Shape Analysis, A Field in Need of Careful Validation

Yi Gao,[1,2]* Tammy Riklin-Raviv,[3] and Sylvain Bouix[4]

[1]Department of Electrical and Computer Engineering, The University of Alabama at Birmingham, Birmingham, Alabama
[2]Neuro-Oncology Program, Comprehensive Cancer Center, The University of Alabama at Birmingham, Birmingham, Alabama
[3]Electrical and Computer Engineering Department, Ben-Gurion University, Beer-Sheva, Israel
[4]Department of Psychiatry, Psychiatry Neuroimaging Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts

◆——————————————————◆

Abstract: In the last two decades, the statistical analysis of shape has become an actively studied field and finds applications in a wide range of areas. In addition to algorithmic development, many researchers have distributed end-user orientated toolboxes, which further enable the utilization of the algorithms in an "off the shelf" fashion. However, there is little work on the evaluation and validation of these techniques, which poses a rather serious challenge when interpreting their results. To address this lack of validation, we design a validation framework and then use it to test some of the most widely used toolboxes. Our initial results show inconsistencies and disagreement among four different methods. We believe this type of analysis to be critical not only for the community of algorithm designers but also perhaps more importantly to researchers who use these tools without knowing the algorithm details and seek objective criteria for tool selection. *Hum Brain Mapp 00:000–000, 2014.* © 2014 Wiley Periodicals, Inc.

Key words: statistical shape analysis; algorithm evaluation; reproducibility

◆——————————————————◆

INTRODUCTION

Motivation and Background

Statistical shape analysis is a widely studied topic [Heimann and Meinzer, 2009], with several proposed algorithms (see Ashburner and Friston, 2000; Cates et al., 2008; Dryden and Mardia, 1998; Gerig et al., 2001; Hufnagel et al., 2007; Kim et al., 2012; Krim and Yezzi, 2007; Miller, 2004; Peter and Rangarajan, 2006; Rao and Suryawanshi, 1996, 1998; Small, 1996; Styner et al., 2003; and Younes, 2010, to name just a few) and with findings in several application areas such as biology [Ziezold, 1994], neuroscience [Qiu et al., 2009], agriculture [Costa et al., 2011], and paleontology [Shen et al., 2009].

Shape analysis has been of particular interest in neuroscience where brain morphometry has been hypothesized to be linked to various neurodevelopmental and/or neurodegenerative brain disorders [Chung et al., 2008; Joseph et al., 2014; Kim et al., 2012; Qiu et al., 2009; Yang et al., 2013; Younes et al., 2014]. To facilitate these analyses, many researchers have translated prototype algorithms into publicly available end-user oriented toolboxes, with the claim that they can accurately detect statistically significant local shape changes between population (e.g., schizophrenia vs. matched controls). As a result, users can easily download and install a shape analysis toolbox "off the

(a) SPHARM-PDM   (b) SPHARM-MAT   (c) ShapeWorks   (d) TBM
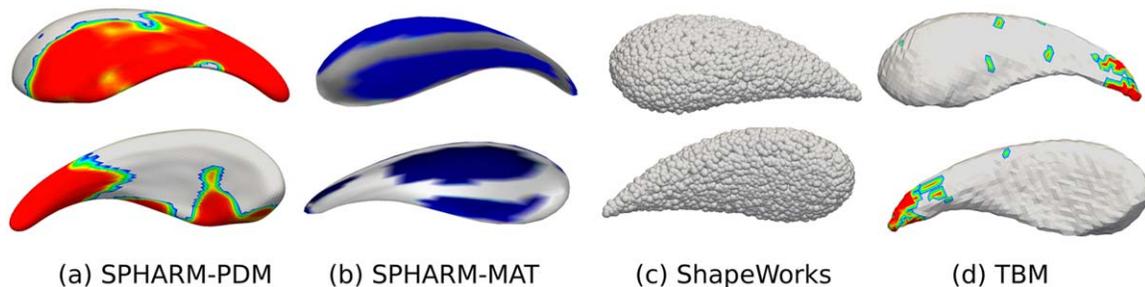
Figure 1.

Shape analysis results of the same sets of shapes using different algorithms. Color indicates the statistically significant different regions found by each method and gray means no difference. Note that SPHARM-MAT uses the jet colormap whereas all the others uses the reversed-jet colormap, also note that the Shape-Works renders the surface with a different (glyph) manner. With the same input shapes, inconsistency among the results demonstrate the necessity of a systematic evaluation for statistical shape analysis methods.

shelf" and apply it to their datasets. Unfortunately, very little evaluation has been done on these pipelines. In fact, when reviewing the literature, we did not find any reports on the evaluation and validation of shape analysis frameworks (although a few studies have evaluated subcomponents of the pipelines [Munsell et al., 2008; Selvarajah and Kodituwakku, 2010; Styner et al., 2007]). This lack of validation raises some serious concerns when using untested shape analysis tools in brain imaging studies.

The goal of our work is to create such an evaluation framework and present some preliminary results on commonly used shape analysis pipelines. Figure 1 shows the output of our analysis of four different toolboxes using two groups of synthesized shapes as input. One group was artificially altered with known local deformations. One can observe that the output varies significantly between methods, which support the necessity of a systematic evaluation of statistical shape analysis algorithms. To the best of our knowledge, this work is the first addressing this problem. We believe that this type of analysis is not only informative for algorithm developers to improve their techniques but also necessary knowledge to be provided to the end-user community.

A shape evaluation platform has to provide two major components: first, a large test dataset with known, measurable, and realistic deviations and second, a metric to evaluate the performance of the algorithm. We address the first problem using manifold learning techniques to generate arbitrarily many new shapes based on a training set of manually drawn neuroanatomical structures from brain MRI data of a normal, healthy population. We then alter some of the generated shapes with smooth and invertible deformations to one or more local regions, consistently located across the "synthetic abnormal" population. Then, the second problem is addressed by defining a metric measuring the matching between the constructed ground truth and the results returned by each algorithm.

A Brief Review of Shape Analysis Methods

Before we present our methodology in more detail, we provide here a brief overview of different classes of shape analysis methods. Note that we only tested a small subset of these methods, as very few are publicly available and reimplementing all these algorithms is beyond the scope of this work.

Surface parametrization based methods

This group of algorithms performs a parametrization of the surfaces, on a simple domain such as the sphere $\mathbb{S}^2$. By treating the surfaces as functions on the simple domain enable using various functional analysis techniques to decompose the surfaces into coefficients of certain basis. For example, in Gerig et al. [2001] and Styner et al. [2006], the spherical harmonics were used. The inputs of the algorithm are two groups of shapes represented by binary volumetric images. After a topology correction step to ensure the input shapes are of spherical topology, images are converted to meshes. Next, an area-preserving, distortion-minimizing spherical mapping is performed to provide a spherical parametrization of the surfaces, from which a spherical harmonics description of the shapes can be computed. The spherical harmonics are then sampled on the triangulated meshes via icosahedron subdivision of the spherical parametrization. Statistical tests are carried out and the P-values are corrected for multiple comparison.

More recently, authors in Chung et al. [2008] use a weighted-SPHARM scheme, where the deformable surface registration in Davatzikos and Bryan [1996] was adopted to map the surface to the sphere. Effectively, the weighted-SPHARM is equivalent to performing a smoothing convolution and it reduces the ringing effect of the SPHARM representation especially for the high frequency components, which correspond to the sharp feature of the

shape. Along similar lines, the spherical wavelets have been used in Nain et al. [2005].

SPHARM-MAT has a similar framework as the SPHARM-PDM but uses a different strategy for spherical parametrization, SPHARM coefficients computation, and alignment [Shen et al., 2009]. In fact, it was one of the design philosophies of this technique to have two similar tools under different platform (Matlab vs. C++) for: "(1) flexibility for users to choose the platform they favor and (2) opportunity for tool comparison and cross validation [Shen, 2010]." In addition, SPHARM-MAT has a slightly different spherical mapping and registration scheme. As a result, evaluating the differences between the two spherical harmonics-based methods would provide insight into how their different components affect the final outcome.

Both SPHARM-PDM and SPHARM-MAT are publicly available and we tested both tools.

