NeuroImage 59 (2012) 2155–2166

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Detecting global and local hippocampal shape changes in Alzheimer's disease using statistical shape models

Kai-kai Shen [a,b,*], Jurgen Fripp [a], Fabrice Mériaudeau [b], Gaël Chételat [c,d], Olivier Salvado [a], Pierrick Bourgeat [a] and The Alzheimer's Disease Neuroimaging Initiative [1]

[a] Australian e-Health Research Centre, CSIRO ICT Centre, Herston, Queensland, Australia
[b] LE2I CNRS UMR 5158, Université de Bourgogne, Le Creusot, France
[c] Department of Nuclear Medicine and Centre for PET, and Department of Medicine, University of Melbourne, Austin Hospital, Melbourne, Victoria, Australia
[d] Inserm-EPHE-Université de Caen/Basse-Normandie, Unité U923, GIP Cyceron, CHU Côte de Nacre, Caen, France

ARTICLE INFO

Article history:
Received 19 January 2011
Revised 9 September 2011
Accepted 3 October 2011
Available online 14 October 2011

ABSTRACT

The hippocampus is affected at an early stage in the development of Alzheimer's disease (AD). With the use of structural magnetic resonance (MR) imaging, we can investigate the effect of AD on the morphology of the hippocampus. The hippocampal shape variations among a population can be usually described using statistical shape models (SSMs). Conventional SSMs model the modes of variations among the population via principal component analysis (PCA). Although these modes are representative of variations within the training data, they are not necessarily discriminative on labeled data or relevant to the differences between the subpopulations. We use the shape descriptors from SSM as features to classify AD from normal control (NC) cases. In this study, a Hotelling's T^2 test is performed to select a subset of landmarks which are used in PCA. The resulting variation modes are used as predictors of AD from NC. The discrimination ability of these predictors is evaluated in terms of their classification performances with bagged support vector machines (SVMs). Restricting the model to landmarks with better separation between AD and NC increases the discrimination power of SSM. The predictors extracted on the subregions also showed stronger correlation with the memory-related measurements such as Logical Memory, Auditory Verbal Learning Test (AVLT) and the memory subscores of Alzheimer Disease Assessment Scale (ADAS).

Introduction

Early detection and diagnosis of Alzheimer's disease (AD) is a challenging task. The hippocampus presents the highest rate of atrophy in the early stage of disease (Gosche et al., 2002), and has therefore become a popular target for early detection of AD through morphological analysis of magnetic resonance (MR) images.

Hippocampal atrophy in neurodegenerative diseases such as AD can be evaluated in terms of the global change in the volume of the hippocampus, as well as global and local changes in its shape. Hippocampal volumetry on MR images has been shown to be a useful tool in AD diagnosis (Chetelat and Baron, 2003; Dubois et al., 2007; Frisoni et al., 2010; Small et al., 2008). In various MRI studies, the tissue loss in the hippocampus was found in AD as well as in mild cognitive impairment (MCI) subjects (for review see Shi et al. (2009)). The hippocampal atrophy rates tracked in longitudinal studies have been found to be more aggressive in AD as compared to the normal aging (Barnes et al., 2007, 2009). Hippocampal volume as quantified by the automatic segmentation has been used to discriminate AD from normal control (NC) subjects, as well as cases with MCI (Chupin et al., 2009).

Volume alone provides significant discrimination ability, but is inadequate to fully describe the effect of the disease on the morphology of hippocampus. In addition to volumetry, hippocampal shape analysis also contributed to the understanding of the development of the disease. Regional differences between hippocampal subfields have been found in the neurodegenerative process of AD, with more pronounced neuron loss reported in CA1 and subiculum subfields (West et al. (1994); Bobinski et al. (1998); Rössler et al. (2002); Mueller and Weiner (2009); for review see Scher et al. (2007)). With the development of brain mapping techniques such as radial atrophy mapping (Frisoni et al., 2008; Thompson et al., 2004), and high-dimensional brain mapping (Csernansky et al., 2000; Wang et al., 2003, 2006), findings from neuroimaging studies into AD and MCI have also corroborated the locality of shape changes in CA1 and subiculum subfields along with global tissue reduction (Apostolova

* Corresponding author at: Australian e-Health Research Centre, CSIRO ICT Centre, Herston, Queensland, Australia.
E-mail address: Kaikai.Shen@gmail.com (K. Shen).

[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). 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://loni.ucla.edu/ADNI/Collaboration/ADNI_Authorship_list.pdf

1053-8119/$ – see front matter.
doi:10.1016/j.neuroimage.2011.10.014

et al., 2006; Chételat et al., 2008; Csernansky et al., 2005; Qiu et al., 2009; Scher et al., 2007). Both shape and volume of hippocampus have been exploited in the detection of AD and to predict the onset of the dementia (Csernansky et al., 2005).

Shape surface represented by spherical harmonics (SPHARM, Brechbühler et al. (1995); Gerig et al. (2001)) has been employed to model the hippocampus and to detect the shape changes caused by AD (Cuingnet et al., 2011; Gerardin et al., 2009; Gutman et al., 2009; Li et al., 2007; Zhou et al., 2009). SPHARM description of shape boundary has been combined with medial representation (Bouix et al., 2005; Pizer et al., 1999; Styner and Gerig, 2001) to obtain more localized shape features (Styner et al., 2004). In order to exploit the localized pattern of hippocampal atrophy, feature selection methodology has been employed to select patches on hippocampal surface thus improving the performance of shape features in diagnosing schizophrenia (Yushkevich et al., 2003) and AD (Zhou et al., 2010). Features characterizing local hippocampal surface deformation with respect to the mean shape have also been used to classify AD (Li et al., 2007).

Statistical shape models (SSMs) are another approach used to model the variability in the hippocampal shapes among the population (Davies, 2002; Davies et al., 2003). They usually rely on principal component analysis (PCA, Cootes et al. (1992)) to determine a lower dimensional subspace that accounts for most of the variability observed in a training set. However, the principal components spanning the subspace for valid shape samples are not necessarily discriminative between the subpopulations of interest. As global shape descriptor, each component models the variation of the whole hippocampal shape which may not be sensitive to local differences. Constraints upon isotropic lesser principal components restricting to the region of interest has been presented to give more meaningful reconstruction (Vermaak and Perez, 2003). In localized components analysis (LoCA, Alcantara et al. (2007)), spatial locality in the shape variation is explicitly optimized, which has been applied to hippocampal shape analysis and shown to produce local shape components strongly correlated with cognitive measurements (Xie et al., 2009).

In this study, our aim is to extract hippocampal shape descriptors capturing both global and local shape changes linked to the AD pathology by SSM, so as to improve the performance of disease classification using additional shape information. We try to model the localized disease-related shape changes by performing the shape analysis upon the regions that are found to be affected by AD. We first identify the hippocampal surface subregions that are significantly different between the AD group and the controls via Hotelling's T^2 test. Then we model the variations on these subregions using SSMs. Compared to the shape analysis of hippocampi via LoCA (Xie et al., 2009) which localizes the shape components while not specifically targeting the area related to the disease, we focus our shape analysis only on the regions that we marked as different, extracting the principal shape components on these regions. With the use of machine learning techniques, classifiers are able to learn the difference between AD and NC from the shape information as quantified by shape models. Extra information concerning AD pathology can be obtained by adding hippocampal shape descriptors in addition to the volumetry. We use the morphological variation on these regions as variables to describe the AD pathology as they improve the discrimination between the classes and are correlated with the measures of memory decline associated with the disease.

Materials and methods

Overview

We model the morphology of hippocampus by SSMs, and use the shape descriptors derived from the SSM to detect the effect of AD on the hippocampal size and its shape. Descriptors produced by different models serve as features for machine learning algorithms and are evaluated in terms of their prediction performance in distinguishing AD from NC.

The method used in this paper to extract relevant shape information from SSMs can be divided into two steps: 1) the localization step, and 2) the shape modeling step. The processing pipeline is shown in Fig. 1. In the localization step, we build up the correspondence on hippocampal surface over a training set of both NC and AD subjects. Once the correspondence problem is solved, all the landmarks on the hippocampal surfaces are aligned by Procrustes analysis (Gower, 1975). A statistical test can be performed on each landmark to evaluate the significance of the difference between the distributions of aligned points of the NC group and the AD group.

The resulting significance map produced by the statistical tests in the localization step can be thresholded to obtain a surface mask of landmarks. In the shape modeling step, we apply these masks to the hippocampal surface to select a subset of hippocampal landmarks separating the subpopulations at a given statistical significance. The selected subregional landmarks are again aligned by Procrustes analysis and a PCA is performed on the aligned subregional landmarks. The coefficients of principal components describing local shape of hippocampus can thus be calculated.

In both the localization step and the shape modeling step, the shapes or selected shape patches are aligned by Procrustes analysis, which can be performed either through rigid-body transformations or through similarity transformations. In our experiments, both alignments are performed at each step and their performances are evaluated and compared. We evaluate the subregional shape models based on their discrimination ability against the control group and the correlation with measurements of memory. The coefficients of modes describing the shape variations are used as features for disease classification, and the evaluation of the subregional models are based on the performance of the classifiers.

Materials

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public–private partnership. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

The Principal Investigator of this initiative is Michael W. Weiner, M. D., VA Medical Center and University of California, San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, ages 55 to 90, to participate in the research—approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years, and 200 people with early AD to be followed for 2 years.

