ROI CONSTRAINED STATISTICAL SURFACE MORPHOMETRY

Chunxiao Zhou[1,2,5], Denise C. Park[3,5], Martin Styner[6], Yongmei Michelle Wang[1,3,4,5]

Departments of [1]Statistics, [2]Electrical&Computer Engineering, [3]Psychology, [4]Bioengineering,
[5]Beckman Institute,
University of Illinois at Urbana-Champaign, IL, USA

[6]Departments of Computer Science and Psychiatry,
University of North Carolina at Chapel Hill, NC, USA

ABSTRACT

This paper presents a novel ROI constrained statistical surface analysis framework that aims to accurately and efficiently localize regionally specific shape changes between groups of 3D surfaces. With unknown distribution of the data, existing shape morphometry analysis involves testing thousands of hypotheses for statistically significant effects through permutation. In this work, we develop a hybrid method to improve the system's efficiency by computing the raw p-values of the nonparametric permutation tests only within a region of interest (ROI) of the surface. The ROI is identified through a parametric Pearson type III distribution approximation. Furthermore, a ROI based adaptive procedure is utilized to control the False Discovery Rate (FDR) for increased power of finding the significance.

Index Terms — Brain morphometry, MRI, permutation, hypothesis testing, FDR.

1. INTRODUCTION

Recent advances in MRI technology have led to increasing interest in statistical shape analysis of brain due to its ability to detect morphological changes in structures of interest for neuroscience research, as well medical diagnosis and treatment. A critical issue in surface morphometry is the shape description and representation. Different strategies have been investigated in the literature [3, 8, 9]. The spherical harmonics (SPHARM) approach using spherical harmonics as basis functions for a parametric surface description was proposed in [2]. The sampled SPHARM-PDM is a smooth, accurate, fine-scale shape representation [7]. The correspondence across different surfaces is established by aligning the parameterizations via the first order ellipsoid. The present work employs the SPHARM-PDM description due to the mentioned advantages and its successful applications [6]. The SPHARM-PDM leads to corresponding location vectors over different surfaces, which are then used in our subsequent statistical analysis.

At each position on the surface, we test whether there is significant mean vector difference between two groups of location vectors. Since the distribution of the location vectors is unknown, a reasonable strategy is to use a nonparametric approach for hypothesis testing. Permutation tests require few assumptions concerning statistical distributions but exchangeability. They belong to the nonparametric "distribution-free" category of hypothesis testing and are thus flexible, and have been used successfully in both structural and functional MR image analysis [5, 6]. There are three main approaches to construct the permutation distribution [4]. First, exact permutation enumerates all possible arrangements. Second, an approximate permutation distribution bases on random sampling from all possible permutations. Third, permutation distribution approximation uses the analytical moments of the exact permutation distribution under the null hypothesis. The computational cost is the main disadvantage of the exact permutation, due to the factorial increase in the number of permutations with the increasing number of subjects. The computation load may be even heavier when we look at multiple factorial analyses. The second technique has the problem of replication and more type I error. The limitation of the last method is whether the moments can be obtained. In this work, we define a simplified T2 test statistic from which the first three moments of the exact permutation distribution can be derived analytically. The regions of interest (ROI) on the surface that possibly have significant shape differences are thus identified by fitting the permutation distribution with the first three moments. The real permutation is then only performed within the ROI rather than over the entire surface. This approach has the advantage of achieving a good comprise between flexibility and efficiency by using a hybrid version of parametric tests and nonparametric permutation tests.

Analysis of the location vectors involves testing thousands of hypotheses, i.e. multiple testing. Simply choosing a type I error threshold, for example with α =0.05 is not appropriate. In this case, 100 false positives would be

expected when 2,000 locations on the surface are compared. False positives must be controlled over all tests. A common technique to handle the multiple testing problem is controlling the familywise error rate (FWER), which however, is too conservative for many applications [1]. False Discovery Rate (FDR) control is becoming more popular due to its power of finding true discoveries through the control of the fraction of false discoveries over total discoveries [1]. The FDR method was initially proposed by Benjamini and Hochberg, so called BH's FDR. More advanced techniques have also been developed recently, such as Storey's pFDR and BH's adaptive FDR [1]. In this paper, we adopt the concept of the adaptive FDR and develop a ROI based adaptive procedure using BH's FDR to enhance the power of finding true discoveries.

