \mathbf{SWM} from the result to obtain the corrected GM.

This method is performed on 3D sets, but for clarity we illustrate it on a 2D reduced example in Figure 2-2. Notice that small black components in Figure 2.2(b) can

correspond to tunnels in the 3D image, thus simple connected component filtering would not give the correct region. Figures 2-3 and 2-4 show further examples in 3D.

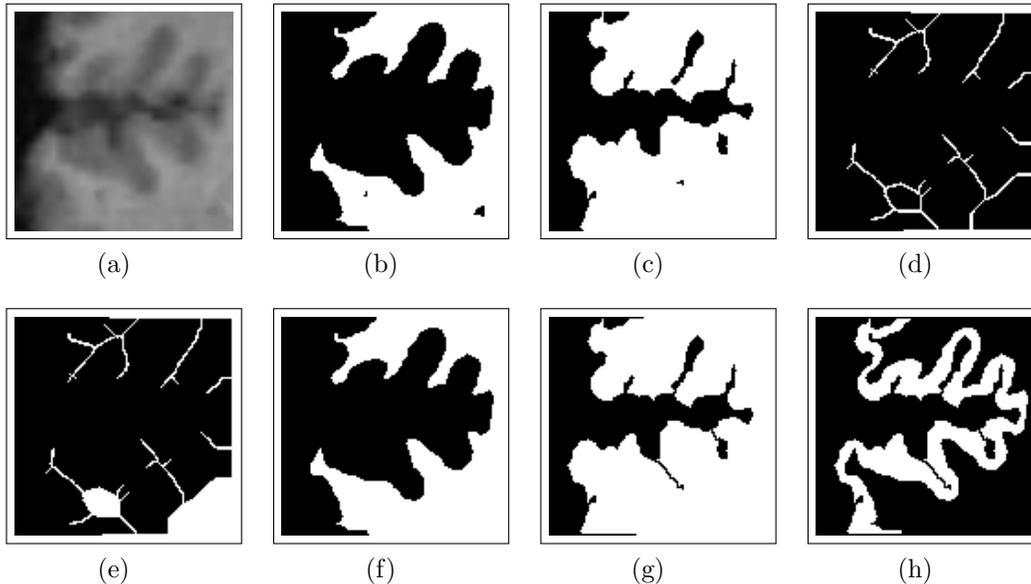


Figure 2-2: (a): original grayscale image. (b): segmented white matter (set \mathbf{WM}). (c): segmented white and gray matter (set $\mathbf{WM} \cup \mathbf{GM}$). (d): surface skeleton of \mathbf{WM} (set \mathbf{SK}). (e): seed set (set \mathbf{S}). (f): corrected white matter (set \mathbf{SWM}). (g): corrected white and gray matter formed by further homotopic dilation. (h): corrected gray matter (final result) formed by subtracted images (g) and (f).

2.2.4 Partial volume relabelling and fractional content

The main contribution of the topology is the relabelling of missegmented GM voxels in hidden sulci as mixed GM/CSF. Once the topologically corrected WM, GM, CSF, WM/GM and GM/CSF segmentations are obtained, the portion of pure tissue, called here fractional content F , is computed for each mixed voxel by estimating the local contribution of each pure tissue. We assume that each voxel contains at most two tissues and the new labelling corresponds only to the mixed voxels WM/GM and GM/CSF. For each mixed voxel, the fractional content F ranges between $[0, 1]$ depending on the amount of pure tissue. Thus, for pure tissue voxels the fractional content F_j are set to 1 for the class j and 0 otherwise. For mixed voxels ($x \in \mathbf{WM/GM}, \mathbf{GM/CSF}$), the fractional content $F_{j/k}$ between both pure tissues j and k is computed using the intensity $I(x)$ of the image and the robust local averages of the closest pure tissue

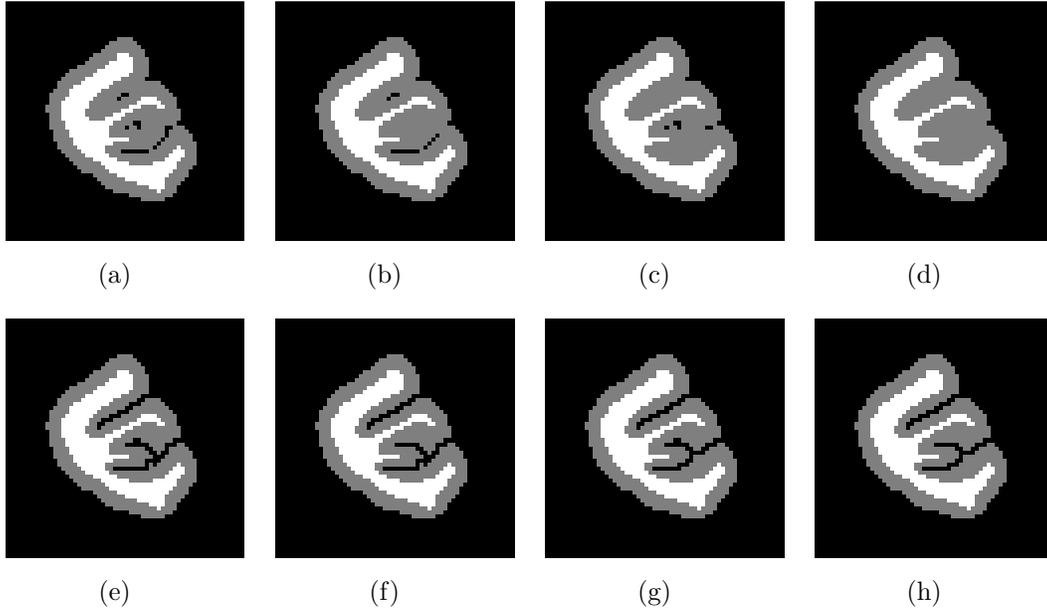


Figure **2-3**: First row (a) - (d): Different initial configurations of a synthetic phantom. Second row (e) - (h): Corresponding topologically corrected WM-GM segmentations.

voxels $\mu_j(x)$ and $\mu_k(x)$, such that:

$$F_{j/k}(x) = U \left(\frac{\mu_k(x) - I(x)}{\mu_k(x) - \mu_j(x)} \right) \quad (2-1)$$

where $U(\cdot)$ is a limiter restricting the range of the fractional content to $[0, 1]$. Unlike [117], which uses the same linear relation between global means of tissues to compute fractional content, we compute μ_k and μ_j as robust local averages rather than global means. This is done by computing the mean of the median 50% of pure tissue intensities (interquartile mean) within a $5mm$ radius sphere, thus rejecting local outliers, over a denoised version of the original MR image. The noise is removed by applying the optimized non-local means method proposed in [31].

Pure tissue voxels are selected by eroding pure tissue segmentations using a $2mm$ radius, therefore reducing the influence of any mixed voxel. Finally, the computed averages are propagated back towards the location of the mixed voxels x , resulting in values of $\mu_j(x)$ and $\mu_k(x)$ that represent the average of the closest pure tissue voxels (Figure **2-5**). The GM fractional content map is eventually defined as $F_{GM/WM} \cup F_{GM} \cup F_{GM/CSF}$. Using a robust local mean overcomes issues related to intensity inhomogeneities and variations of pure tissue signal across the image, weighting accordingly the signal when

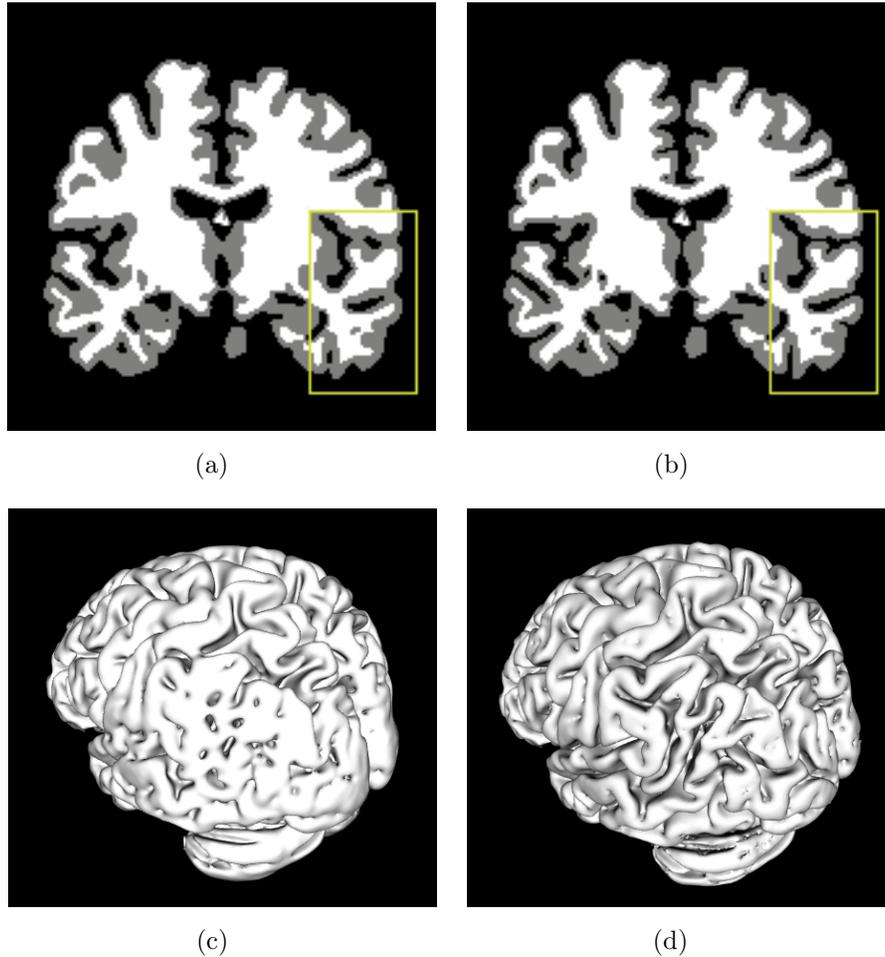


Figure 2-4: (a) Initial and (b) topologically corrected WM-GM segmentations, highlighted within the rectangle; (c) marching cubes reconstruction of GM before and (d) after the topology correction procedure.

computing the fractional content.

Regional differences in the cell structure and the distribution of different layers of the cortex result in variation of regional intensity differences for the same tissue across the brain. These differences produces local variation of contrast between the tissues that might be pronounced with ageing [114]. Whereas global homogeneity assumptions will bias the voxel fractional content estimation, a local computation of intensity averages for pure tissue yields a more accurate value, which accounts for the changes in cytoarchitecture visibles in MR. A local estimate allows also to overcome the issues of intensity inhomogeneities due to the artifacts during the acquisition.

To illustrate the spatial differences in signal, GM intensity was measured over the

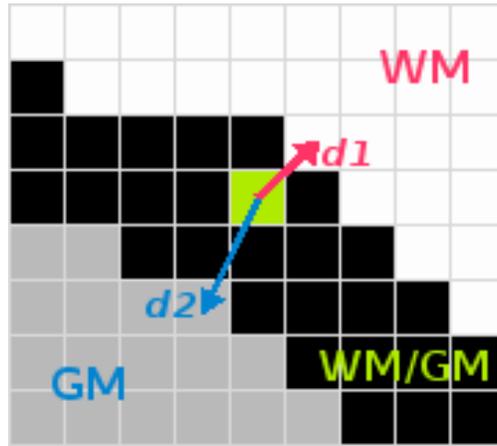


Figure 2-5: Schematic view of the local tissue averages for a given mixed voxel, where $d1$ and $d2$ relates to the closest voxels in the pure tissues.

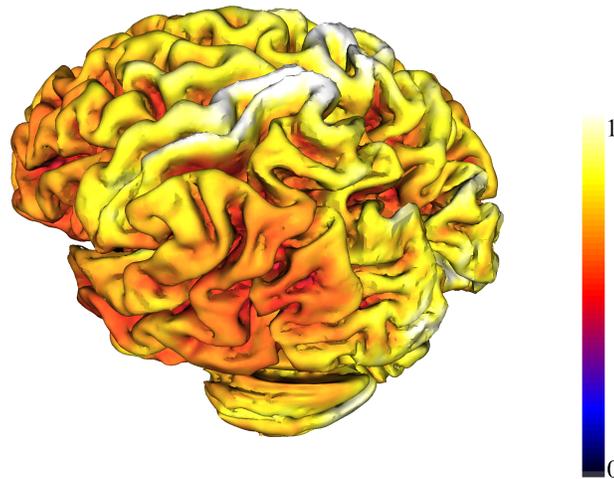


Figure 2-6: Averaged intensity within the connected components of the pure GM, computed as the interquartile mean (IQM) within a $5mm$ radius sphere on an OASIS example data, normalized by the Maximum of intensity. The differences between the regions clearly appear. Thus, GM tissue intensity will be different between the regions and global homogeneity assumptions will slightly bias the computation of partial volume.

population of 20 young adults scans, acquired as described in Section 2.3.3. Figure 2-6 shows the local average intensity of GM across the brain for an individual. In this example, precentral gyrus presented a higher average value than the temporal or

occipital lobe. The same pattern appear in average in all the healthy individuals. The contrast between the tissues has been also measured using the Equation 2-2 as

$$F_{Contrast} = \frac{\mu_{WM} - \mu_{GM}}{\mu_{GM} - \mu_{CSF}} \quad (2-2)$$

where μ_{WM} , μ_{GM} and μ_{CSF} are the regional averages of WM, GM and CSF respectively, which can be considered as a measure of the contrast between WM and GM normalized by the CSF. Figure 2-7 shows the regional differences for the population of 20 young controls.

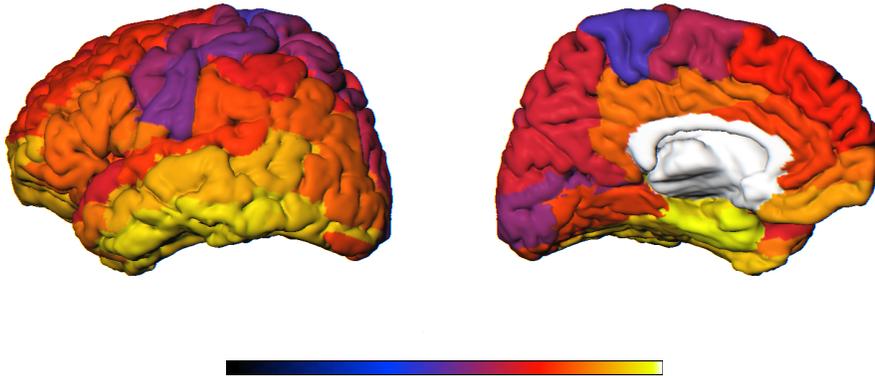


Figure 2-7: AAL template showing the regional differences in contrast between WM and GM over the surface, by calculating the ratio $\frac{\mu_{WM} - \mu_{GM}}{\mu_{GM} - \mu_{CSF}}$. Darkest colours indicate bigger ratios, light colours indicate small values. Left: lateral and Right: medial views.

2.3 Experiments

To evaluate our method, named hereafter as Topologically-corrected Partial Volume (TPV), we used different brain MR data sets including simulated and real images. The purpose was twofold, firstly to illustrate the effect of the topology correction in the estimation of fractional content for mixed voxels, and secondly to compare the obtained results with those publicly available in the area. After that, the method was integrated to our voxel-based cortical thickness estimation pipeline. Experiments demonstrated that the overall method showed a better estimate of thickness and a high reproducibility on real data.

2.3.1 Simulated MR data

A set of 15 simulated MR brain images was obtained from the BrainWeb Simulated Brain Database, maintained by the McConnell Brain Imaging Centre at the Montreal Neurological Institute [26] and available at www.bic.mni.mcgill.ca/brainweb. Each simulation was a $1mm^3$ isotropic T1-weighted MRI volume with dimensions $181 \times 217 \times 181$, generated with varying noise level and intensity inhomogeneity. We tested our method on each combination of 1%, 3%, 5%, 7% or 9% noise levels together with 0%, 20% or 40% intensity nonuniformities. BrainWeb also provides the fuzzy tissue membership volumes, one for each tissue class, together with a discrete anatomical model of the simulated normal brain.

2.3.2 Manually segmented real MR data

20 normal MR brain data sets and their manual segmentations were obtained from the Internet Brain Segmentation Repository (IBSR), provided by the Center for Morphometric Analysis at Massachusetts General Hospital and available at www.cma.mgh.harvard.edu/ibsr. The data sets were acquired along the coronal axis with slice dimension of 256×256 and $1mm^2$ resolution. Interslice distance is $3mm$ and the number of slices for each volume varies between 60 and 65. The data sets have various levels of artifacts, as low contrast and relatively large intensity gradients, that further affects performance of the algorithm. CMA also provides expert tissue labellings of each brain into WM, GM, and CSF, together with reference similarity values for some classification techniques.

2.3.3 Cross sectional series of real MR scans

20 young healthy subjects (12 female, 8 male; age between 19 - 34 years), who underwent 4 scans at baseline and 4 more scans during a subsequent session after a short delay (less than 90 days), were randomly selected from the Open Access Series of Imaging Studies (OASIS) database [89], available at www.oasis-brains.org. For each session, an average motion-corrected image (co-registered average of all available data) was used for our tests. The scans were T1-weighted Magnetization Prepared RApid Gradient Echo (MP-RAGE) in sagittal orientation with isotropic $1mm^3$ resolution ($256 \times 256 \times 128$ pixels). This data was used to assess the precision of the method when classifying partial volume voxels. We also tested the robustness when the method was integrated in our voxel-based cortical thickness estimation pipeline [1], particularly when the detection of deep sulci was improved.

2.3.4 Error and similarity measures

To quantitatively evaluate performance of the method over both simulated and real MR data sets and compare these results with other well-known results, we used two different metrics: the root mean square (RMS) error for comparison of PV classification maps, and the Jaccard similarity measure for comparison of the corresponding crisp tissue segmentations. The RMS error was used to quantify the differences between the fractional content calculated for each tissue and the corresponding values in the ground truth fuzzy membership images. As in [117], the RMS error between two images X and Y is calculated as

$$e_{RMS}(X, Y) = \sqrt{\frac{1}{|\Omega|} \sum_{k \in \Omega} |y_k - x_k|^2}$$

where Ω is the brain region, x_k and y_k are the image intensities at position k .

The Jaccard similarity metric, also known as the Tanimoto coefficient, measures the amount of overlap (agreement) between two images X and Y by taking the ratio between the size of their intersection and the size of their union:

$$J(X, Y) = \frac{|X \cap Y|}{|X \cup Y|}$$

This metric yields values between 0 and 1, where 0 means complete dissimilarity and 1 stands for identical images.

2.4 Results and discussion

2.4.1 BrainWeb

Performance of our TPV method was firstly assessed on the simulated brain images from BrainWeb. One example of the resulting PV maps for WM, GM and CSF, compared with the available ground truth, on the synthetic brain volume, 3% noise level and 20% bias field, is depicted in Figure 2-8. Comparisons between our method and a classical MAP approach are shown in Figure 2-9 for the computed GMPVC fractional content map. It must be noted that compared to a classical MAP approach as in [117], the sulci were better delineated by introducing the topological constraints (Figure 2.9(g)). In this example, a deep sulci voxel with similar intensity to the average GM, will be classified as GM and not as a mixed GM/CSF voxel unless anatomical constraints are introduced. The mean RMS error of fractional content over the entire BrainWeb data set significantly decreased to 6.1% ($p < 0.01$) for the obtained GMPVC map, as compared with the results reported in [117]. Overall, a good agreement was

shown between the computed PV maps and the ground truth, available as fuzzy tissue membership volumes. RMS errors for different noise and intensity nonuniformity levels are shown in Table 2-1. As expected, the computed error was robust to the bias field, which additionally validates the local averaging approach rather than the global one.

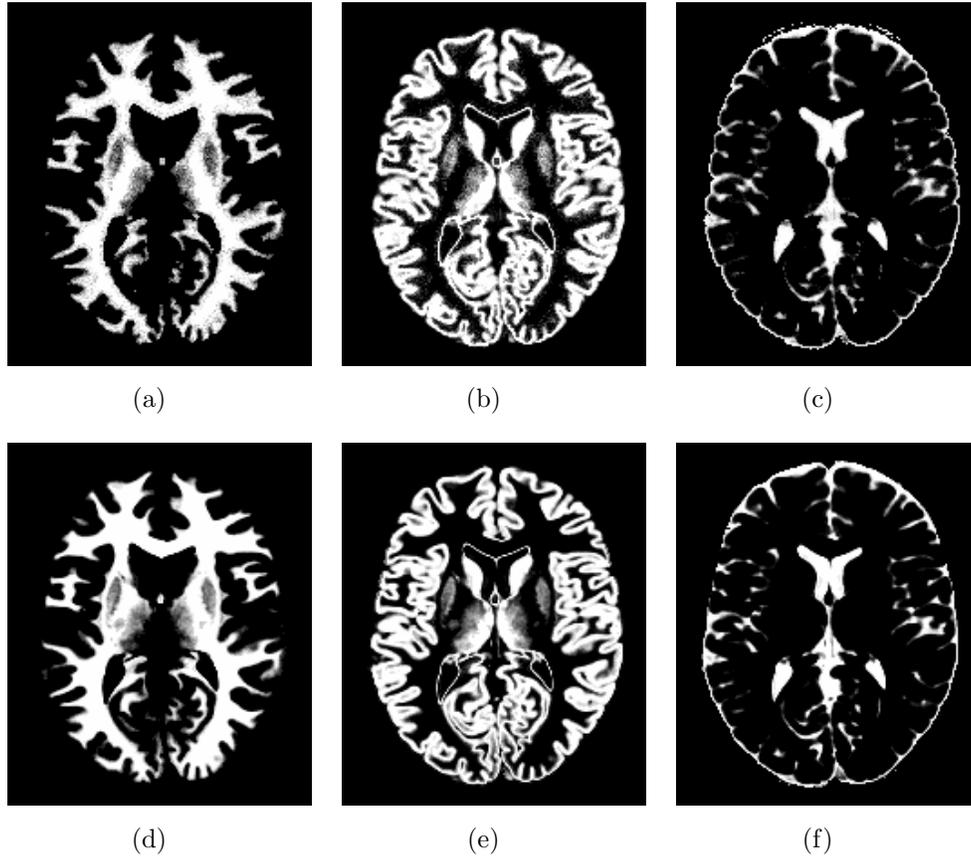


Figure 2-8: Partial volume segmentation of a simulated BrainWeb volume (3% noise, 20% bias field). PV maps for (a) WM, (b), GM (c) and CSF. Ground truth: (d) WM, (e), GM and (f) CSF.

The variability between different regions in the brain may affect the performance of PV classifiers [24]. To illustrate this effect, we used the automated anatomical labeling (AAL) template [123] to calculate the RMS error within each region as in [24]. Averaged results for different levels of noise are shown in Figure 2-10. As a low variability with respect to the bias field was observed, the depicted value corresponds to the average over all the bias field levels (0%, 20% and 40%). The smallest errors appeared in the amygdala (42xx), the insula (30xx), the supplementary motor area (24xx) and the olfactory (25xx); while lower agreement was found in the basal ganglia (70xx), the middle occipital (52xx) and the parietal superior (61xx).

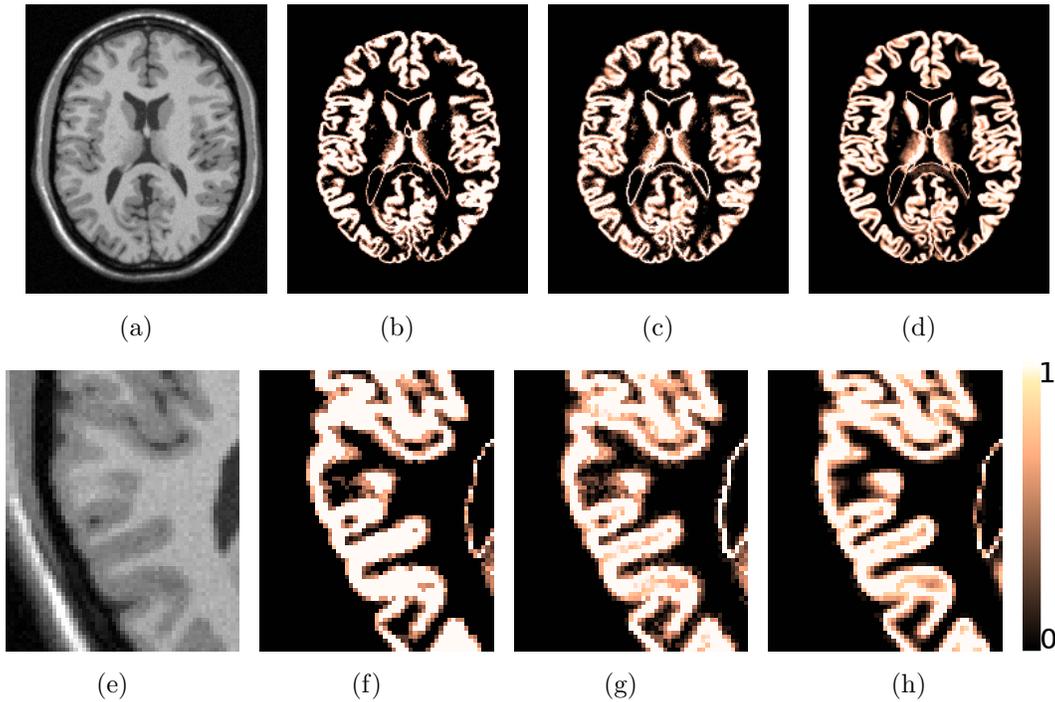


Figure 2-9: Example of PV estimation of a simulated BrainWeb volume (3% noise, 20% bias field). (a),(e) Original image, (b),(f) MAP PV estimation, (c),(g) Topologically-corrected PV, (a),(h) ground truth. In the detailed views we can observe the improvement in deep sulci, (g) relative to (f), brought by the topology correction.

Noise	Intensity inhomogeneity					
	0%		20%		40%	
	WM	GM	WM	GM	WM	GM
1%	0.129	0.130	0.129	0.131	0.125	0.132
3%	0.139	0.142	0.140	0.141	0.140	0.142
5%	0.174	0.174	0.172	0.171	0.170	0.171
7%	0.214	0.216	0.210	0.213	0.208	0.212
9%	0.251	0.261	0.245	0.258	0.242	0.256

Table 2-1: Fractional content RMS error on BrainWeb.

We also compared our TPV method with the results reported by Chiverton et al. [24] (GSR) and Shattuck et al. [117] (SMAP). The results are depicted in Figure 2-11. Evidence suggests that the local average intensity strategy makes the classification more robust to bias field variations, and on average performs better than other methods

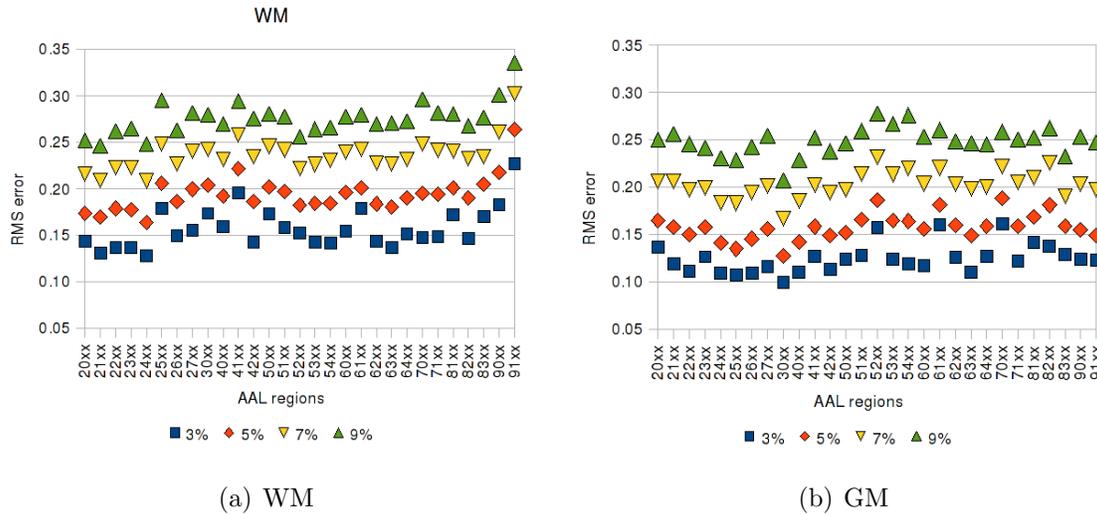


Figure 2-10: RMS error per AAL region (a) WM and (b) GM regions, for different noise levels using the same labels as [24].

for low levels of noise (1% to 7%) and bias field of 20%. We point out the fact that GSR does not explicitly take into account the bias field, hence its effect appears in the reported results.

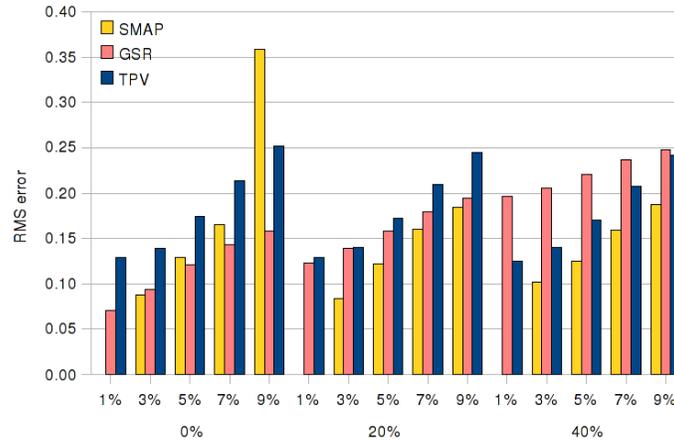
2.4.2 Real MR Data

OASIS

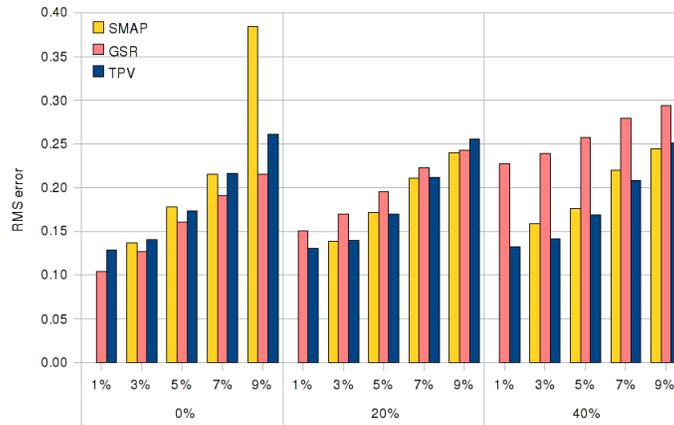
The reproducibility was measured by applying the method to two of the MR scans from the same individual from the OASIS database. We compared the results with the MAP classifier as in [117]. Significant improvements in GM PV estimation were brought by the topology correction. The reproducibility error decreased by 8.8% in GM and 8.5% in WM ($p < 0.001$), measured as the RMS between the PV maps obtained on the rigidly registered baseline and repeat scans. Likewise, when comparing the crisp segmentations obtained by thresholding by 0.5 the baseline and repeat GM PV maps, the Jaccard similarity measure increased by 3.5% in GM. To compute crisp segmentations, each mixed voxel was assigned to the tissue class with the highest fractional content and the obtained segmentation were subsequently compared.

IBSR

Our method was also compared with both TMCD (trimmed minimum covariance determinant) [122] and MMC (mixture model clustering) [93] on the IBSR data sets. Since



(a) WM



(b) GM

Figure 2-11: PV estimation errors for (a) WM and (b) GM on BrainWeb, for different noise and bias field levels. (SMAP results for 1% noise not publicly available)

the ground truth is available as manual segmentations performed by clinical experts, we compared the segmentations obtained from the crisped PV maps. Figure 2-12 shows an example of the ground truth provided by IBSR and a hard segmentation calculated after applying our method. Figure 2.13(b) depicts the results of the comparison for the GM in the 20 normal subjects. As in [24], results of manual expert segmentation and pure tissue classification presented by Ibrahim et al. [64] (HMM, hidden Markov model) were included for reference. Significant improvements in GM classification were demonstrated using the TPV, compared to a MAP classifier. The similarity measure (Jaccard) was improved by 8.7% in GM and 2.6% in WM ($p < 0.001$).

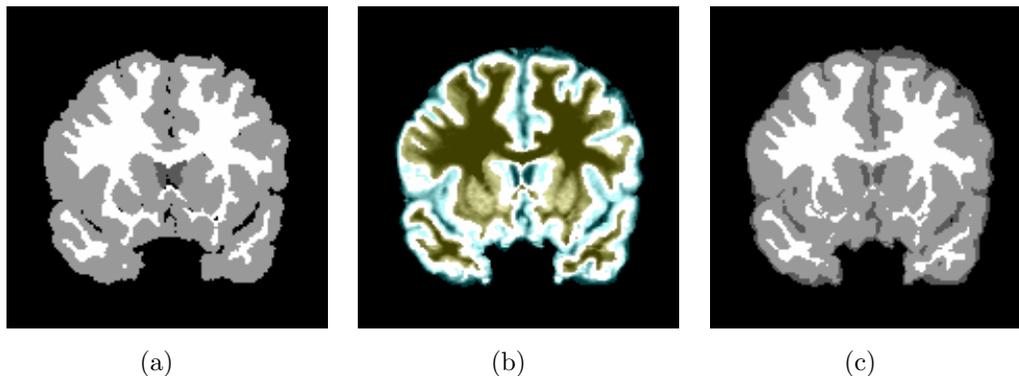


Figure 2-12: (a) IBSR Ground truth pure tissue classification. (b) Estimated PV maps (blue: GM/CSF, white: GM, yellow: GM/WM) and (c) computed crisp segmentation.

Poor similarity results were obtained in 5 cases, which exhibited strong shading artifacts that impeded a reliable GM and WM classification. Similar findings were presented in [93], who excluded them from the analysis. We also observed that the anisotropy in the images biased the computation of the local averages. Table 2-2 summarizes the mean (\pm standard deviation) of the Jaccard similarity values for each method, excluding the volumes with too severe intensity inhomogeneity. In average, our TPV method performed better for WM and GM compared to the others, excepting averaged GM segmentation against [93]. It must be noted that when the PV maps were used to generate the crisp segmentations, the mixed GM/CSF voxels in deep sulci with fractional content above 0.5 might be wrongly reclassified as GM. Under those conditions, the contribution of topology correction in the segmentation can not be fully and accurately validated with this experiment. Nonetheless, we report these results for completeness.

	MMC [93]	TMCD [122]	TPV
WM	0.648 (\pm 0.198)	0.696 (\pm 0.050)	0.701 (\pm 0.042)
GM	0.753 (\pm 0.120)	0.697 (\pm 0.064)	0.708 (\pm 0.045)

Table 2-2: Mean (\pm standard deviation) of Jaccard similarity index for each method.

2.4.3 Computational performance

On each image of the BrainWeb data set, after the initial MAP segmentation, the topology correction and PV fractional content estimation takes less than 10 minutes. For the OASIS data sets, the procedure takes about 9 minutes, while for the IBSR images the topology correction and PV fractional content estimation takes less than 4

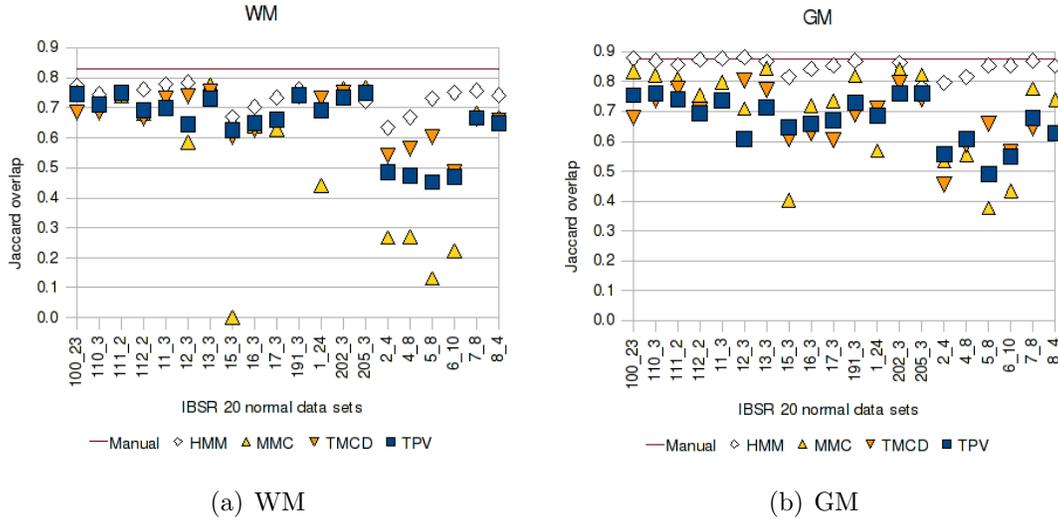


Figure 2-13: Jaccard similarity results for WM (a) and GM (b).

minutes. Operations were encoded in a single-thread application and then executed in a standard Intel Core 2 Duo (3.00GHz, 2 GB RAM) machine running Linux.

2.4.4 Deep sulci cutting and cortical thickness estimation on real data

We integrated the proposed sulci detection and improved partial volume classification methods to our cortical thickness estimation pipeline [1], as depicted in Figure 2-14. Then, we computed the thickness, at two different acquisition times, for the same 20 young healthy subjects from the OASIS database [89] used in the experiment described in Section 2.3.3. The reproducibility was assessed by using the Pearson correlation coefficient for each Region Of Interest (ROI) of the AAL template [123], excluding the cerebellum and subcortical *nuclei* from the analysis.