Surface correspondence based methods

In this group of methods, researchers construct the correspondence between the surfaces of the shapes. In Durrleman et al. [2008], Joshi et al. [1997], Vaillant and Glaunes [2005], and Vaillant et al. [2007], the LDDMM scheme is used for matching surfaces in which the flow field is a geodesic in a Riemannian structure of diffeomorphism. After the correspondences have been established, the discrepancies in the corresponding points on the surfaces are captured using statistical analysis tools. This is applied to the study of the basal ganglia shapes in Qiu et al. [2009] and hippocampus [Yang et al., 2013]. In Pichon et al. [2006], a Laplace equation is solved between the shape surfaces for constructing the correspondence. In both methods, the vector flow field between the surfaces is smooth and smoothly invertible.

Under the discrete setting, shapes have been modeled as points (particles) sets where the particles are distributed on the surfaces of all the shapes by optimizing an entropy-based energy functional ([Cates et al., 2008], Scientific Computing and Imaging Institute [SCI]. ShapeWorks: An open-source tool for constructing compact statistical point-based models of ensembles of similar shapes that does not rely on specific surface parameterization). The resulting particles are corresponding across the population.

The ShapeWorks algorithm is also available publicly ([Cates et al., 2008], Scientific Computing and Imaging Institute [SCI]. ShapeWorks: An open-source tool for constructing compact statistical point-based models of ensembles of similar shapes that does not rely on specific surface parameterization) and was evaluated using our framework. Note, however, that ShapeWorks does not provide statistical testing, so we combined the point-wise correspondence computed by ShapeWorks with the nonparametric statistical analysis of SPHARM-PDM, to produce results, whose formatting is similar to the other methods tested.

Volumetric matching based methods

Building the correspondence lies in the center of the shape analysis tasks. Instead of finding the correspondences between surfaces, some approaches align volumetric representations of shapes [Ashburner and Friston, 2000; Ashburner et al., 1998; Good et al., 2001; Zarei et al., 2011]. The resulting deformation fields are then analyzed (a common measure is the determinant of the Jacobian of the field).

In this work, we propose and implement a variant of the tensor-based morphometry (TBM). TBM has been proposed originally to be applied on gray scale images. As a consequence, unlike surface-based methods, TBM can be applied without a preceding segmentation step. Conversely, nonlinear gray scale image registration being itself an active research topic may bring discrepancy into the analysis when the registration is not perfect. To address such an issue, Wang et al. [2011] proposed to use surface registration to compute the deformation tensor, which effectively incorporates the segmented object surfaces into the registration for higher specificity. Along a similar direction, in this work, we propose to perform the registration on the binary segmented images. This can be considered as applying the registration on the target's density-maps where the decision is binary. By doing this, we are able to specifically focus on the shape of the interest and ignore the influence from the surroundings, nearby or remote. Moreover, registering binary maps can be applied to situations with arbitrary shape topology.

Specifically, first, an unbiased mean shape μ is created by aligning all testing shapes to a mean shape estimate using similarity transforms. The mean shape is then re-estimated and the process is repeated until convergence. Then, each aligned shape is deformably registered to μ. Then, the determinant of the Jacobian of the obtained displacement field is computed. The resulting scalar valued volumes are then mapped on the surface of μ and compared using a two sample student's t-test to check for population differences. False discovery rate (FDR) correction is applied to the resulting P-values to adjust for multiple comparisons [Dudoit and van der Laan, 2008].

The choice of the deformable registration technique to create the deformation field is often viewed as a critical component to the quality of the output. For example, the LDDMM registration is used in Beg et al. [2005], Durrleman et al. [2008], and B-spline-based registration techniques can also be used [Ruckert et al., 1999; Ou et al., 2011]. In this work, we use the symmetric diffeomorphic Demons algorithm [Vercauteren et al., 2009].

Other analysis methods

One common aspect all the methods presented above is that they provide information about the location of shape changes, usually in the form of a P-value map displayed on a template shape. In contrast, other methods provide a single low dimensional feature vector, which describes each shape.

The output of a population study is usually a single number indicating the "distance" between shapes without localizing where the differences may be. For example, in Kim et al. [2012], Niethammer et al. [2007], and Reuter et al. [2009], the eigenvalues of the Laplace–Beltrami operator on the surfaces are used for measuring the distance between any two shapes. Similarly, the LDDMM metric [Beg et al., 2005] and the earth-mover distance have also been used [Osada et al., 2001, 2002]. Another interesting group of methods describe shapes using medial representations [Bouix et al., 2005; Gorczowski et al., 2007; Siddiqi and Pizer, 2008] and provide local statistics on a medial model. We decided not to include these last two classes of methods in the current evaluation as their inherent representation of shape was not directly comparable to the surface based techniques.

In summary, our overarching goal was to provide a framework that can be used to evaluate shape analysis methods with a surface-based representation. We then tested four readily available tools using this framework. We did not perform an exhaustive evaluation of all available tools, but our sampling was sufficient to discover large discrepancies in the output. We hope that such a pilot study along with the evaluation datasets will encourage the community to test newly developed algorithms on controlled datasets and help end-users to better evaluate the tool they select before applying it to their data.

The remainder of the article is organized as follows. In Materials and Methods section, we present the shape analysis evaluation framework, this includes how to generate an arbitrary number of anatomically realistic shapes with known and consistent deformations among them and quantitative evaluation of shape analysis algorithms. Next in Experiments, Validations, and Results section, we perform the evaluation of the four algorithms mentioned above and report the results. Finally, future and ongoing work is discussed in Discussion, Conclusions, and Future Work section.

MATERIALS AND METHODS

In this section, we first present a method to generate the testing data necessary for evaluating shape analysis algorithms, then we present our evaluation criteria and use the full framework on some widely used algorithms/tools.

Shape Synthesis with Known Deformations

To generate many shapes with known deformations, we use a manifold learning strategy for "normal" shape synthesis. We then generate smooth and invertible deformation on a subset of the data to create an "abnormal" population of shapes. A diagram of the shape and deformation synthesis procedures is shown in Figure 2.

Shape sampling via manifold learning

Learning the underlying "manifold" of a set of shapes is an active topic of research and remains an open problem.

In this work, however, we use manifold learning not so much for discovering the underlying structure of the shape space, but rather for providing a large set of realistic test shapes based on a relatively small set of training shapes. We also focus on shapes with relatively simple geometrical complexity (caudate and basal ganglia), for which the use of a standard manifold learning technique (here Local Linear Embedding) is appropriate.

Assume we have a set of anatomical structures from a normal control population represented by binary images: $\tilde{S}_i : \Omega \subset \mathbb{R}^3 \to \{0,1\}$; $i=1,2,\ldots,N$, where Ω is the common domain of the images. From these structures, the goal of this subsection is to generate arbitrarily many similar shapes to enrich the datasets.

Following Dryden and Mardia [1998] and Small [1996], we define "shape" as the geometric features of an object that are invariant to similarity transformations. Therefore, the first step of our pipeline is to align all training images using similarity transformations minimizing the mean-squared-error. We denote the registered training shapes as $S_j := \tilde{S}_j^{\circ} T_j : \Omega \to [0,1]$; $j=1,2,\ldots,N$. We note that we could perform an unbiased population based registration, but our experiments do not show this as a necessary step for the purpose of generating new shapes from a learned manifold.

In a high dimensional shape space, the N training shapes S_j's most likely reside on a manifold rather than in a linear space [Dambreville et al., 2008; Joshi et al., 1997]. To generate arbitrarily many shapes of the same category, one need to "sample" shapes by interpolating the training shapes on the manifold. However, due the high dimensional nature of the manifold, characterizing its topology for interpolation is difficult. To overcome this problem, we apply a local linear embedding method to map the high dimensional manifold to a lower dimensional space, and perform the interpolation therein [Roweis and Saul, 2000]. Let (n_x, n_y, n_z) be the size of the binary volume S_i and let $V_i \in \mathbb{R}^D$, where $D = n_x \cdot n_y \cdot n_z$, be a long vector representation of the binary image. First, we compute the weights $W = \{W_i^j\}$; $j = 1, \ldots, k$ so that each vector V_i can be represented by its k nearest neighbors with minimal error:

$$\Psi(W) := \sum_{i=1,2,\ldots,N} \left| V_i - \sum_{j \in K_i} W_i^j V_j \right|_D^2$$

where K_i is the index set of V_i's k nearest neighbors, and $\sum_j W_i^j = 1$, $\forall i$.

