

A Structural Parametrization of the Brain Using Hidden Markov Models Based Paths in Alzheimer's Disease

F.J. Martínez-Murcia, J.M. Górriz*, J. Ramírez
Dept. of Signal Theory, Networking and Communications
18071 University of Granada, Spain

A. Ortiz
Dpt. of Communications Engineering
29071 University of Málaga, Spain
for the Alzheimer's Disease Neuroimaging Initiative[†]

Abstract

The usage of biomedical imaging in the diagnosis of dementia is increasingly widespread. A number of works explore the possibilities of computational techniques and algorithms in what is called Computed Aided Diagnosis. Our work presents an automatic parametrization of the brain structure by means of a path generation algorithm based on Hidden Markov Models. The path is traced using information of intensity and spatial orientation in each node, adapting to the structural changes of the brain. Each path is itself a useful way to extract features from the MRI image, being the intensity levels at each node the most straightforward. However, a further processing consisting of a modification of the Gray Level Co-occurrence Matrix can be used to characterize the textural changes that occur throughout the path, yielding more meaningful values that could be associated to the structural changes in Alzheimer's Disease, as well as providing a significant feature reduction. This methodology achieves high performance, up to 80.3% of accuracy using a single path in differential diagnosis involving Alzheimer-affected subjects versus controls belonging to the Alzheimer's Disease Neuroimaging Initiative (ADNI).

I. INTRODUCTION

Neurodegenerative diseases are considered one of the biggest challenges of the XXI century, specially in developed countries, where the ageing population is a major concern. According to the World Health Organization, the most common type of dementia is Alzheimer's Disease (AD), with more than 35.6 million people affected, and 7.7 million new cases every year[1].

These diseases have a big impact on today's society, as they involve not only health services, but also a conglomerate of social, psychological and familiar care. Therefore, methods oriented to anticipate and improve the diagnosis are desirable. In this sense, the emergence and expansion of new imaging techniques, such as Magnetic Resonance Imaging (MRI), or Single Photon emission Computed Tomography (SPECT) have meant a breakthrough in the diagnosis of AD. Beyond

*Corresponding author: gorriz@ugr.es

[†]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

manual processing of the images, there exists a trend in the usage of computed-aided systems to help in the differential diagnosis and even predict the conversion from prodromal stages (Mild Cognitive Impairment, or MCI) to AD [2].

Given its high tissue contrast, the computational processing of MRI can be straightforward. A number of approaches have been proposed for the processing of MRI images, most of them using univariate analysis. Univariate analyses are performed using one single feature at a time, and frequently consists of a voxel-wise comparison between intensity levels, either over the images themselves (voxel based morphometry, or VBM[3, 4]), Regions of Interest (ROIs) [5] or models of some brain features, such as cortical thickness[6] or intrinsic curvature[7]. This methodology has been implemented in the most prominent pieces of software used in clinical practice, such as FreeSurfer or SPM[8], and is the most used in epidemiological studies.

However, there is an increasing interest in multivariate approaches. Those, in contrast to univariate, are able to handle regional patterns, texture features and voxel and region-wise relationships [9, 10, 11, 12]. These approaches could reveal other information than the volumetric, complementing and providing new insights into the disease. Among them, some CAD systems based on texture features have been proposed [9, 13, 14]. Particularly, in Ref [13], an easily computed texture descriptor, Local Binary Pattern (LBP) has demonstrated its utility in various high-level brain MR image analysis, and in Ref [15], a Parkinson's Disease diagnosis system based on Haralick Texture Features was proposed.

At the same time, in Ref. [16], a framework for feature extraction in MR Brain Images was proposed. It consists on the computation of different statistical and morphological values along some radia centred in the Anterior Commissure (AC) and spanning over all values of azimuth and elevation angles. Starting at this point, morphological measures such as cortical thickness, surface, or statistical texture features such as average, kurtosis or entropy of the voxels crossed by the mapping vector were computed. This provided a significant feature reduction, as well as some 2D maps representing each of these features, which could be considered a visual way to represent this information. Some of the features have been proved to be related to anatomical changes such as brain atrophy, yielding high performance in differential diagnosis. Furthermore, in Ref. [17] an extension to this algorithm provided characterization of the texture around the radius, with good performance as well.

In this work, a new algorithm to trace paths inside the brain is presented. It is intended to substitute the rectilinear radia in SBM with curvilinear paths that adapt to the structural information present in MRI brain images. Using this paths, the features (either intensity levels or more complex characteristics) will be more descriptive of the underlying brain structure. To complement the methodology, along with the intensity levels in the path, we have developed a radial method to compute the Gray Level Co-occurrence Matrix (GLCM). Firstly developed by Haralick[18], the GLCM is one of the preferred methods to characterize the texture of 2D images, and have been successfully used in the characterization of brain images[15, 19]. In our approach, the original algorithm is modified to compute the GLCM along the extracted curvilinear paths and its neighbourhood, from which a number of descriptors of the underlying texture in each direction can be extracted.

This paper is organized as follows: after this introduction, in Section II the overall methodology is presented, comprising an introduction to SBM, the whole path tracing algorithm using Hidden Markov Models and the computation of the radial GLCM. Later, in Section III, the database used in the article is introduced, along with the experimental setup, and the results obtained in each experiment. In Section IV, these results are discussed and analyzed, and finally, in Section V, some conclusions are drawn.

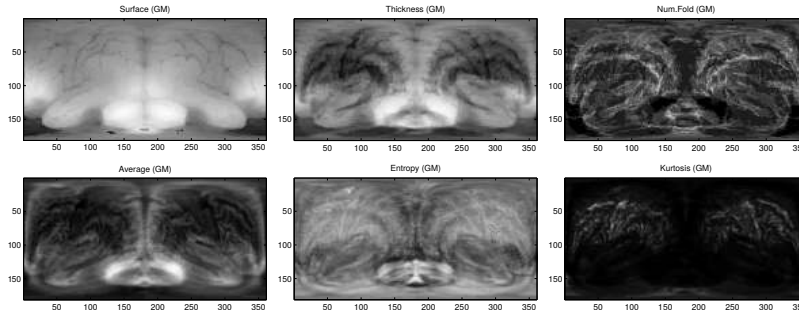


Figure 1: Example of 2D projections of different statistics using the method proposed in [16].

