Hippocampal Shape Modeling Based on a Progressive Template Surface Deformation and its Verification

Jaeil Kim, Maria del C. Valdes-Hernandez, *Member, IEEE*, Natalie A. Royle, Jinah Park*, *Member, IEEE*, The Lothian Birth Cohort 1936 Collaborative Group, and The Alzheimer's Disease Neuroimaging Initiative

Abstract—Accurately recovering the hippocampal shapes against rough and noisy segmentations is as challenging as achieving good anatomical correspondence between the individual shapes. To address these issues, we propose a mesh-to-volume registration approach, characterized by a progressive model deformation. Our model implements flexible weighting scheme for model rigidity under a multi-level neighborhood for vertex connectivity. This method induces a large-to-small scale deformation of a template surface to build the pairwise correspondence by minimizing geometric distortion while robustly restoring the individuals' shape characteristics. We evaluated the proposed method's 1) accuracy and robustness in smooth surface reconstruction, 2) sensitivity in detecting significant shape differences between healthy control and disease groups (mild cognitive impairment and Alzheimer's disease), 3) robustness in constructing the anatomical correspondence between individual shape models, and 4) applicability in identifying subtle shape changes in relation to cognitive abilities in a healthy population. We compared the performance of the proposed method with other well-known methods—SPHARM-PDM, ShapeWorks and LDDMM volume registration with template injection—using various metrics of shape similarity, surface roughness, volume, and shape deformity. The experimental results showed that the proposed method generated smooth surfaces with less volume differences and better shape similarity to input volumes than others. The statistical analyses with clinical variables also showed that it was sensitive in detecting subtle shape changes of hippocampus.

Index Terms—Brain, hippocampus, magnetic resonance imaging (MRI), progressive model deformation, shape analysis.

I. INTRODUCTION

THE HIPPOCAMPUS has attracted researchers' attention because its atrophy has been identified as marker for the development of Alzheimer's disease [1], [2], and has also been related to mild cognitive impairment [2]–[4], schizophrenia [5], [6], and epilepsy [7]. To quantitatively assess hippocampal atrophy, volumetric measurements on magnetic resonance (MR) images [8], [9] have been often applied, as volume changes are features that may explain atrophy or inflammation due to illness. However, volumetric analyses may not reveal the true association between local deformations and disease or risk factors. For example, the increase in volume on enlarged regions can be compensated by the shrinkage on other locations causing little or no change in total volume. Additionally, volumetric analyses do not give information about the precise locations of morphological changes that characterize the appearance and progression of several neurodegenerative diseases. A precise analysis of these changes could provide useful diagnostic information and help to identify individuals at risk. Shape-based morphometry has emerged as the preferred tool for analyzing the structural changes in human subcortical structures including the hippocampus [10]–[13].

Several methods for the regional hippocampal shape analysis have been suggested. Some of these are: medial shape representation [6], [14], spherical harmonic representation (SPHARM) [6], [15], a combination of both (i.e., SPHARM-based medial representation) [16], combination of medial shape surfaces and local width analysis [17], diffeomorphic maps [18] with variations [19]–[22], and minimum description length framework

[23] and entropy-based particle systems [24] for statistical shape modeling. Studies involving healthy adults, participants with mild cognitive impairment (MCI) and Alzheimer's disease (AD) patients have demonstrated the usefulness of shape analysis approaches to study the evolution of hippocampal shapes across the human lifespan [25], identify regionally-specific shape abnormalities of six subcortical structures and lateral ventricles in MCI and AD [26] and contribute predicting the conversion from MCI to AD using 3-D hippocampal morphology [27].

However, despite of the existence of a wide-range of shape modeling and shape-based morphometry approaches, several challenges still remain to be overcome. First of all, the hippocampus shape can be affected by the presence of holes (mainly caused by high cortisol levels in response to stress, perivascular spaces or small infarcts [28]), and elongated shape and sharp and pointy edges (e.g., hippocampal tail). Secondly, the small size of this structure even in normal young adult individuals leads to rough boundaries on the binary masks with increasing effect of the voxel size; and thirdly, image artifacts, presence of noise, slice thickness and inter-slice gaps increase the roughness in the borders of the segmented hippocampi producing also rough surfaces. Some approaches [24], [29] have tried to address these challenges using a fill-hole operation combined with smoothing operations, such as Gaussian blurring and level-set based contour estimation, as preprocessing steps. However, these can produce false hippocampal boundaries depending on the severity of the zigzag effect on the previously delineated boundaries, and this may yield negative effects on the shape-based morphometry. Studies using the diffeomorphic registration [20], [22] showed that the template image registration using large deformation diffeomorphic metric mapping (LDDMM) could trace more accurate and smooth boundaries of the hippocampus against noise and wrong segmentations from automatic methods while establishing the point correspondence between the template and targets via a template mesh injection.

Previous studies of aging have used hippocampal-shape analysis on participants within a wide age range (e.g., 55–90 years old [26] and 18–94 years old [25]), for which morphological hippocampal differences due to age are expected, and with a wide spectrum of cognitive abilities (i.e., from normal healthy to AD patients). To the best of our knowledge, shape analysis methods of subcortical brain structures have only been applied to study differences between diseased and healthy sample populations. Therefore, the sensitivity of these methods to detect early morphological changes in normal older people, and hence their applicability to early diagnosis of neurodegenerative diseases, requires examination. Detecting subtle morphological deformations in the hippocampus of nondiseased subjects rely on the robustness of the individual shape modeling method against noise and rough boundaries.

A. Overview and Contributions

The aim of this paper is to propose a template-based approach for the hippocampal shape modeling that is robust against noise and rough segmentations and sensible to subtle morphological variations. Our shape modeling method performs a mesh-to-image registration, fitting a template surface of fixed point sets and vertex connectivity into the image boundaries while minimizing the geometric distortions of the template model. It is based on a novel Laplacian surface deformation principle, which has a multi-level neighborhood for vertex connectivity and a flexible weighting scheme of the rigidity parameter. These guarantee: 1) anatomical point correspondence between the individual subjects' surface models, and 2) robust reconstruction of individual shape details against large variations of shape and size, noise, and rough boundaries. We describe the details of the proposed method in Section II.

In Sections III–V, we also present various experiments to demonstrate the proposed method's 1) accuracy and robustness in the smooth surface reconstruction with minimal distortion while filtering out high-frequency noises (Section III), 2) sensitivity and consistency in detecting the significant shape differences between healthy control and disease groups (i.e., mild cognitive impairment and Alzheimer's disease) (Section IV), 3) robustness in constructing the anatomical correspondence between individual shape models, and 4) applicability in identifying the subtle shape changes of the hippocampi in relation to cognitive abilities across healthy subjects (Section V). For the experiments, we used three datasets of hippocampi: 1) synthetic images with artificial shape deformation and noises, 2) automatically segmented using the T1-weighted MR images obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, and 3) semi-automatically segmented using the T1-weighted MR images of the Lothian Birth Cohort 1936 (LBC1936) Study. In the experiments, we compare the performance of the proposed method with three well-known shape modeling methods: spherical harmonics followed by a point distribution model (SPHARM-PDM) [29], ShapeWorks [24], [30], and LDDMM volume registration [31], [32] with template injection [20], [22] (LDDMM-TI) using various metrics of shape similarity, surface roughness, volume, and shape deformity to quantitatively investigate how well the methods work in denoising the rough boundaries and tracing the target shape characteristics.

II. Shape Modeling Method

Our shape modeling method generates hippocampal surface models for each individual subject by deforming a template model that encodes the generic shape characteristics of the hippocampus as a triangular mesh. This model deforms in the binary images directly while keeping its anatomical regularities, point distribution and surface smoothness, against the rough boundaries of the binary segmentations. For this purpose, we employ a Laplacian deformation framework supporting the preservation of geometric details (e.g., local curvature and relative point distance) in free-form mesh deformation [33]–[35]. The Laplacian-based deformation is widely used for various graphics applications, owing to its fast computation and robustness in finding a unique global minimum for the quadratic energy function [36]. However, the linear system of the Laplacian-based deformation does not guarantee the optimal minimization of the geometric distortions against "stretching" operations deforming a surface mesh into varied shapes of arbitrary size [36]. To overcome this limitation, various approaches have been proposed in interactive surface editing [37] and mesh parameterization [38]. In this paper, we newly introduce

a simple and efficient approach, called "progressive model deformation," to address the limitation during the process of mesh-to-volume registration for hippocampal shape modeling. The progressive model deformation adjusts the scale of the surface mesh deformation flexibly in accordance with the size and shape variations between the template model and target volumes to minimize the geometric distortions of the template model while mapping its vertices onto image boundaries. In the following sections, we describe our shape modeling method in details.

A. Template Model Representation With Multi-Level Neighborhood and Laplacian Coordinates

The template model is used as shape priors in the individual shape modeling process. The template model can be constructed out of image atlases generated from a representative sample of a population. In this study, we subsequently use marching cubes algorithm [39], mesh smoothing and mesh resampling methods for the template model construction from a given image atlas. Note that our nonrigid shape modeling method is not restricted to a specific image atlas. It is possible to use other hippocampal atlases in public domain. A requisite of our shape modeling method is to construct the template model as manifold without boundary.

Additionally and importantly, the connectivity of multiple neighborhoods between points (vertices) is newly defined in the template model to accommodate them with a progressive surface deformation, which preserves the shape features (i.e., relative triangle size and local curvature) of the template model as rigid as possible to build a dense correspondence between template model and target volume. The multi-level (N-ring) neighborhood of each vertex in the template model is determined as follows [40]:

- 1-ring neighbor vertices of a vertex (v_i) are the vertices of the faces incident on v_i.
- N-ring neighbor vertices of v_i are the union of its (N-1) neighbor vertices and their 1-ring neighbors.

Within the N-ring neighborhood, each vertex is connected directly to its neighbors, and this connectivity composes a volumetric graph for each vertex with the vertices of N-ring neighbors. This extension of the vertex connectivity in the template model provides an explicit control on the range of the feature preservation.

Under the definition of the multi-level neighborhood, we encode the local geometry of the template surface using a discrete Laplacian coordinates (δ_i), which is defined as

$$\delta_i = L(v_i) = \sum_{j \in \mathbb{N}(v_i)} \lambda_{ij}(v_i - v_j), \quad \lambda_{ij} = \frac{\omega_{ij}}{\sum_{j \in \mathbb{N}(v_i)} \omega_{ij}} \quad (1)$$

where $N(v_i)$ is the set of vertex indexes of N-ring neighborhood and $L(.)$ is the discrete Laplacian operator. The weight (w_{ij}) between v_i and its neighbor (v_j) follows the weighting scheme of the mean value coordinates [41]

$$\omega_{ij} = \frac{\tan\left(\frac{\gamma_{j-1}}{2}\right) + \tan\left(\frac{\gamma_j}{2}\right)}{\| v_i - v_j \|}. \quad (2)$$
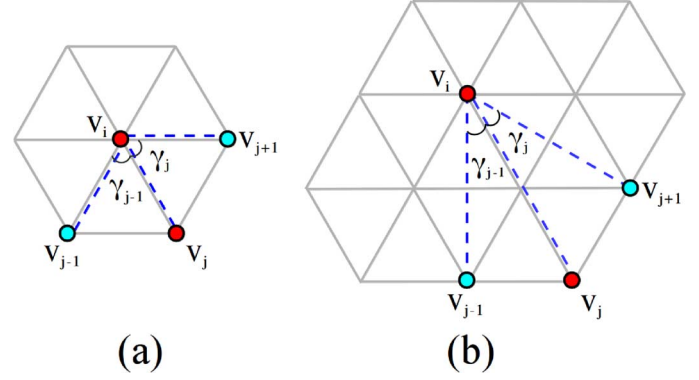

Fig. 1. Graphical description of the vertex connectivity of the multi-level neighborhood. Angles (γ_{i-1} and γ_i) for the weight (w_{ij}) for an edge between v_i and v_j are depicted. (a) 1-ring neighbor and (b) 2-ring neighbor. Dotted lines (blue) indicate the direct connection between a vertex (v_i) and its N-ring neighbors. N-ring neighborhood includes all connectivity of $N-1$ ring neighborhood.

The angles (γ_{j-1} and γ_j), depicted in Fig. 1, are computed from the template models for the different levels of N-ring neighborhood. The mean-value weight function λ_{ij} satisfies the following properties [41]–[44]: 1) affine invariance: $\sum_{j \in \mathbb{N}(v_i)} \lambda_{ij} = 1$, 2) smoothness on the surface, 3) positiveness, and 4) linear precision: $|\delta_i| = 0$ whenever v_i and v_j are coplanar. This mean-value weight function is well defined and smooth for arbitrary shapes, and it helps to smoothly distribute the external forces, which guide the template model to the image boundary at each vertex, across neighbor vertices in our deformation framework.

B. Progressive Model Deformation Framework for Modeling Individuals' Surfaces

Once the template models are constructed, they are used for all individual shape recovery with the pairwise shape correspondence between template and targets. Given a hippocampus binary mask, we first align the template model using an optimal rigid transformation minimizing the geometric differences between the template model and the target hippocampus. This transformation is computed by an iterative closest point algorithm [45] between the template model and the voxel mesh extracted from the binary mask.

The individual shape reconstruction with the dense correspondence between the template and targets is achieved by a Laplacian deformation framework, which finds optimal vertex coordinates minimizing following energy function:

$$E(V') = \sum_{i=1}^{n} \| \alpha_i (L(v_i') - \delta_i) \|^2 + \sum_{i=1}^{n} \| b_i - v_i \|^2. \quad (3)$$

In (3), V' is the set of the optimal vertex coordinates (v_i'), and δ_i is the Laplacian coordinates of each vertex (v_i), which is obtained using the discrete Laplacian operator described in the previous section. α_i is a parameter to control the model rigidity by preserving the Laplacian coordinates (δ_i) during the deformation. b_i is the desired position where v_i will be placed after each iteration. b_i is defined as

$$b_i = \beta \cdot (m_i - v_i) + v_i \quad (4)$$

where m_i is the closest image boundary on the direction normal to the vertex, and β is a weighting term for the stability of model deformation. Simply, the first term of (3) is the function of the changes of the Laplacian coordinates, representing the local geometry of the template model at v_i with respect to its neighbor. The second term of (3) is the function quantifying the discrepancy between the template vertices and the target surface. The optimal positions of the vertices for (3) can be obtained by iteratively solving a linear system in a linear least squares approach [33], [35].

To obtain the point set on the target surface, which satisfies the required smoothness and the pairwise vertex correspondence with the template model, the progressive model deformation derive a large-to-small scale surface deformation in the modeling process using the multi-level neighborhood previously explained and a flexible weighting scheme for the model rigidity parameter (α_i). Following the analogy used by [46], the Laplacian-based deformation framework can be viewed as a mass-spring system of which nodes (vertices) are connected by springs of various strengths. In this system, the external force [i.e., $b_i - v_i$ in (3)] guiding v_i to the image boundary is propagated along the connected vertices after transforming v_i to the optimal position; thus minimizing the changes of the Laplacian coordinates against the external force. The spring strength is determined by the weighting function (w_{ij}) for each edge and the rigidity parameter (α_i). w_{ij} is an intrinsic parameter which depend on the geometric shape and the vertex connectivity of the template model, and α_i is an extrinsic parameter that determines the model rigidity flexibly in proportion to the geometric differences between template and individual targets during the nonrigid deformation.