We then map V_i to $v_i \in \mathbb{R}^d$ using the singular value decomposition, where $d \ll D$ such that the following error is minimized:

$$\Phi(v_i) := \sum_{i=1,2,\ldots,N} \left| v_i - \sum_{j \in K_i} W_i^j v_j \right|_d^2$$

Mapping the manifold to a locally linear lower dimensional space enables us to locally approximate the topology of the manifold with the Delaunay triangulation [Cormen et al., 2001]. The local structure of the manifold

(in low dimension) is thus characterized by the d-simplex (e.g., triangle in 2D or tetrahedron 3D), and we can now generate a sample on the manifold by interpolating it on the d-simplex and then map it back to the high dimensional space. Details on how a new random sample is estimated are as follows. First, a d-simplex, along with its associated shapes (V_i, $i = 1, \ldots, d$), is randomly selected from the manifold, and a d-dimensional random vector $r \in \mathbb{R}^d$ is generated such that each of r's components are uniformly distributed on [0,1]. r is then normalized so that $|r|_1 = 1$. After that, a new shape vector $V^* \in \mathbb{R}^D$ can then be generated as $V^* = \sum_{i=1}^d r_i V_i$ and reformatted to a volumetric image $U^* \in \mathbb{R}^3 \rightarrow [0,1]$. Similar to the k parameter above, larger d results in more shapes contributing in the generation of the new shape, which usually cause the new shape to be smoother. The above process can be repeated infinitely, and allows us to create arbitrarily large test datasets.

It is noted that learning the shape manifold and generating new shapes are topics having been studied by many researchers [Dambreville et al., 2008; Joshi et al., 1997; Kendall, 1984; Siddiqi and Pizer, 2008], and we are not claiming the proposed algorithm being superior to any of the existing ones. In fact, we just find that such generation method is sufficient for the central goal of shape analysis evaluation, with good balance in shape versatility and computation load.

Shape abnormality simulation

In the context of medical studies, most statistical shape analysis methods aim at examining whether there exist statistically significant shape differences between two populations of anatomical structures [Dryden and Mardia, 1998]. The hypothesis is that a disease acts on the morphometry of the anatomy in a consistent manner and can be detected, localized, and measured with the appropriate tools. The method described above is only able to generate large "new" datasets similar to the training dataset. We thus need to design an algorithm to introduce artificial, yet realistic shape deformations to the data to mimic the action of a disease on an anatomical structure. In many cases, the anatomical abnormalities are hypothesized to be (i) located around similar regions across the shape population, and (ii) observable locally as a protuberance/depression or more globally as thinning/thickening of certain regions.

In what follows, we present a scheme that is able to introduce controlled shape changes at similar regions across a population. This is achieved by applying a joint clustering algorithm to define a consistent parcellation across all shapes in the population. The shapes are then deformed using a diffeomorphic framework to simulate smooth and realistic shape abnormalities.

Joint clustering. To find a common location among the population to induce a deformation, a clustering of the points on the shape boundary is performed jointly over all shapes. As we will observe later, this not only addresses the common location requirement but also allows us to control the extent of a shape change over the object.

Let $U_i; i = 1, \ldots, M$ be the generated sample shape binary maps and $C^i = \{c_1^i, \ldots, c_{L_i}^i\} \subset \mathbb{R}^3; i = 1, \ldots, M$ be their corresponding triangulated surface meshes, with L_i the number of vertices in mesh C^i. We first compute the outward normal vector $\mathcal{N}_j^i : C^i \rightarrow \mathbb{R}^3$ for each vertex ($j = 1, \ldots, L_i$). We then represent each surface as a six-dimensional feature vector set $F^i = \{f_j^i; j = 1, \ldots, L_i\}$ where $f_j^i := (c_j^i; \lambda \mathcal{N}_j^i) \in \mathbb{R}^6$ is a feature vector, where λ weighs the contribution of the unit normal in the clustering. Finally, the union of all feature vectors $F := U_M^i F_i$ is clustered using the k-means algorithm [Duda et al., 2001] and the resulting label of each vertex is denoted as $y_j^i \in$. We incorporate the normal vector as a feature to ensure that different "sides" in an object are not clustered together, which could happen in very thin structures. The normal vector is a natural choice to separate two sides of a shape. It also allows to control the side on which an inflation or deflation may be introduced.

Using k-means also allows us to control the scale of the deformation, by simply changing the number of output clusters. If the purpose is to evaluate shape analysis algorithms' capability of discriminating small scale shape difference, the number of cluster parameter is set to be large, usually resulting in small regions on the surface. Conversely, a small cluster number will give large patches on the shape surface, which indicates global scale shape deformation.
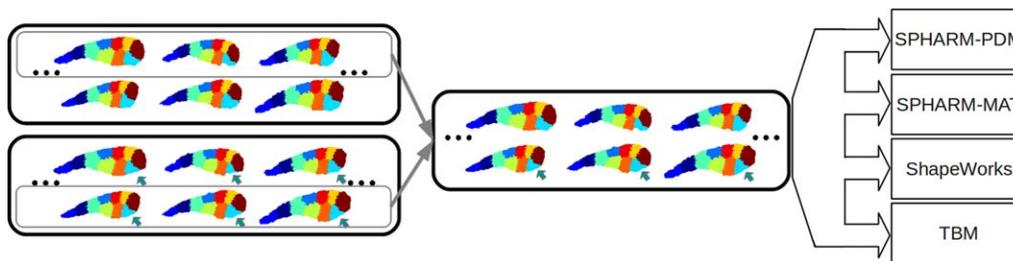
Figure 2.
Evaluate four shape analysis toolboxes using the generated shapes from Figure 2.

Diffeomorphic deformation. Once the clustering is obtained, any cluster (or combination of clusters) can be selected to be deformed. Both inward and outward deformations can be chosen as they can mimic true anatomical phenomena. For instance, to model a local bulge (on a small scale) or thickening (on a large scale), the deformation direction should be aligned with the outward normal. However, simply deforming the patch by pulling all cluster (patch) coordinates by an arbitrary amount in the normal direction has two significant problems. First, the deformation would not be smooth; and second, the deformed patch could collide or merge with other parts of the object, which would be anatomically unrealistic, as shown in Figure 3b. To solve these problems, we use the Log-Euclidean framework of Arsigny et al. [2009] to generate diffeomorphic deformations, which will be smooth and preserve the topology of the object.

Without loss of generality, let $G = \{c_j^i : y_j^i = l_0\} \subset \mathbb{R}^3$ be the patch to be deformed by an affine transformation denoted as a 4×4 homogeneous matrix. Directly applying the affine transformation to the points of G is problematic, because it will cause discontinuity. Instead, we smoothly extend the effect of the deformation to the entire image domain so as to make the deformation a diffeomorphism. This is done by first defining a weighting function: $w : \mathbb{R}^3 \to \mathbb{R}^+$ as

$$w(x) = \exp\left(-\frac{\min_{g \in G}|x-g|^2}{\sigma}\right)$$

The function w indicates how strong the deformation vector will affect a spatial point located at x and the σ parameter controls the smoothness of the deformation.

Although directly using w to modulate the deformation would result in a nondiffeomorphic transformation, one can use the Log-Euclidean approach described in Arsigny et al. [2009], to generate a diffeomorphic mapping. First, the matrix logarithm of H, $\mathscr{H}$, is computed [Golub and Van Loan, 1996] and the weighting function w applied to $\mathscr{H} : \mathscr{L}(x) = w(x) \mathscr{H}$. The diffeomorphic trans-

formation is obtained by applying the matrix exponential to $\mathscr{L} : Z(x) = \exp m(\mathscr{L}(x))$, and all surface points can be diffeomorphically transformed as

$$\tilde{c}_j^i = c_j^i + Z(c_j^i); j = 1, \ldots, L_i; i = 1, \ldots, M$$

for the vertex on the deformed mesh with its corresponding deformed volume

$$\tilde{U}_i = U_i(x - Z(x)); i = 1, \ldots, M$$

Now we have all the elements required to generate synthetic large test datasets for the purpose of evaluating shape analysis tools.

Tests on Statistical Shape Analysis Methods

Having set up the testing shapes, the generated shapes are given as input to several shape analysis algorithms to evaluate their abilities in capturing the shape differences. Schematically, after the shape generation in Figure 4, the nondeformed and deformed shapes are input to four analysis toolboxes, as shown in Figure 4.