II. METHODOLOGY

II.1. Spherical Brain Mapping

In Ref. [16], the Spherical Brain Mapping (SBM), a framework to perform a 2D projection based on the use of spherical coordinates in the brain is proposed. This method computes different statistical measurements across the projecting vector $\mathbf{p}_{\theta,\varphi}$, which is defined in a pair of inclination (θ) and azimuth (φ) angles in the range $[-\pi/2, \pi/2]$ and $[-\pi, \pi]$ respectively, and sampling the nearest voxel. In other words, a 2D projection is obtained by computing a unique value R from the projecting vector $\mathbf{p}_{\theta,\varphi}$ for each (θ, φ) pair of coordinates, what yields a two-dimensional map in the $\theta - \varphi$ plane. As an example, 2D projection of different statistics is shown in Figure 1.

This representation of the brain is able to provide insight into different and useful properties of the anatomical structures such as brain folds or the cortical thickness in a 2D representation. However, the mapping process of reducing from 3D to 2D loses textural information in the vicinity of the vector $\mathbf{p}_{\theta,\varphi}$ that may be important from a discriminative point of view. To avoid this, SBM was extended in Ref. [17] with a method that preserves part of the 3D textural information by means of a so-called Volumetric Radial Local Binary Pattern (VRLBP) descriptor which takes into account not only the voxels in the radius of a specific direction (θ, φ) but also the voxels in a helical neighbourhood.

II.2. Path tracing based on Hidden Markov Models

3D Biomedical Images can be understood as a tuple containing spatial information in the image range ($\mathbf{x} \in \mathbb{I}$, where $\mathbb{I} \subset \mathbb{R}^3$) as well as intensity information ($I(\mathbf{x}) \in \mathbb{R}$). There exist a number of possibilities in the interpretation of intensity data on the images, from plain intensity values to a sampling of the underlying tissue density (and thus, an estimation of the probability of finding a neuron in each position).

Following our Spherical Brain Mapping approach, in which some radia are used to extract relevant statistical features from these images, we formulate a 3D path tracing algorithm suitable for extraction of curvilinear structures from 3D biomedical images, and directly linked to each direction (φ, θ) as in the original work. Our objective is to make these paths as representative of the underlying structure as possible. To do so, we must use both intensity and spatial information to construct maximum intensity-similarity paths oriented in the given direction

Let us note a 3D path in a certain direction (φ, θ) as a Markov Model[20]:

$$\mathbf{X} = \{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \quad (1)$$

Therefore, our optimum path would be the one that maximizes the probability of the path:

$$\mathbf{X}_{opt} = \arg \max_{\mathbf{X}} \{P(\mathbf{X})\} \quad (2)$$

or, similarly, the probability of all the nodes:

$$P(\mathbf{X}) = P(\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \quad (3)$$

where \mathbf{x}_0 is the node located at the Anterior Commissure (AC) and \mathbf{x}_N is the theoretical projection of the current direction (φ, θ) in the limits of the image \mathbb{I} . Setting \mathbf{x}_0 at the AC is the convention when using the Montreal Neurological Institute (MNI) coordinates[21], however it provides one point that is connected with both hemispheres as well. Therefore, that setting ensures continuity during the path tracing. We can assume a first-order Hidden Markov Model (HMM) for the tracing of the path, and the i -th node will be computed as:

$$P(\mathbf{x}_i | \mathbf{x}_{i-1}, \mathbf{x}_{i-2}, \dots, \mathbf{x}_0) \approx P(\mathbf{x}_i | \mathbf{x}_{i-1}) \quad (4)$$

and with this assumption, Eq. (3) becomes:

$$P(\mathbf{X}) = P(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_N) = \prod_{i=1}^N P(\mathbf{x}_i | \mathbf{x}_{i-1}) \quad (5)$$

In our case, we will assume that the hidden state of each node will be its intensity $I(\mathbf{x}_i)$. Let $\mathbf{I} = \{I(\mathbf{x}_0), I(\mathbf{x}_1), \dots, I(\mathbf{x}_N)\}$ be the vector of the intensities in each node. With the introduction of these factors, our optimal path defined in Eq. (2) can be view as:

$$\mathbf{X}_{opt} = \arg \max_{\mathbf{X}} \{P(\mathbf{X} | \mathbf{I})\} \quad (6)$$

$$P(\mathbf{X} | \mathbf{I}) = P(\mathbf{x}_0, \dots, \mathbf{x}_N | I(\mathbf{x}_0), \dots, I(\mathbf{x}_N)) \quad (7)$$

$$= \frac{P(I(\mathbf{x}_0), \dots, I(\mathbf{x}_N) | \mathbf{x}_0, \dots, \mathbf{x}_N) \cdot P(\mathbf{x}_0, \dots, \mathbf{x}_N)}{P(I(\mathbf{x}_0), \dots, I(\mathbf{x}_N))} \quad (8)$$

where:

$$P(I(\mathbf{x}_0), \dots, I(\mathbf{x}_N) | \mathbf{x}_0, \dots, \mathbf{x}_N) = \prod_{i=1}^N P(I(\mathbf{x}_i) | \mathbf{x}_i) \quad (9)$$

and $P(I(\mathbf{x}_0), \dots, I(\mathbf{x}_N))$ is the *a priori* probability of the intensities in the path. We can assume, without lack of generality, that this term is constant along the path, and therefore, it plays no part in the optimization process.

For computational purposes, we will compute all the needed probabilities on a set of candidates $\mathbf{X}_c = \{\mathbf{x}_{c,1}, \mathbf{x}_{c,2}, \dots, \mathbf{x}_{c,M}\}$ defined by \mathbf{x}_{i-1} . This candidates are contained inside the L^2 -norm support ball $B_{2,r}(\mathbf{x} - \mathbf{x}_{i-1})$ of radius r centred in \mathbf{x}_{i-1} .