First, the multi-level neighborhood derives the large-to-small scale deformation for the individual shape modeling as follows. In the higher-level neighborhood (see Section II-A), each vertex is connected to its neighbors of a larger domain as a volumetric graph and the external force is propagated smoothly and largely across the surface (Fig. 2). This behavior derives a large-scale deformation in which the vertex displacement influences a larger number of vertices to be displaced in the direction of the external force to preserve the local geometric shape of the template model. The large scale deformation can preserve the relative areas of each vertex and its neighbors thanks to the weighting function (w_{ij}). However, it cannot reproduce the small shape details of the target surface [Fig. 3(d)]. Our strategy on this matter is to reduce the level of the neighborhood, when the template model is not deformed anymore with a higher-level of neighborhood. We determine this status of the model deformation using a threshold (ε in Algorithm 1) for average vertex displacements. This approach yields to the vertex-wise deformation of the template model in smaller regions (small-scale deformation) and decreases the discrepancy between the vertices and the target surface. Consequently, this large-to-small scale deformation of the template model places each vertex close to the image boundary while preserving the relative positions between the vertices as strong as possible against the anisotropic size and regional shape variations between template and target surface [Fig. 3(e)].
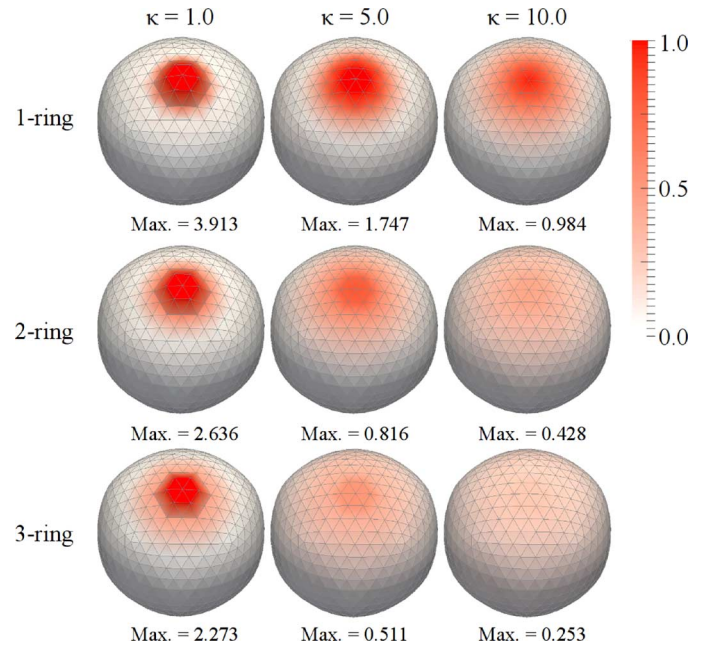
Fig. 2. Vertex displacement with different levels of the neighborhood and model rigidity parameters. Max.: maximum displacement of vertices (mm).
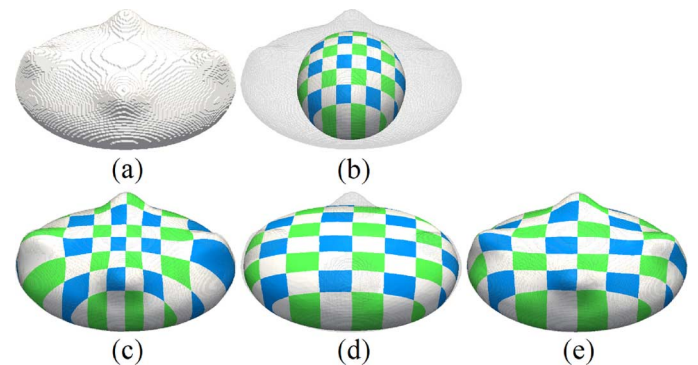
Fig. 3. Example of the progressive model deformation. (a) a synthetic binary mask, (b) initial state of a template model (sphere) with target surface, (c) deformed model with $\kappa = 1$ and 1-ring, (d) deformed model with $\kappa = 5$ and 3-ring, and (e) deformed model with $\kappa = 1$ and 1-ring after deformation with $\kappa = 5$ and 3-to-1 ring. Texture colors indicate the corresponding vertices between surface models.

In addition, the rigidity parameter α_i for each vertex v_i in (3) determines the degree of the preservation of the Laplacian coordinates δ_i. The external force is propagated linearly to the neighbors with respect to α_i. As a larger value is assigned to α_i, the local geometry (δ_i) is preserved more rigidly and larger amounts of the external force are propagated to its neighbors (Fig. 2). Using this characteristic of the Laplacian deformation framework, we synchronize α_i with the magnitude of the external force to preserve the point distribution of the template model against arbitrary shape and size variations. α_i is determined using the following equation:

$$\alpha_i = \kappa \cdot |b_i - v_i| \qquad (5)$$

where κ is a constant weight to control overall rigidity during the model deformation. As the magnitude of the external force at

---

Algorithm 1: The progressive model deformation

Input: Target volume (V) and template model (M)
Output: Deformed template model (M')
  (a) **Template Initialization**;
    1. Align M rigidly to V via iterative closest point algorithm;
    2. Set initial values for the modeling parameters;
      • Constant parameter $\kappa = \kappa_{init}$;
      • Initial level of the neighborhood $N = 3$ (experimental value);
  (b) **Non-rigid model deformation**;
    while $N >= 1$ **do**
      1. Compute vertex normals of M;
      2. Search the closest boundary at v_i of M;
      3. Compute model rigidity weight (α_i);
      4. Solve linear system of the Laplacian deformation;
      5. Regulate vertex displacement into the RSI transformation;
      6. Update the vertex positions of M and compute the average magnitude (d) of the vertex displacements;
      if $d < \varepsilon$ **then**
        $N = N - 1$;
  (c) **Local shape detail restoration**;
    1. Set $\kappa = \kappa_{min}$ and $N = 1$;
    2. Perform the non-rigid model deformation process until $d < \varepsilon$;

---

v_i is larger, more neighbor vertices are moved in the direction of the external force to preserve the local geometry of the template model. And, as v_i is closer to the target boundary, α_i gets smaller and v_i is moved along the displacement of other vertices, which are not fitted into the boundary enough. For κ, we implement a two-step model deformation to control the degree of the shape detail representation explicitly. Simply, we perform the model deformation with an initial κ (κ_{init}) and the multi-level neighborhood. Then, if necessary, we do an additional deformation with κ of a smaller value (κ_{min}) and the vertex connectivity of the 1-ring neighborhood to restore more shape details of the target structures. The shape modeling process based on our progressive model deformation is summarized in Algorithm 1.

C. Rotation-and Scale Invariant (RSI) Transformation

The rough surfaces of hippocampus yield large variations in the magnitude of the external force denoted by $b_i - v_i$ in (3) [Fig. 4(a)]. The variation of the external forces at each vertex is proportional to the quality of the fit between the template and the target. Moreover, large voxel sizes (i.e., low spatial resolution) encompass an increase in the variation of the external factors' magnitude. Fig. 4(a) shows the external forces at each vertex on a section of the template surface of a hippocampal model. The modeling result obtained from applying the two previously explained deformation techniques displays large variations in the local curvature on the hippocampal surface [Fig. 4(b)]. This is a common problem observed in nonrigid surface deformation methods that use vertex normal boundary search.
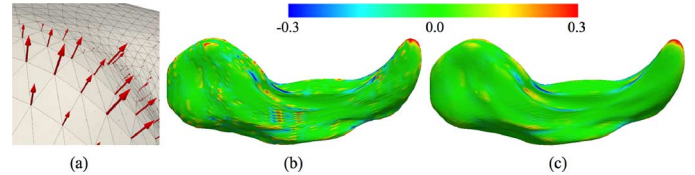

Fig. 4. Variation of external factors (left) and hippocampal surface models reconstructed without/with rotation-and-scale invariant transformation (right). The surface models are color-mapped as shown in the scale, with Gaussian curvature (red: $+0.3$, green: 0.0, blue: -0.3). (a) External Image Factors. (b) Without RSI Transformation. (c) With RSI Transformation.

To overcome this problem, we apply a rotation-and-scale invariant (RSI) transformation to preserve the surface quality of the template model while recovering the small shape details. The RSI transformation, proposed by [34], constrains the vertex transformations only to rotation, isotropic scale and translation. This helps regularizing the individual vertex transformation, derived by external factors, to the transformations of the neighbor vertices using them as reference. The RSI transformation of each vertex is determined by finding the optimal parameters that minimize the differences between the current positions of each vertex (and that of its neighbors) to the newly computed position for each of them via (3) as (6) shows

$$T'_i = \underset{T_i}{argmin} \left(\parallel T_i v_i - v'_i \parallel^2 + \sum_{j \in \mathbb{N}(v_i)} \parallel T_i v_j - v'_j \parallel^2 \right). \tag{6}$$

This quadratic function is solved for each vertex by a linear least squares approach in a similar way to (3). The formula that describes this process can be found in [34]. For hippocampal shape modeling, the application of the RSI transformation successfully maintained the surface smoothness while restoring the local shape characteristics of the targets (i.e., binary masks) in the iterative process [Fig. 4(c)].

III. ROBUSTNESS EVALUATION AGAINST NOISE AND REGIONAL SHAPE DEFORMATIONS ON SYNTHETIC DATA

We evaluated our hippocampal modeling method via the following experiments: A) performance under conditions of noise and extreme deformations on synthetic data, B) sensitivity in detecting the regional shape differences between control and disease groups using data obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, and C) accuracy, applicability and usefulness using data from a study of aging. In the first two experiments, we compared our method with other three modeling methods; SPHARM-PDM (Ver. 1.11, www.nitrc.org/projects/spharm-pdm) [29], ShapeWorks (Ver. 1.1, www.sci.utah.edu/software/shapeworks.html) [24], [30] and LDDMM-TI [22], [31], [32]. For LDDMM volume registration, we used an open source software, uTIlzReg (sourceforge.net/projects/utilzreg). We also evaluated the consistency of findings via statistical analyses comparing our results with those obtained by previous neuroimaging studies about Alzheimer's disease.

The four shape modeling methods evaluated can be grouped differently with respect to data representation, denoising methods and shape correspondence approaches. According

TABLE I
GENERAL PARAMETERS FOR EACH MODELING METHOD

	LDDMM-TI	Progressive	ShapeWorks	SPHARM-PDM
Smoothing Parameter	Gaussian Kernel (σ) 1.5 to 4.5	κ_{min} 2.0 to 20.0	Gaussian Smoothing (σ) 0.1 to 2.0	Gaussian Smoothing (σ) 0.1 to 2.0, maximal degree of SPHARM 10.0 to 20.0
No. Vertices	2463	2463	2048	4002

Legend: LDDMM-TI: LDDMM volume registration with template injection, Progressive: progressive model deformation (i.e. our proposed model), σ: standard deviation
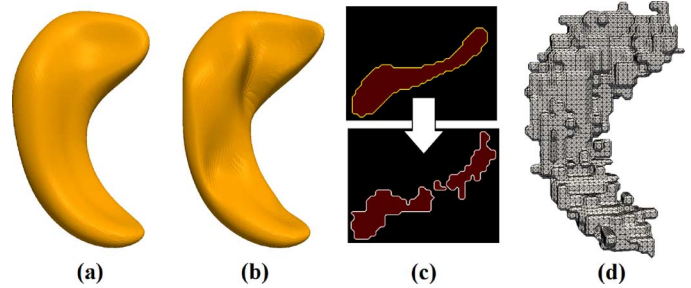

Fig. 5. Synthetic surfaces and noisy binary masks. (a) Synthetic surface with anisotropic scale. (b) Synthetic surface by adding regional deformation to (a). (c) Voxelized surfaces with salt-and-pepper noise addition. (d) Isosurface mesh generated from the noisy binary mask.

to the classification of shape registration methods in [47], our framework performs mesh-to-volume correspondence, SPHARM-PDM and ShapeWorks perform a mesh-to-mesh correspondence, and LDDMM-TI performs a volume-to-volume correspondence. With regards to noise represented by rough boundaries and holes, SPHARM-PDM and ShapeWorks support fill-hole and smoothing operations (anti-aliasing and Gaussian smoothing) on binary masks [24], [29]. These operations help to remove the high-frequency image artifacts and generate smooth surfaces of the target structure. The smoothness or denoising effects are determined by the spread parameter (i.e., standard deviation for Gaussian smoothing). SPHARM-PDM also supports a parameter to control the maximal degree for the SPHARM computation, which determines the level of the shape detail representation [29]. On the contrary, the LDDMM-TI and our method preserve the shape characteristics of a smooth template model, which is represented as an image volume or a surface mesh. In the LDDMM volume registration, the Gaussian kernel is used to describe the velocity fields and generate smooth diffeomorphism in most approaches [32], [48]. The standard deviation of the Gaussian kernel determines the smoothness of the reconstructed surfaces. In our approach, $\kappa_{\min}$ determines the level of shape detail restoration. From the point of view of shape correspondence, SPHARM-PDM, LDDMM-TI, and our method provide pairwise point correspondence between the individual surface and a template surface. ShapeWorks find an optimal solution of group-wise point correspondence using an entropy-based particle system. Table I shows the values of the parameters used for each method unless stated otherwise. We validated the shape correspondence of each method by testing their sensitivity in detecting regional shape differences between hippocampal surfaces.

A. Experiment Design

The goal of this experiment is to evaluate the robustness of the four modeling methods previously mentioned in detecting the regional shape differences against image artifacts, such as sharp edges and holes. For this, we first built a set of synthetic hippocampal surfaces including smooth and locally deformed shapes of hippocampus with different sizes. We transformed a smooth surface mesh of the left hippocampus using anisotropic scale transformation with anisotropic scale

factors ($\mathrm{x, y, z} = 1.2, 1.1, 1.0$). Then, we added local shape deformations to the synthetic surfaces by translating the vertices using surface height maps, which include the known magnitudes of vertex displacement for each vertex along the vertex normals. In the height maps, we set the maximum displacement of vertices as 3 mm. Through this process, we obtained synthetic surfaces of hippocampi with and without local shape deformations with known magnitude. The synthetic surfaces were converted into binary masks of $1.0 \times 1.0 \times 1.0$ mm³ voxel size as inputs of the surface modeling methods.

We simulated image artifacts of noisy surface boundaries in the generated binary masks through the following steps: 1) salt-and-pepper type image noise was added to the binary images (probability $= 0.1$) and 2) the floating particles, not connected to the synthetic surfaces, were removed. This voxel-wise operation for noise generation produced small local perturbations, such as holes and very sharp outliers, as well as rough surfaces in the binary images of the hippocampi. Fig. 5 shows the synthetic data generated via this process. With the "noisy" binary masks, we tested the quality of the denoising approach and the target shape reconstruction of each shape modeling method. For this, we measured surface roughness, shape similarity and surface volume on the modeling results, obtained after performing the shape modeling with sequential values of the parameters that determine the smoothness or the level of shape detail, for each method. We also computed the local shape deformity between the reconstructed shape models; and visually compared the deformity values with the magnitudes of the vertex displacement in the height maps.

