Statistical Group Differences in Anatomical Shape Analysis using Hotelling T2 metric

Martin Styner[a,b], Ipek Oguz[a], Shun Xu[a], Dimitrios Pantazis[c], Guido Gerig[a,b]

[a]Department of Psychiatry, University of North Carolina, Chapel Hill NC, USA;
[b]Department of Computer Science, University of North Carolina, Chapel Hill NC,USA
[c]Signal & Image Processing Inst., University of Southern California, Los Angeles, CA

ABSTRACT

Shape analysis has become of increasing interest to the neuroimaging community due to its potential to precisely locate morphological changes between healthy and pathological structures. This manuscript presents a comprehensive set of tools for the computation of 3D structural statistical shape analysis. It has been applied in several studies on brain morphometry, but can potentially be employed in other 3D shape problems. Its main limitations is the necessity of spherical topology.

The input of the proposed shape analysis is a set of binary segmentation of a single brain structure, such as the hippocampus or caudate. These segmentations are converted into a corresponding spherical harmonic description (SPHARM), which is then sampled into a triangulated surfaces (SPHARM-PDM). After alignment, differences between groups of surfaces are computed using the Hotelling T^2 two sample metric. Statistical p-values, both raw and corrected for multiple comparisons, result in significance maps. Additional visualization of the group tests are provided via mean difference magnitude and vector maps, as well as maps of the group covariance information.

The correction for multiple comparisons is performed via two separate methods that each have a distinct view of the problem. The first one aims to control the family-wise error rate (FWER) or false-positives via the extrema histogram of non-parametric permutations. The second method controls the false discovery rate and results in a less conservative estimate of the false-negatives.

Prior versions of this shape analysis framework have been applied already to clinical studies on hippocampus and lateral ventricle shape in adult schizophrenics. The novelty of this submission is the use of the Hotelling T2 two-sample group difference metric for the computation of a template free statistical shape analysis. Template free group testing allowed this framework to become independent of any template choice, as well as it improved the sensitivity of our method considerably. In addition to our existing correction methodology for the multiple comparison problem using non-parametric permutation tests, we have extended the testing framework to include False Discovery Rate (FDR). FDR provides a significance correction with higher sensitivity while allowing a expected minimal amount of false-positives compared to our prior testing scheme.

1. INTRODUCTION

Quantitative morphologic assessment of individual brain structures is often based on volumetric measurements. Volume changes are intuitive features as they might explain atrophy or dilation due to illness. On the other hand, structural changes at specific locations are not sufficiently reflected in volume measurements. Shape analysis has thus become of increasing interest to the neuroimaging community due to its potential to precisely locate morphological changes between healthy and pathological structures.

One of the first and most influential research in shape analysis was presented by D'Arcy[1] in his groundbreaking book *On Growth and Form*. In more recent years, several researchers proposed shape analysis via deformable registration to a template.[2–5] Inter-subject comparisons are made by analyzing the individual deformable transformations. This analysis of the transformation fields has to cope with the high dimensionality of the transformation, the template selection problem and the sensitivity to the initial position. Nevertheless,

Email: martin_styner@ieee.org, WWW: www.ia.unc.edu

Medical Imaging 2007: Image Processing, edited by Josien P. W. Pluim, Joseph M. Reinhardt,
Proc. of SPIE Vol. 6512, 65123Z, (2007) · 1605-7422/07/$18 · doi: 10.1117/12.708467

several studies have shown stable shape analysis results. Bookstein[6] and Dryden et al[7] presented some of the first mathematical methods for 3D shape analysis based on sampled descriptions. The shape analysis of densely sampled 3D Point Distribution Models (PDM) and their deformations was first investigated by.[8] Inspired by their experiments, Styner et al[9] proposed shape analysis based on a parametric boundary description called SPHARM ([10]). The SPHARM shape analysis approach was extended[11] to use the implied sampled surface (SPHARM-PDM), a method also used by Li.[12] Pizer and later Styner[13,14] and Golland[15] proposed shape analysis on medial shape descriptions in 3D and 2D, respectively. They used a fixed topology sampled model with implicit correspondence that is fitted to the objects.

This manuscript presents a comprehensive methodology for the computation of 3D structural statistical shape analysis on the object boundary via SPHARM-PDM (see Figure 1). It has been applied in several studies on brain morphometry, but can potentially be employed in other 3D shape analysis problems. The main limitations of our shape analysis methodology is the necessity of spherical topology.

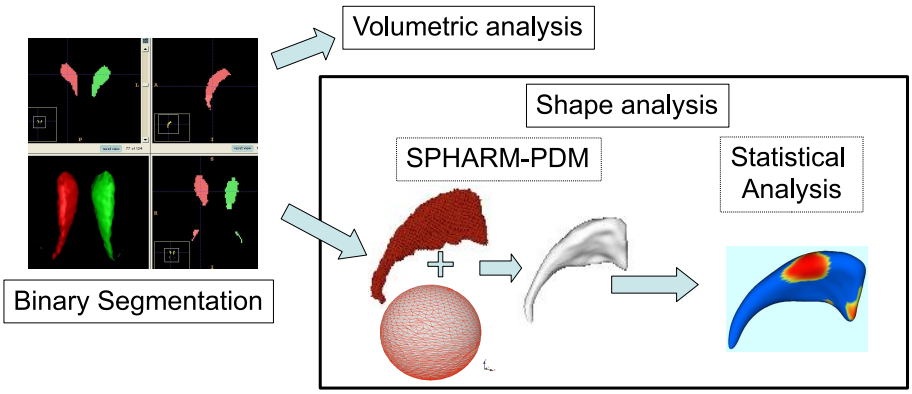

Figure 1. Schematic view of our SPHARM-PDM based shape analysis pipeline.