To estimate the individual probabilities needed for Eq. (9) $P(I(\mathbf{x}_i) | \mathbf{x}_i)$, we can assume a normal distribution of intensities of the candidates with mean $I(\mathbf{x}_{i-1})$ and variance σ_c^2 the variance of the intensities of the candidate set. By assuming this, the probability of a certain intensity in the candidate node \mathbf{x}_i increases as the intensity becomes more similar to the intensity of \mathbf{x}_{i-1} . Therefore, the assumption supports the tracing of minimal intensity difference paths. The probability of the intensity of a candidate \mathbf{x}_i will be:

$$P(I(\mathbf{x}_i) | \mathbf{x}_i) = \frac{1}{\sqrt{2\pi\sigma_c^2}} \exp\left(-\frac{(I(\mathbf{x}_i) - I(\mathbf{x}_{i-1}))^2}{2\sigma_c^2}\right) \quad (10)$$

The last term $P(\mathbf{x}_0, \dots, \mathbf{x}_N)$ to be defined in Eq. (8) is directly related to the Spherical Brain Mapping framework defined before, as it will depend on the radial direction (φ, θ) that we want to “force” in the path. To define this term, we will define an attractor located in the position \mathbf{x}_N (the projection of the current direction (φ, θ) in the limits of the image \mathbb{I}). We assume that this attractor defines the conditionality, that is, that it affects the transition probability between states by means of an isotropic Gaussian radial basis function (RBF), as defined in Eq. (11). This definition helps the attractor to lightly condition the direction of the path at first, and more strongly as the path approaches the cortex, leading to a better representation of the underlying structure.

$$P(\mathbf{x}_0, \dots, \mathbf{x}_N) = P(\mathbf{x}_i | \mathbf{x}_N) \quad (11)$$

$$= \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left(-\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_N)\Sigma^{-1}(\mathbf{x}_i - \mathbf{x}_N)\right) \quad (12)$$

where Σ is the covariance matrix of the given distribution. As we will consider only isotropic gaussian kernel, the matrix can be considered as a 3×3 diagonal matrix whose diagonal elements are a scalar value σ^2 , which we set in each iteration to the euclidean distance between \mathbf{x}_i and \mathbf{x}_N .

Step Size

In this algorithm, instead of using a fixed step size, we evaluate each candidate point $\mathbf{x} \in B_{2,r}(\mathbf{x} - \mathbf{x}_i)$, and thus, the only parameter to regulate is the radius of the L2-norm support ball r . As a trade-off between computational issues and definition of the defined path, we will use a value of $r = 3$, which gives approximately 200 candidate points in each iteration.

Stop Condition

Although the paths could be defined until they reach the last point \mathbf{x}_N (remember, the projection of the general direction onto the limits of \mathbb{I}), our interest is to model the paths inside the brain, and therefore we establish a stop condition using an intensity threshold. This threshold is computed using the entropic thresholding, as defined in Ref. [22]. Let $G_m \equiv \{I_0, I_1, \dots, I_m\}$ denote the set of intensity levels of the whole image. We can compute a histogram to obtain the observed frequencies f_{I_i} , and thus, the observed probability of the different gray levels $p_i = f_{I_i}/N$, where N is the number of voxels in our image \mathbb{I} .

For a given intensity threshold $I_{th} = I_s$, if $\sum_{i=0}^{s-1} p_i$ is larger than zero and smaller than 1, the following distributions can be derived from this distribution after normalization:

$$A \equiv \left\{ \frac{p_0}{P(I_s)}, \frac{p_1}{P(I_s)}, \dots, \frac{p_{s-1}}{P(I_s)} \right\} \quad (13)$$

$$B \equiv \left\{ \frac{p_s}{1 - P(I_s)}, \frac{p_{s+1}}{1 - P(I_s)}, \dots, \frac{p_m}{1 - P(I_s)} \right\} \quad (14)$$

where $P(I_s) = \sum_{i=0}^s p_{I_i}$ is the cumulative density function for the s -th gray level. Therefore, we choose the threshold so that the total amount of information provided by A and B (foreground and background of the image) is maximized. The total information provided by the s -th gray level

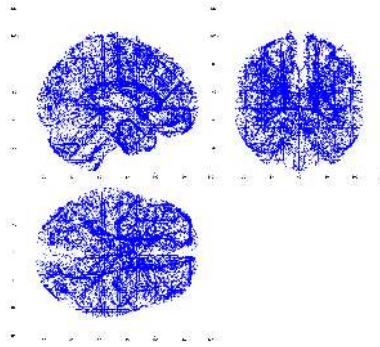


Figure 2: Set of HMM based paths over the MRI DARTEL template.

is:

$$TE(s) = E_A(s) + E_B(s) \quad (15)$$

$$= - \sum_{i=0}^{s-1} \left(\frac{p_i}{P(I_s)} \right) \log \left(\frac{p_i}{P(I_s)} \right) \quad (16)$$

$$- \sum_{i=s}^{m-1} \left(\frac{p_i}{1 - P(I_s)} \right) \log \left(\frac{p_i}{1 - P(I_s)} \right) \quad (17)$$

A summary of our HMM-based path tracing method is shown in Algorithm 1. To illustrate the effect of the algorithm on a real example, Fig. 2 depicts the resulting set of paths computed over the standard MRI DARTEL template.

Algorithm 1: HMM-based Path Creation

input :MRI Brain Image I of size $U \times V \times W$, \mathbf{x}_0

output: List of nodes in the optimum path \mathbf{X}_{opt}

Compute the $I_{th} = I_s$ where s maximizes $TE(s)$;

Set \mathbf{x}_0 to the AC;

Compute the attractor position \mathbf{x}_N in the direction (φ, θ) ;

$\mathbf{x}_i \leftarrow \mathbf{x}_0$;

while $(i < IterLimit) \ \& \ (I(\mathbf{x}_i) > I_{th}) \ \& \ (\mathbf{x}_i \in \mathbb{I})$ **do**

 Get the node candidates $\mathbf{X}_c = \{\mathbf{x}_{c,1}, \mathbf{x}_{c,2}, \dots, \mathbf{x}_{c,M}\}$ where $\mathbf{x}_{c,m} \in B_{2,r}(\mathbf{x}_{c,m} - \mathbf{x}_i)$;

 Get the intensities of the candidates $I(\mathbf{x}_c) \ \forall \mathbf{x}_c \in \mathbf{X}_c$;

foreach $\mathbf{x}_c \in \mathbf{X}_c$ compute $P(\mathbf{x}_c|\mathbf{x}_N)$ and $P(I(\mathbf{x}_c)|\mathbf{x}_i)$;

$\mathbf{x}_{i+1} = \arg \max_{\mathbf{x}_c} [P(I(\mathbf{x}_c)|\mathbf{x}_i) \cdot P(\mathbf{x}_c|\mathbf{x}_N)]$;

$i = i + 1$;

$\mathbf{X}_{opt} \leftarrow \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_N\}$;

II.3. Radial Texture Features

The paths extracted with the aforementioned algorithm present a meaningful representation of the structure of the intensity levels (and therefore, the tissue density) of the MRI brain images. At

this point, one might consider using the intensity values located in those radia as features, and try to characterize each radius' discrimination ability by means of these intensities.