B. Metrics of Surface Roughness and Shape Similarity

To quantify the quality of the reconstructed surface models and trace the denoising patterns of each method, we employed well-known metrics. First, we converted the reconstructed models into binary masks with the same voxel size ($1.0 \times 1.0 \times 1.0$ mm³) to compare the surface quality of the shape reconstruction process, including regional deformations and handling of image artifacts, with the target binary masks, which, in practice, were obtained from the segmentation process.

The surface roughness was assessed by a roughness metric using mean curvature [49]. The surface roughness was determined using the difference $D(v_i)$ between the mean curvature at each vertex (v_i) and the average mean curvature of its N-ring
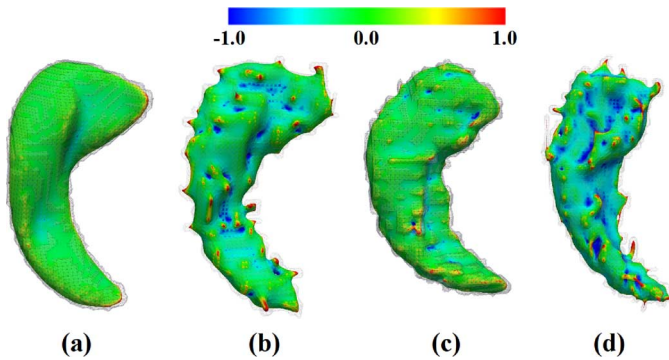
Fig. 6. Maps of the difference $D(v_i)$. (a) and (b) Synthetic surfaces without and with image artifacts. The roughness of (a) is 0.244 and (b) is 1.015. (c) Surface of the minimum roughness (0.506) in a dataset from ADNI database (see Section IV-A). (d) Surface of the maximum roughness (1.115) in the same dataset. Wireframed meshes around the surfaces of colormaps are the isosurface meshes generated from the binary masks.

neighborhood. The equation of the surface roughness of a surface (S) is given by

$$R\left(\mathbb{S}\right) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(D(v_i)\right)^2}. \tag{7}$$

We computed the surface roughness of the target structure from a smooth isosurface generated by marching cubes and mesh smoothing algorithms from any input binary masks. The average mean curvature for $D(v_i)$, computed from all vertices except v_i, is named as global roughness. It describes the average behavior of all vertices in S. Note that the isosurfaces were scaled with respect to the volume of the template model, used for experiments, to compare the roughness between the surfaces of different size. Fig. 6 presents the maps of $D(v_i)$ on the surfaces from the synthetic and ADNI (clinical) dataset. In the former, the artificial image artifacts, which were simulated by salt-and-pepper type noise [Fig. 6(b)], showed similar patterns to the rough boundaries of the binary masks obtained from segmented the hippocampi [Fig. 6(c) and (d)].

The shape similarity between the reconstructed models and the target binary masks was measured using three metrics: a volumetric similarity index [i.e., Dice coefficient (DC)], the symmetric mean distance (MD), and the symmetric Hausdorff distance (HD). Briefly, the DC measures the volume overlap between two volumetric data A and B [50] by

$$\mathrm{DC}\left(A, B\right) = 2 \cdot \frac{n\left\{A \cap B\right\}}{n\left\{A\right\} + n\left\{B\right\}}. \tag{8}$$

In (8) the operator $n\left\{\cdot\right\}$ returns the size of the input data and the operator $\cap$ returns the spatially common (i.e., intersection) volume between A and B. The DC ranges between 0 and 1 representing disjoint and complete overlap, respectively. The MD measures the average distance between two surfaces, and the HD measures the maximum distance between them. We computed the MD and HD between the binary masks using Insight Toolkit (ITK, Ver. 4.5.2, www.itk.org).

C. Experiment Results

Fig. 7 shows the denoising effects of each method by comparing the binary synthetic hippocampi (S) against the binary

"noisy" synthetic hippocampi (S') and/or against the modeled (M) hippocampi. The latter were obtained from applying our progressive model deformation method (a), the LDDMM-TI (b), anti-aliasing and Gaussian smoothing for SPHARM-PDM and ShapeWorks (c), and SPHARM-PDM with different degree of the SPHARM computation (d). Fig. 7(d) was obtained from the binary masks generated by the anti-aliasing and Gaussian smoothing operation of 1.0 standard deviation. The horizontal axis shows the different consecutive values of the smoothing parameters used for each modeling method.

As can be appreciated in Fig. 7, all methods had excellent performance against high frequency noises for medium-to-high values of the smoothing parameters (first column in Fig. 7). However, the patterns of volume and shape changes in the denoising process were very different between the methods. The anti-aliasing and Gaussian smoothing operation for SPHARM-PDM and ShapeWorks showed the best results in denoising at 1.0 standard deviation (higher DC, lower MD and HD, and similar volume to synthetic hippocampi). This result was attributed to the voxel size $(1.0 \times 1.0 \times 1.0 \text{ mm}^3)$ of the synthetic masks. As the size of Gaussian filter increased (i.e., larger than 1.0 mm), the shape similarity was getting worse, due to the volume shrinkage with respect to the Gaussian smoothing operation. For SPHARM-PDM, the degree of the SPHARM computation did not affect the shape similarity and volume of the reconstructed models significantly [Fig. 7(d)]. On the contrary to the Gaussian smoothing operation, the template-based methods (progressive model deformation and LDDMM-TI) generated the converged results in all metrics for medium-to-high values of the smoothing parameters. The lower values of the smoothing parameters for each method generated more rough surfaces. This result indicates that the template-based approaches fitted the template into target surface globally and then filtered out or included the local details of the binary masks. This behavior of the template-based denoising helps to trace the surface of target structure without shrinking the volume, and makes all the modeling process more robust against large variations produced by noise in binary masks from clinical datasets. In addition, our method produced better results than the LDDMM-TI in volume preservation for consecutive values of the smoothing parameters.

Fig. 8 presents the shape deformity measured from the shape models of each method. Since ShapeWorks provides a set of points without connectivity information as an output, we visualized the shape deformity on the spherical particles. By placing the vertices on the target boundary with filtering out the high frequency noise, each method detected the artificial shape deformation added in the binary masks. Our progressive model deformation and the LDDMM-TI showed better performance in localizing the regional shape deformation than SPHARM-PDM and ShapeWorks. SPHARM-PDM and ShapeWorks did not detect correctly the deformation at the lateral anterior part of hippocampus, where the cornu ammonis 1 (CA1) meets with the fimbria. The detected regions of SPHARM-PDM were distorted (stretched). This result may be originated from the artificial oscillations in the surface models by using spherical harmonics (12 degree in our experiment), reported in [51]. In the results from ShapeWorks, we also observed the intersections between
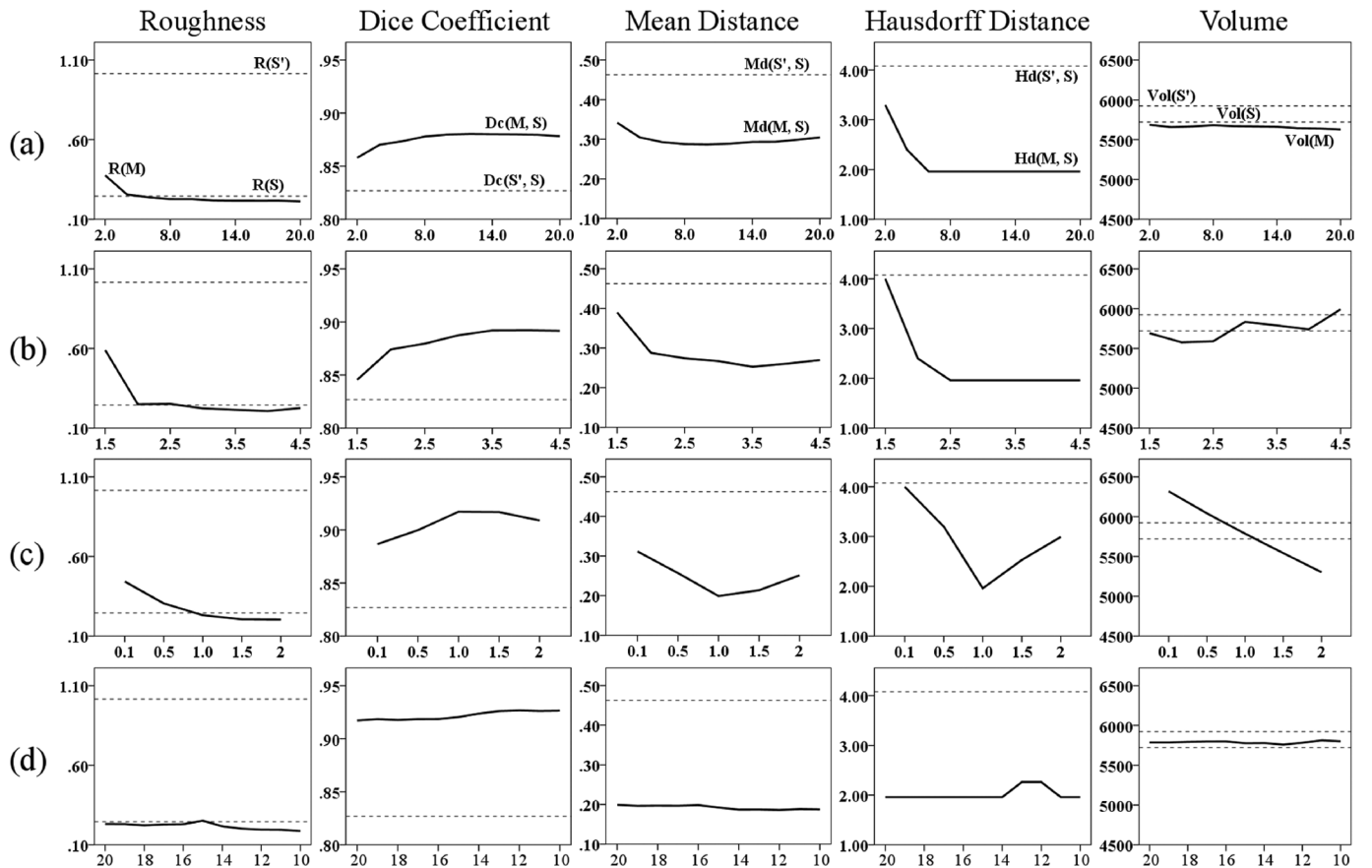
Fig. 7. Denoising effects of each method (a) progressive model deformation with the $\kappa_{\min}$, (b) LDDMM volume registration with template injection (LDDMM-TI) with the standard deviation of the Gaussian kernel, (c) anti-aliasing and Gaussian smoothing with the standard deviation for SPHARM-PDM and ShapeWorks, and (d) SPHARM-PDM with the maximal degree for the SPHARM computation from the binary mask, with anti-aliasing and Gaussian smoothing of 1 mm standard deviation. Horizontal axis of all plots displays the corresponding smoothing parameters for each method, where its increase smooths the shape. Vertical axis is for the measured value—roughness (first column), shape similarity (second to fourth columns), and volume (last column)—from the experiment. Top and bottom dotted lines in the plots of the first and last columns indicate the reference values measured from the binary mask with and without noise, respectively. Dotted lines in the plots of the middle three columns indicate the shape similarity between the binary masks with and without noise. M: the reconstructed models, S: binary mask without noise (gold standard), and S': binary mask with noise.

the difference vectors of the corresponding points, on different binary masks. The intersections indicate the wrong point correspondence. They also caused wrong values of the shape deformity [e.g., green particles among red particles at hippocampal head (last column in Fig. 8)]. This limitation of ShapeWorks have been reported in [52].

IV. SHAPE MODELING AND ANALYSIS ON A DATASET OF THE ALZHEIMER'S DISEASE NEUROIMAGING INITIATIVE

In this section, we present experiments using the surface models, generated by the four shape modeling methods, to evaluate their performance in denoising the rough boundaries and detecting the regional shape differences with respect to clinical factors.

A. Materials and Image Processing

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). It hosts clinical, demographic, cognitive and MRI data from subjects recruited from over 50 sites across the U.S. and Canada. To date it comprises data from over 1500 adults, ages 55–90, who are either cognitively normal older individuals, individuals with early or late MCI, or individuals with early AD. For up-to-date information, see www.adni-info.org.

The database of ADNI gathers cross-sectional and longitudinal data including 1.5T structural MR images from 800 subjects and 3.0T structural MR images from 200 subjects. The MRI protocol of ADNI is described in [53]. For our study, we collected a subset of MR images of 150 subjects, of which 50 were AD patients, 50 were individuals with MCI, and 50 were healthy controls. For reproducibility purposes, the study participant's identifiers used are provided in the Supplementary material. From each subject, we used the T1-weighted structural MR images, acquired at 1.5T, from the screening visit. ADNI performed some postprocessing steps to correct certain image artifacts and to enhance standardization across manufacturers. The postprocessing steps include 3-D gradwarp correction [54], [55], B1 nonuniformity correction [56] and N3 image intensity nonuniformity correction [53]. We use these postprocessed images in this study. The demographic characteristics of the collected subjects are given in Table II. We collected the subjects' information such as MMSE score and intracranial volume from ADNI database. Left and right hippocampi were
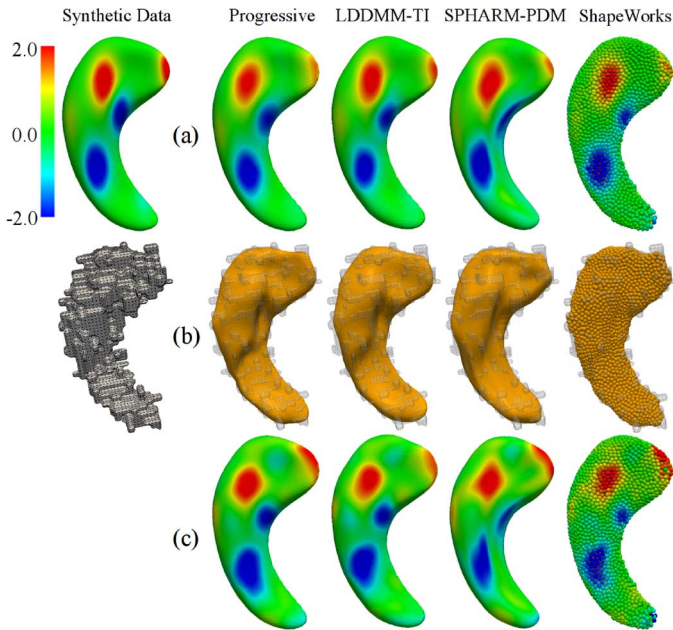
Fig. 8. Comparison of the modeling methods in detecting the regional shape deformation. (a) Shape deformity between the shape models reconstructed from the binary masks without image noise. (b) Shape models reconstructed from "noisy" mask, and (c) shape deformity between the shape models reconstructed from the noisy masks.

TABLE II
DEMOGRAPHIC CHARACTERISTICS OF THE ADNI SUBJECTS
INCLUDED IN THIS STUDY

	N	Gender	Age (year)	MMSE score	LHV (mm^3)	RHV (mm^3)
Control	50	23 F / 27 M	77.1 ± 5.8	29.0 ± 1.0	6997 ± 776	7154 ±1049
MCI	50	16 F / 34 M	75.9 ± 8.3	26.9 ± 1.9	5911 ±1036	6307 ±1130
AD	50	24 F / 26 M	75.7 ± 7.1	23.3 ± 2.1	5475 ± 909	5671 ± 987

Legend: MMSE : Mini Mental State Examination, LHV : left hippocampal volume, RHV: right hippocampal volume, MCI : mild cognitive impairment, AD : Alzheimer's disease

segmented using FIRST software [13] from FMRIB Software Library (FSL, ver. 5.0, fsl.fmrib.ox.ac.uk/fsl/fslwiki/). Visual inspection was performed on the hippocampal binary masks to ensure the quality of the automatic segmentation.