The hippocampal volumes used in this article are semiautomated segmentations provided by ADNI, using high-dimensional brain mapping tool SNT, commercially available from Medtronic Surgical Navigation Technologies (Louisville, CO). SNT hippocampal volumetry has been previously validated on the normal aging, MCI and AD subjects (Hsu et al., 2002). It first uses 22 control points manually placed on the individual brain MRI as local landmarks. Fluid image transformation is then used to match the individual brains to a template brain

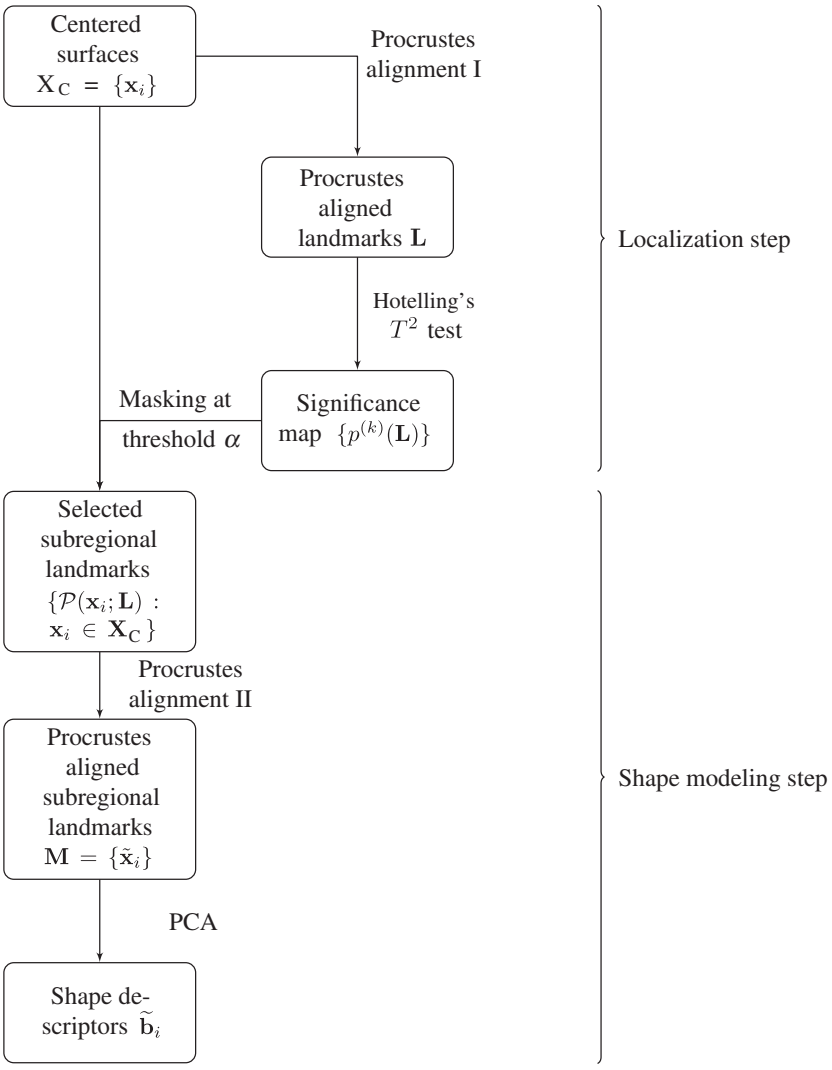

Fig. 1. The pipeline of local shape descriptor extraction. The shape surfaces are first centered and aligned via Procrustes alignment I. A Hotelling's T^2 test is performed on each landmark of the aligned surfaces to obtain a significance mask, which is used to threshold the previously centered shape surfaces. The selected landmarks are aligned by Procrustes alignment II, and a Principal Component Analysis (PCA) is performed to extract the shape descriptors.

(Christensen et al., 1997). The segmentations were manually edited by qualified reviewers if the boundaries delineated by SNT were not accurate.

The hippocampal volumes were divided into a training set and a testing set. The demographic information is listed in Table 1.

Statistical shape models

An SSM is built upon a training set of hippocampal surfaces, in which the n_P landmarks on the surface of hippocampus are reparameterized to avoid the false variation induced by incorrect correspondences. This correspondence problem can be solved by optimization of an information theoretic objective function based on the minimum description length (MDL) of the shape model (Davies et al., 2002). In this study, the hippocampal volumes in the training set are first registered and aligned via rigid body transformations. The correspondence of landmarks over the training set is established by a groupwise optimization and fluid regularization on the shape image (Davies et al., 2008).

Once the correspondence is established, the surfaces can be aligned by using Procrustes analysis, either through rigid-body or similarity transformations. The volume information is preserved after rigid-body transformations. Procrustes analysis aligns the training data via rigid-body transformations to the size-and-shape space $S\Sigma_3^{n_P}$ (see Dryden and Mardia (1998)) in which the variation among the data would be driven by the change in both the size and the shape of hippocampus. If the training samples are aligned via similarity transformations, the surfaces will be rescaled to normalize the hippocampal volume. Thus we have the training set in the shape space $\Sigma_3^{n_P}$. As shape is defined as the remaining information 'when the differences which can be attributed to translations, rotations, and dilations have been quotiented out' (Kendall, 1984), the normalization of the volume by similarity transform enables the SSM to be more specific to the change in shape rather than incorporating variations in the sizes of hippocampi. In this paper, SSMs are built on both rigid-body and similarity aligned surfaces.

Table 1
Demographic information of subjects used in the study.

	Diagnosis	Number (M/F)	Age	Yeas of education
Training set	NC	32/28	76.6(5.2)	16.2(2.7)
	AD	31/29	76.1(7.1)	14.5(2.9)
Testing set	NC	37/41	76.5(4.9)	15.5(3.1)
	AD	20/19	76.5(6.8)	14.6(3.9)
Total	NC	69/69	76.6(5.0)	15.8(2.9)
	AD	51/48	76.2(7.0)	14.6(3.3)

For a set of n_S hippocampal surfaces $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{n_S}\}$ with established correspondence, each sample surface is represented by the coordinates of its n_P landmarks concatenated as a $3n_P$-vector

$$\mathbf{x}_i = \left(x_i^{(1)}, y_i^{(1)}, z_i^{(1)}, x_i^{(2)}, y_i^{(2)}, z_i^{(2)}, \cdots, x_i^{(n_P)}, y_i^{(n_P)}, z_i^{(n_P)}\right)^T \in \mathbb{R}^{3n_P}, \tag{1}$$

where $\mathbf{p}_i^{(k)} = (x_i^{(k)}, y_i^{(k)}, z_i^{(k)})^T \in \mathbf{x}_i$ is the position of the kth landmark point on the $\mathbf{x}_i$. If the surfaces in $\{\mathbf{x}_i\}$ are aligned using Procrustes analysis, either rigidly or via similarity transformations, a PCA can be performed on the data matrix $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{n_S})$, and the shape data can be expressed as

$$\mathbf{x}_i = \mu + \mathbf{W} b_i \tag{2}$$

where μ is the mean of $\{\mathbf{x}_i\}$, $\mathbf{W}$ is the matrix consisting of the eigenvectors of the covariance matrix of the training data, and the elements in vector $\mathbf{b}_i$ are the parameters characterizing the ith shape.

Localization step

In the localization step, we identify the regions affected by atrophy by analyzing the spatial distribution of each landmark across the training examples and the deformation between the two subpopulations. We begin with the set $\mathbf{X}_C$ in which the correspondence over the training set is given and the landmarks are only centered to the origin without further alignment. We use Procrustes analysis to obtain an aligned ensemble $\mathbf{L}$ of surfaces. For the kth landmark in $\mathbf{L}$, a Hotelling's T^2 test (Styner et al., 2007) can be used to assess the statistical significance of the landmark separating the NC from the AD group

$$p^{(k)}\left(\mathbf{L}\right) = P\left(T^2 > \frac{n_{NC} \cdot n_{AD}}{n_{NC} + n_{AD}} \left(\mu_{NC}^{(k)} - \mu_{AD}^{(k)}\right) T\Sigma^{(k)^{-1}} \left(\mu_{NC}^{(k)} - {}_{AD}^{(k)}\right)\right), \tag{3}$$

where $\mu_{NC}^{(k)}$ and $\mu_{AD}^{(k)}$ are the mean location of the kth landmark in each group, and

$$\Sigma^{(k)} = \frac{1}{n_{NC} + n_{AD} - 2} \left(\sum_{i \in NC} \left(\mathbf{p}_i^{(k)} - \mu_{NC}^{(k)}\right) \left(\mathbf{p}_i^{(k)} - \mu_{NC}^{(k)}\right)^T + \sum_{i \in AD} \left(\mathbf{p}_i^{(k)} - \mu_{AD}^{(k)}\right) \left(\mathbf{p}_i^{(k)} - {}_{AD}^{(k)}\right)^T \right) \tag{4}$$

By thresholding the p-map $p^{(k)}(\mathbf{L})$, we select only the subset of landmarks showing significant difference between NC and AD groups. This can be viewed as a projection $\mathcal{P}$ of the shape surface to the regions more relevant to the pathology

$$\mathcal{P}\left(\cdot; \mathbf{L}\right) : \mathbb{R}^{3n_P} \mapsto \mathbb{R}^{3m} \quad \mathbf{x} \mapsto \left(\tilde{x}^{(1)}, \tilde{y}^{(1)}, \tilde{z}^{(1)}, \tilde{x}^{(2)}, \tilde{y}^{(2)}, \tilde{z}^{(2)}, \cdots, \tilde{x}^{(m)}, \tilde{y}^{(m)}, \tilde{z}^{(m)}\right)^T, \tag{5}$$

which consists of $m (< n_P)$ landmarks $\left(\tilde{x}^{(l)}, \tilde{y}^{(l)}, \tilde{z}^{(l)}\right)^T \in \left\{ \left(x^{(k)}, y^{(k)}, z^{(k)}\right)^T \in \mathbf{x} : p^{(k)}(\mathbf{L}) < \alpha \right\}$ found to separate the NC from the disease group at significance level α.

Here the distribution of the surface landmarks depends on how the shapes in the SSM are aligned, and the variation between the subgroups will differ as rigid-body or similarity transformations can be chosen to align the shapes.