2. METHODS

This section describes the SPHARM-PDM shape description, the statistical analysis methodology, as well as issue of alignment and scaling normalization.

In summary, the input of the proposed shape analysis is a set of binary segmentation of a single brain structure, such as the hippocampus or caudate. These segmentations are first processed to fill any interior holes and a minimal smoothing operation. The processed binary segmentations are converted to surface meshes, and a spherical parametrization is computed for the surface meshes using a area-preserving, distortion minimizing spherical mapping. The SPHARM description is computed from the mesh and its spherical parametrization. Using the first order ellipsoid from the spherical harmonic coefficients, the spherical parametrizations are aligned to establish correspondence across all surfaces. The SPHARM description is then sampled into a triangulated surfaces (SPHARM-PDM) via icosahedron subdivision of the spherical parametrization. These SPHARM-PDM surfaces are all spatially aligned using rigid Procrustes alignment. Group differences between groups of surfaces are computed using the standard robust Hotelling T^2 two sample metric. Statistical p-values, both raw and corrected for multiple comparisons, result in significance maps. Additional visualization of the group tests are provided via mean difference magnitude and vector maps, as well as maps of the group covariance information.

The correction for multiple comparisons is performed via two separate methods that each have a distinct view of the problem. The first one aims to control the family-wise error rate (FWER) or false-positives via the extrema histogram of non-parametric permutations. The second method controls the false discovery rate and results in a less conservative estimate of the false-negatives.

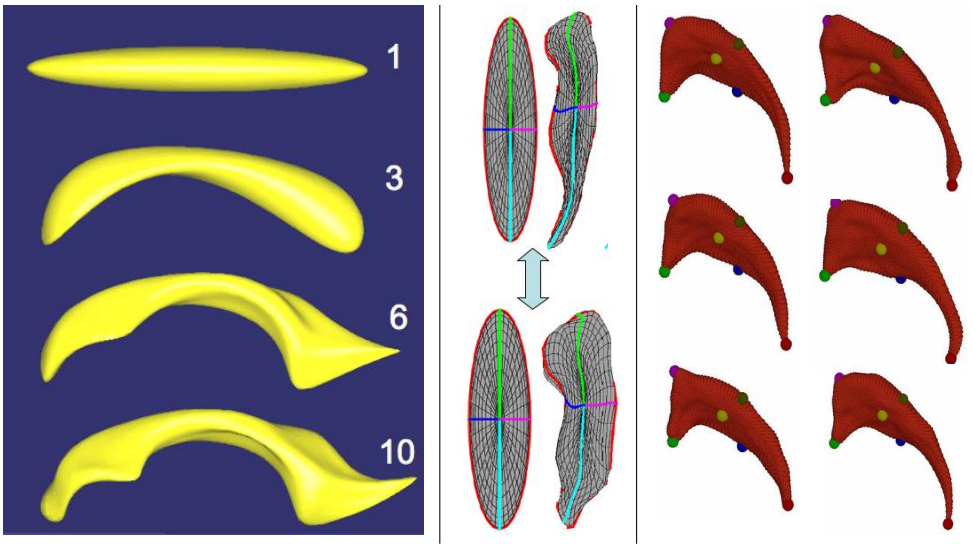

Figure 2. Left: The SPHARM shape description of a human lateral ventricle shown at 4 different degrees (1, 3, 6, 10 harmonics). Middle: Illustration of SPHARM correspondence via alignment of the spherical parametrization using the first order ellipsoid meridian and equator. Right: Set of 6 caudate structures with correspondence shown at selected locations via colored spheres.

2.1. SPHARM-PDM

In summary, the SPHARM description is a hierarchical, global, multi-scale boundary description that can only represent objects of spherical topology.[10] The spherical parameterization is computed via optimizing an equal area mapping of the 3D voxel mesh onto the sphere and minimizing angular distortions. The basis functions of the parameterized surface are spherical harmonics. Each individual SPHARM description is composed of a set of coefficients, weighting the basis functions.[16] demonstrated that SPHARM can be used to express shape deformations. Truncating the spherical harmonic series at different degrees results in object representations at different levels of detail. SPHARM is a smooth, accurate fine-scale shape representation, given a sufficiently high representation level. Based on a uniform icosahedron-subdivision of the spherical parameterization, we obtain a Point Distribution Model (PDM).