2. METHODOLOGY

2.1. Overview

We first use the SPHARM-PDM software [7] to establish surface correspondence and align the surface location vectors over all subjects. The surface shapes of different objects are thus represented by the same number of location vectors. Second, a multivariate simplified T^2 test statistic is defined, and the ROI is identified by using the Pearson type III distribution to approximate the permutation distribution of the normalized test statistic. Third, within the ROI, we permute the corresponding location vectors and measure the p-values of the observed statistics. Finally, a ROI constrained adaptive FDR strategy is utilized to find the regions with significant group shape differences.

Two main contributions of this paper are: i) develop a hybrid method to improve the system's efficiency by computing the nonparametric permutation tests for the raw p-values only within a ROI of the surface, which is identified through a parametric type III distribution approximation; ii) increase the power of the multiple testing by using an ROI constrained adaptive FDR.

2.2. Test Statistic and Region of Interest (ROI)

Suppose we have two groups of location vectors for each corresponding surface position, and the numbers of subjects in the two groups are respectively n_1 and n_2. To analyze the surface shape difference between two groups, we test the null hypothesis of each spatial location one by one independently, and then detect the locations with significant shape changes. The null hypothesis H_0^i is defined as no group mean vector difference at i^{th} location, $i = 1, 2, \cdots, m$, where m is the total number of locations on the surface.

2.2.1. Definition of Test Statistic

The type III distribution is utilized in the present work to approximate the permutation distribution of the test statistic without real permutation, which can be achieved through the use of the first three moments of the exact permutation distribution. Hotelling's T^2 is an optimal multivariate test statistic for mean vector difference testing under multivariate normality. However, it is impossible to compute the pooled covariance matrix without real permutation. We solve this puzzle by replacing the pooled covariance matrix in the Hotelling's T^2 with the sample variance-covariance matrix, S, over all the subjects of the two groups. This leads to a simplified T^2 (denoted as T in this paper), which is the test statistic used here. Since the simplified T^2 is a monotonic function of Hotelling's T^2, this two statistics share the same p-value in permutation tests.

2.2.2. Approximation through Type III Distribution

Let X denote the location vector matrix over all subjects in both groups, i.e., $X = \begin{bmatrix} x_1 & \cdots & x_{n_1} & x_{n_1+1} & \cdots & x_{n_1+n_2} \\ y_1 & \cdots & y_{n_1} & y_{n_1+1} & \cdots & y_{n_1+n_2} \\ z_1 & \cdots & z_{n_1} & z_{n_1+1} & \cdots & z_{n_1+n_2} \end{bmatrix}$. Let $H = I_{n_1+n_2} - \frac{1}{n_1+n_2} 1_{n_1+n_2} 1_{n_1+n_2}{}^T$ denote centering matrix, where $I_{n_1+n_2}$ is a $(n_1+n_2)\times(n_1+n_2)$ identity matrix, and $1_{n_1+n_2}$ is a vector of size n_1+n_2 with all entries 1. Note that H is symmetric and idempotent, thus $HH^T = H$. The $(n_1+n_2)\times(n_1+n_2)$ permutation matrix P randomly assigns n_1 subjects to one group and the rest n_2 subjects to another group. All $\binom{n_1+n_2}{n_1}$ configurations are equally distributed. Let C be the mean difference vector $[\frac{1}{n_1} 1_{n_1}{}^T, \frac{-1}{n_2} 1_{n_2}{}^T]$. We derive the first moment of the exact permutation distribution as following: $E(T) = E(CPX^T S^{-1} XP^T C^T)$