Shape modeling step

In the shape modeling step, instead of performing a PCA on the data consisting of all n_P landmarks of the surface, only the subset of landmarks identified as different between subpopulations in the localization step are used. We mask the centered landmarks in $\mathbf{X}_c$ and re-align the masked landmarks $\{\mathcal{P}(\mathbf{x}_i; \mathbf{L}) : \mathbf{x}_i \in \mathbf{X}_C\}$ by Procrustes analysis to form the training set

$$\mathbf{M} = \{\tilde{\mathbf{x}}_i : \tilde{\mathbf{x}}_i \text{ is Procrustes aligned } \mathcal{P}(\mathbf{x}_i; \mathbf{L}), \mathbf{x}_i \in \mathbf{X}_C\} \tag{6}$$

for the subregional SSM. A PCA is performed on $\mathbf{M}$

$$\tilde{\mathbf{x}}_i = \mu_{\mathbf{M}} + \tilde{\mathbf{W}}\, \tilde{\mathbf{b}}_i \tag{7}$$

where $\mu_{\mathbf{M}}$ is the mean of samples in $\mathbf{M}$, $\tilde{\mathbf{W}}$ is the matrix of eigenvectors describing the variation modes from significantly different landmarks, and $\tilde{\mathbf{b}}_i$ is the vector of coefficients of each mode.

Extrapolation of the SSM to test shapes

For a given target shape not in the training set, the correspondence established when building the SSM needs first to be propagated to the target shape surface, so that the target shape can be represented in the same vector space as the training data. The target shape surfaces are usually generated from the segmentation, which can be smoothed using a windowed sinc function (Taubin et al., 1996). The passband of the sinc filter is set to 0.1. The SSM of the whole hippocampal surface is deformed with respect to the parameters $\mathbf{b}$ to fit the smoothed target surface, minimizing a distance metric between the SSM generated surface and the target.

We use an L_1 distance between two shapes which is less sensitive to outliers and noise (Ruto and Buxton, 2005). Without the assumption of correspondence between two configurations of point set $\mathbf{x}$ and $\mathbf{y}$, the distance between their triangulated surfaces $\mathcal{S}(\mathbf{x})$ and $\mathcal{S}(\mathbf{y})$ can be defined in a symmetrical manner as

$$d_S\left(\mathcal{S}\left(\mathbf{x}\right), \mathcal{S}\left(\mathbf{y}\right)\right) = \sum_{\mathbf{p} \in \mathbf{x}} d_p\left(\mathbf{p}, \mathcal{S}\left(\mathbf{y}\right)\right) + \sum_{\mathbf{q} \in \mathbf{y}} d_p\left(\mathbf{q}, \mathcal{S}\left(\mathbf{x}\right)\right), \tag{8}$$

where $d_p(\mathbf{p}, \mathcal{S})$ is the Euclidean distance from the point $\mathbf{p}$ to the closest point on surface $\mathcal{S}$. Thus we can fit the SSM to the target surface $\mathcal{S}(\mathbf{y})$ by the optimization of parameters using Powell's algorithm (Powell, 1964)

$$(\mathcal{T}_{\mathbf{y}}, \mathbf{b}_{\mathbf{y}}) = \arg\min_{(\mathcal{T}, \mathbf{b})} d_S(\mathcal{S}(\mathcal{T}(\mu + \mathbf{W}\mathbf{b})), \mathcal{S}(\mathbf{y})) \tag{9}$$

where $\mathcal{T}$ is a similarity transformation with 7 degrees of freedom $(\mathbb{R}^3 \rtimes (\mathbb{R}^+ \times SO(3)))$, and $\mathbf{b}$ the parameters describes the deformation of the shape from the mean of the training set according to the SSM. For the kth landmark $\mathbf{p}^{(k)} \in \mathcal{T}_{\mathbf{y}}(\mu + \mathbf{W}\mathbf{b}_{\mathbf{y}})$ generated by SSM, we estimate the closest point on the surface $p'^{(k)} \in \mathcal{S}(\mathbf{y})$ as the corresponding landmark point. By aligning the shape vector $y' = \left(p'^{(1)T}, p'^{(2)T}, \cdots, p'^{(n_P)T}\right)^T$ to the shapes in the SSM through Procrustes analysis, we have the $3n_P$-D shape vector representing the surface $\mathcal{S}(\mathbf{y})$ in the same space as the shapes in the SSM, ready for further analysis. The diagram of this process is shown in Fig. 2.

Classification method

The shape descriptors from SSM are evaluated in terms of their performance when being used as features in classification algorithms. Support vector machines (SVMs, Vapnik (1995)) are widely used in solving general classification problems, and have been applied to AD diagnosis (e.g. Vemuri et al. (2008); Klöppel et al. (2008), for review see Cuingnet et al. (2011)). SVMs typically find the optimal hyperplane with the largest margin separating the classes. The computation of the optimal solution requires only the inner product of feature vectors which can be exploited by the substitution of a kernel for the inner product mapping the feature space to higher dimension (Schölkopf and Smola, 2002). We use the SVM classifier to test the

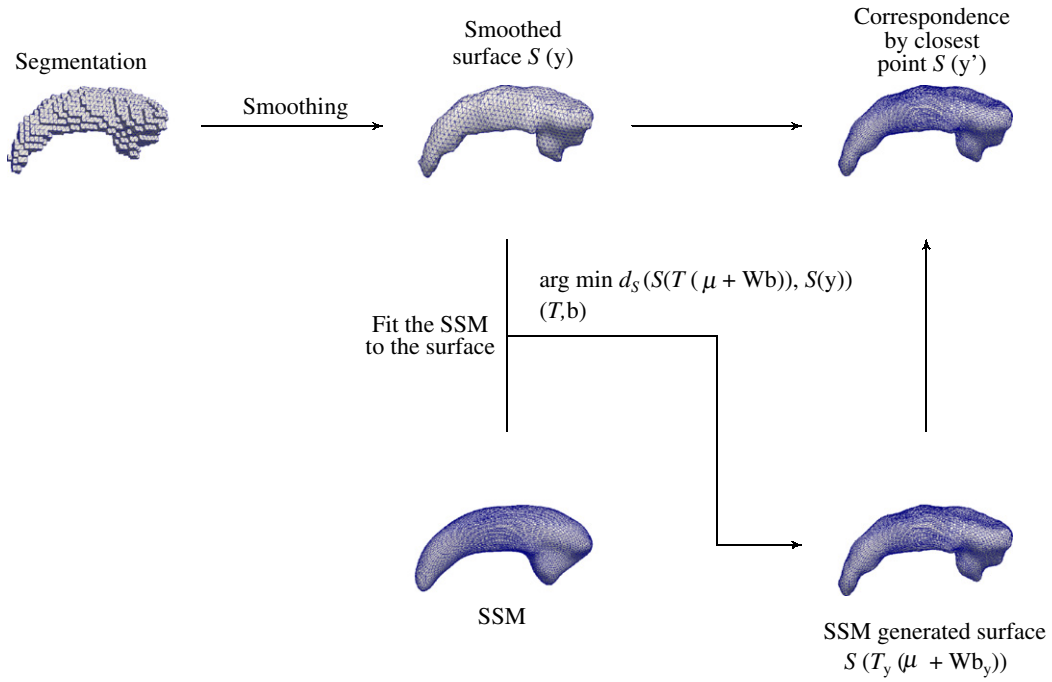

Fig. 2. Representation of a hippocampal shape with Statistical Shape Model (SSM) The marching cube surface is first smoothed, which is fitted by an SSM minimizing the L_1 distance between the model generated surface and the smoothed shape surface. The correspondence is established between the model landmarks and the smoothed surface by the closest point. The corresponding points on the shape surface are used as representation of the shape in the models space.

discrimination ability of the features. In this study, we choose the radial basis function (RBF) as the kernel in SVM. The feature vector of each shape **x** for classification consists of the coefficients of variation modes, as extracted from Eqs. (2) or (7).

In order to avoid modeling noise, less significant components produced by PCA are excluded from the feature set based on the assumption that these components tend to account more for noise than for meaningful shape information. Feature selection is usually required in classification tasks using SVM to reduce variance by eliminating the less relevant or noisier features, and to improve the classifier performance with fewer but more discriminative inputs. Commonly used wrapper methods for feature selection have the potential risk of overfitting the training data (Loughrey and Cunningham, 2005; Reunanen, 2003). The heuristic strategies searching for an optimal subset of features may select a combination not generalizing well on a separate test set. Since the aim of this study is to evaluate the prediction performance of different categories of shape descriptors, the performance of the classifier on a selected subset of features (not guaranteed to be optimal either on the training or the test set) does not represent the overall discriminant ability of the features derived from different shape models. It is therefore preferable to adopt a methodology that takes into account all features in order to assess the respective model that produces these features.

Bootstrap aggregation, or bagging (Breiman, 1996a), is another approach to reduce the variance in the learning algorithm. It is based on the repeating bootstraps of the training set, and trains the classifier using the resampled data. Multiple classifiers learned from the resampled data are often combined by a majority voting. When bagging is used, better performance is usually achieved with larger feature set without discarding weakly informative features (Munson and Caruana, 2009). We thus can train the bagged classifier on the whole set of features to be evaluated and use the prediction performance of the bagged classifier as a measurement of the discriminant ability of the feature set.

In addition to the shape variations, the hippocampal volumes of both hemispheres normalized by the total intracranial volume (TIV) are used as independent features. To compare the discrimination ability of the volume and the shape of the hippocampus, we evaluate the classification performance of shape descriptors both with and without the normalized volume features. The TIV is calculated by summing the volume of white matter (WM), grey matter (GM) and cerebrospinal fluid (CSF). For each subject, the T1- and T2-weighted MR images were classified into WM, GM and CSF using an implementation of the expectation maximization segmentation algorithm (Acosta et al., 2009). In brief, this algorithm models the image intensity histogram using a mixture of Gaussian distribution. The parameters of the Gaussian mixture are iteratively updated using an expectation maximization approach. For this study, 6 Gaussian distributions were used: 1 for each of the main tissue types, namely WM, GM, and CSF, and 3 for non-brain tissues. To provide initialization and spatial consistency to the expectation maximization algorithm, a template and associated probability maps of GM, WM, and CSF were spatially normalized to each individual scan, first through an affine registration using a robust block matching approach (Ourselin et al., 2001) with 12 degrees of freedom, and then using a diffeomorphic demons nonrigid registration (Vercauteren et al., 2009). The expectation maximization algorithm computed probability maps for each tissue type, which were discretized by assigning each voxel to its most likely tissue type.