Fig. 9 presents the volume and surface roughness of the left and right hippocampi segmented from the collected MR images. We also performed a two-tailed Wilcoxon rank sum test to examine the statistical differences in volume and surface roughness between groups and univariate linear regression to evaluate the relationship between the volume and surface roughness across subjects. In this latter test, the hippocampal volume was scaled by ICV. For both left and right hippocampus, the hippocampal volume and surface roughness of control group were significantly different from those in the MCI and AD groups (p < 0.05, Table III). For left hippocampus, the distributions of the volume and roughness were not different between MCI and AD groups. For right hippocampus, the volume was significantly different between MCI and AD groups (p = 0.036), but the surface roughness was not. The surface roughness for all

cases ranged from 0.5 to 1.2, and these values were similar to the roughness of the synthetic data including the high-frequency noises (Fig. 6). The linear regression analysis on the volume and surface roughness showed that the surface roughness was decreased as the volume was increased for both hippocampi (left: $\beta = -0.048$, p < 0.001, right: $\beta = -0.031$, p < 0.001).

B. Experiment Design

To obtain the individual shape models from the collected dataset, we first built average shape atlases of left and right hippocampi from the segmentations following the process described in [57] and [58]. Then, we constructed surface meshes from the average atlases, which were used as template models of the left and right hippocampi for all subjects, in the same way as it is described in Section II-A. The modeling (smoothness) parameters were determined as the values that produced the best shape similarity and similar roughness of the resulting surfaces in the experiment on the synthetic data. For our method, the template models were directly fitted into the individual binary masks. $\kappa_{\min}$ was 9 and κ_{init} was 20. For LDDMM-TI, the deformation fields satisfying the diffeomorphic properties were obtained between the average shape atlases and the individual binary masks first. Then, the template surface models were deformed following the deformation fields, generated via LDDMM volume registration, as the template injection. The standard deviation of the Gaussian kernel for LDDMM volume registration was 4.0 mm. For SPHARM-PDM, we first performed the anti-aliasing and Gaussian blurring with standard deviation 1 mm. The maximal degree of the SPHARM computation was 12. To evaluate the accuracy and denoising effect, we measured the shape similarity, surface roughness and volume from the individual models using the metrics, introduced in Section III-B and we compared the measured values between the shape modeling methods using a Wilcoxon signed-rank test. We also tested the significant differences in the measured values between clinical groups using a two-tailed Wilcoxon rank sum test. For this test, the volumes of the individual models were scaled by the individuals' ICV.

For the statistical analysis on the regional shape deformation, the individual shape models were normalized separately by each shape modeling method via isotropic rescaling of the shape models using intracranial volume and the generalized Procrustes analysis (GPA) [59] to investigate shape differences between the individualized shape models corrected by individuals' head size. Average models for left and right hippocampi were generated after the normalization. Local shape differences of the individual models were determined by the displacement vectors between the corresponding vertices of the individual surface models and the mean surface model. The shape deformity at each vertex was computed as the signed Euclidean norm of the displacement vectors, projected on the vertex normal on the mean surface model to determine the direction of local shape changes. This type of analysis has been applied previously [7], [12].

With the shape deformity maps obtained from the three methods (i.e., LDDMM-TI, our deformation framework and SPHARM-PDM), we carried out the following statistical analyses: 1) representation of the average displacements per
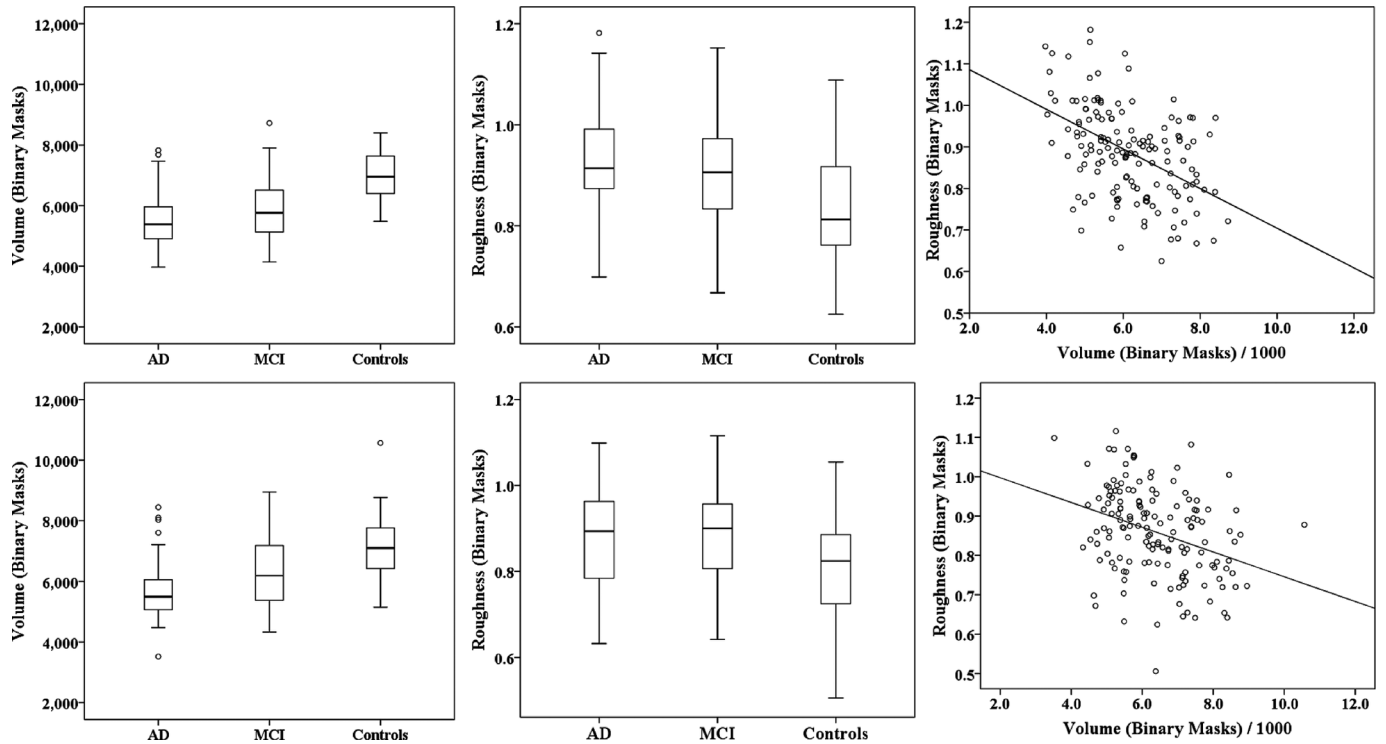
Fig. 9. Volume and roughness of the hippocampal binary masks from a dataset of ADNI. First row gives the box plots of the volume (left), surface roughness (middle) and the fitted line plot between them (right), which were obtained from the individuals' left hippocampus. Second row gives them of the right hippocampus. AD: Alzheimer's disease, MCI: mild cognitive impairment and Controls: healthy controls.

TABLE III
SIGNIFICANT DIFFERENCES IN THE HIPPOCAMPAL VOLUME AND
SURFACE ROUGHNESS BETWEEN GROUPS VIA A TWO-TAILED
WILCOXON RANK SUM TEST

Left Hippocampus			
	Control-MCI	Control-AD	MCI-AD
Scaled Vol.	**< 0.001**	**< 0.001**	0.273
Roughness	**0.002**	**< 0.001**	0.461

Right Hippocampus			
	Control-MCI	Control-AD	MCI-AD
Scaled Vol.	**< 0.001**	**< 0.001**	**0.036**
Roughness	**0.001**	**0.002**	0.967

Significant values (p<0.05) are highlighted as boldface.
Scaled vol. : volume scaled by ICV

group: controls, AD and MCI, to visually compare the average shape deformity of each group with respect to the average model, 2) group-wise comparison using the Wilcoxon rank sum test to determine regions of significant morphological shape differences between groups, and 3) robust univariate linear regression to evaluate the association between local hippocampal shape deformations and cognition [this last represented by the scores of the Mini Mental State Examination (MMSE)] and to determine the strength and significance of these associations using age and gender as covariates. To allow comparability with previously published analyses, we analyzed all subsamples separately and did not combine the results of different methods to conclude about significance on the deformation at any point. Therefore, we did not correct for multiple comparisons. We also reproduced the analysis of covariance (ANCOVA) done in [12] with age and gender as covariates to test the hypothesis

that the shape deformity can discriminate the AD and MCI subjects from healthy controls. As this analysis included correcting for multiple local deformity comparisons, we applied false-discovery-rate adjustment to the p-values produced by ANCOVA reducing the number of false positives.

C. Experiment Results

1) Shape Similarity, Surface Roughness and Volumes: Fig. 10 presents the volume, surface roughness and shape similarity (DC, MD, and HD) of the smooth models, reconstructed by the shape modeling methods. We present the plots of the observed values from the binary masks of right hippocampus only to illustrate the differences between the modeling methods with multiple plots. For the left hippocampus, we obtained similar results to the right hippocampus. For all modeling methods, the volumes and surface roughness of the shape models were significantly different between groups ($p < 0.05$, Table IV). An interesting observation is that the shape models' surface roughness was discriminable between MCI and AD groups ($p < 0.001$). Compared to the roughness of the input binary masks (min.: 0.5, max.: 1.2), the shape models showed the smallest values of surface roughness (min.: 0.1, max.: 0.3). These results indicate that 1) the three shape modeling methods filtered out the high frequency noise in the input binary masks properly and 2) the regional shape variations across hippocampal surface of the shape models were significantly different between the clinical groups.

SPHARM-PDM generated the shape models of largest volume (min.: -29, max.: 483 mm^3) from the input masks and LDDMM-TI generated the shape models of smaller volume
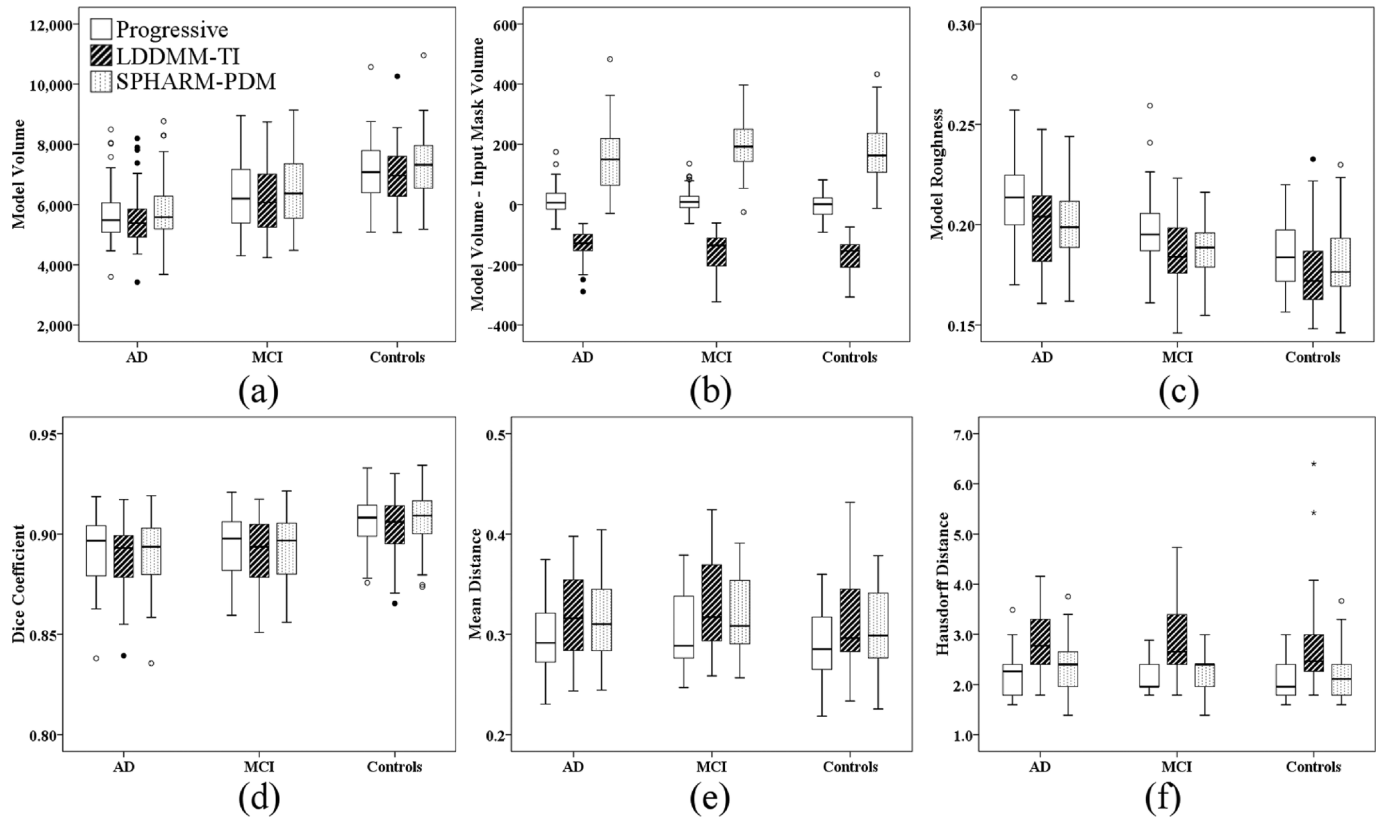
Fig. 10. Volume, surface roughness and shape similarity of the individual shape models of right hippocampus, reconstructed by our progressive model deformation, LDDMM registration + template injection, and SPHARM-PDM. (a) Shape models' volume. (b) Volume difference between the shape models and input masks, (c) models' surface roughness. (d)–(f) Dice coefficient, mean distance and Hausdorff distance between the shape models and input masks. AD: Alzheimer's disease. MCI: mild cognitive impairment and Controls: healthy controls.

TABLE IV
SIGNIFICANT DIFFERENCES IN THE VOLUME AND SURFACE ROUGHNESS OF THE SHAPE MODELS OF THE RIGHT HIPPOCAMPUS, GENERATED BY THE MODELING METHODS, BETWEEN GROUPS VIA A TWO-TAILED WILCOXON RANK SUM TEST

	Control-MCI	Control-AD	MCI-AD
Scaled Vol. (Progressive)	**< 0.001**	**< 0.001**	**0.036**
Scaled Vol. (LDDMM-TI)	**< 0.001**	**< 0.001**	**0.046**
Scaled Vol. (SPHARM-PDM)	**< 0.001**	**< 0.001**	**0.031**
Roughness (Progressive)	**< 0.001**	**< 0.001**	**< 0.001**
Roughness (LDDMM-TI)	**< 0.001**	**< 0.001**	**0.001**
Roughness (SPHARM-PDM)	**0.029**	**< 0.001**	**< 0.001**

Significant values (p<0.05) are highlighted as boldface.
Scaled vol. : volume scaled by ICV

TABLE V
COMPARISON OF THE SHAPE SIMILARITY OF THE SHAPE MODELS TO THE INPUT BINARY MASKS BETWEEN THE MODELING METHODS

	DC	MD	HD
LDDMM-TI vs. Progressive	**-8.887[+], < 0.001**	**-10.602[-], < 0.001**	**-10.031[-], < 0.001**
SPHARM-PDM vs. Progressive	-0.562[+], 0.574	**-10.328[-], < 0.001**	**-5.366[-], < 0.001**

Significant values (p<0.05) are highlighted as boldface.
Values in each cell: Z and p value (two-tailed) of a Wilcoxon signed rank test
[p]: positive ranks, [n]: negative ranks
DC.: Dice coefficient, MD.: mean distance, HD.: Hausdorff distance

shape models and input binary masks than other methods (Table V).