Conversely, the strategy proposed in Ref. [17], which uses Local Binary Pattern (LBP) descriptors in the helical neighbourhood of a rectilinear mapping vector, might be consider useful as well. Due to the HMM paths topology, the helical sampling becomes difficult to compute and not even useful, therefore we propose a modification of the Gray Level Co-occurrence Matrix (GLCM) along the paths.

The GLCM, proposed by Haralick[18], is one of the most widely used methods in texture characterization, and it has been successfully applied to medical imaging in the past[9, 15]. It works by storing the number of voxel-wise correspondences between K grey levels with a certain position offset Δ on a $K \times K$ matrix (C_Δ) along all the image.

Our modification will compute a node-wise GLC matrix, in which the number of grey-level transitions between adjacent nodes, noted as \mathbf{x}_i and \mathbf{x}_{i+1} , is stored along the whole path $\mathbf{X} = \{\mathbf{x}_0, \mathbf{x}_1 \dots \mathbf{x}_N\}$. Mathematically, the computation of the GLCM in each point in the path will be:

$$C_{\Delta_i}(j, k) = \sum_{i=0}^{N-1} \begin{cases} 1 & I(\mathbf{x}_i) = j, I(\mathbf{x}_{i+1}) = k \\ 0 & otherwise \end{cases} \quad (18)$$

where the offset is different for each pair of nodes $\Delta_i = \mathbf{x}_{i+1} - \mathbf{x}_i$.

The definition provided in Eq. (18) is intended for computing the values in each node. However, we can generalize this construction to include not only the nodes, but the intensity information around each node in the computation, which could potentially lead to more significant texture features. Let us note a set containing all the voxels in the closed neighbourhood of \mathbf{x}_i as X_i . Therefore, Eq. (18) can be generalized for any voxel $\mathbf{x} \in X_i$ as:

$$C(j, k) = \sum_{i=0}^{N-1} \sum_{\mathbf{x} \in X_i} \begin{cases} 1 & I(\mathbf{x}) = j, I(\mathbf{x} + \Delta_i) = k \\ 0 & otherwise \end{cases} \quad (19)$$

From the GLCM, a variety of texture descriptors, or features, can be extracted. In this work we will use ten texture features proposed in the original Haralick's article[18] as well as in Refs [23] and [24]: Contrast[18] (CNTR), Correlation[18] (CORR), Dissimilarity[23] (DISS), Energy[18] (ENER), Entropy[23] (ENT), Homogeneity[18] (HOMO), Difference Variance[18] (DVAR), Difference Entropy[18] (DENT), Inverse Difference Normalized[24] (IDN) and Inverse Difference Moment Normalized[24] (IDMN). The mathematical expressions for these features are presented in Equations 20 to 29.

$$\text{CNTR} = \sum_j \sum_k \{(j - k)^2 C(j, k)\} \quad (20)$$

$$\text{CORR} = \frac{\sum_j \sum_k (j - \mu_j)(k - \mu_k) C(j, k)}{(\sigma_j \sigma_k)} \quad (21)$$

$$\text{DISS} = \sum_j \sum_k \{|j - k| C(j, k)\} \quad (22)$$

$$\text{ENER} = \sum_j \sum_k C(j, k)^2 \quad (23)$$

$$\text{ENT} = - \sum_j \sum_k C(j, k) \log(C(j, k)) \quad (24)$$

$$\text{HOMO} = \sum_j \sum_k \frac{C(j, k)}{1 + |j - k|} \quad (25)$$

$$\text{DVAR} = \sum_{j=0}^{N_g-1} j^2 p_{x-y}(j) \quad (26)$$

$$\text{DENT} = - \sum_{j=0}^{N_g-1} p_{x-y}(j) \log\{p_{x-y}(j)\} \quad (27)$$

$$\text{IDN} = \sum_j \sum_k \frac{C(j, k)}{1 + |j - k|/N} \quad (28)$$

$$\text{IDMN} = \sum_j \sum_k \frac{C(j, k)}{1 + (j - k)^2/N^2} \quad (29)$$

II.4. MRI Brain Image Database

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org.

The database used in this article was extracted from the ADNI1: Screening 1.5T (subjects who have a screening data) and contains 1075 T1-weighted MR images: 188 AD, 416 MCI and 201 NC.

In this work, a total number of 360 individuals (180 AD and 180 NC) have been randomly selected to avoid class prevalence. The images were preprocessed, co-registered and segmented using the Statistical Parametric Mapping (SPM) software [8], and the Voxel-Based-Morphometry 8 (VBM8) toolbox [25, 4]. Demographics details about the subjects can be found at Table 1. First, the preprocessing was performed using a nonlinear deformation field using the tissue probability maps provided by the International Consortium for Brain Mapping (ICBM). Secondly, we affine registered the data to the Montreal Neurological Institute (MNI) space and resized to $121 \times 145 \times 121$ voxel. Finally, we performed the co-registration using the high-dimensional Dartel normalization using the standard Dartel template provided by VBM8[4], and a three-dimensional gaussian kernel with SD of 3mm was used to smooth the image maps and account for inter-subject misregistration.

III. EXPERIMENTS AND RESULTS

III.1. Experimental Setup

In order to test our HMM-based path tracing algorithm, we propose the following experiments:

- Firstly, in Sec. III.2 an evaluation of the algorithm over several synthetic 2D and 3D images which are not related to the biomedical task.

Sex	AD			Normal		
	M	F	Total	M	F	Total
N	95	85	180	98	82	180
Age (SD)	75.28 (7.62)	75.24 (7.54)	75.24 (7.57)	75.95 (5.27)	75.80 (5.03)	76.05 (4.98)
MMSE (SD)	23.18 (1.95)	23.31 (2.18)	23.31 (2.04)	29.11 (1.10)	29.17 (0.80)	29.11 (1.03)
CDR (SD)	0.75 (0.25)	0.76 (0.25)	0.74 (0.25)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
GDSCALE (SD)	1.68 (1.44)	1.73 (1.44)	1.67 (1.45)	0.69 (0.97)	0.83 (1.18)	0.86 (1.16)

Table 1: Demographics details of the selected subjects of the dataset. Scores of MMSE, CDR and GDSCALE tests are provided.