In this study, the data set used to train the classifier was resampled with replacement for 25 times, and the ensemble of classifiers consisted of SVMs with RBF kernel each of which were trained on the resampled data. The cost parameter C and the scale parameter of the kernel γ are tuned by a grid search over the range $C=2^{-7}$, 2^{-6}, ..., 2^2, 2^3 and $\gamma=2^{-2}$, 2^{-1}, ..., 2^7, 2^8 using the entire training set with leave-one-out cross validation. The parameters giving the best performance over the training data are chosen for each individual SVM. The predictions of 25 SVMs were combined by a voting rule to produce the final output. The validation on the training set with varying the threshold α, and experiments using separated training and testing sets were performed.

We use the out-of-bag estimation (OOB, Breiman, 1996b) to evaluate the performance of the bagging classifier on the training set. For each resample of the training data with replacement, there are approximately one-third of the training data left out of the bootstrap resampling. A classification of each left-out (i.e. 'out-of-bag') case

can be obtained from the SVM trained using the selected (i.e. 'in-bag') sample. The classification of each case in the training set can be determined by a voting rule after multiple runs of resampling. The OOB estimation can be obtained from average accuracy on the left-out data. In our experiment, the OOB estimates are carried out by 70 resamplings, which result in each training case being left out and classified for approximately 25 times on average.

In order to compare with the disease classification result reported by Cuingnet et al. (2011) using hippocampal shape information (Gerardin et al., 2009), experiments using the linear C-SVM on SSM features were also carried out. The cost parameter C was chosen by searching over the range 10^{-5}, $10^{-4.5}$, ..., $10^{2.5}$, 10^{3} and leave-one-out cross validation on the training set, in the same manner as in Cuingnet et al. (2011).

Correlating the shape variation with memory performance

Hippocampus plays a critical role in the memory formation, and its atrophy is associated with the memory decline in AD (Heun et al., 1997; Petersen et al., 2000). We assessed our global and local hippocampal SSMs in terms of their relevance to the memory decline by the correlation analysis between the shape descriptors derived from these models and the scores from memory tests. The results from Logical Memory test (immediate and delayed recall), Auditory Verbal Learning Test (AVLT, 30-minute delay, Ivnik et al. (1992)), and the memory subscores of Alzheimer Disease Assessment Scale (ADAS, including word recall, delayed word recall, orientation, word recognition and recall instructions) were chosen as the measurements of memory performance. Spearman's correlation coefficient ρ was used to compare the hippocampal shape variation with the memory scores, which were corrected for age, sex and years of education in the experiments. For comparison, we also project the shape samples onto the direction of mean difference $\mu_{NC} - \mu_{AD}$, and perform the correlation analysis between the memory scores and the shape component along the mean difference direction.

Experimental results

SSM

Based on the choice of the transformations aligning the surfaces in the localization step and in the shape modeling step, we can derive the variation modes of the hippocampal shapes in different ways. The combinations of alignment in the SSMs are listed in Table 2.

1. In the localization step, the significance maps are produced from rigidly aligned size-and-shape set $\mathbf{L}_R$, and in the shape modeling step the subregional SSM is built on rigidly aligned the size-and-shape set $\mathbf{M}_R$;
2. in the localization step, the significance maps are produced from rigidly aligned size-and-shape set $\mathbf{L}_R$, and in the shape modeling step the subregional SSM is built on similarity aligned the shape set $\mathbf{M}_S$;
3. in the localization step, the significance maps are produced from similarity aligned shape set $\mathbf{L}_S$, and in the shape modeling step the subregional SSM is built on rigidly aligned the size-and-shape set $\mathbf{M}_R$;
4. in the localization step, the significance maps are produced from similarity aligned shape set $\mathbf{L}_S$, and in the shape modeling step the subregional SSM is built on similarity aligned the shape set $\mathbf{M}_S$.

Table 2
Choices of alignment in SSMs.

		Procrustes alignment II in shape modeling	
		$\mathbf{M}_R$	$\mathbf{M}_S$
Procrustes alignment I in localization	$\mathbf{L}_R$	1	2
	$\mathbf{L}_S$	3	4

In the experiments, the hippocampi on both the right and the left hand sides of 60 NC and 60 AD subjects from ADNI data were used as a training set to build the SSM. Hotelling's T^2 test was performed on each SSM between the AD and NC groups, with the resulting significance maps shown in Fig. 3. The results after the thresholding are shown in Fig. 4.

We vary the thresholds of the significance level when selecting the landmarks at $\alpha = 0.0001$, 0.0002, ..., 0.0009, 0.001, 0.002, ..., 0.009, 0.01, 0.02, ..., 0.09, 0.1, 0.2, ..., 0.9. The number of landmarks thresholded at each significance level is plotted in Supp. Fig. 2. For the purpose of comparison, conventional PCA on all the landmarks was also performed, using models $\mathbf{M}_R$ and $\mathbf{M}_S$.

Disease classification of AD using regional SSM

The first 15 principal components in each of the SSM were used as features for classification, which accounted for approximately 90% of the total variance in the training set (Heimann and Meinzer, 2009). The rest of the modes individually contributed at most 1.0% of variation. (For the proportions of the total variance explained by the first 15 modes, see Supp. Fig. 1.) Thus for a given subject, each choice of SSM yielded 30 variation modes (15 left + 15 right).

The experiments of classification using bagged SVMs on the training set and the separated testing set were both repeated for 20 times. The results of average OOB accuracy on the training set and the results on the separate testing set of bagged SVMs are plotted in Fig. 5. In the experiment on the testing set, the classifier ensemble was trained at each time of the test on the training data with 25-trial bootstrapping, and then tested on the testing set. The average sensitivity (i.e. the proportion in the actual diseased subjects which is correctly identified by the classifier as positive) and specificity (i.e. the proportion in the actual control subjects which is correctly identified by the classifier as negative) for the disease classification are available in Supp. Figs. 3 and 4.

As a baseline, using only the TIV normalized volume gave 81.0% (OOB), 80.6%(testing set) accuracy. The best OOB performance on the training set (81.8% accuracy) is achieved using rigid-body aligned SSM ($\mathbf{M}_R$) on the selected landmarks ($\mathbf{L}_S, \alpha = 0.05$) with additional volume feature. On the separate testing set the highest accuracy is 88.0% ($\mathbf{M}_R, \mathbf{L}_S, \alpha = 0.07$).

With the features produced by SSMs alone, the best OOB accuracy is 80.2%, using the localized SSM ($\mathbf{M}_R, \mathbf{L}_S, \alpha = 0.4$), and the accuracy on the testing set reaches 88.3% ($\mathbf{M}_R, L_S, \alpha = 0.1$) The result of the linear C-SVM with this set of SSM features ($\mathbf{M}_R, \mathbf{L}_S, \alpha = 0.1$) gave 87.2% sensitivity and 79.5% specificity, which is comparable to the result of 69% sensitivity and 84% specificity reported by Cuingnet et al. (2011).

Correlation of hippocampal SSM descriptors with memory scores

The results of the correlation between the memory scores and the hippocampal shape variation over all landmarks, landmarks selected with threshold $\alpha = 0.1$, and threshold $\alpha = 0.01$ are shown in Fig. 6. The component in each SSM with the strongest correlation with the memory scores is plotted. The principal variations in the model on the selected subregions are shown to be better correlated with memory scores. Stronger correlations were found by the models built on $\mathbf{L}_S$ identified regions. The highest correlation between the memory score and the hippocampal shape was captured by similarity aligned SSM ($\mathbf{M}_S$) on the hippocampal subregions masked using $\mathbf{L}_S$ with threshold at $\alpha = 0.01$, in which the first principal component was the most correlated component with AVLT, Logical Memory scores (both

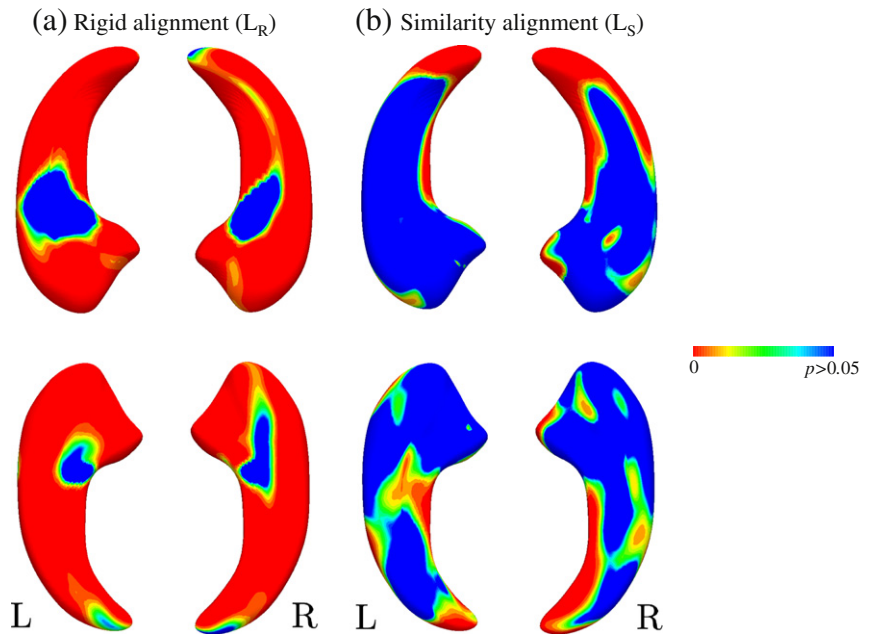

Fig. 3. The significance map by Hotelling's T^2 test. For each landmark in the training set, a Hotelling's T^2 test is carried out between the normal control (NC) group, and the Alzheimer's disease (AD) group, resulting the map of p-values. Top: superior view; bottom: inferior view.

immediate and delayed recall), and memory subscores of ADAS for both right and left hippocampus (visualization in Fig. 7). The shape components extracted from subregions with significant shape differences indicate a better representation of the effect of the disease on the shape of the hippocampus. The mean differences between the NC and AD groups in the training set are shown in Figs. 8 and 9.