2.1.1. SPHARM Correspondence

Correspondence of SPHARM-PDM is determined by normalizing the alignment of the spherical parameterization to an object-specific frame. In the studies presented in this paper, the normalization is achieved by rotation of the parameterization, such that the spherical equator, 0° and 90° longitudes coincide with those of the first order ellipsoid(see Figure 2). After this parametrization normalization, corresponding surface points across different objects possess the same parameterization.

We have compared the SPHARM correspondence to other types of correspondences and it compared favourably to human expert established correspondence.[17]

2.2. Area-preserving Spherical Mapping

The appropriate parameterization of the points of a surface description is a key problem for correspondence finding, as well as for an efficient SPHARM representation. Every point i on the surface is to be assigned a parameter vector (θ_i, ϕ_i) that are located on the unit sphere. A homogeneous distribution of the parameter space is essential for an efficient decomposition of the surface into SPHARM coefficients. We give here a brief summary of the surface parameterization procedure proposed by Brechbühler.[10]

A bijective mapping of the surface to the unit sphere is created, i. e., every point on the surface is mapped to exactly one point on the sphere, and vice versa. The main idea of the procedure is to start with an initial

Recursive Level	Linear Level	Number of Surface Points
1	1	12
2	2	42
3	4	162
4	8	642
5	16	2562

Table 1. Comparison of subdivision levels for recursive and linear subdivision of the icoshedron.

parameterization. This initial parameterization is optimized so that every surface patch gets assigned an area in parameter space that is proportional to its area in object space. The initial mapping from surface to parameter space is constructed using discrete Laplace's equations to solve the corresponding Dirichlet problem. To obtain a homogeneous distribution of the parameter space over the surface, this initial parameterization is modified in a constrained optimization procedure considering two criteria:

1. Area preservation: Every object region must map to a region of proportional area in parameter space.
2. Minimal distortion: Every quadrilateral should map to a spherical quadrilateral in parameter space.

Brechbühler establishes constraints for area preservation, while the distortion of the mesh serves as the objective function during optimization. The optimization solves the resulting system of nonlinear equations by linearizing them and taking Newton steps.

2.3. Surface Models from SPHARM: SPHARM-PDM

From the SPHARM description we can compute triangulated surfaces by sampling the spherical parameterization uniformly on the sphere. Equidistant sampling in the parameter space leads to a dense sampling around the poles ($\theta = 0$, $\theta = \pi$) and a coarse sampling around the equator ($\theta = \pi/2$). This fact can be explained by the poles being mapped to all points having $\phi = 0 \ldots 2\pi$ and $\theta = 0$, or $\theta = \pi$ (see also Fig. 2). Using a linear, uniform icosahedron subdivision shown in Fig. 3, however, we gain a good approximation of a homogeneous sampling of the spherical parameter space and thus also of the object space. The subdivision is linear in the number of subdivisions along each edge of the original icosahedron, rather than the commonly used recursive subdivisions. Any recursive subdivision has a corresponding linear subdivision, but most linear subdivisions have no corresponding recursive subdivision (see Table 1 and Figure 3).

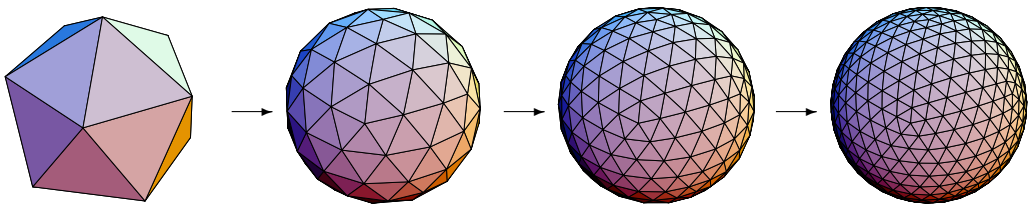

Figure 3. Icosahedron subdivision for different levels of sampling. From left to right: Base icosahedron, subdivision factors 2, 4 and 6.

Using the pre-computed parameter locations (θ_i, ϕ_i) from the icosahedron subdivision, we can compute the PDM directly from the coefficients, as the parameter locations stay constant for all objects at a given subdivision level. In our experience the sampling for the surface of brain structures should be chosen between subdivision level

10 (1002 surface points) and 20 (4002 surface points) depending on the complexity of the objects. For example, hippocampal surfaces are well represented by a subdivision level 10 and caudate surfaces by a subdivision level 15.

2.4. Alignment and Scaling Normalization

Alignment and scaling of the objects are two important issues in shape analysis that are also discussed here.[11] In general, we usually chose rigid-body Procrustes alignment ([18]) followed by no scaling normalization or scaling inversely to the intra-cranial cavity volume (ICV). For a simple normalization with ICV, the scaling information would be $scale_i = ICV_i^{1/3}$. Since shape differences are more easily interpreted in the original unit space (usually millimeter units) the scale information is divided by the mean over all ICV values: $scale_i = ICV_i^{1/3}/\mu_{ICV}^{1/3}$. This is equivalent of scaling each structure such that its overall head size is normalized to a standard head size. Commonly both no scaling and ICV scaling are reported in our shape analysis studies: the original scale analysis represents an unbiased raw analysis, whereas the ICV scale analysis corrects for differences in overall interior head size and thus approximatively normalizes for gender and age differences. In general both types of analysis should yield similar patterns of differences and special care needs to be taken if this is not the case.

