Neurocomputing 175 (2016) 177–187

Contents lists available at ScienceDirect

Neurocomputing

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

Quasi-conformal statistical shape analysis of hippocampal surfaces for Alzheimer's disease analysis

Hei Long Chan [a], Hangfan Li [a], Lok Ming Lui [b,*]

[a] The Chinese University of Hong Kong, 222B, Lady Shaw Building, Shatin, Hong Kong
[b] The Chinese University of Hong Kong, 207, Lady Shaw Building, Shatin, Hong Kong

ARTICLE INFO

Article history:
Received 26 May 2015
Received in revised form
5 October 2015
Accepted 21 October 2015
Communicated by Pingkun Yan
Available online 31 October 2015

Keywords:
Alzheimer's disease
Beltrami coefficient
Hippocampus
p-Value
Shape classification
Surface registration

ABSTRACT

Alzheimer's disease (AD) is a no-cure disease that has been frustrating the scientists for many years. Analyzing the disease has become an important but challenging research topic. The shape analysis of the sub-cortical structure of AD patients has been commonly used to understand this disease. In this paper, we assess the feasibility of using shape information on the hippocampal (HP) surfaces to detect some sub-structural changes in AD patients. We propose a quasi-conformal statistical shape analysis model, which allows us to study local regional geometric changes in the HPs amongst normal control (NC) and AD groups. A shape index defined by the quasi-conformality and surface curvatures is used to characterize region-specific shape variations of the HP surfaces. Feature vectors can be extracted for each HPs, with which a classification model can be built using machine learning methods to classify HPs into NC and AD subjects. Experiments have been carried out on 99 normal controls and 41 patients with AD. Results demonstrate that the proposed quasi-conformal based model is effective for classifying HPs into NC and AD groups with high classification accuracy (with highest overall classification accuracy reaching 87.86% in a leave-one-out experiment using the whole dataset).

1. Introduction

The Alzheimer's disease (AD) is a chronic neurodegenerative disease characterized by a decline in cognitive functions. The cause of AD is poorly understood. It usually starts slowly, gets worsen over time and eventually leads to death. Early detection of AD is thus an important but a challenging task.

Amongst the various subcortical structures, the hippocampus (HP) has demonstrated pronounced shape changes in the early stages of AD. For example, the hippocampal atrophy has been recognized to be more aggressive in AD in comparison with the normal aging [1–6]. The HP surface is therefore amongst the most important biomarker for the early diagnosis of the disease.

HP shape analysis has usually been carried out by studying its global and local shape changes. For global shape analysis, the overall HP volumes are usually evaluated to study global shape differences amongst AD patients. It is believed that HP volumetric decline is correlated to the memory lost [7]. Tissue losses in the HP have also been found in the AD [8]. As a matter of fact, HP volumetry on MR images has been widely used and found helpful for the diagnosis of AD.

Although the HP volume can provide significant information to discriminate AD from normal control subjects, significant regional shape changes in the HP have been observed in the neurodegenerative process of AD [9,10]. For example, neuron loss has commonly been found in CA1 and subiculum subfields [11–13]. In view of this, the examination of the local regional shape changes in the HPs is expected to provide better information to analyze the disease and classify HPs between AD and NC groups. Besides, another potential limitation of the global shape analysis approach is that geometric differences between AD and NC groups may only occur at some specific local regions. Taking into account the overall shape change of the whole HP volume may average out or diminish the discriminative power of the geometric differences amongst the normal and diseased groups, which hinder the shape analysis accuracy. It is therefore desirable to design local shape analysis model that can measure regional shape changes effectively.

In order to analyze the localized pattern of HP shape changes, surface-based morphometry can be employed. A shape index that quantifies regional shape changes is often defined, with which statistical shape analysis can be performed for the HP classification. In this work, we propose a quasi-conformal based shape analysis approach, which allows us to study local regional shape

* Corresponding author.
E-mail addresses: hlchan@math.cuhk.edu.hk (H.L. Chan), hfli@math.cuhk.edu.hk (H. Li), lmlui@math.cuhk.edu.hk (L.M. Lui).

http://dx.doi.org/10.1016/j.neucom.2015.10.047
0925-2312/

differences of the HP amongst NC and AD groups. A shape index based on the quasi-conformality and surface curvatures is applied to characterize region-specific shape variations of the HP surfaces amongst different subjects. The shape index is a positive real-valued function defined on every vertices of the HP. Feature vectors for each HP can be extracted from their shape indices using statistical methodologies. A classification model can then be built using the extracted feature vectors. The proposed quasi-conformal based model is found to be an effective approach to classify HPs into NC and AD groups.

Our experiments are performed on 99 normal controls and 41 patients with AD. The data is obtained using 1.5 T magnetic resonance imaging (MRI) scanner. Using our method, we can accurately achieve a high accuracy of classification between NC and AD groups.

The rest of the paper is organized as follows. In Section 2, we review some previous literatures closely related to this work. Our proposed quasi-conformal shape analysis model will be described in detail in Section 3. Experimental results will be shown in Sections 4 and 5. The paper is concluded in Section 6, in which possible future works are discussed.

2. Related works

The shape analysis of the HP for the disease analysis of AD has been widely studied by various research groups. Different approaches have been developed. For global shape analysis, HP volume has been used for classifying AD subjects and AD diagnosis [5,6,14–17]. In particular, by studying the HP volumes, [18] has reported the classification result between AD and cognitively normal subjects with a success rate of about 72–74% over an Alzheimer's Disease Neuroimaging Initiative (ADNI) database. To further improve the accuracy for the analysis of AD, local shape changes in the HP have been taken into account. Surface-based morphometry of the HP surfaces has been extensively studied. The spherical harmonic (SPHARM) representation of the HP surface has been exploited to extract shape features to quantify shape changes caused by AD [18–22]. Longitudinal approaches which study the HP atrophy rates over times have been proposed for AD classification problems [23,24]. These approaches can often achieve higher classification accuracy than the volumetric approaches (e.g. 82% on 568 images of the ADNI dataset by Wolz et al. [24]). Younes et al. [25] applied the large deformation diffeomorphic metric mapping method for HP surface registration and successfully detected the changing point that indicated the AD. Wang et al. [26] proposed the tensor-based surface morphometry on the HP to analyze shape changes in HPs of AD subjects. To better examine the regional shape changes of the HP, algorithms which segment subfields of the HP have been proposed to detect the local atrophy pattern [27,28]. Lui et al. [29] also proposed to obtain HP registration using Beltrami holomorphic flow. Using the registration, vertex-wise shape changes can be detected and statistically significance map (p-map) can be computed to visualize the regions with significant shape differences.

Statistical shape analysis methods to analyze AD have also drawn much attention recently. For instance, Miklossy et al. [30] used Koch's and Hill's criteria in finding the AD. The analysis of the reviewed data following Koch's and Hill's postulates shows a probable causal relationship between neurospirochetosis and AD. Comelli et al. [31] combined the univariate tests and logistic regression in proposing a therapy for AD. Thompson et al. [32] used the statistical region-of-interest method in assessing the twelve-month metabolic declines in probable AD and Amnestic Mild Cognitive Impairment. It is noteworthy that both the statistical analysis and surface mapping take an important role in human brain analysis especially for the study of AD. Recently, multidimensional classification methods have been widely used for disease classification [19,33–38].

In order to perform local shape analysis, surface registration that captures the one-to-one vertex-wise correspondence between different HP surfaces is crucial. Harmonic surface registration has been widely used [39,40], which produces smooth surface mapping by simply solving an elliptic PDE. Landmark-matching optimized harmonic map has also been proposed to obtain an optimal harmonic map that matches the corresponding landmark features [41–44].

Quasi-conformal theory will be applied in this work. Computational quasi-conformal mapping has been studied recently and applied successfully in the medical imaging field. Lui et al. [29] proposed to obtain quasi-conformal surface registration using the Beltrami holomorphic flow method. The method has been applied to compute HP surface registration [45]. Quasi-conformality has also been utilized to quantify non-isotropic deformations, which can be used to detect abnormal growth or deformation [46]. To deal with higher genus surfaces, different methods have been developed to compute quasi-conformal mappings of surfaces with general topologies [47,48]. Landmark-based surface quasi-conformal registration has also been investigated [47,49].

3. The proposed model

In this section, we will describe our proposed quasi-conformal statistical shape analysis model in detail. Suppose we are given a collection of hippocampal (HP) surfaces of normal controls (NCs) and diseased subjects suffering from Alzheimer's disease (AD). Our goal is to learn a classification model using their shape information, with which a new input HP surface can be classified into either normal or diseased subject. This can potentially assist physicians for the diagnosis of the AD. For this purpose, we propose to combine quasi-conformal theories and statistical tools to develop a shape classification machine. Shape deformation measurement is firstly obtained through quasi-conformal theories, which provide accurate measurement of local geometric differences amongst subjects. A shape classification model can then be learnt through statistical tools and machine learning procedures.

