

Statistical shape analysis of subcortical structures using spectral matching



Mahsa Shakeri^{a,b,*}, Herve Lombaert^c, Alexandre N. Datta^d, Nadine Oser^d, Laurent Létourneau-Guillon^e, Laurence Vincent Lapointe^f, Florence Martin^f, Domitille Malfait^{b,f}, Alan Tucholka^{b,g}, Sarah Lippé^{b,f}, Samuel Kadoury^{a,b}, Alzheimer's Disease Neuroimaging Initiative¹

^a Polytechnique Montréal, Department of Computer and Software Engineering, Centre-ville, Montreal, QC, Canada H3C 3A7

^b Sainte-Justine Hospital, Research Center, 3175 Cote-Sainte-Catherine Rd., Montreal, QC, Canada H3T 1C5

^c Inria Sophia-Antipolis Méditerranée, Asclepios Team, 06902 Sophia-Antipolis, France

^d University of Basel Children's Hospital, Basel, Switzerland

^e Notre-Dame Hospital, Department of Radiology, Montreal, QC, Canada H2L 4M1

^f University of Montreal, Department of Psychology, Montreal, QC, Canada H3C 3J7

^g Barcelona Beta Brain Research Center, Pasqual Maragall Foundation, Barcelona, Spain

ARTICLE INFO

Article history:

Received 12 August 2015

Received in revised form 2 February 2016

Accepted 4 March 2016

Keywords:

Subcortical morphology
Groupwise shape analysis
Spectral matching
Alzheimer's disease

ABSTRACT

Studying morphological changes of subcortical structures often predicate neurodevelopmental and neurodegenerative diseases, such as Alzheimer's disease and schizophrenia. Hence, methods for quantifying morphological variations in the brain anatomy, including groupwise shape analyses, are becoming increasingly important for studying neurological disorders. In this paper, a novel groupwise shape analysis approach is proposed to detect regional morphological alterations in subcortical structures between two study groups, e.g., healthy and pathological subjects. The proposed scheme extracts smoothed triangulated surface meshes from segmented binary maps, and establishes reliable point-to-point correspondences among the population of surfaces using a spectral matching method. Mean curvature features are incorporated in the matching process, in order to increase the accuracy of the established surface correspondence. The mean shapes are created as the geometric mean of all surfaces in each group, and a distance map between these shapes is used to characterize the morphological changes between the two study groups. The resulting distance map is further analyzed to check for statistically significant differences between two populations. The performance of the proposed framework is evaluated on two separate subcortical structures (hippocampus and putamen). Furthermore, the proposed methodology is validated in a clinical application for detecting abnormal subcortical shape variations in Alzheimer's disease. Experimental results show that the proposed method is comparable to state-of-the-art algorithms, has less computational cost, and is more sensitive to small morphological variations in patients with neuropathologies.

© 2016 Elsevier Ltd. All rights reserved.

* Corresponding author at: Polytechnique Montréal, Department of Computer and Software Engineering, Montreal, Quebec, Canada H3C 3A7.

E-mail address: mahsa.shakeri@polymtl.ca (M. Shakeri).

¹ Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

1. Introduction

Quantifying groupwise neuroanatomical shape differences has become an important topic in neuroscience as well as in neuroimaging studies, since brain morphometry has been hypothesized to be linked to various neurological disorders [1]. Recent advances in medical image analysis have led to several morphological studies on different pathologies including schizophrenia [2] and Alzheimer's disease [3]. Early studies on brain morphology were based on volumetric analysis, which had the advantage of simplicity [3,4]. However, these methods lacked regional shape information, which are potentially important for neurological

identification. Thus, shape analysis has emerged as an alternative of evaluating morphology location and magnitude in the brain anatomy, in order to detect the neuroanatomical changes in specific regions of a single structure.

Several works have proposed groupwise shape analysis approaches based on creating a template image and estimating the deformation field between each individual image and the template [5–7]. These methods non-rigidly register each image to a specific template, and evaluate the information contained within the deformation field to detect the location and magnitude of shape changes. In deformation-based approaches the template creation and the choice of the registration technique has a critical impact on the quality of the shape analysis results [8]. Indeed, deformable registration remains a challenging problem, particularly in smaller regions such as subcortical structures, which requires a high level of accuracy to precisely match sub-fields.

Another approach for shape analysis is based on medial surface representations [9–12]. These methods provide a compact parameterization of a volumetric object by extracting the 3D skeleton of a shape. The local positional changes are then quantified by assessing morphological variations of the skeleton across a population of shapes. The medial surface representation has been applied to various subcortical structures including the cross sectional images of the corpus callosum [9] and hippocampus/amygdala complex [11], caudate [12], and lateral ventricle [13], in order to evaluate changes in patients diagnosed with schizophrenia. Since medial representation relies on a coarse-scale sampling of the structure of interest, it may be insensitive to small-scale shape differences, which might be present in a studied population [2].

In addition to the above-mentioned methods, several works have proposed surface parameterization based approaches, in which a parameterization of a shape is performed on a simple domain such as a sphere. One such approach is the spherical harmonics (SPHARM) method by Styner et al. [14], which is used in combination with Point Distribution Models (PDM) to discover structural differences across populations. In this approach, which is called SPHARM-PDM, shapes first represented by binary segmentations are converted into a corresponding spherical harmonic description. Then, the correspondence problem is solved by the alignment of the spherical parameterization using a first-order ellipsoid. In this method, the spherical description of surface meshes is sampled into triangulated surfaces using icosahedron subdivision. These surfaces are then spatially aligned using rigid Procrustes alignment.

The SPHARM-PDM method has been applied in various clinical applications, such as for studying shape variations of the hippocampus [2] and lateral ventricles [13] in schizophrenia, and analyzing caudate morphological changes [15] in bipolar disorder. Although SPHARM-PDM enabled the reliable assessment of local shape variation across populations, it requires establishing correspondence on simplified spherical models of surfaces, which is restricted to surfaces with spherical topology [16,17]. In order to overcome this limitation, combining SPHARM-PDM with an entropy-based particle system correspondence model has been proposed in [16]. Furthermore, Cates et al. [17] proposed to model the shape as sets of particles, where the particles are distributed on the surfaces of the shapes by optimizing an entropy-based energy function. According to a recent evaluation in [8], this method was not able to find the location of group differences in some cases.

Chung et al. [18] proposed weighted-SPHARM, which expresses surface data as a weighted linear combination of spherical harmonics. The weighted-SPHARM method generalizes the traditional SPHARM representation as a special case. This method reduces ringing artifacts observed with the SPHARM representation especially for the high frequency components [18]. However, it is applicable for a limited class of shapes and is only able to detect

the sufficiently large shape difference across populations. The SPHARM-MAT (SPHARM Modeling and Analysis Toolkit) is another shape analysis method, which creates parametric surface models using spherical harmonics [19]. This approach is similar to the SPHARM-PDM framework [14] but applies different spherical parameterization and shape alignment process [19]. A comparison between the two spherical harmonics-based methods showed that SPHARM-PDM could capture the shape differences more accurately rather than the SPHARM-MAT [8].

In general, despite the development of various approaches in the field of groupwise shape analysis, there are still some limitations and challenges that need to be addressed. For instance, the morphometry framework should be robust to different shape topologies and shape segmentation approaches, while from a computational perspective, the total population-wise shape analysis process should be determined in an efficient time manner, in order to be feasibly implemented as a clinical diagnostic tool.

One of the main challenges in a surface-based groupwise shape analysis approach is to establish reliable one-to-one correspondences among the population of surfaces. Among various surface matching algorithms in the literature, the recent spectral matching approach proposed by Lombaert et al. [20,21] have shown promising results, which enables mapping brain cortical surfaces in a computational efficient manner. To the best of our knowledge, no studies have been based on spectral matching theory directly for groupwise shape analysis. In this paper, we employ the spectral matching approach in a groupwise shape analysis pipeline, designed for the population-wise comparison between two groups of subcortical structures. In the proposed framework, the mean curvature feature is used as feature information in the spectral matching process in order to increase the surface matching accuracy. This work presents a complete pipeline that allows the detection, localization, and quantification of statistically significant morphological differences in different subcortical brain structures across various populations.

The main contribution of this paper is to propose a robust and reliable spectral-based shape analysis framework using curvature features for analyzing simple (e.g., putamen) and complex (e.g., hippocampus) subcortical structures. Furthermore, in contrast to the above-mentioned methods, the suggested framework can perform population-wise shape analysis in a computational efficient fashion and integrates the spectral matching approach with a groupwise shape analysis framework. This paper is the first work that integrates spectral matching in to a groupwise subcortical shape analysis pipeline and incorporates curvature features to increase the surface matching accuracy (methodological contribution). Experimental analysis on real clinical datasets show that the extracted group differences are similar to the findings of other clinical studies (clinical contribution). This work would represent a significant forward for providing an alternative for future clinical studies on sub-cortical brain structures.

The proposed shape analysis framework processes two groups of segmented binary images from neuroimaging data (e.g., MRI) as input. After converting the images into triangulated surface meshes and applying the curvature flow smoothing process, the surface correspondence is established between two populations using the spectral matching approach. The mean curvature features are incorporated in the spectral matching process to improve the accuracy of the resulting surface correspondence. Then, two mean shapes for each group are created as the geometric mean of all surfaces, and the distance map between the two mean shapes is computed. To verify for statistically significant differences between two populations, a non-parametric permutation testing scheme [22] is applied, followed an FDR correction [23] for multiple comparisons. The evaluation of the proposed shape analysis framework is performed on three separate neuroanatomical

datasets, and group differences across populations are detected by point-to-point correspondences. The output of these analyses is compared to a state-of-the-art method used for groupwise analysis of anatomical shapes, namely SPHARM-PDM.

A preliminary version of the proposed method has been presented in [24], but was dependent of spherical representations of input shapes, integrated a point-based shape matching method, and with limited validation. The present paper provides the detailed explanation of the methodology and offers the following major contributions: (i) a curvature-based surface smoothing is employed to smoothen extracted 3D surfaces, which has less computational requirements compared to the spherical-based smoothing method used in [24]; (ii) mean curvature feature was added into the spectral matching process to increase the surface matching accuracy; (iii) a statistical permutation-based analysis is applied to detect the significant shape differences across populations; (iv) the evaluation of the proposed pipeline is performed on an increased number of clinical datasets with more samples.