$$= (n_1+n_2-1)E(CPX^T((XH)(XH)^T)^{-1}XP^T C^T)$$
$$= (n_1+n_2-1)E(trace(CPX^T(XHH^T X^T)^{-1}XP^T C^T))$$
$$= (n_1+n_2-1)trace((XHX^T)^{-1}X(\frac{1}{n_1+n_2-1})CC^T HX^T)$$
$$= CC^T trace((XHX^T)^{-1}XHX^T) = dCC^T = 3(\frac{1}{n_1}+\frac{1}{n_2}) \quad (1)$$

where d is the dimension of location vector, i.e. $d = 3$, and the lemma of $E(P^T C^T CP) = CC^T H/(n_1+n_2-1)$ is used. Note that the first moment of permutation distribution turns out not dependent on the location vector matrix X, which then only needs to be computed once for the entire surface.

For mathematical tractability, we adopt the idea of the indictor function in [4] to analytically calculate the second and third moments. The calculation involves measuring the expected values of indictor function in different situations. The final result is a sum of multiple terms with each term a product of a function of location vectors and a coefficient only related to n_1 and n_2. The second and third moments are location based and differ from surface voxel to voxel.

The type III distribution is standardized to have mean zero, variance 1 and skewness γ:

$$f(y)=\frac{(2/\gamma)^{4/\gamma^2}}{\Gamma(4/\gamma^2)}((2+y\gamma)/\gamma)^{(4-\gamma^2)/\gamma^2}\exp(-2(2+y\gamma)/\gamma^2)$$

where $-2/\gamma<y<\infty$ and Γ is the gamma function. We construct this type III function with the derived first three moments to approximate the permutation distribution of the normalized test statistic as below [4]:

$$\gamma=(E(T^3)-3E(T)Var(T)-E(T)^3)/\mathrm{var}(T)^{3/2}$$

In fact, less moments (first two) or more moments may be used to approximate the permutation distribution. The type III distribution using the first three moments achieves a good balance between the model complexity and accuracy.

2.2.3. Identification of ROI

Given an observed statistic T (i.e., simplified T^2), we first normalize it according to $Z=(T-E(T))/\sqrt{Var(T)}$. Then, the corresponding critical value at the pre-chosen significance level α_{ROI} is calculated based on the constructed type III distribution. Finally, the voxel is included into the ROI if its Z value is larger than the critical value.

Under multivariate normality and equal group variance, the Hotelling's T^2 is distributed as $\frac{(n_1+n_2-2)d}{(n_1+n_2-d-1)}F_{d,n_1+n_2-d-1}$. One could define the ROI by simply thresholding the observed values of the Hotelling's T^2 test statistic with this F distribution at α_{ROI}. However, the normality assumption may not be satisfied in real applications, which will lead to inaccurate ROI identification (see Section 3 and Fig. 1).

2.3. Raw p-values through Permutation within ROI

Within the ROI detected by the above permutation distribution approximation, either exact permutations or random permutations with Monte Carlo strategy can be applied to construct the permutation distribution, based on which the p-values of the observed statistics are calculated.

2.4. Adaptive FDR within ROI

The BH's FDR is defined as (see notations in Table 1):

$$FDR=E(\frac{V}{R\vee 1})=E(\frac{V}{R}\Big|R>0)\,P(R>0),\quad \text{where } R\vee 1=\max(R,1).$$

It is more powerful and less stringent than the FWER.

	Accept null	Reject null	Total
Null true	U	V	m_0
Alternative true	G	L	m_1
Total	W	R	m

Table 1. Possible outcomes from *m* hypothesis tests.

In this work, we adopt an adaptive procedure of the BH's FDR [2], with detailed steps as following:

Step 1: For $\{p_i \,|\, location\ i\in ROI\}$, order the set of observed p-values as $p_{(1)}\le p_{(2)}\le\cdots\le p_{(m_r)}$, where m_r is the number of locations within ROI.