Thickness estimation with the proposed method (TPV) showed a higher reproducibility compared with the measure performed after partial volume classification using [117]. As can be seen in Figure 2-15, the differences in cortical thickness between scans were reduced after applying the TPV. The Pearson correlation coefficient was 0.915 in average and a paired t-test did not reveal any significant differences between the two measurements ($p < 0.1$). Also, the difference between scans was decreased by 13.7% in average, as shown in Table 2-3.

By using the proposed method, we found a mean (\pm std. dev.) cortical thickness over the whole brain of $2.08mm (\pm 0.11)$ for all the subjects, which is within the accepted range of cortical thickness for healthy young adults. In previous studies, when the

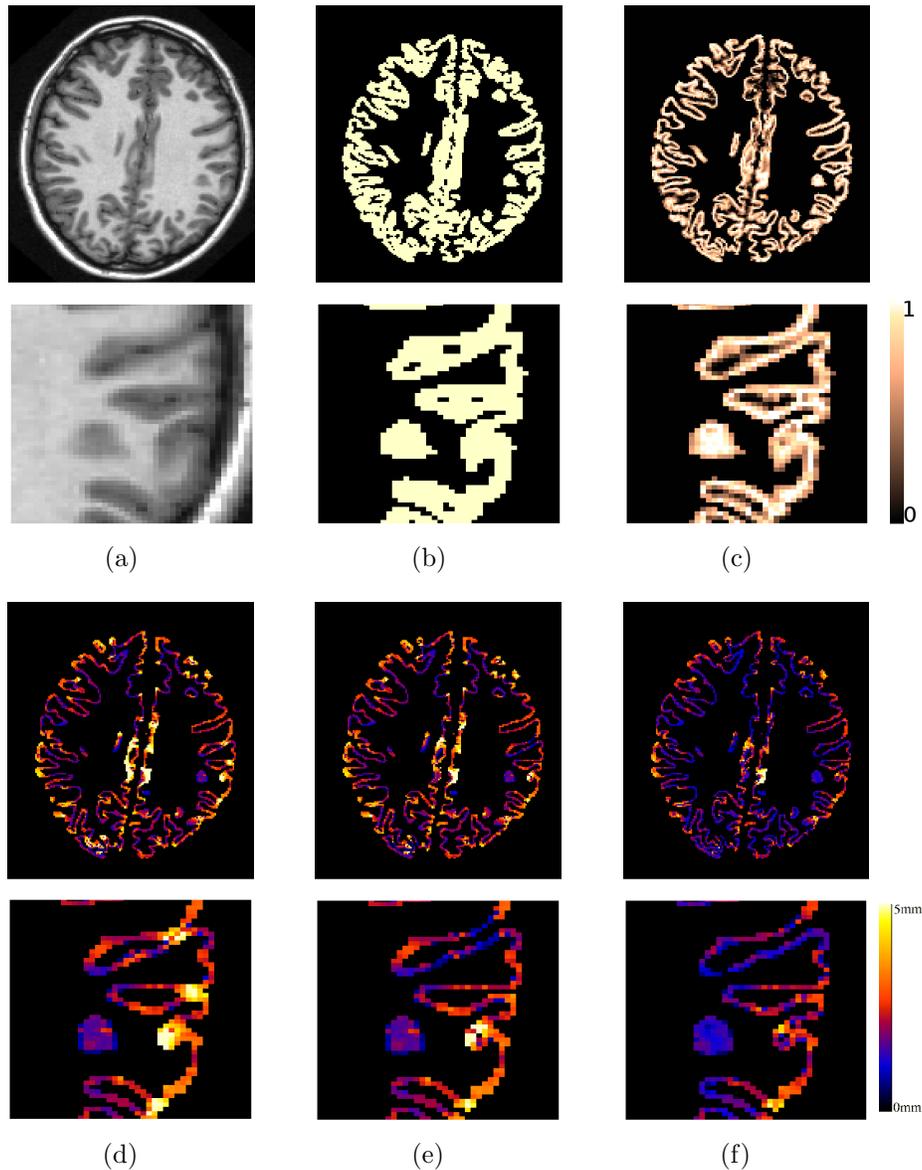


Figure 2-14: Example of cortical thickness estimation from MR. (a) Original T1-W MRI, (b) GM segmentation, (c) Topologically-corrected GM PV map. Cortical thickness maps (d) without any topology modifications, (e) after topology correction only, (f) after TPV. In the detailed views we can observe the improvement brought by the topology to delineate deep sulci zones, which allows an accurate measurement of the cortical thickness.

PV is not taken into account as in [135], the computed mean thickness for the same population was 4.69mm (± 0.11). And when the PV classification method proposed

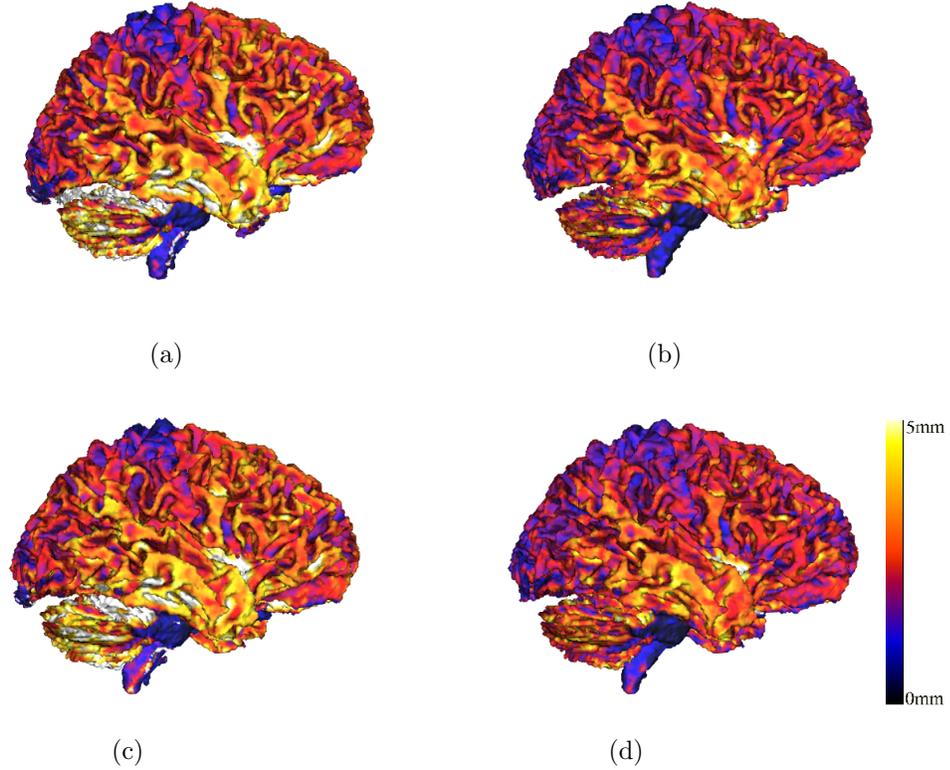


Figure 2-15: Surface representation of cortical thickness, computed at different steps for two scans of a single subject (OASIS). *Top row*: Scan 1, *Bottom row*: Scan 2. (a),(c) Without topology modifications, and (b),(d) with topologically-corrected GM PV map (TPV). Overall, we can observe the high values of thickness corrected with the TPV method.

Brain lobule	Correlation coefficient		Differences between scans	
	SMAP	TPV	SMAP	TPV
Frontal	0.922	0.930	0.090	0.090
Limbic	0.901	0.883	0.158	0.121
Occipital	0.902	0.904	0.101	0.063
Parietal	0.906	0.920	0.058	0.060
Temporal	0.932	0.938	0.105	0.106
Average	0.912	0.915	0.102	0.088

Table 2-3: Pearson correlation coefficient and differences between scans for the OASIS dataset, grouped by brain lobules.

by Shattuck et al. [117] is used, without any topology correction, the computed mean thickness was 3.06mm (± 0.25); using those same PV maps, but correcting the topology problems, decreases the mean thickness to 2.75mm (± 0.17).

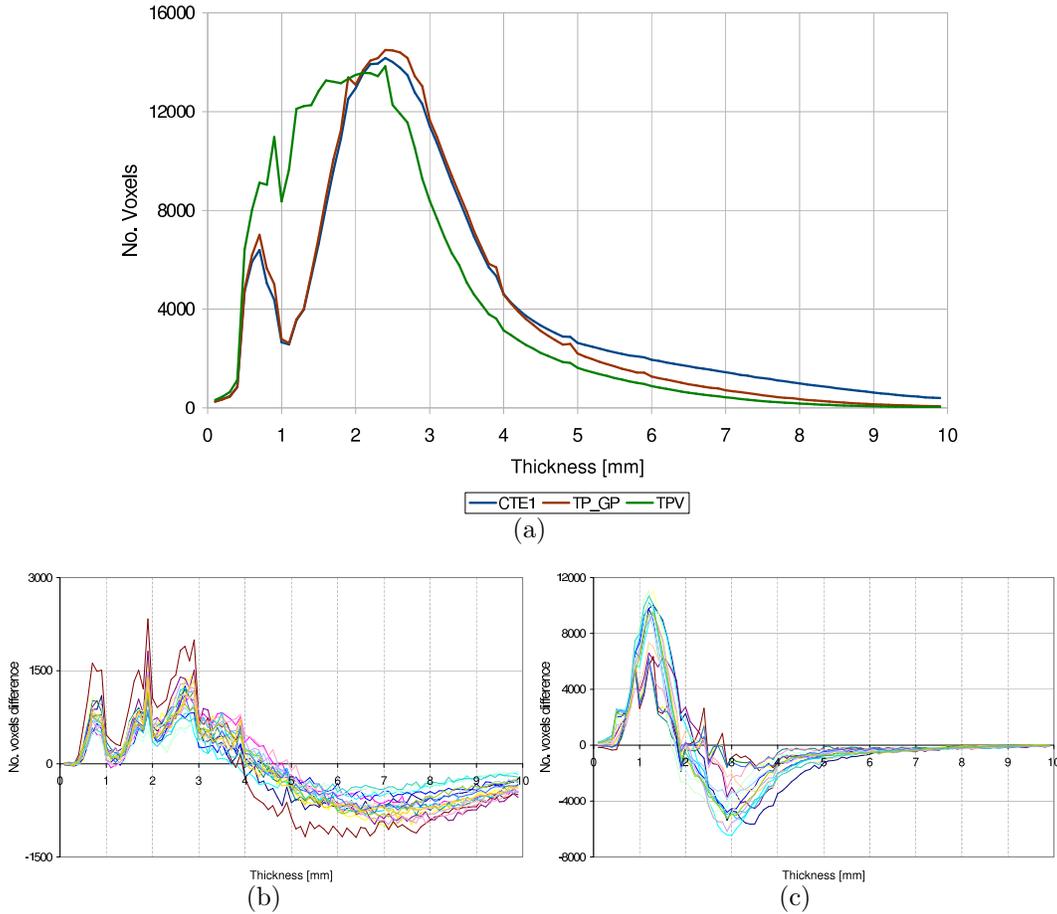


Figure 2-16: (a) Histogram of the average thickness for the 20 MR before topology correction (step 1), after topology correction (step 2) and with TPV. It is shown how the number of higher thickness voxels was reduced. (b) Differences in cortical thickness histograms between steps 1 and 2 for the 20 MR. This figure depicts the improvement after the topology. The number of voxels above 4mm in average has been dramatically reduced. (c) Differences between topology and TPV, in average the number of voxels above 2.5mm has been reduced consolidating the average thickness around 2.5mm (typical value for young adults).

Fig. 2-16 depicts in histograms the impact of the topology correction and the accurate PV estimation on the cortical thickness calculation task. The higher thickness values

produced after the first PV classification disappeared when the topology of GM is corrected and the accurate PV value is computed with the TPV. Fig. 2.16(a) shows the histogram of the average thickness for the 20 MR subjects before any topological modifications, after the topology correction and with TPV. Fig. 2.16(b) depicts the differences for each of the cortical thickness histograms between Step 1 and Step 2, illustrating the improvement after the TPV. The number of voxels above $4mm$ in average has been dramatically reduced. Fig. 2.16(c) shows the differences between topology corrections and TPV, in average in this further step the number of voxels above $2.5mm$ has been reduced.

2.5 Conclusion

We have described a simple and fast technique to improve PV estimation of brain tissues from T1W MRI. It improves the detection of hidden mixed voxels in deep sulci by correcting for the topology errors in the segmentation and uses local averages to better estimate the fractional content. We show that fractional tissue content estimation can be improved for low levels of noise and regardless the intensity inhomogeneity, resulting in superior brain tissue segmentations.

Topology correction improved the classification of mixed voxels in opposed banks of buried sulci by assuming GM as a continuous layer following the WM, with the topology of a filled sphere. Local modelling of tissue intensities helps to overcome the issues related with local intensity inhomogeneity and tissue MR properties across the image. Even with a preprocessing stage to correct the intensity inhomogeneities, pure cortical tissues show different intensity levels in the MRI. This suggests that the tissue properties are different depending on the region of the brain. Accuracy and precision were demonstrated and comparisons with other methods showed comparative performance with simulated and real MR data.

We demonstrated the usefulness of the method to improve the accuracy of the cortical thickness estimation. By labelling mixed GM/CSF voxels in deep sulci and by recomputing a spatially compensated PV map, the measure of thickness in difficult regions is improved. Our method showed a high reproducibility on real data, with an extremely good agreement between the baseline and repeat scans. The computed values of thickness for young adults are similar to the ones reported previously in the literature. In the future, we plan to use our technique on clinical data to study cortical atrophy in Alzheimer's disease and other neurodegenerative diseases. We intend also to develop voxel-based techniques for inter-subject comparisons, a challenging issue given the large anatomical variability between patients.

3 Single-image Super-Resolution of Brain MR Images using Overcomplete Dictionaries

*Resolution in Magnetic Resonance (MR) is limited by diverse physical, technological and economical considerations. In conventional medical practice, resolution enhancement is usually performed with bicubic or B-spline interpolations, strongly affecting the accuracy of subsequent processing steps such as segmentation or registration. This chapter presents a sparse-based super-resolution method, adapted for easily including prior knowledge, which couples up high and low frequency information so that a high-resolution version of a low-resolution brain MR image is generated. The proposed approach includes a whole-image multi-scale edge analysis and a dimensionality reduction scheme, which results in a remarkable improvement of the computational speed and accuracy. The proposed method is shown to obtain accurate high-resolution reconstructions, outperforming the baseline interpolation and a recent state-of-the-art algorithm, and suggesting a substantial impact in voxel-based morphometry studies. The complete content of this chapter has been published as a research article in the **Medical Image Analysis** journal (see [112]).*

3.1 Introduction

Resolution in Magnetic Resonance (MR) is limited by diverse physical, technological and economical considerations. These factors together introduce a series of artifacts, such as the partial volume (PV) effect, affecting the performance of image analysis and post-processing algorithms, and preventing derivation of accurate measurements. In conventional medical practice, interpolation of images to higher resolutions is usually performed by applying standard image processing techniques such as the bicubic or B-spline interpolation. This interpolation has a strong influence on the subsequent processing steps, such as segmentation or registration. Hence, improving image resolution is one of the main challenges in medical image processing. The fundamental problem can be stated as if some high-frequency information (edges) has been lost

during the acquisition process. Therefore, any reconstruction technique should be able to coherently recover these high-frequencies.

Super-Resolution (SR) techniques have been broadly used to increasing medical image resolution [58]. At the beginning, these methods attempted to recover a high-resolution image by combining multiple shifted low-resolution acquisitions. Two kinds of approaches can be identified: one works at the acquisition level over raw data (frequency space), while the others act on the volumetric images (spatial or image space) as an additional processing step. At the acquisition stage, the k -space data can be manipulated and combined to obtain adequate spatial resolution while reducing acquisition time [62]; or parameters can be configured to obtain multiple scans with different slice directions which are then mixed up [118]. Regarding volumetric images, Peled et al. [103] and Greenspan et al. [59] have proposed the first approaches to adapt the iterative back-projection method proposed by Irani et al. [65] to 2D and 3D MR images, respectively; followed by other strategies such as the resolution enhancement method described by Carmi et al. [20]. Recent approaches have changed the classical SR paradigm with multiple images, evolving towards the use of information from a single low-resolution image, but also improving the image information by combining different modalities. Patch-based approaches with non-local regularization frameworks have been proposed by Rousseau [109] and Manjón et al. [88], which have also extended the formulation to super-resolve low-resolution T2w images using high-frequency information from T1w images [87, 110].

A recent trend in signal and image processing is the use of models that exploit the natural redundancy of signals, taking advantage of the fact that media signals, such as audio, images and video can be sparsely represented using transform-domain methods. Many important tasks involving this kind of signals can be better solved as sparse solutions to undetermined systems of linear equations [16]. This kind of models have shown to outperform common approaches for inverse problems and have led to important state-of-the-art results, with successful applications in synthesis [105], denoising [43], restoration [85], reconstruction and compression, among others. One of the classical inverse problems studied under the sparse representation framework is SR, a problem for which important state-of-the-art results have been reported in natural images [133, 137]. In MR images, efforts have been focused in applying or adapting the Compressed (or Compressive) Sensing (CS) approach [38], which is claimed to be able to accurately reconstruct MR images from a small subset of Fourier coefficients (k -space samples), as shown by Lustig et al. [83]. Ravishankar et al. [107] propose a CS modification which uses adaptive dictionaries, by alternating back and forth between image domain and k -space, while Adluru et al. [2] applies a reconstruction method with Total Variation constraints in space and time. Other strategies include the adaptation

of the focal underdetermined system solver (FOCUSS) for MR projection reconstruction [134].

While SR methods working in the frequency domain could be preferred over those working in the image space, given their theoretical simplicity and low computational cost, they also present some important drawbacks. Completion of high-frequencies only implies adding punctual values to the k -space, while in the image space these added values generates aliasing and visual artifacts, according to the point spread function (PSF) of the acquisition. Besides, this PSF can be highly variable under certain subsampling policies, the smaller the number of samples the larger this PSF can be. Therefore, SR methods would need to go back and forth from the frequency to the spatial domain in order to minimize the occurrence of these artifacts. On the other hand, dictionary patches in the image domain provides much more interpretability to medical specialists than particular punctual frequencies in the frequency space. These image patches can be directly related to pathologies or imaging findings and more importantly, they also facilitate the inclusion of a priori medical knowledge.

For achieving minimization of PV effects, we have addressed the problem from a completely new angle: rather than developing PV segmentation algorithms, we chose to use very simple segmentation methods on improved input data. To recover the missing information, we used a sparse representation framework which builded up a high-resolution version from a low-resolution image, as in [133] and [137]. Provided that a straight applicability of this approach was impossible because of the computational time per slice but also because the original approach has been implemented so far only for 2D images, we have adapted the whole method to handle MR brain volumes. First, low-resolution and high-resolution coupled dictionaries were constructed by randomly sampling 3D patches from previously preselected tissue regions in low-resolution and high-resolution images. Then, a multi-scale edge filtering, followed by a dimensionality reduction, are performed on both the low-resolution image and the low-resolution dictionary. Afterwards, the low-resolution image is described as a sparse combination of the patches in the low-resolution dictionary. Finally, the obtained sparse vector is projected onto the high-resolution dictionary to generate the high-resolution reconstruction.

The main differences between our proposal and those from Yang et al. [133] and Zeyde et al. [137] are: an appropriate missing edge analysis for MR images, that uses 3D multi-scale Sobel filters and which acts as an adequate sparsifying transform of the brain boundaries information (the information of interest); a knowledge-driven patch selection criteria, based on previous segmentations of brain tissues, that allows to construct semantic-based dictionaries from brain MR images; and the selection of non-overlapping patches for local reconstruction, which greatly reduces the computational

reconstruction time. We will show that the precision of brain tissue segmentation can be improved by applying this technique, furthermore, increasing the accuracy of brain morphometrical tasks. The proposed method was also compared with a state-of-the-art approach that performs non-local MRI upsampling [88], showing that our proposal produces better results, both in accuracy and execution time. Finally, the method was evaluated in a real scenario: a morphometric study of pathologic subjects versus controls, indicating that both the original and reconstructed images generate a very similar statistical map in a voxel-based morphometry study.

The rest of the chapter is organized as follows. Section 3.2 introduces some generalities about sparse representations, while the mathematical formulation of the SR problem is presented in Section 3.3, and the different stages of the proposed methodology are detailed in Section 3.4. Section 3.5 proposes an extensive validation on different brain MR datasets, and the discussion presented in Section 3.6 concludes the paper.

3.2 Sparse Representations in Image Processing

Image analysis, from its beginning, has gathered concepts from many different domains [56]. Recently, several techniques have exploited the fact that it turns out to be much more difficult to perform an analysis at the level of the whole image structure than at the level of its parts, a fundamental idea that comes from the domain of neurosciences [95]. Briefly stated, Olshausen and Field developed a learning method under the assumption that information is somehow sparsely coded. This method decomposes an image into a set of parts (atoms) that are localized, oriented and frequency band-limited, as performed by the primary visual area V1 of mammalian brains. Based on Barlow's principle of redundancy reduction [8], the representation searches for those atoms with the largest statistical independence, resulting in non-orthogonal elements that cover the space and that grouped together constitute a dictionary, i.e. the basic code behind semantics.

Different image analysis approaches, such as non-negative matrix factorization [79] or sparse and redundant representations [95], use this assumption at their very base. These two methods, in particular, identify the constituent parts of a scene and then, using some of them, the same scene or similar ones may be accurately reconstructed. These parts, denoted as basis functions or atoms, are usually arranged in overcomplete dictionaries with a larger number of elements than the effective dimensionality of the input space, thereby representing a wider range of image phenomena. From this perspective, it can be assumed that there exists a random generation machine \mathcal{M} that returns images following a prior distribution $P(x)$, which is defined by the representation coefficients (see Appendix B).

3.3 Super-Resolution of Brain MR Images

The super-resolution problem can be mathematically stated as follows. Given an original high-resolution image $X \in \mathbb{R}^z$ (represented as a vector of z pixels), its corresponding low-resolution version is denoted as $Y \in \mathbb{R}^q$ (with $q < z$), where the relation between the two images can be modeled as

$$Y = \mathbf{L}_M \mathbf{B} X \quad (3-1)$$

where $\mathbf{B} : \mathbb{R}^z \rightarrow \mathbb{R}^z$ is a linear (blurring) filter and $\mathbf{L}_M : \mathbb{R}^z \rightarrow \mathbb{R}^q$ is the down-sampling operator (by a factor M). \mathbf{B} attenuates high-frequency image information, while \mathbf{L}_M is a decimation operator (defined as taking each M th value starting from zero in each dimension). This means that the observed low-resolution image Y is a blurred and downsampled version of X .

In the single image super-resolution problem, the goal is to approximately recover a high-resolution image $\hat{X} \in \mathbb{R}^z$ given its blurred and down-sampled version Y , such that $\hat{X} \approx X$. To obtain a feasible solution to this problem, we have chosen to work at the level of small patches and to apply the sparse representation framework, aiming to represent each patch from the images by using a linear combination of some atoms from a dictionary.

Denote as $R_c : \mathbb{R}^q \rightarrow \mathbb{R}^n$ the operator that allows to extract a low-resolution patch $p_Y^c = R_c Y \in \mathbb{R}^n$ of size $\sqrt[3]{n} \times \sqrt[3]{n} \times \sqrt[3]{n}$ from the image Y around location $c = (i, j, k)$. With the sparse generative model, each patch p_Y^c can be projected over the dictionary $D_\ell \in \mathbb{R}^{n \times a}$, which characterizes the low-resolution patches. This projection produces a sparse representation of p_Y^c via $\alpha^c \in \mathbb{R}^n$, namely

$$p_Y^c = D_\ell \alpha^c$$

where $\|\alpha^c\|_0 \ll n$.

The corresponding high-resolution patch $p_X^d = S_d X \in \mathbb{R}^m$, with size $\sqrt[3]{m} \times \sqrt[3]{m} \times \sqrt[3]{m}$ (where $m = nM^3$), is then extracted from the image X around the corresponding location $d = Mc$. Applying again the sparse generative model we have

$$p_X^d = D_h \alpha^d$$

where $D_h \in \mathbb{R}^{m \times a}$ is the dictionary that characterizes the high-resolution patches, and is coupled to D_ℓ through the relation $D_\ell = \mathbf{L}_M \mathbf{B} D_h$. This means that each atom in D_h has its corresponding low-resolution version in D_ℓ and viceversa. Given the relation between D_ℓ and D_h , and c and d , it can be assumed that the sparse representation of a low-resolution patch in terms of D_ℓ can be directly used to recover the corresponding high-resolution patch from D_h , namely, that $\alpha^c = \alpha^d$.

With this in mind, the reconstructed high-resolution image \hat{X} can be built up by applying the sparse representation to each p_Y^c and then using the estimated α^c with D_h to obtain each \hat{p}_X^d , which together form the image \hat{X} .

3.3.1 Dictionary Considerations

As mentioned in Appendix B, the sparse representation approach is based on the assumption that the dictionaries D_ℓ and D_h are overcomplete, with more atoms than the signal dimensions, allowing to represent a wide range of signal phenomena. However, the choice of an optimum dictionary for a given task, despite many good approximations proposed so far, can still be considered as an open problem. In particular, Yu et al. [136] indicates that, for obtaining precise and stable sparse super-resolution estimates, the chosen dictionaries should meet some necessary qualitative conditions:

- **Sparsity:** D_h provides a sparse representation for p_X^d .
- **Recoverability:** The dictionary atoms D_h^a have non negligible norms $\|\mathbf{U}D_h^a\|^2 \gg 0$, where \mathbf{U} represents a degradation operator, in this case $\mathbf{U} = \mathbf{L}_M\mathbf{B}$.
- **Stability:** The transformed dictionary $D_\ell = \mathbf{L}_M\mathbf{B}D_h$ is incoherent enough, in the sense that the columns in D_ℓ are not too similar between them.

3.4 Proposed Methodology

The proposed method consists of two separate stages. First, the coupled low-resolution and high-resolution dictionaries must be constructed from training images (Figure 3-1). Then, a two-step reconstruction algorithm is applied to a new low-resolution image to recover its estimated high-resolution version. This strategy makes use of a local model with the sparse prior to recover lost high-frequencies in a patch-by-patch basis (Figure 3-2 in page 43), followed by a global correction which removes discontinuity effects and ensures consistency and naturalness of the final result (Figure 3-3 in page 43). The complete super-resolution process is summarized in Algorithm 1.

3.4.1 Dictionary Construction

Dictionaries are constructed (as depicted in Figure 3-1) starting from a training set, composed by some high-resolution images $\{X^j\}_j$. The corresponding low-resolution image set $\{Y^j\}_j$ is constructed by blurring and down-sampling, by a factor M , each training image (where $Y^j = \mathbf{L}_M\mathbf{B}X^j$). Finally, the upsampled set $\{Y_X^j\}_j$ is constructed by scaling-up back again each low-resolution image to the original high-resolution size

Algorithm 1 Image SR via Patch-based Sparse Representation

Require: semantic-based dictionaries D_h and D_ℓ , a low-resolution image Y

Upsample low-resolution image Y by $Y_X = \mathbf{H}_M Y$

Apply the multi-scale edge analysis on Y_X (Equation 3-2 in page 41)

for each patch $p_{Y_X}^d$ of Y_X , at locations d taken starting from the anterior-upper-left corner **do**

 Solve the optimization problem for $p_{Y_X}^d$ (find α^c through Equation 3-3 in page 42)

 Generate the high-resolution patch with $\hat{p}_X^d = D_h \alpha^d$

 Place the patch \hat{p}_X^d into the high-resolution image \hat{X}_0

end for

Apply the discontinuity correction method (Equation 3-5 in page 44) to find the closest image (\hat{X}) to \hat{X}_0 which satisfies the reconstruction constraint (Equation 3-4 in page 44)

return the super-resolution image \hat{X}

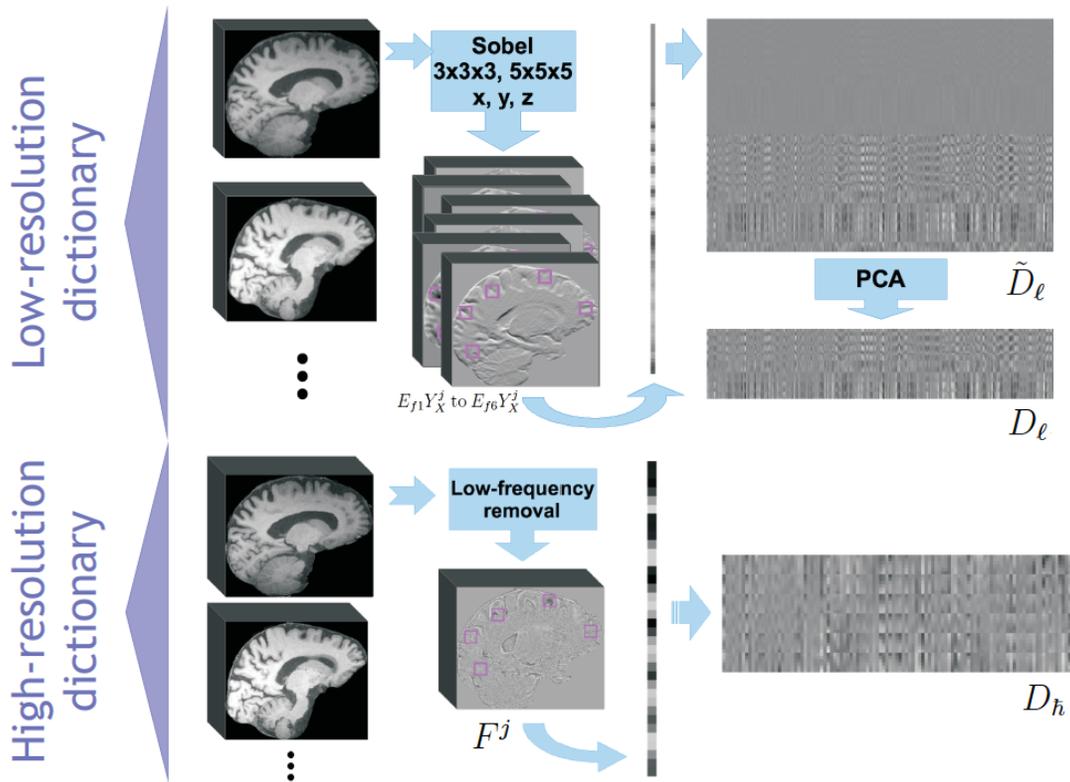


Figure 3-1: Illustration of low- and high-resolution dictionary construction.

using \mathbf{H}_M , the upsampling operator implemented as a bicubic interpolator (where

$$Y_X^j = \mathbf{H}_M Y^j.$$

As proposed by Zeyde et al. [137], the high-resolution images are processed to obtain only the high-frequency information, by using the upsampled images to remove their low-frequencies through $F^j = X^j - Y_X^j$. On the upsampled images, the preprocessing step that we propose involves a multi-scale edge analysis, where a series of 6 different filters (Sobel kernels, size $3 \times 3 \times 3$ and $5 \times 5 \times 5$, in x , y and z directions) are applied to the upsampled image set $\{Y_X^j\}_j$. In brain MR images, this multi-scale edge analysis make explicit certain aspects of the missing information, namely, boundaries with structural regularity and directionality. The Sobel operator combines a perpendicular smoothing, denoted as h_3 for size $3 \times 3 \times 3$ and h_5 for size $5 \times 5 \times 5$, with a simple central difference, denoted as h'_3 and h'_5 , respectively:

$$h_3(-1) = 1; h_3(0) = 2; h_3(1) = 1;$$

$$h'_3(-1) = 1; h'_3(0) = 0; h'_3(1) = -1;$$

$$h_5(-2) = 1; h_5(-1) = 4; h_5(0) = 6; h_5(1) = 4; h_5(2) = 1;$$

$$h'_5(-2) = 1; h'_5(-1) = 2; h'_5(0) = 0; h'_5(1) = -2; h'_5(2) = -1;$$

With this definitions, the 6 different kernels applied for the multi-scale image analysis have the form:

$$\begin{aligned} E_{f1} &= h'_3(x)h_3(y)h_3(z), & E_{f4} &= h'_5(x)h_5(y)h_5(z) \\ E_{f2} &= h_3(x)h'_3(y)h_3(z), & E_{f5} &= h_5(x)h'_5(y)h_5(z) \\ E_{f3} &= h_3(x)h_3(y)h'_3(z), & E_{f6} &= h_5(x)h_5(y)h'_5(z) \end{aligned} \quad (3-2)$$

where x , y and z correspond to a certain location within the image volume. With these kernels, 6 different filtered images ($E_{f1}Y_X^j$ to $E_{f6}Y_X^j$) are obtained.

Prior knowledge related to brain tissues is introduced at this stage by processing also the high-resolution images $\{X^j\}_j$ with standard medical image processing tools, which remove the skull and skin, and produce binary image segmentations of the three brain tissues (gray matter X_{GM}^j , white matter X_{WM}^j and cerebrospinal fluid X_{CSF}^j).

Following the preprocessing tasks described before, D_ℓ and D_h dictionaries are constructed by collecting image patches only in a predetermined number of random image locations d (arranged in a set Φ) from the $\{Y_X^j\}_j$ images. The working area is restricted to the voxels belonging to a (2-voxel) dilated version of the binary gray matter segmentations $\{X_{GM}^j\}_j$, thereby working with patches that mainly belong to the interfaces between tissues (WM-GM or GM-CSF), i.e., with important edge information. At each location d of the high-frequency image F^j , a patch p_X^d of size $\sqrt[3]{m} \times \sqrt[3]{m} \times \sqrt[3]{m}$ is extracted, while the corresponding low-resolution patches are extracted from the

filtered images ($E_{f1}Y_X^j$ to $E_{f6}Y_X^j$), at the same location and with the same size. Low-resolution patches are then concatenated into one vector $\tilde{p}_{Y_X}^d$ of length $6m$. Then, the high-resolution dictionary D_h is built up by collecting together all patches $\{p_X^d\}_d, d \in \Phi$, and the low-resolution dictionary \tilde{D}_ℓ is comprised of all patches $\{\tilde{p}_{Y_X}^d\}_d, d \in \Phi$. Finally, the dimensionality of \tilde{D}_ℓ may be reduced to speed up the subsequent computations, given the intrinsic redundancy of the multi-scale edge analysis (as 6 different filters are applied to the same image, resulting in complementary but redundant information of the image edges). For doing so, a Principal Component Analysis (PCA) is applied to this matrix, searching for a set of projection coefficients that represents at least the 90% of the original variance. With the selected coefficients, a projection operator $\mathbf{P} \in \mathbb{R}^{m_\ell \times 6m}$ can be defined as the one that transforms the patch $\tilde{p}_{Y_X}^d \in \mathbb{R}^{6m}$ to its reduced feature vector $p_{Y_X}^d \in \mathbb{R}^{m_\ell}$, that is to say, $p_{Y_X}^d = \mathbf{P}\tilde{p}_{Y_X}^d$. All patches $p_{Y_X}^d$ are collected together to form the reduced low-resolution dictionary D_ℓ , whereby the number of atoms in the dictionary has not changed.

3.4.2 Local reconstruction by sparsity

Once the dictionaries D_ℓ and D_h are constructed, the next stage is related with the estimation of a high-resolution version \hat{X} from a given low-resolution image volume Y . The reconstruction stage involves two steps: first, a local reconstruction is made for each patch, and then, a global image regularization is performed.

First, the low-resolution image needs to be preprocessed to extract the edge features. So, Y is scaled up by a factor of M using \mathbf{H}_M , resulting in Y_X . This upsampled image is multi-scale filtered, as described before (Equation 3-2), leading to 6 images $E_{f1}Y_X$ to $E_{f6}Y_X$.

Then, the local reconstruction step works at the level of patches in the upsampled image, as depicted in Figure 3-2. The filtered images $E_{f1}Y_X$ to $E_{f6}Y_X$ are divided into a grid of regular image patches, and the feature information that corresponds to the same location d is concatenated to form a patch vector $\tilde{p}_{Y_X}^d$. Each patch vector is multiplied by the projection operator \mathbf{P} , for dimensionality reduction, resulting in $p_{Y_X}^d$. Then, a sparse representation for each reduced patch vector $p_{Y_X}^d$ is found by solving

$$\alpha^d = \arg \min_{\alpha} \lambda \|\alpha\|_1 + \frac{1}{2} \|D_\ell \alpha - p_{Y_X}^d\|_2^2 \quad (3-3)$$

where λ balances sparsity of the solution and approximation fidelity (see [16, 44] for additional details). The α^d vector obtained is then multiplied by the D_h dictionary to obtain the reconstructed patch \hat{p}_X^d . Finally, each reconstructed patch is placed in the corresponding location d of the high-resolution image \hat{X}_0 .