Mathematically, first, all the M newly generated shapes, $U_i : \mathbb{R}^3 \to [0, 1]; i = 1, \ldots, M$, are deformed in a consistent way and result in deformed shapes $\bar{U}_i; i = 1, \ldots, M$. Then, without loss of generality, half of the nondeformed shapes $U_i : \mathbb{R}^3 \to [0, 1]; i = 1, \ldots, M/2$ are considered as the "normal group," whereas the other half from the deformed group, $\bar{U}_i; i = M/2, \ldots, M$, are regarded as the "abnormal group."

The two groups of shapes are input to four well-known and publicly available statistical shape analysis algorithms reviewed in A Brief Review of Shape Analysis Methods section: SPHARM-PDM, SPHARM-MAT, ShapeWorks, and TBM.

Qualitative and Quantitative Analysis

To visualize the results, the *P*-value maps computed by each of the four shape analysis methods are plotted on
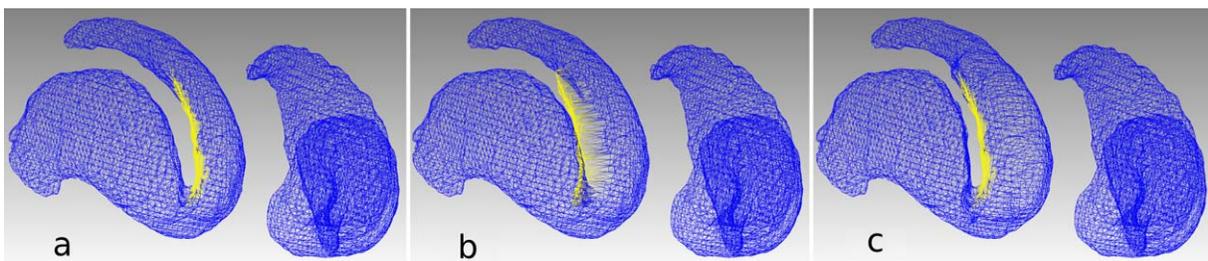
Figure 3.
In order to model a lateral swelling of the right caudate head, the yellow region in the original surface, shown in (**a**), is to be deformed. If an expansion vector field is directly applied, the surface protrudes into the putamen (**b**). Using the proposed method, the expansion becomes diffeomorphic (**c**) and the putamen retreats smoothly to avoid the collision.
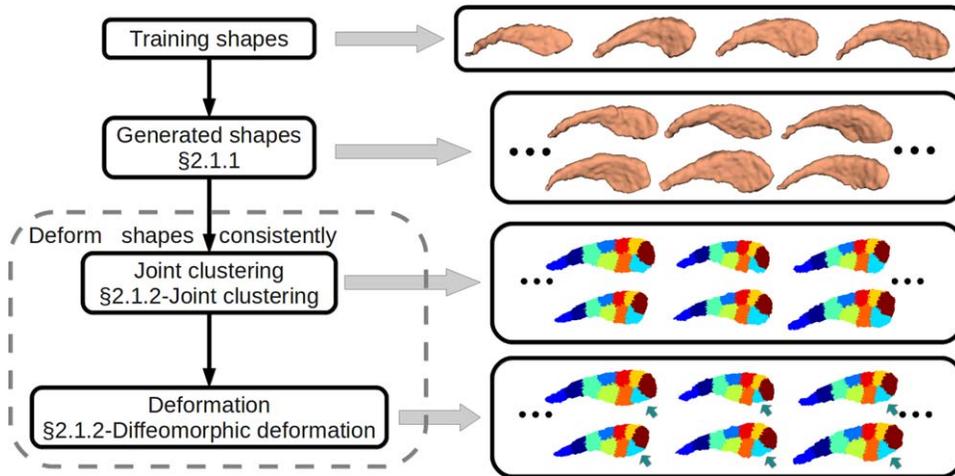
Figure 4.
Diagram of shape and deformation synthesis. Starting from a set of training shapes, the generate shape step uses the manifold learning to generate an arbitrary number of similar shapes. A joint clustering parcellates shapes into regions consistent across the population. Finally, the deformation step deforms the shapes using diffeomorphic transformations.

their respective mean shape. These colored mean shapes allow us to locate areas of differences between population as detected by the methods tested.

In addition to the visual illustration of the location and the scale of the detected shape difference region, we further quantitatively evaluate the quality of each tool by computing the ratio between the area of deformation and the area of the entire shape. This ratio is computed for each analysis methods and compared to a gold standard ratio as described below.

First, one has to make a clear definition for the "region of deformation" (ROD). For the "ground truth," due to the fact that we have the nondeformed control for each of the deformed shapes, we define the ROD as the region on the nondeformed control shape where the distance to the deformed counterpart exceeds a given threshold. Formally, given a shape pair U_i and its deformed version $\tilde{U}_i$, a distance map is defined on the 0.5-isosurface of U_i, ∂U_i, as the shortest distance to the 0.5-isosurface of $\tilde{U}_i$. Then, the distance is sorted and the ROD is defined as the region whose distance value exceeds 5% of all the distances. The area of ROD is divided by the total surface area of ∂U_i, resulting for each U_i an area ratio value a_i. The mean of a_i's, $i = 1, \ldots, M$, now serves as the gold standard ratio.

The ROD on the mean shapes resulting from the shape analysis methods is defined as the region whose P-value (after FDR correction) is less or equal than 0.05. The area of this region, divided by the total surface area of the mean shape, is also recorded and denoted as a'. One special case is the results of the ShapeWorks where no surface is defined. To resolve that problem, the ratio is taken between the number of points whose P-value is less or equal than 0.05 with the total number of points.

While we realize that this ratio measure has clear limitations, as the same ratio could stem from many different spatial configurations, we believe that this number in combination with the P-value plots is an informative way to evaluate shape analysis output. Ideally, one would want to be able to measure the overlap between the ground truth deformation and the detected deformation, as is often done with the Dice coefficient for evaluating image segmentation [Dice, 1945]. However, the P-value maps are defined on different template/mean shapes for different methods and would require building a "between-mean-shape" correspondence, which is in fact implicitly another shape analysis algorithm, and will almost inevitably introduce bias in the results.

The entire shape evaluation framework is given in Algorithm 1.

Algorithm 1. Shape analysis evaluation

1. Generate M shapes U_i $i = 1, \ldots, M$ from N training shapes
2. Deform shapes U_i to $\tilde{U}_i$ to represent the abnormal shapes
3. Input shapes U_i; $i = 1, \ldots, M/2$ and $\tilde{U}_i$; $i = M/2, \ldots, M$ to a shape analysis method
4. Display the mean shape and the P-value map on the mean shape surface
5. Extract ROD from the U_i and $\tilde{U}_i$ pairs (ground truth)
6. Extract ROD from P-value map
7. Compute ratio of areas
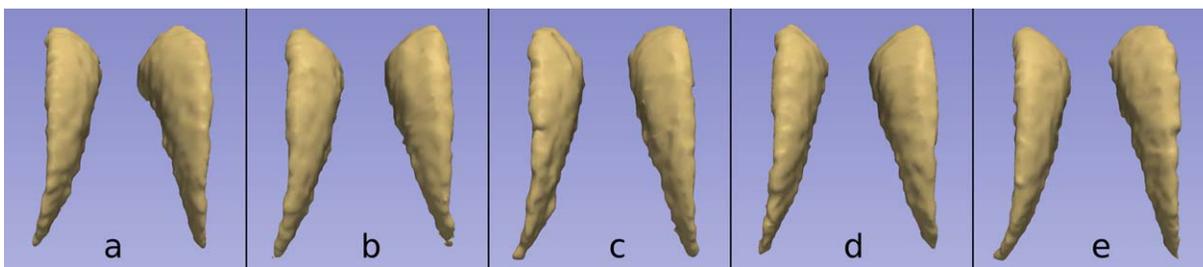
Figure 5.

An example set of left and right caudates (**a**) manually traced and used for training, and four synthetic shapes (**b–e**) generated through manifold learning.

EXPERIMENTS, VALIDATIONS, AND RESULTS

In this section, we first demonstrate the output of the shape generation algorithm based on two neuroanatomical structures. One is the caudate nucleus, and the other is the striatum (which includes the caudate). The newly generated shapes used to evaluate the four shape analysis algorithms are presented in Evaluation of Statistical Shape Analysis Methods section.

Shape Generation

We demonstrate the shape generation algorithm applied to binary maps representation of the caudate nucleus and the full striatum. All binary maps were selected from manual tracings of neuroanatomical structures in our MRI database of healthy controls. The training sample sizes were 30 for the caudate and 27 for the striatum. Based on Gao and Bouix [2012], the k, d, and σ parameters, there are fixed to 5, 3, and 2 for all the tests. Figure 5 shows four generated caudate shapes and Figure 6 shows four generated striatum shapes.