2.5. Local Testing of Group Mean Difference using Hotelling T^2 metric

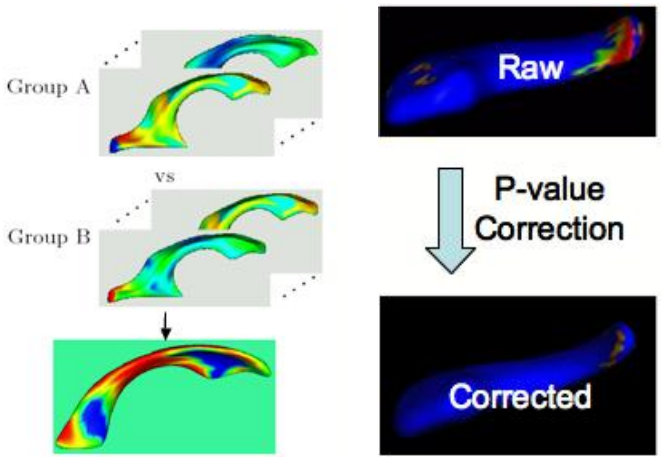

Figure 4. Left: Schematic view of testing procedure: A set of features per surface point (shown as colormaps) are analyzed using uni- or multi-variate statistics, resulting in a raw significance map (bottom) .Right: The raw significance maps are overly optimistic and need to be corrected to control for the multiple comparison problem. The corrected significance maps on the other hand are commonly pessimistic estimates.

After shape description, correspondence establishment, alignment and scaling normalization, the next step in the shape analysis pipeline consists of testing for differences between groups at every surface location (see Figure 4). This can be done in 2 main fashions:

- *Analyzing the magnitude of the local difference to a template:* For this option, a template needs to be first selected, usually this is the common mean of the 2 groups or the mean of separate control group. The magnitude of the difference is easily understood and results in difference maps for each subject. Also, the resulting univariate statistical analysis is quite well known and local significance can be easily computed using the Student t metric. The main disadvantage of this method is the need to select a template, which introduces an additional bias into the statistical analysis, as well as the possible loss of statistical sensitivity by compressing the difference information into a simple magnitude measurement. We applied this method in our earlier studies of hippocampi[19] and ventricles.[20]

- *Analyzing the spatial location of each point:* For this option, no template is necessary and multivariate statistics of the (x, y, z) location is necessary. We have chosen to use the Hotelling T^2 two sample difference metric as a measurement of how 2 groups locally differ from each other. The standard Hotelling T^2 is defined as $T^2 = (\mu_1 - \mu_2)'(\Sigma(\frac{1}{n_1} + \frac{1}{n_2}))^{-1}(\mu_1 - \mu_2)$, where $\Sigma = (\Sigma_1(n_1 - 1) + \Sigma_2(n_2 - 1))/(n_1 + n_2 - 2)$ is the pooled covariance matrix. An alternative modified Hotelling T^2 metric is less sensitive to group differences of the covariance matrixes and the number of samples: $T^2 = (\mu_1 - \mu_2)'(\Sigma_1 \frac{1}{n_1} + \Sigma_2 \frac{1}{n_2})^{-1}(\mu_1 - \mu_2)$. All our current studies are based on this modified Hotelling T^2 metric.

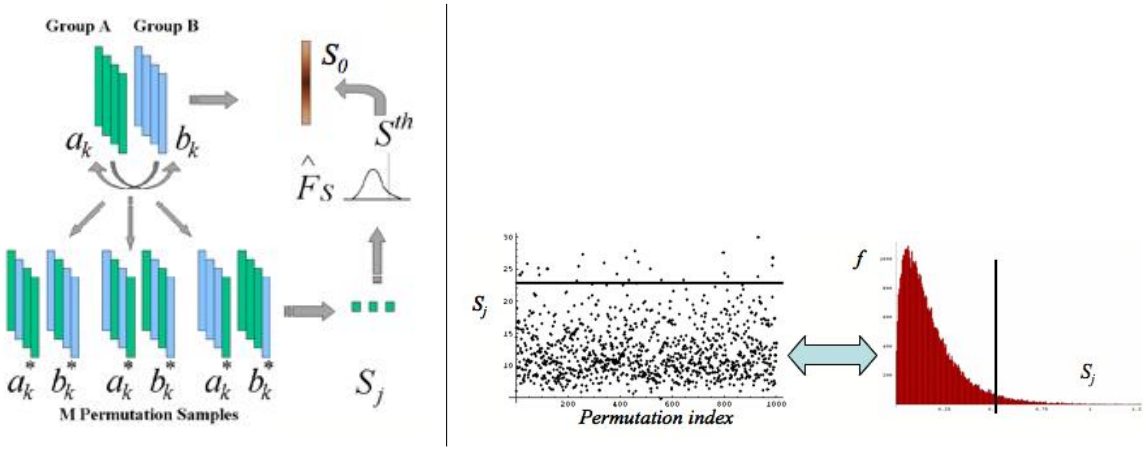

Figure 5. Left: Scheme of P-value computation via permutation tests. The real group difference S_0 (e.g. Hotelling T^2) is compared to that the group differences S_j of random permutations of the group labels. The quantile in the S_j histogram associated with S_0 is the p-value. Right: Example run on test data with 1000 random permutations. On the left the individual S_j are plotted and on the left the corresponding S_j histogram is shown. The black line indicates the value of S_0.