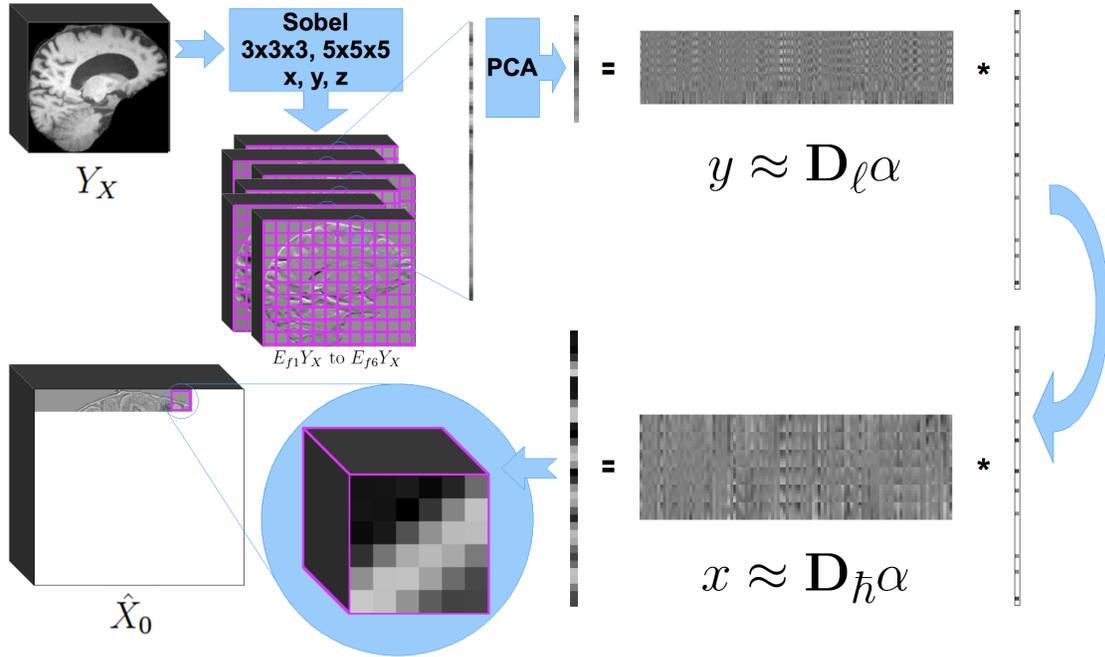


Figure 3-2: Illustration of patch-based local super-resolution reconstruction.

3.4.3 Global regularization by back-projection

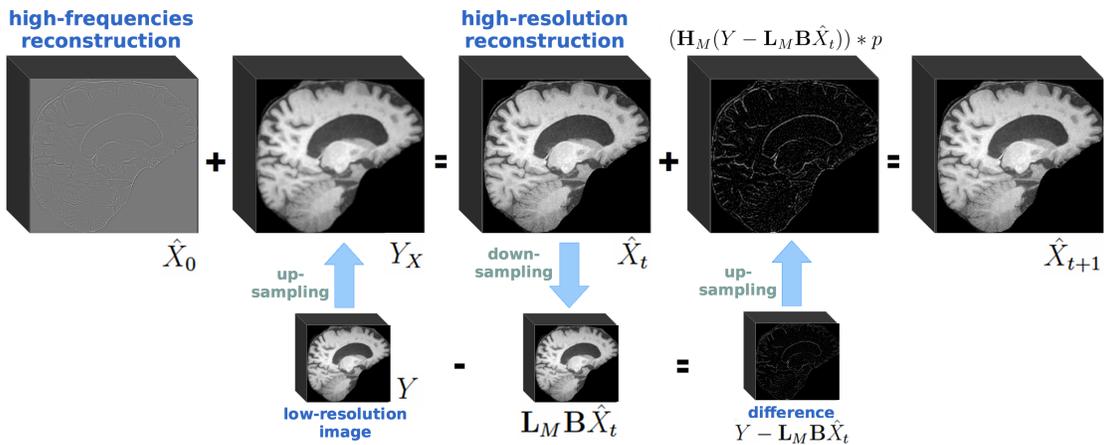


Figure 3-3: Illustration of global regularization.

As the local reconstruction process is locally-oriented and completely independent, no continuity conditions are imposed in the boundaries between patches. The entire high-resolution image \hat{X}_0 (produced by the local sparse representation approach) should thus be further regularized and refined by using the reconstruction constraint (Equation 3-

1). A simple implementation of this idea [65, 133] consists in back-projecting the error as in computerized tomography, through the calculation of

$$X^* = \arg \min_X \|X - X_0\| \quad \text{s.t.} \quad \mathbf{L}_M \mathbf{B} X = Y \quad (3-4)$$

The solution to this optimization problem (illustrated in Figure 3-3) can be roughly described as an iterative calculation of the residual image (difference between the reconstructed and original low-resolution images), which is then convolved with a back-projection kernel, warped back into the super-resolution frame (upsampling) and finally used to update the high-resolution estimated image. This process can be written as

$$\hat{X}_{t+1} = \hat{X}_t + (\mathbf{H}_M(Y - \mathbf{L}_M \mathbf{B} \hat{X}_t)) * p \quad (3-5)$$

where \hat{X}_t is the estimate of the high-resolution image after the t -th iteration, p is the back-projection filter (closely related to the blurring filter B) and $*$ is the convolution operator. This process is iteratively repeated until the Euclidean norm of the difference between consecutive images is less than a given μ . A detailed description and analysis of this scheme and its convergence can be found in [19, 65].

3.5 Experimental Results

To demonstrate the benefits of applying the proposed SR approach to structural 3D brain MR images, we have designed an extensive set of validation experiments, divided in four parts. In the first one, aspects related with the process of dictionary construction were studied and validated. Then, with the appropriate dictionaries, we studied the influence of different parameters in the SR proposed scheme. Once the parameters of our complete approach are tuned, we performed a comparison with recently proposed methods for super-resolution of natural and MR images. Finally, we investigated the influence of the proposed approach in common medical post-processing tasks.

3.5.1 Implementation Details

The image super-resolution algorithm has been implemented in MATLAB R14, running on a Linux PC with 2 Intel Quad Core i7 at 3.07GHz and 24GB of RAM, and using the SparseLab¹ library that provides a set of solvers for the optimization problem. From this library, we have chosen the Basis Pursuit solver, designed to find an approximated solution to the optimization problem of Equation 3-3. Other optimization toolboxes,

¹<http://sparselab.stanford.edu/>

such as SPAMS (SPArse Modeling Software)² could also be used, however, we have experienced reduced computational times (up to 78% less) using the SparseLab toolbox on the same Linux machine.

We have chosen to use the Basis Pursuit solver because it is a kind of swap-down approach, which starts from a sort of full model which is iteratively improved by swapping useless elements with new useful ones, guided by an optimization goal. Matching Pursuit approaches, in contrast, add one element at a time, applying a simple rule repeatedly. In initial experiments, solvers based on Matching Pursuit (MP, OMP, StOMP) did not converge for our approach, while the Basis Pursuit solver converged in few iterations.

The linear filter \mathbf{B} was implemented as the convolution with a Gaussian kernel of size $3 \times 3 \times 3$ and standard deviation 1. In spatial domain, the best way to approximate a continuous signal is the use of a base of sinc functions. However, in practice, it is impossible to obtain an actual sinc function since it requires an infinite support. As the truncated version introduces a ringing pulse in the frequency domain, it is then quite frequent to decently approximate the sinc function with a Gaussian (Normal) distribution or even a triangle, with finite extents and weights greater than or equal to zero. On the other hand, the downsampling operator \mathbf{L}_M implies only taking each M th value starting from zero in each dimension. The Gaussian blurring operator was firstly used, followed by the subsampling operator. The combination of these two processes (in this order) is known as decimation, and it guarantees anti-aliasing after the downsampling. These operators are commonly used in SR approaches to describe the observation model of the image [44, 109, 133], as it accounts for degradation effects and sub-sampling. For upsampling, the \mathbf{H}_M operator was implemented as a bicubic spline interpolator. The main advantages of using splines are their smoothness, their robustness to noise and their approximation accuracy for they are considered as discrete representations of a continuous function. Also, the magnification factor was set to 2, with a patch size of $3 \times 3 \times 3$ in low-resolution and $6 \times 6 \times 6$ in high-resolution.

For all experiments, the λ parameter was set to 0.01. A sensitivity analysis for this parameter showed that values between 0.01 and $50 \times \dim(\text{feature patch})$ in the local reconstruction only generated a variation of about 0.0075 in the root mean squared error and so of 0.1651 dB in the peak signal-to-noise ratio. These variations show the small dependence of the problem on this parameter, so we decided to use 0.01 as had already been reported in the literature [133].

²<http://www.di.ens.fr/willow/SPAMS/>

3.5.2 Brain MR Data Sets

To evaluate the super-resolution algorithm we used different brain MR data sets, including simulated and real images of normal and pathologic subjects.

- **Base data set:** Twenty-nine T1-weighted brain MR images were selected as the evaluation data set. Volumes were acquired in a 3 Tesla General Electric Signa II scanner, at the Alzheimer’s Research Center of Fundación Reina Sofía in Madrid, using a SPGR sequence in sagittal view, with a slice thickness of $1.0mm$, a slice dimension of 512×512 and pixel size of $0.469mm \times 0.469mm$. The number of slices per volume varies between 144 and 168. The data set includes images from control subjects, as well as patients suffering from mild cognitive impairment (MCI) and Alzheimer’s disease.
- **BrainWeb Simulated Brain Database:** Simulated brain MRI data was obtained from the BrainWeb [27] digital brain phantom³, maintained by The McConnell Brain Imaging Centre at Montreal Neurological Institute. The simulation provides volumes acquired in the axial plane with slice dimension of 181×217 and $1mm^2$ resolution. For evaluating with different slice thicknesses, we selected interslice distances of $1mm$ and $3mm$, with a number of slices for each volume of 181 and 60, respectively. For analysis of noise influence, we selected noise percentages of 1%, 3%, 5%, 7% and 9%; while for testing influence of intensity inhomogeneities we selected INU (Intensity Non-Uniformity) percentages of 20% and 40%.
- **MCI data set:** Forty T1-weighted brain MR images comprise this data set used for a statistical study in MCI, with results already reported [3]. Volumes were acquired in a 3 Tesla General Electric Signa II scanner, at the Alzheimer’s Research Center of Fundación Reina Sofía in Madrid, using a FSPGR sequence in axial view, with a slice dimension of 512×512 , a pixel size of $0.469mm \times 0.469mm$, a slice thickness of $1.0mm$ and 158 as the number of slices per volume. The data set includes images from 18 control subjects and 22 patients suffering from MCI.
- **Multicenter data set:** Eight images were collected from two different centers and scanners. At Hospital Internacional Ruber (Center 1) in Madrid, 5 volumes were acquired in a 3 Tesla General Electric Genesis Signa scanner, using a 3D SPGR sequence in axial view, with a slice thickness of $1.0mm$, a slice dimension of 512×512 , a pixel size of $0.469mm \times 0.469mm$ and a varying number of slices (between 98 and 176). At Hospital 12 de Octubre (Center 2) in Madrid,

³available at <http://mouldy.bic.mni.mcgill.ca/brainweb>

3 volumes were acquired in a 1.5 Tesla Philips Achieva scanner, using a SENSE SPGR sequence in axial view, with a slice dimension of 512×512 , a pixel size of $0.357mm \times 0.357mm$, a slice thickness of $1.0mm$ and 150 as the number of slices per volume.

3.5.3 Volume Processing Tools

For the base, BrainWeb and multicenter data sets, we applied the Brain Extraction Tool (BET) [119] to every volume, aiming at leaving only the brain tissues, i.e. removing the skull and skin. After skull stripping and super-resolution reconstruction, we used the FMRIB's Automated Segmentation Tool (FAST) [138] to segment the three main brain tissues (white matter, gray matter, cerebrospinal fluid), and obtain the corresponding binary segmentations.

For the MCI data set, skull stripping, brain tissue segmentation and statistical analysis have been performed using the tools provided in the Statistical Parametrical Mapping (SPM) software toolkit. Then, voxel-based morphometry (VBM) was performed using Diffeomorphic Anatomical Registration Through Exponential Lie Algebra (DARTEL) [4].

3.5.4 Quantitative Performance Measures

To quantitatively evaluate the performance of the reconstruction over the different brain data sets, we used four different metrics for comparison of the reconstructed images with the original ones:

- **Root Mean Square Error (RMSE):** quantifies the pixel intensity differences between the original high-resolution image (A) and its corresponding super-resolution reconstruction (B), using

$$\text{RMSE}(A, B) = \sqrt{\frac{1}{|\Omega|} \sum_{k \in \Omega} |b_k - a_k|^2}$$

where Ω is the brain region, a_k and b_k are the image intensities at position k .

- **Peak Signal-to-Noise Ratio (PSNR):** measures the reconstruction accuracy, expressed in terms of the logarithmic decibel scale. The PSNR was calculated as

$$\text{PSNR}(A, B) = 10 \cdot \log_{10} \left(\frac{\text{MAX}_I^2}{\text{RMSE}(A, B)^2} \right)$$

where MAX_I is the maximum pixel value. Typical values for the PSNR are between 25 dB and 50 dB, where higher is better.

- **Structural Similarity Index (SSIM) [128]:** measures the similarity between two images, in a way which is more consistent with the human visual system and perception. The SSIM was calculated on image windows (size $8 \times 8 \times 8$), where the similarity between windows a and b was measured as

$$\text{SSIM}(a, b) = \frac{(2\mu_a\mu_b + c_1)(2\sigma_{ab} + c_2)}{(\mu_a^2 + \mu_b^2 + c_1)(\sigma_a^2 + \sigma_b^2 + c_2)}$$

where μ_a and μ_b are the mean values of images a and b , σ_a and σ_b are the standard deviation of images a and b , σ_{ab} is the covariance of a and b , $c_1 = (k_1L)^2$ and $c_2 = (k_2L)^2$ (with L being the maximum pixel value, $k_1 = 0.01$ and $k_2 = 0.03$). The resultant SSIM index is a decimal value between -1 and 1, where 1 is only reachable in the case of two identical images.

- **Jaccard similarity index [67]:** measures the overlap (agreement) between two binary images A and B , by taking the ratio between the size of their intersection and the size of their union:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

This metric yields values between 0 and 1, where 0 means complete dissimilarity and 1 stands for identical images.

3.5.5 Tests on Dictionary Construction

The dictionary construction step described in Subsection 3.4.1 requires some analyses in order to verify that the dictionaries D_ℓ and D_h comply with the conditions described in 3.3.1. Also, it is important to explore the influence of important parameters such as the dictionary size and the sampling strategy.

For all experiments in this paper (unless otherwise stated), dictionaries were constructed using images only from the base data set, by sampling 150 random image locations from 28 images, resulting in 29 complete dictionaries of 4200 atomic patches. With this configuration, each dictionary was constructed in about 18 minutes. The multi-scale edge analysis on the low-resolution images leads to a atom size of $6m = 1296$. With this, D_ℓ size is 1296×4200 , while D_h size is 216×4200 .

Atom Correlation - Stability Condition

Correlation of atoms in the dictionaries D_h , \tilde{D}_ℓ and D_ℓ was evaluated by randomly dividing each dictionary in two different groups of atoms and then performing a one-by-one comparison of each atom in one group with all atoms in the other group, using the

Pearson’s correlation coefficient. The correlation coefficient was computed assuming atoms A and B as random variables, and using

$$\rho_{A,B} = \frac{\text{cov}(A, B)}{\sigma_A \sigma_B}$$

The experiment was repeated several times, at least 20 times per dictionary, and correlation values obtained at each repetition were averaged. The results obtained reveals very low correlation values for all dictionaries: D_h (0.00001 ± 0.1524), \tilde{D}_ℓ (0.0007 ± 0.3287) and D_ℓ (-0.00002 ± 0.3849).

Projection Sparsity - Sparsity Condition

The sparse representation approach (Equation B-2) was used to reconstruct a set of randomly sampled patches from high-resolution images with different D_h dictionaries, aiming to verify the sparsity of the projection vector α . In average, the optimization method generates α vectors with a number of nonzero values between 400 and 850, representing the 10% – 20% of the original dictionary atoms.

Reconstruction with Different Dictionaries - Recoverability Condition

First, a small experiment with a single dictionary was performed in order to find the corresponding lower bound in the condition $\|\mathbf{U}D_h^a\|^2 \gg 0$. Given that, for each dictionary atom, $\mathbf{U}D_h^a = D_\ell^a$, and the multi-scale edge information is the one that composes the low-resolution dictionary, we have obtained edge values that range from 0 to 10000 at each dictionary atom. The minimum norm value for these atoms results in about 3000, corresponding to the 30% of the original range, and being about three orders of magnitude greater than zero, which corroborates the recoverability condition of our dictionaries.

Then, the local representation step (Subsection 3.4.2) was applied to super-resolve a set of randomly sampled patches from low-resolution images using several D_ℓ dictionaries, with the objective of verifying that the different dictionaries could produce almost the same reconstruction for a given patch. The different high-resolution reconstructions obtained for a fixed patch were compared using the RMSE, showing in average a variation around 0.4% in the reconstruction error. Figure 3-4 presents an example of the obtained results, showing a slice of the low-resolution patch, the corresponding high-resolution patch, and some SR reconstructions obtained with different dictionaries. Differences between the obtained reconstructions are hardly noticeable, denoting that the dictionary construction process do not bias or deviate the SR results.

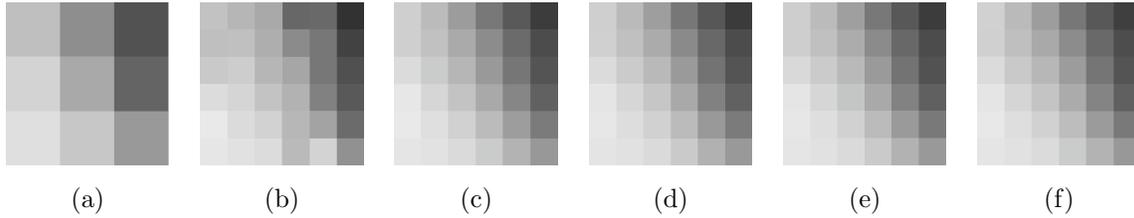


Figure 3-4: Comparison of local super-resolution of a low-resolution patch using several dictionaries. (a) low-resolution patch, (b) original high-resolution patch, (c) - (f) high-resolution reconstructions with different dictionaries.

Dictionary Size

The number of atoms in dictionaries D_ℓ and D_h has impact on two important aspects of the proposed SR method: reconstruction accuracy and reconstruction time. Larger dictionaries include more image patterns, therefore they are entailed with much more information samples, yielding more accurate super-resolved images. Side effects of using a large number of atoms are that both the computational cost of solving the optimization problem and the time involved in patch extraction increase. Here, we evaluate the effect of dictionary size on our SR approach. Two different strategies were tested: number of used images and number of patches sampled per image. First, we fix a number of 150 patches per image, and construct 5 different dictionaries using 1, 2, 5, 10 and 20 different images, obtaining dictionaries of 150, 300, 750, 1500 and 3000 atoms. Then, we fix to 28 the number of used images, and construct 5 different dictionaries by extracting 10, 25, 50, 75 and 100 patches per image, yielding dictionaries of 280, 700, 1400, 2100 and 2800 atoms.

In terms of dictionary construction times, for the first strategy this value ranges from 1 minute for a 150-size dictionary to 18 minutes for a 3000-size dictionary, while for the second strategy the time ranged from 20 minutes for a 280-size dictionary to 24 minutes for a 2800-size dictionary. Building times for the second strategy included the image loading time. This is why times were much less variable in the second strategy when comparing with the first. In terms of image reconstruction times, smaller dictionaries, such as those with 150 and 280 atoms, yielded reconstructions in an average time of 11.59 minutes, while larger dictionaries, such as those with 2800 and 3000 atoms, yielded image reconstructions in an average time of 19.63 minutes. Accuracy reconstruction measurements (RMSE / PSNR / SSIM) have little but possibly significant variations among all dictionary sizes: for smaller dictionaries these values were about 3.95 / 36.18 dB / 0.957, while for larger dictionaries these values were in average 3.82 / 36.49 dB / 0.958.

Atom Sampling Strategy

A main contribution of our proposal with respect to the dictionary construction was the introduction of prior knowledge to determine the image region where the dictionary atoms will be extracted from. As mentioned in Subsection 3.4.1, a previous GM segmentation of the brain volume was morphologically dilated using a 2-voxel-size structuring element, and patches were randomly sampled from this region. Provided that an usual GM segmentation results in a tissue width of about 3 to 4 voxels, the dilation operation transforms this width into 7 to 8 voxels, so that most of the atoms mainly belong to the interfaces (boundaries) between GM-WM and GM-CSF. To investigate the impact of this sampling strategy we have generated SR reconstructions using our segmentation-based dictionaries and also using dictionaries sampled from the whole image, with different sizes (number of atoms per dictionary between 150 and 4200). When the dictionary size was set to 150, a maximum PSNR difference of 0.5 dB was found between our segmentation-based dictionaries and whole image based dictionaries, and as long as the number of patches in the dictionary increased, this difference decreased. In contrast, with dictionaries of 4200 patches, no difference between both approaches was observed.

Dictionary Construction Strategy

We have chosen to construct dictionaries only by a random sampling of image patches, in contrast to the current trend of learning smaller dictionaries from a larger set of sampled atoms, thus obtaining compact and descriptive dictionaries. This can be seen as a sub-optimal choice. However, in our particular case, training dictionaries built up from 3D patches takes a considerable time, given the dimensionality of the patches. To justify this choice, we have downloaded the MATLAB code provided by Yang et al. [133] for dictionary learning⁴, and used it to learn dictionaries for our SR method. With this code, from initial sets of about 10000 3D (vectorized) patches, we have learned compact dictionaries of 512 patches in about 1.2 hours each, in contrast with our random sampling, that takes in average 18 minutes per dictionary. In terms of accuracy, both SR reconstructions, from a learned dictionary and from a sampled dictionary, perform similarly, with very little differences in PSNR (0.21%), RMSE (0.46%) and SSIM (0.02%) values.

⁴<http://www.ifp.illinois.edu/~jyang29/ScSR.htm>

3.5.6 Tests on the Super-Resolution Approach

Once the important parameters for dictionary construction were tested and fixed, the next step is to validate the proposed super-resolution approach. Important aspects to take into account in this validation are related with the influence of image artifacts and features (noise, intensity non-uniformity, slice thickness), and SR parameters such as scaling factor, non-overlapping local reconstruction, Sobel multi-scale analysis and PCA dimensionality reduction.

Noise and Intensity Non-Uniformity (INU) Sensitivity

The proposed super-resolution approach did not take into account image artifacts such as noise or intensity inhomogeneities. Overall, a reconstruction procedure aims to get data the more similar to artifact free data, however the intention in our particular MR application is to reconstruct images with a better resolution than that delivered by a MR scanner, but without exposing the patient to larger acquisition times. We are not specially focused in improving the reconstructed image by removing noise and/or correcting intensity non-uniformities. However, it is important to identify the impact of such factors in the final reconstruction result. To do so, simulated brain MR images with different noise and INU percentages were selected from the BrainWeb database, and corresponding low-resolution versions of each were constructed by blurring and downsampling with a factor of 2. Then, low-resolution noisy images were denoised using MNLM3D [30] and super-resolved with our approach, a nearest-neighbor interpolator, a bicubic interpolator and the non-local approach of Manjón [88] (without the denoising step), and finally compared with the original image with 0% noise and 0% INU. Tables 3-1 and 3-2 presents the obtained accuracy reconstruction values in terms of RMSE, PSNR and SSIM. These tables shows that our approach performs better in all cases than the nearest-neighbor and bicubic interpolation and also than the non-local approach [88].

Influence of Slice Thickness

As mentioned in Section 3.1, the partial volume (PV) effect depends directly on the spatial resolution of the acquisition. In particular, the PV effect is stronger as the inter-slice distance increases. To study the effect of slice thickness on the proposed super-resolution algorithm, two simulated brain MR images with different inter-slice distances were selected from the BrainWeb database. The original size of the volumes (see Subsection 3.5.2) was modified to odd numbers ($180 \times 216 \times 180$ for $1mm$ thickness, $180 \times 216 \times 60$ for $3mm$ thickness) to facilitate downsampling and further super-resolution reconstruction. Again, the low-resolution versions of these images were

		% noise	0%	1%	3%	5%	7%	9%
Nearest-neighbor interpolation	RMSE		17.12	17.36	18.89	19.63	20.29	22.23
	PSNR (dB)		23.40	23.18	21.90	21.18	20.75	19.61
	SSIM		0.895	0.803	0.635	0.562	0.521	0.490
Bicubic interpolation	RMSE		14.19	14.64	16.83	18.13	19.11	21.33
	PSNR (dB)		25.03	24.66	22.90	21.87	21.26	19.97
	SSIM		0.922	0.829	0.657	0.581	0.538	0.505
Non-local approach	RMSE		13.78	14.17	16.18	17.53	18.57	20.85
	PSNR (dB)		25.28	24.94	23.24	22.17	21.51	20.17
	SSIM		0.932	0.839	0.668	0.592	0.548	0.514
Proposed approach	RMSE		11.44	9.55	10.20	10.66	11.37	13.11
	PSNR (dB)		26.89	28.37	27.25	26.48	25.78	24.20
	SSIM		0.977	0.888	0.709	0.632	0.589	0.560

Table 3-1: Accuracy super-resolution values under influence of noise.

constructed by blurring and downsampling by a factor of 2 each volume, resulting in a slice dimension of 90×108 pixels, and number of slices per volume of 90 ($1mm$) and 30 ($3mm$). Figure 3-5 presents the super-resolution reconstructions obtained for the different slice thicknesses, compared with the corresponding interpolated images. RMSE / PSNR values obtained for the bicubic interpolation were 13.7 / 25.4 dB in $1mm$ and 13.5 / 25.5 dB in $3mm$, which were improved by the super-resolution reconstruction with values of 5.6 / 33.2 dB in $1mm$ and 10.9 / 27.4 dB in $3mm$.

Effect of the Sobel Operator

We have tested the influence of selecting the Sobel operator as feature detector, by comparing the results obtained with the local reconstruction step (Equation 3-3) over the 29 brain MR images of the base data set and using three different edge extraction operators: the first- and second-order derivatives proposed by Yang et al. [133], the 3×3 and 5×5 Sobel kernels and the 3×3 and 5×5 Prewitt kernels. Differences between the three approaches are imperceptible for the human eye, as can be noticed in Figure 3-6, but in average, the Sobel operator reduces the RMSE in about 30%, when compared with an standard bicubic interpolation, while the Prewitt operator and the proposal of Yang et al. [133] only reduces the error in about 26%.

Effect of dimensionality reduction

To test the influence of the dimensionality reduction step in the reconstruction results, different amounts of PCA coefficients were selected: $m_l = 12, 51, 116, 193$, which cor-

		% INU	0%	20%	40%
Nearest-neighbor interpolation	RMSE		17.12	19.71	25.04
	PSNR (dB)		23.40	21.67	19.54
	SSIM		0.895	0.887	0.870
Bicubic interpolation	RMSE		14.19	17.47	23.48
	PSNR (dB)		25.03	22.71	20.10
	SSIM		0.923	0.913	0.898
Non-local approach	RMSE		13.78	16.89	22.92
	PSNR (dB)		25.28	23.01	20.31
	SSIM		0.932	0.924	0.906
Proposed approach	RMSE		11.44	14.99	22.40
	PSNR (dB)		26.89	24.04	20.51
	SSIM		0.977	0.969	0.948

Table **3-2**: Accuracy super-resolution values under influence of intensity non-uniformities.

responds to 1%, 4%, 9% and 15% of the original atom size (1296), respectively. An analysis of the coefficients obtained after application of PCA, on the different constructed dictionaries, reveals that selecting the first 12 coefficients explains 90% of the original variance, 36 coefficients explains 99% of the original variance, and 100% is approximately explained by selecting the first 162 coefficients. The reconstruction algorithm was then tested using a conventional leave-one-out on the base data set, leading to 29 different experiments, where the dictionary construction process was performed using 28 images and the reconstruction was performed on the remaining one. Low-resolution versions of these images were constructed by blurring and downsampling by a factor of 2 each high-resolution image. Each leave-one-out experiment was then executed using the different percentages of PCA coefficients, and the obtained results were compared with a version of the method where no dimensionality reduction was performed before the super-resolution reconstruction.

Panels (a), (b), (c) and (d) in Figure **3-7** show the evolution in average of each metric (RMSE, PSNR, SSIM and execution time, respectively). As the charts show, the larger the number of coefficients the smaller the achieved accuracy (up to a limit, see Appendix C), even though these differences can be considered as non significant given the small variations (0.05 in RMSE, 0.1 dB in PSNR, 0.0005 in SSIM, calculated between 1% and 15% of PCA coefficients). However, the main difference appears in the reconstruction time, that ranges between 26 minutes for 1% to 92 minutes for 15%. Compared with a reconstruction without the dimensionality reduction step, the PCA analysis provides clear benefits, by reducing the execution time to nearly 3.7% of the

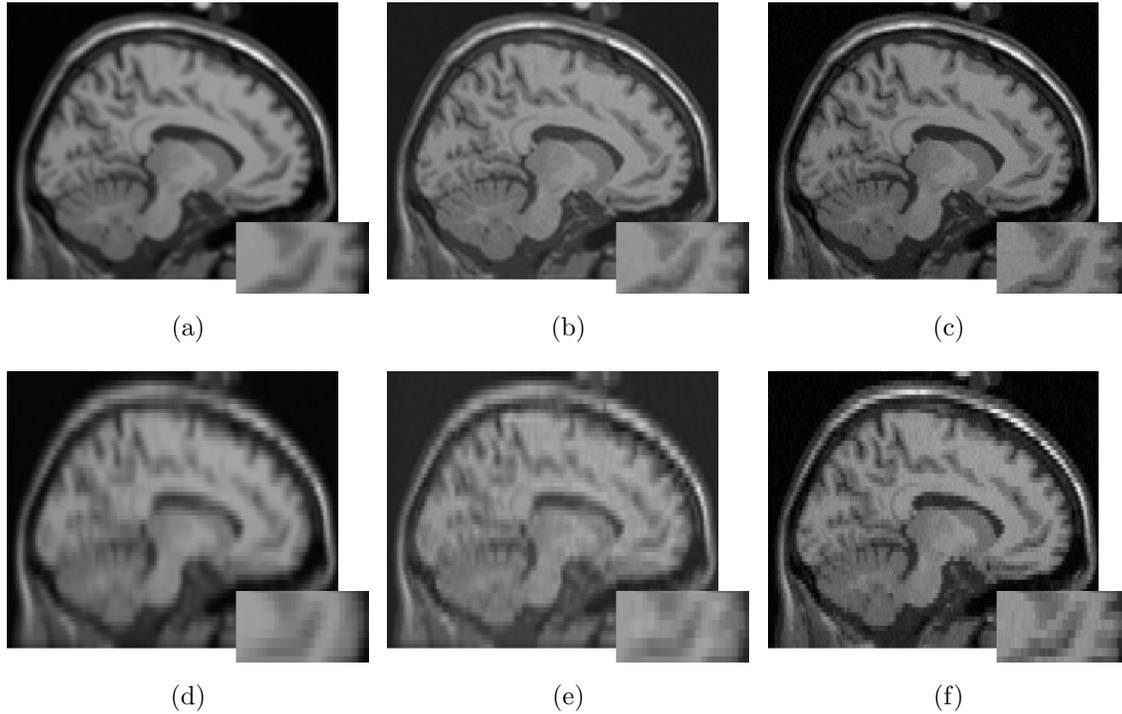


Figure 3-5: Results of the reconstruction with 1mm (*top*) and 3mm (*bottom*) slice thicknesses. *Column 1* (a),(d): Bicubic interpolation of low-resolution image. *Column 2* (b),(e): Super-resolution reconstruction. *Column 3* (c),(f): Original high-resolution image.

original time and the reconstruction error in about a 4%.

Influence of Non-Overlapping Local Reconstruction

An important contribution of the proposed approach is the use of non-overlapping patches when solving the optimization problem per patch, whereby the scheme becomes fully-parallelizable. The introduction of this strategy allows to process the entire volume in a reduced time. To evaluate the real impact of this selection, we have compared our proposed approach with a modified version which uses a 1-voxel overlap per patch (resembling [133]). In overlapping regions, multiple reconstruction values are just averaged to obtain the final reconstruction. All images in the base data set were super-resolved using both strategies, and reconstruction time and accuracy were recorded and compared. RMSE / PSNR / SSIM values for the non-overlapping approach were in average 4.17 / 36 dB / 0.967, and slightly better values were obtained using the overlapping approach: 4.14 / 36.06 dB / 0.967. However, the main difference is observed in the reconstruction time, which was about 26 minutes for the

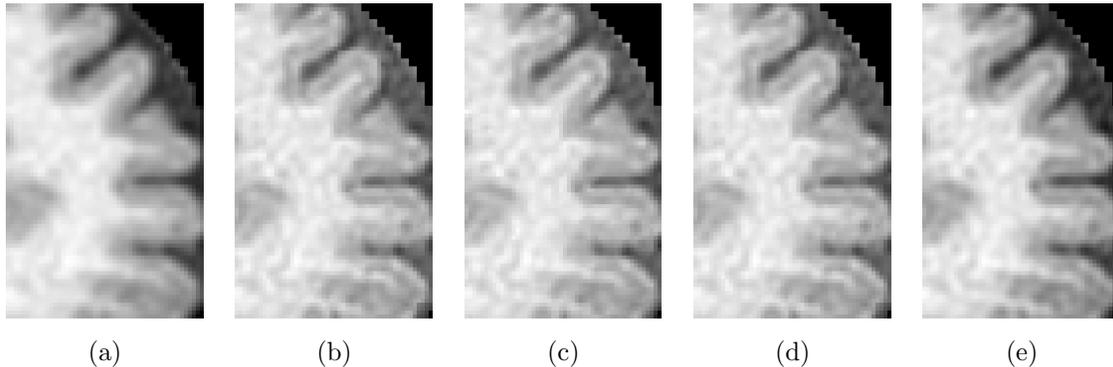


Figure **3-6**: Comparison of the local high-resolution reconstruction using different edge extraction operators. Detailed views of (a) bicubic interpolation of low-resolution image, (b) local reconstruction with the proposal of [133], (c) local reconstruction with Sobel operator, (d) local reconstruction with Prewitt operator, and (e) original high-resolution image. The differences between the three proposals result imperceptible for the human eye.

non-overlapping approach and 75 minutes in average for the overlapping. Given this result, we can state that our non-overlapping approach reconstructs image versions as accurate as those obtained with the overlapping strategy, in a considerably smaller interval of time, and with the side effect that the reconstruction is fully-parallelizable in this case.

Influence of the Scaling Factor

In all experiments, the scaling factor was set to 2. However, it could be interesting to test the effect of using a larger factor, given that it implies that more high-resolution complex patterns are associated to a very simple low-resolution patch. To do so, we have set the scaling factor to 4, and with this we constructed low-resolution versions of the base data set high-resolution images, resulting in an in-plane resolution of 128×128 and a number of slices ranging from 36 to 42. Afterwards, we have constructed new dictionaries using this low- and high-resolution images, following the same procedure (random sampling of 150 patches per image, using 28 images at a time, obtaining dictionaries of 4200 atoms) and the SR approach have been applied to the low-resolution images in a leave-one-out fashion. Figure **3-8** presents an example of the reconstruction using a scaling factor of 4, and the comparison with a bicubic interpolation and the original high-resolution image. Using a bicubic interpolation on the low-resolution images, the accuracy reconstruction values (RMSE / PSNR / SSIM) obtained were in average 16.71 / 24.13 dB / 0.879, while with our approach we achieved average values



Figure 3-7: Comparison of super-resolution reconstruction with different amounts of PCA coefficients (1%, 4%, 9% and 15%) and without dimensionality reduction (noPCA), in terms of (a) RMSE, (b) PSNR (in dB), (c) SSIM, and (d) reconstruction time (in hours).

of 10.93 / 27.93 dB / 0.887. This represents a reduction in the reconstruction error of 34.6%, and increments in the PSNR of 15.8% and in the SSIM of 0.91%.