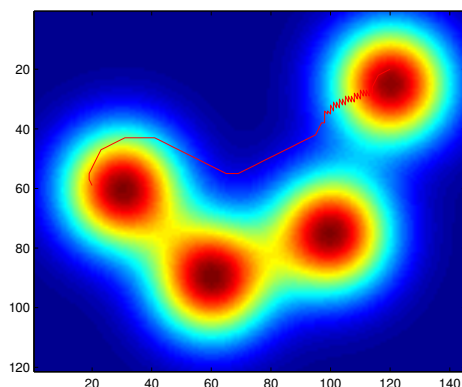


Figure 3: Path traced over a gaussian mixture distribution of 4 isotropic gaussian kernels.

- Secondly, in Sec. III.3, the HMM paths created using the DARTEL template will be evaluated in the differential diagnosis (NOR versus AD).
- Finally, in Sec. III.4, the proposed texture feature maps computed along the HMM paths will be evaluated in a differential diagnosis as well.

III.2. 2D and 3D demonstrations

A demonstration of the ability of our HMM path tracing algorithm can be found in Figures 3, 4 and 5. In Fig. 3, the path tracing algorithm has been tested over a synthesized gaussian mixture probability density function using four isotropic gaussian kernels. The initial point was located at $\mathbf{x}_0 = (120, 20)$ and the attractor at $\mathbf{x}_N = (20, 60)$. The resulting path maximizes both the orientation of the path (towards \mathbf{x}_N) and the minimum change in the intensity values, which is specially visible in the last nodes of the path, where it approaches \mathbf{x}_N surrounding the nearby kernel. In this case, the chosen L2-norm of the support ball has been $r = 3$.

The algorithm has been tested on a three-dimensional, helix-shaped point distribution as well (Fig. 4). The tracing algorithm needs per-voxel intensity (or probability) values, therefore we have estimated the probability distribution of the points as the number of points within each voxel over the total number of points. Using \mathbf{x}_0 as the point with minimum z coordinate in the data distribution and \mathbf{x}_N the one with maximum z , the resulting path follows the data distribution consistently until it reaches the attractor.

Finally, we have tested the algorithm on a real world example, using a digital elevation model

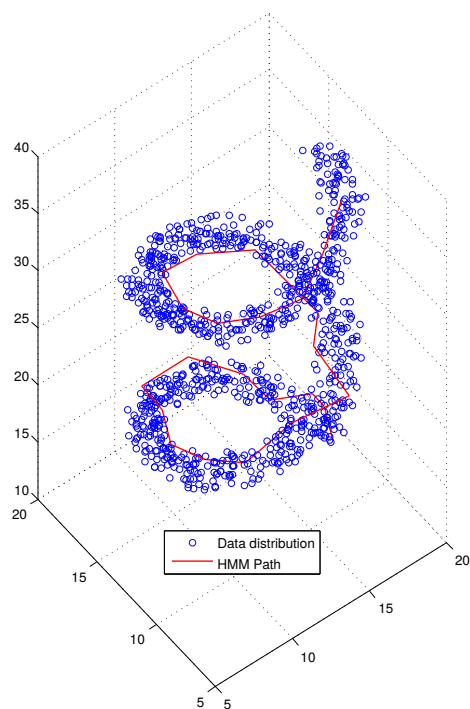


Figure 4: HMM path computed inside a density distribution defined by an helix.

(DEM) of the Iberian Peninsula, generated by the LANDSAT SRTM30+ mission (see Fig. 5). We have tested a multiple path tracing by establishing sequentially \mathbf{x}_0 and \mathbf{x}_N in ten cities. The resulting paths optimize both the distance and height variation, as well as resembling -in most cases- the roads that connect these cities in the real world. Given the dimensions of the image, in this case, the L2-norm of the support ball has been set to $r = 30$.

III.3. Intensity paths

In this section, we present the results of the first experiment involving paths in MRI brain images. First, the experiment performs an estimation of all the paths in each spatial direction (φ, θ) , with $\varphi \in [0, 360]$ and $\theta \in [-90, 90]$ over the DARTEL template[4] (from now on, DARTEL paths). After that, the paths computed in the template are applied to all images in the database in different ways.

The first approximation is to compute the performance values of each path computed at each spatial direction (φ, θ) . Thereby, the feature vector for each subject will be the intensity of the MRI brain image at the nodes of the path, which we have evaluated using a 10-Fold cross validation strategy. The performance is shown as the colour of each path traced in the DARTEL template in Figure 6. The higher accuracy obtained using only one path is 0.8028 ± 0.0873 , and correspond to the light green paths that cross the temporal lobe.

It is interesting to question if the performance of this differential diagnosis could be improved using the information contained in more than one path at a time. To this end, we can select the higher accuracy paths according to the aforementioned performance measures, and select use all the voxels in these paths as features in the SVC. Additionally, we can use a t -test over the set

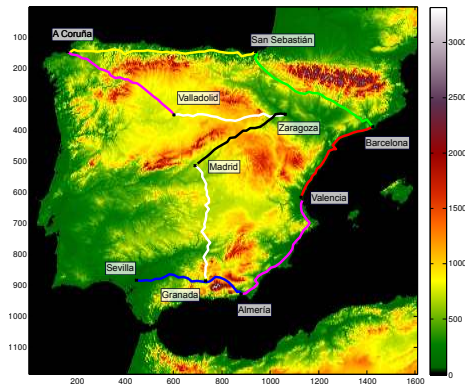


Figure 5: Simulation of the HMM-based path tracing over an Iberian Peninsula height map, interconnecting different cities.

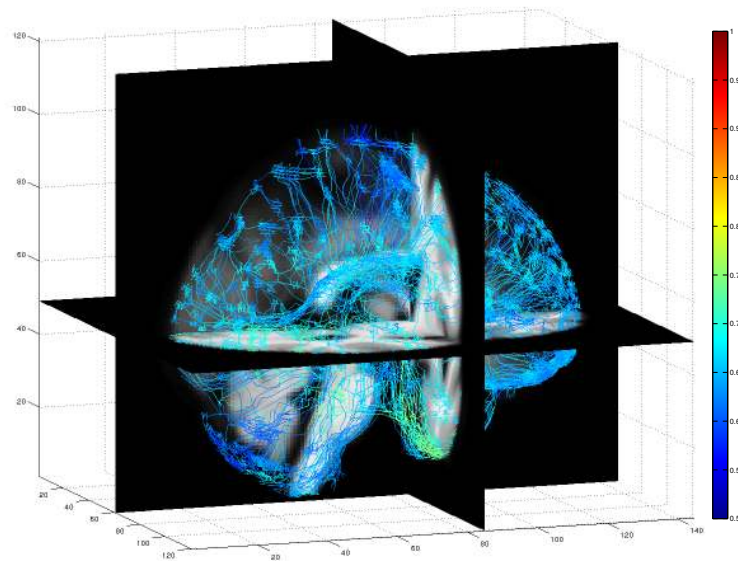


Figure 6: DARTEL paths computed in each direction (φ, θ). Each path's colour represent the accuracy in a differential diagnosis. Only one in every five paths are shown for clarity purposes.