Figure 7 shows examples of deformed striatum shapes mimicking a thickening or thinning of the head of the caudate. Such shapes can then be used for testing and evaluating statistical shape analysis algorithms.

Evaluation of Statistical Shape Analysis Methods

In this section, we use the synthetically generated caudate nuclei and the striatums to test the ability of four statistical shape analysis methods to identify the location and magnitude of the shape differences.

The caudate shape is used due to its relative simplicity, spherical topology, and its shape being related to many brain disorders (e.g., Levitt et al., 2009). $M = 80$ shapes are generated and 40 are randomly selected to be deformed. In Figure 8, three examples of "before and after" shape pairs for to different deformations are shown. Using these pairs, we can compute the ground truth ROD and corresponding gold standard ratio, which in combination characterize the region, scale, and magnitude of the deformation. The color map on the shapes in the first row shows the point-wise distance to the corresponding shapes in the second row. Looking at Figure 8, we anticipate that detecting the synthesized atrophy at the caudate head may be a relative easy task while the detection of the small bend at the tail on the right caudates will be more challenging.

For the second set of experiments, we use the shape of the striatum because it is a more complex structure and a more challenging task for the shape analysis tools. Similarly to the caudate, we generate $M = 80$ new shapes and deform half of them. Different deformations are applied independently to the right and left striatum. Examples of deformed striatums are shown in Figure 9. On the left striatum, the medial side of the putamen was "pushed in." The deformation was significant enough that the lateral side of the putamen was slightly inflated due to the diffeomorphic nature of the deformation. Effectively, this deformation can be interpreted as atrophy on the medial side, a slight swelling on the lateral side, and a slight bending of
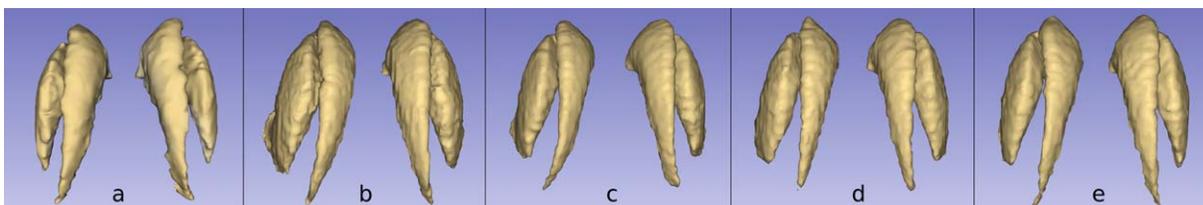
Figure 6.

A sample striatum shape used for training (**a**) and four striatum shapes (**b–e**) generated through manifold learning.
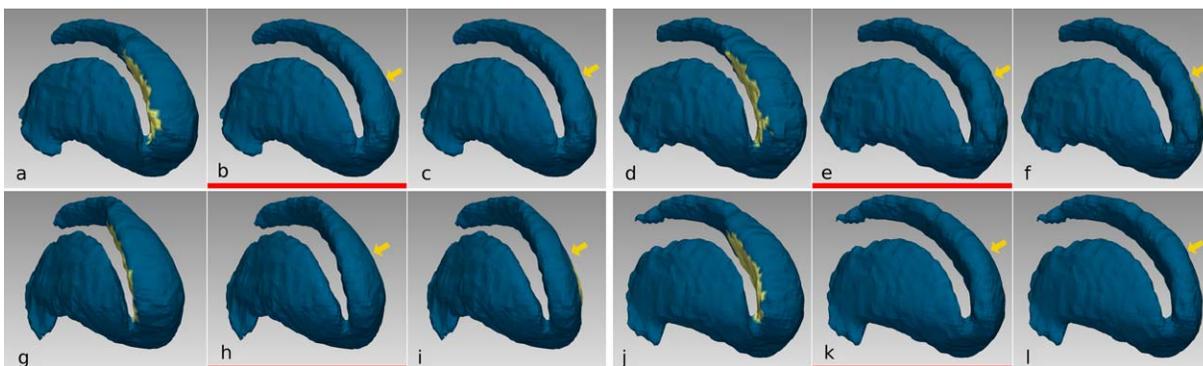
Figure 7.
Four sets of deformed shapes. The four subfigures underlined with a red bar (**b**, **e**, **h**, **k**) are four synthetic "normal" samples. Those to their left, are significantly thickened on the lateral side of the right caudate in the area highlighted in yellow. Those to their right, are more moderately thinned on the medial side of the right caudate, indicated by the yellow arrow.

the putamen toward the lateral direction. For the right striatum, we simulated a smaller atrophy on the putamen near the connecting region with the caudate.

Furthermore, note that the left and right sides of both the caudate and the striatum were analyzed independently, as many of the methods tested can only handle objects of spherical topology. In what follows, the findings of the four statistical shape analysis methods are reported and their performances are evaluated.

Analysis on generated non-deformed shapes

First, we checked that no statistically significant differences were found between groups of synthetically generated shapes regions which did not undergo any artificial deformation. To test this, we input U_i; $i=1,\ldots,M/2$ and U_i; $i=M/2,\ldots,M$ as two different groups to the four algorithms. All shape analysis algorithms tested found no statistically significant differences between these two "normal" groups.

This result suggests that our manifold learning procedure indeed produces shapes of similar morphometry as all shape analysis methods produce the same negative finding. Furthermore, it also increases our confidence that

when we do introduce a deformation, the shape analysis method tested, if accurate, should only detect that deformation and nothing else.

Evaluation results

As outlined in the Materials and Methods section, we tested four different shape analysis pipelines. Details about their respective settings are given below

- SPHARM-PDM: As suggested by the software manual, for the caudate, we set the icosahedral subdivision level to 21 so that the resulting surface has 4,412 vertices. Moreover, the maximum spherical harmonics degree is set to 15. For the striatum, the level of icosahedral subdivision is set to 23 and the maximal degree of spherical harmonics is set to 15. After the spherical mapping and spherical harmonics decomposition, an optional step consists of performing alignment of the shapes. Due to the fact that our input shapes have been registered before they are generated, this step is not performed. Of note, the spherical mapping did not successfully complete for three
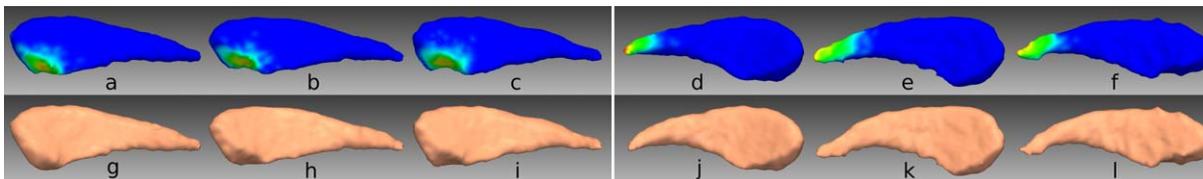
Figure 8.
Example of benchmark caudate shapes with know deformation. Top row: examples of the generated left caudates (**a**, **b**, **c**) and right caudates (**d**, **e**, **f**). Colormap indicates the regions of deformations with respect to the bottom row, which shows the respectively deformed shapes.
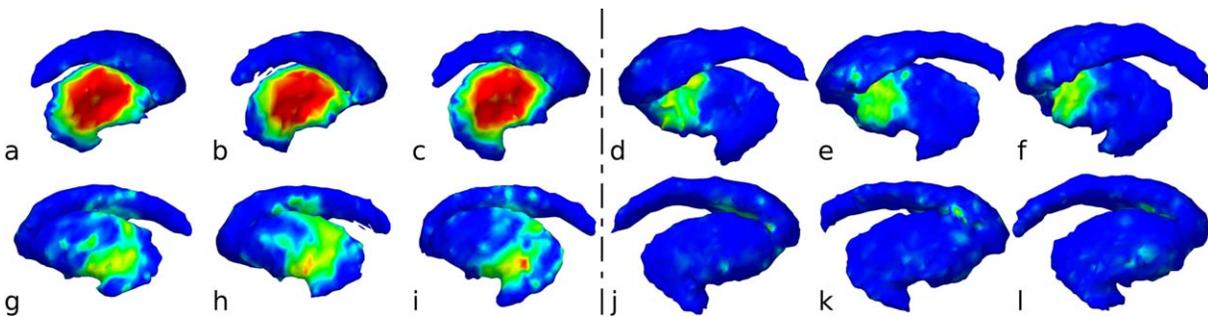
Figure 9.