The remainder of this paper is organized as follows. In Section 2, the methodology is presented, including the overview of the framework, the spectral surface matching approach, and the detailed description of the proposed population-wise shape analysis framework. The quantitative evaluation of the shape analysis method is provided in Section 3. Section 4 presents a discussion on the results obtained from different neuroimaging datasets, followed by a conclusion in Section 5.

2. Materials and methods

2.1. Framework overview

In this section, an overview of the proposed groupwise shape analysis pipeline is presented (Fig. 1). The inputs of the proposed framework are two sets of segmented binary maps of a particular structure of interest from medical images, obtained either manually or automatically. The output of the proposed pipeline is the extracted local and global shape differences between two sets.

In the first step, the segmented binary maps are processed to the same image orientation and isotropic voxel sizes, and then converted into 3D triangulated surfaces using the marching cube algorithm [25]. A smoothing process is subsequently applied on each surface in order to remove surface noise. Then, a reference surface is defined in an iterative process, and all triangulated surfaces are aligned to this reference using a rigid registration algorithm. In order to establish the point-to-point correspondences across all surfaces, each mesh is matched to a selected reference surface using the spectral matching algorithm. The vertices of all surfaces are rearranged to create the new reconstructed meshes with consistent vertex ordering. This enables to create a mean shape of each study group, and detect any morphological variations between two groups. A distance map between two mean shapes is computed to capture the local group differences across populations. Moreover, average and maximum distances, as well as Dice volume overlap are calculated to indicate the global group differences between two study groups.

2.2. Surface spectral matching

This section presents the theoretical framework of computing the spectral matching between two surfaces in the spectral domain. The matching between two surfaces S_1 and S_2 is conducted in a two-step process (Fig. 2a and b). In the first step, an initial map is calculated between the two surfaces [20]. This initial map is then used in the second step to establish a smooth map between the two meshes [21].

For each surface mesh $S_i (i \in \{1, 2\})$ composed of the set of vertices V_i and edges E_i (neighboring points in mesh faces), the corresponding graph $G_i = \{V_i, E_i\}$ is built. Then, the weighted adjacency matrix W_i is created based on a distance between connected vertices and the graph Laplacian matrix L_i [26] is defined as:

$$L_i = G_i(D_i - W_i) \quad (1)$$

where, D_i is a diagonal matrix with elements given by the degree of vertices. G_i is another diagonal matrix, which could be considered as $G_i = D_i^{-1}$ or any other (positive valued) vertex weights [20,27]. The eigendecomposition of the Laplacian matrix L_i computes the eigenvalues and the associated eigenmodes. The spectrum (spectral representation) \tilde{S}_i is defined as the first k eigenmodes associated with the non-zero eigenvalues.

Once the spectrums \tilde{S}_1 and \tilde{S}_2 are computed, the reordering and sign adjustment [20] process are performed and the resulting spectrums are aligned using the Coherent Point Drift (CPD) method [28]. Then, the correspondence map c between two surfaces S_1 and S_2 is established with a simple nearest-neighbor search in the spectral domain. An overview of the procedure to find the correspondence map c is shown in Fig. 2a.

Given the initial map c between two surfaces, the final map φ (smooth match) is obtained according to Fig. 2b. In this procedure, an association graph $G_a = \{V_{1,2}, E_{1,2,c}\}$ composed of the set of vertices and edges of S_1 and S_2 with the initial links c between the two surface meshes is created. Then, the Laplacian matrix is created as $L_a = G_a(D_a - W_a)$, where, W_a , G_a , and D_a could be defined in the similar way as Eq. (1). The eigendecomposition of the Laplacian matrix L_a produces a shared set of eigenvectors that enables a direct mapping $\varphi_{1 \rightarrow 2}$ between two meshes (see [21] for more details).

According to [20,27], considering higher node weights in the graph G_a could improve the spectral matching precision between two surfaces. This could be achieved by adding extra features on node weights G_a [20,27]. In this paper, we propose to use the mean curvature [29] as an additional feature, due to its ability to represent the pointwise characteristic within a surface. In fact, the exponential of the mean curvature is added to the node weights in the Laplacian graph L_a to assign higher weights to each vertex. In order to compute the mean curvature at each vertex v , the principal curvatures $C_{\min}(v)$ and $C_{\max}(v)$ [29] are calculated as the minimum and maximum curving degrees of the surface S_i . Then, the mean surface curvature C at each vertex v is defined as:

$$C(v) = \frac{1}{2}(C_{\min}(v) + C_{\max}(v)) \quad (2)$$

Hence, the mean curvature of surface S_i is computed as $\{C(1), C(2), \dots, C(n)\}$, where, n is the total number of vertices in the surface S_i . The additional information is incorporated in the weighting of the nodes by computing the exponential of the mean curvature, and defining the diagonal node weighting matrix G_a as:

$$G_a = D_a^{-1}(\exp(\text{diag}(C(1), C(2), \dots, C(n)))) \quad (3)$$

The effect of the curvature feature on the matching accuracy will be assessed in Section 3.

2.3. Groupwise shape analysis

2.3.1. Preprocessing and surface representation

Let $\{I_i^A\}_{i=1, \dots, N_A}$ and $\{I_i^B\}_{i=1, \dots, N_B}$ be two separate sets of segmented binary maps of a particular subcortical structure, with N_A and N_B as the number of samples in each set, respectively. To extract smooth 3D triangulated surface meshes from input binary maps, the labeled datasets are first reoriented in the same image orientation and resampled to isotropic voxel sizes of $1 \times 1 \times 1 \text{ mm}^3$. Then, the isosurfaces of the segmentation are extracted using the

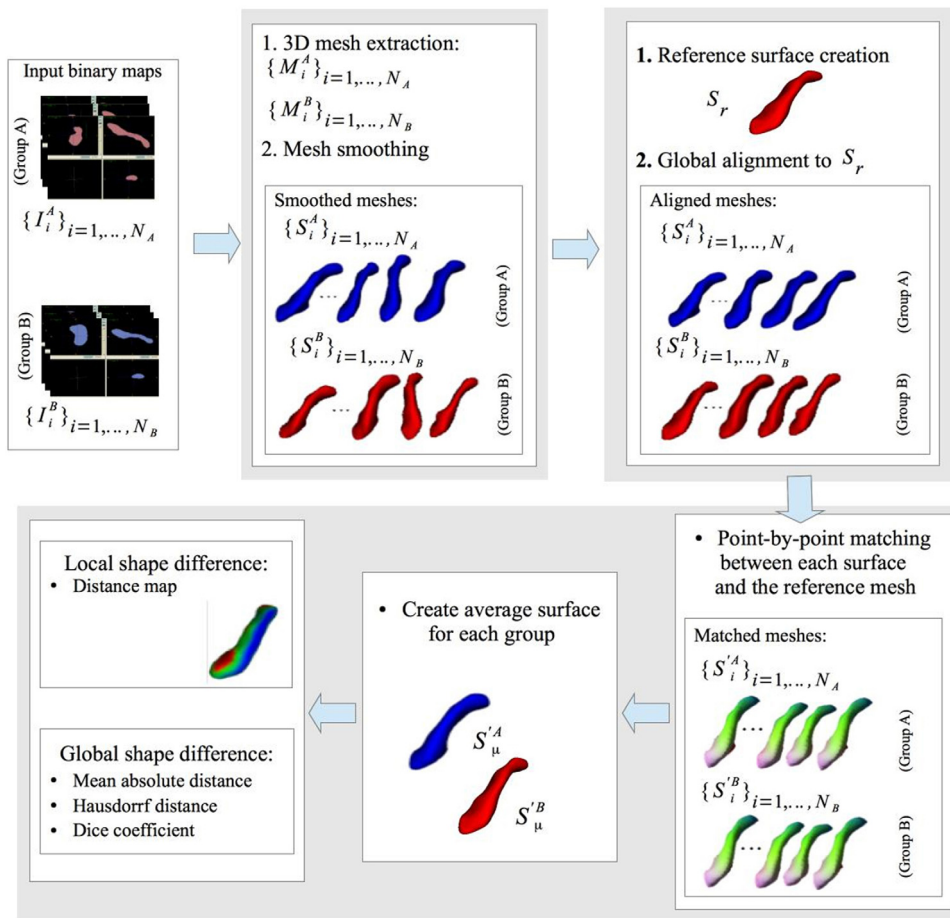


Fig. 1. Shape analysis pipeline based on spectral decomposition between two groups of subjects ($\{I_i^A\}_{i=1, \dots, N_A}$ and $\{I_i^B\}_{i=1, \dots, N_B}$, where N_A and N_B are the number of samples in each set). Once 3D triangulated surfaces $\{M_i^A\}_{i=1, \dots, N_A}$ and $\{M_i^B\}_{i=1, \dots, N_B}$ are extracted, a mesh smoothing process is performed to create surfaces $\{S_i^A\}_{i=1, \dots, N_A}$ and $\{S_i^B\}_{i=1, \dots, N_B}$. Then, the reference surface S_r is created and all surface meshes are aligned to the reference. The point-to-point matching between all surfaces is computed using the spectral matching algorithm to generate the surfaces $\{S_i^{\prime A}\}_{i=1, \dots, N_A}$ and $\{S_i^{\prime B}\}_{i=1, \dots, N_B}$. Finally, the mean surfaces $S_\mu^{\prime A}$ and $S_\mu^{\prime B}$ are created and local and global shape differences are computed.

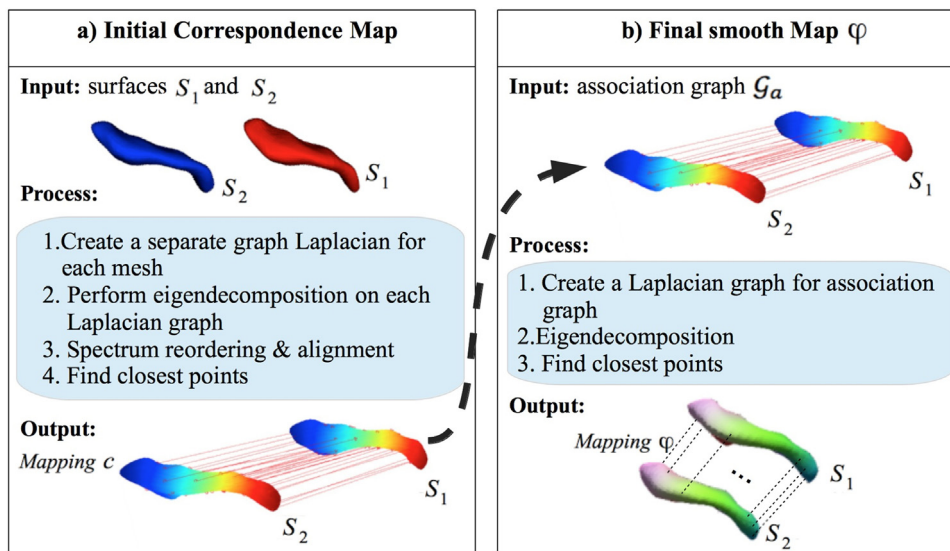


Fig. 2. Surface matching between two surfaces. (a) Initial matching of two meshes. (b) Final correspondence mapping between two surfaces based on diffeomorphic spectral matching approach.

marching cube algorithm [25] to generate 3D surface meshes denoted as $\{M_i^A\}_{i=1,\dots,N_A}$ and $\{M_i^B\}_{i=1,\dots,N_B}$ for the two study groups.