Relation between Low-resolution and High-resolution Acquisition

In a different SR experiment, images from the same subject, with different resolutions ($256 \times 256 \times 156$ with voxel size $0.94mm \times 0.94mm \times 1mm$ and $512 \times 512 \times 156$ with voxel

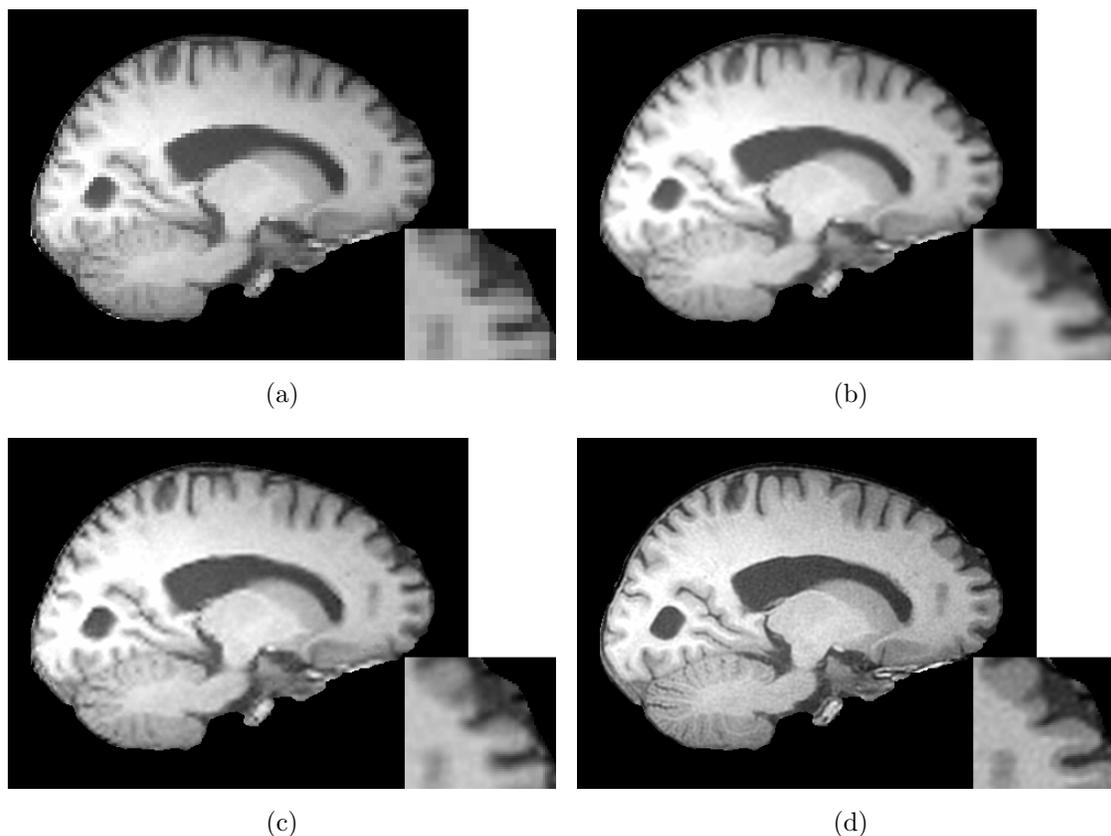


Figure 3-8: Comparison of super-resolution of a low-resolution image using a scaling factor of 4. (a) low-resolution image, (b) bicubic interpolation of low-resolution image, (c) high-resolution reconstruction, and (d) original high-resolution image.

size $0.47mm \times 0.47mm \times 1mm$) were acquired in the same scan session, with the aim of comparing the super-resolved image (reconstructed from the low-resolution image) with the corresponding high-resolution image produced by the MR equipment. When the resolution is increased in MR, the size of the pixel is reduced, so that the intensity of the image is also reduced. On the other hand, as long as the detail increases, the tissues show a different texture (these effects are shown in Figure 3-9). A quantitative comparison between both images is thus not direct. Panel 3.9(c) shows a slice of the reconstructed image after our algorithm is applied (using a dictionary constructed with images of the base data set), and Panel 3.9(d) the corresponding slice of the high-resolution image. As can be seen, both images are visually very similar in terms of the brain structures, although the tissue intensities and textures are not exactly the same. To perform a real quantitative evaluation, a previously downsampled version of the

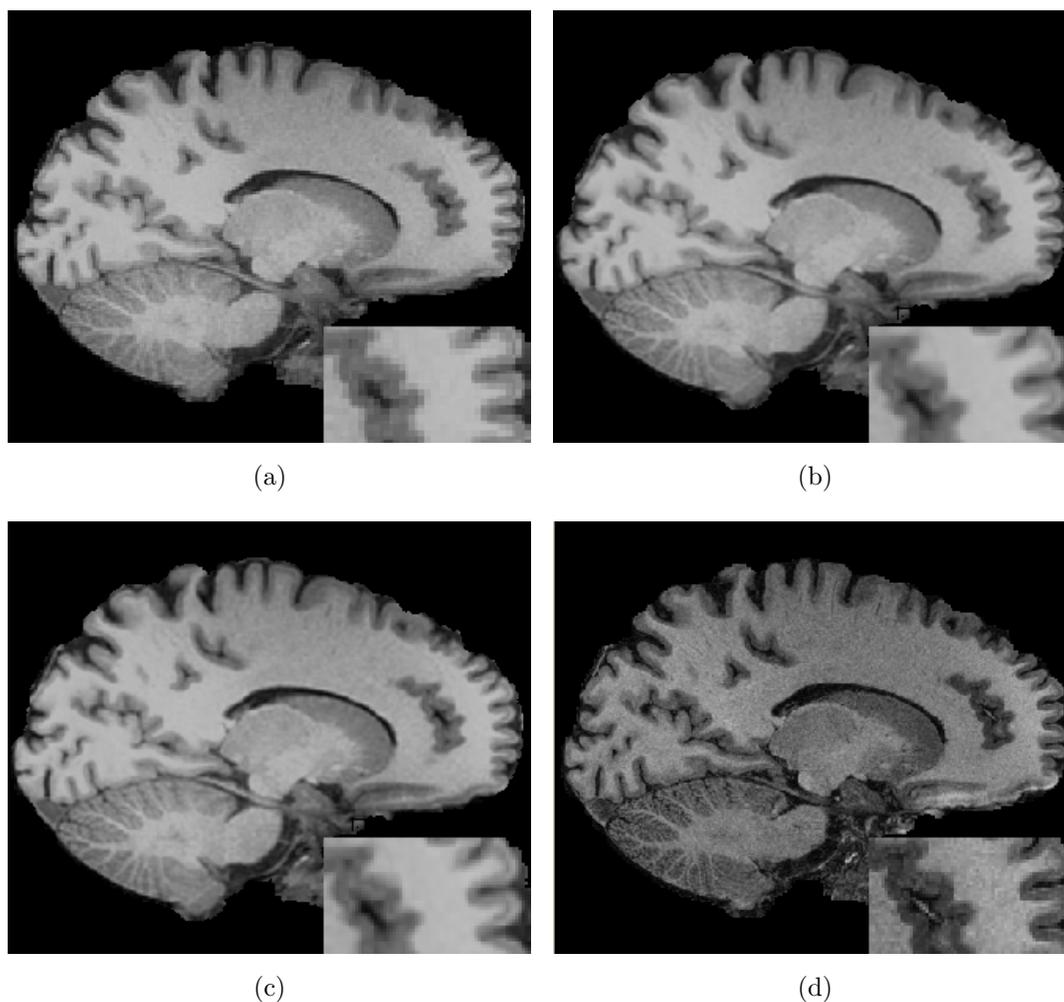


Figure 3-9: Comparison of super-resolution of a low-resolution image with the corresponding high-resolution image acquired from the same subject in the same scan session. (a) low-resolution image, (b) bicubic interpolation of low-resolution image, (c) high-resolution reconstruction, and (d) original high-resolution image.

high-resolution image was super-resolved by applying the proposed methodology (again with a dictionary constructed with images of the base data set) and also by using a standard bicubic interpolation. The obtained results are presented in Figure 3-10. As can be visually noted, the proposed method was able to better reconstruct the downsampled volume, producing better defined boundaries between tissues. Using the RMSE and the PSNR, these results were compared with the original high-resolution image. The proposed method achieved a RMSE of 4.46 and a PSNR of 35.15 dB, while

the bicubic interpolation results in a RMSE of 11.95 and a PSNR of 26.58 dB. This represents an increment of about 9 dB and a reduction of 63% in the reconstruction error by applying our super-resolution algorithm.

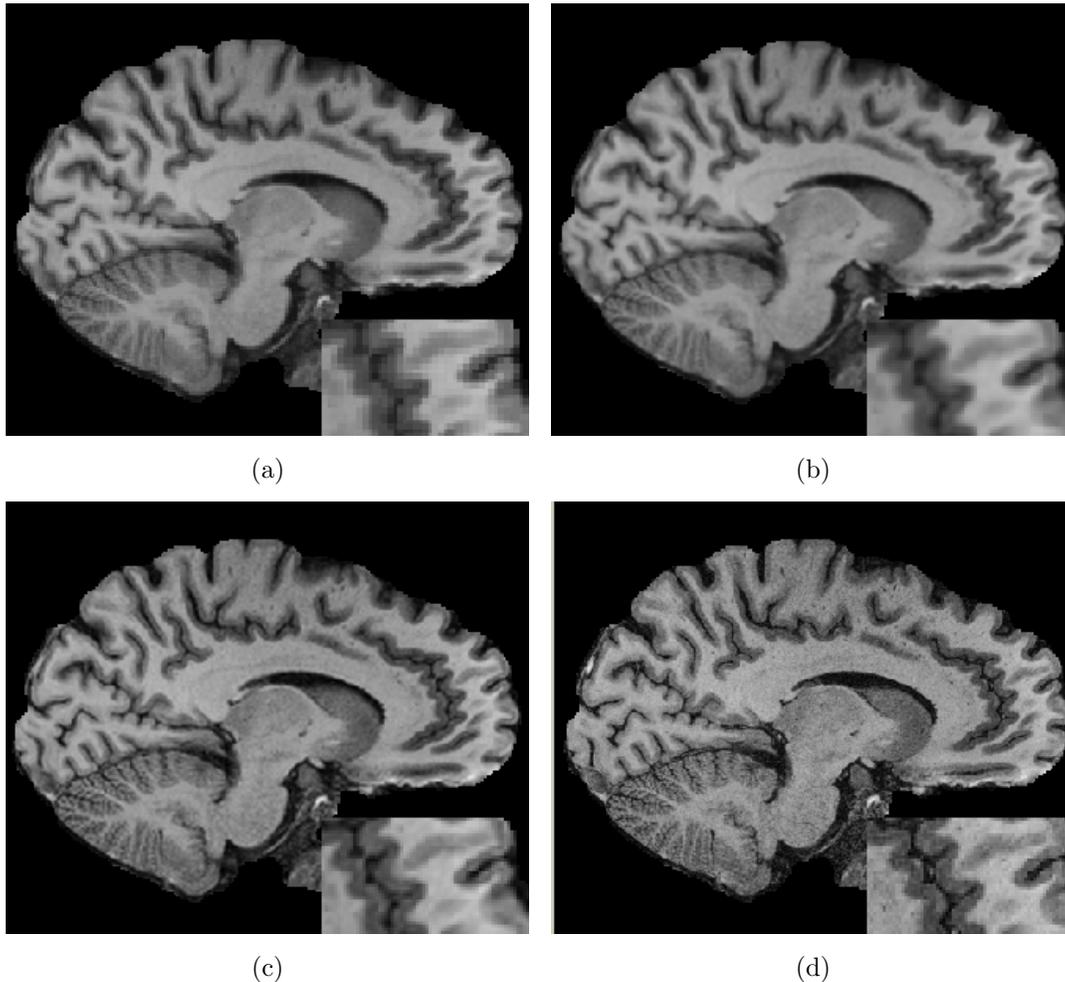


Figure 3-10: Comparison of super-resolution of a previously downsampled low-resolution image with the corresponding high-resolution image. (a) low-resolution image, (b) bicubic interpolation of low-resolution image, (c) high-resolution reconstruction, and (d) original high-resolution image.

3.5.7 Comparison with Other Approaches

Once our super-resolution approach was tuned, the next step was to compare its performance with other interpolation and SR approaches, to clearly identify the advantages

and improvements introduced. These other methods includes: the 2D SR method proposed by Yang et al. [133], a 3D standard bicubic interpolation and the 3D non-local approach of Manjón et al. [88].

Comparison with 2D Super-Resolution [133]

The proposed SR approach has some elements in common with the method presented by Yang et al. [133], although their method has been implemented so far on 2D images. To demonstrate the impact of the introduced modifications, we have prepared a version of our approach which super-resolves 2D images, in order to compare against [133], whose MATLAB code is available on the Web⁵. In this case, one slice per volume in the base data set was selected as the high-resolution images, and the corresponding low-resolutions versions were obtained after blurring and downsampling by a factor of 2 each slice. Afterwards, 29 different dictionaries were constructed, in our case by randomly sampling 150 patches per image with 28 images at a time to obtain dictionaries of 4200 2D patches, and in Yang et al. [133] approach by selecting initial sets of 100000 patches from 28 images at a time and then learning compact dictionaries of 512 atoms. Finally, the SR approaches were applied to the 29 low-resolution versions to generate high-resolution approximations of the slices, and reconstruction results were compared in terms of accuracy and computational time. In average, the approach by Yang et al. [133] yielded RMSE / PSNR / SSIM values of 10.12 / 28.24 dB / 0.922, while our approach obtained more accurate values: 7.46 / 31.12 dB / 0.930. Figure 3-11 presents one of the reconstructed slices with both methods, and the original high-resolution image as reference. In terms of reconstruction time, in average our approach takes 151 seconds per 2D image, while [133] takes 72 seconds per image.

Comparison with Standard Bicubic Interpolation

Figure 3-12 presents one slice of the obtained results after applying the proposed method (using 1% of PCA coefficients) to one brain MR volume of the base data set, and also the corresponding comparison with a standard bicubic interpolation and the original image. In the low-resolution image (panel (a)) a series of small blocky artifacts can be observed in the boundary between the gray matter and cerebrospinal fluid, specially in the parietal and occipital lobes. These artifacts are not longer present in the high-resolution reconstruction (panel (b)), indicating that the super-resolution method yields sharper edges. Detailed views in panels (c) to (f) allow a better visual comparison and show that boundaries between tissues, which are somehow blocky in the low-resolution and fuzzy in the bicubic interpolation, have been sharply recovered.

⁵<http://www.ifp.illinois.edu/jyang29/ScSR.htm>

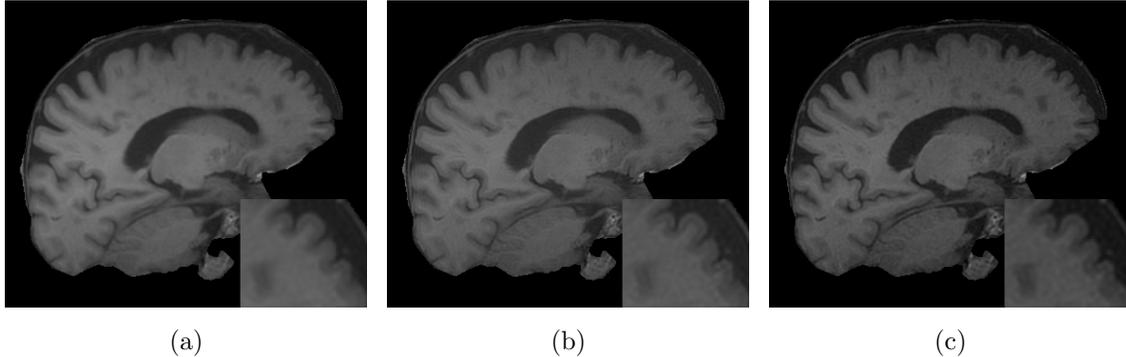


Figure 3-11: Comparison of super-resolution results obtained after applying our proposal and the method in [133]. *Left*: reconstruction produced by [133]. *Middle*: reconstruction produced by our algorithm. *Right*: original high-resolution image.

Differences between the reconstructed and the original high-resolution images are not visually noticeable, showing that the proposed method does not introduce noise or other artifacts in the reconstruction.

Each reconstruction and its corresponding interpolated version were compared with the original high-resolution image using the RMSE, the PSNR and the SSIM. Applying a standard bicubic interpolation technique, an average PSNR of 25.71 dB was obtained, together with an average RMSE value of 14.01 and an average SSIM value of 0.94. The proposed super-resolution algorithm achieved an average PSNR of 36.00 dB, an average RMSE of 4.17 and an average SSIM of 0.97, representing an increment of about 10.3 dB, an increment of 2.4% in the similarity and a reduction of 70% ($p < 0.001$) in the reconstruction error with respect to the interpolated images.

Comparison with a Non-local Approach [88]

Recently, Manjón et al. [88] have proposed a new upsampling method that recovers some of the missing high-frequency information in brain MR images by using an iterative scheme that combines a data-adaptive patch-based reconstruction with a subsampling coherence constraint. This approach, based on an adaptation of the non-local means algorithm [17, 31], is shown to outperform classical interpolation methods (nearest neighbor, trilinear, cubic and B-spline interpolation) using synthetic (T1w BrainWeb digital brain phantom⁶ [27]) and real brain MR data (T2w low-resolution and high-resolution images). As the MATLAB source code is freely available on the

⁶available at <http://mouldy.bic.mni.mcgill.ca/brainweb>

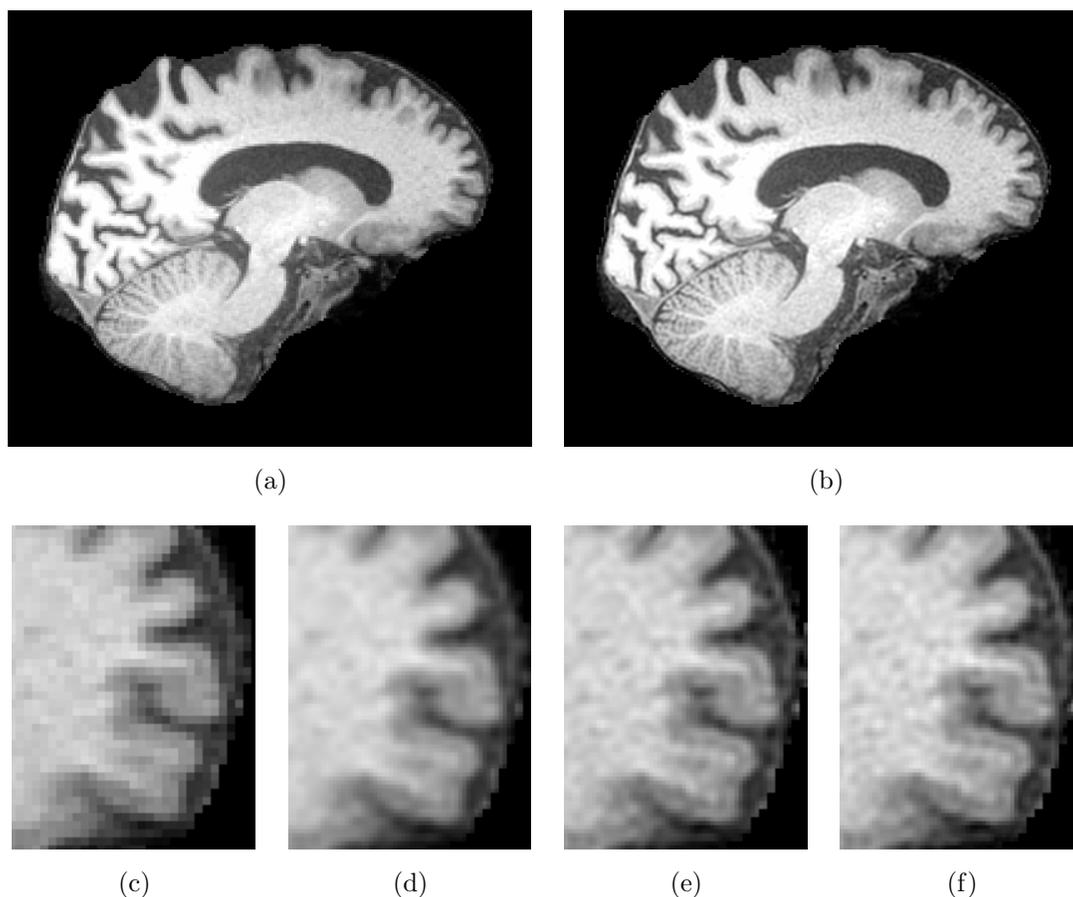


Figure 3-12: Reconstruction with 4200 dictionary patches from multiple MR brain images. *Top row*: (a) low-resolution image, (b) reconstruction by super-resolution. *Bottom row*: Detailed views of (c) low-resolution image, (d) bicubic interpolation of low-resolution image, (e) reconstruction by super-resolution, and (f) original high-resolution image. Blocky artifacts in low-resolution image and smoothness in the interpolation have been eliminated in the high-resolution reconstruction, obtaining sharper edges as the ones present in the original image.

Web⁷, we compared the results obtained with our method to those obtained with the method by Manjón et al. [88].

Manjón et al. [88] proposal includes two main steps: an initial image denoising using MNLM3D [30], followed by the super-resolution reconstruction. As our approach does not include a denoising step, two different configurations for both methods were tested:

⁷<http://personales.upv.es/jmanjon/reconstruction/upsampling.htm>

first, denoising was disabled in the Manjón et al. [88] approach, allowing to compare the sparse-based approximation versus the nonlocal-based one; and then, the denoising step was performed for both methods, to identify the impact of this preprocessing in the reconstruction results.

The 29 images in the base data set (without any preprocessing) were then reconstructed using both algorithms: our proposal was performed with 1% of PCA coefficients, while the Manjón et al. [88] approach was performed without the denoising step; on a Linux PC with 2 Intel Quad Core i7 at 3.07GHz and 24GB of RAM, and compared in terms of quantitative measures (RMSE, PSNR, SSIM) and execution time. While our proposal, in average, achieved RMSE / PSNR / SSIM values of 4.17 / 36 dB / 0.967, the non-local method obtained an average score of 14.81 / 26.12 dB / 0.945, indicating that our strategy produces more accurate images with sharper edges. This fact can also be verified by visually comparing the results obtained by both methods (examples shown in Figure 3-13). In terms of processing time, our approach takes in average 0.44 hours (26 minutes) to yield a high-resolution reconstruction, while the non-local technique takes 0.72 hours (43 minutes).

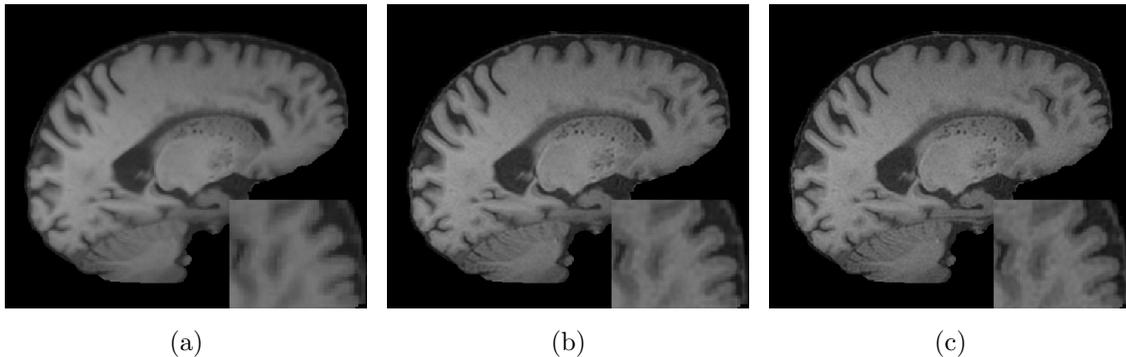


Figure 3-13: Comparison of super-resolution results obtained after applying our proposal and the method in [88]. *Left*: reconstruction produced by the non-local algorithm [88]. *Middle*: reconstruction produced by our algorithm. *Right*: original high-resolution image.

After this, the 29 images were then preprocessed using the MNLM3D [30] denoising algorithm, and these images were then reconstructed using both algorithms, in the same conditions as before. With the denoised images, our approach achieved in average RMSE / PSNR / SSIM values of 8.25 / 30.3 dB / 0.952, while the non-local method reaches in average values of 15.39 / 25.64 dB / 0.938. In this case, our proposal also produces more accurate reconstructions. Figure 3-14 presents an example of the results on denoised images.

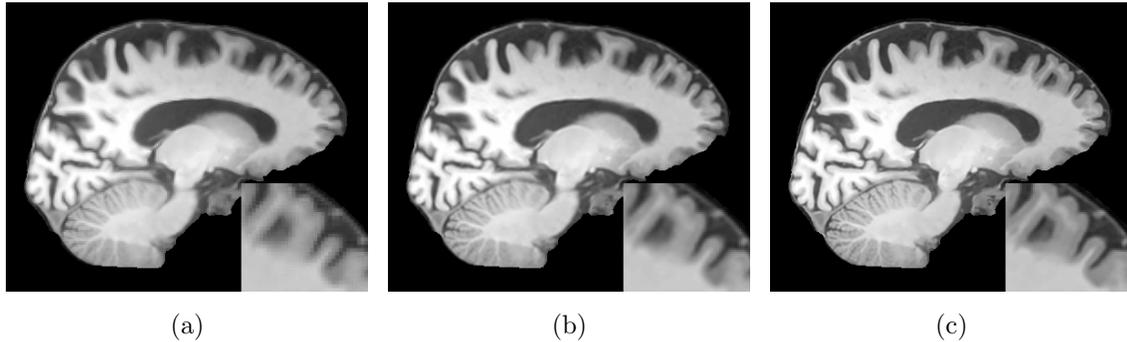


Figure 3-14: Comparison of super-resolution results obtained on denoised images with our proposal and the method in [88]. *Left*: reconstruction produced by the non-local algorithm [88]. *Middle*: reconstruction produced by our algorithm. *Right*: original high-resolution denoised image.

3.5.8 Impact on Post-processing Tasks

To place the proposed SR approach in a medical application context, we have tested its influence in different medical post-processing tasks, such as segmentation, voxel-based morphometry and multicentric studies.

Impact on Brain Tissue Segmentation

Figure 3-15 illustrates the improvements introduced by the super-resolution method on brain MR gray matter segmentations, with enlarged views of the low-resolution MR image and its corresponding high-resolution reconstruction (GM segmentations of images in Figure 3-12). An ellipse highlights the cortical folds that are misdeteched in the low-resolution version of the image, while in the high-resolution reconstruction these folds are better delineated. In this particular case, the reconstruction outperforms the original image in the segmentation task, by detecting completely a convolution (panel (b)) that is segmented as fractionated in the original image (panel (c)).

Compared with other approaches, such as the bicubic interpolation and the proposal of Manjón et al. [88], our method allows to obtain more accurate GM segmentations, as illustrated in Figure 3-16. In terms of the Jaccard index, gray matter segmentation of reconstructed images in average shows larger overlapping (0.85) with the original image, compared with the interpolated image (0.75) and the result obtained with the nonlocal-based reconstruction (0.71), representing an increment of about 9% ($p < 0.001$).

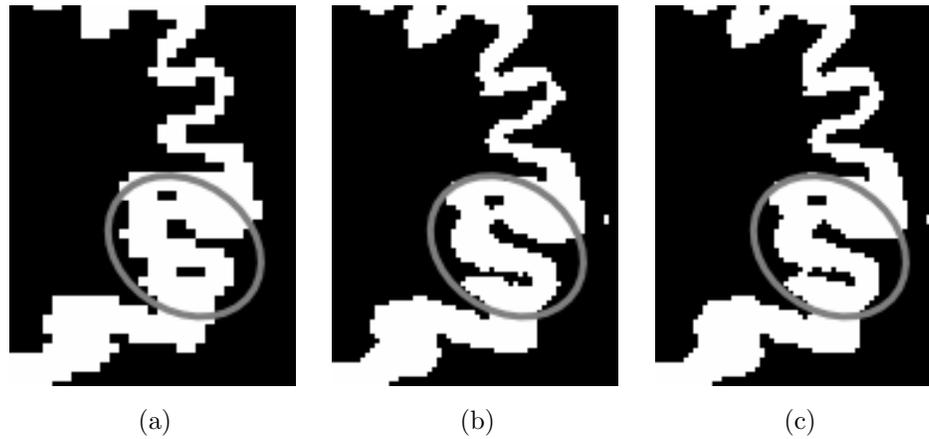


Figure 3-15: Improvement of gray matter segmentation using the proposed method. Detailed views of (a) low-resolution image, (b) high-resolution reconstruction, and (c) original high-resolution image. Ellipses highlight the delineation of cortical folds, which are accurately segmented in the high-resolution reconstructed image.

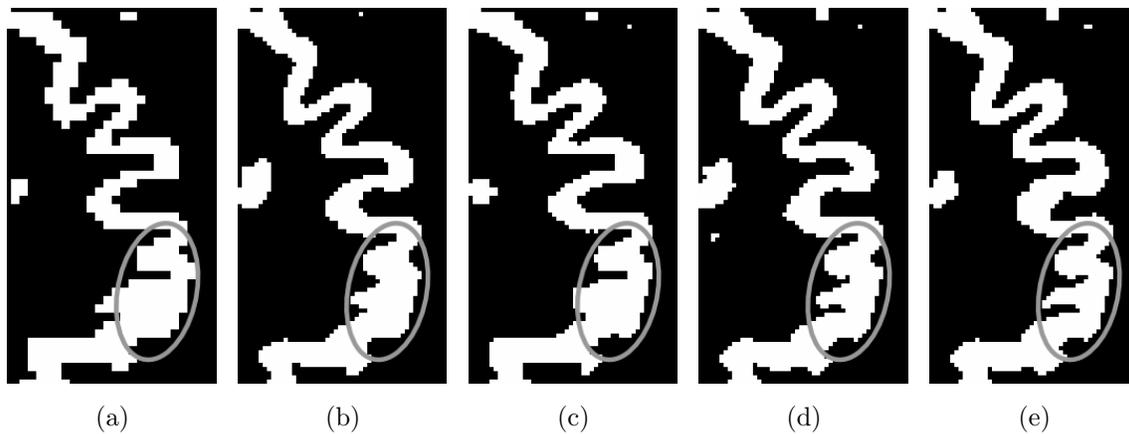


Figure 3-16: Comparison of improvement in gray matter segmentation. Detailed views of (a) low-resolution image, (b) bicubic interpolation of low-resolution image, (c) nonlocal-based reconstruction [88], (d) high-resolution reconstruction [88], and (e) original high-resolution image. Ellipses in the segmentation highlight the accurate delineation of cortical folds in the high-resolution reconstructed image.

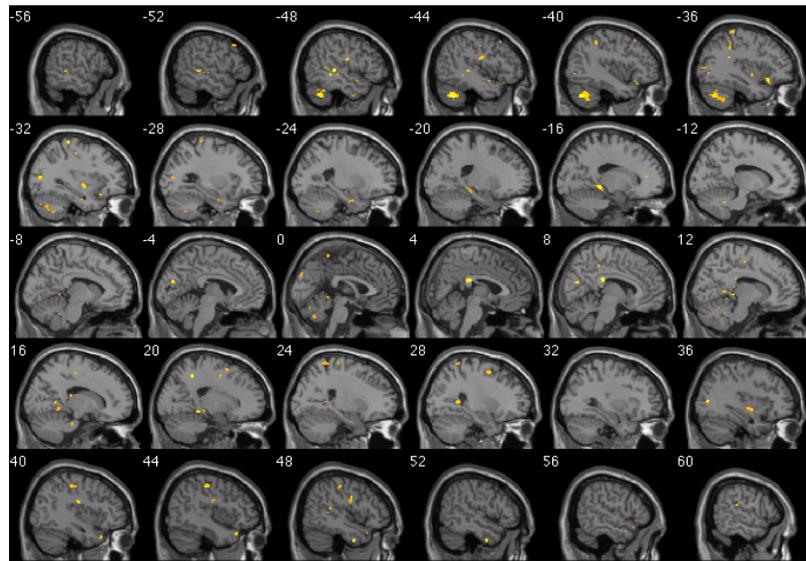
Results on a real study of Mild Cognitive Impairment

To evaluate the effect of the proposed approach on morphometric analysis of real clinical data, we collected a set of 40 high-resolution images, acquired in the axial direction, which includes 8 normal controls and 22 patients suffering different degrees of MCI (4 amnesic, 9 non-amnesic and 9 multi-domain). A previous morphometrical study on this data set [3] (using VBM and DARTEL) has revealed volume reductions in the parahippocampal gyrus, lingual gyrus and cerebellum for amnesic MCI patients; in the posterior cingulate gyrus for the non-amnesic MCI group, and in the posterior cingulate and parahippocampal gyrus in multi-domain MCI. We wanted to reproduce these results by applying VBM on reconstructed high-resolution volumes, obtained after applying the proposed method to low-resolution versions (obtained by blurring and downsampling, volume size: $256 \times 256 \times 79$) of the original images.

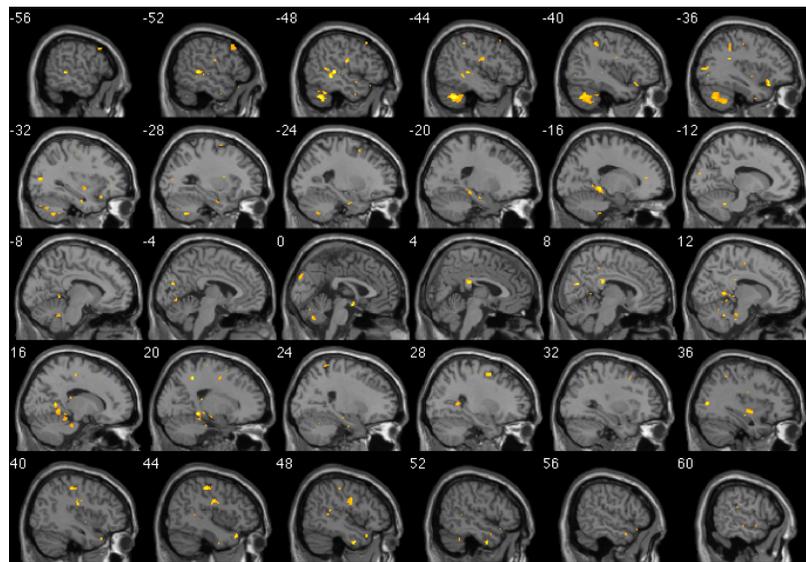
Panel (a) in Figure 3-17 presents slices of the statistical map obtained after VBM analysis of the original data set ($p < 0.01$, comparison of control subjects versus MCI patients), while panel (b) presents the statistical map obtained by using the reconstructed high-resolution images for VBM analysis. A visual comparison of the regions with significative differences in both statistical maps, reveals that the VBM analysis on the reconstructed images produces similar results than VBM applied to the high-resolution original images.

Influence on Multicentric Studies

To demonstrate the capability of super-resolving images from other centers (different scanners) using the previously constructed dictionaries, we have collected a set of images from two different centers and scanners, as described in Subsection 3.5.2. Low-resolution versions of these images were constructed by blurring and downsampling each high-resolution volume. D_ℓ and D_h dictionaries, constructed using images from the base data set, were selected for super-resolving the low-resolution images, and the results were compared with a bicubic interpolation and the original images through the accuracy measurements (RMSE / PSNR / SSIM). For images from Center 1 the bicubic interpolation reaches average scores of 7.03 / 31.3 dB / 0.977, while our proposal largely outperforms these results: 2.20 / 41.8 dB / 0.990. A similar behavior has been observed with images from Center 2, where the interpolation obtains in average 3.15 / 38.4 dB / 0.978 and our approach reaches 1.51 / 44.8 dB / 0.984. Figure 3-18 shows two examples of the obtained results, each image coming from a different center.



(a)



(b)

Figure 3-17: Statistical map ($p < 0.01$) of comparison between control subjects and MCI patients, obtained using (a) the original and (b) the reconstructed high-resolution images of MCI data set.

3.6 Discussion

We have presented a powerful technique which reconstructs high-resolution brain MR images from low-resolution images using a sparse representation. Basically, the missing

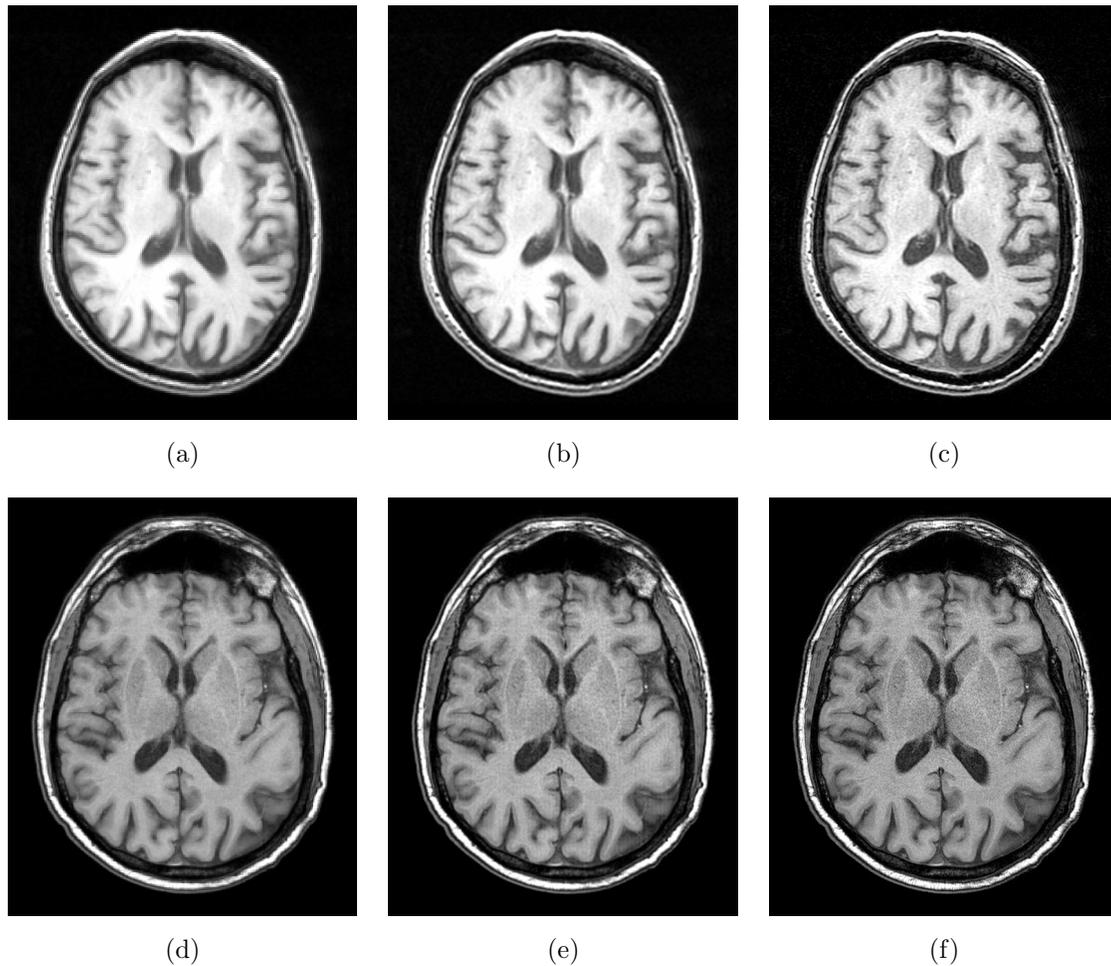


Figure 3-18: Reconstruction of images from different centers (different scanners and acquisition parameters). Top: Example from Center 1, Bottom: Example from Center 2. (a), (d) low-resolution image; (b), (e) high-resolution reconstruction; (c), (f) original high-resolution image.

edge information was inferred from a multi-scale edge analysis and used to reconstruct a high-resolution version by including prior knowledge from high-resolution images. A sparse representation can be thought of as a decomposition of the input image as a linear combination of statistically semi-independent image patches, which make up a dictionary. Low- and high-resolution dictionaries for the reconstruction were built up by randomly sampling patches from low- and high-resolution images, which were previously cropped to leave only the three main brain tissues (WM, GM and CSF). The introduction of the PCA, as a dimensionality reduction technique, allows to project the data in a new coordinate system which better captures relevant information, whereby

only few dimensions are enough. The combination of a multi-scale analysis together with the semantic-based dictionaries and the dimensionality reduction scheme led to a useful technique that improves the quality of reconstruction in about 10 dB, compared with interpolation approaches.