Step 2: Starting with $p_{(m_r)}$ in decreasing order, find the largest k for which $p_{(k)}\le k\frac{q}{m}$, where q is the preset FDR error rate and m is the total number of location vectors on the surface.

Step 3: Estimate the number of null locations by using $\hat{m}_0=(1+q)(m-k)$.

Step 4: Calculate the "BH-FDR-adjusted p-values" for locations within ROI through $p_{(i)}^{BH}=\min\{p_{(j)}\hat{m}_0/j\,|\,j\ge i\}$.

Step 5: Set $p_i^{BH}=1,\ if\ location\ i\notin ROI$.

Step 6: Compare all p_i^{BH} with desired level of FDR q; all $p_i^{BH}\le q$ are declared as significant.

This adaptive FDR control is more powerful than the conventional BH's FDR. It can find more significant areas while preserve the same desired FDR rate. The ROI constrained adaptive scheme in this work is more efficient since only the p-values within the ROI need to be sorted. We set the p-values outside the ROI to 1 because it is unlikely to change the ranking of the p-values of true significant, therefore, this procedure won't change the significance detection result with the adaptive FDR strategy.

3. RESULTS

We applied the method to the MRI hippocampi[1] that were semi-automatically segmented by human expert raters and manually grouped into 2 groups (group A and group B) with 21 subjects in each group. This dataset serves as a testing dataset for methodology validation and as an example for all users of the SPHARM-PDM software.

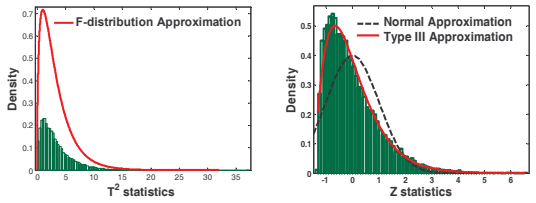

Fig.1. Left: *F*-distribution approximation of the Hotelling's T^2 statistic permutation distribution. Right: Normal and Type III approximation of our *Z* statistic permutation distribution.

Fig. 1 demonstrates the performance of the type III distribution approximation, by showing the fitting results of a randomly selected voxel on the hippocampus surface. The green graph on the left is the histogram of the Hotelling's T^2 statistic, and the one on the right is the histogram of our Z statistic (simplified and normalized T^2 statistic) Fig. 1 right shows that approximating the permutation distribution of

[1] http://www.ia.unc.edu/dev/download/shapeAnalysis/

this normalized test statistic with type III distribution (using the first three moments) is more precise than with a normal distribution (using only the first two moments). In addition, the fitting of the Hotelling's T^2 permutation distribution by $\frac{(21+21-2)\times 3}{(21+21-3-1)}F_{3,38}$ distribution (Fig.1 left red) is obviously unsuitable for this voxel and for this dataset.

Evaluation of our ROI identification method on the example hippocampus dataset is shown in Fig. 2. It can be seen that compared with approximating the permutation distribution using parametric F distribution, the type III distribution approximation using the first three moments leads to less discrepancy with the raw p-value map from real permutation. This further confirms that the present ROI identification strategy is reasonable and accurate. Note that at the same significance level α_{ROI}, even if a few significant surface locations based on the real permutation ROI identification are missed in our ROI identification, these voxels would have raw p-values very close to α_{ROI} due to the precise type III distribution approximation, and would be shown as non-significant after false positive error control (such as FDR). Alternatively, we can set α_{ROI} slightly larger than the subsequent preset DFR rate, α_{FDR} (for example, choose α_{ROI}=0.055 for α_{FDR}=0.05), to guarantee that all locations with raw p-values smaller than or equal to α_{FDR} (based on real permutation) are included in our ROI.

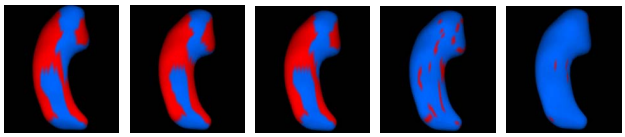

Fig.2. Comparison of techniques in ROI identification (at significance level 0.05). a): ROI (in red) detected by the raw p-map through real permutation. b): ROI (in red) detected by F distribution. c): ROI (in red) detected by type III distribution. d): The difference of b) and a). e): The difference of c) and a).