We then want to test the two groups for differences in the means of the selected difference metric (univariate: Student t, multivariate: Hotelling T^2) at each spatial location. Permutation tests are a valid and tractable approach for such an application, as they rely on minimal assumptions and can be applied even when the assumptions of the parametric approach are untenable. Non-parametric permutation tests are exact, distribution free and adaptive to underlying correlation patterns in the data. Further, they are conceptually straightforward and, with recent improvements in computing power, are computationally tractable.

Our null hypothesis is that the distribution of the locations at each spatial element is the same for every subject regardless of the group. Permutations among the two groups satisfy the exchangeability condition, i.e. they leave the distribution of the statistic of interest unaltered under the null hypothesis. Given n_1 members of the first group a_k, $k = 1 \ldots n_1$ and n_2 members of the second group b_k, $k = 1 \ldots n_2$, we can create $M \leq \binom{n_1 + n_2}{n_2}$ permutation samples. A value of M from 20000 and up should yield results that are negligibly different from using all permutations.[21]

2.6. Correction for Multiple Comparison

The local shape analysis involves testing from a few to many thousands of hypothesis (one per surface element) for statistically significant effects. It is thus important to control for the multiple testing problem (see Figure 4), and the most common measure of multiple false positives is the familywise error rate (FWER). The multiple testing problem has been an active area of research in the functional neuroimaging community. The most widely used methods in the analysis of neuroimaging data use random field theory[22][23] and make inferences based on the maximum distribution. In this framework, a closed form approximation for the tail of the maximum distribution is available, based on the expected value of the Euler characteristic of the thresholded image.[23] However, random field theory relies on many assumptions such as the same parametric distribution at each

spatial location, a point spread function with two derivatives at the origin, sufficient smoothness to justify the application of the continuous random field theory, and a sufficiently high threshold for the asymptotic results to be accurate.

In our work, we are employing non-parametric permutation tests and false discovery rate as two alternative correction methods for the multiple comparison problem.

2.6.1. Minimum over Non-parametric permutation tests

Non-parametric methods deal with the multiple comparisons problem[24] and can be applied when the assumptions of the parametric approach are untenable. Similar methods have been applied in a wide range of functional imaging applications.[25]

Details of this method are described in[24] and only a summary is provided here. The correction method is based on computing first the local p-values using permutation tests. The minimum of these p-values across the surface is then computed for every permutation. The appropriate corrected p-value at level α can then be obtained by the computing the value at the α-quantile in the histogram of these minimum values. Using the minimum statistic of the p-values, this method correctly controls for the FWER, or the false positives, but no control of the false negatives is provided. The resulting corrected local significance values can thus be regarded as pessimistic estimates akin to a simple Bonferroni correctection.

2.6.2. False Discovery Rate

Additionally to the non-parametric permutation correction, we have also implemented and applied a False Discovery Rate Estimation (FDR)[26,27] method. The innovation of this procedure is that it controls the expected proportion of false positives only among those tests for which a local significance has been detected. The FDR method thus allows an expected proportion (usually 5%) of the FDR corrected significance values to be falsely positive. The correction using FDR provides an interpretable and adaptive criterion with higher power than the non-parametric permutation tests. FDR is further simple to implement and computationally efficient even for large datasets.

The FDR correction is computed as follows:[27]

1. Select the desired FDR bound q, e.g. 5%. This is the maximum proportion of false positives among the significant tests that you are willing to tolerate (on average).
2. Sort the p-values smallest to largest.
3. Let p_q be the p-value for the largest index i of the sorted p-values $p_{sort,i} \leq q \cdot i/N$, where N is the number of vertices.
4. Declare all locations with a p-value $p \leq p_q$ significant.

2.7. Output of Statistical Analysis

There are two main types of results from the statistical testing step: a) descriptive group statistics, such as the mean and covariance information, and b) group mean difference hypothesis testing. The descriptive statistics are used for quality control, sanity check, as well as these are necessary to make sense of the computed significance maps ("Is that significant region enlarged in the patient population?"). The significance maps show the regions of significant difference of the Hotelling T^2 metric with 3 different maps: the raw p-value map, the FDR corrected p-value map, and the permutation based corrected p-value map. The value of the Hotelling T^2 metric is also useful to visualize the full range of the effect size.