Discussion

Shape model and correspondences

In our current setting, the correspondence over the training set is MDL optimized, and propagated to the testing set via closest point. This is an economical solution while suboptimal to the optimization of MDL over the testing set. In practice, however, we found that the classification accuracy was not lowered when using the correspondences by closest point on the separate testing set as compared to the cross validation on training set equipped with MDL-optimized correspondence. Therefore we opted for the closest-point as a more practical method to process the correspondences on the testing set.

Identification of atrophy affected subregions

The results of the Hotelling's T^2 test superimposed on the hippocampal surface were compared to the maps showing the location of the different hippocampal subfields (La Joie et al., 2010). Regions of

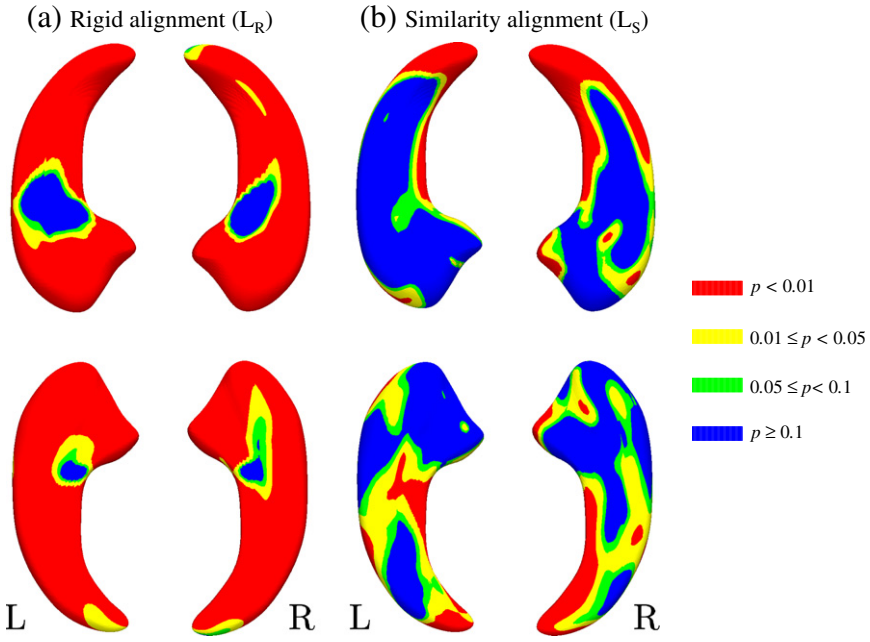

Fig. 4. Thresholded significance map, by thresholding p-values in the significance map Fig. 3.

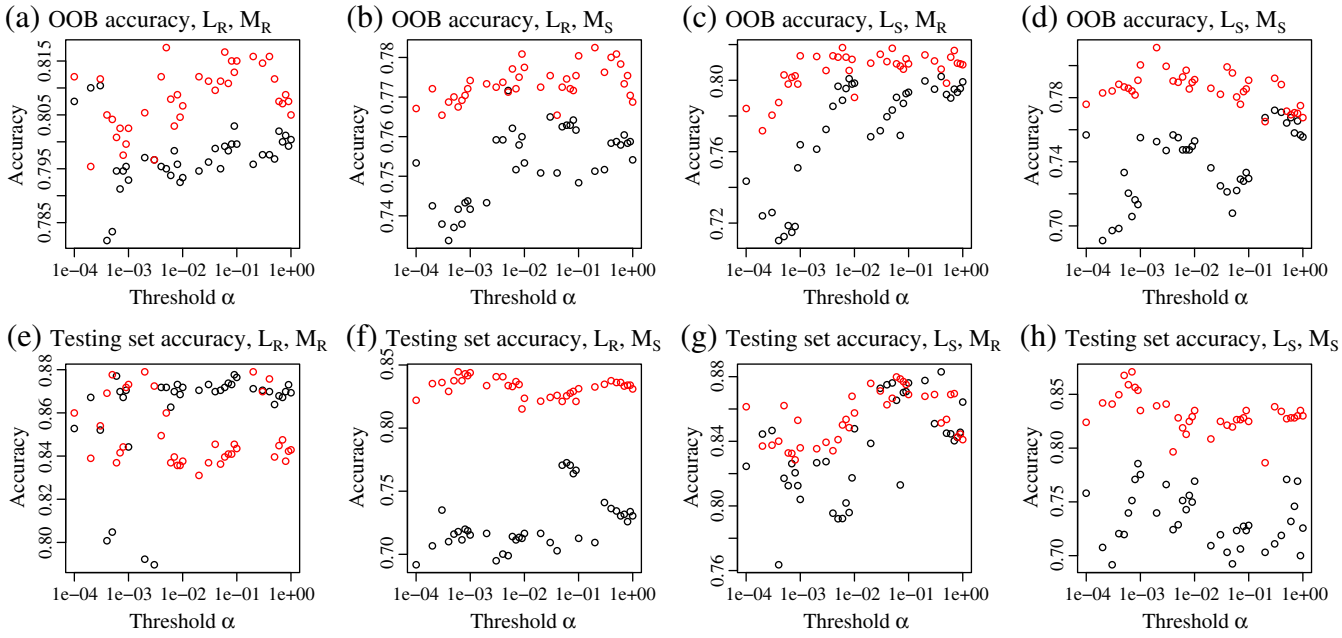

Fig. 5. The accuracy of the disease classification using bagged support vector machines (SVMs) with varying thresholds in the landmark selection. The out-of-bag (OOB) estimates is used for the measure of the training error (the first row), and the performance of the bagged classifier on a separated testing set is also presented. Black circles: shape features from the Statistical Shape Model (SSM) only; red circles: shape features with additional volume features.

significant differences between AD and NC were found mainly located in the CA1 and subiculum subfields (Fig. 3(a)), which is consistent with previous studies (Apostolova et al., 2010; Chételat et al., 2008; Mueller et al., 2010). When the effect of volume change was removed in SSM by similarity alignment (Fig. 3(b)), the regions of significant shape difference between the two groups revealed a different profile with more localized effects mainly located in the posterior hippocampus. Because global changes were mainly driven by the predominant effects in the CA1 and subiculum subfields, it is not surprising that the changes independent from these main effects reveals a different pattern.

It should be pointed out that the hippocampal subregions affected by AD were identified in this work by comparing the AD subpopulation with NC, which did not take into account the association of AD and its cognitive outcome with other factors such as age and level of education. Hippocampal atrophy has been shown to develop during the course of normal aging (Chételat et al., 2008; Jack et al., 1998; La Joie et al., 2010; Malykhin et al., 2008). The shape changes and variations modeled in this work reflect the combined effect of aging and AD, which may limit the discrimination ability of the shape analysis. In the future work it will be of interest to explicitly control the effects of normal aging on the hippocampus and other variables such as years of education, in order to isolate the disease from these effects.

Disease classification using SSM

Using the SSM descriptors combining both size and shape information provided better discrimination than using only hippocampal volume to classify AD from NC. In general the descriptors of size-and-shape from $\mathbf{M}_R$ outperformed features produced by $\mathbf{M}_S$, because volume alone is a good discriminant. Since the changes modeled by $\mathbf{M}_R$ were driven by both size and shape, adding volume to the features extracted using $\mathbf{M}_R$ and $\mathbf{M}_S$ increased both of their accuracy, but to a much less extent for $\mathbf{M}_R$ compared to $\mathbf{M}_S$. Using shape information therefore provides additional discrimination power to volumetry.

Restricting the analysis to the subregions affected by the disease increased the discrimination ability of the SSM approach by capturing localized differences between the subpopulations. Whole surface SSMs are able to describe the global shape or size-and-shape of the biological object, but are not sensitive to the deformations limited to specific areas on the object surface. Localizing the PCA to subregions with significant shape difference ($\mathbf{L}_S$) on the surface produced overall better discrimination between NC and AD than using all the points. When TIV normalized volume was added as additional features to the shape features, the best classification results were obtained using the SSM built using $\mathbf{M}_R$ on the hippocampal subregions selected by $\mathbf{L}_S$.

In particular, using $\mathbf{L}_S$ in the localization step gave more informative surface masks than $\mathbf{L}_R$ when describing the atrophy pattern in the disease classification. Subregional masks derived from rigid-body aligned localization model $\mathbf{L}_R$ tended to be predominantly representing changes in global scale due to the volume reduction. Since the sizes of samples in $\mathbf{L}_S$ were normalized, the global difference in the size between subpopulations was filtered out. The resulting subregions found by $\mathbf{L}_S$ to be significantly different between AD and controls were more localized to subfields such as CA1 and subiculum.

The best classification performance was obtained by the models describing local size-and-shape variations ($\mathbf{L}_S + \mathbf{M}_R$). It is noticed that the shape changes may appear on a local scale as deformation involving both size and shape of the subregion. The localized size-and-shape model $\mathbf{M}_R$ on hippocampal subregions can therefore detect the shape changes in the form of local volume change which were filtered out in $\mathbf{M}_S$. This may explain the better performance local size-and-shape variations ($\mathbf{L}_S + \mathbf{M}_R$) than the shape models ($\mathbf{L}_S + \mathbf{M}_S$) on the local scale, in addition to the discriminant ability of the size factor present in $\mathbf{M}_R$.

The current results demonstrate the potentiality in improving discrimination ability of SSM by localizing the model to hippocampal subregions. More targeted localization of the atrophy than simple thresholding may be developed in the future work for the optimal selection of landmarks. In addition to the statistical difference in the landmark position, histological and anatomical knowledge of hippocampus may be incorporated into the identification of regions in order to capture the shape variation that is more closely linked to the AD pathology.