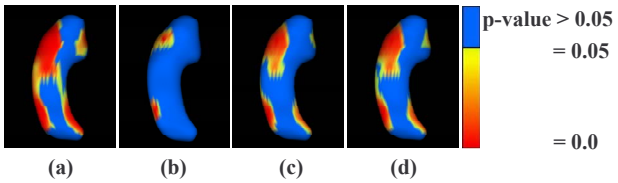

Fig.3. Comparison of multiple testing techniques. a): Raw p-map. b): FWER corrected p-map. c): BH's FDR corrected p-map. d): ROI constrained adaptive FDR corrected p-map.

The false positive error control results are shown in Fig. 3. The raw p-map has the largest significance region, including numerous false positives. Since FWER control is conservative, it leads to the smallest size of significance areas. Our ROI constrained adaptive FDR method discovers more significant locations than the conventional BH's FDR method under the same FDR level (0.05).

4. CONCLUSION

A new ROI constrained surface morphometry analysis method is developed and presented. The ROI is identified by using the analytically derived first three moments of the permutation distribution. The real permutation is then performed only within the ROI rather than the entire object surface for considerably reduced computation cost. Our hybrid strategy takes advantage of nonparametric permutation tests and parametric type III distribution approximation to achieve both accuracy/flexibility and efficiency. Furthermore, a ROI based adaptive procedure is employed to control the False Discovery Rate (FDR) for increased power of the multiple testing. Experimental results demonstrate the effectiveness of the present statistical analysis method. Compared with the conventional FWER and BH's FDR control, our ROI based adaptive FDR control is more powerful while keeping the false positive rate below the desired level as well.

5. ACKNOWLEDGEMENTS

This work was supported in part by grants from the Beckman Institute (seed proposal), the NIH/NIA (2 R37 AG006265-21), the University of Illinois Research Board (Award # 05037), and the National Alliance for Medical Image Computing (NAMIC, NIH U54 EB005149).

REFERENCES

[1] Y. Benjamini *et. al*., "Adaptive linear step-up false discovery rate controlling procedures," *Biometrika*. 93(3): 491-507, 2006.

[2] C. Brechbühler, G. Gerig, and O. Kübler, "Parametrization of closed surfaces for 3-D shape description," *Comp. Vision, Graphics, and Image Proc.,* 61: 154–170, 1995.

[3] P. Golland, W. Grimson, and R. Kikinis, "Statistical shape analysis using fixed topology skeletons: Corpus callosum study," *Info. Proc. In Med. Imag.*, pp. 382–388. 1999.

[4] L. Hubert, *Assignment methods in combinatorial data analysis,* Marcel Dekker, Inc., 1987.

[5] T.E. Nichols, and A.P. Holmes, "Nonparametric permutation tests for functional neuroimaging: A primer with examples," *Human Brain Mapping* 15: 1–25, 2001.

[6] D. Pantazis, R.M. Leahy, T.E. Nichols, and M. Styner, "Statistical surface-based morphometry using a non-parametric approach," *IEEE Int. Sym. Biomedical Imaging,* pp. 1283-1286, 2004.

[7] M. Styner, I. Oguz, *et. al*., "Framework for the statistical shape analysis of brain structures using SPHARM-PDM," *Open Science Workshop at MICCAI 2006, Insight Journal*, Dspace link http://hdl.handle.net/1926/215.

[8] P.M. Thompson, *et. al.*, "Mapping hippocampal and ventricular change in Alzhimer disease," *Neuroimage* 22: 1754-1766, 2004.

[9] Y. Wang, and L.H. Staib, "Boundary finding with prior shape and smoothness models," *IEEE Trans. on Pattern Analysis and Machine Intelligence* 22(7): 738-743, 2000.