Ground truth of three generated left striatum shapes (three left columns) and three right striatum shapes (three right columns) with know deformation region. The top row: the views from the medial side and the bottom row shows the lateral views. Colormap on the shapes indicates the location and magnitude of the artificial deformations with respect to their corresponding shapes. Blue means no deformation. By design, the deformation on the left striatum is larger and stronger than that on the right side.

caudate cases and they were excluded from further processing.

- SPHARM-MAT: The settings are kept consistent with SPHARM-PDM. Specifically, we set MeshGridSize to 50, MaxSPHARMDegree to 15, Tolerance to 2, Smoothing to 2, Iteration to 100, and LocalIteration to 10. Two caudate shapes and one striatum shape whose spherical mappings are not correct are removed from further processing. Moreover, the FDR correction is added to the end of the processing to correct the output P-values (The original software does not perform FDR correction). In the output display, SPHARM-MAT uses a color map where blue indicates close-to-zero P-values, which is opposite to the other methods.
- ShapeWorks: The binary images are input to the algorithm. The parameters antialias_iterations and blur_sigma are set to 20 and 2.0, respectively. For each shape, 4,096 particles are used. We tested the algorithm using more particles but no visible improvement was observed. The output of the ShapeWorks algorithm is a set of independent particles with no information about connectivity or topology. Surface reconstruction is available in the software, but our experience is that it can be unstable and may add an unnecessary variable to our evaluation. We thus decided to only use the particles and display them as small spherical glyphs with the P-values on the glyphs. This results in "rugged" figures, although we believe they are more representative of the method than surfaces.
- TBM: The binary images are input to the algorithm. The ITK software package is used to perform the volumetric registration [Ibanez et al., 2003]. Then, the mean shape surface is extracted using the marching cube algorithm [Lorensen and Cline, 1987] on which

the tensor fields are extracted from the registration deformation field and the statistical analysis is performed.

The results of the evaluations for the caudate shape are summarized in Figure 10. SPHARM-PDM presents many false-positive regions, even after the P-value maps have been FDR corrected. ShapeWorks, after FDR correction, correctly captures the ROD on the left caudate. However, FDR corrected P-values maps do not show any significant regions for the right caudate. Before FDR correction, although some scattered locations are detected, they are not consistent with the ground truth. For SPHARM-MAT, one can observe that the region of shape change on the left caudate, though smaller than anticipated, is captured. This improved accuracy over SPHARM-PDM may be explained by the different spherical mapping scheme used to minimize both angle and area distortion [Shen and Makedon, 2006]. However, on the right side the method underperforms, as too many regions are considered to be statistically different. Finally, TBM does not correctly identify the deformation for the left caudate. Instead, the result indicates that there are deformations on the medial and lateral sides near the caudate head. However, regions identified as different by TBM coincide with the ground truth very well on the right caudate.

Next, testing was performed using the striatum and the results of the evaluations are summarized in Figure 11.

One can observe that the atrophy in the right striatum is nicely captured by SPHARM-PDM. With FDR correction, ShapeWorks does not give any significant region. Without FDR correction (displayed in Fig. 11), the regions with significant differences on the left striatum reside in the putamen region; and the regions reaching significance differences on the right striatum are mostly around the region connecting the putamen and the caudate. Such
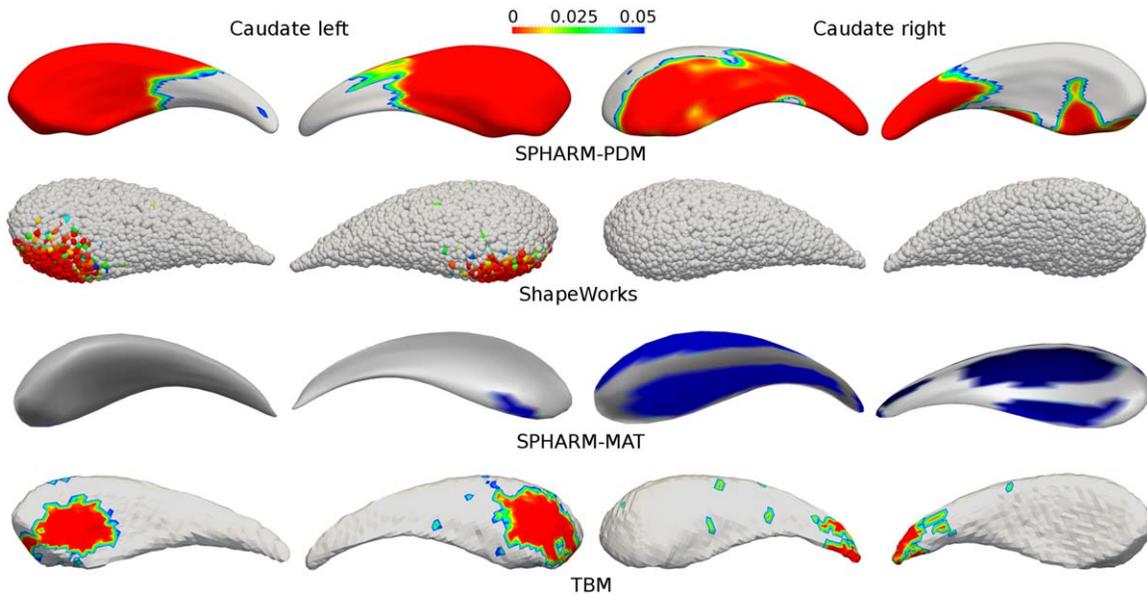
Figure 10.

Results of the caudate. Shown here are the FDR corrected *P*-value map plotted on the mean shapes computed by different algorithms. *P*-values larger than 0.05 are shown in gray. Please refer to text for algorithm specific settings.
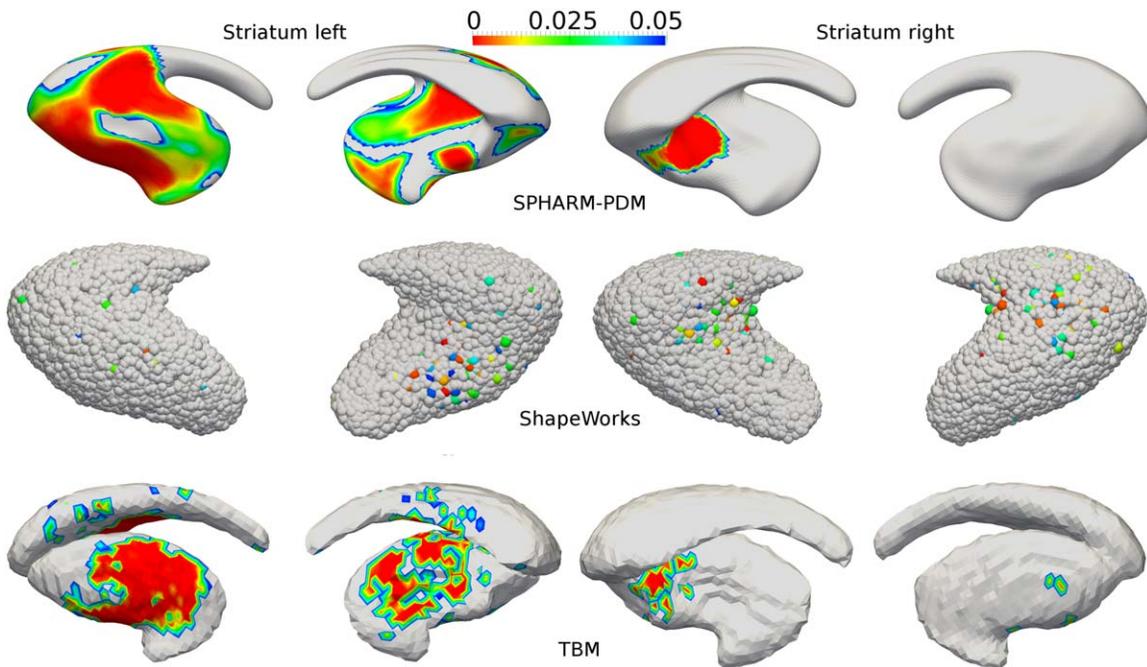
Figure 11.