Denote the collection of HP surfaces of m normal controls and n diseased subjects by $\tilde{N} = \{S_i^t\}_{i=1}^{m}$ and $\tilde{A} = \{S_i^t\}_{i=m+1}^{m+n}$, respectively, where $t = 0$ or 1. The HP surface of each subject was captured at the base-time $t = 0$ and after one year $t = 1$. Their surfaces are denoted by S_i^0 and S_i^1, respectively. Our proposed quasi-conformal statistical shape analysis can be summarized as follows, which consists of five main steps.

1. *Surface registration:* For each subject i, the deformation $f_i : S_i^0 \to S_i^1$ of subject i is obtained. Also, surface registrations $g_{ij} : S_i^0 \to S_j^0$ are computed. These registrations give point-wise correspondence between subjects for further shape analysis.
2. *Shape deformation measurement:* For each subject i, measure the shape deformation at each vertex of the HP from $t = 0$ to $t = 1$ using quasi-conformal theories. A shape index $E_{shape}^i : S_i^0 \to \mathbb{R}^+$ is obtained that measures the degree of deformation at each vertex.
3. *Extraction of statistically significant regions:* Statistically significant p-map is obtained based on the shape index computed for each subject. A statistically significant region Ω can be extracted to obtain more accurate classification results.
4. *Extraction of feature vector:* The shape index ${E_{shape}}^i$ together with the statistically significant region Ω gives rise to a discriminative feature vector $\tilde{\mathbf{c}}_i$ for each subject. A mean feature $\tilde{\mathbf{c}}_{mean}^{NC}$ amongst the normal control group can also be extracted. Distance d_i between each feature vector $\tilde{\mathbf{c}}_i$ and the mean feature

vector $\tilde{\mathbf{c}}_{mean}^{NC}$ can be computed, which can be used to build the shape classification machine.

5. *Building the classification model:* Using the discriminative feature vector, a classification model is to be built to classify a new input subject into either NC group or AD group.

We will now explain each step in detail.

3.1. Surface registration

Registration between HP surfaces must be computed to obtain point-wise correspondences between surfaces. With the registration, vertex-wise geometric difference between subjects can be measured and local shape analysis can be carried out. In this work, we apply the radial registration, a non-rigid transformation model, about the centerlines of the HP surfaces to obtain the surface registration.

Given two HP surfaces S_i^0 and S_j^0, the centerlines of each surface are computed via level-set based algorithm ([50]). Let $l_i : (0,1) \to \mathbb{R}^3$ and $l_j : (0,1) \to \mathbb{R}^3$ be the centerlines of S_i^0 and S_j^0, respectively. For each $x \in (0,1)$, $l_i'(x)$ and $l_j'(x)$ are the tangent vectors of l_i and l_j at x, respectively. Let $N_i(x)$ and $N_j(x)$ be the collection of all vectors, starting at $l_i(x)$ and $l_j(x)$, respectively, perpendicular to $l_i'(x)$ and to $l_j'(x)$, respectively. Then, for each $n_i(x) \in N_i(x)$ and $n_j(x) \in N_j(x)$, $n_i(x)$ is a ray hitting S_i^0 exactly once at $p_{n_i}(x)$ and $n_j(x)$ similarly hits S_j^0 at $p_{n_j}(x)$. Write

$$P_{n_i}^x = \{p_{n_i}(x) \in S_i^0 : n_i(x) \in N_i(x)\}, \tag{1}$$

$$P_{n_j}^x = \{p_{n_j}(x) \in S_j^0 : n_j(x) \in N_j(x)\}. \tag{2}$$

$P_{n_i}^x$ and $P_{n_j}^x$ extract radial loops on S_i^0 and on S_j^0, respectively, at level x.

$P_{n_i}^x$ and $P_{n_j}^x$ can be viewed as a curve, and can be parameterized using the angular parametrization. That is

$$P_{n_i}^x : \theta \in (0, 2\pi) \to S_i^0, \tag{3}$$

$$P_{n_j}^x : \theta \in (0, 2\pi) \to S_j^0, \tag{4}$$

are curves and we register $P_{n_i}^x(\theta)$ to $P_{n_j}^x(\theta)$ for any $\theta \in (0, 2\pi)$ and for any $x \in (0,1)$. Combining the registrations of every pairs of radial loops, the surface registration $g_{ij} : S_i^0 \to S_j^0$ between S_i^0 and S_j^0 can be obtained. Similarly, the surface registration $f_i : S_i^0 \to S_i^1$ between the HP surfaces of a subject measured at different times can be computed. For details of the registration algorithm, we refer the readers to [19].

3.2. Shape deformation measurement

The surface registration $f_i : S_i^0 \to S_i^1$ for each subject allows us to analyze the shape deformation at each vertex of the HP surface. It is believed that significant shape deformations (atrophy) occur on HP surfaces of patients suffering from AD [2,11–15]. Understanding the shape deformation pattern of the HP surface over time can possibly provide information to classify HP surfaces into normal and AD groups. To quantitatively measure the degree of shape deformation, we compute a shape index using the quasi-conformality and curvatures to measure subtle deformation of the HP surface. Given the surface registration $f_i : S_i^0 \to S_i^1$, the shape index $E_{shape}^i : S_i^0 \to \mathbb{R}^+$ is a positive real-valued function on S_i^0. In practice, HP surfaces are represented discretely by triangular meshes. $E_{shape}{}^i$ is defined on each vertex $\mathbf{v}_i^j$ of S_i^0 as follows.

Definition (*Shape index*). Given the surface registration $f_i : S_i^0 \to S_i^1$, the shape index $E_{shape}^i(\mathbf{v}_i^j)$ at the vertex $\mathbf{v}_i^j$ of S_i^0 is defined as:

$$E_{shape}^i(\mathbf{v}_i^j) = \gamma |\mu(f_i)(\mathbf{v}_i^j)| + \alpha |H_0(\mathbf{v}_i^j) - H_1(f_i(\mathbf{v}_i^j))| + \beta |K_0(\mathbf{v}_i^j) - K_1(f_i(\mathbf{v}_i^j))| \tag{5}$$

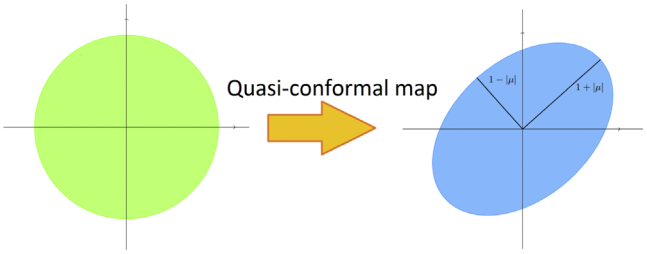

Fig. 1. Infinitesimal behaviour of quasi-conformal map.

where α, β and γ are real positive scalar parameters. $|\mu(f_i)(\mathbf{v}_i^j)|$ is the complex dilation defined by:

$$|\mu(f_i)(\mathbf{v}_i^j)| = \left|\frac{\partial f_i}{\partial \bar{z}}(\mathbf{v}_i^j)\right| \Big/ \left|\frac{\partial f_i}{\partial z}(\mathbf{v}_i^j)\right|. \tag{6}$$

H_1 and H_2 are the mean curvature on S_i^0 and S_i^1, respectively. K_1 and K_2 are the Gaussian curvature on S_i^0 and S_i^1, respectively.

The complex dilation $|\mu(f_i)(\mathbf{v}_i^j)|$ measures the conformality distortion of the deformation at the vertex $\mathbf{v}_i^j$. Intuitively, a general deformation maps an infinitesimal circle at $\mathbf{v}_i^j$ to an infinitesimal ellipse at $f_i(\mathbf{v}_i^j)$. The distortion from the small circle to the small ellipse can be measured by $|\mu(f_i)(\mathbf{v}_i^j)|$. More specifically, the maximal stretching and shrinkage can be measured by $1+|\mu(f_i)(\mathbf{v}_i^j)|$ and $1-|\mu(f_i)(\mathbf{v}_i^j)|$, respectively see Fig. 1. In particular, the deformation is conformal or locally isotropic at $\mathbf{v}_i^j$ if $|\mu(f_i)(\mathbf{v}_i^j)| = 0$. The partial derivatives $\frac{\partial f_i}{\partial \bar{z}}(\mathbf{v}_i^j)$ and $\frac{\partial f_i}{\partial z}(\mathbf{v}_i^j)$ are defined using the coordinate charts of S_i^0 and S_i^1. Let $\phi_j : U_j \subset S_i^0 \to \mathbb{C}$ and $\varphi_j : V_j \subset S_i^1 \to \mathbb{C}$ be coordinate charts of S_i^0 and S_i^1 around $\mathbf{v}_i^j$ and $f_i(\mathbf{v}_i^j)$, respectively. The partial derivatives can be defined as:

$$\frac{\partial f_i}{\partial \bar{z}}(\mathbf{v}_i^j) := \frac{\partial \varphi_j \circ f_i \circ \phi_j^{-1}}{\partial \bar{z}}(\phi_j(\mathbf{v}_i^j)); \quad \frac{\partial f_i}{\partial z}(\mathbf{v}_i^j) := \frac{\partial \varphi_j \circ f_i \circ \phi_j^{-1}}{\partial z}(\phi_j(\mathbf{v}_i^j)), \tag{7}$$

where $\frac{\partial}{\partial \bar{z}} = \frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})$ and $\frac{\partial}{\partial z} = \frac{1}{2}(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y})$.