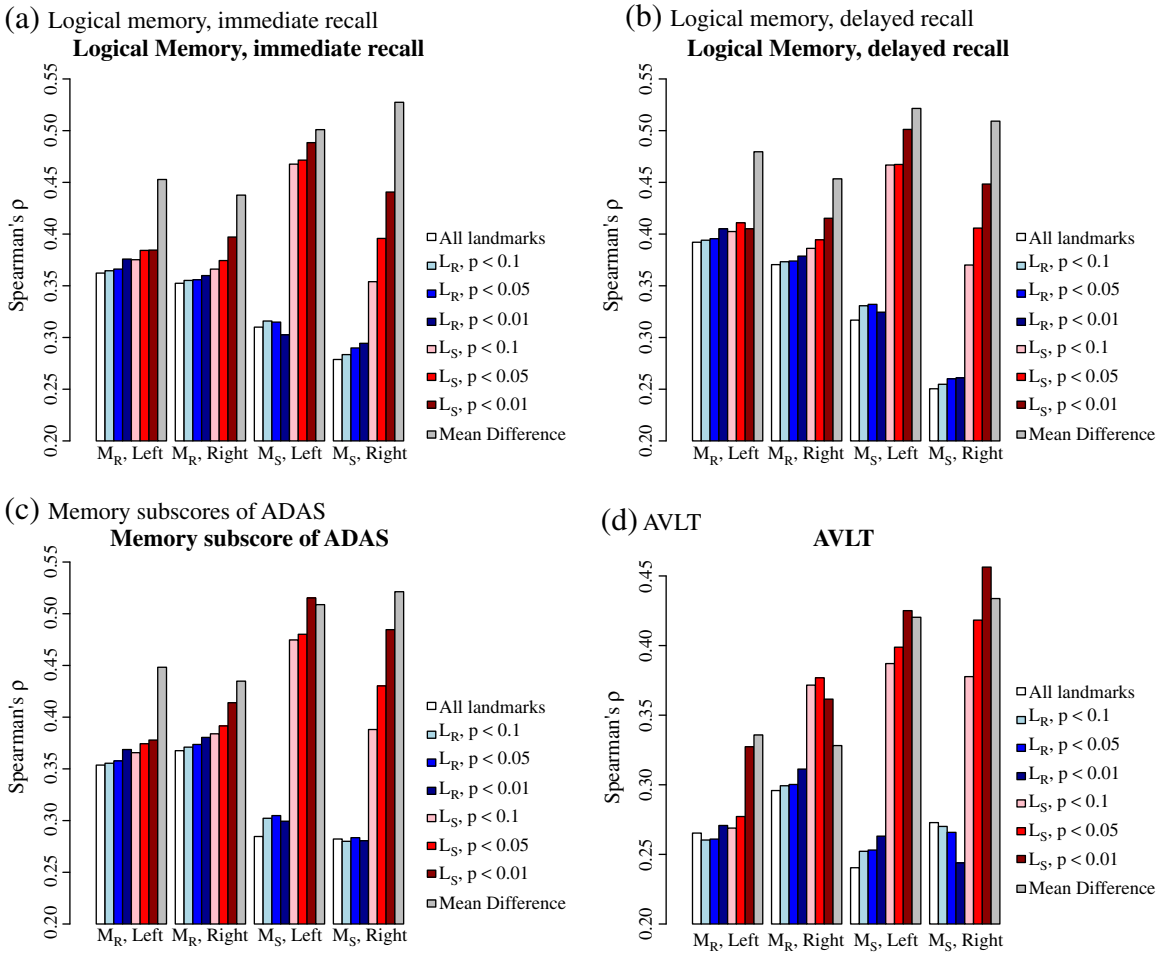

Fig. 6. Bar plot of Spearman's ρ of the correlation between the memory subscores and the principal components from each model. The correlation analysis is performed on the modes extracted from all landmarks, landmarks selected with thresholds at $\alpha = 0.1$, $\alpha = 0.05$, and $\alpha = 0.01$. The component with strongest correlation with the memory score is plotted. The correlation of the shape component along the mean difference direction is also plotted.

Correlation analysis

Since the accuracy of classification could possibly be the results of overfitting of the given data, the correlation analysis connects the link between shape components explaining most shape variations on the selected subregions and the memory decline in the disease, confirming that the subregional shape components displaying higher discrimination ability are related to the development of the disease. The correlation between subregional shape components and the memory scores are comparable to the correlation between the mean difference and memory scores. The variation modes extracted on the hippocampal subregions (Fig. 7) are also consistent with the pattern along the mean difference direction in Fig. 8.

In the correlation analysis, the $\mathbf{M}_S$ models were able to identify the shape components with the maximum correlation higher than other shape or size-and-shape models, despite of the superiority of $\mathbf{M}_R$ to $\mathbf{M}_S$ in the disease classification. It is because more principal components from $\mathbf{M}_R$ models are found to correlate with the memory decline, while to a less degree comparing to the fewer but higher correlations of $\mathbf{M}_S$ components. Therefore using the features from $\mathbf{M}_R$ though less correlated with the memory decline may contribute to the higher classification accuracy. The mean difference direction in the $\mathbf{M}_S$ model is also more correlated with the memory scores than that of the rigidly aligned model.

The correlation analysis was not designed to specify the involvement of one particular subfield in episodic memory processes, for instance in encoding and retrieval processes that are thought to preferentially engage different hippocampal subfields (Eldridge et al., 2005; Henson, 2005). Our findings however confirm the preferential implication of both CA1 and subiculum changes in episodic memory deficits in AD. The regions with variation of most significant correlation with the episodic memory indices also mainly matches to the CA1 and subiculum subfields.

Conclusion

The shape of the hippocampus can provide valuable information for the diagnosis of AD. The principal components of the hippocampus among the population as modeled by the SSM can be used to classify AD against NC. The conventional PCA in SSM is performed on all the landmark points on the surface, which represents the original shape data in a lower dimensional subspace, but may be not discriminative between two groups. The whole-surface description of variation incorporating both hippocampal shape among the population may be not sensitive to the changes induced by the disease. By applying a statistical test on the landmark points in the SSM, we can identify the regions on the hippocampal surface which display more significant effect of the disease on the morphology and thus are more discriminative between AD and NC groups. The PCA performed on this subset produced variation modes which were used as features for the classification between these two groups. The principal variation modes on these regions were more sensitive to the shape change from NC to AD, and better correlated with the measurements of memory decline.

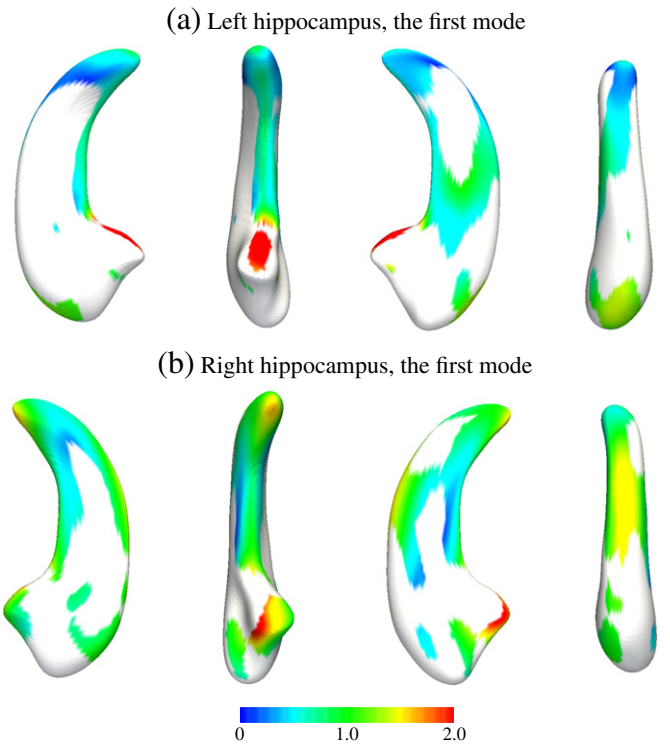

Fig. 7. Variation modes captured by Principal Component Analysis (PCA) on hippocampal subregions best correlated to memory indices (logical memory, AVLT, ADAS memory), color-coded map showing the magnitude of variation described by the mode. $\mathbf{L}_S$, $\alpha = 0.01$, $\mathbf{M}_S$, mode 1. (From left to right: superior, medial, inferior, lateral).

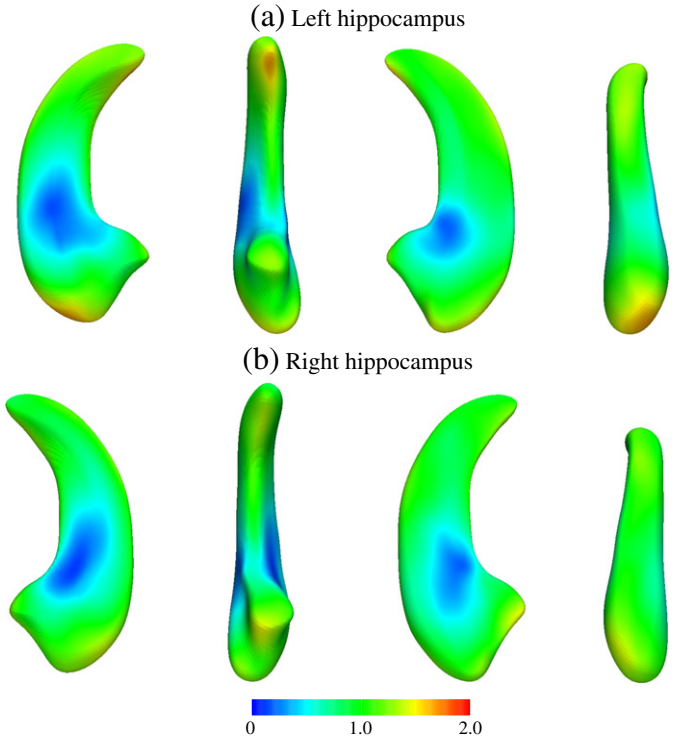

Fig. 8. The variation along the direction of mean difference $\mu_{NC} - \mu_{AD}$ between the normal control (NC) group and the Alzheimer's disease (AD) group, color-coded map showing the magnitude of variation on the rigidly aligned model ($\mathbf{M}_R$). (From left to right: superior, medial, inferior, lateral).