Fig. 11 presents shape models of large volume differences with the voxel meshes of the input binary masks to explain the limitations of the modeling methods. The Gaussian smoothing operation for SPHARM-PDM was not good to preserve the hippocampal shape details while removing the high frequency noises. It just smoothed out the local shape features and generated a blunt shape of hippocampus (regions where red arrows point in Fig. 11). This tendency was also confirmed via a linear regression between the surface roughness and the volume difference. From the regression analysis, we found that the volume difference is in direct ratio to the surface roughness ($\beta = 381.364$, p < 0.001). The volume differences between LDDMM-TI models and input masks were originated from insufficient template propagation to the image boundary (fourth

(min.: -323, max.: -61 mm^3). On the contrary, our progressive model deformation produced similar volume to the input binary masks in a small range of volumetric difference (min.: -92, max.: 175 mm^3) [Fig. 10(b)] against the large variations in the volume and image quality across the input binary masks. Our progressive model deformation also showed better shape similarity (higher DC and lower MD and HD) between the
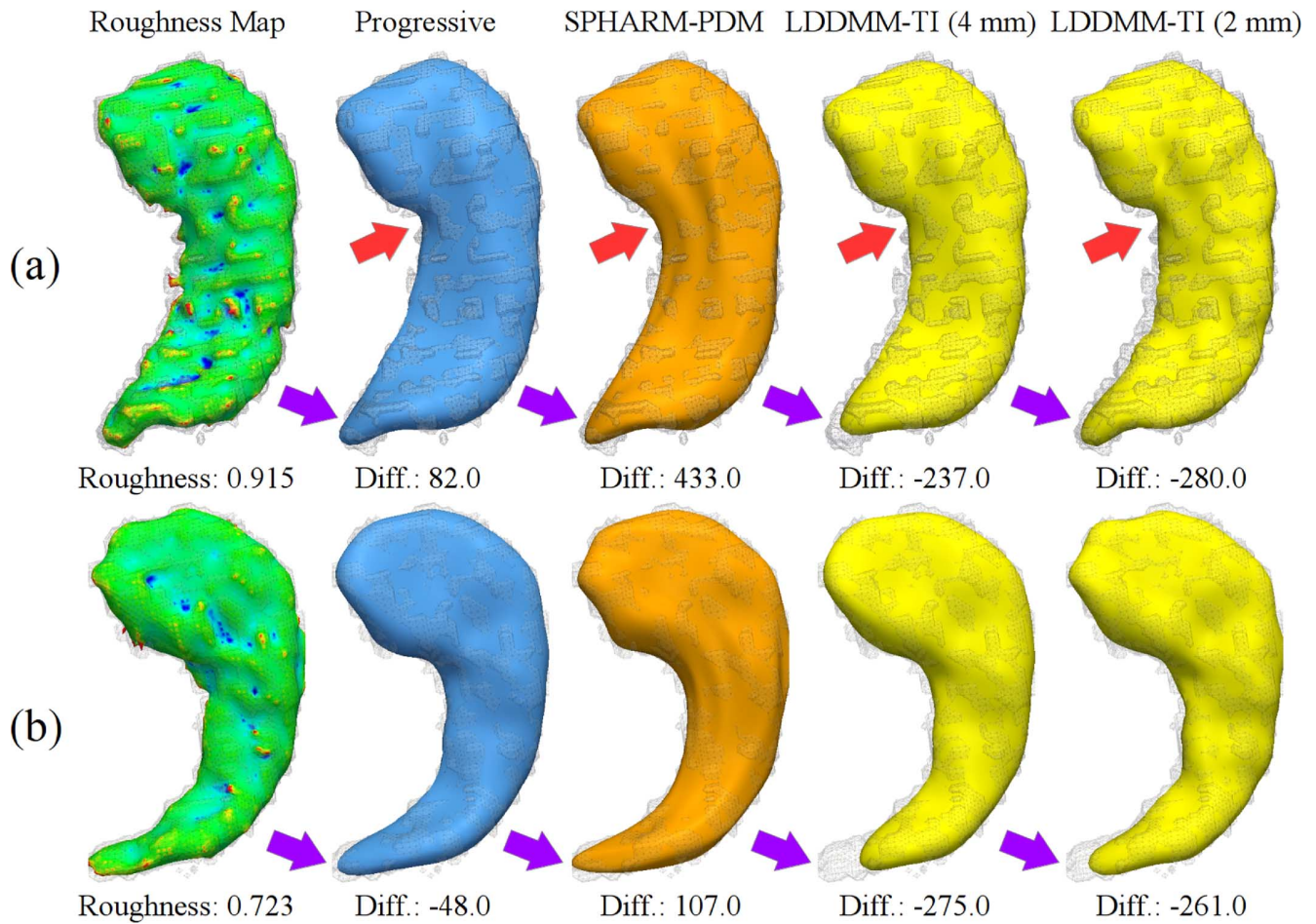
Fig. 11. Example of hippocampal shape models of large volume differences to the input masks. Diff.: model volume–volume of input masks (mm^3).

and fifth columns in Fig. 11). Because the LDDMM volume registration regulates the smoothness of the velocity field using the size of Gaussian kernel, it was expected that smaller values for the standard deviation of the Gaussian kernel can produce better shape similarity. The lower values for Gaussian kernel size produced more deformation of the template volume to unrestored parts of the hippocampus (e.g., hippocampal tail (purple arrows) in Fig. 11), but at the same time, other regions of the shape model were getting more rough by propagating the template model closer to the rough image boundary. This characteristic of the LDDMM volume registration makes hard to control the overall smoothness of the shape model. On the contrary, our progressive model deformation reconstructed the hippocampal shape features properly while filtering out the sharp outliers, owing to the large-to-small scale deformation and vertex-wise boundary search. Fig. 12 showed that our method restored the major features (e.g., hippocampal head and tail) of hippocampus in a large-scale deformation using large values of $\kappa_{\min}$ and that the local details can be reconstructed via a local vertex deformation across the surface using small values of $\kappa_{\min}$. The average computation time of our method using the template model of 2463 vertices and three-ring as the maximum level of the neighborhood on a PC with Intel Core Quad Core 2.3 GHz and 4 GB memory for all subjects was 1.12 ± 0.38 min. The computation time of our method is pro-
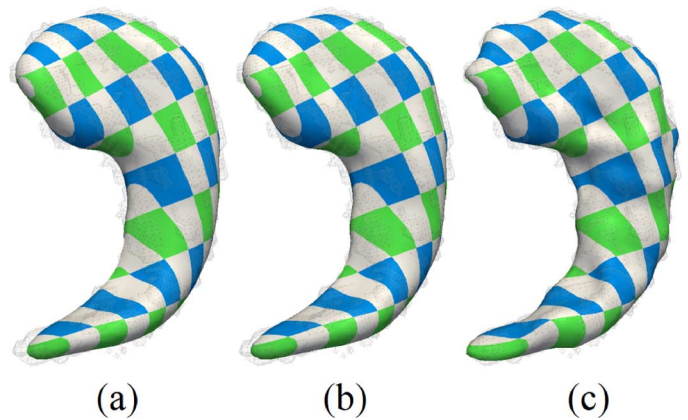

Fig. 12. Hippocampal shape models with different $\kappa_{\min}$. For all cases, κ_{init} was 20.0. (a) $\kappa_{\min} = 20.0$ (roughness: 0.167), (b) $\kappa_{\min} = 9.0$ (roughness: 0.190), and (c) $\kappa_{\min} = 2.0$ (roughness: 0.237). Texture colors indicate the corresponding vertices between surface models.

portional to the number of vertices and the maximum level of the neighborhood, due to the linearity of the Laplacian surface deformation. When we increased the number of vertices in the template model as 4002, the computation time was increased about two times. For SPHARM-PDM, the time was 1.31 ± 0.25 and for LDDMM-TI it was 4.68 ± 1.91 min .
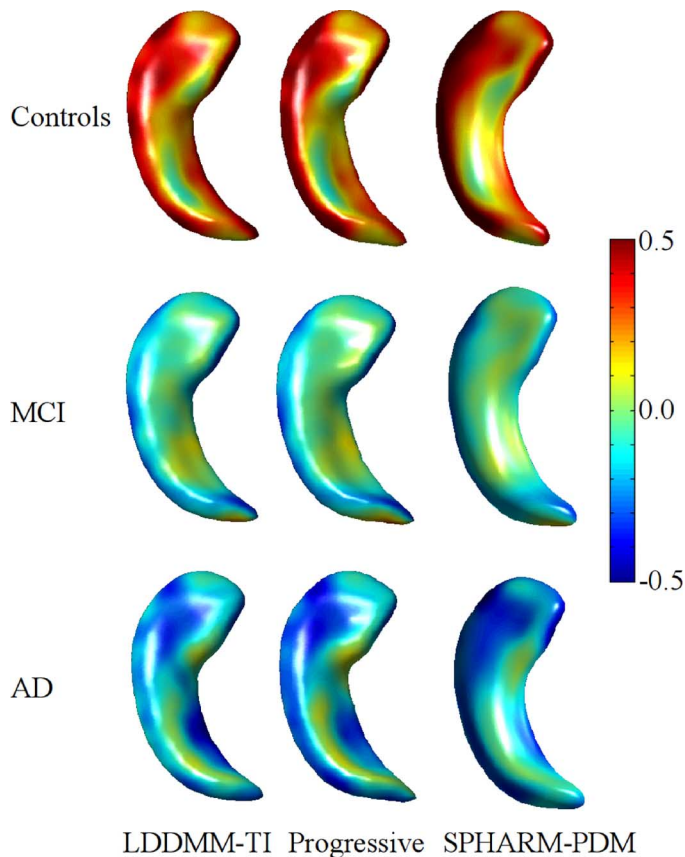
Fig. 13. Distributional shape deformation average maps of the left hippocampi for cognitively normal controls, MCI and AD individuals, obtained (from left to right) from LDDMM volume registration with template injection (LDDMM-TI), our progressive model deformation and SPHARM-PDM.
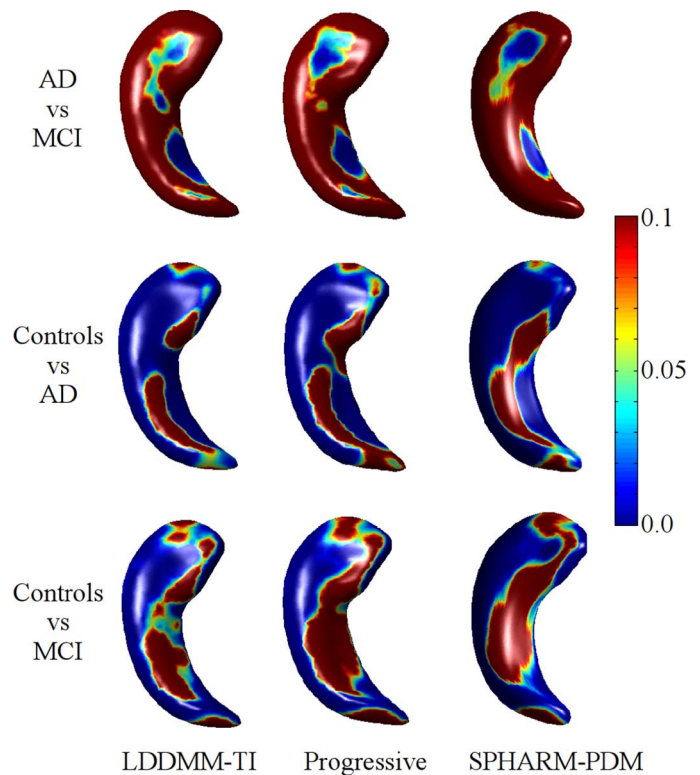

Fig. 14. Results of the Wilcoxon's rank sum test for the left hippocampi, obtained from LDDMM volume registration with template injection (LDDMM-TI), our progressive model deformation and SPHARM-PDM (from left to right, respectively).

2) Average Displacements per Group: Fig. 13 shows the patterns of the average displacements with respect to mean shape for the left hippocampi of individuals from the ADNI database grouped according the Clinical Demential Rating Scale in normal controls, MCI and AD obtained, from left to right, by LDDMM-TI, our progressive model deformation and SPHARM-PDM. The three methods produced almost identical results. The same pattern was observed for the right hippocampi.

3) Nonparametric Analysis of Significant Deformations Between Groups (Wilcoxon's Rank Sum Test): Significant deformations between the AD and MCI groups were only observed in the antero lateral region of the CA1 and in the subiculum, with similar patterns obtained from LDDMM-TI, our progressive model deformation and SPHARM-PDM. The regions with significant displacement differences between controls and MCI and controls and AD, were predominant and excluded the antero-ventral and postero-lateral (fimbria) regions for both (i.e., right and left) hippocampi. Fig. 14 illustrates the results obtained for the left hippocampi.

4) Robust Univariate Linear Regression: Fig. 15 shows, for each hippocampi, the distributional map of the nonstandardized beta coefficients (upper rows) and p-values (bottom rows) obtained by the three methods. Although the patterns of significance obtained by the three methods are similar, our method produced results slightly more similar to LDDMM-TI

than to SPHARM-PDM. The distribution map of standard errors was the same for all methods and ranged uniformly from 0.15 in the dorsal and ventral regions to 0.5 in the lateral regions. The p-value maps were similar to those obtained from the Wilcoxon's rank sum test between control and AD groups.

Interestingly, in all cases, positive associations represented by beta coefficients of nearly 1 (i.e., outwards displacements with respect to the mean shape and higher MMSE scores or inwards displacements and lower MMSE scores) were significant, whilst the absence of associations (i.e., beta coefficients nearly 0) was nonsignificant.

5) ANCOVA Followed by False-Discovery Rate Adjustment: This test produced similar patterns as those obtained from the Wilcoxon's rank sum test (Figs. 14 and 16) with the exception of when AD and MCI groups were compared. In this case, no significance was observed on the deformations between those two groups for any point with our method and neither with SPHARM-PDM. LDDMM-TI showed slight significance ($p = 0.04$) on a reduced region of the postero-ventral subiculum on the left hippocampus, but not at all on the right hippocampus.