This shape index has been used to formulate an energy functional over all possible surface mappings to solve the geometric matching surface registration problem [45]. Note that $E_{shape}{}^i$ is a complete shape index measuring different kinds of distortions of the deformation. A different combination of parameters gives rise to different shape indices measuring different kinds of distortions, which can be summarized as below:

- $\alpha = \beta = 0$ and $\gamma \neq 0$: $E_{shape}{}^i$ measures the conformality distortion of the deformation. In other words, $E_{shape}^i \equiv 0$ if the deformation is locally isotropic everywhere.
- $\alpha = 0$, $\beta \neq 0$ and $\gamma \neq 0$: $E_{shape}{}^i$ measures the isometric distortion of the deformation. An isometric deformation preserves the metric (both length and angle). In other words, $E_{shape}^i \equiv 0$ if the deformation deforms the shape while keeping the length and angle structure of the shape.
- $\gamma = 0$, $\alpha \neq 0$ and $\beta \neq 0$: $E_{shape}{}^i$ measures the curvature deviation of the deformation.
- $\alpha \neq 0$, $\beta \neq 0$ and $\gamma \neq 0$: $E_{shape}{}^i$ measures all kinds of distortions of the deformation. In particular, $E_{shape}^i \equiv 0$ under the deformation f_i if and only if S_i^0 and S_i^1 are equal up to a rigid motion.

In this work, we set $\alpha \neq 0$, $\beta \neq 0$ and $\gamma \neq 0$ to measure all kinds of distortions of the deformation. The shape index functions can be computed for every subjects. A feature matrix can then be built to learn a classification machine. Since all HP surfaces are registered, we assume that all HP surface meshes have the same number of vertices and connectivity. Suppose the vertices of S_i^0 are ordered and denoted by $\{v_i^1, v_i^2, \ldots, v_i^N\}$. Without loss of generality, we may assume that v_i^k and v_j^k correspond to each others for every

$1 \le k \le N$. Using the shape index function $E_{shape}{}^{i}$, a feature vector $\mathbf{c}_i$ can be computed for each subject:

$$\mathbf{c}_i = (c_{i1}, c_{i2}, \ldots, c_{ij}, \ldots, c_{iN}) \tag{8}$$

where $c_{ij} = E^{i}_{shape}(\mathbf{v}^{j}_{i})$. Combining all feature vectors, we obtain a feature matrix C, which can be used to study the geometric difference of the HP deformations and build the classification model:

$$C = \begin{pmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_i \\ \vdots \\ \mathbf{c}_{m+n} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1j} & \cdots & c_{1N} \\ \vdots & \vdots & & \vdots & & \vdots \\ c_{i1} & c_{i2} & \cdots & c_{ij} & \cdots & c_{iN} \\ \vdots & \vdots & & \vdots & & \vdots \\ c_{m+n,1} & c_{m+n,2} & \cdots & c_{m+n,j} & \cdots & c_{m+n,N} \end{pmatrix} \tag{9}$$

Without loss of generality, we may assume that the first m rows of C are the feature vectors of the HP surfaces from the normal control group. The last n rows are the feature vectors of the HP surfaces from the AD group. Each column of C captures the degrees of distortions of the HP deformations at a particular (corresponding) vertex of all subjects. Since AD is believed to be related to the shape changes of the HP, the feature matrix provides full information of the shape deformation of the HP and hence can be used to develop a classification model.

3.3. Extraction of statistically significant regions

The feature matrix C gives full information about the shape deformation at every vertices of the HPs. In real situation, AD may have effect only on some particular regions or positions of the HP. Utilizing the full information of the deformation at every vertices to build the statistical model may hinder the classification accuracy. To solve this issue, we propose to extract significant regions and analyze their shape deformations to enhance the classification accuracy.

More specifically, we proceed to look for a set of vertices $\{\mathbf{v}_i^{j_1}, \ldots, \mathbf{v}_i^{j_k}, \ldots, \mathbf{v}_i^{j_l}\}$, whose shape deformations give the most important information for the classification of the HPs. Each column j of C captures the degree of distortions of the HP deformation at a particular vertex. To determine the importance of the shape deformation information at this vertex, we quantitatively measure its statistical significance. We perform a t-test on each column to get a p-value p_j at the vertex j. p_j reports the probability of the geometric difference in the deformation at vertex j occurring from the same distribution by chance. The smaller the p_j is, the higher the probability that the shape deformation at vertex j can distinguish between the normal and AD groups.

In order to stabilize and advance the discriminative power of the selected feature points, we make use of the famous bagging predictors [51]. We compute P_{ij}, the p-value computed at vertex j, using all the $m+n$ subjects from the database except for the i^{th} one. Then we assign

$$\tilde{p}_j = \min_{i = 1, \ldots, m+n} P_{ij} \tag{10}$$

to the vertex j. Hence, we obtain a p-map $\tilde{p} : \tilde{S} \to [0, 1]$, where $\tilde{S}$ is a template mesh having the same number of vertices and connectivity as the HP dataset, such that $\tilde{p}(v_j) = \tilde{p}_j$.

From the p-map $\tilde{p}$, the statistically significant regions can be extracted:

$$\Omega := \bigcup_{j \in I_{sig}} \{\tilde{\mathbf{v}}^{j}\} \subseteq \tilde{S}, \tag{11}$$

where $\tilde{\mathbf{v}}^{j}$ is the jth vertex of the surface mesh $\tilde{S}$ and the index set I_{sig} is defined as:

$$I_{sig} = \{j_1, j_2, \ldots, j_l : \tilde{p}_{j_k} \le p_{cut} \text{ for } 1 \le k \le l\}. \tag{12}$$

Here, p_{cut} is some constant threshold. In other words, we extract all vertices having a p-value less than or equal to p_{cut} as statistically significant regions for our investigation.

3.4. Extraction of discriminative feature vectors

Once the statistically significant region Ω is extracted, the shape deformations at statistically significant vertices can be analyzed and the classification model can be built. The discriminative feature vector $\tilde{\mathbf{c}}_i$ can be computed for each subject i as follows:

$$\tilde{\mathbf{c}}_i = (c_{ij_1}, \ldots, c_{ij_k}, \ldots, c_{ij_l}). \tag{13}$$

Combining discriminative feature vectors of all subjects together gives the discriminative feature matrix $\tilde{C}$ defined as:

$$\tilde{C} = \begin{pmatrix} c_{1j_1} & c_{1j_2} & \cdots & c_{1j_k} & \cdots & c_{1j_l} \\ \vdots & \vdots & & \vdots & & \vdots \\ c_{ij_1} & c_{ij_2} & \cdots & c_{ij_k} & \cdots & c_{ij_l} \\ \vdots & \vdots & & \vdots & & \vdots \\ c_{m+n,j_1} & c_{m+n,j_2} & \cdots & c_{m+n,j_k} & \cdots & c_{m+n,j_l} \end{pmatrix} \tag{14}$$

The discriminative feature matrix captures distortions of HP deformations at vertices which provide the most significant information to classify HP surfaces into normal and AD groups. The mean discriminative feature vectors of the normal group can also be constructed:

$$\tilde{\mathbf{c}}^{NC}_{mean} = \sum_{i=1}^{m} \frac{1}{m} \tilde{\mathbf{c}}_i \tag{15}$$

3.5. Building the classification model

The discriminative feature vectors can be used to build a classification machine. In this work, both for simplicity and efficiency, we use a simple L^2 classification model, which is based on the mean discriminative feature vector $\tilde{\mathbf{c}}^{NC}_{mean}$ of the NC group. For each subject i, we compute the distance of $\tilde{\mathbf{c}}_i$ from $\tilde{\mathbf{c}}^{NC}_{mean}$ to get:

$$d_i = \| \tilde{\mathbf{c}}_i - \tilde{\mathbf{c}}^{NC}_{mean} \|_2. \tag{16}$$

We make the assumption that the deformation patterns of the HPs in the normal group are similar to each others. Thus, d_i is assumed to be small if subject i is in the NC group. We proceed to search for a parameter δ to classify the HP surfaces into either NC or AD groups. More specifically, we conclude that the HP is from the NC group if the distance of its discriminative feature vector from $\tilde{\mathbf{c}}^{NC}_{mean}$ is less than δ. Otherwise, we conclude that the HP comes from the AD group.