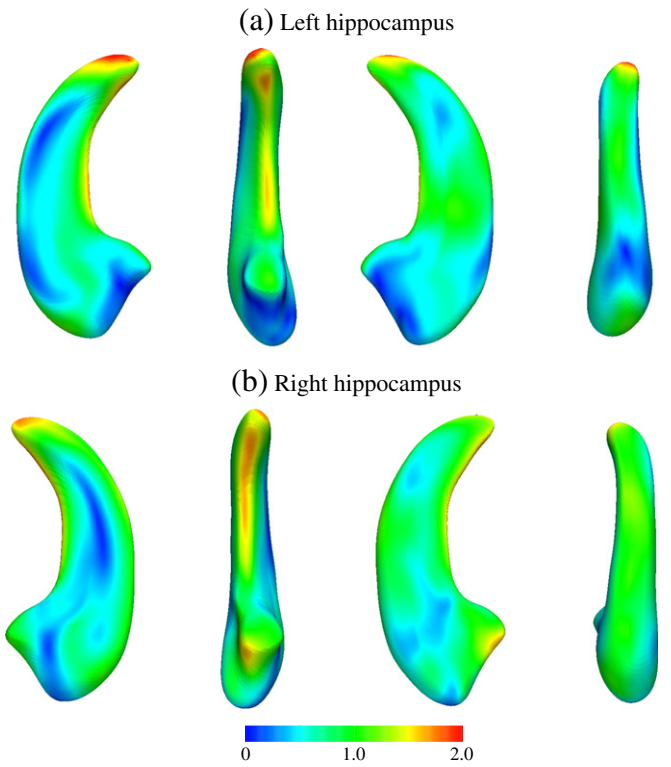

Fig. 9. The variation along the direction of mean difference $\mu_{NC} - \mu_{AD}$ between the normal control (NC) group and the Alzheimer's disease (AD) group, color-coded map showing the magnitude of variation on the similarity aligned model ($\mathbf{M}_S$). (From left to right: superior, medial, inferior, lateral).

Supplementary materials related to this article can be found online at doi: 10.1016/j.neuroimage.2011.10.014.

Acknowledgments

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI 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 California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation.

References

Acosta, O., Bourgeat, P., Zuluaga, M.A., Fripp, J., Salvado, O., Ourselin, S., 2009. Automated voxel-based 3D cortical thickness measurement in a combined Lagrangian–Eulerian PDE approach using partial volume maps. Med. Image Anal. 13 (5), 730–743 Oct.

Alcantara, D., Carmichael, O., Delson, E., Harcourt-Smith, W., Sterner, K., Frost, S., Dutton, R., Thompson, P., Aizenstein, H., Lopez, O., Becker, J., Amenta, N., 2007. Localized components analysis. Proceedings of the 20th international conference on Information processing in medical imaging. Springer-Verlag, Kerkrade, The Netherlands, pp. 519–531.

Apostolova, L.G., Dinov, I.D., Dutton, R.A., Hayashi, K.M., Toga, A.W., Cummings, J.L., Thompson, P.M., 2006. 3D comparison of hippocampal atrophy in amnestic mild cognitive impairment and Alzheimer's disease. Brain 129 (11), 2867–2873 Nov.

Apostolova, L.G., Thompson, P.M., Green, A.E., Hwang, K.S., Zoumalan, C., Jack, C.R., Harvey, D.J., Petersen, R.C., Thal, L.J., Aisen, P.S., Toga, A.W., Cummings, J.L., DeCarli, C.S., 2010. 3D comparison of low, intermediate, and advanced hippocampal atrophy in MCI. Hum. Brain Mapp. 31 (5), 786–797.

Barnes, J., Godbolt, A.K., Frost, C., Boyes, R.G., Jones, B.F., Scahill, R.I., Rossor, M.N., Fox, N. C., 2007. Atrophy rates of the cingulate gyrus and hippocampus in AD and FTLD. Neurobiol. Aging 28 (1), 20–28.

Barnes, J., Bartlett, J.W., van de Pol, L.A., Loy, C.T., Scahill, R.I., Frost, C., Thompson, P., Fox, N.C., 2009. A meta-analysis of hippocampal atrophy rates in Alzheimer's disease. Neurobiol. Aging 30 (11), 1711–1723.

Bobinski, M., de Leon, M.J., Tarnawski, M., Wegiel, J., Bobinski, M., Reisberg, B., Miller, D.C., Wisniewski, H.M., 1998. Neuronal and volume loss in CA1 of the hippocampal formation uniquely predicts duration and severity of Alzheimer disease. Brain Res. 805 (1–2), 267–269 Sep.

Bouix, S., Pruessner, J.C., Collins, D.L., Siddiqi, K., 2005. Hippocampal shape analysis using medial surfaces. NeuroImage 25 (4), 1077–1089 May.

Brechbühler, C., Gerig, G., Kübler, O., 1995. Parametrization of closed surfaces for 3-D shape description. Comput. Vision Image Understanding 61 (2), 154–170.

Breiman, L., 1996a. Bagging predictors. Mach. Learn. 24 (2), 123–140.

Breiman, L., 1996b. Out-of-bag estimation. Tech. rep., Statistics Department, University of California.

Chetelat, G., Baron, J.-C., 2003. Early diagnosis of Alzheimer's disease: contribution of structural neuroimaging. NeuroImage 18 (2), 525–541.

Chételat, G., Fouquet, M., Kalpouzos, G., Denghien, I., la Sayette, V.D., Viader, F., Mézenge, F., Landeau, B., Baron, J., Eustache, F., Desgranges, B., 2008. Three-dimensional surface mapping of hippocampal atrophy progression from MCI to AD and over normal aging as assessed using voxel-based morphometry. Neuropsychologia 46 (6), 1721–1731 May.

Christensen, G., Joshi, S., Miller, M., 1997. Volumetric transformation of brain anatomy. IEEE Trans. Med. Imaging 16 (6), 864–877.

Chupin, M., Gérardin, E., et al., 2009. Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. Hippocampus 19 (6), 579–587.

Cootes, T., Taylor, C., Cooper, D., Graham, J., 1992. Training models of shape from sets of examples. Proceedings of British Machine Vision Conference, pp. 9–18.

Csernansky, J.G., Wang, L., Joshi, S., Miller, J.P., Gado, M., Kido, D., McKeel, D., Morris, J.C., Miller, M.I., 2000. Early DAT is distinguished from aging by high-dimensional mapping of the hippocampus. Neurology 55 (11), 1636–1643 Dec.

Csernansky, J.G., Wang, L., Swank, J., Miller, J.P., Gado, M., McKeel, D., Miller, M.I., Morris, J.C., 2005. Preclinical detection of Alzheimer's disease: hippocampal shape and volume predict dementia onset in the elderly. NeuroImage 25 (3), 783–792 Apr.

Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M.-O., Chupin, M., Benali, H., Colliot, O., 2011. The Alzheimer's Disease Neuroimaging Initiative, Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. NeuroImage 56 (2), 766–781. ISSN 1053-8119, doi:10.1016/j.neuroimage.2010.06.013.

Davies, R. H., 2002. Learning shape: optimal models for analysing natural variability. Ph.D. thesis, University of Manchester.

Davies, R.H., Twining, C.J., Cootes, T.F., Waterton, J.C., Taylor, C.J., 2002. A minimum description length approach to statistical shape modeling. IEEE Trans. Med. Imaging 21 (5), 525–537.

Davies, R., Twining, C., Allen, P., Cootes, T., Taylor, C., 2003. Shape discrimination in the hippocampus using an MDL model. Information Processing in Medical Imaging, pp. 38–50.

Davies, R.H., Twining, C.J., Taylor, C., 2008. Groupwise surface correspondence by optimization: representation and regularization. Med. Image Anal. 12 (6), 787–796 Dec.

Dryden, I.L., Mardia, K.V., 1998. Statistical Shape Analysis, 1st Edition. Wiley. Sep.

Dubois, B., Feldman, H.H., Jacova, C., Dekosky, S.T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P.J., Scheltens, P., 2007. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol. 6 (8), 734–746 Aug.

Eldridge, L.L., Engel, S.A., Zeineh, M.M., Bookheimer, S.Y., Knowlton, B.J., 2005. A dissociation of encoding and retrieval processes in the human hippocampus. J. Neurosci. 25 (13), 3280–3286 Mar.

Frisoni, G.B., Ganzola, R., Canu, E., Rub, U., Pizzini, F.B., Alessandrini, F., Zoccatelli, G., Beltramello, A., Caltagirone, C., Thompson, P.M., 2008. Mapping local hippocampal changes in Alzheimer's disease and normal ageing with MRI at 3 Tesla. Brain 3266–3276 Nov.

Frisoni, G.B., Fox, N.C., J. Jr., C.R., Scheltens, P., Thompson, P.M., 2010. The clinical use of structural MRI in Alzheimer disease. Nat. Rev. Neurol. 6 (2), 67–77.

Gerardin, E., Chételat, G., et al., 2009. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. NeuroImage 47 (4), 1476–1486 Oct.

Gerig, G., Styner, M., Jones, D., Weinberger, D., Lieberman, J., 2001. Shape analysis of brain ventricles using SPHARM. Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA'01). IEEE Computer Society, p. 171.

Gosche, K.M., Mortimer, J.A., Smith, C.D., Markesbery, W.R., Snowdon, D.A., 2002. Hippocampal volume as an index of Alzheimer neuropathology: findings from the nun study. Neurology 58 (10), 1476–1482 May.

Gower, J., 1975. Generalized procrustes analysis. Psychometrika 40, 33–51.

Gutman, B., Wang, Y., et al., 2009. Disease classification with hippocampal shape invariants. Hippocampus 19 (6), 572–578.