The accomplished dimensionality reduction when applying a technique such as PCA in the compression of the low-resolution dictionary suggests the existence of anatomical patterns, a remarkable agreement with the fact that the cortical folding patterns are finite and redundant [96]. In consequence, brain MR image analysis should fully exploit the redundancy and finite variability of these patterns, i.e. a sparse representation framework which captures the relevant information. The example-based dictionary construction allows to collect multi-scale edge information of groups of folding patterns, thereby including a complementary source of redundancy: edge information in different directions and sizes. This fact implies that the relevant information characterizing the image patches, is encoded in few directions that can be effectively identified with a general technique such as PCA. Taking a large number of PCA coefficients increases the reconstruction error, issue that can be attributed to the fact that PCA eliminates many statistical dependences by projecting information into the principal directions, so any additional directions may behave as noise. Nevertheless, it is important to study if other approaches such as Non-negative Matrix Factorization (NMF) [79] or Probabilistic Latent Semantic Analysis (PLSA) [63], commonly used for matrix decomposition under certain constraints, would improve the reached performance presented here.

The principal element for sparsely representing a signal, the dictionary used to describe each image patch, has shown to comply (Section 3.5.5) with the desired conditions identified in SR problems: sparsity, recoverability and stability [136]. In addition, high-resolution reconstructions have shown to be nearly invariant to the chosen dictionary, indicating that our proposal does not introduces bias or deviations with respect to the images used to construct the dictionaries. It is important to mention that in these experiments the dictionaries were constructed using images from the base data set, which mixes brain volumes from normal and pathological subjects, so that normal and pathological patterns are present in the same dictionary. Also, the sparsity of the projection vector α in all experiments indicates that reconstructions are using at most 20% (850) of the atomic dictionary patches for representing each input patch, and thereby information from several bases are being combined at each time to estimate the SR corresponding patch.

One of the main contributions of the proposed approach is the reconstruction using non-overlapping patches when solving the optimization problem per patch. The reconstruction using overlapping patches involves an additional time (given that it is necessary to use a larger number of patches to cover the same region in the image),

and also requires a strategy (such as averaging) to handle multiple reconstruction values in overlapping areas. In average, reconstruction of a single image without overlapping takes about 26 minutes (needless to mention that in this case the algorithm is fully-parallelizable and therefore the processing time can be of the order of seconds), while using overlapping information increases this time to about 70 minutes (and this approach is hardly parallelizable). In terms of accuracy and similarity, our exhaustive experimentation indicates that non-overlapping and overlapping strategies are equivalent, that is to say, they give nearly the same image reconstruction errors. Given this result, we can state that our non-overlapping approach reconstructs image versions as accurate as those obtained with the overlapping strategy, in a considerably smaller interval of time. From a theoretical point of view, this can be considered as a result of the sparsification introduced by the Sobel transform and the fact that each voxel is described by a set of coupled high-low frequency information, obtained from neighbouring patches, enforcing coherence at boundaries between patches.

Another important contribution of this work is the use of a multi-scale edge analysis to estimate the missing high-frequency information. Representing brain MR image information in terms of edges for the SR problem has brought two major advantages. Firstly, introduction of multiplicative and high-frequency noise is clearly avoided. High-frequency information, related to true edges, has more influence in the reconstruction than other kinds of data, allowing at the same time to infer missing edge parts. Secondly, it is not strictly necessary to enforce image coherence and regularity by means of patch overlapping, as usually needed for natural images [133]. Indeed, in the present work each high-resolution patch was independently reconstructed, followed by a global regularization filter that corrects possible coherence problems that could appear between patches. This will lead us to naturally process the entire image in a parallel way, using a grid, cluster or any other distributed computing approach, decreasing in many orders of magnitude the computational time of reconstructing the entire image so as to use this approach in actual clinical scenarios.

It is important to note that, as our approach did not perform any preprocessing of the images, such as denoising or correction of intensity inhomogeneities, the super-resolved version of a noisy image will be noisy too. The main point is that we did not want to super-resolve and enhance an image at the same time, we are focused on the super-resolution, so that we are simply restoring an image as if it were generated by the MR equipment. Experiments on simulated and real brain MR volumes with different parameters (noise, intensity non-uniformities, slice thickness, scaling factor) have shown the adequate performance of our proposal under different scenarios. Regarding noise influence, the proposed approach has shown to work better when the image noise level is at 1%-3% than when the image is noiseless. This can be attributed to the fact that the

test image is synthetic and therefore this image is actually noiseless, but not the images used for constructing the dictionary, even though the original images were filtered out attempting to get a noiseless dictionary. The Rician noise, present in these images, is usually modeled as an additive noise, meaning for this case that when a noisy image is reconstructed with a noisy dictionary, this works better for the noisy image than for the noiseless one. Even though, even with few images acquired in diverse scanners and under varying protocols and acquisition parameters, such as in multicentric studies, our proposal is able to deliver meaningful high-resolution reconstructions to be further used in morphometric and statistical studies.

Our method is based on constructing dictionaries from previously acquired images. This fact (use of training images) can be seen as a drawback of our proposal when compared to other approaches that only uses information from the same image, in terms of additional image acquisition time, cost and specific MR equipment configuration. However, in our case, the acquisition of a set of few images for training has a non-significant raise of time or cost, when applied to a large study. The same scanner used to acquire the images in a study can be used to acquire the training data. Also, as we have demonstrated, meaningful dictionaries can be constructed even with a single image. Regarding MR sequence development, the idea of our method is to obtain the maximum resolution of the scanner from a lower resolution image, likewise it is always possible to acquire a high-resolution image to use as training data. Besides, the use of dictionaries for performing the super-resolution has other advantages, namely: it allows a better clinical interpretability, since the dictionaries can be built from pathological subjects. The pattern analysis of these pathologies may define anatomical markers (biomarkers) that could be useful not only for the diagnosis but also for the prognosis and follow-up. Finally, the distance of a particular image to a set of dictionaries could actually constitute a morphometry index.

The proposed method has demonstrated, visually and quantitatively, a substantial impact in brain tissue segmentation, representing an important step towards MR image analysis of anatomical details. However, immediate research efforts must be focused on determining the optimal number of raw sample patches required to generate compact semantic-based dictionaries, attempting to reduce the computational complexity and the time involved in solving the optimization problem. We have demonstrated that, in contrast with the actual trend that learned dictionaries will reduce the reconstruction time without losing accuracy, in our particular case, training of dictionaries built up from 3D patches takes a considerable time, given the atom dimensionality. The great computational time involved was the main reason for not choosing to learn the dictionaries, but instead to use the basic random sampling on specific regions of the images. The results obtained so far with the proposed approach have been generated

using a dictionary size of 4200 patches, however, two different approaches for reducing the dictionary size were tested: using less images but extracting the same number of patches from each, and extracting less patches per image but using the same number of images. Visual inspection of the results obtained using these dictionary sizes do not reveal noticeable differences (radiologist examination), but quantitatively, larger dictionaries always yield smaller RMSE. Nevertheless, the computational cost and time is approximately linear to the size of the dictionary [133], implying that larger dictionaries requires heavier computation. In fact, meaningful dictionaries can be constructed even with a single image, and this dictionary can still achieve good accuracy reconstructions. Our SR approach has been compared to a state-of-the-art algorithm, recently proposed by Manjón et al. [88], outperforming that strategy in a data set of real MR images. The main difference between this upsampling method and our SR method is that we perform a prior learning step from data sets of real brain MR images for appropriate dictionary construction, while the non-local approach of Manjón et al. [88] only uses information from the image to be reconstructed. In other contexts, such as in real-time applications, avoiding this learning step can be critical; however, in most clinical applications, where a large number of images are produced daily, this knowledge needs to be extracted only once to construct meaningful dictionaries adapted to specific characteristics of MRI sequences and equipments. We confirmed that the inclusion of a prior learning step on similar images represents an advantage in single-image SR, and increases both the accuracy as well as the computational speed of the reconstruction. The proposed approach can be explored to work with multimodal reconstructions, as studied before by other researchers [87, 110]. Following this, we have performed a preliminary experiment where 5 different low-resolution T2w images (size: 256x256x36) were super-resolved using dictionaries trained with the T1w images from the base dataset. As the original high-resolution T2w images are not available, quantitative information about reconstruction performance can not be reported. However, by visual inspection and comparison with a bicubic interpolation, our approach can provide better definition at image boundaries, in despite of some noise introduced in the reconstruction. With this experiment, we can state that our approach could handle multimodal reconstructions, however, it will require some additional work to improve performance.

Results on tissue segmentation have shown that the super-resolution reconstruction is a promising methodology for increasing the accuracy of morphometric analysis. Here, a preliminary study using SPM and DARTEL allowed us to verify that the reconstructed images can have nearly the same statistical power in Voxel Based Morphometry analysis of patient populations. However, more experiments are required to clearly identify the impact in statistical analyses and the relations with image acquisition parameters.

4 Extracting Brain Patterns using Visual Saliency for Imaging-Based Classification of Neurodegenerative Diseases

*Neurodegenerative diseases comprise a wide variety of disorders with different neurological dysfunction and a very irregular evolution. Currently, an objective method that helps to correlate the clinical onset with the radiological signs, is not available and the whole interpretation is dependent on the radiologist's skills. From a structural point of view, a main problem is that clinical analysis is not directly related to the visual analysis made by radiologists on brain images. Radiologists can hardly quantify systematic differences in these stages and current brain morphometry automatic analyses, that perform this quantification, do not allow a clinical useful interpretation. This paper presents a new fully automatic classification method that finds discriminative brain patterns associated, mining systematic differences and therefore grading objectively any neurodegenerative disorder. This is accomplished by a fusion strategy that mixes together bottom-up and top-down information flows. Bottom-up information comes from a multiscale analysis of different image features, while the top-down stage includes learning and fusion strategies formulated as a max-margin multiple-kernel optimization problem. Comparison of the classification performance of different configurations of the proposed approach in a public brain MR dataset (OASIS) with patients diagnosed with Alzheimer's disease, reveals an increment varying from 0.05 to 0.09 in the equal error rate measure for four different experimental groups, with respect to what has been reported by a state-of-the-art method. In terms of the anatomical analysis, relevant regions found by the proposed approach highly correlates to what has been reported in clinical studies of Alzheimer's disease. A short version of this chapter has been submitted for publication to the **Human Brain Mapping** journal.*

4.1 Introduction

Existing studies suggest neuroimaging may become a valuable tool in the early diagnosis of neurodegenerative diseases by extracting anatomical patterns and revealing hidden relations from structural Magnetic Resonance (MR) images. The value of neuroimaging against clinical, neuropsychological and biochemical analysis remains to be demonstrated in large representative populations, yet there exists sufficient evidence in small series of patients with different states of neurodegenerative disorders. The usual examination workflow is performed by expert neurologists or radiologists that are able to figure out complex pathological conditions and subtle changes with clinical meaning. The process that an expert follows when examining a particular case involves two different kinds of tasks: those related with image perception, such as visual search or exploration paths, and others associated with cognitive skills, mainly related to diagnostic reasoning and decision making [14]. An expert structures a diagnosis by using contextual knowledge and fusing information from different sources, a process that has been recently under study [14].

At analyzing structural brain MR images, a main aim is to find anatomical changes, either local or global, related to functional disturbances. In particular, radiologists examine images by looking at distinctively regions and compare them by searching differences [14]. In the computational attempt of emulating the human vision process –a synchronized collaborative work between the brain and low level visual mechanisms– the concept of visual attention has introduced a generation of techniques that are able to transform an image into a hierarchy of relevant regions, known as salient regions. Relevant regions in radiological terms may be defined as those image areas that are visually altered and are entailed with a certain degree of clinical interpretability. Nevertheless, most methods used to compare brains establish local rather than regional (salient) differences.

Currently, a morphometric brain analysis consists of a set of strategies aimed to extract and quantify anatomical differences between groups of subjects. Commonly, this analysis comprises two main processes: first, all images are warped or registered together to a common reference frame or template, and second, a quantification of the estimated local deformation required to register is computed, producing specific measurements of interest. Voxel-Based Morphometry (VBM) [5] and Deformation-Based Morphometry (DBM) [6] are currently the most used techniques to compare populations. In VBM, local differences, found in brain tissue segmentations, are voxel-by-voxel statistically analyzed, while DBM statistically compares information coming from the deformations fields obtained after registration to the template. With these methods, one-to-one correspondences between subjects are assumed and statistics are computed for the same

voxel across all subjects. However, conclusions are limited when the same structure may be partially present, or when a single anatomical region may exhibit multiple shapes across the population. On the other hand, some pathologies may affect not only a unique anatomical structure or even contiguous regions, but localized structures separated from each other. These kinds of patterns are difficult to find and analyze with these classic morphometric techniques. A recent proposal, the Feature-Based Morphometry (FBM) [121], copes with these issues by modeling the image as a collage of local scale-invariant features and by learning, from them, a probabilistic model that reflects group-related anatomical characteristics. However, these approaches disregard the local statistical dependences, and then subtle changes are hardly detected; exactly the opposite strategy used by the radiologists, who analyses regions rather than pixels [76].

In recent years, there has been an increasing interest in using analytical methods to improve inferences using a small set of individuals. These methods include, among others, supervised machine learning techniques and supervised pattern recognition algorithms, which are able to automatically extract information from data. Most of them are used for classification, while they attempt to automatically discover data patterns. The most popular technique has been by far the Support Vector Machine (SVM), which has been applied to classifying individuals with several neurological disorders. A complete review and comparison of about 40 SVM-based approaches for classification of neurological and psychiatric diseases using neuroimaging data can be found in [97], specifically from structural MR brain images. Classifications of patients with depressive disorders, psychosis, schizophrenia, Alzheimer's disease and Parkinson's disease, among others, have been presented in this review. Useful information for classification can be extracted from the whole brain volume or from specific Regions of Interest (ROIs), manually delineated or registered to a parcellated atlas. The SVM classifier is usually fed with features such as intensity [99, 113], textural and statistical information [36, 54, 82], binary tissue segmentations [73, 81, 132] or cortical thickness estimations [130]. Overall, a dimensionality reduction technique helps to decrease the computational time and the presence of irrelevant and noisy features. Recently, it has been shown that the analysis using only ROIs [40, 55, 84] outperforms any of the other methods. This statement points out the fact that a biased analysis, performed specifically on the known anatomical disease locations, systematically leads to stronger and more significant conclusions. However, most neurological disorders have a very variable clinical and pathological presentation, whereby subtle patterns or the atypical disease entities can be easily missed and, on the other hand, previous ROIs delineations are highly time-consuming and expert-dependent.

Among the neurodegenerative diseases studied using structural MR image information,

the most known are the Alzheimer’s disease (AD), the Mild Cognitive Impairment (MCI) and the schizophrenia. In particular, AD is the most common type of dementia, affecting over 20 millions of people in the world. In the clinical practice, a probable diagnosis is reached with specific neuropsychological tests, clinical examinations and particular conditions of relatives. Complementary sources of information, relevant for accurate AD diagnosis, come from different medical imaging techniques, such as structural and functional Magnetic Resonance Imaging (MR) and positron emission tomography (PET). At present, brain atrophy, hypometabolism and quantification of specific proteins measured using these techniques have been proved to be sensitive to AD. In the particular case of structural MR images, early diagnosis of AD turns out to be a challenging task [39], basically because the atrophy patterns associated with aging can be confounded with complex patterns associated with the pathology. The necessity of arriving to accurate diagnoses, given these difficulties, has increased the interest in structural neuroimaging.

Several studies, by taking structural MRI volumetric measurements of specific brain regions, have demonstrated significant differences between patients with probable AD and normal controls [23, 33, 55, 104, 106, 129–131]. These studies have reported several signs of the disease progression by using exclusively anatomical relationships, i.e., from MCI to AD. In particular, it has been classically described that AD early stages are characterized by premature atrophic changes along the perforant hippocampal pathway, including the entorhinal cortex (a portion of the anterior parahippocampal gyrus), the hippocampus and the posterior cingulate cortex [41, 50], and, in some cases, along the banks of the superior temporal sulcus [39, 70]. Later on, larger regions such as the temporal, parietal and frontal neocortices may also be compromised and exhibit neuronal loss. In terms of the disease progression, Frisoni et al. [50] have established the morphometric protocol when measuring the whole-brain, paying particularly attention to the entorhinal cortex, hippocampus and temporal lobe volume estimations, as well as to the associated ventricular enlargement percentages. From the diverse automatic classification approaches that have been applied to the study of AD, the relevant anatomical regions usually corresponds to those clinically described: hippocampus, amygdala, entorhinal cortex, temporal gyrus and parahippocampal gyrus [81, 84, 121, 130].

This paper proposes an automatic image analysis method inspired by the radiologist visual perception. The method builds on a visual saliency model and extends it to involve a learning process that mimics the adaptation of a radiologist visual perception. The method performs a multiscale analysis of saliency maps that are optimally combined. This method is able to map any brain to a set of visual patterns that previously have been learned as associated to the pathological or normal condition. This is not about certain salient points but salient regions, whereby the whole brain structure

results classified either as pathological or normal. The proposed method has been validated by accurately classifying patients from a public brain MR dataset (OASIS) [89] as probable AD subjects or normal controls.

A main contribution of this paper is a fusion strategy that learns, from training data, the discriminant structural patterns of neurological disorders, in particular, the Alzheimer's disease. Another important contribution is the model interpretability, since the learned patterns can be mapped to the original brain and used to quantitatively estimate the importance of each region for the final classification, thereby improving the current understanding of the disease. Technical contributions include: the use of a 3D multiscale analysis of the brain saliency inspired by what radiologists do when examining cases, the use of low-level features that sparsify data, and the formulation of the model adaptation and fusion strategies as a max-margin multiple-kernel optimization problem. An extensive parameter analysis of the influence of the image features as discriminative factors is also carried out. The classification accuracy between normal controls and probable AD subjects is improved by applying this approach, outperforming a recently proposed technique (FBM). To the best of our knowledge, this kind of visual-saliency-based pattern extraction approach has not been previously investigated for AD characterization and classification in structural MR images.

The rest of the chapter is organized as follows. First, Sections 4.2 and 4.3 presents the methodological considerations for modeling the problem and the proposed solution. Then, the description of the dataset and the proposed saliency-based classification framework is presented in Section 4.4, together with extensive experiments and comparison with another approach on the OASIS dataset presented in Section 4.5. Finally, the conclusions and future work are discussed in Section 4.6.

4.2 Representation of Relevant Structural Regions by their Saliency

Salient regions can be thought of as those image regions considered as relevant in a particular context. Identification of such regions is associated with structuration of the image information following bottom-up or top-down organizational flows. Bottom-up approaches emulate the preattentive vision mechanisms and are independent of the problem knowledge [53, 66, 78]. In contrast, top-down strategies are task guided and aim to reach a global understanding of the image contents [49, 52, 94]. Such strategies are broadly used in artificial vision to determine areas of interest, but as long as we know this type of analysis has not been applied to figure out particular image patterns, specially in the medical context.

At analysing structural brain MR images, a main aim is to find anatomical changes related to functional disturbances that may result in local or global morphological alterations. This problem can be formulated as a search of changes of particular patterns among the anatomical areas and can be considered as equivalent to figure out the preferential information flux through a net of nodes belonging to a fully-connected graph, being each node a particular anatomical region and each edge a similarity (or dissimilarity) measure. Similar approaches have been successfully used to determine a privileged path when an user is browsing Web pages, namely the well known Google's PageRank [100] or HITS [72] algorithms. If an image is partitioned and its parts are somehow connected together, the interaction process between an user and the image can be modeled as a fully connected graph. In particular, radiologists usually analyze images by looking at distinctively regions and compare them by searching differences [14]. Relevant regions in radiological terms may be defined as those image areas with clinical meaning that are visually altered and correspond to the salient image features. There exist different approaches to calculate the image saliency in natural images, but, to our knowledge, none of them has been applied to medical images.

Some approaches have taken into account the possibility of using graphs and induced Markov chains to model saliency and attentional fixations in natural images. Typically, to construct a graph associated to an image, the image pixels (or patches) are represented as vertices, and the relations between them are encoded within the graph edges, in some cases with a specific weight associated to each connection. Global relations are represented when interconnecting all vertices, forming a fully-connected graph, while local relations can be analyzed by connecting each vertex only to its immediate neighbors. Graph relations are summarized in the adjacency matrix A , for which each position i, j corresponds to the edge connection weight between the pair of vertices i and j . Algebraic operations of the adjacency matrix A allow to select important information about the graph structure. For example, the centrality measure, proposed by Freeman [48], allows to estimate the global importance of each vertex, if importance is defined in terms of how "central" the vertex is within the graph structure. An application of such measures (centrality degree) to modeling visual saliency in natural images was proposed by Pal et al. [101].

Once the image graph is constructed, a Markov chain allows to analyze the graph flow by associating the graph vertices to states, and the edge weights to transition probabilities. The adjacency matrix A of the graph corresponds now to the stochastic matrix M of the Markov chain, and under some appropriate assumptions (irreducibility of M due to a strongly connected graph) it has been proved that the Markov chain tends to a unique stationary probability distribution. This stationary (equilibrium) distribution represents the amount of time (or frequency of visits) that a random walker

would spend at each state if he were allowed to walk for an infinitely long time. This can be interpreted as a proportional measure of the vertex importance, with respect to all other vertices in the chain. This measure has been shown to be equivalent to the eigenvector centrality described by Freeman [48], where the vertex importance is calculated in a self-referential way: a vertex is important because it is adjacent to other important vertices. Costa et al. [29] has used this as a saliency measure, however, his proposal lacks of an extensive validation over real images (only saliency results of two synthetic images were reported). On the other hand, Gopalakrishnan et al. [57] proposed a robust approach where global and local image properties are computed from Markov random walks on a complete graph and a sparse k -regular graph, respectively. From this information, a small number of the most salient (object) and background nodes are set and labeled as object or background. Then, a semi-supervised learning technique uses them to determine the labels of the remaining image nodes (patches) and yield a binary saliency map. The combination of global and local properties allows, according with these authors, to better identify salient regions that correspond to objects in the image, discarding cluttered backgrounds with high local contrasts that can result in larger saliencies.

A similar approach, known as Graph-Based Visual Saliency (GBVS), was proposed by Harel et al. [61], case in which the graph connections encode dissimilarities between feature pixels modulated by a closeness measure, i.e., salient pixels are the most dissimilar in a local context. This method aims to find areas of human visual fixation rather than salient objects in a scene (as attempted by Costa et al. [29] and Gopalakrishnan et al. [57] approaches), and therefore is closer to a radiologist's diagnostic process [14]. Unlike other approaches that connect graph vertices in terms of similarity of image features, the local dissimilarity between image pixels in the GBVS method can be seen as an approximation to the visual analysis made by radiologist when studying a medical image. The method therein introduced was adapted to extract the relevant patterns involved in the diagnostic process, as further described in Section 4.4.

4.3 Adaptive Learning of Salient Regions

Medical images are particularly challenging for saliency modeling since the relevant information is a complex mixture of local and global patterns that are somehow altered by a particular pathological process. The human visual system (HVS) is not naturally trained to find such patterns. Radiologists undergo a long learning process that trains their visual system to find patterns associated with particular physiological or pathological conditions, detecting salient areas with clinical meaning. This condition is however limited in quantitative terms since radiologists can naturally determine where

the differences are, if they are notorious enough, but they can hardly establish how important these differences may become, among others because of the high anatomical and pathological variability of any neurological disorder, in particular the Alzheimer's disease. The interest then on revealing such salient meaningful clinical locations is that an objective strategy could set a hierarchy of differences, thereby decreasing the inter-expert variability. The whole task nevertheless consists not only in exposing those relevant areas but also in finding an adaptive strategy that picks those regions and sorts them out by an importance order with clinical meaning.

The fact that brain anatomical and pathological variabilities can easily hinder subtle structural changes makes that any learning strategy should be adaptive and non linear. If some salient features are less informative than others, it should be possible to down-weight their importance. If a disorder causes atrophic and hypertrophic changes simultaneously, the linear model would only be able to encode one mode of variability, but a non-linear model may be able to capture both modes [7]. From a clinical standpoint, in terms of the disease knowledge, it is fundamental to determine the minimum number of morphological regions that better correlates with a pathological stage, in which case the learning strategy should also be sparse. In consequence, a main goal of the present investigation was to build a fusion strategy, i.e., an hybrid bottom-up-top-down computer saliency model that were able to find visual patterns in MR images associated to the presence of Alzheimer's disease. Fulfilling such goal requires: (1), to use a set of visual features which are appropriate for modeling the visual content of MR images (bottom-up) and, (2), to properly involve the problem knowledge in an adaptive representation scheme (top-down).

In theory, a proper characterization of the anatomical differences between two groups passes by determining the specific dissimilar areas. The performance of any model in such a task is simply measured as how good this model is to separating the two groups. Different techniques have allowed to approach this problem, from standard multivariate models, PCA or MANOVA to complex pattern recognition approaches. Two different approximations, the discriminative and generative strategies, have been used, being the discriminative models more robust except for small population groups [7]. Overall, these discriminative techniques are non-linear models that project the data into a space with a higher number of dimensions, where data can be linearly fitted. From the broad spectrum of existing non-linear models, Support Vector Machine (SVM) has been used as a paradigm in many types of problems with separable and non-separable data. SVM uses a kernel function $k(\mathbf{x}_i, \mathbf{x}_j)$ which computes the similarity between samples \mathbf{x}_i and \mathbf{x}_j . This kernel is nothing but a projection of the data to another space in which data can be linearly separable. The success of SVM is dependent on the choice of good kernels which are typically hand-crafted and known in advance. In practice, learning

problems involve multiple, heterogeneous data sources that hide complex statistical dependences and therefore are really difficult to separate in consistent classes. A better alternative consists then in somehow mixing up multiple kernels into a single kernel, that is to say a transformation that represents an optimal combination of what the individual kernels make to a each part of the data space. In particular, Multiple Kernel Learning (MKL) learns the kernel from training data, it focuses on how the optimal kernel can be learnt as a linear combination of given base kernels. It should be strengthened out that a linear combination of base kernels corresponds to concatenation of individual kernel feature spaces, whereby the type of information fusion is non-linear and therefore the relevant features of each space can be fully conserved.

4.4 Materials and Methods

4.4.1 OASIS public data set

A set of 198 brain MR images from healthy (98) and pathological (100) subjects, extracted from the OASIS (Open Access Series of Imaging Studies) database [89], were used to evaluate the performance of the proposed approach. Each subject has been previously analyzed with a Mini-Mental State Examination (MMSE) and a Clinical Dementia Rating (CDR), and diagnosed as normal controls (NC) or with probable Alzheimer’s disease (AD) using the scores obtained in the MMSE and CDR tests. As described in [89], a set of 3-4 images were acquired per each subject on a 1.5T Vision scanner (Siemens, Erlangen, Germany), using a T1-weighted magnetization prepared rapid gradient-echo (MP-RAGE) sequence, on a single imaging session. Images were first spatially warped into the 1988 atlas space of Talairach and Tournoux, using a 12-parameter affine transformation as described by Buckner et al. [18]. Images per each subject were then averaged, obtaining a single, high-contrast MP-RAGE image in atlas space per subject, and skull-stripped by application of a loose-fitting atlas mask, as described in [89]. Finally, images were gain-field corrected, based on a fitted quadratic inhomogeneity model introduced in [120]. For the sake of a comparison with the FBM technique, which has been tested on the same dataset (Toews et al., 2010), results are reported for four different groups:

- **Group 1:** 86 subjects, aged between 60 to 80 years: includes 66 healthy controls and 20 patients suffering only mild AD (CDR=1).
- **Group 2:** 126 subjects, aged between 60 to 96 years: includes 98 healthy controls and 28 patients suffering only mild AD (CDR=1). This group results after

including in the Group 1 elderly subjects (>80 years) to increase the difficulty level, given that aging patterns can be confounded with the disease patterns.

- **Group 3:** 136 subjects, aged between 60 to 80 years: includes 66 healthy controls and 70 patients suffering both very mild and mild AD (CDR=0.5,1): This group results after including in Group 1, patients with very mild AD to increase the difficulty level, given that in this case the very mild AD patients could not necessarily present visual differentiating patterns.
- **Group 4:** 198 subjects, aged between 60 to 96 years: includes 98 healthy controls and 100 patients suffering very mild, mild and moderate AD (CDR=0.5,1,2). This group includes all subjects available in the OASIS data set, and can be categorized as the most difficult to classify, given that it mixes both elderly subjects and different stages of the disease.

As pointed out by Toews et al. [121], analysis of the classification performance must take into account the clinical and demographic information of subjects in the dataset, given that it is more difficult to discriminate between elderly normal and pathological subjects, or between healthy subjects and patients with very mild AD. The four dataset groups are proposed to illustrate the influence of these aspects.

4.4.2 Proposed Approach

The proposed method is based on a two-phase visual saliency model that combines a bottom-up and top-down approaches to achieve accurate classification of brain MR images into normal controls or probable AD subjects. The bottom-up phase performs a multiscale analysis of different basic image characteristics, similar to what was previously described [61], and fuse them to feed a discriminative model. The top-down phase uses high level knowledge, represented by the labels assigned to the training brain MR volumes, to adapt the parameters of the bottom-up saliency model using a Multiple Kernel Learning strategy. A graphical overview of the proposed approach is presented in Figure 4-1.

Calculation of Saliency Maps

Calculation of saliency information starts by extracting a set of feature maps from a given image volume x , $\{T^\phi(x)\}_\phi$, where $\phi \in \Phi$ indicates a visual feature at a particular scale, as depicted in Figure 4-2. For MR images, selected features includes intensity, orientation and edges. Intensity information corresponds to the individual gray value of each voxel; orientation information is calculated using a bank of Gabor

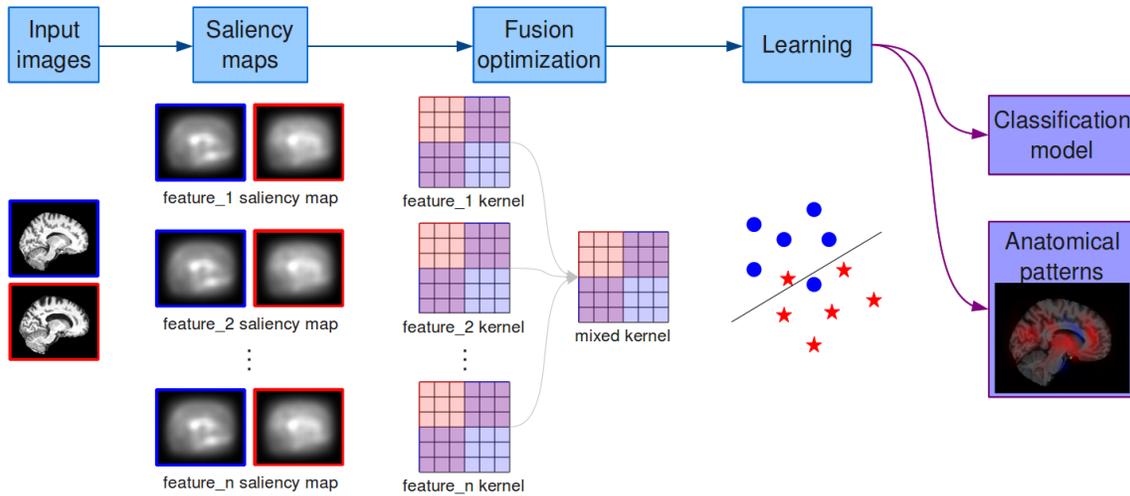


Figure 4-1: Graphical overview of the proposed method. Input brain MR images are processed to extract different feature-scale saliency maps that highlight different brain features, then a learning algorithm fuses optimally this information to feed a SVM classifier. The model produces both a classification model as well as maps of anatomical regions where differences were established.

filters with four different orientations (0° , 45° , 90° and 135°); and finally, edge information is extracted by applying a Sobel operator, with a kernel size of $5 \times 5 \times 5$, for the 3 different orthogonal directions. Projection of the original data into these three different spaces attempts to sparsify the raw brain data since sparser representations facilitate dimensionality reduction [37], a crucial factor for the success of any classification strategy. Furthermore, the selected features aim to approximate the sparsity of the Human Visual System, a concept illustrated recently by Olshausen et al. [95], who have shown that sparse coding of images produces Gabor-like oriented filters that resemble the receptive fields of simple cells in the visual cortex. The feature maps are also calculated at different scales, by subsampling the volume to $1/4$, $1/8$ and $1/16$ of the original size. The motivation for including such a multi-scale analysis comes from the observation that objects in a scene may appear in different ways depending upon the scale of observation [80], but their information is proportional to the coherence through the different scales. In summary, a set of 18 different 3D feature maps at various scales (3 for intensity, 12 for orientation and 3 for edges) is finally collected. Subsequently, a per-slice-fully-connected graph G_A^ϕ is defined on each feature map $T^\phi(x)$, where vertices correspond to image pixels and edges store information of a regional dissimilarity between nodes. As proposed in [61] and depicted in Figure 4-3,

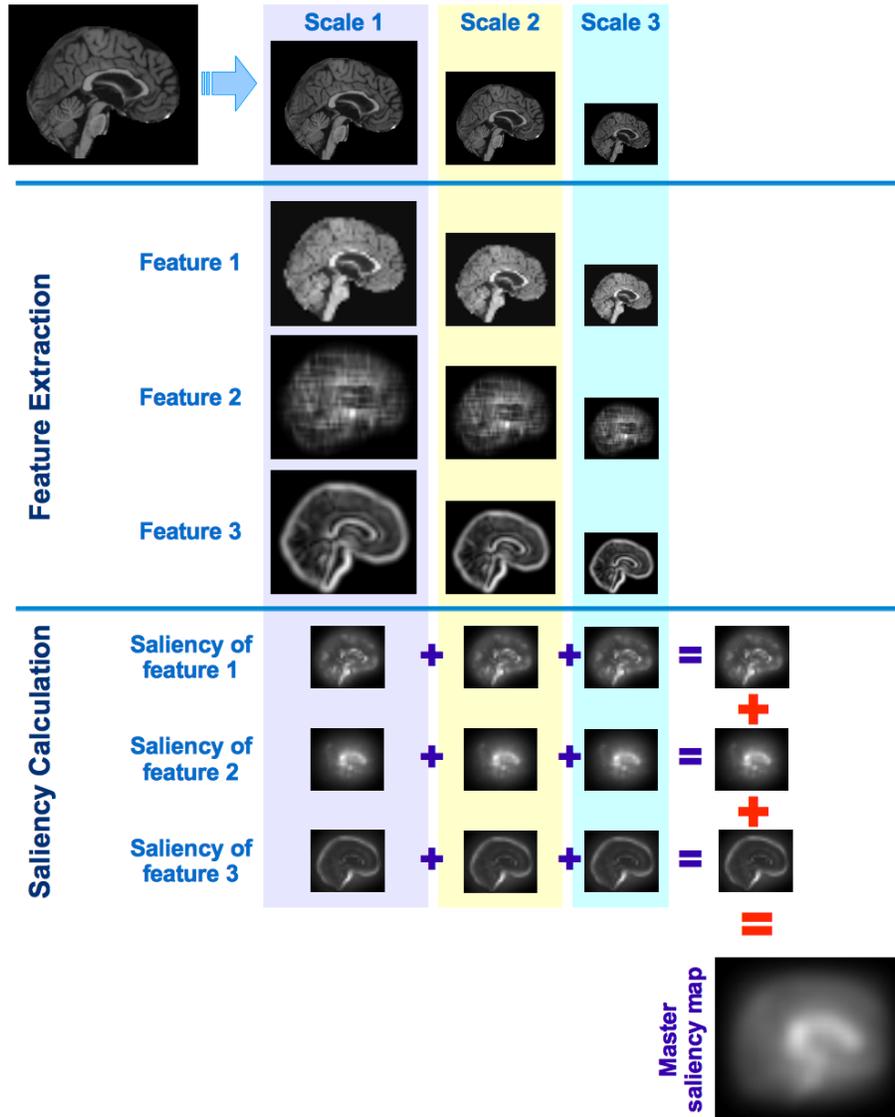


Figure 4-2: Construction of the saliency maps: first, the input image is decomposed into three different scales. Second, different features are extracted from the scaled images, namely intensity, orientation and Sobel edges. Finally, for each feature and scale, the obtained maps are down-sampled to set maps with the same size. The master saliency map is built by adding up all scales and features into one single saliency map.

the edge weight between graph nodes $g_{i,j}^A$ and $g_{p,q}^A$ is calculated with

$$w_A(g_{i,j}^A, g_{p,q}^A) = d(g_{i,j}^A, g_{p,q}^A) \cdot F(i - p, j - q)$$

where $d(g_{i,j}^A, g_{p,q}^A)$ encodes the dissimilarity (in terms of the respective feature informa-

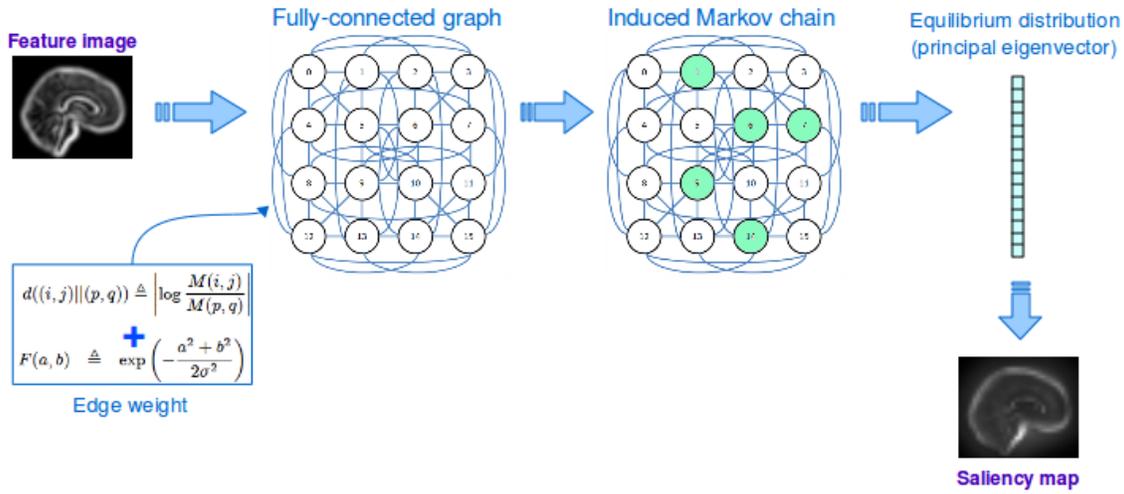


Figure 4-3: Overview of the graph-based approach for computing saliency. Once a set of characteristics are calculated, this feature image is represented using a fully-connected graph, which encodes the closer dissimilarity of the previously found features. A Markov chain is then induced over the graph and its associated equilibrium distribution sets the more salient nodes as those with the larger probability of being visited if a random walker were allowed to wander around. This saliency is coded in the first eigenvector of the adjacency matrix and superimposed upon the feature image as the output saliency

tion) and $F(i - p, j - q)$ encodes the spatial closeness between nodes. Dissimilarity is calculated as

$$d(g_{i,j}^A, g_{p,q}^A) = \left| \log \frac{T^\phi(v)_{i,j}}{T^\phi(v)_{p,q}} \right|$$

, where the inclusion of a logarithmic metric guarantees that larger feature dissimilarities pop out easily while similar feature information have little impact in the edge weight. On the other hand, the closeness is measured with

$$F(a, b) = \exp \left(-\frac{a^2 + b^2}{2\sigma^2} \right),$$

where σ is a free parameter of the GBVS algorithm, fixed to 0.15. This means that feature dissimilarity information is penalized with the spatial distance between nodes, thus encoding regional dissimilarity information at the graph edges.