V. MODEL VALIDATION ON A DATASET OF A HEALTHY AGING POPULATION

With a clinical dataset from a study of aging, we performed two experiments: 1) evaluation of the accuracy in representing the variations of the hippocampal shape and the anatomical shape correspondence between the individualized models, and 2) evaluation of the sensitivity of the proposed modeling
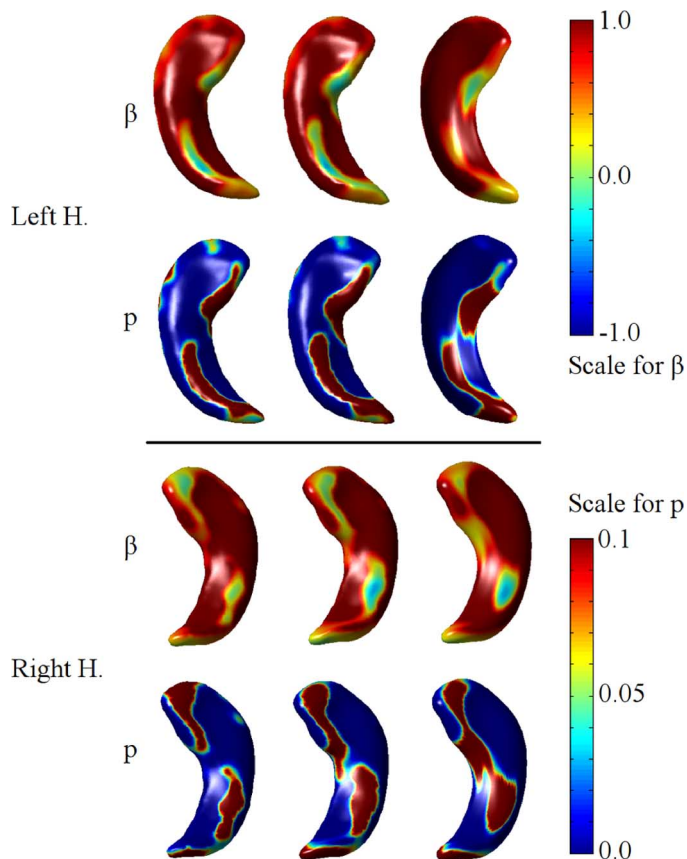
Fig. 15. Maps of the strength and direction [rows above on each panel, represented by the nonstandardized beta coefficients (β)] and significance (rows below on each panel, represented by the p-value) of the association between cognition (represented by the MMSE scores), and local hippocampal shape deformation patterns, for left (upper panel) and right (lower panel) hippocampi obtained, from left to right, by LDDMM volume registration with template injection, our progressive model deformation and SPHARM-PDM. Age and gender were used as covariates.
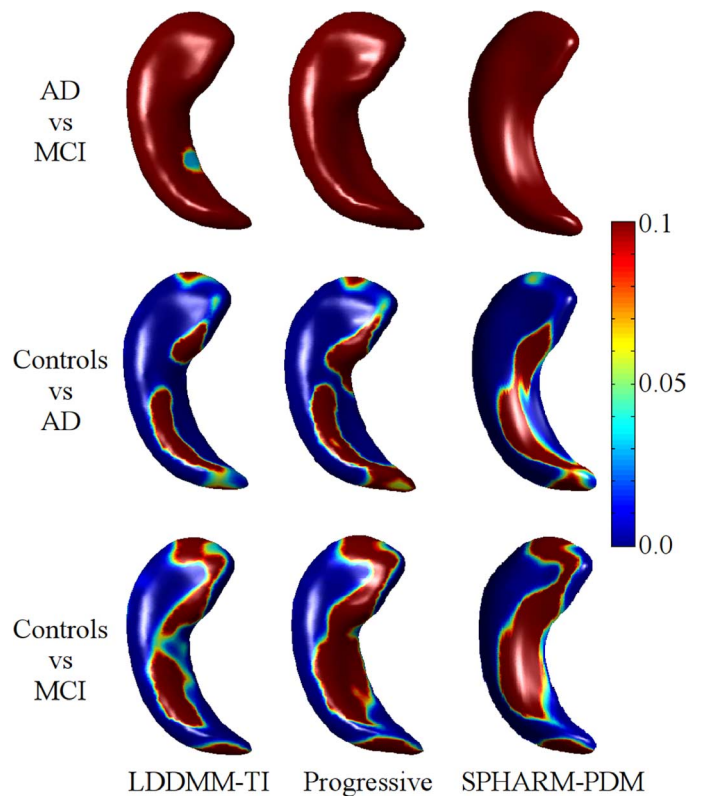

Fig. 16. Results from the analysis of covariance (ANCOVA) followed by false discovery rate correction for the left hippocampi, obtained from LDDMM volume registration with template injection (LDDMM-TI), our progressive model deformation and SPHARM-PDM (from left to right respectively). Age and gender were used as covariates.

method in showing differences related to cognitive performance in nondiseased older people. In this section, we first describe the characteristics of the sample, imaging protocol, data processing and clinical variables of the study of aging. Then, we present the experiment for the model validation of accuracy and shape correspondence on the clinical data. In the next section, we demonstrate the statistical shape analysis of hippocampal morphology with the cognitive variables.

A. Materials and Image Processing

We used MR images and cognitive data from 51 participants (33 women and 18 men) randomly selected from The Lothian Birth Cohort 1936 (LBC1936) Study [60] (www.lothianbirthcohort.ed.ac.uk). The data used in our experiments was obtained at individuals' mean age of 72.7 years (standard deviation = 0.7 years). All study participants included in our sample had Mini Mental State Examination scores > 24 (i.e., limit considered to be indicative of possible pathological cognitive impairment). Written informed consent was obtained from all participants under protocols approved by the Lothian (REC 07/MRE00/58) and Scottish Multicentre (MREC/01/0/56) Research Ethics Committees. Study participants were imaged

in a GE Signa HDxt 1.5T clinical scanner (General Electric, Milwaukee, WI, USA) with a manufacturer supplied 8-channel phased-array head coil following a scanning protocol described in detail elsewhere [61]. The T1-weighted MR sequence, used to generate the hippocampal binary masks, was acquired in coronal orientation and had inversion, echo and repetition times of 500/4/9.8 ms, respectively, flip angle of $8°$, slice thickness of 1.2 mm, bandwidth of 122 Hz/pixel and voxel dimension of $1 \times 1 \times 1.3$ mm^3

A first approximation of the left and right hippocampal segmentation was obtained fully automatically in the same way as it was done for the ADNI dataset (Section IV-B). The results were visually assessed by a trained image analyst and manually corrected using Analyze 10.0 software (www.analyzedirect.com) when required [61]. The resulting segmented areas were saved as binary masks and the volumes were quantified and expressed in mm^3. We also measured the head size of individuals. Intracranial volume (i.e., contents within the inner skull table including brain tissue, cerebrospinal fluid, veins, and dura), with an inferior limit on the axial slice just superior to the tip of the odontoid peg at the foramen magnum was extracted semi-automatically using the T2*-weighted sequence, with the Object Extraction Tool in Analyze followed by manually editing [62].

All study participants underwent a wide range of cognitive tests as described in [60]. We used the following cognitive data: general factors of cognitive ability (g), processing speed (g-speed), and memory (g-memory) at mean age 72.7

TABLE VI
Shape Similarity of the Shape Models, Reconstructed From a Dataset of a Healthy Aging Population (mean ± SD)

	Left Hippocampus	Right Hippocampus
Dice Coefficient	0.951 ± 0.008	0.952 ± 0.007
Mean Distance	0.439 ± 0.028	0.451 ± 0.033
Hausdorff Distance	2.384 ± 0.713	2.533 ± 0.663

(±0.7 SD) years. These were generated using principal component analysis by members of The LBC1936 Study Group (see author's list) as described elsewhere [63]. Briefly, to derive g six subtests of the WAIS-III^{UK} [64] (Digit Symbol, Digit Span Backward, Symbol Search, Letter-Number Sequencing, Block Design & Matrix Reasoning) were used. g-memory was derived from five subtests from the WMS-III^{UK} [65] (Logical Memory Total Immediate & Delayed Recall, Verbal Paired Associates Immediate & Delayed Recall & Spatial Span Total Score) and two subtests from the WAIS-III^{UK} (Letter-Number Sequencing & Digit Span Backward). g-speed was obtained from two reaction time tests (Simple Reaction Time & Choice Reaction Time), an Inspection Time test and two WAIS-III^{UK} subtests (Digit Symbol & Symbol Search). These cognitive variables (i.e., g, g-memory, and g-speed) were continuous and normally distributed throughout the sample. Gender and age in days at the time of the cognitive tests and at the time of the MRI scan were also used in the statistical analysis. For the hippocampal binary masks of the selected subjects, we built the template models for left and right hippocampi following the process described in Sections II-A and IV-B. Then, we performed the shape modeling using our progressive model deformation technique and measured the shape similarity of the shape models with the input binary masks. Table VI showed the measured values of the shape similarity. The volume overlap between the surface models and the target structures in the segmentations was 95% for left and right hippocampus. The average MD from all subjects was 0.442 mm, i.e., less than half the voxel spacing of the segmentations ($1.0 \times 1.0 \times 1.3$ mm^3). The average HD was 2.458 mm.

B. Evaluation of Anatomical Shape Correspondence

To investigate the shape correspondence between the individualized surface models, we performed an anatomical landmark test that included manual landmark assignments and fiducial localization error (FLE) [66] measurements across subjects. For this landmark test, we selected seven representative datasets: three datasets from subjects who had the largest hippocampal volume, three with the smallest and one with medial hippocampal volume. A trained analyst manually placed six representative anatomical landmarks on each of the left and right hippocampal models of the selected subjects (seven subjects × 12 landmarks). Each time, the landmark placement was done blind to the placements on previous subjects. To account for intra-operator variability the analyst performed the landmark assignment three times, and the average positions of each landmark were used for this evaluation. We determined two sources of errors: 1) intra-operator variability and 2) inter-surface variability. The intra-operator variability was

measured by the FLE expression as (9). The six positional coordinates of the corresponding landmarks on each of the surface models can be expressed as: $\mathbb{L}_i = \{l_i^1, \ldots, l_i^7\}$ where l_i^j is the average position of ith landmark on the surface model of the jth subject. To assess the inter-surface variability, we projected the landmarks $\mathbb{L}_i$ onto the surface models of the selected subjects (seven subjects) and the mean surface model (this last obtained from all subjects) separately. Then we measured the FLE with respect to the projected points of the landmarks $\mathbb{L}_i$. This FLE value indicated the positional variability of the landmarks on each surface model, and was formulated as

$$\text{FLE}_i = \sqrt{\frac{1}{k-1} \sum_{j=1}^{k} \|p_j - \bar{p}_j\|^2} \qquad (9)$$

where p_j are the positional coordinates (x, y, z) of a landmark on a surface model and $\bar{p}_j$ is $(1/k) \sum_{j=1}^{k} p_j$. For the intra-operator variability, k is the number of manual placements ($k = 3$), and p_j are the positional coordinates of the landmark l_i at the jth manual assignment. For the inter-surface variability, k is the number of subjects ($k = 7$), and p_j is the projected position of a landmark (l_i^j) of the jth subject in the set $\mathbb{L}_i$. For the inter-subject variability, the FLE value for $\mathbb{L}_i$ on a surface model represents the localization error of the landmark assignment as the average distance from the projected points (p_j) of the landmarks in $\mathbb{L}_i$ to their centroid ($\bar{p}_j$) on the surface model and is related to the operator's observation of the various shapes of the hippocampal surface models. To quantitatively investigate the similarity of FLEs between the surface models, we also measured the difference between the maximum and minimum FLEs of $\mathbb{L}_i$ across the surface models. This difference was considered as the localization error caused by the hippocampal modeling method. Consequently, it can be a measure of the robustness of our modeling method in preserving the correspondence of the anatomical landmarks in $\mathbb{L}_i$ between subjects. As FLE is based on Euclidean distance, the size of the hippocampal surface models can lead to bias while investigating the inter-surface variability. To avoid this bias, we used the surface models that were normalized by the hippocampal volumes.

Fig. 17(a) presents 12 anatomical landmarks on left and right hippocampal surface models (i.e., six on each hippocampi) of a subject with large hippocampal volume. In order to access the intra-operator variability, we measured the FLE of each manual landmark assignment for 14 hippocampal surface models (left and right surface models of seven subjects). The average FLE was 0.560 ± 0.282 mm. The error range of the manual assignment was from 0.076 mm to 1.839 mm. The average FLE for each landmark are listed in Table VII. For the inter-surface variability, Fig. 17 shows the projected landmarks (yellow dots) and their average positions (blue dots) on the surface models. The positions of the projected landmarks in each set ($\mathbb{L}_i$) are close to each other and to their average positions. This result was consistent across the surface models. The maximum, minimum and error range of the projected landmarks are listed in Table VIII. The minimum FLE for the landmark projection on the surface models was observed at the tip of the right hippocampal tail (0.733 mm, R#1 in Table VIII). The maximum value was observed at the dorsal joint between the cornu ammonis (CA1)
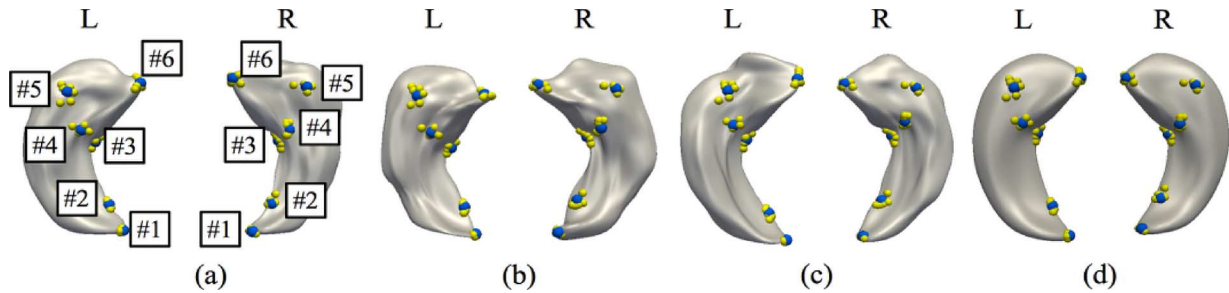
Fig. 17. Anatomical landmarks manually assigned to investigate the inter-subject shape correspondence, placed on four representative surfaces. Landmarks are located in: tip of the tail (L#1), inferior joint between cornu ammonis 1 (CA1) and subiculum (L#2), ventral joint between CA1 and subiculum (L#3) at the base of the hippocampus head, dorsal joints between CA1, subiculum, and fimbria (L#4 and L#5), tip at the top of the hippocampal head in the boundary between presubiculum and CA1 (L#6). A trained analyst performed the manual landmark assignment three times. Yellow dots show the average position of the assigned landmarks on a surface model, and blue dots are the centers for calculating the fiducial localization error values. (a) Large #01. (b) Median. (c) Small #01. (d) Mean Surface.