Our goal is to search for the optimal classification parameter δ^{opt} that yields the best classification accuracy. We first rearrange the distances d_i's in an ascending order:

$$d_{i_1} \le d_{i_2} \le \cdots \le d_{i_{m+n}}. \tag{17}$$

For each δ, we can obtain the true positive set (TPS(δ)) and true negative set (TNS(δ)) with respect to δ as follows:

$$\begin{aligned} \text{TPS}(\delta) &= \{S_i^0 : S_i^0 \text{ is classified as a normal subject and } 1 \le i \le m\}; \\ \text{TNS}(\delta) &= \{S_i^0 : S_i^0 \text{ is classified as an AD subject and } m \\ &\quad + 1 \le i \le m+n\}. \end{aligned} \tag{18}$$

The true positive rate (TPR(δ)) and the true negative rate (TNR (δ)) with respect to δ can be defined as follows:

$$\mathrm{TPR}(\delta)=|\mathrm{TPS}(\delta)|/m;\quad \mathrm{TNR}(\delta)=|\mathrm{TNS}(\delta)|/n \tag{19}$$

The optimal classification parameter δ^{opt} can then be computed by solving the following optimization problem:

$$\delta^{opt}=\underset{\delta}{\operatorname{argmax}}\left\{\frac{|\mathrm{TPS}(\delta)|+|\mathrm{TNS}(\delta)|}{m+n}\right\} \tag{20}$$

In practice, the number of HP surfaces are finite. Therefore, the optimization problem (20) above is optimized over a finite set. More specifically, we define:

$$P=\left\{\tilde{\delta}:\tilde{\delta}=\frac{d_{i_j}+d_{i_{j+1}}}{2}\text{ for some } j \text{ such that } d_{i_j}<d_{i_{j+1}}\right\}. \tag{21}$$

Then, P is a finite set and the optimization problem (20) can be formulated as:

$$\tilde{\delta}^{opt}=\underset{\tilde{\delta}\in P}{\operatorname{argmax}}\left\{\frac{|\mathrm{TPS}(\tilde{\delta})|+|\mathrm{TNS}(\tilde{\delta})|}{m+n}\right\}. \tag{22}$$

It becomes an optimization problem over a finite set, which can be solved by maximizing amongst all possible choices.

Once the optimal classification parameter is computed, the classification model can be built. Suppose a new input HP surface S is given, S will be registered to the template surface $\tilde{S}$. The discriminative feature vector $\tilde{c}_S$ of S will be constructed. The distance between $\tilde{\mathbf{c}}_S$ and $\tilde{\mathbf{c}}^{NC}_{mean}$ will then be computed: $d_S=\|\tilde{\mathbf{c}}_S-\tilde{\mathbf{c}}^{NC}_{mean}\|_2$. If $d_S\le\tilde{\delta}^{opt}$, we conclude that S belongs to the NC group. Otherwise, we conclude that S belongs to the AD group.

Apart from the simple L^2 classification, we can also use other clustering methods, like the K-mean clustering and the Support Vector Machine (SVM) clustering.

If we turn to use K-mean clustering, then after applying p-value test to select feature points, we split those NC HP into K_{NC} sub-groups and those AD HP into K_{AD} sub-groups, by packing those with similar shape index on the statistically significant region together. A mean discriminative feature vectors of each sub-group is built. Whenever a new subject is imported, we compute its distance from all the $K_{NC}+K_{AD}$ sub-groups similarly as (16). If the smallest distance is recorded to any sub-group of NC HP models, we classify the new subject as NC. Otherwise we classify it as AD.

This K-mean clustering is more advanced than the simple L^2 classification since we have in total $K_{NC}+K_{AD}$ sub-groups instead of just one for each class. However, we should note that firstly, the number K_{NC} and K_{AD} must be selected carefully. Too few sub-groups may have only limited help to boost up the classification accuracy of our algorithm, while too many sub-groups will induce over-determining effect and hence causing inaccuracy. Also, the proportion between the two class of data is important. Unbalanced proportion of data will significantly induce bias to our algorithm and hence limit its performance.

On the other hand, we can also choose to use the famous SVM clustering to replace the simple L^2 classification above. In SVM clustering, we compute a hyperplane which best separates the NC HPs and the AD HPs into zones concerning the shape index E_{shape} on the region Ω. When a new subject is imported, we compute E_{shape} on this subject and see which zone, NC or AD, it belongs to, so as to classify it.

3.6. Overall algorithm

The overall algorithm can be described as follows.

Algorithm 1. Quasi-conformal statistical shape analysis.

Input: Training data: NC HP $\{S_i^t\}_{1\le i\le m,t\in\{0,1\}}$; AD HP $\{S_i^t\}_{m+1\le i\le m+n,t\in\{0,1\}}$. Input surface S to be classified.

Output: Classification value

$$\mathcal{C}(S)=\begin{cases}0 & \text{if } S \text{ is NC}\\ 1 & \text{if } S \text{ is AD}\end{cases}$$

Step 1: Compute the deformations $f_i: S_i^0\to S_i^1$ and the pairwise registrations $g_{ij}: S_i^0\to S_j^0$.

Step 2: For each S_i, compute ${E_{shape}}^i$ that measures the distortion of the shape deformation of ${S_i}^0$.

Step 3: Extract the statistical significant region Ω.

Step 4: Compute the discriminative feature vector $\tilde{\mathbf{c}}_i$ for each ${S_i}^0$ and the mean discriminative feature vector of the normal control group $\tilde{\mathbf{c}}^{NC}_{mean}$.

Step 5: Compute $d_i=\|\tilde{\mathbf{c}}_i-\tilde{\mathbf{c}}^{NC}_{mean}\|_2$ for $1\le i\le n+m$. Compute the optimal classification parameter δ^{opt}.

Step 6: Compute the discriminative feature vector $\tilde{\mathbf{c}}_S$ of S. Compute $d_S=\|\tilde{\mathbf{c}}_S-\tilde{\mathbf{c}}^{NC}_{mean}\|_2$. If $d_S\le\delta^{opt}$, set $\mathcal{C}(S)=0$. If $d_S>\delta^{opt}$, set $\mathcal{C}(S)=1$.

Remark. Step 5 and Step 6 can be replaced by other classification models such as the K-mean clustering or the Support Vector Machine (SVM) clustering. In this work, we use a simple L^2 thresholding classification model, which can be computed efficiently. We have found that even with such a simple model, the classification accuracy is reasonably good. We have also compared this simple classification model with the K-mean clustering and the SVM clustering in Section 5.

4. Experiments

We are given a data set consisting of 99 NC HP models and 41 AD HP models from the ADNI database. Segmentation of the HP surfaces is based on a semi-automatic multi-atlas segmentation method (see [52,53]), and the whole segmentation process is accomplished and validated by neo-scientists.

In this section, our first focus is on analyzing the accuracy of our proposed model via different experiments. We employ the leave-one-out cross validation, with successful classification of an NC (AD, resp.) model being counted as true positive (true negative, resp.). The importance of each component in the shape index will be verified. Parameters, including α, β and γ, the weighting function of shape index, and p_{cut}, the threshold cutting in the feature selection, will be varied to analyze the sensitivity of our algorithm to them. Also, variation of the database will be conducted to check the reliance of our model to the given database.

Our model does not only contribute to the numerical classification of NC/AD HPs but visualization of the shape index is also possible. We plot the template mesh $\tilde{S}$ with colors indicating the p-value of the shape index at each vertex. Therefore, the geometric region being included in Ω in the feature selection process can be better understood for further medical and statistical research.

Among all the experiments, we take the following setup:

$$\alpha=3.05,\quad \beta=0.35,\quad \gamma=1,\quad p_{cut}=0.03 \tag{23}$$

4.1. Elements of the shape index

The shape index is constructed by three terms, namely, complex dilation, mean curvature and Gaussian curvature. By setting their corresponding weighting parameters to be 0, respectively, the importance of each element can be studied. The result is reported in Table 1.

Table 1
Importance of each energy element (0 for excluding the corresponding element).

α	β	γ	*TPR*	*TNR*	Total accuracy
0	0.35	1.00	0.8687	0.7805	0.8429
3.05	0	1.00	0.8990	0.7805	0.8643
3.05	0.35	0	0.8586	0.5610	0.7714
3.05	0.35	1.00	0.9091	0.8049	0.8786

Table 2
Classification results with different α.

α(mean curvature)	β(Gaussian curvature)	*TPR*	*TNR*	Total accuracy
0.00	0.35	0.8687	0.7805	0.8429
0.35		0.8687	0.7805	0.8429
0.80		0.8485	0.7561	0.8214
1.25		0.8586	0.7561	0.8286
1.70		0.8384	0.8049	0.8286
2.15		0.8586	0.7561	0.8286
2.60		0.8990	0.8049	0.8714
3.05		0.9091	0.8049	0.8786
3.50		0.8687	0.7805	0.8429

Table 3
Classification results with different β.