2.8. Computation Scheme

The general workflow starts from the segmented datasets with binary labels for each structure of interest. The first preprocessing program extracts a single binary label and applies heuristic methods to ensure the spherical topology of the segmentation. In the next step, the surface mesh is computed from the segmentation, as well as a spherical parametrization of this surface mesh via area preserving mapping (see Section 2.2). The SPHARM-PDM shape representation is then computed. These three steps have to be run on all datasets in a shape analysis study. Often they are applied in parallel to all datasets in a cluster environment or a multi-CPU processing server. After all SPHARM-PDM representations have been computed, the statistical testing step computes the difference maps, the significance maps, as well as the descriptive statistics such as the mean surfaces and covariance matrixes.

3. RESULTS

The first tests we computed on a schizophrenia study of the hippocampus in order to assess the differences of our template-free analysis methodology to the prior analysis measuring differences to a template.[19] 54 schizophrenis subjects (age: 30.1y (11.9), male gender) and 26 healthy control (age: 31.2y (10.7), male gender) subjects were analyzed. The groups are matched for age and ethnicity. All subjects were scanned on the same 1.5 T scanner and with the same protocol (IR-Prepped SPGR, axial, 0.9375x0.9375x1.5mm^3).

The hippocampi were analyzed with the analysis framework described in this paper, as well as with the prior method that consisted of initially computing an correctly group weighted average and then assessing the univariate signed distance magnitude of each surface to this template.[19] The results are visualized in Figure 6. The correlation between the two methods is clearly visible, as well as the gain in statistical sensitivity. The Hotelling T^2 based analysis shows enlarged regions of significance as well as a new larger region of significance in the hippocampus head.

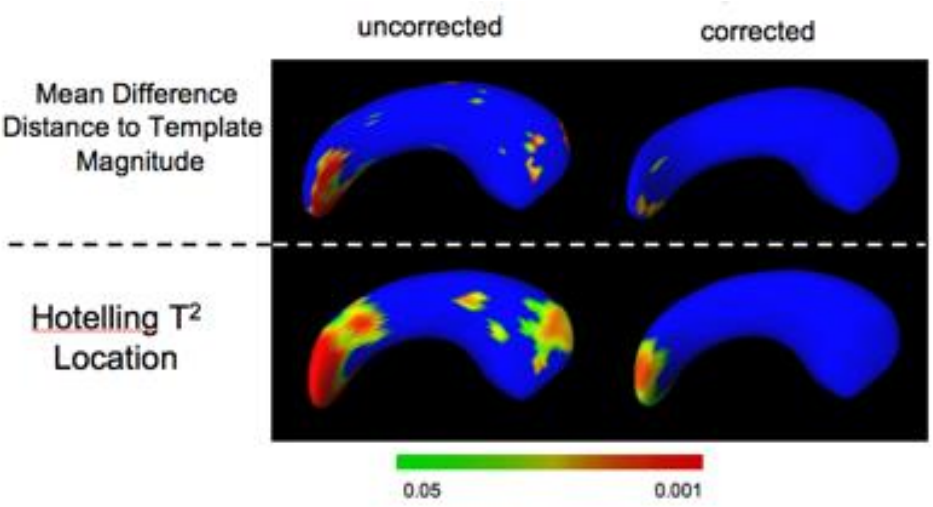

Figure 6. Statistical maps of example shape analysis study on the hippocampus (schizophrenics vs controls) using the univariate mean difference of the distance magnitude to a template as a metric (top) or the template free Hotelling T2 on the location vector of each mesh point. The correction for multiple tests was performed using the pessimistic permutation based method.

As an additional test study we applied our shape analysis framework to a schizo-typal personality disorder (SPD) study on the caudate brain structure that was available to us through the NAMIC network.[28] The data consists of structural SPGR MRI datasets from 15 right handed medication-free male SPD patient, as well as

from 14 matched healthy control subjects. Although both left and right hemispheric caudate were outlined on the MRI data, only the right caudate is presented here as a test study.

The correspondence of the SPHARM-PDM description was visually controlled for all datasets. The descriptive group statistics (see Figure 7) show that the mean objects of the 2 groups are quite similar. The differences between the mean are overall small (max at 1.6 millimeter), but most pronounced in the anterior head region, the superior body region and the posterior tail region *. The visualization of the covariance ellipsoids show that the variability is reduced in the body section compared to the anterior head and especially the posterior tail section.