Afterward, activation maps A^ϕ that expose connected regions of high dissimilarity, are determined by constructing a Markov Chain on each G_A^ϕ and estimating its equilibrium

distribution as the principal eigenvector of the stochastic matrix, using the Power Iteration Method [100]. Once activation maps are computed, a normalization step is required to guarantee that these maps concentrate the activation (saliency) only in a few key locations [61]. To do so, the same Markovian approach is applied to each activation map A^ϕ , using a new graph G_N^ϕ with image pixels as vertices, but edges now storing information about regional activation:

$$w_N(g_{i,j}^N, g_{p,q}^N) = A^\phi(p, q) \cdot F(i - p, j - q)$$

so the equilibrium distribution of a new Markov chain on each G_N^ϕ highlights pixels with high activation (saliency). Finally, the feature saliency map $S^\phi(x)$ for the whole volume x is constructed by stacking the 2D-per-slice saliency maps.

Bottom-up Saliency Fusion

The model described in the previous section provides a set of saliency maps for a volume x , $\{S^\phi(x)\}_\phi$, that encompasses different scales, different types of basic visual features (color, edges etc), and different volume orientations (sagittal, coronal or axial). Each saliency map is calculated by a function $S^\phi : I \rightarrow [0, 1]^{m \times n \times l}$, where $\phi \in \Phi$ indicates a particular combination of scale, visual feature and volume orientation, and I indicates the volume representation space which is usually $\{0, \dots, 255\}^{m \times n \times l}$, with (m, n, l) the volume size.

Computational visual saliency models use different strategies to fuse information from saliency corresponding to different visual features. A common strategy is to weight the maps and then sum them up to calculate an overall saliency map [49]:

$$S^*(x) := \sum_{\sigma, \phi} \omega_{\sigma, \phi} S_\sigma^\phi(x)$$

The problem with this strategy is that important information from the individual features may be lost when somehow linearly filtering out the maps. The proposed model uses a different strategy that keeps all the information from the different saliency maps and only fuse them when a decision needs to be made. Specifically, the saliency maps are used as input to a discriminant function, $g_W(x)$, that indicates to which extend a particular volume x corresponds to a probable AD case:

$$g_W(x) = \langle W, (S^1(x), \dots, S^{|\Phi|}(x)) \rangle,$$

where $W \in \mathbb{R}^{|\Phi| \times m \times n \times l}$ is the vector of parameters that indicates the relative importance of each voxel extracted from each saliency map $S^\phi(x)$. The value of $g_W(x)$ is

expected to satisfy $g(x) \geq 0$ if x corresponds to an AD case and $g_W(x) < 0$ if x corresponds to a healthy subject. This model can be extended by introducing a saliency mapping function as follows:

$$g_{W,\Gamma}(x) = \langle W, (\gamma_1 \Psi(S^1(x)), \dots, \gamma_{|\Phi|} \Psi(S^{|\Phi|}(x))) \rangle, \quad (4-1)$$

where $\Psi : [0, 1]^{m \times n} \rightarrow F$ is a function that maps each saliency map $S^\phi(x)$ to a feature space F , $W \in \mathbb{R}^{\dim(F) \times |\Phi|}$ and the parameters $\Gamma = \{\gamma_\phi\}$ indicate the relative importance of each saliency map $S^\phi(x)$. This formulation has the advantage of improving the flexibility by allowing the model to account for potentially complex non-linear interactions between original single saliency values corresponding to different features and scales.

An important parameter of the model in Equation 4-1 is the function Ψ , which maps a saliency map to a new feature space. The main aim of such representation transformation is that complex non-linear patterns in the original space become linear in the new, potentially high-dimensional, feature space. This is a well known strategy used in kernel methods, the *kernel trick*, where the mapping Ψ is implicitly induced by a kernel function. A kernel is a function $k : X \times X \rightarrow \mathbb{R}$ associated to a mapping $\Psi : X \rightarrow F$ such that $\forall x, y \in X, k(x, y) = \langle \Psi(x), \Psi(y) \rangle_F$, i.e., k calculates the dot product in F . Intuitively, a kernel may be seen as a function that measures the similarity between two objects from the input space. In the proposed model, the input space is the space of saliency maps, so a kernel function measures the similarity between saliency maps.

Top-down Learning

The goal of the top-down model is to adapt the parameters of the bottom-up model in such a way that the most discriminative image features, represented in the different saliency maps, receive a higher weight value. Specifically, the top-down model uses domain knowledge codified as a set of labeled training volumes, to find optimal values for the parameters W and Γ that maximize the discriminative ability of the model. This is formulated as the following min-max-margin-discrimination optimization problem:

$$\begin{aligned} \min_{W,\Gamma} \quad & C \sum_{i=1}^N \max(0, 1 - y_i g_{W,\Gamma}(x_i)) + \|W\|_2^2 + \|\Gamma\|_1 \\ \text{s.t.} \quad & \Gamma \geq 0 \end{aligned} \quad (4-2)$$

where x_i represents a training volume, $y_i \in \{-1, 1\}$ represents the corresponding label, N is the number of training samples, and C controls the regularization of the model.

The first term of the objective function in Equation 4-2 is a loss function that penalizes the wrong classification of training samples, the second and the third terms are regularizers of the W and Γ parameters respectively. It is important to notice the difference between the regularizer of W (l_2 norm) and Γ (l_1 norm). The regularization of W is associated to finding a max margin classifier in the same way as it is done for support vector classification [28], while the regularization of Γ aims to find the sparsest set of feature weights. In this context, sparsity is motivated by the goal of finding a reduced set of saliency maps that better codify visual patterns to discriminate probable AD cases from normal controls.

Regarding the function Ψ , in this work we used two different kernels: the linear kernel that simply calculates the dot product between its inputs and corresponds to an identity map $\Psi(x) = x$, and the histogram intersection kernel defined as:

$$k_{hi}(s_p, s_q) = \sum_i \sum_j \sum_k \min(s_p(i, j, k), s_q(i, j, k)).$$

The histogram intersection kernel requires the input saliency maps to be normalized in such a way that all their values add up to 1. This is inspired by the fact that a saliency map may be seen as a saliency probability distribution over the voxels of a volume.

A saliency map kernel k may be extended to a volume kernel as follows:

$$k_\phi(x_p, x_q) := k(S^\phi(x_p), S^\phi(x_q))$$

These kernels can be combined in a single kernel $k^* = \sum_\phi \gamma_\phi k_\phi$. It is not difficult to show that the kernel k^* is associated to the mapping:

$$\begin{aligned} \Psi^* : I &\rightarrow F^{|\Phi| \times |\Sigma|} \\ x &\mapsto (\gamma_1 \Psi(S^1(x)), \dots, \gamma_{|\Phi|} \Psi(S^{|\Phi|}(x))) \end{aligned}$$

that maps a volume to the feature space where the discriminant function in Equation 4-1 is defined. In other words, the proposed model fuses the information from the different saliency maps by combining the respective kernels instead of directly adding the saliency maps themselves. This means that the optimization problem in Equation 4-2 can be seen as a Multiple Kernel Learning (MKL) problem, where in addition to finding a good discriminant hyperplane determined by the parameters W , the contribution of each kernel, determined by coefficients in Γ , must be also found. Several MKL formulations have been proposed so far, the approach herein used was proposed by Varma et al. [126], and describes a Generalized Multiple Kernel Learning (GMKL), suitable for learning different combinations of kernels.

Three different kernel groups were evaluated to test the influence of the different features, scales and orientations:

- **Kernel group 1 (KG 1):** 3 kernels calculated using saliency maps coming from only intensity, orientation or Sobel edge information
- **Kernel group 2 (KG 2):** 9 kernels calculated using saliency maps coming from only intensity, orientation or Sobel edge information evaluated in the three acquisition planes (sagittal, coronal and axial)
- **Kernel group 3 (KG 3):** 18 kernels calculated using saliency maps coming from only intensity, orientation (each orientation angle evaluated separately) or Sobel edge information evaluated at three different image scales (1/4, 1/8, 1/16)

At each group, saliency maps per feature, scale and orientation were individually compared using the histogram intersection to construct a kernel matrix per each. Then, the model parameters (W and Γ) are learned by solving the optimization problem in Equation 4-2 over a set of labeled training volumes. Cross validation over a subset of training images was used to find an optimal value for the regularization parameter C . Finally, with the optimal C , the final classification of test subjects is performed.

In all experiments, we used the GMKL source code posted in <http://research.microsoft.com/en-us/um/people/manik/code/gmkl/download.html>, while the SVM classifier is implemented using the LIBSVM toolbox [21], with precomputed kernel matrices.

4.4.3 Performance Evaluation

Classification of each group is performed in a leave-one-out manner, for which one subject at a time is set aside during the training phase and then classified using the SVM model trained with the remaining subjects. Classification performance was validated using the following metrics:

- Accuracy (Acc) = $\frac{TP+TN}{TP+TN+FP+FN}$
- Sensitivity (Sens) = $\frac{TP}{TP+FN}$
- Specificity (Spec) = $\frac{TN}{FP+TN}$
- Balanced Accuracy (BAC) = $\frac{Sens+Spec}{2}$
- Equal Error Rate (EER): the point on a ROC (Receiving Operating Characteristic) curve where the false positive rate and false reject rate (1- true positive rate) are equal.

where TP stands for true positives (AD individuals correctly classified), TN for true negatives (NC individuals correctly classified), FP for false positives (NC individuals misclassified) and FN for false negatives (AD individuals misclassified).

4.4.4 Anatomical Interpretation

A main contribution of this work is that the method allows not only to classify structural MRI brain images but rather to highlight the anatomical areas related with a particular diagnosis, thereby achieving clinical interpretability. The method actually finds a quantitative estimate of the found brain differences, an important issue in terms of the clinical management of the Alzheimer's disease.

The solution of the max-margin classification problem defined in Eq 4-2, requires to find both an optimal coefficient vector W , which indicates the relative importance of each voxel of each saliency map in the saliency map space, and an optimal vector Γ , which indicates the contribution of each individual kernel. Provided that the transformation used effectively separates the space into two classes, the saliency maps voxels with largest absolute coefficients can be considered as those that contain the most relevant information characterizing a class, while the saliency maps voxels with smallest coefficients are entailed with less related information.

In the proposed approach, the pre-defined kernels convert the input image into individual feature saliency maps, whose voxels corresponds to dimensions of the saliency map space. The vector Γ allows to construct a master saliency map that incorporates the relevant feature information, by combining the individual feature saliency maps according with the learned kernel weights. Therefore, the vector W defining the separating hyperplane allows to identify the most relevant regions for AD discrimination, i.e., to set a particular set of saliency map voxels as important for discriminating AD class from the NC class, by using the coefficient value (positive or negative) associated to each master saliency map voxel. Those relevant regions can then be visualized in an overall discrimination relevance map by performing a linear combination of the master saliency maps and their corresponding coefficients, as illustrated in Figure 4-4.

By aligning these discrimination relevance maps to some previously delineated brains in standard atlases, it is possible to accurately identify the specific anatomical brain areas involved in the identification of normal controls and probable AD patients. In the present investigation, 96 cortical and 21 subcortical structural areas obtained from the Harvard-Oxford atlas [35] (see Figure 4-5), have been used to label the Regions of Interest. For each anatomical region, the maximum value of its discrimination relevance map is stored, allowing to compare relevance values of the different anatomical regions. This scale allows to set different disease patterns at quantifying the importance of each of the anatomical areas.

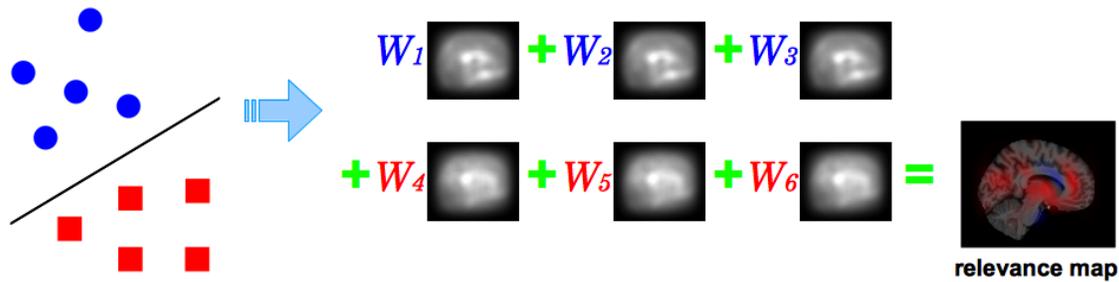


Figure 4-4: Schematic description of the relevance weight map construction. The optimal hyperplane, defined by W , found by solving the max-margin optimization problem (Eq 4-2) is visualized directly over the brain volume. Voxels with the highest positive or negative coefficients are colored according to the class: red for AD and blue for NC.

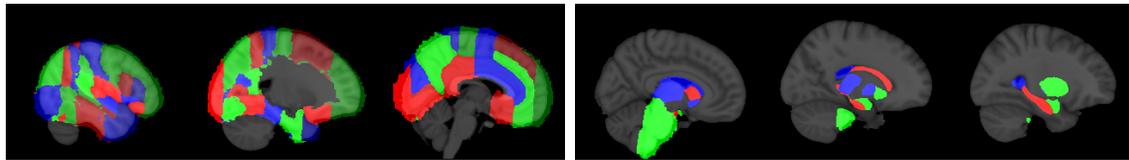


Figure 4-5: Illustration of Harvard-Oxford cortical (left) and subcortical (right) atlases.

4.5 Experimental Results

The proposed method was assessed with respect to its capability to discriminate different experimental groups of the dataset as well as its ability to determine the important regions and to weight them in diagnostic terms.

4.5.1 Classification Results

First of all, the obtained performance was compared with a state-of-the-art method, the FBM framework [121], that reports classification results over the same OASIS groups using the Equal Error Rate (EER). Table 4-1 shows the EER values obtained with the different kernel groups and the values reported in [121] for the same classification groups.

The presented results show the performance for the three different kernel groups (3 features, 9 feature-plane combinations and 18 feature-scale combinations) and the reported classification values of the FBM framework. A single kernel and a conventional SVM algorithm (without MKL), reach a similar performance to what has been re-

	Baseline	Our approach			
	FBM	Single kernel	KG 1	KG 2	KG 3
Group 1	0.80	0.80	0.82	0.83	0.86
Group 2	0.70	0.71	0.76	0.73	0.79
Group 3	0.71	0.71	0.73	0.74	0.76
Group 4	0.65	0.68	0.70	0.69	0.69

Table 4-1: Comparison of classification performance (using the EER) of the different configurations tested in the proposed approach (use of a single kernel against different combinations of feature, scale and plane kernels) with respect to the FBM approach of Toews et al. [121]. Best classification performance per each OASIS group is highlighted in bold.

ported by Toews et al. [121] ($p = 0.25$). This performance improves when introducing the learning step and, as long as information is segregated into different kernels, classification values ameliorate even more. For classification of Group 1, the KG 3 with 18 dimensions reports 0.86 and outperforms the other learned kernels, while the KG 2 with 9 dimensions reports 0.83 and the KG 1 with 3 dimensions reaches 0.82. Overall, this trend is observed through the whole experimental setup: the performance increases in about 7.5% for Group 1, 13% for Group 2, 7% for Group 3 and 7.6% for Group 4 with respect to the baseline ($p < 0.05$). The last group, which corresponds to the whole dataset, mixes different disease conditions, and therefore the particular patterns result hindered by the large anatomical variability. In that case, the mechanism of segregating information is insufficient to capture the discriminant features and then a more general analysis classifies slightly better, i.e., KG 1 moderately outperforms KG 2 and KG 3. For all these cases, the presented method is more accurate than the baseline.

Evaluation has been extended to other type of measurements, not only to the reported EER, as introduced in previous Section, but also to the full set of biomedical measures: accuracy, sensitivity, specificity and balanced accuracy, aiming to obtain a more precise experimental description. In addition, we have tested the influence of the particular type of metric (l_1 or l_2) used by the kernel regularizer. Overall, one expects that l_1 norm would prefer sparsest solutions for the optimization problem. These results are further explored hereafter.

Classification performance of KG 1

KG 1 is composed of 3 feature kernels: intensity, orientation and Sobel edges, extracted by following the sagittal plane. This plane is commonly used by the radiologist in

the diagnosing workflow since, in general, the MR scanner acquires images in this plane. Table 4-2 presents the classification values for the KG 1 using both l_1 and l_2 regularizers.

		Acc	Sens	Spec	BAC	EER
Group 1	l_1	86.04	65	92.42	78.71	0.76
	l_2	82.56	85	81.82	83.41	0.82
Group 2	l_1	77.78	71.43	79.59	75.51	0.76
	l_2	76.98	75	77.55	76.28	0.76
Group 3	l_1	71.32	67.14	75.76	71.45	0.71
	l_2	72.06	70	74.24	72.12	0.73
Group 4	l_1	70.20	71	69.39	70.19	0.70
	l_2	70.71	70	71.43	70.71	0.70

Table 4-2: Classification performance measures for KG 1 (3 feature kernels) for the different OASIS subject groups, using both l_1 and l_2 regularizations. Best classification performance per each OASIS group is highlighted in bold.

Whilst all l_2 sensitivity values are larger for the first and second experimental groups, this pattern is not observed at all for the l_1 norm. The specificity is nevertheless larger for the l_1 norm and deteriorates from the first to the fourth experimental group, while with the l_2 norm, sensitivity progressively decays from the first group to the fourth. Overall, results show that in all dataset groups, the best values are obtained when using the l_2 regularizer, probably because of the tight correlation between the three information sources so that it results more optimal to combine them. These observations are corroborated by the learned kernel weights presented in Table 4-3.

Feature	Group 1		Group 2		Group 3		Group 4	
	l_1	l_2	l_1	l_2	l_1	l_2	l_1	l_2
Intensity	11.38	1.99	19.08	2.44	20.33	2.05	27.63	3.22
Orientation	7.78	2.05	6.09	2.28	11.94	2.12	13.77	3.11
Sobel edges	0	1.86	0	2.09	0.21	1.91	0	2.83

Table 4-3: Kernel weights learned for KG 1 (3 feature kernels, each in one row) on the different OASIS subject groups, using both l_1 and l_2 regularizations.

Table 4-3 shows the learned kernel weights, per each dataset group, using both regularizers. When using the l_1 norm, information coming from Sobel edges is usually discarded, while the l_2 norm uniformly weights the three sources. In this case, it can be observed that in Groups 1 and 3 (with younger subjects) the orientation information

is a little bit more important, whilst in Groups 2 and 4 the intensity information shows the largest weight. Overall, the edge information counts less in either approach.

Classification performance of KG 2

KG 2 is composed of 9 feature-plane kernels: intensity, orientation and Sobel edges, each extracted following the three planes: coronal, sagittal and axial. These features have been included to investigate the discriminability of the information calculated from the three different views and their relative importance when they are integrated within a single strategy. Table 4-4 presents the classification metrics obtained for the KG 2, in terms of the l_1 and l_2 regularizers.

		Acc	Sens	Spec	BAC	EER
Group 1	l_1	87.21	80	89.39	84.70	0.82
	l_2	84.88	90	83.33	86.67	0.83
Group 2	l_1	73.81	82.14	71.43	76.79	0.73
	l_2	77.78	64.29	81.63	72.96	0.72
Group 3	l_1	72.79	75.71	69.70	72.71	0.74
	l_2	72.79	74.29	71.21	72.75	0.74
Group 4	l_1	68.69	66	71.43	68.71	0.67
	l_2	69.19	71	67.35	69.17	0.69

Table 4-4: Classification performance measures for KG 2 (9 feature-plane kernels) for the different OASIS subject groups, using both l_1 and l_2 regularizations. Best classification performance per each OASIS group is highlighted in bold.

The obtained sensitivity follows no pattern at all, the value is very good for the first experimental group with both norms, but abruptly falls down for the second group when using the l_2 norm. Notice this value is very similar in case of the l_1 norm for the first and second groups, and slightly degrades for the third and fourth groups, with a larger sensitivity value for the l_2 norm of the last group. These values are consistent with the obtained accuracy since high values indicate a high capability of detecting each of the two classes in the set of four groups. In case of this kernel, in spite of the larger number of information sources, it was not possible to identify which regularizer performs the best in all dataset groups. For Groups 1 and 4, the BAC and the EER indicate that the best choice is the l_2 norm, however, performance difference is really small for Group 3. In contrast, the best values are obtained using the l_1 norm for Group 2. The learned weights for the nine kernels are presented in Table 4-5.

Unlike the previous kernel group, the number of used feature-plane kernels allows in this case to identify the sparsity introduced by the l_1 norm regularizer, contrasting

Feature	Plane	Group 1		Group 2		Group 3		Group 4	
		l_1	l_2	l_1	l_2	l_1	l_2	l_1	l_2
Intensity	Sagittal	0	1.26	4.23	1.44	0	1.27	9.59	1.73
	Coronal	5.08	1.33	3.47	1.45	0	1.31	4.23	1.72
	Axial	0	1.23	0	1.34	0	1.23	0	1.59
Orientation	Sagittal	8.09	1.29	7.14	1.45	0	1.27	14.79	1.75
	Coronal	1.95	1.29	2.33	1.43	8.66	1.45	8.73	1.79
	Axial	0	1.22	0	1.34	0	1.20	0	1.56
Sobel edges	Sagittal	3.14	1.28	5.59	1.45	0	1.26	2.40	1.69
	Coronal	0	1.23	0	1.38	0	1.23	0	1.66
	Axial	0	1.22	0	1.37	0	1.20	0	1.59

Table 4-5: Kernel weights learned for KG 2 (9 feature-plane kernels, each in one row) on the different OASIS subject groups, using both l_1 and l_2 regularizations.

with the similar weights learned by the l_2 norm regularizer. With the l_1 norm, it is observed that the axial plane is completely ignored, while the sagittal plane is the one with the largest weights in almost all features, thereby suggesting that relevant information is captured from this single acquisition plane (as it is commonly done by expert radiologists). In case of the l_2 norm, the axial plane has always the lowest weights, while the sagittal and coronal planes reach higher values. In Groups 2 and 4 (when oldest subjects are included), the observed similar weighting pattern gives more priority to the sagittal and coronal planes. In Group 3, it seems that the orientation information at the coronal plane is the most relevant to discriminate normal controls from patients with probable mild and very mild AD.

Classification performance of KG 3

KG 3 is composed of 18 feature-scale kernels: intensity, orientation (at 4 different angles) and Sobel edges, each extracted at 1/4, 1/8 and 1/16 of the original volume size; all of them picked from the sagittal plane. Information from multiple scales has been introduced here to investigate the influence of the image scale-space representation, i.e., a one-dimensional family of images with various levels of spatial scales. Additionally, the orientation feature has been disaggregated into 4 different angles (0° , 45° , 90° and 135°) to better identify the preferential orientations, aiming to detect volumetric changes into the transversal and diagonal planes. The classification performance obtained with the proposed method for the KG 3 is presented in Table 4-6.

The first and third groups are better separated by the l_1 norm in terms of sensitivity while the first two groups are better classified regarding specificity with the same norm,

		Acc	Sens	Spec	BAC	EER
Group 1	l_1	86.05	85	86.36	85.68	0.86
	l_2	82.56	80	83.33	81.67	0.80
Group 2	l_1	80.16	75	81.63	78.32	0.79
	l_2	76.19	75	76.53	75.77	0.76
Group 3	l_1	75.74	87.14	63.64	75.39	0.76
	l_2	76.47	82.86	69.70	76.28	0.74
Group 4	l_1	70.20	67	73.47	70.23	0.69
	l_2	69.20	70	68.37	69.18	0.68

Table 4-6: Classification performance measures for KG 3 (18 feature-scale kernels) for the different OASIS subject groups, using both l_1 and l_2 regularizations. Best classification performance per each OASIS group is highlighted in bold.

indicating that the discriminant capability of this kernel is easily adaptable, yet this ability decreases when groups are mixed and the classifier is not class-specific anymore. Specificity is nevertheless better for the l_2 norm in the third experimental group, but this is reversed for the fourth group, showing that when groups are very different, the anatomical and pathological variability end up by hiding the characteristic pattern of each group.

This set of measurements shows in general that the multiscale analysis improves the classification rates for every group and that the learning strategy effectively chooses those characteristics with more discriminant power. The larger number of information sources allowed to more precisely find the particular kernels to be combined. This was evidenced in the classification results since best values were obtained using the l_1 norm regularizer for the whole set of experimental groups. Nevertheless, as long as the group size increases and the experimental group includes both very mild versions of the disease and elderly subjects, the difference between the results obtained with both norms decreases. This sparsity condition emerges when increasing the number of analysis sources and introducing the l_1 norm regularizer, thereby facilitating the search of the particular anatomical areas with more important differences. The sparsity of the found solutions can be verified by looking at the learned kernel weights in Table 4-7.

This table shows that the relevant information for discriminating AD patients from NC subjects is mainly due to the 0° , 45° and 135° orientations at the three different scales. Interestingly, intensity information, as well as Sobel edges and 90° orientations are always disregarded. Basically, this amounts to consider horizontal and diagonal changes of a bi-dimensional brain view in a multi-scale analysis, a topic on which we

Feature	Scale	Group 1		Group 2		Group 3		Group 4	
		l_1	l_2	l_1	l_2	l_1	l_2	l_1	l_2
Intensity	1/4	0	1.03	0	1.05	0	1.03	0	1.08
	1/8	0	1.04	0	1.05	0	1.03	0	1.09
	1/16	0	1.04	0	1.07	0	1.04	0	1.12
Orientation 0°	1/4	0	1.06	0	1.09	0	1.05	0	1.19
	1/8	0.22	1.06	0.27	1.09	0	1.05	0	1.15
	1/16	1.97	1.08	0.57	1.10	2.66	1.09	2.46	1.18
Orientation 45°	1/4	4.67	1.08	6.46	1.13	2.26	1.06	1.50	1.20
	1/8	0	1.06	0.62	1.10	0	1.06	1.65	1.19
	1/16	0	1.04	0	1.08	0	1.04	0	1.14
Orientation 90°	1/4	0	1.07	0	1.10	0	1.05	0	1.18
	1/8	0	1.06	0	1.09	0	1.05	0	1.16
	1/16	0	1.05	0	1.08	0	1.05	0	1.14
Orientation 135°	1/4	4.30	1.08	6.68	1.13	10.68	1.08	4.63	1.23
	1/8	0.60	1.06	0.20	1.09	1.37	1.07	1.17	1.18
	1/16	0	1.05	0	1.07	0	1.05	0	1.14
Sobel edges	1/4	0	1.04	0	1.06	0	1.04	0	1.09
	1/8	0	1.03	0	1.04	0	1.03	0	1.08
	1/16	0	1.03	0	1.03	0	1.02	0	1.07

Table 4-7: Kernel weights learned for KG 3 (18 feature-scale kernels, each in one row) on the different OASIS subject groups, using both l_1 and l_2 regularizations

will come back later in this paper.

4.5.2 Anatomical Patterns and Saliency

Several studies, by taking structural MRI volumetric measurements of specific brain regions, have demonstrated significant differences between patients with probable AD and normal controls. These studies have reported several signs of the disease progression by using exclusively anatomical relationships, i.e., from Mild Cognitive Impairment (MCI) to AD. In particular, it has been described that AD early stages are characterized by premature atrophic changes along the perforant hippocampal pathway, including the entorhinal cortex (a portion of the anterior parahippocampal gyrus), the hippocampus and the posterior cingulate cortex [50], and, in some cases, along the banks of the superior temporal sulcus [70]. Later on, larger regions such as the temporal, parietal and frontal neocortices may also be compromised and exhibit neuronal loss. In terms of the disease progression, Frisoni et al. [50] have established the morphometric protocol when measuring the whole-brain, paying particularly attention to the entorhinal cor-

tex, hippocampus and temporal lobe volume estimations, as well as to the associated ventricular enlargement percentages.

The anatomical analysis carried out with the discrimination relevance maps, as described in Section 4.4.4, remarkably agrees with these very known anatomical findings [50, 70]. Figure 4-6 illustrates, per each OASIS subject group, some sagittal slices of the corresponding discrimination relevance maps constructed upon the classification model trained using the KG 3 and the l_1 norm as regularizer. Those regions, systematically similar in case of pathological condition, are drawn in red, while those areas that remain similar for the normal brains, are colored in blue, all of them overlaid upon a structural brain MR image. The obtained relevance patterns are consistent through the four experimental groups, with a high coincidence level of discriminant regions, as observed in Figure 4-6. Notice that most important structures in terms of differences are mainly located at the subcortical level, specially the hippocampus cortex. However, as long as elderly patients and different AD conditions and evolutions are included into the groups, the discrimination capability of the relevant regions diminishes, as expected and broadly documented in several clinical studies [50, 70].

By registering each of the maps presented in Figure 4-6 to the Harvard-Oxford cortical and subcortical atlases, the discrimination values per region can be better identified. As described in Section 4.4.4, the relevance maps contain both positive and negative values, associated to the AD and NC classes, respectively. Then, using the atlases, the largest positive and negative relevance values enclosed within each anatomical region are then selected for the two different classes, producing a quantitative indicator that allows to identify those regions that enclose the most relevant patterns for discrimination of the AD and NC classes. Thus, this procedure can deliver high relevance values for both classes simultaneously, leading to find the same anatomical region as relevant for classification for both AD and NC classes.

Separation between cortical and subcortical regions is performed with these two atlas to facilitate the analysis and visualization of the relevance values. Figure 4-7 presents some sagittal slices of the relevance values assigned to each cortical and subcortical region for all OASIS subject groups. Notice that some regions appear systematically as AD discriminant for the four experimental groups, namely the anterior division of the left and right parahippocampal gyrus (left and right entorhinal cortices), the left precuneous cortex, the left and right amygdalas and the left hippocampus. The discriminant level is obviously not the same but this difference can be used to characterize different levels of AD, a side result of the presented method. In contrast, normal subjects are mainly discriminated with very different regions, including the temporooccipital part of the right middle temporal gyrus, the inferior division of the right lateral occipital cortex, the right angular gyrus and the right lateral ventricle. The differences in the

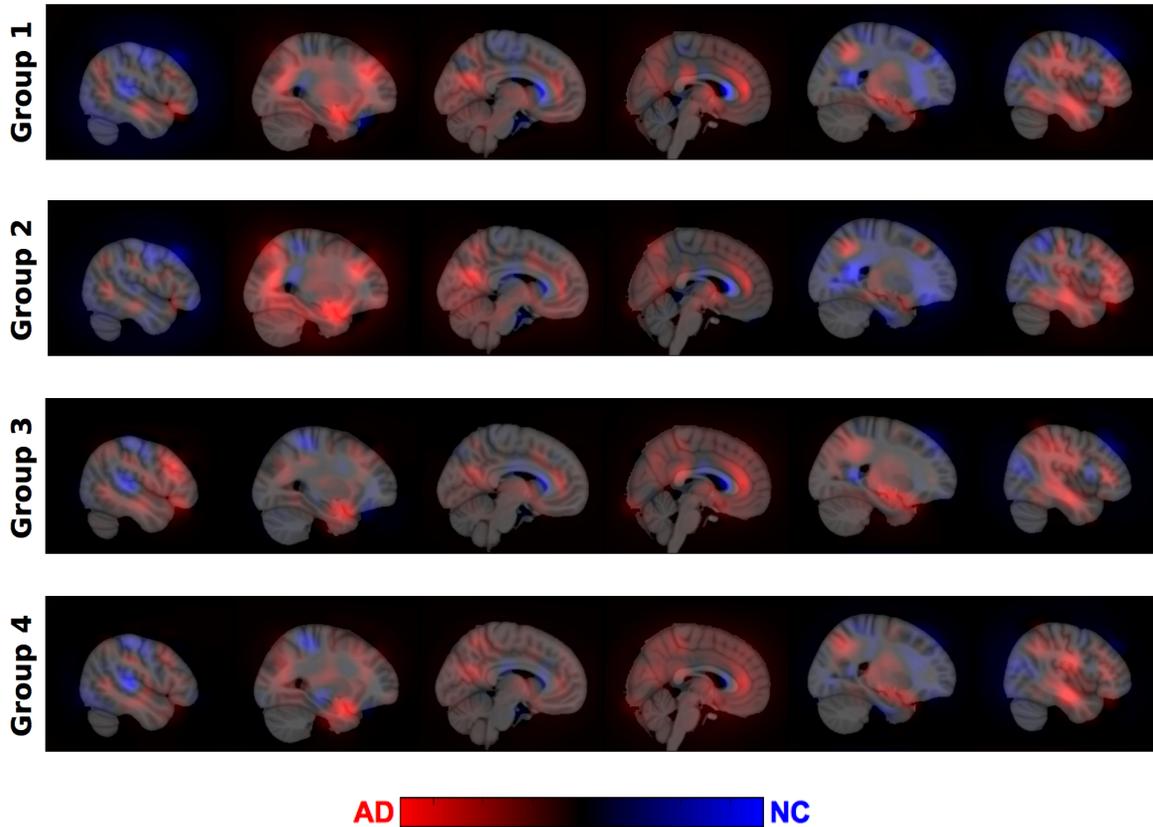


Figure 4-6: Relevance maps for AD classification. Rows: OASIS subject groups, Columns: selected sagittal slices (same slice for all groups). Blue regions are associated to NC class and red regions are associated to AD class.

relevant patterns found among the four experimentation groups may be attributed to two different factors: the presence or absence of elderly subjects and the different stages of the disease. In the first case, it is possible to identify, for example, certain correlations between the patterns of Groups 1 and 3, those ones that do not include very elderly subjects, in particular for those regions relevant for NC discrimination. In the second case, the relevance values of specific regions, used for AD discrimination, decay as long as different stages of the disease are included in the classification groups, for example when comparing the AD relevant regions in Groups 1 and 3.