α(mean curvature)	β(Gaussian curvature)	*TPR*	*TNR*	Total accuracy
3.05	0.00	0.8990	0.7805	0.8643
	0.35	0.9091	0.8049	0.8786
	0.80	0.8990	0.8049	0.8714
	1.25	0.8889	0.8049	0.8643
	1.70	0.8889	0.8049	0.8643
	2.15	0.8788	0.8049	0.8571
	2.60	0.8889	0.8049	0.8643
	3.05	0.8889	0.8049	0.8643
	3.50	0.8687	0.7805	0.8429

An important observation is that among all three elements, the complex dilation plays the most important role in the classification model, as without it the accuracy drops very significantly (10%). This can be explained by the ability of the complex dilation term to capture local and non-conformal shape deformation, inferring it with a bounded real number μ so that simple comparison can be conducted directly. On the other hand, the mean curvature and the Gaussian curvature are less essential but still contributes to a higher accuracy.

Our next goal is to investigate the sensitivity of our algorithm to the weighting of the elements of the shape index. From the previous experiment, the complex dilation is undoubtedly the most important term in the shape index. However, the sensitivity of our model to the weighting of the curvature terms is still unclear. Hence, an investigation will be conducted. Under normalization of the shape index, we can always assume $\gamma = 1$ if the complex dilation is included and $\gamma = 0$ if not. We will vary α and β, respectively, so that performance of our model under different parameters setting can be compared. The result is reported in Tables 2 and 3. The graph of TA (total classification accuracy) versus alpha and versus beta is plotted in Fig. 2.

From the results, by a slight variation of α, the performance of our model is fluctuated to give 84–91% TPR, 76–80% TNR, and 82–88% overall accuracy. And by varying β slightly, the performance varies between 88–91% TPR, 78–80% TNR, and 86–88% overall accuracy. Therefore, our model is more sensitive to the variation of α than that of β.

4.2. Threshold cutting in feature selection

In our model, feature selection is applied for data filtering, highlighting only the significant region for further statistical analysis. A vertex on HPs is selected if the p-value of the shape index at that vertex is less than p_{cut}. Small p_{cut} infers a strict criterion that only a few vertices are included in Ω. On the contrary, if p_{cut} is large, a lot of vertices is included in Ω but too many feature points may weaken the significance of those critical vertices. In the following experiment, p_{cut} will be varied and the performance of our model will be compared to seek for a better trade-off for p_{cut}. The result is recorded in Table 4 and Fig. 3.

The result suggests that setting p_{cut} to be 0.03 is the best trade-off, for if $p_{cut} < 0.03$, too few vertices are included such that the performance of our model is hindered, and if $p_{cut} > 0.03$, too many vertices are selected and this downgrades the performance of our model. In particular, if we do not perform feature selection (i.e. select every vertex as feature points), the model reports TPR=95%, TNR=44%, and total accuracy=80%. Therefore, applying a suitable feature selection is essential to maintain the discriminative power of the proposed shape index.

4.3. Variation of database

Sensitivity of our model to variations in the database should also be analyzed, so that the reliance of our algorithm to the given database can be checked. Since we are given just a few AD HP models, we will pick randomly, from those NC HP models, 10–90 models with step size 10 each time, and compare the performance of our algorithm using a different database. The total number of HP models involved varies from 51 to 140. In each setting, parameters of the shape index are trained to optimize the performance and the threshold cutting p_{cut} is set to be 0.03 constantly. The result is recorded in Table 5 and Fig. 4.

When the total number of HPs involved is low, our model shows itself to be very accurate. However, this may be just a consequence of unstable fluctuation when a machine learning algorithm is tested by an extremely small-size database. And when the total number of HPs involved is beyond 81, our model becomes stable and achieves 87–88% overall accuracy. This shows that our model is accurate not only with respect to a specific database.

Another observation is that when the total number of HPs involved increases, TPR generally increases while TNR generally decreases. This can be explained by the variation of the proportion between NC HPs and AD HPs throughout the experiment. When there are only 10 NC HPs involved, the corresponding ratio is about 1:4. But when the whole database is used, the corresponding ratio elevates to about 2:1. Besides, TPR is about the same as TNR when the ratio is about 1:1. Therefore, though the overall accuracy of our model is stable when a large size database is being used, a balanced database is better to remain our model unbiased.

4.4. Visualization of the statistical significant region

To facilitate further medical and statistical analyses, our model is made to allow the visualization of the statistically significant region of the shape index. Under the setting $p_{cut} = 0.03$, a total of 1600 vertices out of 6002 vertices are selected as significant features in the feature selection process. We plot the template mesh $\tilde{S}$ with colors indicating the p-value at the corresponding vertex in Fig. 5.

Fig. 5 (a) and (b) shows a smooth color plot of the p-value at each vertex from the front view and back view, respectively. Fig. 5(c) and (d) shows the highlighting (in red) of vertices with p-values less than 0.03 from the front view and back view, respectively.

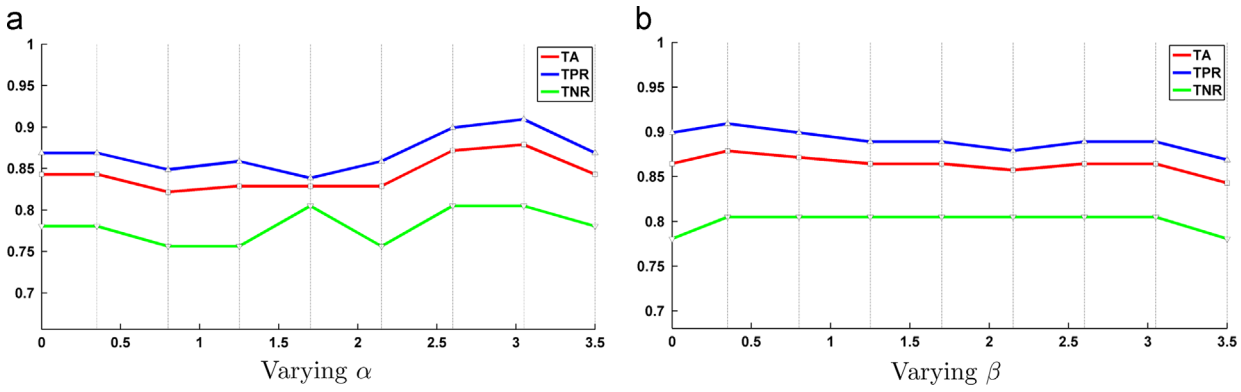

Fig. 2. Plot of TA (total accuracy) versus (a) α and (b) β.

Table 4
Classification results with different p_{cut}.

p_{cut}	*TPR*	*TNR*	Total accuracy	p_{cut}	*TPR*	*TNR*	Total accuracy
0.01	0.8788	0.8049	0.8571	0.11	0.9192	0.5366	0.8071
0.02	0.8586	0.8049	0.8429	0.12	0.9293	0.5122	0.8071
0.03	0.9091	0.8049	0.8786	0.13	0.9293	0.5366	0.8143
0.04	0.8788	0.7805	0.8500	0.14	0.9293	0.5366	0.8143
0.05	0.8788	0.7317	0.8357	0.15	0.9192	0.5366	0.8071
0.06	0.8687	0.6829	0.8143	0.16	0.9192	0.5366	0.8071
0.07	0.8586	0.5610	0.7714	0.17	0.9192	0.5366	0.8071
0.08	0.9192	0.5854	0.8214	0.18	0.9192	0.5366	0.8071
0.09	0.9596	0.5854	0.8500	0.19	0.9394	0.5366	0.8214
0.10	0.9596	0.5854	0.8500	0.20	0.9697	0.5366	0.8429

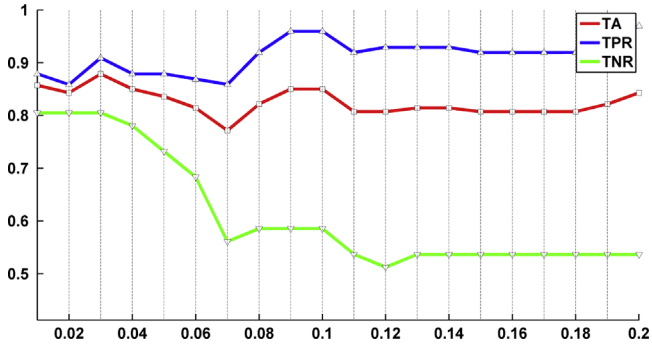

Fig. 3. Plot of TA versus different p_{cut}.

Table 5
Classification results with different numbers of HP models used.

No. of NC HP	No. of AD HP	*TPR*	*TNR*	Total accuracy
10		0.8000	1.0000	0.9608
20		0.7500	1.0000	0.9180
30		0.8000	0.9756	0.9014
40		0.8000	0.9268	0.8642
50	41	0.9000	0.8537	0.8791
60		0.9167	0.8283	0.8812
70		0.9143	0.8293	0.8829
80		0.8875	0.8293	0.8678
90		0.9111	0.7805	0.8702
99		0.9091	0.8049	0.8786

5. Evaluation of the model

In this section, we want to further evaluate the performance of our algorithm. In the previous section, the tuning scheme of those parameters is to optimize the performance of our model in the leave-one-out cross validation. However, such scheme may cause overfitting of parameters to the given database. That means, the accuracy of our algorithm may decrease if another database is being used while those parameters are kept unchanged. To check if overfitting happens, we perform the following tests.