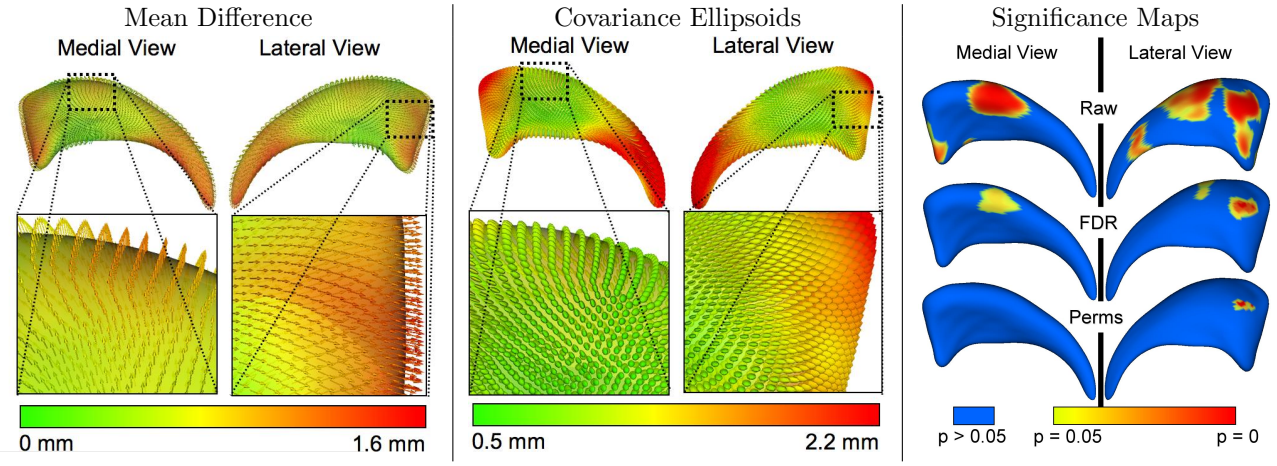

Figure 7. Statistics of right caudate analysis. The left visualization shows the mean difference as a colormap from green (0 mm difference) to red (1.6 mm difference) as well an vectorfield on the overall mean surface mesh. For two regions with large differences a zoomed view is provided. The middle visualization shows the covariance ellipsoid colored with the magnitude of the largest ellipsoid axis (axis of largest covariance). The same two zoomed views are provided. The variance in the first zoom section is markedly reduced compared to the second zoom section.

The results of the mean difference hypothesis tests are visualized in Fiigure 7. The raw significance maps displays an overly optimistic estimate with the superior body, as well as anterior head region to be significantly different. The conservative, permutation based correction map provides a pessimistic estimate with only a single significant region in the anterior head region. The result of the false discovery rate correction is in-between the raw and the permutation based correction. On one hand, significant parts in the anterior tail section are visible in the raw significant map, but are not preserved in the FDR corrected map. On the other hand, significant parts in the superior body as well as the anterior head region are present in the FDR corrected maps. The global shape difference has a p-value of $p = 0.009$. Based on these results, a significant shape difference between the two populations on the right caudate is clearly suggested.

4. CONCLUSION

We have presented a comprehensive, local shape analysis framework for brain structures founded on the SPHARM-PDM shape representation and local statistical tests. The whole software is available as open source and is currently in use by several research groups outside of UNC. It has been applied in several studies on brain morphometry at UNC, as well as by several outside collaborators at the Brigham and Womens Hospital (Boston), GeorgiaTech (Atlanta), University of Illinois (Champaign-Urbana) and several more. The methodology can potentially be employed in 3D shape analysis problems other than brain structure analysis. The main limitation of our shape analysis is the necessity of spherical topology.

The presented studies illustrate the applicability of our testing framework and the comparison to the prior template-based method show clear correlation with enhanced sensitivity for the template-free method.

*The naming of caudate head and tail regions in this section are not based on anatomical labels

The shape analysis framework described in this paper is accessible as part of the UNC NeuroLib library.[29] Instructions for the download precompiled executable as well as the compilation process of the open-source UNC NeuroLib can be found in the tutorial section of the webpage www.ia.unc.edu/dev. The binary distribution also contains sample data, and sample scripts.

We presented research and software development in progress, and many additional enhancements of the current tool base are being developed or are planned. These enhancements include the incorporation of subject covariates, such as age, gender, handedness etc, as well as an user-interface based control tool for the whole shape analysis.

5. ACKNOWLEDGMENT

We are thankful to Steve Pizer and Sarang Joshi for highly valuable discussions in regard to shape analysis as well as to Christian Brechbuehler for providing the initial SPHARM parametrization software. The caudate datasets were provided by Jim Levitt and Martha Shenton, Harvard Medical School. This research has been/is supported by the UNC Neurodevelopmental Disorders Research Center HD 03110, NIH Roadmap for Medical Research, Grant U54 EB005149-01, NIH NIBIB grant P01 EB002779, EC-funded BIOMORPH project 95-0845.