The tessellation of the discrimination relevance maps allows to quantify directly the importance of each anatomical region in the classification. Different patterns can be obtained by setting the number of regions responsible for a given percentage of the found differences, value which defines also the discrimination degree of these regions. The regions with higher relevance values are then identified per each class and compared

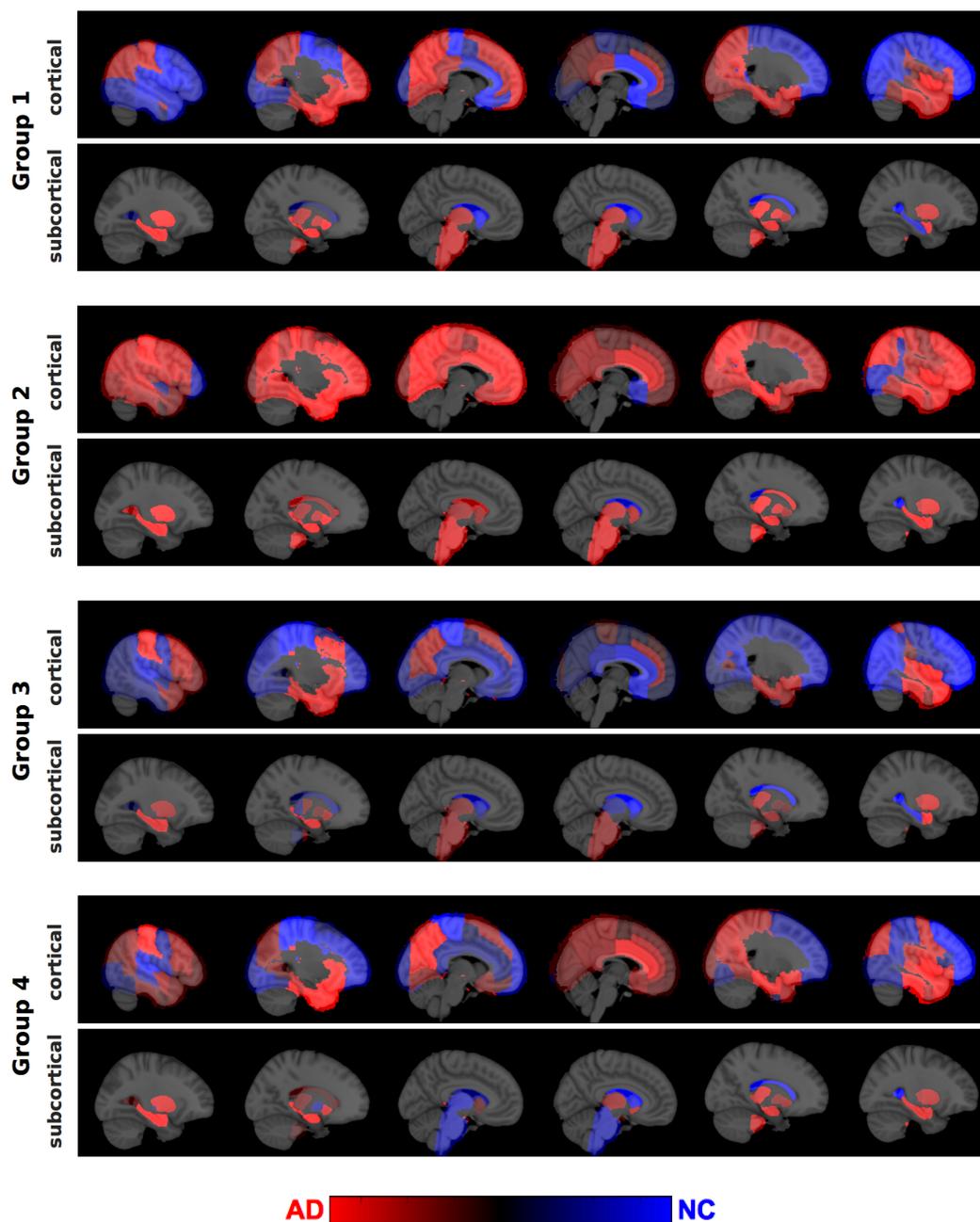


Figure 4-7: Anatomical relevance maps for AD classification. Rows: OASIS subject groups and structural regions (cortical and subcortical), Columns: selected sagittal slices (same cortical or subcortical slice for all groups). Blue regions are associated to NC class and red regions are associated to AD class.

through the different configurations tested. In this case, up to the eight most relevant cortical structures and the six most relevant subcortical structures, responsible for about the 60% of the found differences, are selected, however, in some cases the number of relevant subcortical structures are less than six, revealing that only few regions encode the largest percentage of differences. Tables 4-8 to 4-11 present the ID numbers of the cortical and subcortical regions identified per each OASIS group, kernel group and subject class, with the regularizer that brings the best classification results at each kernel group. The corresponding anatomical region names are provided in Tables D-1 and D-2 in the Appendix D.

		Class	Relevant Regions							
Group 1	Cortical	AD	C33	C27	C24	C47	C31	C34	C77	C8
		NC	C29	C77	C60	C58	C61	C68	C27	C52
	Subcortical	AD	S10	S12	S13	S11	S18	S26		
		NC	S43	S4						
Group 2	Cortical	AD	C65	C55	C33	C77	C29	C90	C54	C19
		NC	C58	C60	C68	C79	C61	C84	C72	C55
	Subcortical	AD	S10	S12						
		NC	S43	S4	S50	S16				
Group 3	Cortical	AD	C77	C29	C34	C50	C82	C19	C20	C43
		NC	C77	C29	C54	C53	C55	C49	C52	C79
	Subcortical	AD	S18	S10	S12	S54	S17	S53		
		NC	S43	S4	S50	S58				
Group 4	Cortical	AD	C77	C29	C90	C19	C34	C55	C65	C54
		NC	C79	C72	C58	C71	C84	C95	C77	C70
	Subcortical	AD	S18	S17	S10	S12	S53	S51		
		NC	S43	S4	S50	S58				

Table 4-8: List of relevant cortical and subcortical regions highlighted by the classification process using a single kernel, discriminated for the different OASIS subject groups. Red rows correspond to regions relevant for AD class and blue rows corresponds to regions relevant for NC class. Correspondant anatomical areas are provided in Tables D-1 and D-2 in the Appendix D.

Relevant cortical and subcortical regions, shown in Tables 4-8 to 4-11, are consistent with what has been described in clinical studies, previously discussed at the beginning of this Section. When considering the whole set of four groups and the different analysis kernels, discrimination of probable AD patients was mainly due to specific anatomical regions, namely the anterior division of the right cingulate gyrus, the anterior division of the right parahippocampal gyrus (entorhinal cortex), the left intracalcarine cortex, the anterior division of the left cingulate gyrus, the right insular cortex and the right postcentral gyrus; while the most used subcortical regions were the left and right putamen, the left and right hippocampus, the left thalamus, caudate and pallidum, and the

		Class	Relevant Regions							
Group 1	Cortical	AD	C19	C10	C29	C77	C46	C65	C20	C24
		NC	C7	C54	C61	C75	C60	C27	C4	C52
	Subcortical	AD	S13	S10	S12	S11	S52	S18		
		NC	S4	S43	S16	S54				
Group 2	Cortical	AD	C65	C29	C77	C55	C19	C10	C28	C46
		NC	C58	C60	C61	C68	C8	C78	C79	C56
	Subcortical	AD	S49	S52						
		NC	S43	S4	S16	S17				
Group 3	Cortical	AD	C65	C67	C19	C68	C17	C46	C10	C84
		NC	C54	C53	C55	C52	C49	C29	C77	C1
	Subcortical	AD	S10	S11	S13	S53	S12	S43		
		NC	S4	S43						
Group 4	Cortical	AD	C19	C65	C77	C29	C46	C10	C45	C2
		NC	C58	C67	C53	C54	C60	C94	C8	C49
	Subcortical	AD	S10	S12	S53	S49	S51	S17		
		NC	S43	S4	S50					

Table 4-9: List of relevant cortical and subcortical regions highlighted by the classification process using the KG 1 (3 feature kernels), discriminated for the different OASIS subject groups. Red rows correspond to regions relevant for AD class and blue rows corresponds to regions relevant for NC class. Correspondant anatomical areas are provided in Tables D-1 and D-2 in the Appendix D.

left and right amygdala. On the other hand, cortical regions systematically unchanged in the group of normal subjects were the left subcallosal cortex, the right frontal pole, the right middle frontal gyrus, the pars triangularis and opercularis of the right inferior frontal gyrus, the right precentral gyrus, the posterior division of the right superior temporal gyrus and the right precuneous cortex; while the subcortical regions involved were basically the left and right lateral ventricles, the brainstem and the right caudate. These findings indicate that frontal and temporal cortex and global subcortical structures are fundamentally altered in AD patients, presenting a complex composition of different levels of local alterations which are therefore difficult to characterize and manage.

In Table 4-8, the regions that always appear as discriminant, using a single kernel, include the anterior division of the right cingulate gyrus, the left putamen and the left thalamus for AD class, and the left and right lateral ventricles for NC class. The left frontal cortex can be considered possibly related with presence of mild AD (CDR = 1), as it appears as discriminant in Groups 1 and 2. In contrast, left and right hippocampus are found to be relevant for discrimination when mixed stages of AD are present (Groups 3 and 4). Similar regional patterns are found for Groups 2 and

		Class	Relevant Regions							
Group 1	Cortical	AD	C96	C69	C70	C68	C48	C59	C66	C62
		NC	C34	C56	C82	C85	C75	C8	C14	C27
	Subcortical	AD	S10	S13	S53	S12	S49	S11		
		NC	S16	S53	S54	S17	S43	S4		
Group 2	Cortical	AD	C22	C23	C77	C29	C25	C28	C68	C69
		NC	C85	C82	C86	C70	C71	C52	C31	C47
	Subcortical	AD	S52	S49	S51	S16				
		NC	S16	S4	S54	S53	S17			
Group 3	Cortical	AD	C68	C65	C66	C69	C67	C50	C82	C17
		NC	C33	C8	C1	C53	C52	C49	C54	C55
	Subcortical	AD	S10	S13	S11	S52	S54	S53		
		NC	S16	S43	S4					
Group 4	Cortical	AD	C65	C18	C17	C68	C67	C66	C70	C52
		NC	C49	C1	C33	C79	C95	C8	C29	C58
	Subcortical	AD	S49	S16	S53	S52	S10	S51		
		NC	S43	S16	S50	S17	S18	S4		

Table 4-10: List of relevant cortical and subcortical regions highlighted by the classification process using the KG 2 (9 feature-plane kernels), discriminated for the different OASIS subject groups. Red rows correspond to regions relevant for AD class and blue rows corresponds to regions relevant for NC class. Correspondant anatomical areas are provided in Tables D-1 and D-2 in the Appendix D.

4 (same regions marked as relevant), suggesting that those regions are responsible for the increasing in classification performance of the single kernel configuration, when compared to the FBM framework (See Table 4-1). Regions such as the anterior division of the left and right cingulate gyrus and the right precentral gyrus appears in Table 4-8 as simultaneously discriminant for AD and NC classes, suggesting that within these regions there exist distinctive relevance patterns for both classes.

For the first kernel group (KG 1, Table 4-9) similar regional results are found for the NC class, given that the left and right lateral ventricles are those that appear as discriminant in the four OASIS groups, while for discrimination of AD class the most relevant regions include the posterior division of the left superior temporal gyrus, the anterior division of the left supramarginal gyrus, the left planum temporale and the right postcentral gyrus. In OASIS groups with mainly mild AD patients (Groups 1 and 2), the right pallidum always appears as discriminant, possibly relating this region to early stages of AD; while in those groups that include different AD stages (Groups 3 and 4), the right hippocampus appears as the most discriminant for the disease. The only region with simultaneous relevant AD and NC patterns is the right lateral ventricle, which appears in Group 3.

		Class	Relevant Regions							
Group 1	Cortical	AD	C77	C76	C50	C82	C8	C34	C24	C47
		NC	C27	C49	C52	C29	C79	C95	C77	C72
	Subcortical	AD	S54	S51	S53	S18	S12	S17		
		NC	S4	S43	S50	S11	S26			
Group 2	Cortical	AD	C33	C8	C34	C50	C79	C24	C77	C17
		NC	C79	C95	C72	C78	C84	C58	C94	C27
	Subcortical	AD	S18	S51	S53	S54	S17	S16		
		NC	S43	S4	S50					
Group 3	Cortical	AD	C50	C82	C56	C86	C4	C92	C5	C6
		NC	C49	C52	C79	C78	C51	C84	C77	C17
	Subcortical	AD	S54	S53	S51	S18	S17			
		NC	S43	S50	S4	S53				
Group 4	Cortical	AD	C82	C86	C92	C50	C34	C85	C63	C8
		NC	C17	C18	C49	C52	C94	C68	C79	C91
	Subcortical	AD	S53	S51	S54	S18	S17	S16		
		NC	S43	S50	S53	S4				

Table 4-11: List of relevant cortical and subcortical regions highlighted by the classification process using the KG 3 (18 feature-scale kernels), discriminated for the different OASIS subject groups. Red rows correspond to regions relevant for AD class and blue rows corresponds to regions relevant for NC class. Correspondant anatomical areas are provided in Tables D-1 and D-2 in the Appendix D.

In the second kernel group (KG 2, Table 4-10), only the posterior division of the right supramarginal gyrus appears as the most discriminant for the AD class, while the left lateral ventricle and the brainstem are the most relevant regions for discrimination of NC. In OASIS Groups 3 and 4, where very mild, mild and moderate stages of the AD are present, the most discriminant regions include the left and right postcentral gyrus and the anterior division of the right supramarginal gyrus. The right putamen can be considered as a region possibly relevant for discrimination of AD in elderly patients, located in Groups 2 and 4, while the left caudate and the left pallidum can be related as relevant for “younger” AD subjects (Groups 1 and 3). Combinations of relevant patterns for both classes are found in the brainstem and the right hippocampus, regions that appear at the same time as discriminant for AD and NC in Groups 1, 3 and 4.

Finally, if we center our analysis only in the kernel group that obtained the best classification results (KG 3, Table 4-11), we can see that the right insular cortex, the left and right hippocampus, the left and right amygdalas and the right putamen always appear as discriminant for AD, while the right precuneous cortex, the left and right lateral ventricles and the right caudate always pop out as discriminant for NC in all subject groups. Regions such as the left temporal pole, the left intracalcarine cortex

and the anterior division of the left parahippocampal gyrus (left entorhinal cortex) can be closely related with presence of mild AD, given that they are mainly present for Groups 1 and 2, which enclose only pathological subjects with mild AD ($CDR = 1$). In contrast, the posterior division of the right temporal fusiform cortex and the right planum polare emerge as significant for Groups 3 and 4, where mixed stages of AD (very mild, mild, moderate) are present, suggesting that changes in these structures can generally indicate presence of the pathology. The right hippocampus is the only region in Table 4-11 present in Groups 3 and 4 as relevant for both AD and NC classes.

4.6 Discussion

This paper has introduced a fully automatic strategy that reveals structural brain patterns associated to the presence of the Alzheimer's disease in a public dataset of brain MR images. The underlying idea behind this proposal is that it is possible to find the discriminant patterns that an expert clinician might discover in similar images. This is accomplished using a fusion strategy that mixes together bottom-up and top-down information flows, achieving accurate classifications of probable AD patients or healthy controls. The bottom-up representation is given by a visual saliency method that automatically highlights relevant regions correlated with the AD diagnosis, using contributions from different multi-scale visual features. On the other hand, the top-down scheme allows to adaptively select the meaningful part of the representation, identifying patterns associated to pathological stages. The whole strategy allows to find anatomical regions with clinical meaning that can be quantitatively related to the diagnosis, and therefore, may be suitable for an objective graduation and understanding of the different AD stages.

Morphometrical analysis of groups of subjects, for identifying discriminant patterns associated to diverse pathologies (mainly neurodegenerative diseases), is currently a wide and active research area [7]. As mentioned in the Introduction (Section 4.1), information coming from voxel intensities (VBM), deformation fields (DBM), spatial locations and 3D-reconstructed boundaries (SBM) is commonly used for statistical identification of anatomical between-group-differences. However, so far such analyses are only able to establish very localized differences that can not be systematically found at exactly the same place along an experimental group, whereby their clinical meaning is still limited. In addition to their anatomical inconsistency, these analyses can hardly identify complex relations between these local differences. In the search of morphometrical methods, more robust to the anatomical variability and to the lack of one-to-one correspondence between all subjects, the analysis paradigm has evolved from the very local approaches to the identification of distinctive and reproducible patterns,

represented by scale-invariant salient features, along with a probabilistic framework that together permit to evaluate the significance and differentiation degree of salient features. This is the main idea behind the FBM approach [121], which automatically discovers sets of anatomical features which are consistent with clinically established differences between normal controls and probable AD patients. These sets of features, with their appearance and geometric information, are then considered as group-related anatomical patterns, suitable to be used as image biomarkers [121]. Nevertheless, this analysis is really far from a usual diagnostic analysis, in which case anatomophysiological correlations are required to determine the physiopathology of a particular disease. In contrast, the approach herein described can be seen as an improvement to the state-of-the-art analysis of localized salient features and much closer to the clinical interpretation of a pathological finding and therefore more suitable to support any diagnostic decision. By combining local and global visual analyses, the presented method allows to extract anatomic relevant regions and weight their contribution to the differentiation of pathological stages.

The present investigation has included an extensive validation and parameter study, evaluating both its accuracy for discriminating different experimental groups and its capacity of determining the relevant anatomical regions together with their weights. Regarding discriminative power, different parameters involved in the top-down and bottom-up information flows, were assessed in terms of classification accuracy, allowing to identify the influence of the different visual features, acquisition planes and image scales in the final discrimination between AD and NC classes. The simpler version of our proposal (combining a single saliency-based kernel with a SVM learning) has reached an equivalent performance to a state-of-the-art approach (FBM proposed by Toews et al. [121]). Comparisons between three different kernel groups (KG 1, KG 2 and KG 3) and the single kernel approach, have shown that the segregation of information into different feature-scale kernels, improves the classification performance in all OASIS subject groups, reaching an average increment of 8.8% in the EER measure, with respect to what was reported by the FBM technique. The presented approach allows to identify in addition that the most relevant information for AD classification comes from the orientation feature, specially at 0° , 45° and 135° , extracted in the sagittal and coronal planes, and at the three different scales. This result illustrates that the learning technique herein used is able to separately explore the parameter space and to optimally combine or fusion each part of the complete parameter space.

One of the main contributions of this work has been the design of a fully automatic method which is also completely interpretable and consistent with what radiologists diagnose. The optimal fusion of different features and their learned discriminative power, facilitates an objective understanding and localization of pathological differences. In

the first place, the sagittal plane, identified after the proposed method as the most discriminant, is the one mostly used in acquisition and analysis of brain MR images, and also corresponds to the acquisition plane of the herein used OASIS dataset. In the second place, the optimal fusion of the relevant features, through the l_1 or l_2 norm regularizers, brings out at the end those anatomical areas with systematically important differences, with respect to the available feature information. Results shown in Section 4.5.1 reveal that when the information sources are few (such as in kernel groups KG 1 and KG 2), the non-sparse l_2 norm regularizer delivers the best classification results, since it attempts to better combine the small number of available features. In contrast, when the information is more segregated and there exists a large number of features (KG 3), the sparse l_1 norm regularizer selects only a small set of relevant sources, disregarding those that may provide redundant or divergent information for the classification.

In terms of the anatomical analysis, the regions found with the proposed approach as systematically relevant for discrimination of AD patients, include the cingulate gyrus, the anterior division of the parahippocampal gyrus (entorhinal cortex) and subcortical structures such as the putamen, amygdala and hippocampus; results completely coherent to what has been reported by clinical studies of AD [50, 70]. Likewise, this analysis has determined that main discriminative features are orientations, in particular, systematic changes were mainly detected at the horizontal and diagonal directions. Yet the number of cases is not enough as to statistically conclude that this finding is disease-related, similar outcomes have been described in other global neurological disorders. It has been observed for instance that patients with schizophrenia exhibit faster volume decline in regions like the right frontal gray matter and the bilateral posterior superior temporal gray matter, that is to say global changes in anterior or posterior brain regions [90]. In this study we have found that main changes are located in horizontal and diagonal directions, an indirect evidence that changes occur very likely in oriented areas but not precisely located in a particular region. From the clinical perspective, the proposed strategy follows the visual analysis made by radiologists when diagnosing medical images, allowing in addition a quantitative determination of the brain anatomical regions which are different between experimental groups. With an adequate and exhaustive evaluation in larger data sets, containing sufficient examples of the different AD stages, this method can be also used as a second diagnostic opinion in the current clinical practice.

5 Conclusions and Perspectives

This thesis has developed and validated novel techniques and strategies to address the problem of automatically extracting relevant medical information from brain MR images, aiming to improve the accuracy and interpretability of the morphometric measures and comparisons. The thesis has introduced and adapted biologically inspired methods for identification of diagnostic-relevant image regions in a very complex and challenging problem, the Alzheimer’s disease (AD). The automatic strategies herein developed have included prior anatomical and medical knowledge within the morphometrical analysis. The set of proposed tools constitute an innovative framework in the context of anatomical studies: sparse-based representations and visual attention methods, together with machine learning techniques, provide efficient representations of the image content in terms of visual features, leading to the discovery of visual patterns directly related with a specific pathology.

The techniques introduced in this thesis have led to an innovative perspective of the brain morphometric analyses: the clinical interpretability. This condition can be considered as an important requirement of any software package intended to be used in the actual clinical practice, as it assures a simple integration with common diagnostic tasks and routines. In this particular case, the clinical interpretability comes from two important sources: the efficient incorporation of prior medical knowledge and the meaningful information extracted by means of visual features. These sources, combined with biologically inspired techniques, have allowed to identify and infer pathology-related patterns for discrimination of neurological diseases, in particular, the AD. Analysis of these patterns may help to define anatomical biomarkers, useful for diagnosis, prognosis and follow-up. In fact, with a dataset including both normal subjects and patients with different stages of AD, the proposed techniques have shown to be capable of finding anatomical regions with clinical meaning while at the same time they permit to quantitatively relate these findings to the diagnosis, a very promising contribution towards discriminating pathological stages.

5.1 Perspectives

This thesis was developed attempting to contribute to the construction of a computational framework for morphometrical analyses of medical images. It aims to provide a set of analysis, interpretation and visualization tools that serve as a support for diagnosis, training and research processes. For a physician, the possibility of obtaining quantitative and complementary information about the patient condition, allows to refine the medical management, decreases the diagnostic variability and results in more accurate treatments, impacting directly the patient quality-of-life. Also, compilation and comparison of this knowledge within different populations affect positively the development of public and preventative health policies. We expect that the proposed computational tools, in the future, will contribute to these medical advances. In the meanwhile, some work must be performed to prepare these tools to be used by medical experts in a daily basis. Some of this work includes:

1. **Fusion of multimodal information.** For complex diseases such as the Alzheimer's disease, a large set of diagnostic information sources, such as neuropsychological tests, structural images, metabolic images, functional images, genetic information and so on, are currently available. As these sources provide complementary information for the decision making process, a computational framework for diagnosis support could benefit from this multimodality by effectively integrating these diverse sources of information. For each source, a weighting factor should be identified and assigned, according with its importance in the pathological diagnosis, and a preprocessing step should be applied to reduce the amount of information only to a small set of relevant features. Then, all multimodal features should be interpreted in terms of a common reference frame to facilitate further analyses and inferences.
2. **Performance validation with larger and complex data sets.** Evaluation of the proposed techniques have been performed mainly with data sets that are widely available, to facilitate comparison with other approaches and their published results. However, these data sets need not necessarily include enough examples of the different pathological conditions to diagnose, or they do not provide an adequate number of cases to reach significant statistical conclusions. To increase the sensitivity and specificity of clinical studies, and also to serve as a second diagnostic opinion in the current clinical practice, the proposed computational tools must be exhaustively evaluated in larger data sets.
3. **Mining complex visual patterns for diagnosis support.** At conducting group-based analyses of pathological conditions in structural medical images, a

main aim is to find anatomical changes related to functional disturbances, that may result in a complex mixture of local and/or global morphological alterations. This problem can be formulated as the systematic search of visual patterns associated with particular pathological conditions, and the formulation of an adaptive strategy for classification of these patterns according to their clinical meaning. Differential patterns of subject groups can thus serve as a meaningful input to a knowledge extraction framework, applying data mining techniques to infer unknown or latent relations among them, and thus serving as an important information source for diagnosis support.

4. **Towards a new evidence-based medicine perspective.** Medical specialists undergo a long training process to be able to find clinically-relevant patterns defining particular physiological or pathological conditions. However, they can hardly quantify systematic differences, that is to say, to establish the relative importance of the differences and to organize them as a hierarchical structure. A computational strategy capable of automatically extracting pathological-related visual patterns and estimating their quantitative importance can provide the basis for this new perspective in evidence-based medicine, yielding objective estimates of risks and benefits that may aid to strength the medical act in terms of the patient diagnosis, prognosis and treatment.

A Topology Preservation and Homotopic Transformations

Homotopic transformations are topology-preserving procedures that consist of sequentially deleting or adding *simple points*. This operation works only on binary images, such as the pure tissue segmentations, where each voxel is considered as a point. Informally, a simple point of an object X is a point that can be added or removed from X without changing the topological characteristics of X . It is possible to locally characterize simple points in 3D using two *topological numbers* T and Tb [12].

Thus, skipping some technical details, let $A(x)$ be the set of points of $X \setminus \{x\}$ lying in a neighborhood of x , and let $Ab(x)$ be the set of points of the complement of X (background) lying in a neighborhood of x . Then, $T(x)$ (resp. $Tb(x)$) is the number of connected components of $A(x)$ (resp. $Ab(x)$). A point x is simple if and only if $T(x) = Tb(x) = 1$. Topological numbers are useful for classifying points of an object X based on local topological characteristics: for example, a point x such that $Tb(x) > 1$ characterizes a region of the object which separates (locally) its background into several parts.

Based on these notions, given an object X , a subset I of X and a priority function P , Algorithm 2 computes an *homotopic erosion of X constrained by I* , that is, an object that is topologically equivalent to X , that contains I and that has no simple point outside I . In this algorithm, the priority function P is usually chosen as the inverse of the distance to I , in order to select in the first place the points that are farthest to the set I . This choice will be assumed in the remaining operations.

Algorithm 2 Homotopic erosion of X constrained by I
with priority P

repeat

 Select $x \in X \setminus I$ such that $P(x)$ is minimal

if x is simple for X **then**

$X = X \setminus \{x\}$

end if

until stability

Applying Algorithm 2 to the complementary sets of X and I , then inverting the result, yields an *homotopic dilation of X constrained by I* . In a similar way, Algorithm 3 [11] computes a *surface skeleton of X* which contains medial surfaces of the original object (provided that the priority function P is a distance map of X).

Algorithm 3 Surface skeleton of X with priority P

Let C be a null image

repeat

 Select $x \in X$ such that x is simple for X ,

$C(x) == \emptyset$ and $P(x)$ is minimal

$X = X \setminus \{x\}$

for all y in the neighborhood of x **do**

if $Tb(y) > 1$ **then**

$C(y) = 1$

end if

end for

until stability

B The Super-Resolution Generative Model

Consider a $n \times a$ matrix D , where each column is a possible 3D image in \mathbb{R}^n with size $\sqrt[3]{n} \times \sqrt[3]{n} \times \sqrt[3]{n}$ (each image is thus known as an atom) and $a \gg n$, so that D becomes an overcomplete dictionary of atoms. The projection of an image x onto the space spanned by D yields a weighting vector α ($x = D\alpha$). Furthermore, if α is sparse (with $k_0 \ll m$ nonzeros), this produces a linear combination of k_0 atoms with varying weights. Nonzero values are located randomly within the vector, and their values are drawn independently from the zero-mean, τ -variance, Gaussian distribution $\mathcal{N}(0, \tau)$ [44]. A random perturbation is usually introduced, in terms of a noise vector $e \in \mathbb{R}^n$ with bounded power $\|e\|_2 \leq \epsilon$, such that $x = D\alpha + e$. This gives us a probabilistic generative model for images, denoted by $\mathcal{M}(D, k_0, \tau, \epsilon)$ [44].

Suppose we have a 3D image x , which we assume to have been generated by the model $\mathcal{M}(D, k_0, \tau, \epsilon)$, and that the parameters of the model are known. To determine the underlying vector α which generates x (atomic decomposition), we need to solve the problem $\mathcal{P}_0(D, x, \delta)$, which has the form

$$\mathcal{P}_0(D, x, \delta) : \min_{\alpha} \|\alpha\|_0^0 \text{ subject to } \|x - D\alpha\|_2 \leq \delta$$

The solution to this problem consists in finding the sparsest vector α that weights x as a linear combination of atoms from D with an error no larger than δ [44].

Unfortunately, this turns out in an NP-hard problem. Given that the base equation $x = D\alpha$ is undetermined for the unknown coefficients α , solving this problem amounts to a combinatorial optimization process. Different alternative techniques for approximating the solution to this problem have been recently proposed, detailed descriptions and references can be found in [16]. One of these approaches is based on relaxing (convexification) the sparsity restriction, so that the ℓ_0 penalty is replaced by the ℓ_1 penalty, whereby the $\mathcal{P}_0(D, x, \delta)$ problem is read as:

$$\mathcal{P}_1(D, x, \delta) : \min_{\alpha} \|\alpha\|_1^1 \text{ subject to } \|x - D\alpha\|_2 \leq \delta \tag{B-1}$$

This problem is currently known in the literature as *Basis Pursuit Denoising* (BPDN) [16, 22]. Furthermore, these two norms are shown to be equivalent for "sufficiently sparse" solutions [42].

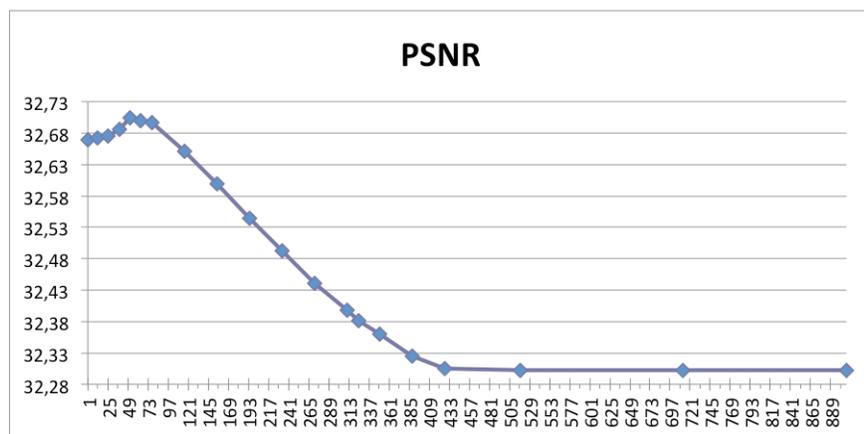
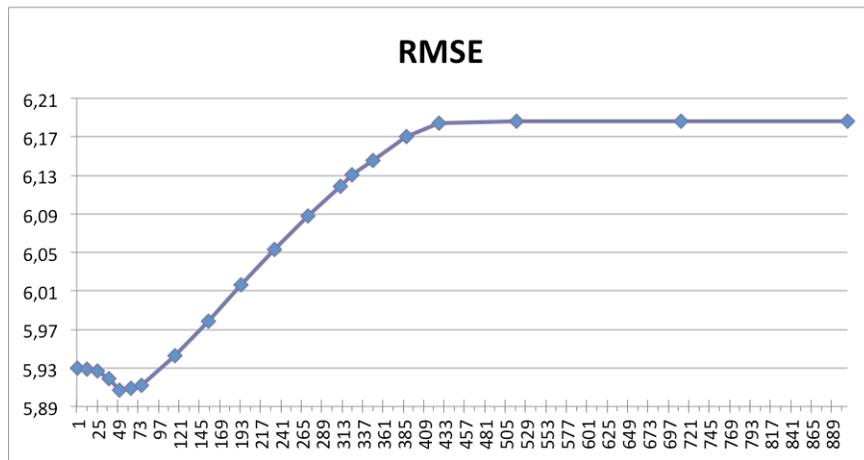
For an appropriate Lagrange multiplier λ (function of D , x and δ), the solution to the following problem

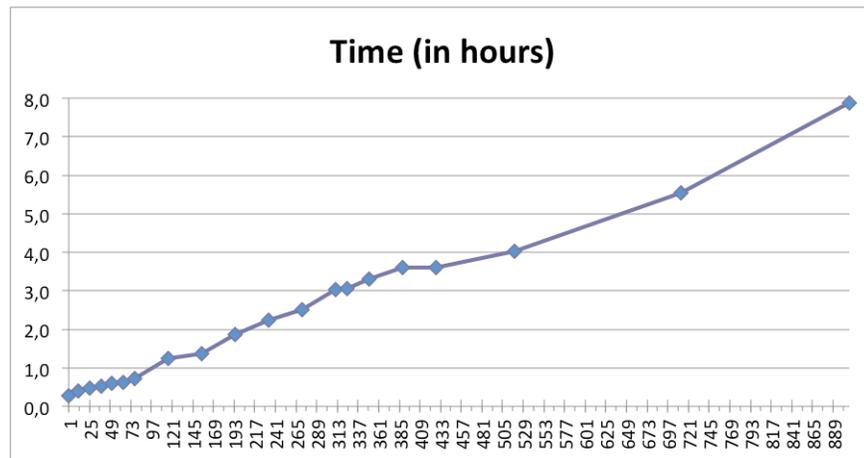
$$\mathcal{G}_1(D, x, \lambda) : \min_{\alpha} \lambda \|\alpha\|_1 + \frac{1}{2} \|x - D\alpha\|_2^2 \quad (\text{B-2})$$

is equivalent to the solution of $\mathcal{P}_1(D, x, \delta)$ [16].

C Influence of Dimensionality Reduction in Super-Resolution

We have performed a simple experiment using only one image of the base data set, and applying the proposed super-resolution approach using different amounts of PCA coefficients, to illustrate the performance of our proposal in terms of RMSE, PSNR and reconstruction time. The following graphs shows the behavior of each measure with respect to the PCA selected coefficients, which ranges between 1 and 907, corresponding to up to 70% of the original dimensions.





These figures shows that although the optimum value is around 50 coefficients, the difference between using 50 coefficients and 12 (which is the amount that we have selected, corresponding to 1% of the original dimensions) in terms of RMSE is about 0.02 and in PSNR is about 0.03 dB, difference that can be considered as negligible. However, in terms of reconstruction time, the difference between reconstructing with 50 coefficients and 12 coefficients is an increment in about 0.2 hours (12 minutes), representing an increment of 150% in the reconstruction time for this particular image. This is the reason why we have chosen to work with 1% of the PCA coefficients (12 coefficients), which also explains 90% of the original variance.

D Cortical and Subcortical Regions in Harvard-Oxford Atlases

Tables D-1 and D-2 provides the correspondant names per each anatomical region code presented in Tables 4-8 to 4-11.

C1	Left Frontal Pole	C53	Right Inferior Frontal Gyrus, pars triangularis
C2	Left Insular Cortex	C54	Right Inferior Frontal Gyrus, pars opercularis
C4	Left Middle Frontal Gyrus	C55	Right Precentral Gyrus
C5	Left Inferior Frontal Gyrus, pars triangularis	C56	Right Temporal Pole
C6	Left Inferior Frontal Gyrus, pars opercularis	C58	Right Superior Temporal Gyrus, posterior division
C7	Left Precentral Gyrus	C59	Right Middle Temporal Gyrus, anterior division
C8	Left Temporal Pole	C60	Right Middle Temporal Gyrus, posterior division
C10	Left Superior Temporal Gyrus, posterior division	C61	Right Middle Temporal Gyrus, temporooccipital part
C14	Left Inferior Temporal Gyrus, anterior division	C62	Right Inferior Temporal Gyrus, anterior division
C17	Left Postcentral Gyrus	C65	Right Postcentral Gyrus
C18	Left Superior Parietal Lobule	C66	Right Superior Parietal Lobule
C19	Left Supramarginal Gyrus, anterior division	C67	Right Supramarginal Gyrus, anterior division
C20	Left Supramarginal Gyrus, posterior division	C68	Right Supramarginal Gyrus, posterior division
C22	Left Lateral Occipital Cortex, superior division	C69	Right Angular Gyrus
C23	Left Lateral Occipital Cortex, inferior division	C70	Right Lateral Occipital Cortex, superior division
C24	Left Intracalcarine Cortex	C71	Right Lateral Occipital Cortex, inferior division
C25	Left Frontal Medial Cortex	C72	Right Intracalcarine Cortex
C27	Left Subcallosal Cortex	C75	Right Subcallosal Cortex
C28	Left Paracingulate Gyrus	C76	Right Paracingulate Gyrus
C29	Left Cingulate Gyrus, anterior division	C77	Right Cingulate Gyrus, anterior division
C31	Left Precuneous Cortex	C78	Right Cingulate Gyrus, posterior division
C33	Left Frontal Orbital Cortex	C79	Right Precuneous Cortex
C34	Left Parahippocampal Gyrus, anterior division	C82	Right Parahippocampal Gyrus, anterior division
C43	Left Parietal Operculum Cortex	C84	Right Lingual Gyrus
C45	Left Heschl's Gyrus (includes H1 and H2)	C85	Right Temporal Fusiform Cortex, anterior division
C46	Left Planum Temporale	C86	Right Temporal Fusiform Cortex, posterior division
C47	Left Supracalcarine Cortex	C90	Right Central Opercular Cortex
C48	Left Occipital Pole	C91	Right Parietal Operculum Cortex
C49	Right Frontal Pole	C92	Right Planum Polare
C50	Right Insular Cortex	C94	Right Planum Temporale
C51	Right Superior Frontal Gyrus	C95	Right Supracalcarine Cortex
C52	Right Middle Frontal Gyrus	C96	Right Occipital Pole

Table D-1: Names of relevant cortical regions listed in Tables 4-8 to 4-11.