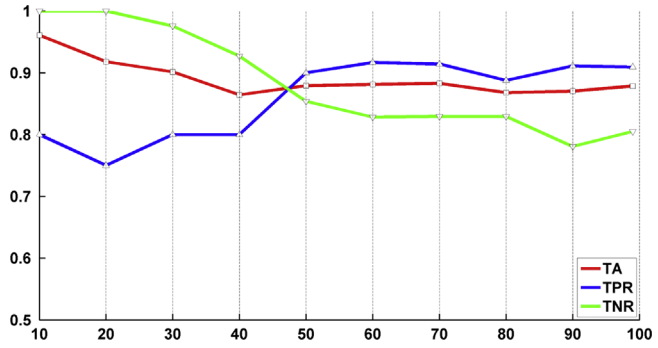

Fig. 4. Plot of TA versus different numbers of HP models used.

5.1. A reliability test

Recall that we are given altogether 140 subjects in our database. Now we perform 140 experiments, and in the i-th experiment, we remove the i-th subject out of our database, and use the remaining 139 subjects as database to do a leave-one-out test. So ultimately we would have done 140 leave-one-out tests. And throughout all the tests, those parameters are set constantly as in (23). In this way, the setting of the parameters is not optimal to any of the 140 different database. We compute the statistical distribution of these 140 classification accuracies. Fig. 6 shows the histogram of the 140 values. The mean, median and standard deviation are 87.20%, 87.77% and 0.81% respectively. And the 95% confidence interval is $(87.07\%, 87.34\%)$.

The result suggests that our algorithm is very stable to reach a classification accuracy over 87%, even though the parameters are not tuned to fit the database in the best way.

5.2. Separation of training data and testing data

Another test is done to further validate the performance of our algorithm. We randomly pick partial data from our database as training data, leaving the remaining as testing data. Parameters are tuned using only the training data, then we perform a leave-one-out cross validation using only the testing data. Hence, the setting of those parameters is completely independent of the testing data set. The result is recorded in Table 6.

The result shows that even if we separate the database into one training set and one testing set, the accuracy of our model still remains over 86%. Comparing with the 87.86% accuracy we got in

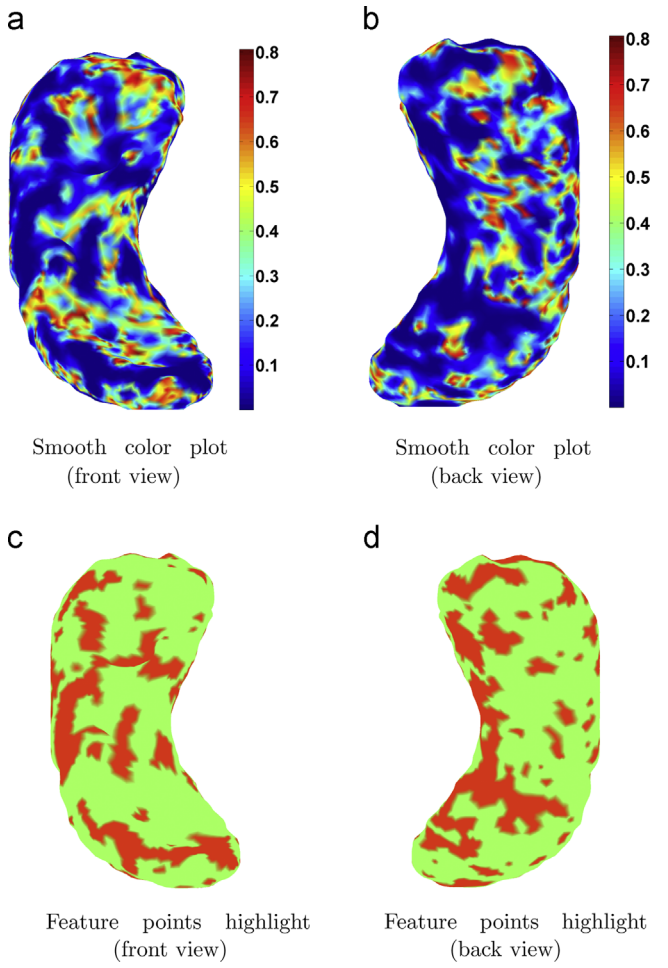

Fig. 5. Color-map of HP surface indicating p-values at each vertex. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

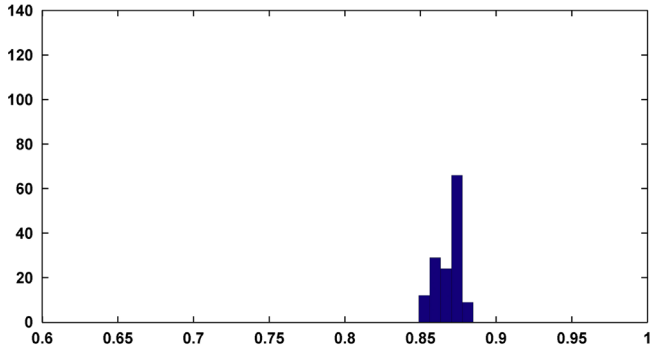

Fig. 6. Histogram of the classification accuracy of our algorithm under the optimal parameters. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

the previous section, a slight decrease in accuracy is recorded when the size of training data set is small. This is mainly a consequence of having insufficient training data. And when the number of data in the training set grows, the accuracy of our model generally increases. In particular, we even recorded an accuracy of over 92% in the test. Therefore, we are confident that the performances of our algorithm throughout all the previous experiments are reliable, being bias to neither the database setting nor the parameters setting.

6. Comparison and discussion

In our work, a simple L^2 classification model is applied to classify AD. In this section, we investigate the classification accuracies when other novel classification methods are used, namely, 1. K-mean clustering and 2. Support Vector Machine (SVM) clustering. Also, different approaches to classify Alzheimer's disease are to be compared with our model.

6.1. Comparison with K-mean clustering and Support Vector Machine

We modify our algorithm, using the K-mean clustering and then the Support Vector Machine clustering to replace the L^2 distance, respectively, and compare it with our original algorithm. The accuracies are compared with respect to the variation in the total number of HPs involved. In particular, when using K-mean clustering, we set $K_{NC} = K_{AD} = 2$, respectively. The result is recorded in Tables 7 and 8, and in Figs. 7 and 8.

From the result, the accuracy using the K-mean clustering method is comparable to that using the L^2 classification model. And when the database is more balanced, the overall performance of the modified algorithm is comparable to that of the original model. But when the whole database is used, using the L^2 classification model is better than using the K-mean clustering method. Nevertheless, this motivates us to try replacing the L^2 classification method by the K-mean clustering method when a large database is available.

For SVM clustering, we see that it generally helps our algorithm reach significantly higher overall accuracy. However, when the ratio of NC HPs to AD HPs is high, TNR drops sharply to even lower than 60%. This can be explained by the sensitivity of SVM clustering to the balance of the data set. Therefore, we are motivated by the result to replace the simple L^2 classification method by the more advanced SVM clustering method, provided once again we have a larger and balanced data size.

Table 6
Classification results with different numbers of HP models used as training data (NTNC=No. of training NC HP, NTAD=No. of training AD HP).

NTNC	NTAD	α	β	γ	p_{cut}	Total accuracy
15	15	0.7267	0.0015			0.8727
20	20	8.9796	8.9796			0.8600
25	25	10.2739	7.4650	1	0.03	0.8889
30	30	6.0064	1.9514			0.9250

Table 7
Classification results with different numbers of HP models used (K-mean clustering).

No. of NC HP	No. of AD HP	*TPR*	*TNR*	Total accuracy
10		0.9000	1.0000	0.9804
20		0.8000	1.0000	0.9344
30		0.7667	0.9756	0.8873
40		0.8750	0.8537	0.8642
50	41	0.9200	0.8537	0.8901
60		0.8833	0.8293	0.8614
70		0.9714	0.7805	0.9009
80		0.9750	0.7561	0.9008
90		0.9667	0.6829	0.8779
99		0.8990	0.7073	0.8429

Table 8
Classification results with different numbers of HP models used (SVM clustering).