REFERENCES

1. D. Thomson, *On Growth and Form*, Cambridge University Press, second edition, 1942.
2. C. Davatzikos, M. Vaillant, S. Resnick, J. Prince, S. Letovsky, and R. Bryan, "A computerized method for morphological analysis of the corpus callosum," *J. of Comp. Assisted Tomography* **20**, pp. 88–97, Jan./Feb 1996.
3. S. Joshi, M. Miller, and U. Grenander, "On the geometry and shape of brain sub-manifolds," *Pat. Rec. Art. Intel.* **11**, pp. 1317–1343, 1997.
4. J. Csernansky, S. Joshi, L. Wang, J. Haller, M. Gado, J. Miller, U. Grenander, and M. Miller, "Hippocampal morphometry in schizophrenia via high dimensional brain mapping," *Proc. Natl. Acad. Sci. USA* **95**, pp. 11406–11411, September 1998.
5. J. Csernansky, L. Wang, D. J. Jones, D. Rastogi-Cru, G. Posener, J.A.and Heydebrand, J. Miller, U. Grenander, and M. Miller, "Hippocampal deformities in schizophrenia characterized by high dimensional brain mapping," *Am. J. Psychiatry* **159**, pp. 1–7, December 2002.
6. F. Bookstein, "Shape and the Information in Medical Images: A Decade of the Morphometric Synthesis," *Comp. Vision and Image Under.* **66**, pp. 97–118, May 1997.
7. I. Dryden and K. Mardia, "Multivariate shape analysis," *Sankhya* **55**, pp. 460–480, 1993.
8. T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models - their training and application," *Comp. Vis. Image Under.* **61**, pp. 38–59, 1995.
9. G. Gerig, M. Styner, M. Shenton, and J. Lieberman, "Shape versus size: Improved understanding of the morphology of brain structures," in *MICCAI*, pp. 24–32, 2001.
10. C. Brechbühler, G. Gerig, and O. Kübler, "Parametrization of closed surfaces for 3-D shape description," *Comp. Vision, Graphics, and Image Proc.* **61**, pp. 154–170, 1995.
11. G. Gerig, M. Styner, D. Jones, D. Weinberger, and J. Lieberman, "Shape analysis of brain ventricles using spharm," in *MMBIA*, pp. 171–178, IEEE press, 2001.
12. L. Shen, J. Ford, F. Makedon, and A. Saykin, "Hippocampal shape analysis surface-based representation and classification," in *SPIE-Medical Imaging*, 2003.
13. S. Pizer, D. Fritsch, P. Yushkevich, V. Johnson, and E. Chaney, "Segmentation, registration, and measurement of shape variation via image object shape," *IEEE Trans. Med. Imaging* **18**, pp. 851–865, Oct. 1999.
14. M. Styner, G. Gerig, J. Lieberman, D. Jones, and D. Weinberger, "Statistical shape analysis of neuroanatomical structures based on medial models," *Medical Image Analysis* **7**(3), pp. 207–220, 2003.
15. P. Golland, W. Grimson, and R. Kikinis, "Statistical shape analysis using fixed topology skeletons: Corpus callosum study," in *IPMI*, pp. 382–388, 1999.
16. A. Kelemen, G. Székely, and G. Gerig, "Elastic model-based segmentation of 3d neuroradiological data sets," *IEEE Trans. Med. Imaging* **18**, pp. 828–839, October 1999.

17. M. Styner, K. Rajamani, L. Nolte, G. Zsemlye, G. Szekely, C. Taylor, and R. H. Davies, "Evaluation of 3d correspondence methods for model building," in *IPMI*, pp. 63–75, July 2003.
18. F. Bookstein, *Morphometric Tools for Landmark Data: Geometry and Biology*, Cambridge University Press, 1991.
19. M. Styner, J. Lieberman, D. Pantazis, and G. Gerig, "Boundary and medial shape analysis of the hippocampus in schizophrenia," *Medical Image Analysis* **8**(3), pp. 197–203, 2004.
20. M. Styner, J. Lieberman, R. McClure, D. Weingberger, D. Jones, and G. Gerig, "Morphometric analysis of lateral ventricles in schizophrenia and healthy controls regarding genetic and disease-specific factors," *Proceedings of the National Academy of Science* **102**, pp. 4872–4877, March 2005.
21. E. S. Edgington, ed., *Randomization Tests*, Academic Press, 1995.
22. M. K. Chung, *Statistical Morphometry in Neuroanatomy*. PhD thesis, McGill University, Montreal, 2001.
23. K. J. Worsley, S. Marrett, P. Neelin, A. C. Vandal, K. J. Friston, and A. C. Evans, "A unified statistical approach for determining significant signals in images of cerebral activation," *Human Brain Mapping* **4**, pp. 58–73, 1996.
24. D. Pantazis, R. Leahy, T. Nichol, and M. Styner, "Statistical surface-based morphometry using a nonparametric approach," in *Int. Symposium on Biomedical Imaging(ISBI)*, pp. 1283–1286, April 2004.
25. T. Nichols and A. Holmes, "Nonparametric analysis of pet functional neuroimaging experiments: A primer with examples," *Hum Brain Mapp* **1**, pp. 1–25, Jan 2001.
26. Y. H. Y Benjamini, "Controlling false discovery rate: A practical and powerful approach to multiple testing," *Stat. Soc. Ser. B* (57), pp. 289–300, 1995.
27. C. Genovese, N. Lazar, and T. Nichols, "Thresholding of statistical maps in functional neuroimaging using the false discovery rate," *NeuroImage* (15), pp. 870 – 878, 2002.
28. J. Levitt, R. McCarley, C. Dickey, M. Voglmaier, M. Niznikiewicz, L. Seidman, Y. Hirayasu, A. Ciszewski, R. Kikinis, F. Jolesz, and M. Shenton, "Mri study of caudate nucleus volume and its cognitive correlates in neuroleptic-naive patients with schizotypal personality disorder," *Am J Psychiatry* **159**(7), pp. 1190–1197, 2002.
29. M. Styner, M. Jomier, and G. Gerig, "Closed and open source neuroimage analysis tools and libraries at unc," in *IEEE Symposium on Biomedical Imaging ISBI*, pp. 702–705, 2006.