TABLE VII
FLE FOR INTRA-OPERATOR VARIABILITY

	L#1	L#2	L#3	L#4	L#5	L#6
FLE	0.386 ± 0.250	0.506 ± 0.246	0.435 ± 0.213	0.482 ± 0.248	0.814 ± 0.579	0.557 ± 0.141

	R#1	R#2	R#3	R#4	R#5	R#6
FLE	0.491 ± 0.124	0.521 ± 0.256	0.769 ± 0.295	0.554 ± 0.256	0.666 ± 0.176	0.541 ± 0.169

Data for FLE: mean ± SD.
Unit of FLE: mm

TABLE VIII
FLE FOR INTER-SURFACE VARIABILITY

	L#1	L#2	L#3	L#4	L#5	L#6
Max.	1.446	1.421	2.046	1.597	2.275	1.838
Min.	0.806	1.107	1.569	1.327	2.150	1.502
Diff.	0.640	0.314	0.477	0.270	0.125	0.336
Avg.	1.049 ± 0.201	1.264 ± 0.129	1.739 ± 0.197	1.434 ± 0.094	2.203 ± 0.041	1.677 ± 0.117

	R#1	R#2	R#3	R#4	R#5	R#6
Max.	1.003	1.672	1.713	1.691	1.659	1.723
Min.	0.733	1.317	1.345	1.375	1.456	1.298
Diff.	0.270	0.355	0.368	0.316	0.203	0.425
Avg.	0.849 ± 0.079	1.518 ± 0.127	1.531 ± 0.106	1.517 ± 0.100	1.560 ± 0.069	1.494 ± 0.175

Max.: Maximum FLE, Min.: Minimum FLE, Diff.: Maximum FLE – Minimum FLE, and Avg.: Average FLE
Data for Avg.: mean ± SD.
Unit of FLE: mm

and the fimbria on the left hippocampus (2.275 mm, L#5 in Table VIII). The FLE difference ranged from 0.125 mm (L#5 in Table VIII) to 0.640 mm (R#1 in Table VIII). The FLE difference for the inter-surface variability was similar to the average FLE for the intra-operator variability (i.e., operator error during the manual placement).

C. Hippocampal Shape Analysis With Cognitive Variables

1) Statistical Analysis Design: To assess the feasibility of our method in the detection of subtle morphological deformations of the hippocampus on normal older individuals, we performed statistical shape analyses using the individualized models and the cognitive parameters. For this analysis, we performed the individual shape modeling, model normalization using scale transformation and GPA, and shape deformity computation.

With the shape deformity maps for individuals, we performed two analyses: 1) Robust univariate linear regression to evaluate the association between local hippocampal shape deformations and cognitive abilities, and determine the strength and significance of these associations (identical as statistical test three using the ADNI dataset, see Section IV-C4), and 2) Comparative analysis on extreme cognitive groups to explore the magnitude and direction of these deformation patterns (similar as statistical test 2 using the ADNI dataset, Section IV-C3). For the first analysis, which included the whole sample, we considered a regression model to determine how much cognition at age 72 can be explained by local hippocampal deformations. This model considered as independent variable the hippocampal shape deformities at each vertex normalized by intracranial volume, and as dependent variable a cognitive measure determined at age 72 years. As the three cognitive variables evaluated (namely g, g-speed, and g-memory) have high degree of collinearity, they were analyzed separately. Age and gender were used as covariates. The robust multilinear regression was performed using iteratively reweighted least squares with a bisquare weighting function through the function "robustfit" from MATLAB Statistical Toolbox.

For the second analysis, we selected the subjects for which the values of the cognitive parameters were greater or less than a standard deviation (SD) from the mean (i.e., poor versus good cognitive performance within the normal range). The characteristics of the total sample and number of subjects on each group appear in Table IX. A p-value map was generated using the function "ranksum" from MATLAB Statistical Toolbox to determine whether the local shape differences between groups of extreme cognitive performance were or not significant. This function performs a two-sided rank sum test equivalent to a Mann-Whitney U-test. We also tested for significance in the volumetric difference of each hippocampus (i.e., left and right) between each group using the Mann-Whitney U-test from the statistical package PASW Statistics 18.

D. Regression Analysis in the Whole Sample

Local hippocampal deformations were a strong predictor of cognitive abilities in old age. While local predictors of cogni-
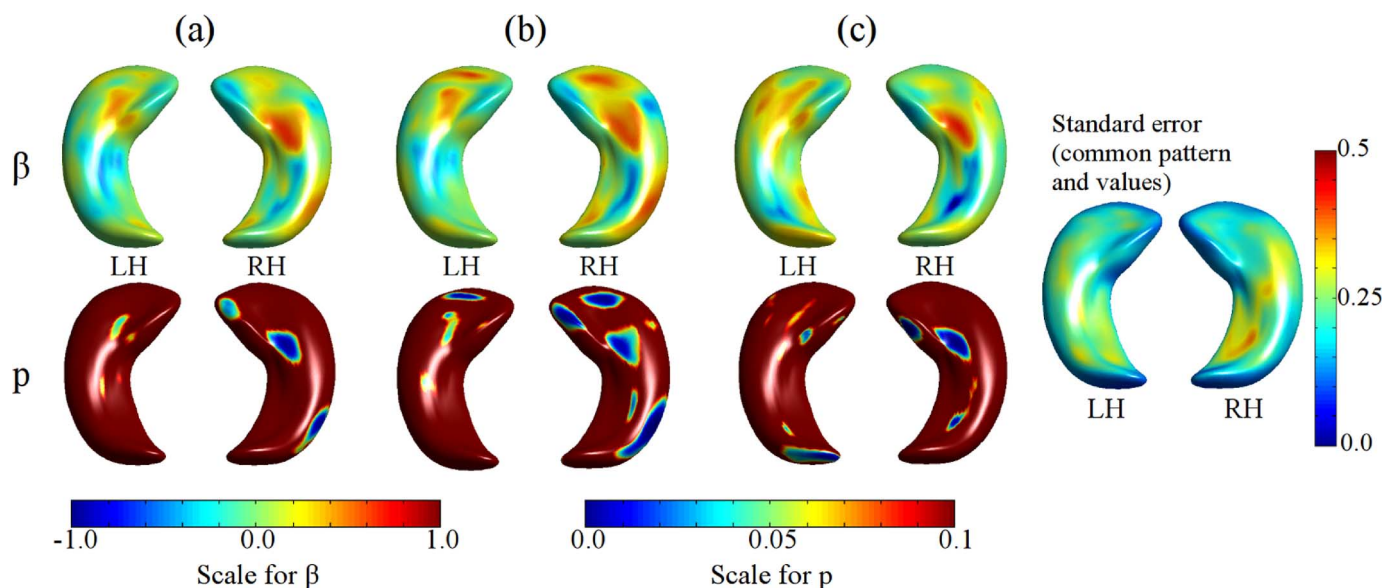
Fig. 18. Maps of the strength and direction (above, represented by the nonstandardized beta coefficients) and significance (below, represented by the p-value) of the association between the cognitive abilities at age 72, represented by g, g-memory and g-speed, and local hippocampal shape deformation patterns, for left (LH) and right (RH) hippocampi. Distribution map of standard errors was the same across all tests.
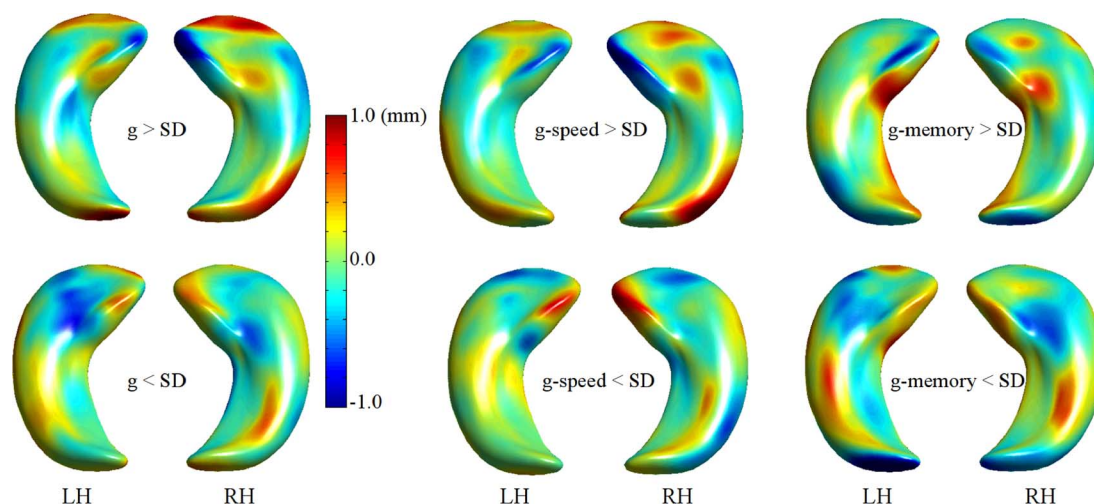

Fig. 19. Maps of the deformation patterns of the hippocampal shape for the subject groups with different cognitive abilities at old age. These were obtained correcting for hippocampal size. Color scale represents the magnitude of the deformation (inwards blue and outwards red) with respect to the mean shape.

TABLE IX
SAMPLE CHARACTERISTICS AND NUMBER OF SUBJECTS ON EACH TEST GROUP

Variable name	mean	Standard deviation	No. of subjects with [variable] < 1 SD	No. of subjects with [variable] > 1 SD
g	0.086	0.950	8	8
g-memory	-0.043	0.944	7	9
g-speed	0.057	0.939	9	8

TABLE X
RANGE OF THE NONSTANDARDIZED REGRESSION BETA COEFFICIENTS
(β) AND STANDARD ERRORS (STD ERROR) OBTAINED FROM
EACH REGRESSION ANALYSIS

Association examined	Left Hippocampus	Right Hippocampus
Local deformations and general cognitive abilities at age 72 (g)	$-0.556 \leq \beta \leq 0.469$ $0.088 \leq$ std error $\leq$ 0.502	$-0.659 \leq \beta \leq 0.658$ $0.093 \leq$ std error $\leq$ 0.500
Local deformations and processing speed at age 72 (g-speed)	$-0.501 \leq \beta \leq 0.614$ $0.085 \leq$ std error $\leq$ 0.481	$-1.074 \leq \beta \leq 0.588$ $0.089 \leq$ std error $\leq$ 0.500
Local deformations and general memory at age 72 (g-memory)	$-0.599 \leq \beta \leq 0.638$ $0.086 \leq$ std error $\leq$ 0.473	$-1.055 \leq \beta \leq 0.703$ $0.094 \leq$ std error $\leq$ 0.500

tive performance were accentuated in the dorso-lateral region of the head for the left hippocampus, they were consistently accentuated in the fimbria and antero-ventral region in the right hippocampus as Fig. 18 shows. Results from the group analysis show consistency and coincidence on the anatomical localization of the regions where the local deformations were a strong predictor of the cognitive variables (Figs. 18 and 19). The ranges

for the nonstandardized beta coefficients and standard errors for each analysis are given in Table X.

E. Shape Analysis in Groups With Extreme (Opposite) Cognitive Performance

Hippocampal shapes from the extreme groups within the normal aging sample showed strong differences on specific regions of both hippocampi (left and right) for all cognitive variables at age 72 years. The patterns were similar when the shape models were corrected by hippocampal or head size. The lateral anterior part of the joint between the CA1 with the fimbria on the right hippocampus had pronounced deformations for subjects with low versus high cognitive performance at old age. For the right hippocampus, the CA1 deformations were opposite and accentuated in the fimbria and for the dorso-lateral tail and ventral and dorsal head, these last mainly for g and g-speed. The direction and magnitude of the deformations are shown in Fig. 19 while the strength and significance of their association with cognitive abilities can be appreciated in Fig. 18. Regions of significance obtained from the two-sided rank sum test were, as expected, coincident with those obtained from the univariate regression analysis on the whole sample. For the left hippocampus, significant deformations in specific head (CA1) regions were similar for g and g-speed. The regions of the left fimbria that had accentuated deformations for g-speed were not the same as those for g-memory. The hippocampal volumes (left, right and total) of participants with g, g-speed, and g-memory less than a SD from the mean were not significantly different from those for which these cognitive parameters were higher than a SD from the mean.

VI. DISCUSSION

Our shape modeling framework allows accurate adjustments of the hippocampal surface models on a population of healthy aging people with an accuracy of 95% volume overlap. On the synthetic dataset, our method was comparable with SPHARM-PDM, ShapeWorks and LDDMM-TI, although with our method better results were obtained across a wider range of the smoothing parameter than with these methods (see Fig. 7). On the ADNI dataset, the volumetric difference between the shape models generated by our method and the binary masks was nearly 0 for controls, MCI and AD groups. LDDMM-TI and SPHARM-PDM showed higher volumetric differences across a wider range of values for all groups. Our model-to-image registration method optimally reproduces the shape of the hippocampus minimizing the geometrical differences with the target structures obtained from the segmentations.

Hippocampal segmentations were affected by image noise (i.e., holes), rough boundaries, sharp edges and inaccuracies due to MRI artifacts, slice thickness and inter-slice gaps, all which may distort the reproduction of the local shape details and mask the true differences among shapes. SPHARM-PDM applies the fill-hole and Gaussian smoothing operations to obtain the initial surface models with spherical topology from the segmentations [29]. However, these operations could not separate the shape noises from the hippocampal surfaces effectively, and they generated smooth hippocampal surface models embracing the shape noises and the rough boundaries. This partially explains why the surface models of SPHARM-PDM were less similar to the manual segmentations than the models generated by LDDMM-TI and our method [see Fig. 10(b)].

A previous study showed that the volumetric changes on hippocampal atrophy are modest with rate of progression over time relatively slow, depending mainly on age differences [67]. In agreement with this study, in our age-homogeneous sample we did not find significant differences in hippocampal volumes between subjects with extreme cognitive outcomes. With the radial distance mapping technique used by [67], the morphological shape variations between Alzheimer's, mild cognitive impaired and normal subjects were diffusively distributed throughout the volumes and it was not possible to conclude on local shape differences amongst groups of subjects. The regions where our model identified statistically significant differences amongst groups of dissimilar cognitive performance are included within the regions with similar differences on [67]. Moreover, our model shows a neater local deformation pattern, similar to the deformation patterns obtained from using LDDMM-TI and SPHARM-PDM.

In the experiments, we reproduced the ANCOVA analysis published by Cho *et al.* [12] on the ADNI dataset, the results obtained, although similar, are not identical. Several factors caused these differences. Firstly, we did not use the segmentations provided by the ADNI group. These were done in standard space and there is evidence that quantitative analysis on standard space introduces more partial volume effect errors than if the analysis is done on native space, biasing the results [68]. Secondly, we obtained the hippocampal binary masks on this dataset using FSL-FIRST followed by manual rectification of the incorrectly segmented boundaries. This method has been extensively tested and validated, and was applied to the aging sample from the LBC1936 Study, allowing comparability of the results. And thirdly, it is not guaranteed that we used the same sample as the one used by Cho *et al.* [12].

The strengths of this study are: 1) the shape modeling development that allowed to robustly preserve the individual shape details across the large variability characteristic of an aging population while preserving the anatomical correspondence across subjects, 2) the analysis performed on a normal and relatively healthy aging sample of community-dwelling individuals that allowed us to test the sensitivity of the shape modeling framework developed, 3) the analyses using a subsample from the ADNI dataset, which facilitates the comparison of the results with those obtained by other studies and 4) the fitting of the mesh model to the reference binary masks was done fully automatically, which guarantees the reproducibility of this approach to hippocampal shape analysis on other datasets. This shape modeling approach could be applied to other subcortical structures, such as caudate and thalamus.

Due to sample size constraints, the patterns of hippocampus local deformations found here for each subgroup cannot be generalized. Nevertheless our results show that the shape modeling framework developed here can be used in older people to detect subtle specific hippocampal regional morphological variations.