No. of NC HP	No. of AD HP	*TPR*	*TNR*	Total accuracy
10		0.9000	1.0000	0.9804
20		0.7500	1.0000	0.9180
30		0.9000	0.9756	0.9437
40		1.0000	0.9024	0.9506
50	41	0.9600	0.8049	0.8901
60		0.9333	0.8293	0.8911
70		0.9571	0.7561	0.8829
80		0.9750	0.7073	0.8843
90		0.9444	0.6829	0.8626
99		0.9798	0.5854	0.8642

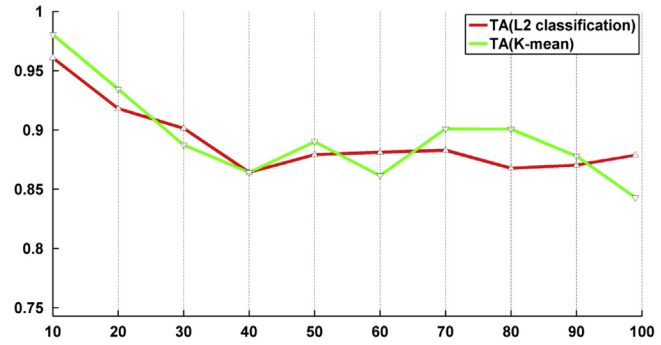

Fig. 7. Plots of TA versus different numbers of HP models using L^2 classification and K-mean clustering. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

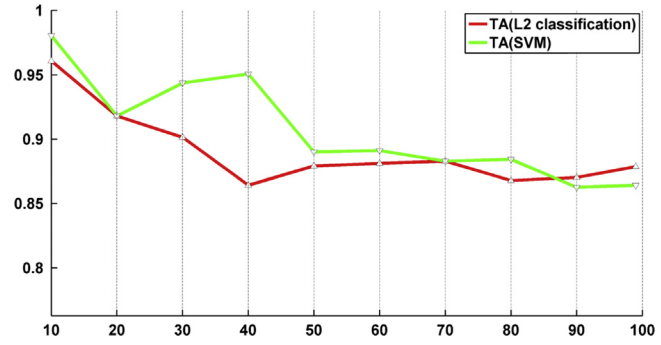

Fig. 8. Plots of TA versus different numbers of HP models using L^2 classification and SVM clustering. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

6.2. Comparison with other surface-based models

In our work, we propose to study the complex dilations supplied with curvature deformations on the hippocampus surface to classify the Alzheimer's disease. Experiments suggest that, with altogether 140 HP models, our model is confident to classify AD HPs from NC HPs with over 87% accuracy. In advance, by using 60 HP models as training data and the remaining 80 HP models as testing data, a leave-one-out cross validation on the testing data shows our approach being possible to reach over 92% accuracy.

In this section, we compare the accuracy of our model with that of other existing models. Since we are given the surface of hippocampus only, we will focus on other hippocampal-surface-based models to perform the comparison. And to promote fairness, for each model we use the leave-one-out cross validation on our given database.

Global shape analysis on HPs is a common approach to study AD. Some recent methods include, for example, in [14], volumes of the hippocampus are normalized and then the mean of the volumes for the NC group and the AD group are computed. A new subject is classified based on the difference between its volume of the mean volume of each group. And in [15], the authors propose a fully automatic method for hippocampus segmentation and volume-based AD classification algorithm. We apply these two algorithms on our database under the leave-one-out cross validation and the result is reported in Table 9.

Table 9
Comparison of classification accuracies between different models under leave-one-out cross validation on the whole given database (Parameters for each model, if any, are set according to the suggestion from the corresponding paper).

Method	Total accuracy
Volume-based [14]	0.7071
Volume-based [15]	0.7643
SPHARM-based [20]	0.8029
SPHARM-based [19]	0.8643
Our method	0.8786

The SPHARM approach is also a famous branch, in which the shape change of each HP is described by a series of spherical harmonics. Analysis is done on the coefficients of the spherical harmonics to capture the details of the shape deformation process. In [20], spherical conformal parametrization is used to compute the spherical harmonics coefficients, which are then summed up as shape invariants for AD classification. In [19], spherical harmonics coefficients are used, with the aid of bagging predictors for feature selection, to classify AD. We apply these two algorithms on our database under the leave-one-out cross validation and the result is reported in Table 9.

For global shape analysis based algorithms, results show that the accuracies of them are around 70–77%. In addition, in [18], authors tested two models using volume-based approach to classify AD. Although the tests are performed on a different database (which contain 162 NC subjects and 137 AD subjects from ADNI database), results show that the two methods report 72–74% overall accuracy, which is significantly lower than that of our algorithm. Comparing all these models with our model, which achieves beyond 87% overall accuracy, it is persuasive that our model is better than those models. This can be explained as the difference between global measurements (volume) and local measurements (quasi-conformality, curvature) on HPs. More specifically, the locally defined quasi-conformality is able to capture fine details about the HP deformation which globally defined measurements (e.g. volumes) may not capture. It is believed that this is the major factor that contributes to the difference in the classification accuracies. As for the spherical harmonics approach, information may be lost when selecting a finite number of elements out of the spherical harmonic series, whereas our approach uses quasi-conformality and curvature to capture local shape deformations exactly. Therefore, our method generally performs better than those SPHARM-based methods.

6.3. Discussion on other non-surface-based models

The study of AD classification is not limited to HP analysis. There are models that concern not only the hippocampus but also other features such as cortical thickness, probability of different tissue classes in a given voxel, or even the volumetric whole brain data are used to classify AD. For example, [18] tested 4 categories of voxel-based approach and they reported 71–88% overall accuracy. Besides, [18] also tested 3 categories of cortical-thickness-based approach and they turned out to have 83–85% overall accuracy. On the other hand, based on multiple atlases that capture information of volumetric brain changes related to AD, [38] suggested a criterion and a model to select the best atlases to

classify AD and this method reached 91% accuracy with 128 NC data and 97 AD data. Although we cannot perform a direct comparison between these models with ours due to the lack of non-hippocampal-surface information, our model, which achieves over 87% accuracy in a leave-one-out test using the given full data set and got over 92% accuracy under different setting of the data set, is reliable to give a fair accuracy compared to these models.

7. Conclusion and future work

In this paper, we proposed a quasi-conformal based statistical shape analysis model on hippocampal surfaces to study Alzheimer's disease. Given a set of HP surfaces as training data, our algorithm computes a shape index measuring local regional geometric changes, including quasi-conformality and curvature changes, on each vertex of the surface. Based on the shape index, an automatic machine learning algorithm is then built to classify any newly imported HP surface into either NC or AD group. Experimental results show that, in particular, the quasi-conformality term plays a crucial role in the classification accuracy of our proposed model. According to the results, the maximal total classification accuracy of our algorithm is 87.86% in a leave-one-out cross validation using the whole given database. In addition, the results show that there is a possibility to further boost up the classification accuracy of our algorithm using more advanced clustering methods, provided that the database is balanced and large enough.

There are several directions which are still under investigation. Firstly, our current algorithm fixes the parameters, such as α and β in defining ${E_{shape}}^i$ and p_{cut}. In the future, we aim to develop an algorithm to simultaneously search for the optimal parameters in our classification model. Secondly, the current classification algorithm utilizes a simple L^2 threshold clustering model. While this simple model together with the quasi-conformality gives reasonably satisfactory classification accuracy, a more advanced classification model may further boost up the classification accuracy. Other than the K-mean clustering and the SVM clustering, another approach may be the following. Instead of studying the L^2 distance of the shape index between the input subject and the template mean subject, the diffeomorphic deformation from the shape index of the input subject to that of the template mean subject can be studied, and hence a more robust and accurate shape-based geometric distance classification model may be built, based on the analysis on the deformation. Hence, our future works also include the investigation of more classification models and combining them into our current model to improve the performance of our method. Finally, we will examine our method on a larger and more balanced dataset, and consider including more sub-cortical brain structures as inputs.

Acknowledgment

Data used in preparation of this article were obtained from the Alzheimers 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 provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found here: ADNI Acknowledgement List.

The authors would like to thank Prof. Paul M. Thompson for his valuable advices in developing the model. Lok Ming Lui is supported by RGC GRF (GRF Project ID: 2130271; CUHK Project ID: 401811) and CUHK FIS Grant (Project ID: 1902036).