The 3D triangulated surface meshes are then smoothed using mean curvature flow algorithm [30], in order to create the smoothed meshes $\{S_i^A\}_{i=1,\dots,N_A}$ and $\{S_i^B\}_{i=1,\dots,N_B}$. This smoothing step is an iterative process, which removes surface irregularities and improves the appearance of surfaces. Given a triangulated surface M_i and for each mesh vertex x , the mean curvature flow is computed using the following explicit vertex-updating scheme:

$$x_{\text{new}} \leftarrow x_{\text{old}} + \lambda H(x_{\text{old}}) \quad (4)$$

where, the step-size λ is a small positive number. This parameter should be chosen small enough to keep the smoothing process stable. Here, $H(x_{\text{old}})$ is a discrete approximation of the mean curvature vector at a mesh vertex x , which is defined as:

$$H(x_{\text{old}}) = \frac{1}{4\theta} \sum_i (\cot \alpha_i + \cos \beta_i (Q_i - x_{\text{old}})) \quad (5)$$

where, θ is the sum of the areas of the triangles surrounding x and $\{Q_i\}$ are the neighbors of the vertex x_{old} . Here, α_i and β_i are the two angles opposite to the edge $Q_i x_{\text{old}}$. The process defined in Eq. (4) is applied to every point of each triangulated surface M_i iteratively to further smooth the mesh. High iterations produce smoother mesh, but it might cause the loss of fine surface details. Removing fine surface information could mislead a groupwise morphological study in capturing small shape differences. In this paper, the 3D surfaces are smoothed after five iterations. This is the least smoothing that is required to avoid failure in spectral matching process. The parameter λ is set to 1.0 as suggested in [30].

2.3.2. Reference surface and alignment

As a prerequisite for any shape analysis study, objects have to be normalized with respect to a reference coordinate frame. In order to define the reference surface S_r , an arbitrary surface from the input dataset can be chosen randomly, and used as a reference mesh. Alternatively, an average template could be created in an iterative process and employed as a reference surface. In this study, these two different reference selection approaches are tested, in order to evaluate their impact on the performance of the proposed shape analysis framework.

The step-by-step procedure of the average reference computation is presented in Algorithm 1. This algorithm is based on the approach proposed by Guimond et al. [31], in which an average image was created by alternating between pairwise registrations and updates of the average image. The input to Algorithm 1 is the union of the two sets $\{S_i^A\}_{i=1,\dots,N_A}$ and $\{S_i^B\}_{i=1,\dots,N_B}$, which is defined as $\{S_i\}_{i=1,\dots,(N_A+N_B)}$. Starting from an initial reference surface (e.g., $S_r = S_1$), all other surfaces S_i are rigidly aligned to the selected reference mesh. This rigid alignment is performed using the iterative closest point (ICP) algorithm [32], which finds the optimal rigid transformation between each mesh and the reference S_r . In the next step, the point-to-point mapping $\varphi_{i \rightarrow r}$ between each surface S_i and S_r is computed based on the spectral matching framework presented in Section 3. In the final step, the geometrical mean shape is estimated from all matched structures, and used as an updated reference shape. This process terminates when no more changes in S_r obtained or the maximum number of iterations is achieved. According to preliminary experiments, the maximum iteration value is set as 10.

Once the reference surface S_r is defined, using either random selection or average surface creation, all surfaces S_i are registered rigidly to S_r . This global surface alignment between each surface S_i and reference surface S_r consists of rigid transformations.

Algorithm 1 (:). Average template

Input: $N = (N_A + N_B)$ triangulated surfaces S_i , and an initial reference surface (e.g., $S_r = S_1$)

repeat

1. Align all surfaces $S_i - S_r$ using rigid ICP
2. Find point-to-point mapping $\varphi_{i \rightarrow r}$ between each surface S_i and S_r
3. Update reference $S_r = \frac{1}{N} \sum_{i=1}^N (S_i \circ \varphi_{i \rightarrow r})$

until No more changes in S_r obtained or the maximum number of iterations is reached

Output: average template S_r

2.3.3. Groupwise shape analysis

Given two sets of aligned surface meshes $\{S_i^A\}_{i=1,\dots,N_A}$ and $\{S_i^B\}_{i=1,\dots,N_B}$, the point-to-point correspondences between each surface and the reference S_r is established using the spectral matching approach presented in Section 2.2. Then, the mesh vertices of all surfaces are rearranged to have the consistent ordering across the populations. The resulting reconstructed meshes are denoted as $\{S_i^A\}_{i=1,\dots,N_A}$ and $\{S_i^B\}_{i=1,\dots,N_B}$ for the two study groups.

The mean shapes S_μ^A and S_μ^B are computed by averaging the 3D coordinates of corresponding surface points across each group:

$$S_\mu^A = \frac{1}{N_A} \sum_{i=1}^{N_A} S_i^A, \quad \text{where } S_i^A = S_i^A \circ \varphi_{i \rightarrow r} \quad (6)$$

$$S_\mu^B = \frac{1}{N_B} \sum_{i=1}^{N_B} S_i^B, \quad \text{where } S_i^B = S_i^B \circ \varphi_{i \rightarrow r}$$

with $\varphi_{i \rightarrow r}$ the mapping from instance S_i to reference S_r . The local shape differences between groups can be detected by computing a distance map between two mean shapes S_μ^A and S_μ^B . This will provide the visual assessment of the location and magnitude of the differences between groups A and B. We use a standard, non-parametric Hotelling T^2 test [22] to check for significant group differences, with the null hypothesis that the two groups are drawn from the same distribution. The resulting raw p -values are then corrected for multiple comparisons by the false discovery rate (FDR) approach [33].

The performance of the method is further evaluated by significant vertex ratio metric, which is estimated by computing the ratio of the number of statistically significant vertices over the total number of surface points of the mesh [8].

The global shape difference between two groups is computed by averaging the distances between the mean surfaces, which results in the mean absolute distance (MAD). The mean absolute distance (MAD) attempts to estimate the average distance of surfaces S_μ^A and S_μ^B by projecting the vertices of the first surface on the second one [34]. Moreover, the Hausdorff distance [35] is calculated as the maximum distance between two mean shapes. This measure has high sensitivity to positional differences between two surfaces. In addition, Dice coefficient [36] is reported as a global measure, which indicates the amount of volume overlap between two mean shapes. The 3D Dice coefficient metric between two surfaces S_μ^A and S_μ^B is defined as the intersection divided by the mean volume of the two surfaces:

$$\text{Dice}(S_\mu^A, S_\mu^B) = \frac{|S_\mu^A \cap S_\mu^B|}{\frac{|S_\mu^A| + |S_\mu^B|}{2}} \quad (7)$$

For identical surfaces, Dice coefficient achieves its maximum value of 1, with decreasing values indicating less volume overlap.

3. Results

In this section, we will first evaluate the accuracy of the proposed method under different configurations and initial conditions. Then, the performance of the shape analysis approach will be evaluated on different subcortical structures (hippocampus and putamen) from various neuroimaging datasets. Finally, the proposed methodology is applied on a clinical application of Alzheimer's disease. This disorder is one of the most widespread diseases in the elderly population, which gradually damages the brain regions including hippocampus. Here, we use our framework to quantify subcortical morphological changes in patients diagnosed with Alzheimer's and mild cognitive impairment.

3.1. Datasets

For accuracy and performance evaluation, two separate neuroimaging datasets were used. The first dataset includes putamen structures (PT dataset) and the second dataset consists of hippocampus shapes (HPC dataset).

The PT dataset includes 36 T1-weighted MR images from healthy controls. Subjects were recruited from a pediatric brain study (mean age 11, 24 males). Acquisition was performed on a 3.0 T MRI system (Philips Acheiva or Siemens) with an in-plane resolution of 256×256 , slice thickness of 1 mm, and voxel size ranging between $1 \times 0.9677 \times 0.9677 \text{ mm}^3$ and $1 \times 1 \times 1 \text{ mm}^3$. In this dataset, the left putamen was segmented using FSL-FIRST automatic segmentation toolbox [37] and corrected by an experienced user. The segmented putamen structures in the dataset have been randomly separated into two groups to create two different groups A and B.

The HPC dataset contains 42 hippocampus shapes obtained from schizophrenic patients and matched healthy controls (mean age 32, all male gender) [2]. The hippocampi were manually segmented from IR-Prepped SPGR (Inversion Recovery-Prepared Spoiled Gradient Echo) data segmented originally at $0.9375 \times 0.9375 \times 1.55 \text{ mm}^3$ resolution. In the HPC dataset, all cases were randomized and group association was performed to create two different groups (group A and group B). The SPHARM-PDM toolbox developers prepared this dataset at the UNC Neuro Image Analysis Laboratory (see www.nitrc.org/projects/spharm-pdm), as part of the public UNC Shape Tool distribution for shape analysis studies. Since the HPC dataset was provided by the SPHARM-PDM software developers, it is relevant to assess the performance of the proposed shape analysis method compared to the SPHARM-PDM approach on this dataset.