Results of the striatum. Shown here are *P*-value map plotted on the mean shapes computed by the different algorithms. *P*-values larger than 0.05 are shown in gray. All *P*-value maps presented underwent FDR correction, except for ShapeWorks, as no regions survived the FDR threshold. The maps for this method are of raw *P*-values.

	Ground truth	SPHARM-PDM	ShapeWorks	SPHARM-MAT	TBM
Left caudate	0.17	0.81	0.18	0.09	0.17
Right caudate	0.13	0.65	0	0.74	0.03
Left striatum	0.33	0.61	0	N/A	0.17
Right striatum	0.15	0.10	0	N/A	0.01

All *P* values have been FDR corrected.

results are consistent with the ground truth presented in Figure 9, although spatially sparse and with weak *P*-values. SPHARM-MAT was not able to converge and no results could be displayed. TBM captured differences consistent with the ground truth for both location and size.

Quantitative Evaluation of the Results

While in the above section, we provide a visual illustration of the location and the scale of the detected shape difference region, in this section, we quantitatively evaluate the ROD and the results are given in Table I.

One should note that the numbers in the table should be analyzed in conjunction with the *P*-value maps. Indeed, even though the area ratios of significance can be identical, the locations may differ. Overall, TBM has significance ratios closer to the ground truth, although it underestimates the areas of change. SPHARM-PDM, conversely, overestimates the amount of true deformation. Other kinds of quantitative measures need to be investigated and we discuss this in Discussion, Conclusions, and Future Work section.

DISCUSSION, CONCLUSIONS, AND FUTURE WORK

In this work, we first present a framework for the generation of synthetic datasets for the purpose of evaluating shape analysis algorithms. A manifold learning technique allows us to generate arbitrarily many shapes, and the combination of a joint clustering and diffeomorphic trans-formation generation can simulate realistic, robust, and controlled deformations. Then, we perform an evaluation on well known statistical shape analysis algorithms using synthetically generated shapes.

For future work, we will investigate the influences of two factors in the shape analysis. First, it can be observed that registration plays an important role in the shape analysis framework. Due to the complexity of the shape that may be encountered in various applications, state-of-the-art registration techniques should be used accordingly. Moreover, based on the shape representation (volumetric/surface/points) adopted by different methods, the corresponding registration techniques should be selected. Indeed, different shape analysis algorithms have different ways to represent the shape, and therefore different subsequent processes. In fact, a precise quantitative analysis of shape study result is still an area of active research for us. However, we do think that the approximate location and magnitude of the known shape change should be consistent across methods, which can be qualitatively evaluated by the human eye with the figures we present in the article. This is particularly true when the group-wise deformations are relatively small. When the deformation is large, however, there may be ambiguity in how to represent and interpret the final comparison results.

For example, the two shapes in Figure 12a. Depending on different shape representation and subsequently different registration scheme being used, the algorithm may report that the orange one has a bended tail Figure 12b; the orange one has a bended head Figure 12c; they have deformation all over the entire shape Figure 12d. In such large and wide-spread deformation cases, we think
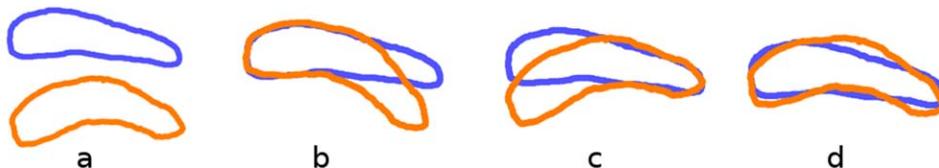
Figure 12.
Ambiguity in case the deformation is large/global. (**a**) Two different shapes with large difference. (**b**) If the heads are aligned, the algorithm will report the orange one has a bended tail. (**c**) If the tails are aligned, the algorithm will report the orange one has a bended head. (**d**) If they are aligned by minimizing L2 distance, the algorithm will report the orange one has deformation everywhere.

algorithms giving results as in Figure 12b–d are all "correct" by emphasizing certain aspects of the shapes. Consequently, we may not expect these to be comparable to the same standard.

The second important factor is the construction of the correspondence. Most shape analysis methods, explicitly or tacitly, build the correspondence among the shapes for comparison. The influence of this step has been studied in Munsell et al. [2008] and Styner et al. [2007]. With the proposed framework, we will further perform quantitative evaluation of different correspondence construction methods.

One important future direction of the proposed framework is to provide better quantitative measures for the evaluation. Our use of the area ratio together with the visual assessment of P-value maps give an indication of how various shape analysis methods perform. However, the area ratio alone is not sufficient to quantify the spatial accuracy of the methods. Ideally, one would want to be able to measure the overlap between the ground truth deformation and the detected deformation. This poses a serious challenge as the ground truth and algorithm output would have to be put in correspondence, which in itself would be another shape analysis algorithm. Another approach is to transfer scalar information about distortion on each shapes surface (see Fig. 8) through the analysis pipelines and evaluate the P-value map against this ground truth. This involves modifying the shape analysis techniques, and currently, we are actively working on finding better ways to quantitatively evaluate both location and extent accuracy.

ACKNOWLEDGEMENT

REFERENCES

Arsigny V, Commowick O, Ayache N, Pennec X (2009): A fast and log-Euclidean polyaffine framework for locally linear registration. J Math Imaging Vis 33:222.

Ashburner J, Friston K (2000): Voxel-based morphometry—The methods. Neuroimage 11:805–821.

Ashburner J, Hutton C, Frackowiak R, Johnsrude I, Price C, Friston K (1998): Identifying global anatomical differences: Deformation-based morphometry. Hum Brain Mapp 6:348–357.

Beg MF, Miller MI, Trouvé A, Younes L (2005): Computing large deformation metric mappings via geodesic flows of diffeomorphisms. Int J Comput Vis 61:139–157.

Bouix S, Siddiqi K, Tannenbaum A (2005): Flux driven automatic centerline extraction. Med Image Anal 9:209–221.

Cates J, Fletcher P, Styner M, Hazlett H, Whitaker R (2008): Particle-based shape analysis of multi-object complexes. Med Image Comput Comput Assist Interv 11:477–485.

Chung MK, Dalton KM, Davidson RJ (2008): Tensor-based cortical surface morphometry via weighted spherical harmonic representation. IEEE Trans Med Imaging 27:1143–1151.

Cormen T, Leiserson C, Rivest R (2001): Introduction to Algorithms. Cambridge, MA: MIT press.

Costa C, Antonucci F, Pallottino F, Aguzzi J, Sun D.-W, Menesatti P (2011): Shape analysis of agricultural products: A review of recent research advances and potential application to computer vision. Food Bioprocess Technol 4:673–692.

Dambreville S, Rathi Y, Tannenbaum A (2008): A framework for image segmentation using shape models and kernel space shape priors. IEEE Trans Pattern Anal Machine Intell 30:1385.

Davatzikos C, Bryan N (1996): Using a deformable surface model to obtain a shape representation of the cortex. IEEE Trans Med Imaging 15:785–795.

Dice L (1945): Measures of the amount of ecologic association between species. Ecology 26:297–302.

Dryden I, Mardia K (1998): Statistical Shape Analysis. New York: John Wiley & Sons.

Duda R, Hart P, Stork D (2001): Pattern Classification. John Wiley & Sons.

Dudoit S, van der Laan M (2008): Multiple Testing Procedures with Applications to Genomics. New York: Springer.

Durrleman S, Pennec X, Trouvé A, Thompson P, Ayache N (2008): Inferring brain variability from diffeomorphic deformations of currents: An integrative approach. Med Image Anal 12:626–637.

Gao Y, Bouix S (2012): Synthesis of realistic subcortical anatomy with known surface deformations. In: Levine JA, Paulsen RR, Zhang Y, editors. Mesh Processing in Medical Image Analysis 2012. Berlin Heidelberg: Springer, pp 80–88.

Gerig G, Styner M, Jones D, Weinberger D, Lieberman J (2001): Shape analysis of brain ventricles using SPHARM. In: IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, 2001. IEEE, pp 171–178. Meeting Date: 09 Dec 2001-10 Dec 2001 Print ISBN: 0-7695-1336-0 INSPEC Accession Number: 7190648.

Golub G, Van Loan C (1996): Matrix Computations. Baltimore, MD: Johns Hopkins University Press.

Good CD, Johnsrude IS, Ashburner J, Henson RNA, Friston KJ, Frackowiak RSJ (2001): A voxel-based morphometric study of ageing in 465 normal adult human brains. NeuroImage 14:21–36.