Heimann, T., Meinzer, H., 2009. Statistical shape models for 3D medical image segmentation: a review. Med. Image Anal. 13 (4), 543–563 Aug.

Henson, R., 2005. A mini-review of fMRI studies of human medial temporal lobe activity associated with recognition memory. Q. J. Exp. Psychol. B 58 (3–4), 340–360 Oct.

Heun, R., Mazanek, M., Atzor, K.R., Tintera, J., Gawehn, J., Burkart, M., Gnsicke, M., Falkai, P., Stoeter, P., 1997. Amygdala–hippocampal atrophy and memory performance in dementia of Alzheimer type. Dement. Geriatr. Cogn. Disord. 8 (6), 329–336 Dec.

Hsu, Y., Schuff, N., Du, A., Mark, K., Zhu, X., Hardin, D., Weiner, M.W., 2002. Comparison of automated and manual MRI volumetry of hippocampus in normal aging and dementia. J. Magn. Reson. Imaging 16 (3), 305–310.

Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, E.G., Petersen, R.C., Kokmen, E., Kurland, L.T., 1992. Mayo's older Americans normative studies: updated AVLT—norms for ages 56 to 97. Clin. Neuropsychol. 6 (1, Supplement 1), 83–104.

Jack, C.R., Petersen, R.C., Xu, Y., O'Brien, P.C., Smith, G.E., Ivnik, R.J., Tangalos, E.G., Kokmen, E., 1998. The rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease. Neurology 51 (4), 993–999 Oct.

Kendall, D.G., 1984. Shape manifolds, procrustean metrics, and complex projective spaces. Bull. Lond. Math. Soc. 16 (2), 81–121 Mar.

Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., Fox, N.C., Jack, C.R., Ashburner, J., Frackowiak, R.S.J., 2008. Automatic classification of MR scans in Alzheimer's disease. Brain 131 (3), 681–689 Mar.

La Joie, R., Fouquet, M., Mézenge, F., Landeau, B., Villain, N., Mevel, K., Pélerin, A., Eustache, F., Desgranges, B., Chételat, G., 2010. Differential effect of age on hippocampal subfields assessed using a new high-resolution 3T MR sequence. NeuroImage 53 (2), 506–514 Nov.

Li, S., Shi, F., Pu, F., Li, X., Jiang, T., Xie, S., Wang, Y., 2007. Hippocampal shape analysis of Alzheimer disease based on machine learning methods. AJNR Am. J. Neuroradiol. 28 (7), 1339–1345 Aug.

Loughrey, J., Cunningham, P., 2005. Overfitting in Wrapper-Based feature subset selection: The harder you try the worse it gets. Research and Development in Intelligent Systems XXI, pp. 33–43.

Malykhin, N.V., Bouchard, T.P., Camicioli, R., Coupland, N.J., 2008. Aging hippocampus and amygdala. NeuroReport 19 (5), 543–547.

Mueller, S.G., Weiner, M.W., 2009. Selective effect of age, Apo e4, and Alzheimer's disease on hippocampal subfields. Hippocampus 19 (6), 558–564.

Mueller, S.G., Schuff, N., Yaffe, K., Madison, C., Miller, B., Weiner, M.W., 2010. Hippocampal atrophy patterns in mild cognitive impairment and Alzheimer's disease. Hum. Brain Mapp. 31 (9), 1339–1347.

Munson, M.A., Caruana, R., 2009. On feature selection, Bias-Variance, and bagging. Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases: Part II. Springer-Verlag, Bled, Slovenia, pp. 144–159.

Ourselin, S., Roche, A., Subsol, G., Pennec, X., Ayache, N., 2001. Reconstructing a 3D structure from serial histological sections. Image Vision Comput. 19 (1–2), 25–31.

Petersen, R., Jack, C., Xu, Y., Waring, S., Oâ€™Brien, P., Smith, G., Ivnik, R., Tangalos, E., Boeve, B.F., Kokmen, E., 2000. Memory and MRI-based hippocampal volumes in aging and AD. Neurology 54 (3), 581 Feb.

Pizer, S., Fritsch, D., Yushkevich, P., Johnson, V., Chaney, E., 1999. Segmentation, registration, and measurement of shape variation via image object shape. IEEE Trans. Med. Imaging 18 (10), 851–865.

Powell, M.J.D., 1964. An efficient method for finding the minimum of a function of several variables without calculating derivatives. Comput. J. 7 (2), 155–162 Jan.

Qiu, A., Fennema-Notestine, C., Dale, A.M., Miller, M.I., 2009. Regional shape abnormalities in mild cognitive impairment and Alzheimer's disease. NeuroImage 45 (3), 656–661 Apr.

Reunanen, J., 2003. Overfitting in making comparisons between variable selection methods. J. Mach. Learn. Res. 3, 1371–1382.

Rössler, M., Zarski, R., Bohl, J., Ohm, T.G., 2002. Stage-dependent and sector-specific neuronal loss in hippocampus during Alzheimer's disease. Acta Neuropathol. 103 (4), 363–369 Apr..

Ruto, A., Buxton, B., 2005. Application of a robust and efficient ICP algorithm for fitting a deformable 3D human torso model to noisy data. Digital Image Computing: Techniques and Applications, p. 62.

Scher, A.I., Xu, Y., Korf, E.S.C., White, L.R., Scheltens, P., Toga, A.W., Thompson, P.M., Hartley, S.W., Witter, M.P., Valentino, D.J., Launer, L.J., 2007. Hippocampal shape analysis in Alzheimer's disease: a population-based study. NeuroImage 36 (1), 8–18 May.

Schölkopf, B., Smola, A.J., 2002. Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT Press.

Shi, F., Liu, B., Zhou, Y., Yu, C., Jiang, T., 2009. Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: meta-analyses of MRI studies. Hippocampus 19 (11), 1055–1064.

Small, G.W., Bookheimer, S.Y., Thompson, P.M., Cole, G.M., Huang, S.-C., Kepe, V., Barrio, J.R., 2008. Current and future uses of neuroimaging for cognitively impaired patients. Lancet Neurol. 7 (2), 161–172.

Styner, M., Gerig, G., 2001. Medial models incorporating object variability for 3D shape analysis. Proceedings of the 17th International Conference on Information Processing in Medical Imaging. Springer-Verlag, pp. 502–516.

Styner, M., Lieberman, J.A., Pantazis, D., Gerig, G., 2004. Boundary and medial shape analysis of the hippocampus in schizophrenia. Med. Image Anal. 8 (3), 197–203 Sep.

Styner, M., Oguz, I., Xu, S., Pantazis, D., Gerig, G., 2007. Statistical group differences in anatomical shape analysis using Hotelling T^2 metric. In: Pluim, J.P.W., Reinhardt, J.M. (Eds.), Medical Imaging 2007: Image Processing, Vol. 6512. SPIE, San Diego, CA, USA, pp. 65123Z–Z. Mar.

Taubin, G., Zhang, T., Golub, G., 1996. Optimal surface smoothing as filter design. ECCV '96, pp. 283–292.

Thompson, P.M., Hayashi, K.M., Zubicaray, G.I.D., Janke, A.L., Rose, S.E., Semple, J., Hong, M.S., Herman, D.H., Gravano, D., Doddrell, D.M., Toga, A.W., 2004. Mapping hippocampal and ventricular change in Alzheimer disease. NeuroImage 22 (4), 1754–1766 Aug.

Vapnik, V., 1995. The Nature of Statistical Learning Theory. Springer-Verlag.

Vemuri, P., Gunter, J.L., Senjem, M.L., Whitwell, J.L., Kantarci, K., Knopman, D.S., Boeve, B.F., Petersen, R.C., J. Jr., C.R., 2008. Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. NeuroImage 39 (3), 1186–1197.

Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2009. Diffeomorphic demons: efficient non-parametric image registration. NeuroImage 45 (1, Supplement 1), S61–S72 Mar.

Vermaak, J., Perez, P., 2003. Constrained subspace modeling. Computer Vision and Pattern Recognition, 2003: Proceedings. 2003 IEEE Computer Society Conference on, Vol. 2, pp. 106–113.

Wang, L., Swank, J.S., Glick, I.E., Gado, M.H., Miller, M.I., Morris, J.C., Csernansky, J.G., 2003. Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. NeuroImage 20 (2), 667–682 Oct.

Wang, L., Miller, J.P., Gado, M.H., McKeel, D.W., Rothermich, M., Miller, M.I., Morris, J.C., Csernansky, J.G., 2006. Abnormalities of hippocampal surface structure in very mild dementia of the Alzheimer type. NeuroImage 30 (1), 52–60 Mar..

West, M.J., Coleman, P.D., Flood, D.G., Troncoso, J.C., 1994. Differences in the pattern of hippocampal neuronal loss in normal ageing and Alzheimer's disease. Lancet 344 (8925), 769–772 Sep.

Xie, J., Alcantara, D., Amenta, N., Fletcher, E., Martinez, O., Persianinova, M., DeCarli, C., Carmichael, O., 2009. Spatially localized hippocampal shape analysis in late-life cognitive decline. Hippocampus 19 (6), 526–532.

Yushkevich, P., Joshi, S., Pizer, S.M., Csernansky, J.G., Wang, L.E., 2003. Feature selection for shape-based classification of biological objects. Information Processing in Medical Imaging, LNCS 2732, 114–125. doi:10.1007/978-3-540-45087-0_10 (http://www.springerlink.com/content/vf6p19yqlqjtdjpk/).

Zhou, L., Lieby, P., Barnes, N., Rglade-Meslin, C., Walker, J., Cherbuin, N., Hartley, R., 2009. Hippocampal shape analysis for Alzheimer's disease using an efficient hypothesis test and regularized discriminative deformation. Hippocampus 19 (6), 533–540 Jun.

Zhou, L., Wang, L., Shen, C., Barnes, N., 2010. Hippocampal shape classification using redundancy constrained feature selection. In: Jiang, T., Navab, N., Pluim, J.P., Viergever, M.A. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010. : Vol. 6362 of Lecture Notes in Computer Science. Springer, pp. 266–273.