References

[1] G. Chetelat, J.C. Baron, Early diagnosis of Alzheimer's disease: contribution of structural neuroimaging, Neuroimage 18 (2003) 525–541.
[2] B. Dubois, et al., Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria, Lancet Neurol. 6 (August (8)) (2007) 734–746.
[3] G.B. Frisoni, et al., The clinical use of structural MRI in Alzheimer disease, Nat. Rev. Neurol. 6 (2) (2010) 67–77.
[4] G.W. Small, et al., Current and future uses of neuroimaging for cognitively impaired patients, Lancet Neurol. 7 (2) (2008) 161–172.
[5] J. Barnes, et al., Atrophy rates of the cingulate gyrus and hippocampus in AD and FTLD, Neurobiol. Aging 28 (1) (2007) 20–28.
[6] J. Barnes, et al., A meta-analysis of hippocampal atrophy rates in Alzheimer's disease, Neurobiol. Aging 30 (11) (2009) 1711–1723.
[7] MonyJ. De Leon, et al., Early marker for Alzheimer's disease: the atrophic hippocampus, The Lancet 334 (1989) 672–673.
[8] F. Shi, et al., Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: meta-analyses of MRI studies, Hippocampus 19 (11) (2009) 1055–1064.
[9] G.B. Frisoni, et al., Mapping local hippocampal changes in Alzheimer's disease and normal ageing with MRI at 3 Tesla, Brain November (2008) 3266–3276 , Nov.
[10] P.M. Thompson, et al., Mapping hippocampal and ventricular change in Alzheimer disease, NeuroImage 22 (August (4)) (2004) 1754–1766.
[11] M. Bobinski, et al., Neuronal and volume loss in CA1 of the hippocampal formation uniquely predicts duration and severity of Alzheimer disease, Brain Res. 805 (September (12)) (1998) 267–269.
[12] M. Rssler, R. Zarski, J. Bohl, T.G. Ohm, Stage-dependent and sector-specific neuronal loss in hippocampus during Alzheimer's disease, Acta Neuropathol. 103 (April (4)) (2002) 363–369.
[13] S.G. Mueller, M.W. Weiner, Selective effect of age, Apo e4, and Alzheimer's disease on hippocampal subfields, Hippocampus 19 (6) (2009) 558–564.
[14] O. Colliot, et al., Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus, Radiology 248 (2008) 194–201.
[15] M. Chupin, E. Grardin, et al., Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI, Hippocampus 19 (6) (2009) 579–587.
[16] J.H. Morra, et al., Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation, IEEE Trans. Med. Imaging 29 (2010) 30–43.
[17] S.G. Mueller, et al., Hippocampal atrophy patterns in mild cognitive impairment and Alzheimer's disease, Hum. Brain Mapp. 31 (2010) 1339–1347.
[18] R. Cuingnet, et al., 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) (2011) 766–781, http://dx.doi.org/10.1016/j.neuroimage.2010.06.013, ISSN 1053-8119.
[19] E. Gerardin, G. Chtelat, et al., Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging, NeuroImage 47 (October (4)) (2009) 1476–1486.
[20] B. Gutman, Y. Wang, et al., Disease classification with hippocampal shape invariants, Hippocampus 19 (6) (2009) 572–578.
[21] S. Li, et al., Hippocampal shape analysis of Alzheimer disease based on machine learning methods, Am. J. Neuroradiol. 28 (August (7)) (2007) 1339–1345.
[22] L. Zhou, et al., Hippocampal shape analysis for Alzheimer's disease using an efficient hypothesis test and regularized discriminative deformation, Hippocampus 19 (June (6)) (2009) 533–540.
[23] W.J. Henneman, et al., Hippocampal atrophy rates in Alzheimer disease: added value over whole brain volume measures, Neurology 72 (2009) 999–1007.
[24] R. Wolz, et al., Measurement of hippocampal atrophy using 4D graph-cut segmentation: application to ADNI, NeuroImage 52 (2010) 109–118.
[25] L. Younes, M. Albert, M.I. Miller, Inferring changepoint times of medial temporal lobe morphometric change in preclinical Alzheimer's disease, NeuroImage: Clinical 5 (2014) 178–187.
[26] Y. Wang, et al. Multivariate tensor-based brain anatomical surface morphometry via holomorphic one-forms, in: Medical Image Computing and Computer-Assisted Intervention, MICCAI, Springer, Berlin, Heidelberg, 2009, pp. 337–344.
[27] K. Van Leemput, et al., Automated segmentation of hippocampal subfields from ultra-high resolution in vivo MRI, Hippocampus 19 (2009) 549–557.
[28] P.A. Yushkevich, et al., Nearly automatic segmentation of hippocampal subfields in in vivo focal T2-weighted MRI, NeuroImage 53 (2010) 1208–1224.
[29] L.M. Lui, et al., Optimization of surface registrations using Beltrami holomorphic flow, J. Sci. Comput. 50 (3) (2012) 557–585.
[30] Judith Miklossy, Alzheimer's disease—a neurospirochetosis. Analysis of the evidence following Koch's and Hill's criteria, J. Neuroinflammation 8 (2011) 90.
[31] M. Comelli, U. Lucca, A. Spagnoli, Statistical analysis of the clinical trial of a therapy for Alzheimer's disease. Univariate tests and logistic regression, Acta Neurol. (Napoli) 12 (June (3)) (1990) 222–230.
[32] K. Chen, et al., The Alzheimer's Disease Neuroimaging Initiative (2010): twelve-month metabolic declines in probable Alzheimer's disease and amnestic mild cognitive impairment assessed using an empirically

pre-defined statistical region-of-interest: findings from the Alzheimer's Disease Neuroimaging Initiative, NeuroImage 51 (2) (2010) 654–664.
[33] Y. Fan, et al., COMPARE: classification of morphological patterns using adaptive regional elements, IEEE Trans. Med. Imaging 26 (2007) 93–105.
[34] Y. Fan, et al., Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline, NeuroImage 39 (2008) 1731–1743.
[35] S. Kloppel, et al., Automatic classification of MR scans in Alzheimer's disease, Brain 131 (2008) 681–689.
[36] Z. Lao, et al., Morphological classification of brains via high-dimensional shape transformations and machine learning methods, NeuroImage 21 (2004) 46–57.
[37] P. Vemuri, et al., Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies, NeuroImage 39 (2008) 1186–1197.
[38] Rui Min, et al., Maximum-margin based representation learning from multiple atlases for Alzheimers disease classification, Hum. Brain Mapp. (2015) 1847–1865 36.5.
[39] J. Eells, L. Lemaire, A report on harmonic maps, Bull. Lond. Math. Soc. 10 (1978) 1–68.
[40] J. Eells, L. Lemaire, Another report on harmonic maps, Bull. Lond. Math. Soc. 20 (1988) 385–524.
[41] L.M. Lui, et al., Landmark constrained genus zero surface conformal mapping and its application to brain mapping research, Appl. Numer. Math. 57 (2007) 847–858.
[42] Y. Wang et al. Optimization of brain conformal mapping with landmarks, in: Proceedings of the Medical Image Computing and Computer Assisted Intervention Part II, October 2005, pp. 675–683.
[43] L.M. Lui, et al., Optimized conformal surface registration with shape-based landmark matching, SIAM J. Imaging Sci. 3 (1) (2010) 52–78.
[44] T.W. Wong, et al., Instrinic feature extraction and Hippocampal Surface Registration using Harmonic eigenmap, SIAM J. Imaging Sci. 5 (2) (2012) 746–768.
[45] L.M. Lui, et al., Hippocampal shape registration using Beltrami holomorphic flow, in: Medical Image Computing and Computer Assisted Intervention (MICCAI), Part II, Lecture Notes in Computer Science, vol. 6362, 2010, pp. 323–330.
[46] L.M. Lui, et al., Detecting shape deformations using yamabe flow and beltrami coefficents, J. Inv. Problem Imaging 4 (2) (2010) 311–333.
[47] W. Zeng, et al., Quasiconformal maps using discrete curvature flow, Numer. Math. 121 (2012) 671–703.
[48] L.M. Lui, C.F. Wen, Geometric registration of high-genus surfaces, SIAM J. Imaging Sci. 7 (1) (2014) 337–365.
[49] L.M. Lui, et al., Teichmuller mapping (T-Map) and its applications to landmark matching registrations, SIAM J. Imaging Sci. 7 (1) (2014) 391–426.
[50] A. Telea, A. Vilanova, A robust level-set based algorithm for centerline extraction, in: Proceedings of the Symposium on Data Visualisation 2003, Eurographics Association, Grenoble, France, 2003, pp. 185–194.
[51] Breiman Leo, Bagging predictors, Mach. Learn. 24 (2) (1996) 123–140.
[52] P. Aljabar, et al., Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy, NeuroImage 46 (2009) 726–738.
[53] Jyrki M.P. Lotjonen, et al., Fast and robust multi-atlas segmentation of brain magnetic resonance images, NeuroImage 49 (2010) 2352–2365.

Hei Long Chan received the B.Sc. degree from the Department of Mathematics, the Chinese University of Hong Kong, Hong Kong in 2014. He is currently a M. Phil. student in the Chinese University of Hong Kong, Department of Mathematics. His research interests include computational conformal and quasi-conformal geometry, medical imaging, shape analysis and image segmentation.

Hangfan Li received the B.Sc. degree from the Department of Mathematics, the Hong Kong Baptist University, Hong Kong in 2012. He is currently a M.Phil. student in the Chinese University of Hong Kong, Department of Mathematics and is going to graduate. His research interests include computational conformal and quasi-conformal geometry, medical imaging and shape analysis.

Lok Ming Lui received the Ph.D. degree in applied mathematics from the University of California at Los Angles, Los Angeles, CA, USA, in 2008. He is an Assistant Professor with the Department of Mathematics, Chinese University of Hong Kong (CUHK), Hong Kong. Before joining CUHK, he was a Post-Doctoral Scholar for 2 years with the Department of Mathematics, Harvard University, Cambridge, MA, USA. His current research interests include computational conformal and quasi-conformal geometry, Teichmuller theory, surface registration, medical imaging, and shape analysis.