Side	Accuracy	Sensitivity	Specificity
Both	0.806 ± 0.069	0.733 ± 0.073	0.878 ± 0.097
Left	0.769 ± 0.035	0.717 ± 0.061	0.822 ± 0.057
Right	0.792 ± 0.080	0.706 ± 0.120	0.878 ± 0.101
Both	0.806 ± 0.067	0.761 ± 0.074	0.850 ± 0.098
Left	0.772 ± 0.090	0.728 ± 0.115	0.817 ± 0.131
Right	0.758 ± 0.051	0.689 ± 0.060	0.828 ± 0.109

Table 2: Performance values (\pm STD) for the selected paths as features, and using t -test to select the voxels.

Feature	Accuracy	Sensitivity	Specificity
CNTR	0.733 ± 0.060	0.689 ± 0.126	0.778 ± 0.105
CORR	0.672 ± 0.068	0.672 ± 0.112	0.672 ± 0.100
DISS	0.711 ± 0.085	0.678 ± 0.110	0.744 ± 0.102
ENER	0.689 ± 0.061	0.700 ± 0.115	0.678 ± 0.073
ENT	0.675 ± 0.101	0.672 ± 0.115	0.678 ± 0.159
HOMO	0.697 ± 0.058	0.700 ± 0.115	0.694 ± 0.106
DVAR	0.736 ± 0.070	0.683 ± 0.098	0.789 ± 0.090
DENT	0.725 ± 0.122	0.683 ± 0.176	0.767 ± 0.114
IDN	0.719 ± 0.065	0.683 ± 0.108	0.756 ± 0.105
IDMN	0.717 ± 0.076	0.678 ± 0.125	0.756 ± 0.084

Table 3: Performance values (\pm STD) for each of the 10 texture features.

of voxels selected by these paths, to further reduce the set to those voxels that have significant ($p < 0.05$) t -values ($|t| > 1.96$). The performance values for the experiment involving all voxels in the paths (first row) and the one that uses only those significant voxels (second row) are presented in Table 2.

III.4. Texture features

The second experiment proposed is based on the extraction of texture features from the DARTEL paths in each image. With this approach to the problem, we obtain one single value of a texture feature for each path in the subject. That is, each subject will be characterized by 2D, 361×181 array of scalars, one for each texture feature applied to the paths. Performance values for the nine texture features maps (Contrast (CNTR), Correlation corrm (CORR), dissimilarity (DISS), Energy (ENER), Entropy (ENT), Homogeneity (matlab) (HOMO), difference variance (DVAR), difference entropy (DENT), inverse difference normalized (IDN) and inverse difference moment normalized (IDMN)) are presented in Table 3.

The first thing to note in this table is that the maximum accuracy obtained using any texture map is 0.736 ± 0.070 , obtained by Difference Variance. The performance values of the different texture features, all obtaining accuracies higher than 65% (most of them above 70%) reveal the discrimination abilities of these textures, although these values are not as good as those obtained using the plain paths, as in the previous section.

IV. DISCUSSION

We have proposed a Hidden Markov Model based description of similar intensity paths to parametrize the internal structure of the brain. These paths have been using throughout this work as a feature selection method, either by selecting voxels or computing texture features. Later, we have tested the two approaches in a differential diagnosis using MRI brain images of controls and AD affected patients.

Each path is constructed so that it follows a general direction set by the spherical coordinate pair (φ, θ) as well as keeping in a minimum intensity variation in the choice of the next node. This makes the paths very representative of the structure of the MRI brain image. There exists two possible approaches to the path creation, the first of them being the one used in this article. Our approach traces the HMM paths over the DARTEL template (the one used in the registration step) and then uses these paths to select the intensity levels in all the images in the dataset, and then perform several computations, including the characterization of texture along these paths. A second approach would compute the HMM paths independently on each subject in the dataset, and then morphologically compare the path features. However, this last approach is far more computationally expensive and does not allow a direct anatomical relationship between the paths in each subject and their discriminative power, and therefore, we have opted on the first.

Using the first of the proposed approaches, we have compared the discriminative power of the paths by using the set of intensity values at each node as the feature vector in a classifier, and then evaluating the performance in a differential diagnosis (AD versus NOR). In Fig. 7, the paths that achieved best performance (accuracy higher than 0.75) in this differential diagnosis are shown along with some structures rendered from the Automated Anatomical Labeling (AAL) brain atlas[26].

The paths that obtain higher accuracy are those that cross structures such as the Hippocampus, Amygdala, Thalamus, Fusiform and Inferior Temporal Gyrus. Particularly, grey matter loss in the Hippocampus has been described in the NINCDS-ADRDA criteria for AD diagnosis[27] and is widely accepted[28, 29, 30]. Furthermore, the evidence suggest that atrophy affects the surrounding structures (Amygdala, Parahippocampal and Fusiform Gyrus) as well[28, 29]. Some studies have found significant atrophy in the Thalamus and Putamen in early AD[30] as well. Generally, in advanced AD, most of the neocortex and grey matter suffer from atrophy[28, 29, 30], which explains why most of the paths that involve the neocortex in Fig. 6 obtain accuracy rates around 0.7.

Regarding the texture analysis, one can see the discrimination ability of the features in their performance results (Table 3), where most of them surpass the barrier of 70% accuracy. While these results are poorer than in the case of single or combined paths, it is probably due to the definition of the paths itself: the paths are traced using a minimum intensity change criterium (the one that states the probability of the intensity of the next node using a RBF function centered in the current intensity value, as described in Eq. (10)). Therefore, the textural changes between the different nodes in the path should be minimal.

The usage of the 1-voxel closed neighbourhood of the nodes of the path might be a bad choice as well, specially regarding the spatial smoothing applied to the images to account for inter-subject misregistration. Consequently a wider neighbourhood could be more indicative of the underlying texture, allowing a better characterization of the atrophy of the different structures in the brain.