In order to validate the proposed framework on clinical applications, a popular brain imaging dataset in Alzheimer's disease, namely the Alzheimer's disease Neuroimaging Initiative (ADNI) was used. Capturing the structural morphometry have always been a target of interest in Alzheimer's disease studies, since evaluating the structural changes could provide a clue for early detection of the pathology. The ADNI database (adni.loni.usc.edu) was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org. The database of ADNI consists of cross-sectional and longitudinal data including 1.5 or 3.0 T structural MR images. The detailed description of the MRI protocol of ADNI is provided in [38]. For this study, a subset of screening 1.5 T MR images is used including 47 normal controls (NC), 47 AD patients, and 47 individuals with MCI. The three groups are matched approximately by age and gender (NC with mean age of 76.7 ± 5.4 , 23 males; AD

with mean age of 77.4 ± 7.2 , 21 males; and MCI with mean age of 75.0 ± 6.9 , 28 males). For the purpose of reproducibility, the list of the study participants' identifiers is provided in the Supplementary material. Contributors of ADNI performed additional post-processing steps on MR images to correct certain image artifacts and to enhance standardization across sites and platforms. The post-processing steps include gradient non-linearity correction [39], intensity inhomogeneity correction [38], bias field correction [40], and phantom-based geometrical scaling to remove calibration errors [41]. Here, we use these processed images. Left and right hippocampi were segmented using FSL-FIRST automatic segmentation software package [37] and visual inspection was performed on the output binary masks to ensure the quality of the automatic segmentation.

3.2. Validation methodology

The proposed framework was first tested under different configurations using both PT and HPC datasets. The minimum number of spectral coordinates, which is required for having a stable output, was determined. Then, the impact of curvature features and different reference surfaces on the accuracy of the proposed framework was tested. In each case, mean absolute distance (MAD) measure, along with Hausdorff distance [35] and Dice coefficient [36] was computed to assess the group differences between mean shape A and B. To evaluate the accuracy of the proposed framework, the performance of the spectral approach under different configurations was compared with a state-of-the-art method used for group-wise analysis of anatomical shapes, namely SPHARM-PDM [14]. The maximum spherical harmonics degree of the SPHARM-PDM method is set to 15 for putamen structures and 12 for hippocampus shapes.

Once the configuration of the proposed framework was completed, the proposed framework was validated on PT and HPC subcortical datasets. In these experiments, both local and global group differences were analyzed and compared to SPHARM-PDM. Finally, the proposed methodology was applied for quantifying hippocampal morphological variations in Alzheimer's disease (AD). Hippocampal morphometry have been increasingly used in the AD research in the perspective of early detection of the pathology and future treatments. Here, the proposed framework was applied for detecting the morphological changes in individuals with Alzheimer's disease, in order to verify the accuracy of the spectral based shape analysis approach in the statistical studies of this clinical application.

3.3. Framework configurations

3.3.1. Number of spectral coordinates

The number of eigenvectors used to model surfaces in the spectral domain has an impact on the accuracy of the established surface correspondences in the spectral matching process. In this experiment, we examine the effect of the number of eigenvectors used as spectral coordinates on the groupwise shape analysis output. The eigendecomposition process is performed twice in the spectral matching approach: once for generating the initial map c , and once for establishing the final map φ (see Section 2.2). As mentioned in Section 2.2, the first k eigenvectors associated with the smaller non-zero eigenvalues are extracted as spectral coordinates. The rationale for this choice is that the eigenvectors associated with the lower non-zero eigenvalues represent coarse intrinsic geometric properties of a shape (i.e., depicting the global shape model). However, the question remains how many eigenvectors is enough to represent the surfaces in the initial step, as well as the association graph G_a in the final process.

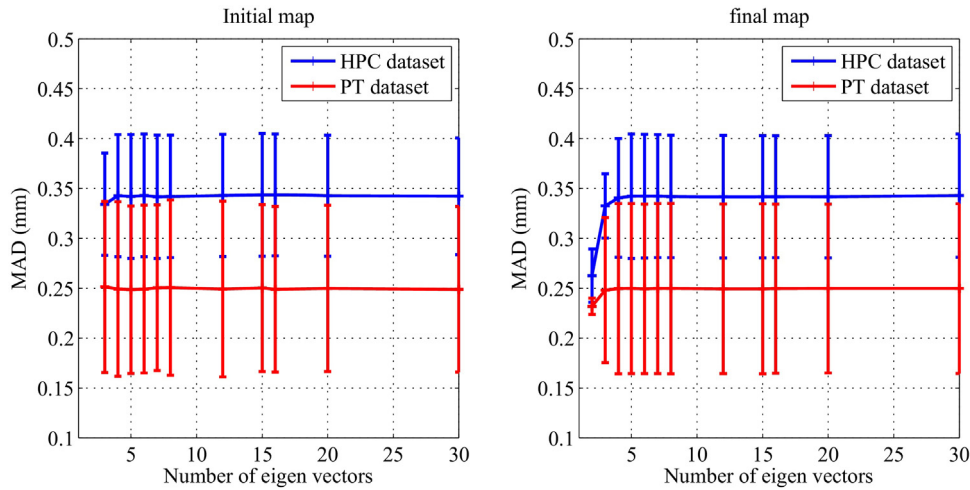


Fig. 3. Performance of the shape analysis framework under varying number of eigenvectors used as spectral coordinates. Left: average mean absolute distance obtained for different number of eigenvectors in creating the initial map. Right: average mean absolute distance computed for different number of eigenvectors in creating the final map. The framework result becomes stable with three eigenvectors for PT dataset and four eigenvectors for HPC dataset.

Here, we apply our shape analysis framework on both PT and HPC datasets with varying number of eigenvectors, and evaluate the framework output by measuring the mean absolute distance between two groups A and B. In this experiment, 10 separate study groups A and B are created from each dataset, and the average mean absolute distance is reported under varying number of eigenvectors (Fig. 3). This experiment demonstrates that the framework output becomes stable at $k=3$ for PT dataset and $k=4$ for HPC dataset, and gains no further significant changes beyond these values. The difference in the sufficient numbers of eigenvectors between two datasets is due to the higher complexity of hippocampus structures compared to putamen surfaces. Thus, more eigenvectors are required to model hippocampus surfaces. Therefore, we consider $k_{\text{initial}} = k_{\text{final}} = 3$ for PT, and $k_{\text{initial}} = k_{\text{final}} = 4$ for HPC dataset as the number of eigenvectors for the remainder of the experiments.

3.3.2. Incorporating curvature feature

In this section, the impact of integrating additional features on the performance of the spectral groupwise approach is first evaluated to understand the value of surface curvature in the matching process. Then, the robustness of the subcortical surface matching process, both with and without additional features, is assessed by applying synthetic deformations to the surfaces.

As mentioned previously in Section 2.2, one can include additional information in the spectral matching process, by establishing a meaningful node weight for each vertex. Here, we compute the exponential of the mean curvature at each vertex (Eq. (3)), and use the combination of curvature and spectral coordinates for surface matching.

In order to evaluate the performance improvement after incorporating additional information, 10 separate subsets are created

from 36 putamen structures of PT dataset. Each subset includes 10 different surfaces, i.e., five surfaces are considered as the member of the study group A and the rest are marked as group B. The proposed shape analysis framework is then applied on each subset with and without curvature features. In addition, the accuracy of the two implementation of the shape analysis framework is compared with the state-of-the-art groupwise shape analysis method, SPHARM-PDM [14]. Table 1 provides the average Dice coefficient, Hausdorff distance, and mean absolute distance in 10 separate subsets. Comparing our framework to the SPHARM-PDM approach revealed that the groupwise shape analysis is more accurate using curvature features ($\text{MAD} = 0.25 \pm 0.09$ mm rather than using the spectral coordinate alone ($\text{MAD} = 0.37 \pm 0.14$ mm)).

We then tested the proposed shape analysis pipeline on the HPC dataset with and without curvature features, in order to investigate the influence of incorporating additional features. The extracted group differences show that the groupwise shape analysis integrated with curvature features (Dice=0.93, Hausd. Dist.=1.16 mm, $\text{MAD} = 0.25 \pm 0.23$ mm) is more accurate rather than using the spectral coordinate alone (Dice=0.70, Hausd. Dist.=2.56 mm, $\text{MAD} = 0.95 \pm 0.57$ mm). Hence, we can conclude that adding mean curvature has a significant positive effect on the framework's precision, when matching the hippocampus structures.

Now, we evaluate the accuracy of the subcortical surface matching process with and without additional features by testing for robustness to synthetic deformations on the PT dataset. To achieve this, we synthetically deform putamen surfaces, and match each putamen structure to its deformed instance. Since the vertex indexing remains the same, we are able to establish a ground truth for correspondence maps. Thus, we measure the average distance

Table 1
Shape differences between mean shape A and mean shape B with and without incorporating curvature features. The Spectral method with out additional information shows significantly different results compared to the SPHARM-PDM.

	Spectral method		SPHARM-PDM	
	No additional features	With additional features		
Dice coefficient	0.93 ± 0.02	0.94 ± 0.02	$p = 0.02$	0.95 ± 0.02
Hausd. Dist. (mm)	2.08 ± 0.90	1.15 ± 0.47	$p < 0.01$	1.04 ± 0.34
MAD (mm)	0.37 ± 0.14	0.25 ± 0.09	$p = 0.03$	0.25 ± 0.09

The average Dice coefficient, Hausdorff distance, and MAD are reported on 10 subsets of PT dataset. p -values of t -test between the spectral method and SPHARM-PDM framework are also shown. Spectral method without additional features produces significantly different results compared to SPHARM-PDM, while spectral framework integrated with curvature features generates similar results to the spherical method.

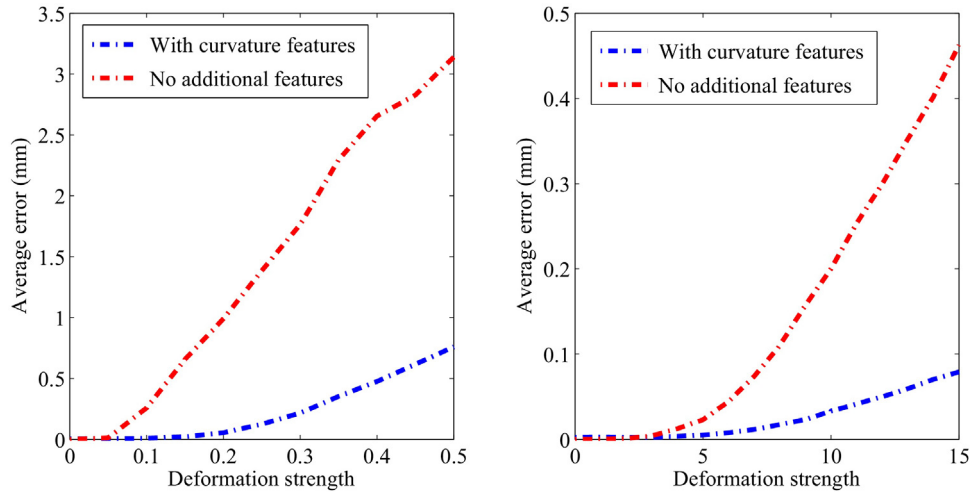


Fig. 4. Testing the robustness of the spectral matching approach with respect to added deformations. Left: deformation in the z direction, simulating surface compression; Right: radial distortion. The average error at each deformation is shown for two separate implementations of the spectral approach, one with and another without incorporating additional features (curvature). This shows that using additional features improves the precision of the surface matching.