VII. CONCLUSION

In this paper, we propose a mesh-to-volume registration framework based on the progressive model deformation, built-up on a multi-level neighborhood and flexible weighting scheme for modeling the hippocampal shapes. It incorporates a

large-to-small scale deformation to allow robustly restore the individual shape details of the hippocampus while maintaining the anatomical point-to-point correspondence across subjects. Moreover, it can be argued that the large-to-small scale deformation helps to control the surface quality (the level of shape detail representation) across surface while restoring the major shape features of hippocampus properly. Our experiment results indicate that our modeling method is accurate and robust against noisy segmentations and large size and shape variations in hippocampus, and is sensitive in detecting subtle morphological changes on the structure.

REFERENCES

[1] D. Y. Lee *et al.*, "Sub-regional hippocampal injury is associated with fornix degeneration in Alzheimer's disease," *Front. Aging Neurosci.*, vol. 4, pp. 1–10, 2012.

[2] P. A. Thomann *et al.*, "Hippocampal morphology and autobiographic memory in mild cognitive impairment and Alzheimer's disease," *Curr. Alzheimer Res.*, vol. 9, no. 4, pp. 507–515, May 2012.

[3] F. Bai *et al.*, "Aberrant hippocampal subregion networks associated with the classifications of aMCI subjects: A longitudinal resting-state study," *PLoS ONE*, vol. 6, no. 12, p. e29288, 2011.

[4] D. V. Moretti *et al.*, "Specific EEG changes associated with atrophy of hippocampus in subjects with mild cognitive impairment and Alzheimer's disease," *Int. J. Alzheimers Dis.*, vol. 2012, p. 253153, 2012.

[5] F. M. Benes, I. Sorensen, and E. D. Bird, "Reduced neuronal size in posterior hippocampus of schizophrenic patients," *Schizophr. Bull.*, vol. 17, no. 4, pp. 597–608, 1991.

[6] M. Styner, J. A. Lieberman, D. Pantazis, and G. Gerig, "Boundary and medial shape analysis of the hippocampus in schizophrenia," *Med. Img. Anal.*, vol. 8, no. 3, pp. 197–203, Sep. 2004.

[7] R. E. Hogan *et al.*, "MRI-based high-dimensional hippocampal mapping in mesial temporal lobe epilepsy," *Brain*, vol. 127, pp. 1731–1740, Aug. 2004.

[8] T. M. Doring *et al.*, "Evaluation of hippocampal volume based on MR imaging in patients with bipolar affective disorder applying manual and automatic segmentation techniques," *J. Magn. Reson. Imag.*, vol. 33, no. 3, pp. 565–572, Mar. 2011.

[9] T. Gattringer *et al.*, "Vascular risk factors, white matter hyperintensities and hippocampal volume in normal elderly individuals," *Dement Geriatr Cogn. Disord*, vol. 33, no. 1, pp. 29–34, 2012.

[10] H. Kim, P. Besson, O. Colliot, A. Bernasconi, and N. Bernasconi, "Surface-based vector analysis using heat equation interpolation: A new approach to quantify local hippocampal volume changes," *Med. Image Comput. Comput. Assist. Intervent.*, vol. 11, pp. 1008–1015, 2008.

[11] G. B. Frisoni *et al.*, "Mapping local hippocampal changes in Alzheimer's disease and normal ageing with MRI at 3 Tesla," *Brain*, vol. 131, no. Pt 12, pp. 3266–3276, Dec. 2008.

[12] Y. Cho *et al.*, "A multi-resolution scheme for distortion-minimizing mapping between human subcortical structures based on geodesic construction on Riemannian manifolds," *NeuroImage*, vol. 57, no. 4, pp. 1376–1392, Aug. 2011.

[13] B. Patenaude, S. M. Smith, D. N. Kennedy, and M. Jenkinson, "A Bayesian model of shape and appearance for subcortical brain segmentation," *NeuroImage*, vol. 56, no. 3, pp. 907–922, Jun. 2011.

[14] G. Gerig *et al.*, "Age and treatment related local hippocampal changes in schizophrenia explained by a novel shape analysis method," *Med. Image Comput. Comput. Assist. Intervent.*, vol. 2879, pp. 653–660, 2003.

[15] L. Shen, J. Ford, F. Makedon, and A. Saykin, "Hippocampal shape analysis: Surface-based representation and classification," in *Proc. SPIE Med. Imag.*, 2003, pp. 253–264.

[16] M. Styner, G. Gerig, J. Lieberman, D. Jones, and D. Weinberger, "Statistical shape analysis of neuroanatomical structures based on medial models," *Med. Img. Anal.*, vol. 7, no. 3, pp. 207–220, Sep. 2003.

[17] S. Bouix, J. C. Pruessner, D. Louis Collins, and K. Siddiqi, "Hippocampal shape analysis using medial surfaces," *NeuroImage*, vol. 25, no. 4, pp. 1077–1089, May 2005.

[18] J. G. Csernansky *et al.*, "Preclinical detection of Alzheimer's disease: Hippocampal shape and volume predict dementia onset in the elderly," *NeuroImage*, vol. 25, no. 3, pp. 783–792, Apr. 2005.

[19] N. A. Lord, J. Ho, B. C. Vemuri, and S. Eisenschenk, "Simultaneous registration and parcellation of bilateral hippocampal surface pairs for local asymmetry quantification," *IEEE Trans. Med. Imag.*, vol. 26, no. 4, pp. 471–478, Apr. 2007.

[20] A. Qiu and M. I. Miller, "Multi-structure network shape analysis via normal surface momentum maps," *NeuroImage*, vol. 42, no. 4, pp. 1430–1438, Oct. 2008.

[21] J. Pluta *et al.*, "Appearance and incomplete label matching for diffeomorphic template based hippocampus segmentation," *Hippocampus*, vol. 19, no. 6, pp. 565–571, Jun. 2009.

[22] P. R. Raamana *et al.*, "Three-class differential diagnosis among Alzheimer disease, frontotemporal dementia, and controls," *Front. Neurol.*, vol. 5, pp. 1–15, 2014.

[23] R. H. Davies, C. J. Twining, P. D. Allen, T. F. Cootes, and C. J. Taylor, "Shape discrimination in the hippocampus using an MDL model," *Inf. Process. Med. Imag.*, vol. 18, pp. 38–50, Jul. 2003.

[24] J. Cates, P. T. Fletcher, M. Styner, M. Shenton, and R. Whitaker, "Shape modeling and analysis with entropy-based particle systems," *Inf. Process. Med. Imag.*, vol. 20, pp. 333–345, 2007.

[25] X. Yang, A. Goh, S.-H. A. Chen, and A. Qiu, "Evolution of hippocampal shapes across the human lifespan," *Hum. Brain Mapp.*, vol. 34, no. 11, pp. 3075–3085, Nov. 2013.

[26] A. Qiu, C. Fennema-Notestine, A. M. Dale, and M. I. Miller, "Regional shape abnormalities in mild cognitive impairment and Alzheimer's disease," *NeuroImage*, vol. 45, no. 3, pp. 656–661, Apr. 2009.

[27] S. G. Costafreda *et al.*, "Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment," *NeuroImage*, vol. 56, no. 1, pp. 212–219, May 2011.

[28] S. J. van Veluw *et al.*, "Hippocampal T2 hyperintensities on 7 Tesla MRI," *Neuroimage Clin.*, vol. 3, pp. 196–201, 2013.

[29] M. Styner *et al.*, "Framework for the statistical shape analysis of brain structures using SPHARM-PDM," *Insight J.*, no. 1071, pp. 242–250, 2006.

[30] J. Cates, M. Meyer, T. Fletcher, and R. Whitaker, "Entropy-based particle systems for shape correspondence," in *Proc. 1st MICCAI Workshop Math. Foundat. Computat. Anat.: Geometr., Stat. Registrat. Methods Model. Biol. Shape Variabil.*, Oct. 2006, pp. 90–99.

[31] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *Int. J. Comput. Vis.*, vol. 61, no. 2, pp. 139–157, Nov. 2004.

[32] L. Risser, F.-X. Vialard, R. Wolz, M. Murgasova, D. D. Holm, and D. Rueckert, "Simultaneous multi-scale registration using large deformation diffeomorphic metric mapping," *IEEE Trans. Med. Imag.*, vol. 30, no. 10, pp. 1746–1759, Oct. 2011.

[33] M. Alexa, "Differential coordinates for local mesh morphing and deformation," *Vis. Comput.*, vol. 19, pp. 105–114, 2003.

[34] O. Sorkine *et al.*, "Laplacian surface editing," in *Proc. 2004 Eurograph./ACM SIGGRAPH Symp. Geometry Process.*, New York, Jul. 2004, pp. 175–184.

[35] O. Sorkine, "Differential representations for Mesh processing," *Comput. Graph. Forum*, vol. 25, no. 4, pp. 789–807, 2006.

[36] M. Botsch and O. Sorkine, "On linear variational surface deformation methods," *IEEE Trans. Visualizat. Comput. Graph.*, vol. 14, no. 1, pp. 213–230, Jan./Feb. 2008.

[37] O. Sorkine and M. Alexa, "As-rigid-as-possible surface modeling," in *Eurograph. Symp. Geomet. Process.*, 2007, pp. 109–116.

[38] L. Liu, L. Zhang, Y. Xu, C. Gotsman, and S. J. Gortler, "A local/global approach to mesh parameterization," *Comput. Graph. Forum*, vol. 27, no. 5, pp. 1495–1504, 2008.

[39] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3-D surface construction algorithm," in *Proc. 14th Annu. Conf. Comput. Graph. Interact. Techn.*, Aug. 1987, pp. 163–169.

[40] N. Ray, D. Wang, X. Jiao, and J. Glimm, "High-order numerical integration over discrete surfaces," *SIAM J. Numerical Analys.*, vol. 50, no. 6, pp. 3061–3083, 2012.

[41] M. S. Floater, "Mean value coordinates," *Comput. Aided Geometric Design*, vol. 20, no. 1, pp. 19–27, Mar. 2003.

[42] T. Ju, S. Schaefer, and J. Warren, "Mean value coordinates for closed triangular meshes," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 561–566, 2005.

[43] M. Wardetzky, S. Mathur, and F. Kälberer, "Discrete Laplace operators: No free lunch," in *Eurograph. Symp. Geometry Process.*, 2007, pp. 33–37.

[44] Y. Lipman, J. Kopf, D. Cohen-Or, and D. Levin, "GPU-assisted positive mean value coordinates for mesh deformations," in *Proc. 5th Eurograph. Symp. Geomet. Process.*, Jul. 2007, pp. 117–123.

[45] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.

[46] H. Zhang, "Discrete combinatorial Laplacian operators for digital geometry processing," in *Proc. SIAM Conf. Geometric Design Comput.*, 2003, pp. 575–592.

[47] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3-D medical image segmentation: A review," *Med. Img. Anal.*, vol. 13, no. 4, pp. 543–563, Aug. 2009.

[48] S. Sommer, M. Nielsen, F. Lauze, and X. Pennec, "A multi-scale kernel bundle for LDDMM: Towards sparse deformation description across space and scales," *Inf. Process. Med. Imag.*, vol. 22, pp. 624–635, 2011.

[49] R. Kushunapally, A. Razdan, and N. Bridges, "Roughness as a shape measure," *Computer-Aided Design Appl.*, vol. 4, no. 1–4, pp. 295–310, 2007.

[50] D. W. Shattuck, G. Prasad, M. Mirza, K. L. Narr, and A. W. Toga, "Online resource for validation of brain segmentation methods," *NeuroImage*, vol. 45, no. 2, pp. 431–439, Apr. 2009.

[51] Y. Shi, R. Lai, J. H. Morra, I. Dinov, P. M. Thompson, and A. W. Toga, "Robust surface reconstruction via Laplace-Beltrami Eigen-projection and boundary deformation," *IEEE Trans. Med. Imag.*, vol. 29, no. 12, pp. 2009–2022, Dec. 2010.

[52] Y. Gao, T. Riklin-Raviv, and S. Bouix, "Shape analysis, A field in need of careful validation," *Hum. Brain Mapp.*, vol. 35, no. 10, pp. 4965–4978, Oct. 2014.

[53] C. R. Jack *et al.*, "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *J. Magn. Reson. Imag.*, vol. 27, no. 4, pp. 685–691, Apr. 2008.

[54] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Phys. Med. Biol.*, vol. 46, no. 3, pp. R1–R45, Mar. 2001.

[55] J. Jovicich *et al.*, "Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data," *NeuroImage*, vol. 30, no. 2, pp. 436–443, Apr. 2006.

[56] P. A. Narayana, W. W. Brey, M. V. Kulkarni, and C. L. Sievenpiper, "Compensation for surface coil sensitivity variation in magnetic resonance imaging," *Magn. Reson. Imag.*, vol. 6, no. 3, pp. 271–274, May 1988.

[57] A. Guimond, J. Meunier, and J.-P. Thirion, "Average brain models: A convergence study," *Comput. Vis. Image Understand.*, vol. 77, no. 2, pp. 192–210, Feb. 2000.

[58] G. Heitz, T. Rohlfing, and C. R. Maurer, Jr., "Statistical shape model generation using nonrigid deformation of a template mesh," in *Proc. SPIE Med. Imag.*, 2005, pp. 1411–1421.

[59] J. C. Gower, "Generalized Procrustes analysis," *Psychometrika*, vol. 40, no. 1, pp. 33–51, 1975.

[60] I. J. Deary *et al.*, "The Lothian Birth Cohort 1936: A study to examine influences on cognitive ageing from age 11 to age 70 and beyond," *BMC Geriatr.*, vol. 7, p. 28, 2007.

[61] J. M. Wardlaw *et al.*, "Brain aging, cognition in youth and old age and vascular disease in the Lothian Birth Cohort 1936: Rationale, design and methodology of the imaging protocol*," *Int. J. Stroke*, vol. 6, no. 6, pp. 547–559, 2011.

[62] M. C. Valdés Hernández *et al.*, "Color fusion of magnetic resonance images improves intracranial volume measurement in studies of aging," *Open J. Radiol.*, vol. 2, pp. 1–9, 2012.

[63] L. Penke *et al.*, "Brain iron deposits are associated with general cognitive ability and cognitive aging," *Neurobiol. Aging*, vol. 33, no. 3, pp. 510–517.e2, Mar. 2012.

[64] D. Wechsler, WAIS-III$^{\text{UK}}$ *Administration and Scoring Manual*. London, U.K.: Psychol. Corp., 1998.

[65] D. Wechsler, WMS-III$^{\text{UK}}$ *Administration and Scoring Manual*. London, U.K.: Psychol. Corp., 1998.

[66] J. M. Fitzpatrick *et al.*, "Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain," *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 571–585, Aug. 1998.

[67] L. G. Apostolova *et al.*, "Hippocampal atrophy and ventricular enlargement in normal aging, mild cognitive impairment (MCI), and Alzheimer disease," *Alzheimer Dis. Assoc. Disord.*, vol. 26, no. 1, pp. 17–27, Jan. 2012.

[68] B. S. Aribisala, J. He, and A. M. Blamire, "Comparative study of standard space and real space analysis of quantitative MR brain data," *J. Magn. Reson. Imag.*, vol. 33, no. 6, pp. 1503–1509, Jun. 2011.