Gorczowski K, Styner M, Jeong J.-Y, Marron J, Piven J, Hazlett HC, Pizer SM, Gerig G (2007): Statistical shape analysis of multi-object complexes. In: IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR'07. Minneapolis, MN: IEEE, pp 1–8.

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

Hufnagel H, Pennec X, Ehrhardt J, Handels H, Ayache N (2007): Shape analysis using a point-based statistical shape model built on correspondence probabilities. Med Image Comput Comput Assist Interv, pp 959–967.

Ibanez L, Johnson HJ, McCormick M, Ibanez L (2003): The ITK software guide. Kitware.

Joseph J, Warton C, Jacobson SW, Jacobson JL, Molteno CD, Eicher A, Marais P, Phillips OR, Narr KL, Meintjes EM (2014): Three-dimensional surface deformation-based shape analysis of hippocampus and caudate nucleus in children with fetal alcohol spectrum disorders. Hum Brain Mapp 35:659–672.

Joshi SC, Miller MI, Grenander U (1997): On the geometry and shape of brain sub-manifolds. Int J Pattern Recognit Artifi Intell 11:1317–1343.

Kendall DG (1984): Shape manifolds, procrustean metrics, and complex projectivespaces. Bull Lond Math Soc 16:81–121.

Kim S.-G, Chung MK, Schaefer SM, van Reekum C, Davidson RJ (2012): Sparse shape representation using the laplace-beltrami eigenfunctions and its application to modeling subcortical structures. In: IEEE Workshop on IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA), 2012. Breckenridge, CO: IEEE, pp 25–32.

Krim H, Yezzi A Jr (2007): Statistics and analysis of shape. Control Cybern 36:303.

Levitt JJ, Styner M, Niethammer M, Bouix S, Koo MS, Voglmaier MM, Dickey CC, Niznikiewicz MA, Kikinis R, McCarley RW, Shenton ME (2009): Shape abnormalities of caudate nucleus in schizotypal personality disorder. Schizophr Res 110:127–139.

Lorensen WE, Cline HE (1987): Marching cubes: A high resolution 3d surface construction algorithm. In: ACM Siggraph Computer Graphics, Vol. 21. New York, NY, USA: ACM, pp 163–169.

Miller M (2004): Computational anatomy: Shape, growth, and atrophy comparison via diffeomorphisms. NeuroImage 23:S19–S33.

Munsell B, Dalal P, Wang S (2008): Evaluating shape correspondence for statistical shape analysis: A benchmark study. IEEE Trans Pattern Anal Machine Intell 30:2023–2039.

Nain D, Haker S, Bobick A, Tannenbaum AR (2005): Multiscale 3d shape analysis using spherical wavelets. In: Medical Image Computing and Computer-Assisted Intervention—MICCAI 2005. Berlin, Heidelberg: Springer, pp 459–467.

Niethammer M, Reuter M, Wolter F.-E, Bouix S, Peinecke N, Koo M.-S, Shenton ME (2007): Global medical shape analysis using the laplace-beltrami spectrum. In: Medical Image Computing and Computer-Assisted Intervention—MICCAI 2007. Berlin, Heidelberg: Springer, pp 850–857.

Osada R, Funkhouser T, Chazelle B, Dobkin D (2001): Matching 3d models with shape distributions. In: International Conference on Shape Modeling and Applications, SMI 2001. New York, NY, USA: ACM, IEEE, pp 154–166.

Osada R, Funkhouser T, Chazelle B, Dobkin D (2002): Shape distributions. ACM Trans Graph 21:807–832.

Ou Y, Sotiras A, Paragios N, Davatzikos C (2011): Dramms: Deformable registration via attribute matching and mutual-saliency weighting. Med Image Anal 15:622–639.

Peter A, Rangarajan A (2006): Shape analysis using the fisher-rao riemannian metric: Unifying shape representation and deformation. In: IEEE ISBI. IEEE, pp 1164–1167.

Pichon E, Nain D, Niethammer M (2006): A Laplace equation approach for shape comparison. In: Kevin R. Cleary; Robert L. Galloway, Jr. Proceedings of SPIE, International Society for Optics and Photonics, San Diego, CA, pp. 614119–614119.

Qiu A, Crocetti D, Adler M, Mahone EM, Denckla MB, Miller MI, Mostofsky SH (2009): Basal ganglia volume and shape in children with attention deficit hyperactivity disorder. Am J Psychiatry 166:74.

Rao CR, Suryawanshi S (1996): Statistical analysis of shape of objects based on landmark data. Proc Natl Acad Sci USA 93:12132–12136.

Rao CR, Suryawanshi S (1998): Statistical analysis of shape through triangulation of landmarks: A study of sexual dimorphism in hominids. Proc Natl Acad Sci USA 95:4121–4125.

Reuter M, Wolter F.-E, Shenton M, Niethammer M (2009): Laplace-beltrami eigenvalues and topological features of eigenfunctions for statistical shape analysis. Comput-Aided Des 41:739–755.

Roweis S, Saul L (2000): Nonlinear dimensionality reduction by locally linear embedding. Science 290:2323.

Ruckert D, Sonoda LI, Hayes C, Hill DL, Leach M, Hawkes DJ (1999): Nonrigid registration using free-form deformations: Application to breast MR images. IEEE Trans Med Imaging 18:712–721.

Selvarajah S, Kodituwakku S (2010): Performance evaluation of shape analysis techniques. ARPN J Syst Softw 1:12–18.

Shen L (2010): SPHARM-MAT v1.0.0 documentation. http://www.nitrc.org/projects/spharm-mat.

Shen L, Farid H, McPeek M (2009): Modeling three-dimensional morphological structures using spherical harmonics. Evolution 63:1003–1016.

Shen L, Makedon F (2006): Spherical mapping for processing of 3d closed surfaces. Image Vis Comput 24:743–761.

Siddiqi K, Pizer SM (2008): Medial Representations: Mathematics, Algorithms and Applications, Vol. 37. Springer, Science+Business Media, pp 127–153.

Small C (1996): The Statistical Theory of Shape. New York, USA: Springer.

Styner M, Gerig G, Lieberman J, Jones D, Weinberger D (2003): Statistical shape analysis of neuroanatomical structures based on medial models. Med Image Anal 7:207–220.

Styner M, Oguz I, Xu S, Brechbühler C, Pantazis D, Levitt J, Shenton M, Gerig G (2006): Framework for the statistical shape analysis of brain structures using SPHARM-PDM. Insight J 1071:242.

Styner M, Xu S, El-Sayed M, Gerig G (2007): Correspondence evaluation in local shape analysis and structural subdivision. In: IEEE ISBI. Arlington, VA: IEEE, pp 1192–1195.

Vaillant M, Glaunes J (2005): Surface matching via currents. In: Information Processing in Medical Imaging. Berlin, Heidelberg: Springer, pp 381–392.

Vaillant M, Qiu A, Glaun's J, Miller MI (2007): Diffeomorphic metricsurface mapping in subregion of the superior temporal gyrus. NeuroImage 34:1149–1159.

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

Wang Y, Panigrahy A, Shi J, Ceschin R, Nelson MD, Gutman B, Thompson PM, Lepore N. Surface multivariate tensor based morphometry on premature neonates: A pilot study. In: Proceedings of the MICCAI Workshop on Image Analysis of Human Brain Development (IAHBD 2011). 2011.

Yang X, Goh A, Chen S.-HA, Qiu A (2013): Evolution of hippocampal shapes across the human lifespan. Hum Brain Mapp 34:3075–3085.

Younes L (2010): Shapes and Diffeomorphisms, Vol. 171. Springer.

Younes L, Ratnanather JT, Brown T, Aylward E, Nopoulos P, Johnson H, Magnotta VA, Paulsen JS, Margolis RL, Albin RL, Miller MI (2014): Regionally selective atrophy of subcortical structures in prodromal HD as revealed by statistical shape analysis. Hum Brain Mapp 35:792–809.

Zarei M, Mataix-Cols D, Heyman I, Hough M, Doherty J, Burge L, Winmill L, Nijhawan S, Matthews PM, James A (2011): Changes in gray matter volume and white matter microstructure in adolescents with obsessive-compulsive disorder. Biol Psychiatry 70:1083–1090.

Ziezold H (1994): Mean figures and mean shapes applied to biological figure and shape distributions in the plane. Biometrical J 36:491–510.