However the real utility of these texture features might be related to the evolution of the disease and, therefore, its application in longitudinal studies, as it has been recently proven[19]. It is very convenient to use a single number to characterize a measure (in our case, texture features) in each direction. The texture obtained in each session can be used to construct a function of

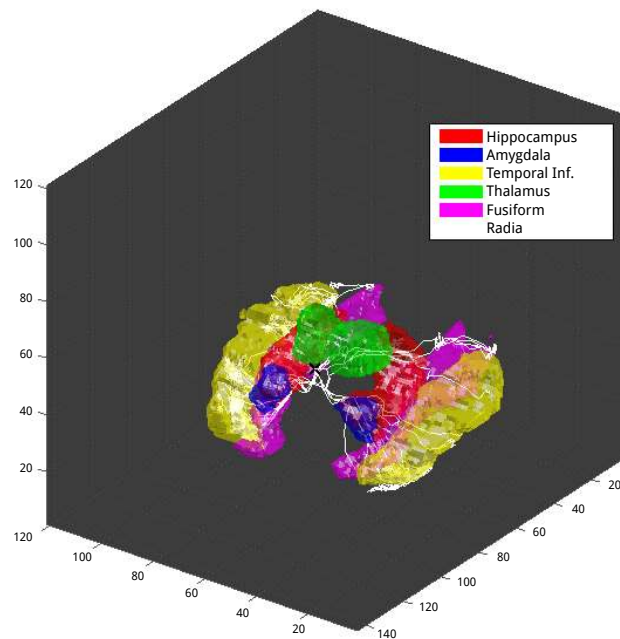


Figure 7: Paths that obtain more than 75% accuracy, and a three-dimensional representation of the structures crossed by them.

neurodegeneration that allows the exploration of the different stages of the disease as the changes in the brain texture along the time within a single patient.

Finally, it might be straightforward to compare our HMM based paths to Diffusion Tensor Imaging (DTI) fibers, which has indeed proved a good tool for the diagnosis of AD[31, 32]. Some could even argue that the resulting DTI tract reconstruction look similar to our computed paths (Figs. 2 and 6). However, they are based on very different principles and sources. DTI fibers are the result of a further processing over diffusion images that contain tensors that quantify the water molecule motion -in both direction and average magnitude- at the voxel level. Conversely, the HMM paths are extracted from the static, structural MRI images by means of their grey levels, and are used just as a tool to characterize the structure and perform feature extraction. Therefore, it is just the idea of “connectivity” what makes both strategies look similar, but each of them focus on very different aspects of the brain.

As we commented previously, the presented work uses the paths computed over the DARTEL template to extract features in all brains. It could be much more representative to compute the paths at a subject levels. Consequently, each image could be modelled by its own paths, and the further processing (in terms of intensity levels or texture features) would be far more specific, although extraordinarily computationally expensive. This will be an approach that we will probably adopt in future work.

V. CONCLUSIONS

In this work, a three dimensional path tracing algorithm based on Hidden Markov Models, where each node’s location depends on the intensity levels of a 3D image and the relative position to an

attractor. It has been proposed as a method to study the structure of the brain using Magnetic Resonance Imaging (MRI). The resulting paths, starting at the Anterior Commissure, generate a number of nodes that adapt to the distribution of intensities inside the MRI images, while maintaining a certain direction. After generating these paths, the values of the intensity in the nodes can be used as features. Furthermore, some further texture features can be extracted using the neighbourhood of the nodes in the path, using what we called radial Gray Level Co-occurrence Matrix and Haralick Texture Analysis. The resulting texture measures can be assembled to form a texture map in each direction of the space.

We have evaluated the discrimination ability of both the intensities that fall in the nodes of the image and the texture features extracted from the paths, obtaining up to a 0.8028 ± 0.0873 accuracy in the differential diagnosis of AD when using one single path. The described texture features ability in the differential diagnosis is somehow less powerful than those of the single intensity values at the nodes, but their utility can reside in the fact that they model structural changes that could be applied to longitudinal studies. These features could model in future work a neurodegeneration function that characterizes the texture at different stages of the disease, providing a different approach for an anatomical model of AD to those currently being used.

ACKNOWLEDGMENTS

This work was partly supported by the MINECO under the TEC2008-02113 and TEC2012-34306 projects and the Consejería de Economía, Innovación, Ciencia y Empleo (Junta de Andalucía, Spain) under the Excellence Projects P09-TIC-4530 and P11-TIC-7103, as well as the “Programa de fortalecimiento de las capacidades de I+D+I en las Universidades 2014-2015”, cofunded by the European Regional Development Fund (ERDF) under Project FC14-SAF-30.

Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimers Association; Alzheimers Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; MesoScale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer’s Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

REFERENCES

- [1] W. H. Organization *et al.*, *Dementia: a public health priority* (World Health Organization, 2012).