Table 2

Shape differences between mean shape A and mean shape B.

	Different initial references ($n = 36$)	Average template
Dice coefficient	0.97 ± 0.001	0.97
Hausd. Dist. (mm)	0.83 ± 0.12	0.83
MAD (mm)	0.14 ± 0.005	0.14 ± 0.12

The average Dice coefficient, Hausdorff distance, and MAD are reported for 36 separate runs by choosing 36 different initial references. In addition, Dice coefficient, Hausdorff distance, and MAD are listed for a shape analysis test using average template as a reference surface.

between the locations of corresponding points found with spectral matching and the real locations of corresponding points defined by the ground truth.

Here, each vertex $x = (X, Y, Z)$ of the surface S_i is deformed by applying the transformation $Z' = (1 + \alpha)Z$. This deformation provides surface compression without changing the topology of the mesh. All 36 putamen surfaces are deformed by varying α in the range of $[0, 0.4]$. Fig. 4 left shows the accuracy of the surface matching with and without additional curvature features. At maximal deformation ($\alpha = 0.4$), the spectral matching method achieves an average error of 0.47 ± 0.10 mm with curvature features, and 2.65 ± 0.87 mm without extra information. In the next experiment, a radial distortion of $X' = X + \beta(X^2 + Y^2)^2 / \max((X^2 + Y^2)^2)$ is applied to each vertex $x = (X, Y, Z)$ of all 36 putamen surfaces. This deformation simulates a drastic change in a surface shape, while preserving the mesh topology. The controlling parameter β is varied in the range of $[0, 15]$ and the surface matching accuracy is assessed as shown in Fig. 4 right. These experiments show the performance of the point-to-point spectral matching under synthetic deformation and indicate that using additional curvature features enhances the surface matching accuracy.

3.3.3. Reference surface selection

In this section, we verify that the proposed shape analysis approach is unbiased to the choice of the reference surface within a study group. To this end, an arbitrary initial reference is selected as one of the surfaces in the PT dataset. Then, all surfaces are matched to this reference surface and shape analysis is conducted to estimate group differences. We performed 36 separate tests by iteratively selecting a new putamen surface as reference shape and computing the average mean distance in multiple runs (see Table 2). The average mean difference for the whole dataset

is 0.14 ± 0.005 mm. In order to further examine the impact of reference surface on the shape analysis accuracy, in the next experiment, an average template is created in an iterative process, as described in Section 4.2. This template is applied as a reference surface in a shape analysis study on PT dataset, yielding a MAD of 0.14 ± 0.12 mm (see Table 2). These results suggest that the method produces near identical shape differences even by varying the choice of the initial reference. Therefore, if reducing the computational time is important in a study, it is possible to choose one of the input surfaces as a reference surface, instead of creating an average template.

3.4. Accuracy on subcortical shapes

In this section, we validate the performance of our proposed methodology on localizing structural morphologies across populations. The spectral based shape analysis framework is applied on both PT and HPC datasets to detect shape differences between groups A and B in each dataset. The performance of our spectral-based framework is compared with the state-of-the-art groupwise shape analysis method, SPHARM-PDM [14].

For putamen shapes, the local group differences produced by both methods are illustrated in Fig. 5, which shows that both methods capture similar shape differences. However, no significant difference is found after applying statistical analysis (see vertex ratio in Table 3). Since PT dataset includes healthy control subjects, detecting no significant differences between two groups was expected. The Dice coefficients between A and B (Dice = 0.97) and the mean absolute distances (MAD = 0.14 ± 0.11 mm in spectral framework, and MAD = 0.14 ± 0.12 mm in SPHARM-PDM) were the same for both methods (see Table 3). Also, the Hausdorff distances were 0.87 mm and 0.74 mm for spectral-based method and SPHARM-PDM, respectively. In order to further verify the accuracy of the proposed framework in comparing two groups of healthy subjects, we perform another experiment, in which the proposed pipeline is applied on different randomly separated datasets (i.e., five study group pairs, where each group contain 18 random putamen structures). The significant vertex ratio is computed on all five separate tests, which results in an average vertex ratio of zero. This experiment shows that the proposed method is working well by not detecting any significant shape variations between two healthy groups.

Table 3

Evaluation of shape differences between mean shape A and mean shape B obtained with spectral method and SPHARM-PDM for the two datasets (HPC and PT).

	Hippocampus (HPC)		Putamen (PT)	
	Spectral method	SPHARM-PDM	Spectral method	SPHARM-PDM
Dice coefficient	0.93	0.92	0.97	0.97
Hausd. Dist. (mm)	1.16	1.12	0.87	0.74
MAD (mm)	0.25 ± 0.23	0.28 ± 0.23	0.14 ± 0.11	0.14 ± 0.12
Vertex ratio	0.16	0.23	0	0

Results are compared using the Dice coefficient, the Hausdorff distance (Hausd. Dist.), the mean absolute distance (MAD), and significant vertex ratio.

For hippocampus shapes, the location and magnitude of group differences is consistent between two methods (see Fig. 6). The p -values maps show a strong significance in tail region for both methods, while the head region does not show the same significant difference. On one hand, the subjects in HPC dataset are randomly divided into two groups A and B, thus we are not able to compare the pattern of hippocampal deformation with the one reported in previous studies on schizophrenia. On the other hand, inconsistencies between groupwise shape analysis approaches have been reported in the literature [8,42]. This might be due to the use of different strategies in extracting surface meshes and establishing surface correspondences between methods. The global shape differences show considerable agreement between the spectral framework (Dice=0.93, Hausd. Dist. = 1.16 mm, and MAD = 0.25 ± 0.23 mm) and SPHARM-PDM (Dice = 0.92, Hausd. Dist. = 1.12 mm, and MAD = 0.28 ± 0.23 mm) (see Table 3). In general the two methods are consistent in global group difference measures and distance maps, although they show a small difference in p -value maps.

The computational time of both methods was measured on a 3.2GHz Core i5 computer with 32GB of RAM. The computational cost of spectral method (PT: 5 min and HPC: 54 min) is significantly lower than SPHARM-PDM (PT: 14 min and HPC: 242 min (or 4 h)), especially for more complex shapes, like the hippocampus. This significant speed advantage was achieved without using parallel programming. These results suggest that the proposed spectral framework not only produces similar groupwise shape differences as the SPHARM-PDM, but it also has a clear speed advantage.

3.5. Application to Alzheimer's disease

Alzheimer's disease (AD) [43] is the most common form of dementia in the elderly population. There is evidence that the neuropathological changes in AD damage the hippocampus structure, which is a brain region crucial to various cognitive functions [43,44]. Mild cognitive impairment (MCI) [45] is considered as a transition state between normal aging and dementia. The cognitive deficits in MCI patients are not as severe as those seen in people with AD. However, MCI has a 10-fold risk of a transition to early Alzheimer's disease [45]. According to neuroanatomical studies, MCI patients have shown hippocampal deformation compared with healthy elderly people [43,46]. Here, our groupwise shape analysis is applied on the ADNI dataset including 47 AD patients, 47 individuals with MCI, and 47 matched normal controls (NC). The shape variation of left and right hippocampi was studied between AD and NC, as well as MCI and NC. To correct for head size differences across populations, each hippocampi surface was scaled by the individual's total intracranial volume. The total intracranial volumes were collected from ADNI database (adni.loni.usc.edu).

Fig. 7(a) and (b) shows shape variations between AD and normal controls for the left and right hippocampi produced by the spectral and SPHARM-PDM methods. In both methods,

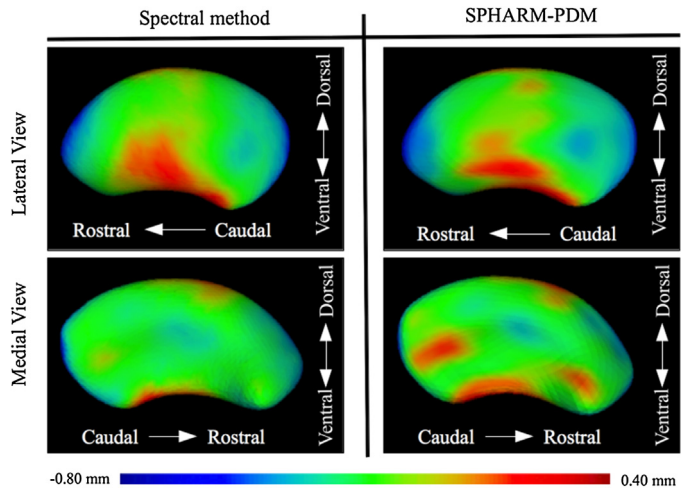


Fig. 5. Putamen shape analysis. Distance maps were computed using the proposed spectral method on the left and SPHARM-PDM on the right. The resulting distance maps are shown in lateral and medial views. Statistical analysis showed that none of the captured group differences were significant.

hippocampal surfaces of AD patients show inward local deformity mainly in lateral zones (CA1) and inferior-medial zones (subiculum), which is consistent with findings from previous studies [44,46,47]. Table 4 shows the global shape differences between the spectral framework (Left hippocampi: MAD = 0.58 ± 0.26 mm and Right hippocampi: MAD = 0.55 ± 0.29 mm) and SPHARM-PDM (Left hippocampi: MAD = 0.52 ± 0.28 mm and Right hippocampi: MAD = 0.55 ± 0.28 mm). The p -values maps are almost similar between the two methods, however, in general both methods limit the shape changes to CA1 and subiculum. In addition, the ratio of the statistically significant vertices over the total number of surface points shows agreement between both methods.