S4	Left Lateral Ventricle	S43	Right Lateral Ventricle
S10	Left Thalamus	S49	Right Thalamus
S11	Left Caudate	S50	Right Caudate
S12	Left Putamen	S51	Right Putamen
S13	Left Pallidum	S52	Right Pallidum
S16	Brain-Stem	S53	Right Hippocampus
S17	Left Hippocampus	S54	Right Amygdala
S18	Left Amygdala	S58	Right Accumbens
S26	Left Accumbens		

Table D-2: Names of relevant subcortical regions listed in Tables 4-8 to 4-11.

Bibliography

- [1] ACOSTA, O. ; BOURGEAT, P. ; ZULUAGA, M. A. ; FRIPP, J. ; SALVADO, O. ; OURSELIN, S.: Automated voxel-based 3D cortical thickness measurement in a combined lagrangian-eulerian PDE approach using partial volume maps. In: *Medical Image Analysis* 13 (2009), S. 730–743
- [2] ADLURU, G. ; MCGANN, C. ; SPEIER, P. ; KHOLMOVSKI, E.G. ; SHAABAN, A. ; DiBELLA, E.V.R.: Acquisition and Reconstruction of Undersampled Radial Data for Myocardial Perfusion Magnetic Resonance Imaging. In: *Journal of Magnetic Resonance Imaging* 29 (2009), Nr. 2, S. 466–473
- [3] ALVAREZ-LINERA, J. ; HERNANDEZ-TAMAMES, J.A. ; DOBATO, J.L. ; DIAZ, M.C. ; HERRERA, A.G. Seco d. ; GARCIA, S. ; PERAITA, H.: Neuropsychological and imaging (MR) biomarkers in the early detection of mild cognitive impairment (MCI). In: *3rd World Congress on Controversies in Neurology (CONy)*, 2009
- [4] ASHBURNER, J.: A fast diffeomorphic image registration algorithm. In: *NeuroImage* 38 (2007), Nr. 1, S. 95–113
- [5] ASHBURNER, J. ; FRISTON, K.J.: Voxel-based morphometry: the methods. In: *Neuroimage* 11 (2000), Nr. 6, S. 805–821
- [6] ASHBURNER, J. ; HUTTON, C. ; FRACKOWIAK, R. ; JOHNSRUDE, I. ; PRICE, C. ; FRISTON, K. u. a.: Identifying global anatomical differences: deformation-based morphometry. In: *Human Brain Mapping* 6 (1998), Nr. 5-6, S. 348–357
- [7] ASHBURNER, J. ; KLÖPPEL, S.: Multivariate models of inter-subject anatomical variability. In: *Neuroimage* 56 (2011), Nr. 2, S. 422–439
- [8] BARLOW, H.: Possible principles underlying the transformations of sensory messages, MIT Press, 1961, S. 217–234
- [9] BAZIN, P.-L. ; PHAM, D. L.: Topology correction using fast marching methods and its application to brain segmentation. In: *Proceedings of the Medical Image Computing and Computer-Assisted Intervention Conference*, Springer Verlag, 2005, S. 484–491

-
- [10] BAZIN, P.-L. ; PHAM, D. L.: Topology-preserving tissue classification of magnetic resonance brain images. In: *IEEE Transactions on Medical Imaging* 26 (2007), Nr. 4, S. 487–496
- [11] BERTRAND, G. ; COUPRIE, M.: Transformations topologiques discrètes. In: COEURJOLLY, D. (Ed.) ; MONTANVERT, A. (Ed.) ; CHASSERY, J.-M. (Ed.): *Géométrie discrète et images numériques*, Hermès, 2007, S. 187–209
- [12] BERTRAND, G. ; MALANDAIN, G.: A new characterization of three-dimensional simple points. In: *Pattern Recognition Letters* 15 (1994), Nr. 2, S. 169–175
- [13] BESAG, J.: On the statistical analysis of dirty pictures. In: *Journal of the Royal Statistical Society* 48 (1986), S. 259–302
- [14] BEUTEL, J. ; KUNDEL, H.L. ; VAN METTER, R.L.: *Handbook of Medical Imaging. Vol. 1: Physics and Psychophysics*. SPIE Press, 2000
- [15] BROWN, Matthew S. ; MCNITT-GRAY, Michael F.: Medical Image Interpretation. In: SONKA, Milan (Ed.) ; FITZPATRICK, J. M. (Ed.): *Handbook of Medical Imaging, Vol. 2: Medical Image Processing and Analysis*, SPIE Press, 2000, S. 399–445
- [16] BRUCKSTEIN, A.M. ; DONOHO, D.L. ; ELAD, M.: From sparse solutions of systems of equations to sparse modeling of signals and images. In: *SIAM Review* 51 (2009), Nr. 1, S. 34–81
- [17] BUADES, A. ; COLL, B. ; MOREL, J.M.: A non local algorithm for image denoising. In: *IEEE International Conference on Computer Vision and Pattern Recognition* 2 (2005), S. 60–65
- [18] BUCKNER, R.L. ; HEAD, D. ; PARKER, J. ; FOTENOS, A.F. ; MARCUS, D. ; MORRIS, J.C. ; SNYDER, A.Z.: A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. In: *NeuroImage* 23 (2004), Nr. 2, S. 724–738
- [19] CAPEL, D.: *Image mosaicing and super-resolution*. Springer-Verlag, 2004
- [20] CARMİ, E. ; LIU, S. ; ALON, N. ; FIAT, A. ; FIAT, D.: Resolution enhancement in MRI. In: *Magnetic Resonance Imaging* 24 (2006), Nr. 2, S. 133–154

-
- [21] CHANG, Chih-Chung ; LIN, Chih-Jen: LIBSVM: A library for support vector machines. In: *ACM Transactions on Intelligent Systems and Technology* 2 (2011), S. 27:1–27:27. – Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [22] CHEN, S.S. ; DONOHO, D.L. ; SAUNDERS, M.A.: Atomic decomposition by basis pursuit. In: *SIAM Journal of Scientific Computing* 20 (1998), Nr. 1, S. 33–61
- [23] CHINCARINI, Andrea ; BOSCO, Paolo ; CALVINI, Piero ; GEMME, Gianluca ; ESPOSITO, Mario ; OLIVIERI, Chiara ; REI, Luca ; SQUARCIA, Sandro ; RODRIGUEZ, Guido ; BELLOTTI, Roberto u. a.: Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer’s disease. In: *Neuroimage* 58 (2011), Nr. 2, S. 469–480
- [24] CHIVERTON, J. ; WELLS, K.: Adaptive partial volume classification of MRI data. In: *Physics in Medicine and Biology* 53 (2008), Nr. 20, S. 5577–5594
- [25] CHOI, H. ; HAYNOR, D. ; KIM, Y.: Partial volume tissue classification of multi-channel magnetic resonance images - a mixel model. In: *IEEE Transactions on Medical Imaging* 10 (1991), Nr. 3, S. 395–407
- [26] COCOSCO, C. ; KOLLOKIAN, V. ; KWAN, R.-S. ; EVANS, A.: Brainweb: Online interface to a 3D MRI simulated brain database. In: *NeuroImage* 5 (1997), Nr. 4, S. S425
- [27] COLLINS, D.L. ; ZIJDENBOS, A.P. ; KOLLOKIAN, V. ; SLED, J.G. ; KABANI, N.J. ; HOLMES, C.J. ; EVANS, A.C.: Design and construction of a realistic digital brain phantom. In: *IEEE Transactions on Medical Imaging* 17 (1998), Nr. 3, S. 463–468
- [28] CORTES, C. ; VAPNIK, V.: Support-vector networks. In: *Machine Learning* 20 (1995), Nr. 3, S. 273–297
- [29] COSTA, L.F.: Visual saliency and attention as random walks on complex networks. In: *Arxiv preprint physics/0603025* (2006)
- [30] COUPÉ, P. ; HELLIER, P. ; PRIMA, S. ; KERVRANN, C. ; BARILLOT, C.: 3D wavelet subbands mixing for image denoising. In: *International Journal of Biomedical Imaging* 2008 (2008), S. 1
- [31] COUPE, P. ; YGER, P. ; PRIMA, S. ; HELLIER, P. ; KERVRANN, C. ; BARILLOT, C.: An optimized blockwise nonlocal means denoising filter for 3D magnetic

- resonance images. In: *IEEE Transactions on Medical Imaging* 27 (2008), Nr. 4, S. 425–441
- [32] DALE, A. ; FISCHL, B. ; SERENO, M.: Cortical surface-based analysis I: Segmentation and surface reconstruction. In: *NeuroImage* 9 (1999), Nr. 2, S. 179–194
- [33] DAVATZIKOS, Christos ; FAN, Yong ; WU, Xiaoying ; SHEN, Dinggang ; RESNICK, Susan M.: Detection of prodromal Alzheimer’s disease via pattern classification of magnetic resonance imaging. In: *Neurobiology of aging* 29 (2008), Nr. 4, S. 514–523
- [34] DEQUARDO, J.R. ; KESHAVAN, M.S. ; BOOKSTEIN, F.L. ; BAGWELL, W.W. ; GREEN, W.D.K. ; SWEENEY, J.A. ; HAAS, G.L. ; TANDON, R. ; SCHOOLER, N.R. ; PETTEGREW, J.W.: Landmark-based morphometric analysis of first-episode schizophrenia. In: *Biological Psychiatry* 45 (1999), Nr. 10, S. 1321–1328
- [35] DESIKAN, Rahul S. ; SÉGONNE, Florent ; FISCHL, Bruce ; QUINN, Brian T. ; DICKERSON, Bradford C. ; BLACKER, Deborah ; BUCKNER, Randy L. ; DALE, Anders M. ; MAGUIRE, R. P. ; HYMAN, Bradley T. ; ALBERT, Marilyn S. ; KILLIANY, Ronald J.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. In: *NeuroImage* 31 (2006), Nr. 3, S. 968–980
- [36] DOAN, Nhat T. ; LEW, Baldur A. ; LELIEVELDT, Boudewijn ; BUCHEM, Mark A. ; REIBER, Johan H. ; MILLES, Julien: Deformation texture-based features for classification in Alzheimer’s disease. In: *SPIE Medical Imaging International Society for Optics and Photonics*, 2013, S. 866927–866927
- [37] DONOHO, David L. ; HUO, Xiaoming: Uncertainty principles and ideal atomic decomposition. In: *Information Theory, IEEE Transactions on* 47 (2001), Nr. 7, S. 2845–2862
- [38] DONOHO, D.L.: Compressed sensing. In: *IEEE Transactions on Information Theory* 52 (2006), Nr. 4, S. 1289–1306
- [39] DUBOIS, Bruno ; FELDMAN, Howard H. ; JACOVA, Claudia ; CUMMINGS, Jeffrey L. ; DEKOSKY, Steven T. ; BARBERGER-GATEAU, Pascale ; DELACOURTE, André ; FRISONI, Giovanni ; FOX, Nick C. ; GALASKO, Douglas u. a.: Revising the definition of Alzheimer’s disease: a new lexicon. In: *The Lancet Neurology* 9 (2010), Nr. 11, S. 1118–1127

-
- [40] DUCHESNE, Simon ; CAROLI, Anna ; GEROLDI, C ; BARILLOT, Christian ; FRISONI, Giovanni B. ; COLLINS, D L.: MRI-based automated computer classification of probable AD versus normal controls. In: *IEEE Transactions on Medical Imaging* 27 (2008), Nr. 4, S. 509–520
- [41] DUYNCKAERTS, Charles ; DELATOUR, Benoît ; POTIER, Marie-Claude: Classification and basic pathology of Alzheimer disease. In: *Acta neuropathologica* 118 (2009), Nr. 1, S. 5–36
- [42] ELAD, M.: *Sparse and redundant representations: from theory to applications in signal and image processing*. Springer-Verlag, 2010
- [43] ELAD, M. ; AHARON, M.: Image denoising via sparse and redundant representations over learned dictionaries. In: *IEEE Transactions on Image Processing* 15 (2006), Nr. 12, S. 3736–3745
- [44] ELAD, M. ; FIGUEIREDO, M. ; MA, Y.: On the role of sparse and redundant representations in image processing. In: *Proceedings of the IEEE* 98 (2010), Nr. 6, S. 972–982
- [45] FISCHL, B. ; DALE, A.: Measuring the thickness of the human cerebral cortex from magnetic resonance images. In: *Proceedings of the National Academy of Sciences of the United States of America* Bd. 97, 2000, S. 11050–11055
- [46] FISCHL, B. ; SERENO, M. ; DALE, A.: Cortical Surface-Based Analysis II: Inflation, Flattening, and a Surface-Based Coordinate System. In: *NeuroImage* (1999), S. 195–207
- [47] FLORENT SÉGONNE, J. P. ; FISCHL, B.: Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. In: *IEEE Transactions on Medical Imaging* 26 (2007), Nr. 4, S. 518–529
- [48] FREEMAN, L.C.: Centrality in social networks conceptual clarification. In: *Social networks* 1 (1979), Nr. 3, S. 215–239
- [49] FRINTROP, S. ; BACKER, G. ; ROME, E.: Goal-directed search with a top-down modulated computational attention system. In: *Pattern Recognition* (2005), S. 117–124
- [50] FRISONI, G.B. ; FOX, N.C. ; JACK, C.R. ; SCHELTENS, P. ; THOMPSON, P.M.: The clinical use of structural MRI in Alzheimer disease. In: *Nature Reviews Neurology* 6 (2010), Nr. 2, S. 67–77

- [51] FRISTON, K.J. ; ASHBURNER, J.T. ; KIEBEL, S.J. ; NICHOLS, T.E. ; PENNY, W.D.: *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, 2011
- [52] GAO, D. ; VASCONCELOS, N.: Discriminant saliency for visual recognition from cluttered scenes. In: *Advances in Neural Information Processing Systems 17* (2004), Nr. 481-488, S. 1
- [53] GAO, D. ; VASCONCELOS, N.: Bottom-up saliency is a discriminant process. In: *IEEE 11th International Conference on Computer Vision*. IEEE, 2007, S. 1–6
- [54] GARCÍA-SEBASTIÁN, Maite ; SAVIO, Alexandre ; GRAÑA, Manuel ; VILLANÚA, Jorge: On the use of morphometry based features for Alzheimer’s disease detection on MRI. In: *Bio-Inspired Systems: Computational and Ambient Intelligence*. Springer, 2009, S. 957–964
- [55] GERARDIN, Emilie ; CHÉTELAT, Gaël ; CHUPIN, Marie ; CUINGNET, Rémi ; DESGRANGES, Béatrice ; KIM, Ho-Sung ; NIETHAMMER, Marc ; DUBOIS, Bruno ; LEHÉRICY, Stéphane ; GARNERO, Line u. a.: Multidimensional classification of hippocampal shape features discriminates Alzheimer’s disease and mild cognitive impairment from normal aging. In: *Neuroimage* 47 (2009), Nr. 4, S. 1476
- [56] GONZÁLEZ, R.C. ; WOODS, R.E.: *Digital Image Processing*. Prentice Hall, 2008
- [57] GOPALAKRISHNAN, V. ; HU, Y. ; RAJAN, D.: Random walks on graphs for salient object detection in images. In: *Image Processing, IEEE Transactions on* 19 (2010), Nr. 12, S. 3232–3242
- [58] GREENSPAN, H.: Super-Resolution in Medical Imaging. In: *The Computer Journal* 52 (2008), Nr. 1, S. 43–63
- [59] GREENSPAN, H. ; OZ, G. ; KIRYATI, N. ; PELED, S.: MRI inter-slice reconstruction using super-resolution. In: *Magnetic Resonance Imaging* 20 (2002), Nr. 5, S. 437–446
- [60] HAN, X. ; XU, C. ; BRAGA-NETO, U. ; PRINCE, J.: Topology correction in brain cortex segmentation using a multiscale, graph-based algorithm. In: *IEEE Transactions on Medical Imaging* 21 (2002), Nr. 2, S. 109–121
- [61] HAREL, J. ; KOCH, C. ; PERONA, P.: Graph-based visual saliency. In: *Advances in Neural Information Processing Systems* Bd. 19, 2007, S. 545

-
- [62] HERMENT, A. ; ROULLOT, E. ; BLOCH, I. ; JOLIVET, O. ; DE CESARE, A. ; FROUIN, F. ; BITTOUN, J. ; MOUSSEAU, E.: Local reconstruction of stenosed sections of artery using multiple MRA acquisitions. In: *Magnetic Resonance Imaging* 49 (2003), Nr. 4, S. 731–742
- [63] HOFMANN, T.: Probabilistic latent semantic analysis. In: *Uncertainty in Artificial Intelligence*, 1999, S. 289–296
- [64] IBRAHIM, M. ; JOHN, N. ; KABUKA, M. ; YOUNIS, A.: Hidden markov models-based 3D MRI brain segmentation. In: *Image and Vision Computing* 24 (2006), Nr. 10, S. 1065–1079
- [65] IRANI, M. ; PELEG, S.: Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency. In: *Journal of Visual Communication and Image Representation* 4 (1993), Nr. 4, S. 324–335
- [66] ITTI, L. ; KOCH, C. ; NIEBUR, E.: A model of saliency-based visual attention for rapid scene analysis. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 20 (1998), Nr. 11, S. 1254–1259
- [67] JACCARD, P.: Etude comparative de la distribution florale dans une portion des Alpes et des Jura. In: *Societe Vaudoise des Sciences Naturelles* 37 (1901), S. 547–579
- [68] JAUME, S. ; RONDAO, P. ; MACQ, B.: Open Topology: A Toolkit for Brain Isosurface Correction. In: *MICCAI Open Source Workshop*, 2005
- [69] JENKINSON, Mark ; BECKMANN, Christian F. ; BEHRENS, Timothy E. ; WOOLRICH, Mark W. ; SMITH, Stephen M.: FSL. In: *NeuroImage* 62 (2012), Nr. 2, S. 782–790
- [70] KILLIANY, R.J. ; GOMEZ-ISLA, T. ; MOSS, M. ; KIKINIS, R. ; SANDOR, T. ; JOLESZ, F. ; TANZI, R. ; JONES, K. ; HYMAN, B.T. ; ALBERT, M.S. u. a.: Use of structural magnetic resonance imaging to predict who will get Alzheimer’s disease. In: *Annals of Neurology* 47 (2000), Nr. 4, S. 430–439
- [71] KIM, J. ; SINGH, V. ; LEE, J. ; LERCH, J. ; AD-DAB’BAGH, Y. ; MACDONALD, D. ; LEE, J. ; KIM, S. ; EVANS, A.: Automated 3D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. In: *NeuroImage* 27 (2005), Nr. 1, S. 210–221
- [72] KLEINBERG, J.M.: Authoritative sources in a hyperlinked environment. In: *Journal of the ACM (JACM)* 46 (1999), Nr. 5, S. 604–632

- [73] KLÖPPEL, Stefan ; STONNINGTON, Cynthia M. ; CHU, Carlton ; DRAGANSKI, Bogdan ; SCAHILL, Rachael I. ; ROHRER, Jonathan D. ; FOX, Nick C. ; JACK, Clifford R. ; ASHBURNER, John ; FRACKOWIAK, Richard S.: Automatic classification of MR scans in Alzheimer's disease. In: *Brain* 131 (2008), Nr. 3, S. 681–689
- [74] KOUWE, A.J.W. van d. ; BENNER, T. ; SALAT, D.H. ; FISCHL, B.: Brain morphometry with multiecho MPRAGE. In: *NeuroImage* 40 (2008), Nr. 2, S. 559–569
- [75] KRIEGESKORTE, N. ; GOEBEL, R.: An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical MR volumes. In: *NeuroImage* 14 (2001), S. 329–346
- [76] KUNDEL, H.L. ; NODINE, C.F. ; THICKMAN, D. ; TOTO, L.: Searching for Lung Nodules a Comparison of Human Performance with Random and Systematic Scanning Models. In: *Investigative Radiology* 22 (1987), Nr. 5, S. 417–422
- [77] LAIDLAW, D. ; FLEISCHER, K. ; BARR, A.: Partial-volume Bayesian classification of material mixtures in MR volume data using voxel histograms. In: *IEEE Transactions on Medical Imaging* 17 (1998), Nr. 1, S. 74–86
- [78] LE MEUR, O. ; LE CALLET, P. ; BARBA, D. ; THOREAU, D.: A coherent computational approach to model bottom-up visual attention. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 28 (2006), Nr. 5, S. 802–817
- [79] LEE, D.D. ; SEUNG, H.S.: Learning the parts of objects by non-negative matrix factorization. In: *Nature* 401 (1999), Nr. 6755, S. 788–791
- [80] LINDBERG, Tony: Feature detection with automatic scale selection. In: *International Journal of Computer Vision* 30 (1998), Nr. 2, S. 79–116
- [81] LIU, Manhua ; ZHANG, Daoqiang ; YAP, Pew-Thian ; SHEN, Dinggang: Hierarchical Ensemble of Multi-level Classifiers for Diagnosis of Alzheimer's Disease. In: *Machine Learning in Medical Imaging*. Springer, 2012, S. 27–35
- [82] LIU, Yanxi ; TEVEROVSKIY, Leonid ; CARMICHAEL, Owen ; KIKINIS, Ron ; SHENTON, Martha ; CARTER, Cameron S. ; STENGER, V A. ; DAVIS, Simon ; AIZENSTEIN, Howard ; BECKER, James T. u. a.: Discriminative MR image feature analysis for automatic schizophrenia and Alzheimer's disease classification, Springer, 2004

- [83] LUSTIG, M. ; DONOHO, D. ; PAULY, J.M.: Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging. In: *Magnetic Resonance in Medicine* 58 (2007), Nr. 6, S. 1182–1195
- [84] MAGNIN, Benoît ; MESROB, Lilia ; KINKINGNÉHUN, Serge ; PÉLÉGRINI-ISSAC, Mélanie ; COLLIOT, Olivier ; SARAZIN, Marie ; DUBOIS, Bruno ; LEHÉRICY, Stéphane ; BENALI, Habib: Support vector machine-based classification of Alzheimer’s disease from whole-brain anatomical MRI. In: *Neuroradiology* 51 (2009), Nr. 2, S. 73–83
- [85] MAIRAL, J. ; ELAD, M. ; SAPIRO, G.: Sparse representation for color image restoration. In: *IEEE Transactions on Image Processing* 17 (2008), Nr. 1, S. 53–69
- [86] MANGIN, J.-F. ; FROUIN, V. ; BLOCH, I. ; RÉGIS, J. ; LÓPEZ-KRAHE, J.: From 3D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. In: *Journal of Mathematical Imaging and Vision* 5 (1995), Nr. 4, S. 297–318
- [87] MANJÓN, J.V. ; COUPÉ, P. ; BUADES, A. ; COLLINS, D.L. ; ROBLES, M.: MRI Superresolution Using Self-Similarity and Image Priors. In: *International Journal of Biomedical Imaging* 2010 (2010), S. 1
- [88] MANJÓN, J.V. ; COUPÉ, P. ; BUADES, A. ; FONOV, V. ; COLLINS, D.L. ; ROBLES, M.: Non-local MRI Upsampling. In: *Medical Image Analysis* 14 (2010), Nr. 6, S. 784–792
- [89] MARCUS, D.S. ; WANG, T.H. ; PARKER, J. ; CSERNANSKY, J.G. ; MORRIS, J.C. ; BUCKNER, R.L.: Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. In: *Journal of Cognitive Neuroscience* 19 (2007), Nr. 9, S. 1498–1507
- [90] MATHALON, D.H. ; SULLIVAN, E.V. ; LIM, K.O. ; PFEFFERBAUM, A.: Progressive brain volume changes and the clinical course of schizophrenia in men: a longitudinal magnetic resonance imaging study. In: *Archives of General Psychiatry* 58 (2001), Nr. 2, S. 148
- [91] MCINERNEY, T. ; TERZOPOULOS, D.: Topology Adaptive Deformable Surfaces for Medical Image Volume Segmentation. In: *IEEE Transactions on Medical Imaging* 18 (1999), Nr. 10, S. 840–850

- [92] MIETCHEN, D. ; GASER, C.: Computational morphometry for detecting changes in brain structure due to development, aging, learning, disease and evolution. In: *Frontiers in Neuroinformatics* 3 (2009), Nr. 25
- [93] NOE, A. ; GEE, J.: Partial volume segmentation of cerebral MRI scans with mixture model clustering. In: *Proceedings of the International Conference on Information Processing in Medical Imaging* Bd. 2082, Springer, 2001, S. 423–430
- [94] OLIVA, A. ; TORRALBA, A. ; CASTELHANO, M.S. ; HENDERSON, J.M.: Top-down control of visual attention in object detection. In: *Proceedings of the International Conference on Image Processing*. Bd. 1 IEEE, 2003, S. I–253
- [95] OLSHAUSEN, B.A. ; FIELD, D.J.: Sparse coding with an overcomplete basis set: A strategy employed by V1? In: *Vision Research* 37 (1997), Nr. 23, S. 3311–3325
- [96] ONO, M. ; KUBIK, S. ; ABERNATHEY, C.: *Atlas of the cerebral sulci*. Thieme, 1990
- [97] ORRÙ, Graziella ; PETTERSSON-YEO, William ; MARQUAND, Andre F. ; SARTORI, Giuseppe ; MECHELLI, Andrea: Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. In: *Neuroscience & Biobehavioral Reviews* 36 (2012), Nr. 4, S. 1140–1152
- [98] OURSELIN, S. ; ROCHE, A. ; SUBSOL, G. ; PENNEC, X. ; AYACHE, N.: Reconstructing a 3D structure from serial histological sections. In: *Image and Vision Computing* 19 (2001), Nr. 1, S. 25–31
- [99] PADILLA, P ; LÓPEZ, M ; GÓRRIZ, JM ; RAMÍREZ, J ; SALAS-GONZÁLEZ, D ; ÁLVAREZ, I: NMF-SVM Based CAD Tool Applied to Functional Brain Images for the Diagnosis of Alzheimer’s Disease. In: *IEEE Transactions on Medical Imaging* 31 (2012), Nr. 2, S. 207–216
- [100] PAGE, L. ; BRIN, S. ; MOTWANI, R. ; WINOGRAD, T.: The PageRank citation ranking: bringing order to the web. (1999)
- [101] PAL, R. ; MUKHERJEE, A. ; MITRA, P. ; MUKHERJEE, J.: Modelling visual saliency using degree centrality. In: *IET Computer Vision* 4 (2010), Nr. 3, S. 218–229
- [102] PANTAZIS, D. ; LEAHY, R.M. ; NICHOLS, T.E. ; STYNER, M.: Statistical surface-based morphometry using a nonparametric approach. In: *IEEE International Symposium on Biomedical Imaging: Nano to Macro*. IEEE, 2004, S. 1283–1286

- [103] PELED, S. ; YESHURUN, Y.: Superresolution in MRI: application to human white matter fiber tract visualization by diffusion tensor imaging. In: *Magnetic Resonance in Medicine* 45 (2001), Nr. 1, S. 29–35
- [104] PENNANEN, Corina ; KIVIPELTO, Miia ; TUOMAINEN, Susanna ; HARTIKAINEN, Päivi ; HANNINEN, Tuomo ; LAAKSO, Mikko P. ; HALLIKAINEN, Merja ; VANHANEN, Matti ; NISSINEN, Aulikki ; HELKALA, Eeva-Liisa u. a.: Hippocampus and entorhinal cortex in mild cognitive impairment and early AD. In: *Neurobiology of aging* 25 (2004), Nr. 3, S. 303–310
- [105] PEYRE, G.: Sparse Modeling of Textures. In: *Journal of Mathematical Imaging and Vision* 34 (2009), Nr. 1, S. 17–31
- [106] PLANT, Claudia ; TEIPEL, Stefan J. ; OSWALD, Annahita ; BÖHM, Christian ; MEINDL, Thomas ; MOURAO-MIRANDA, Janaina ; BOKDE, Arun W. ; HAMPEL, Harald ; EWERS, Michael: Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer’s disease. In: *Neuroimage* 50 (2010), Nr. 1, S. 162
- [107] RAVISHANKAR, S. ; BRESLER, Y.: MR Image Reconstruction from Highly Undersampled k-Space data by Dictionary Learning. In: *IEEE Transactions on Medical Imaging* 30 (2011), Nr. 5, S. 1028–1041
- [108] ROMERO, E. ; GONZÁLEZ, F.: From Biomedical Image Analysis to Biomedical Image Understanding Using Machine Learning. In: GONZÁLEZ, F. (Ed.) ; ROMERO, E. (Ed.): *Biomedical Image Analysis and Machine Learning Technologies: Applications and Techniques*, IGI Global, 2009, S. 1–26
- [109] ROUSSEAU, F.: Brain hallucination. In: *Proceedings of the European Conference on Computer Vision: Part I*, 2008, S. 497–508
- [110] ROUSSEAU, F. ; INITIATIVE, The Alzheimer’s Disease N.: A non-local approach for image super-resolution using intermodality priors. In: *Medical Image Analysis* 14 (2010), Nr. 4, S. 594–605
- [111] RUEDA, Andrea ; ACOSTA, Oscar ; COUPRIE, Michel ; BOURGEAT, Pierrick ; FRIPP, Jurgen ; DOWSON, Nicholas ; ROMERO, Eduardo ; SALVADO, Olivier: Topology-corrected segmentation and local intensity estimates for improved partial volume classification of brain cortex in MRI. In: *Journal of Neuroscience Methods* 188 (2010), Nr. 2, S. 305–315

- [112] RUEDA, Andrea ; MALPICA, Norberto ; ROMERO, Eduardo: Single-image super-resolution of brain MR images using overcomplete dictionaries". In: *Medical Image Analysis* 17 (2013), Nr. 1, S. 113–132
- [113] SALAS-GONZALEZ, D ; GÓRRIZ, JM ; RAMÍREZ, J ; LÓPEZ, M ; ALVAREZ, I ; SEGOVIA, F ; CHAVES, R ; PUNTONET, CG: Computer-aided diagnosis of Alzheimer's disease using support vector machines and classification trees. In: *Physics in Medicine and Biology* 55 (2010), Nr. 10, S. 2807
- [114] SALAT, D.H. ; LEE, S.Y. ; KOUWE, A.J. van d. ; GREVE, D.N. ; FISCHL, B. ; ROSAS, H.D.: Age-associated alterations in cortical gray and white matter signal intensity and gray to white matter contrast. In: *NeuroImage* 48 (2009), Nr. 1, S. 21–28
- [115] SANTAGO, P. ; GAGE, H.: Quantification of MR brain images by mixture density and partial volume modeling. In: *IEEE Transactions on Medical Imaging* 12 (1993), Nr. 3, S. 566–574
- [116] SÉGONNE, F.: Active contours under topology control - genus preserving level sets. In: *International Journal of Computer Vision* 79 (2008), Nr. 2, S. 107–117
- [117] SHATTUCK, D. ; SANDOR-LEAHY, S. ; SCHAPER, K. ; ROTTENBERG, D. ; LEAHY, R.: Magnetic resonance image tissue classification using a partial volume model. In: *NeuroImage* 13 (2001), Nr. 5, S. 856–876
- [118] SHILLING, R. ; ROBBIE, T. ; BAILLOEUL, T. ; MEWES, K. ; MERSEREAU, R. ; BRUMMER, M.: A Super-Resolution Framework for 3-D High-Resolution and High-Contrast Imaging Using 2-D Multislice MRI. In: *IEEE Transactions on Medical Imaging* 28 (2009), Nr. 5, S. 633–644
- [119] SMITH, S.: Fast robust automated brain extraction. In: *Human Brain Mapping* 17 (2002), Nr. 3, S. 143–155
- [120] STYNER, Martin ; BRECHBUHLER, Christian ; SZCKELY, G ; GERIG, Guido: Parametric estimate of intensity inhomogeneities applied to MRI. In: *Medical Imaging, IEEE Transactions on* 19 (2000), Nr. 3, S. 153–165
- [121] TOEWS, M. ; WELLS, W. ; COLLINS, D.L. ; ARBEL, T.: Feature-based morphometry: Discovering group-related anatomical patterns. In: *NeuroImage* 49 (2010), Nr. 3, S. 2318–2327

- [122] TOHKA, J. ; ZIJDENBOS, A. ; EVANS, A.: Fast and robust parameter estimation for statistical partial volume models in brain MRI. In: *NeuroImage* 23 (2004), Nr. 1, S. 84–97
- [123] TZOURIO-MAZOYER, N. ; LANDEAU, B. ; PAPATHANASSIOU, D. ; CRIVELLO, F. ; ETARD, O. ; DELCROIX, N. ; MAZOYER, B. ; JOLIOT, M.: Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. In: *NeuroImage* 15 (2002), Nr. 1, S. 273–289
- [124] VAN LEEMPUT, K. ; MAES, F. ; VANDERMEULAN, D. ; SUETENS, P.: Automated model-based bias field correction of MR images of the brain. In: *IEEE Transactions on Medical Imaging* 18 (1999), Nr. 10, S. 885–896
- [125] VAN LEEMPUT, K. ; MAES, F. ; VANDERMEULEN, D. ; SUETENS, P.: A unifying framework for partial volume segmentation of brain MR images. In: *IEEE Transactions on Medical Imaging* 22 (2003), Nr. 1, S. 105–119
- [126] VARMA, M. ; BABU, B.R.: More generality in efficient multiple kernel learning. In: *Proceedings of the 26th Annual International Conference on Machine Learning* ACM, 2009, S. 1065–1072
- [127] VERCAUTEREN, T. ; PENNEC, X. ; PERCHANT, A. ; AYACHE, N.: Non-parametric diffeomorphic image registration with the DEMONS algorithm. In: *Proceedings of the Medical Image Computing and Computer-Assisted Intervention Conference* Bd. 4792, Springer Berlin / Heidelberg, 2007, S. 319–326
- [128] WANG, Z. ; BOVIK, A.C. ; SHEIKH, H.R. ; SIMONCELLI, E.P.: Image quality assessment: From error visibility to structural similarity. In: *IEEE Transactions on Image Processing* 13 (2004), Nr. 4, S. 600–612
- [129] WEE, Chong-Yaw ; YAP, Pew-Thian ; SHEN, Dinggang: Prediction of Alzheimer’s disease and mild cognitive impairment using cortical morphological patterns. In: *Human Brain Mapping* (2012)
- [130] WESTMAN, Eric ; MUEHLBOECK, J ; SIMMONS, Andrew u. a.: Combining MRI and CSF measures for classification of Alzheimer’s disease and prediction of mild cognitive impairment conversion. In: *Neuroimage* (2012)
- [131] WHITWELL, Jennifer L. ; PRZYBELSKI, Scott A. ; WEIGAND, Stephen D. ; KNOPMAN, David S. ; BOEVE, Bradley F. ; PETERSEN, Ronald C. ; JACK, Clifford R.: 3D maps from multiple MRI illustrate changing atrophy patterns

- as subjects progress from mild cognitive impairment to Alzheimer's disease. In: *Brain* 130 (2007), Nr. 7, S. 1777–1786
- [132] WILSON, Stephen M. ; OGAR, Jennifer M. ; LALUZ, Victor ; GROWDON, Matthew ; JANG, Jung ; GLENN, Shenly ; MILLER, Bruce L. ; WEINER, Michael W. ; GORNO-TEMPINI, Maria L.: Automated MRI-based classification of primary progressive aphasia variants. In: *Neuroimage* 47 (2009), Nr. 4, S. 1558–1567
- [133] YANG, J. ; WRIGHT, J. ; HUANG, T. ; MA, Y.: Image super-resolution via sparse representation. In: *IEEE Transactions on Image Processing* 19 (2009), Nr. 11, S. 2861–2873
- [134] YE, J.C. ; TAK, S. ; HAN, Y. ; PARK, H.W.: Projection Reconstruction MR Imaging using FOCUSS. In: *Magnetic Resonance in Medicine* 57 (2007), Nr. 4, S. 764–775
- [135] YEZZI, A.J. ; PRINCE, J.L.: An Eulerian PDE approach for computing tissue thickness. In: *IEEE Transactions on Medical Imaging* 22 (2003), Nr. 10, S. 1332–1339
- [136] YU, G. ; SAPIRO, G. ; MALLAT, S.: Solving inverse problems with piecewise linear estimators: from Gaussian mixture models to structured sparsity. In: *Arxiv preprint arXiv:1006.3056* (2010)
- [137] ZEYDE, R. ; ELAD, M. ; PROTTER, M.: On Single Image Scale-Up using Sparse-Representations. In: *Curves & Surfaces*, 2010, S. 711–730
- [138] ZHANG, Y. ; BRADY, M. ; SMITH, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. In: *IEEE Transactions on Medical Imaging* 20 (2001), Nr. 1, S. 45–57
- [139] ZHOU, Q.-Y. ; JU, T. ; HU, S.-M.: Topology repair of solid models using skeletons. In: *IEEE Transactions on Visualization and Computer Graphics* 13 (2007), Nr. 4, S. 675–685