- [2] M. L. Schroeter, T. Stein, N. Maslowski and J. Neumann, Neural correlates of alzheimers disease and mild cognitive impairment: A systematic and quantitative meta-analysis involving 1351 patients, *NeuroImage* **47** (Oct 2009) p. 11961206.
- [3] A. Shiino, T. Watanabe, K. Maeda, E. Kotani, I. Akiguchi and M. Matsuda, Four subgroups of Alzheimer's disease based on patterns of atrophy using VBM and a unique pattern for early onset disease, *NeuroImage* **33** (October 2006) 17–26.
- [4] J. Ashburner, A fast diffeomorphic image registration algorithm., *Neuroimage* **38** (October 2007) 95–113.
- [5] N. Ayache, Analyzing 3D Images of the Brain, *NeuroImage* **4**(3) (1996) S34–S35.
- [6] X. Han, J. Jovicich, D. Salat, A. van der Kouwe, B. Quinn, S. Czanner, E. Busa, J. Pacheco, M. Albert, R. Killiany *et al.*, Reliability of mri-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer, *Neuroimage* **32**(1) (2006) 180–194.
- [7] L. Ronan, R. Pienaar, G. Williams, E. Bullmore, T. J. Crow, N. Roberts, P. B. Jones, J. Suckling and P. C. Fletcher, Intrinsic curvature: a marker of milimeter-scale tangential cortico-cortical connectivity?, *International Journal of Neural Systems* **21**(05) (2011) 351–366.
- [8] K. Friston, J. Ashburner, S. Kiebel, T. Nichols and W. Penny, *Statistical Parametric Mapping: The Analysis of Functional Brain Images* (Academic Press, 2007).
- [9] V. A. Kovalev, F. Kruggel, H.-J. Gertz and D. Y. von Cramon, Three-dimensional texture analysis of mri brain datasets, *Medical Imaging, IEEE Transactions on* **20**(5) (2001) 424–433.
- [10] Y. Fan, H. Rao, H. Hurt, J. Giannetta, M. Korczykowski, D. Shera, B. B. Avants, J. C. Gee, J. Wang and D. Shen, Multivariate examination of brain abnormality using both structural and functional MRI, *NeuroImage* **36** (July 2007) 1189–1199.
- [11] A. Ortiz, J. M. Górriz, J. Ramírez and F. Martínez-Murcia, Lvq-SVM based CAD tool applied to structural MRI for the diagnosis of the alzheimers disease, *Pattern Recognition Letters* **34** (Oct 2013) p. 17251733.
- [12] U. Yoon, J.-M. Lee, K. Im, Y.-W. Shin, B. H. Cho, I. Y. Kim, J. S. Kwon and S. I. Kim, Pattern classification using principal components of cortical thickness and its discriminative pattern in schizophrenia, *NeuroImage* **34** (February 2007) 1405–1415.
- [13] D. Unay, A. Ekin, M. Cetin, R. Jasinschi and A. Ercil, Robustness of local binary patterns in brain mr image analysis, *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Aug 2007).
- [14] U. Acharya, R. Yanti, J. Zheng, M. Krishnan, J. Tan, R. Martis and C. Lim, Automated diagnosis of epilepsy using cwt, hos and texture parameters., *International Journal of Neural Systems* **23**(03) (2013) p. 1350009.
- [15] F. Martinez-Murcia, J. Górriz, J. Ramírez, M. Moreno-Caballero, M. Gómez-Río, P. P. M. Initiative *et al.*, Parametrization of textural patterns in 123i-ioflupane imaging for the automatic detection of parkinsonism, *Medical physics* **41**(1) (2014) p. 012502.

- [16] F. J. Martínez-Murcia, J. M. Górriz, J. Ramírez, I. Alvarez Illán, D. Salas-González, F. Segovia and A. D. N. I. , Projecting mri brain images for the detection of alzheimer's disease., *Stud Health Technol Inform* **207** (2015) 225–233.
- [17] F. Martinez-Murcia, A. Ortiz, J. M. Górriz, J. Ramírez and I. Illán, A volumetric radial lbp projection of mri brain images for the diagnosis of alzheimers disease, *Artificial Computation in Biology and Medicine*, (Springer, 2015), pp. 19–28.
- [18] R. Haralick, K. Shanmugam and I. Dinstein, Textural features for image classification, *IEEE Transactions on Systems, Man and Cybernetics* **3**(6) (1973) 610–621.
- [19] M. Sikiö, K. K. Holli-Helenius, L. C. Harrison, P. Ryymin, H. Ruottinen, T. Saunamäki, H. J. Eskola, I. Elovaara and P. Dastidar, Mr image texture in parkinsons disease: a longitudinal study, *Acta Radiologica* **56**(1) (2015) 97–104.
- [20] H. Chen, M. Goldberg, M. Magdon-Ismael and W. Wallace, Reverse engineering a social agent-based hidden markov model - visage, *International Journal of Neural Systems* **18**(06) (2008) 491–526.
- [21] A. C. Evans, D. L. Collins, S. Mills, E. Brown, R. Kelly and T. M. Peters, 3d statistical neuroanatomical models from 305 mri volumes, *Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record., IEEE1993*, pp. 1813–1817.
- [22] J.-C. Yen, F.-J. Chang and S. Chang, A new criterion for automatic multilevel thresholding, *Image Processing, IEEE Transactions on* **4** (Mar 1995) 370–378.
- [23] L.-K. Soh and C. Tsatsoulis, Texture analysis of sar sea ice imagery using gray level co-occurrence matrices, *Geoscience and Remote Sensing, IEEE Transactions on* **37**(2) (1999) 780–795.
- [24] D. A. Clausi, An analysis of co-occurrence texture statistics as a function of grey level quantization, *Canadian Journal of remote sensing* **28**(1) (2002) 45–62.
- [25] E. Luders, C. Gaser, L. Jancke and G. Schlaug, A voxel-based approach to gray matter asymmetries, *NeuroImage* **22** (Jun 2004) 656–664.
- [26] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer and M. Joliot, Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain., *Neuroimage* **15** (Jan 2002) 273–289.
- [27] B. Dubois, H. H. Feldman, C. Jacova, S. T. DeKosky, P. Barberger-Gateau, J. Cummings, A. Delacourte, D. Galasko, S. Gauthier, G. Jicha and et al., Research criteria for the diagnosis of alzheimers disease: revising the nincdsadrda criteria, *The Lancet Neurology* **6** (Aug 2007) p. 734746.
- [28] D. Chan, N. C. Fox, R. I. Scahill, W. R. Crum, J. L. Whitwell, G. Leschziner, A. M. Rossor, J. M. Stevens, L. Cipelotti and M. N. Rossor, Patterns of temporal lobe atrophy in semantic dementia and alzheimer's disease, *Annals of neurology* **49**(4) (2001) 433–442.
- [29] J. C. Baron, G. Chételat, B. Desgranges, G. Perchey, B. Landeau, V. de la Sayette and F. Eustache, In vivo mapping of gray matter loss with voxel-based morphometry in mild alzheimer's disease., *Neuroimage* **14** (Aug 2001) 298–309.

- [30] L. W. de Jong, K. van der Hiele, I. M. Veer, J. J. Houwing, R. G. J. Westendorp, E. L. E. M. Bollen, P. W. de Bruin, H. A. M. Middelkoop, M. A. van Buchem and J. van der Grond, Strongly reduced volumes of putamen and thalamus in alzheimers disease: an MRI study, *Brain* **131** (Nov 2008) p. 32773285.
- [31] M. Graña, M. Termenon, A. Savio, A. Gonzalez-Pinto, J. Echeveste, J. Pérez and A. Besga, Computer aided diagnosis system for alzheimer disease using brain diffusion tensor imaging features selected by pearsons correlation, *Neuroscience Letters* **502** (Sep 2011) p. 225229.
- [32] D. A. Medina and M. Gavia, Diffusion tensor imaging investigations in alzheimer's disease: the resurgence of white matter compromise in the cortical dysfunction of the aging brain., *Neuropsychiatr Dis Treat* **4** (Aug 2008) 737–742.