The result of the same comparison performed between MCI patients and normal controls is shown in Fig. 8 and Table 5. The distance maps produced by both methods are almost similar, while the p -value maps do not agree. Indeed, the spectral matching approach found significant inward variation in CA1 and subiculum regions of left and right hippocampi, however, no significant region was captured by SPHARM-PDM. The amount of vertex ratio is interestingly different between spectral approach (Left hippocampi = 0.17 and Right hippocampi = 0.19) and SPHARM-PDM (Left hippocampi and Right hippocampi = 0). However, since there is no ground truth for the clinical datasets and the previous neuroanatomical studies on these pathologies used different input samples, it is unclear which method can be considered as the gold-standard. However, both methods provided almost the same global regional differences (Dice, Hausd. Dist., and MAD) and similar distance maps. In addition both methods are in line with the previous neuroanatomical studies [46,48], which reported capturing weaker group differences between MCI and normal individuals rather than AD and normal controls.

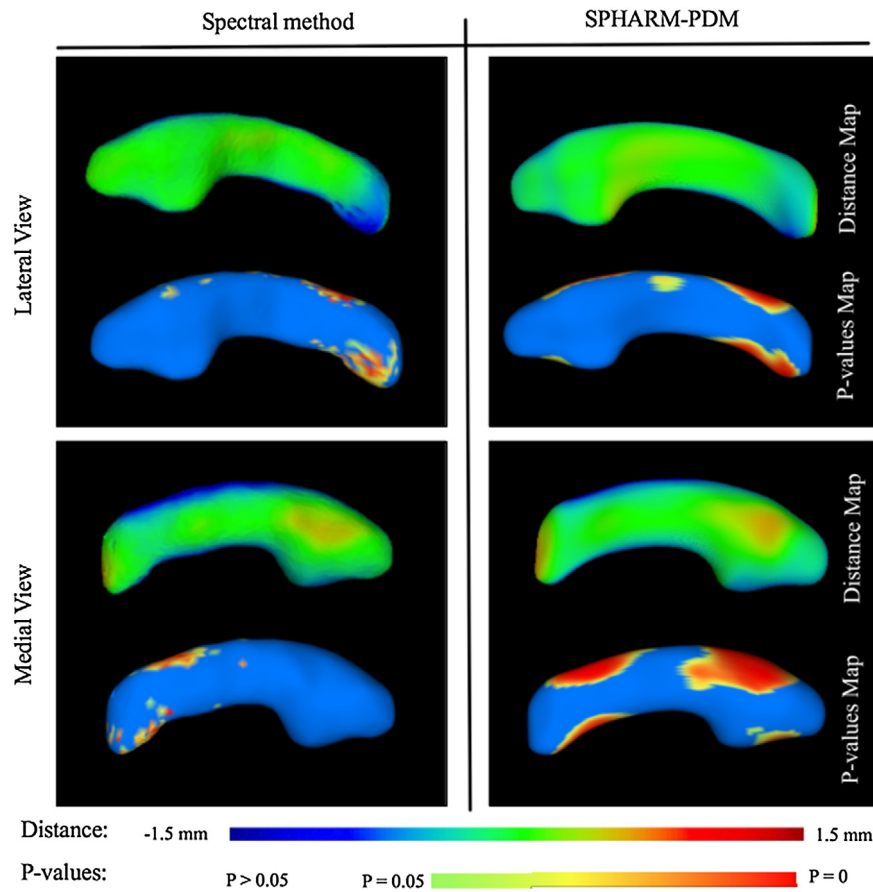


Fig. 6. Hippocampus shape analysis. Distance maps are computed using the spectral method and the SPHARM-PDM. The resulting distance maps are shown in lateral and medial views. The respective FDR-corrected *p*-value maps are shown below distance maps.

Table 4
Evaluation of shape differences between AD and control groups obtained with spectral method and SPHARM-PDM for the left and right hippocampus.

	Left hippocampus		Right hippocampus	
	Spectral method	SPHARM-PDM	Spectral method	SPHARM-PDM
Dice coefficient	0.85	0.86	0.86	0.86
Hausd. Dist. (mm)	1.49	1.50	1.39	1.50
MAD (mm)	0.58 ± 0.26	0.55 ± 0.28	0.52 ± 0.29	0.55 ± 0.28
Vertex ratio	0.45	0.40	0.46	0.45

Results are compared using the Dice coefficient, the Hausdorff distance (Hausd. Dist.), the mean absolute distance (MAD), and significant vertex ratio.

4. Discussion

In this work, a new framework for groupwise shape analysis of subcortical structures, such as the hippocampus or putamen, is proposed in order to detect regional morphological alterations of subcortical structures from patients with neurological conditions, such as schizophrenia and Alzheimer’s. The proposed method extracts the 3D meshes from input binary maps using the

Marching cubes algorithm, which is one of the most popular 3D modeling algorithms in medical visualization. Since this approach was not able to generate high quality smoothed surfaces out of subcortical structures, the mean curvature flow algorithm was employed to smoothen the 3D surface meshes. Mean curvature flow smoothes a surface mesh by moving the surface nodes along the normal direction and achieves the best smoothing result with respect to the geometric information of the surface. In the next step,

Table 5
Evaluation of shape differences between MCI and control groups obtained with spectral method and SPHARM-PDM for the left and right hippocampus.

	Left hippocampus		Right hippocampus	
	Spectral method	SPHARM-PDM	Spectral method	SPHARM-PDM
Dice coefficient	0.95	0.95	0.95	0.95
Hausd. Dist. (mm)	1.1	1.0	1.17	1.64
MAD (mm)	0.20 ± 0.18	0.21 ± 0.21	0.20 ± 0.20	0.24 ± 0.24
Vertex ratio	0.17	0	0.19	0

Results are compared using the Dice coefficient, the Hausdorff distance (Hausd. Dist.), the mean absolute distance (MAD), and significant vertex ratio.

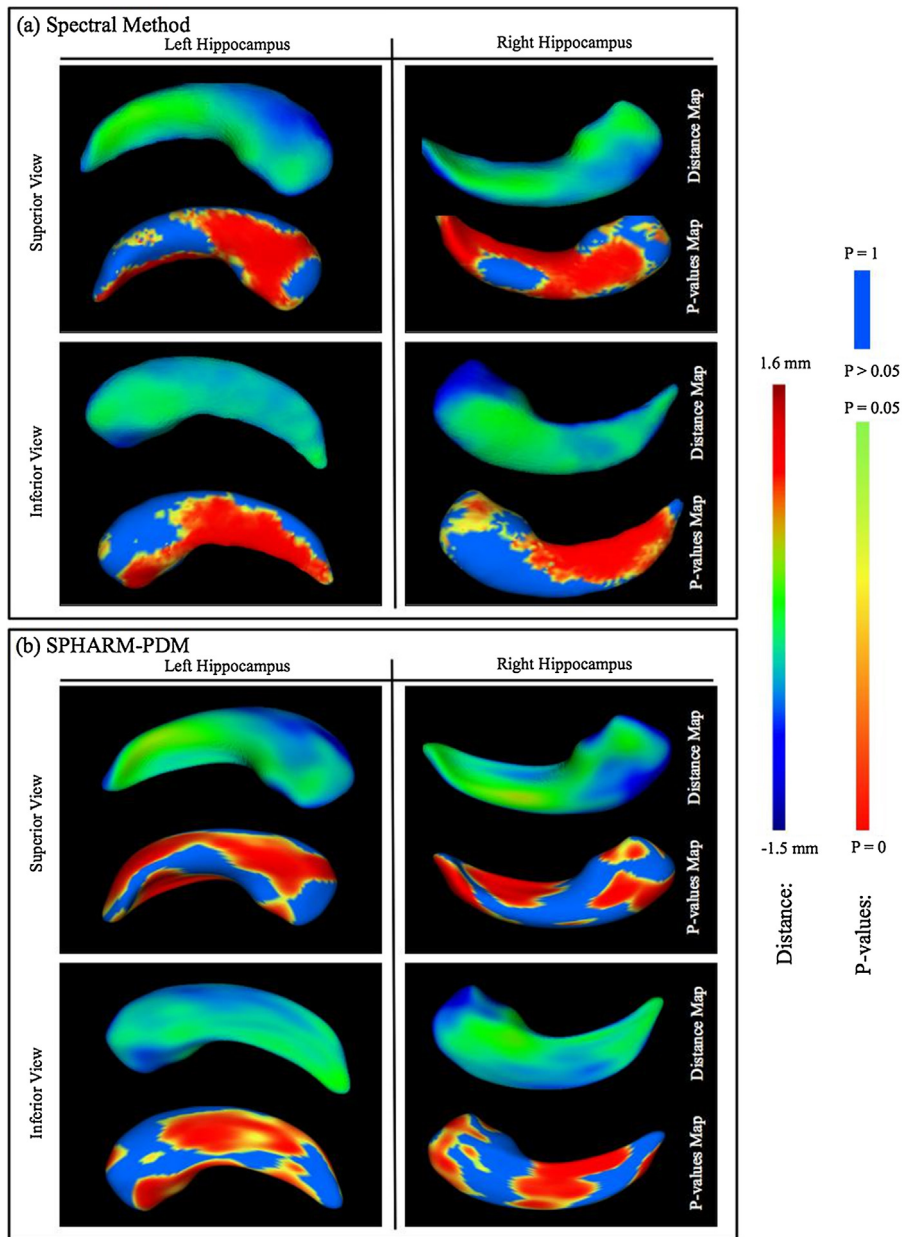


Fig. 7. Left and right hippocampal shape deformations in AD patients compared with normal controls. Distance maps are computed along with the respective FDR-corrected p -values maps using spectral method (a) and SPHARM-PDM (b).

all surfaces are rigidly registered to a reference surface and mean curvature is computed at each node of all surfaces. Once all shapes are globally aligned, the spectral matching approach as proposed in [20,21] is applied in order to match each surface to the reference mesh. The proposed pipeline then creates two mean shapes as the geometric mean of all surfaces in each study group and a standard non-parametric permutation test is applied at every vertex point to detect significant local group differences. The framework terminates by applying various metrics in order to assess local and global shape differences across populations.

As suggested by the authors of the spectral matching approach [20,21], adding extra features, such as texture information, anatomical information, or landmark positions could lead to a better surface correspondence. For instance, in the case of matching cortical surfaces, the technique in [21] proposed to use sulcal depth as additional feature, which provides information about the depth of the cerebral cortex. Therefore, the choice of additional features

and assessing the effect of incorporating them into the matching process depends completely on the type of structures under study. In this work, we included the exponential of the mean curvature features in the node weights of the Laplacian matrix (Eq. (2)), since it intrinsically describes the local shape information of a 3D surface, without being too application specific. Experimental analysis in Section 3.3.2 revealed that embedding this additional feature significantly improves the matching accuracy to a greater extent. For instance, adding mean curvature has a significant positive effect on point correspondence performance, when matching hippocampus shapes. However, this feature is not crucial for analyzing putamen shapes. Hence, one might conclude that incorporating curvature features has more relevance to evaluate complex structures.

In order to verify the performance of the proposed approach, the robustness of the shape analysis framework was evaluated using two reference selection approaches. According to these experiments, the use of an average reference shape gave slightly better

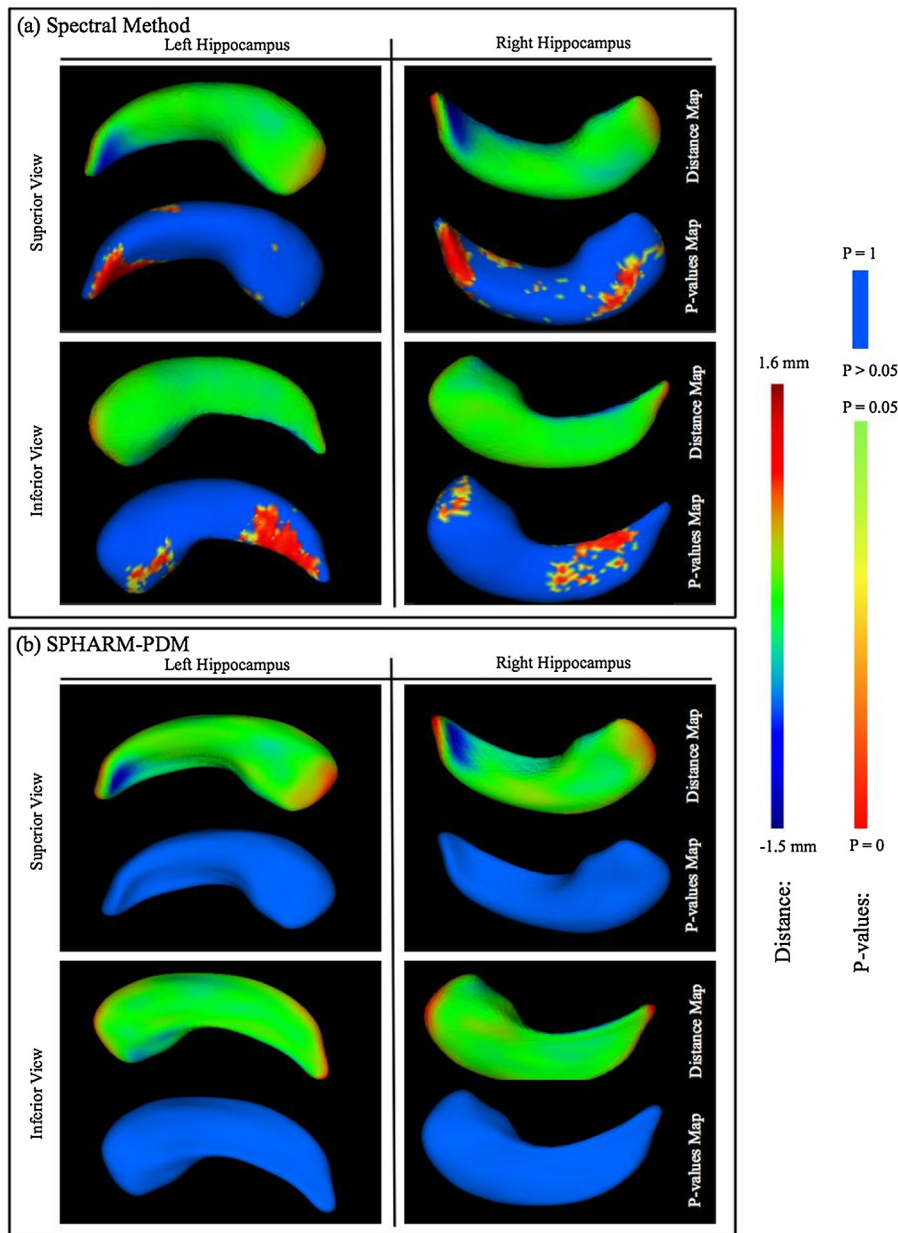


Fig. 8. Left and right hippocampal shape deformations in MCI patients compared with normal controls. Distance maps are computed along with the respective FDR-corrected p -values maps using spectral method (a) and SPHARM-PDM (b).

results than using a single reference shape. However, the difference between average and randomly selected initial shape was not noticeable, indicating that the proposed framework is not significantly sensitive to this initial choice. Therefore, since computing an average reference is time-consuming, it is preferable to choose an arbitrary reference surface from input surfaces if the computational time is a critical issue in a particular study.

A series of experiments were conducted in this study to analyze the performance of the proposed groupwise shape analysis approach. According to these experiments, the proposed group analysis framework and the state-of-the-art method, the SPHARM-PDM have considerable agreement in most of the cases. However, from a computational perspective, the proposed spectral method has a clear speed advantage over the spherical harmonics based method. This speed advantage becomes more noticeable for more complex structures, such as for example analyzing a complex shape

like the hippocampus, with a fourfold decrease in time compared to SPHARM-PDM. According to the experiments, the total running time of SPHARM-PDM could be reduced, by employing lower number of spherical harmonic coefficients. For instance, groupwise shape analysis on HPC dataset took on average 5.5 h, when the structures were represented by 15 spherical coefficients, compared to 4 h when 12 spherical coefficients were used. The problem is that the sufficient number of spherical coefficients must be provided in order to obtain an appropriate representation of a structure. In this paper, the number of spherical harmonic descriptions has been chosen in the range of 12–15, according to the SPHARM-PDM software manual (www.nitrc.org/projects/spharm-pdm) [14].

Aside from the consistency in the group distance maps produced by the two shape analysis methods, spectral and SPHARM-PDM methods, the generated p -value maps did not agree in some experiments. This type of inconsistency, which have been previously

reported in the literature [8,42], mainly comes from the specific surface representation used by each method. A method like SPHARM-PDM generates smooth surfaces based on spherical harmonics, while our approach keeps the original meshes and applies only a slight smoothing process to remove surface irregularities. Severe smoothing can cause a method to miss small surface deformations in a groupwise morphological study. This might be the reason why the SPHARM-PDM does not capture any region of group differences in the comparison between MCI and healthy controls. Another reason for inconsistency in capturing significant groupwise variations is the choice of surface correspondences across populations. This influence seems to be higher in studies with lower number of samples, especially when there is large shape variability due to the high age range and gender differences. Therefore, using relatively higher number of samples in the input populations, could reduce the inconsistency in groupwise morphological frameworks. Since there is no ground truth for the clinical datasets, it is not clear which one of the two methods (spectral group analysis or SPHARM-PDM) are more accurate. Nonetheless, one should keep in mind that both methods produce similar distance maps and close global group differences, which shows that the two methods are consistent in general. In addition, both methods did not detect any significant shape variations when the two groups of controls were used as inputs. This indicates that the two methods are working well.

As an application, the proposed framework was applied on the ADNI dataset including AD, MCI, and normal controls. In previous neuroimaging studies, it has been reported that subjects with AD and MCI could be discriminated from healthy subjects by examining the pattern of hippocampal surface variations [43–48]. Hippocampal atrophy begins in lateral zones (CA1) and spreads to inferior-medial zones (subiculum) [48]. The superior zones including CA2–4 and Gyrus Dentatus are preserved [48]. We validated our scheme through a statistical shape analysis on hippocampal surface deformity in ADNI dataset, by demonstrating a consistency with previous clinical findings. These results suggest that the proposed spectral-based shape analysis framework could allow for quantitative assessment of variations in subcortical structures, associated with a neurological disorder, which leads to the better understanding of a pathology.

An important issue in any surface-based shape analysis approach is the reliance on the segmentation accuracy. This step has a crucial impact on a shape analysis pipeline, since its quality could affect the accuracy of the detected shape variations. The segmentation process becomes more challenging, when we attempt to capture slight group differences across a population, since small shape variations have greater chances of being lost in a segmentation process. In this paper, we used an openly-available automatic segmentation tool (FSL-FIRST [37]) for segmenting structures in PT and ADNI datasets. However, in order to ensure about the quality of the structural delineations, we asked a trained expert user to correct the segmentations. These corrected labels were then used as an input in the proposed shape analysis framework. As future work, it would be helpful to incorporate a more accurate segmentation process in the pipeline, which is able to correctly define the boundary of each structure and extract smoothed meshes directly from MR images.

In summary, the proposed groupwise morphological framework with its methods for surface extraction, smoothing, matching, and statistical analysis is novel, and produced comparable results to the state-of-the-art algorithm, while being faster. This paper was the first work that integrated curvature-based spectral matching in to a groupwise shape analysis pipeline. Experimental analysis on real clinical dataset even showed that the extracted group differences were in line with the findings in other clinical studies. This paper proposed an alternative to the current state-of-the-art groupwise

shape analysis framework and provides a new pipeline for future clinical studies.

5. Conclusion

In this paper, a new groupwise shape analysis framework is proposed for subcortical surfaces based on spectral matching theory. This spectral matching process is able to establish reliable correspondences between different surface meshes and may help in the investigation of groupwise structural differences between two study groups. From a clinical perspective, the proposed method can contribute to the diagnosis of subcortical shape variations in different pathologies, like Alzheimer's disease.

One important future direction of the proposed framework is to extend the available approach for diagnosis of different pathologies, which cause morphological variations in subcortical structures. This could be done by modeling the existing shape variations in patients using a classification approach, and try to distinguish the pathological subjects from normal individuals. Such a framework would become a complementary technique to other available diagnosis approaches, which are currently used for identifying different diseases. However, this kind of diagnosis framework requires incorporating larger datasets to achieve high and stable classification accuracy.

Acknowledgments

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

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.compmedimag.2016.03.001>.

References

- [1] Nitzken MJ, Casanova MF, Gimel MDG, Inanc T, Zurada JM, Fellow L, et al. Shape analysis of human brain: a brief survey. *IEEE J Biomed Health Inform* 2014;18:1337–54, <http://dx.doi.org/10.1109/JBHI.2014.2298139>.
- [2] Styner M, Lieberman JA, Pantazis D, Gerig G. Boundary and medial shape analysis of the hippocampus in schizophrenia. *Med Image Anal* 2004;8:197–203, <http://dx.doi.org/10.1016/j.media.2004.06.004>.
- [3] Shi F, Liu B, Zhou Y, Yu C, Jiang T. Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: meta-analyses of MRI studies. *Hippocampus* 2009;19:1055–64, <http://dx.doi.org/10.1002/hipo.20573>.
- [4] Hastings RS, Parsey RV, Oquendo MA, Arango V, Mann JJ. Volumetric analysis of the prefrontal cortex, amygdala, and hippocampus in major depression. *Neuropsychopharmacology* 2004;29:952–9, <http://dx.doi.org/10.1038/sj.npp.1300371>.
- [5] Allasonnière S. Toward a coherent statistical framework for dense deformable template estimation. *J R Stat Soc* 2007;69:3–29.
- [6] Marsland S, Twining CJ. Constructing diffeomorphic representations of non-rigid registrations of medical images. *IEEE Trans Med Imaging* 2004;23:1006–20.
- [7] Rueckert D, Frangi AF, Member A, Schnabel JA. Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. *IEEE Trans Med Imaging* 2003;22:1014–25.
- [8] Gao Y, Riklin-Raviv T, Bouix S. Shape analysis, a field in need of careful validation. *Hum Brain Mapp* 2014;35:4965–78, <http://dx.doi.org/10.1002/hbm.22525>.
- [9] Joshi S, Pizer S, Member S, Fletcher PT, Yushkevich P, Thall A, et al. Multiscale deformable model segmentation and statistical shape analysis using medial descriptions. *IEEE Trans Med Imaging* 2002;21:538–50.
- [10] Bouix S, Pruessner JC, Louis Collins D, Siddiqi K. Hippocampal shape analysis using medial surfaces. *Neuroimage* 2005;25:1077–89, <http://dx.doi.org/10.1016/j.neuroimage.2004.12.051>.
- [11] Styner M, Gerig G, Lieberman J, Jones D, Weinberger D. Statistical shape analysis of neuroanatomical structures based on medial models. *Med Image Anal* 2003;7:207–20.
- [12] Terriberry TB, Damon JN, Pizer SM, Joshi SC, Gerig G. Population-based fitting of medial shape models with correspondence optimization. *Inf Process Med Imaging* 2007;20:700–12.
- [13] Styner M, Lieberman JA, McClure RK, Weinberger DR, Jones DW, Gerig G. Morphometric analysis of lateral ventricles in schizophrenia and healthy controls regarding genetic and disease-specific factors. *Proc Natl Acad Sci USA* 2005;102:4872–7.
- [14] Styner M, Oguz I, Xu S. Framework for the statistical shape analysis of brain structures using SPHARM-PDM. *Insight J* 2006;242–50.
- [15] Ong D, Walterfang M, Malhi GS, Styner M, Velakoulis D, Christos P. Size and shape of the caudate nucleus in individuals with bipolar affective disorder. *Aust N Z J Psychiatr* 2012;46:340–51, <http://dx.doi.org/10.1016/j.biotechadv.2011.08.021.Secreted>.
- [16] Paniagua B, Bompard L, Cates J, Whitaker R, Datar M, Vachet C, et al. Combined SPHARM-PDM and entropy-based particle systems shape analysis framework. *Biomed Opt Imaging* 2012;1:1–10, <http://dx.doi.org/10.1117/12.911228.Combined> [SPIE Conf. 8317 Prog.].
- [17] Cates J, Fletcher PT, Styner M, Hazlett HC. Particle-based shape analysis of multi-object complexes. *Med Image Comput Assist Interv* 2009;11:477–85.
- [18] Chung MK, Worsley KJ, Nacewicz BM, Dalton KM, Davidson RJ. General multivariate linear modeling of surface shapes using SurfStat. *Neuroimage* 2010;53:491–505, <http://dx.doi.org/10.1016/j.neuroimage.2010.06.032>.
- [19] Shen L, Farid H, McPeake MA. Modeling three-dimensional morphological structures using spherical harmonics. *Evolution (N Y)* 2009;63:1003–16, <http://dx.doi.org/10.1111/j.1558-5646.2008.00557.x.MODELING>.
- [20] Lombaert H, Grady L, Polimeni JR, Chieriet F. FOCUSR: feature oriented correspondence using spectral regularization—a method for precise surface matching. *IEEE Trans Pattern Anal Mach Intell* 2013;35:2143–60, doi:15A24E1C-FA85-48DC-A556-C6526C479AB2.
- [21] Lombaert H, Sporring J, Siddiqi K. Diffeomorphic spectral matching of cortical surfaces. In: *Information processing in medical imaging*. Berlin, Heidelberg: Springer; 2013. p. 376–89.
- [22] Pantazis D, Leahy RM, Nichols TE, Styner M. Statistical surface-based morphometry using a non-parametric approach. *Int Symp Biomed Imaging* 2004;2:1283–6.
- [23] Styner M, Oguz I, Xu S, Brechb C, Pantazis D, Gerig G. Statistical shape analysis of brain structures using SPHARM-PDM. *Med Image Comput Assist Interv* 2006;1:1–7 [OpenSource Work].
- [24] Shakeri M, Lombaert H, Lippé S, Kadoury S. Groupwise shape analysis of the hippocampus using spectral matching. *Proc SPIE Med Imaging* 2014;9034, <http://dx.doi.org/10.1117/12.2043615>.
- [25] Lorensen WE, Cline HE. Marching cubes: a high resolution 3D surface construction algorithm. *Proceedings of the 14th annual conference on computer, SIGGRAPH* 1987. *Graphics and interactive techniques*, 1987. p. 163–169.
- [26] Grady L, Polimeni J. *Discrete calculus: applied analysis on graphs for computational science*. 2010th ed. London: Springer-Verlag; 2010.
- [27] Lombaert H, Grady L, Polimeni JR, Chieriet F. Fast brain matching with spectral correspondence. *Inf Process Med Imaging* 2011;22:660–73.
- [28] Myronenko A, Song X, Miguel AC. Non-rigid point set registration: coherent point drift. *IEEE Trans Pattern Anal Mach Intell* 2009;32:2262–75.
- [29] Cohen-Steiner D, Morvan J-M. Restricted delaunay triangulations and normal cycle. In: *Proceedings of the nineteenth international conference on computational geometry, SCG '03*. New York, USA: ACM Press; 2003. p. 312–21, <http://dx.doi.org/10.1145/777837.777839>.
- [30] Desbrun M, Meyer M, Schr P, Barr AH. Implicit fairing of irregular meshes using diffusion and curvature flow. *SIGGRAPH*, 1999, pp. 317–324.
- [31] Guimond A, Meunier J, Thirion J-P. Average brain models: a convergence study. *Comput Vis Image Underst* 2000;77:192–210.
- [32] Besl JP. A method for registration of 3D shapes. *IEEE Trans Pattern Anal Mach Intell* 1992;14:239–56.
- [33] Styner M, Oguz I, Xu S, Brechb C, Pantazis D, Gerig G. Statistical shape analysis of brain structures using SPHARM-PDM. *Insight J* 2006;1–7 [MICCAI 2006 OpenSource Work].
- [34] Gerig G, Jomier M, Chakos M. Valmet: a new validation tool for assessing and improving 3D object segmentation. *Med Image Comput Assist Interv* 1992;2208:516–23 [MICCAI 2001].
- [35] Aspert N, Santa-Cruz D, Ebrahimi T. MESH: measuring errors between surfaces using the Hausdorff distance. *Proceedings of the IEEE international conference on multimedia and expo*, 2002, pp. 705–708. doi:10.1109/ICME.2002.1035879.
- [36] Dice LR. Measures of the amount of ecologic association between species. *Ecology* 1945;26:297–302.
- [37] Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* 2011;56:907–22, <http://dx.doi.org/10.1016/j.neuroimage.2011.02.046>.
- [38] Jack Jr CR, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J Magn Reson* 2008;27:685–91, doi:10.1002/jmri.21049.The.
- [39] Jovicic J, Czanner S, Greve D, Haley E, Van Der Kouwe A, Gollub R, et al. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *Neuroimage* 2006;30:436–43, <http://dx.doi.org/10.1016/j.neuroimage.2005.09.046>.
- [40] Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 1998;17:87–97, <http://dx.doi.org/10.1109/42.668698>.
- [41] Gunter J, Bernstein M, Borowski B, Felmlee J, Blezek D, Mallozzi R, et al. Validation testing of the MRI calibration phantom for the Alzheimer's disease neuroimaging initiative study. *Proceedings of the 14th scientific meeting of the international society of magnetic resonance in medicine*, 2006.
- [42] Styner M, Xu S, El-sayed M, Gerig G, Science C, Carolina N, et al. Correspondence evaluation in local shape analysis and structural subdivision. 2007 IEEE international symposium on biomedical imaging from nano to macro, 2007, pp. 1192–1195.
- [43] Du AT, Schuff N, Amend D, Laakso MP, Hsu YY, Jagust WJ, et al. Magnetic resonance imaging of the entorhinal cortex and hippocampus in mild cognitive impairment and Alzheimer's disease. *J Neurol Neurosurg Psychiatr* 2001;71:441–7, <http://dx.doi.org/10.1136/jnnp.71.4.441>.
- [44] Cho Y, Seong JK, Shin SY, Jeong Y, Kim JH, Qiu A, et al. A multi-resolution scheme for distortion-minimizing mapping between human subcortical structures based on geodesic construction on Riemannian manifolds. *Neuroimage* 2011;57:1376–92, <http://dx.doi.org/10.1016/j.neuroimage.2011.05.066>.
- [45] Petersen R, Smith G, Waring S, Ivnik R, Tangalos E, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol* 1999;56:303–8.
- [46] Kim J, Valdes-Hernandez MDC, Royle Na, Park J. Hippocampal shape modeling based on a progressive template surface deformation and its verification. *IEEE Trans Med Imaging* 2014;34:1242–61, <http://dx.doi.org/10.1109/TMI.2014.2382581>.
- [47] Wang L, Miller JP, Gado MH, McKeel DW, Rothermich M, Miller MI, et al. Abnormalities of hippocampal surface structure in very mild dementia of the Alzheimer type. *Neuroimage* 2006;30:52–60, <http://dx.doi.org/10.1016/j.neuroimage.2005.09.017>.
- [48] Xie J, Fletcher E, Singh B, Carmichael O. Robust measurement of individual localized changes to the aging hippocampus. *Comput Vis Image Underst* 2013;117:1128–37, <http://dx.doi.org/10.1016/j.cviu.2012.12.007>.