

# PCA-based filtering for hypothesis testing on hippocampus shape differences between schizophrenics and normals

S. Olmos<sup>a</sup>, M. Bossa<sup>a</sup> and M Styner<sup>b\*</sup>

<sup>a</sup>I3A, Zaragoza University, Spain. <sup>b</sup>Computer Science Dep., North Carolina University, USA

**Abstract.** Statistical shape analysis has become of increasing interest to the neuroimaging community due to its potential to precisely locate morphological changes and thus potentially discriminate between healthy and pathological structures. A common shape description is based on the sampled boundary defined by a spherical harmonic SPHARM description. Shape analysis is computed by means of hypothesis testing of differences from an average template structure using multivariate group mean difference tests. The presence of noise in the landmark location or correspondence may obscure the statistical significance at certain locations of the structure. This work shows that statistical significance is improved when hypothesis testing is done on reconstructed shapes by means of PCA. Performance evaluation was done on a hippocampus study for schizophrenic patients, confirming deformations previously found in the head and tail regions of the right hippocampus.

## 1 Introduction

Obtaining knowledge about the morphological characteristics of brain structures is an important and challenging problem in medical image analysis. One approach is volumetric analysis, which is based on measuring the volume of a structure of interest. The main advantage of volumetric analysis is its simplicity; however, many structural differences may be overlooked. On the other hand, shape analysis has the potential to provide important information above and beyond simple volume measurements and may characterize abnormalities in the absence of volume differences.

Shape analysis of densely sampled 3D point distribution models (PDM) and their deformations was first investigated by Cootes et al. [1]. Later, Gerig et al. proposed shape analysis based on a parametric boundary description called SPHARM [2]. Several studies used later the implied PDM in the SPHARM approach [3,4].

Several researchers have focused their analysis on the hippocampus [4–7]. Some authors search shape differences at each boundary point by means of hypothesis testing mean difference [7], yielding a regional shape analysis. On the other hand Shen et al. [4] formulate a classification problem in a feature space defined by a subset of selected coefficients obtained by principal component analysis (PCA). These features are a global shape characterization.

PCA is a very well-known technique for multivariate statistical analysis, very often used for shape analysis [1]. The basis functions are the dominant eigenvectors of the covariance matrix. The subspaces spanned by the dominant eigenvectors explain most of the shape variance. In this way, PCA is often used for filtering purposes [8].

The goal of this work was to evaluate the effect of PCA-based filtering on local hypothesis testing of mean shape differences between healthy controls and schizophrenic patients.

## 2 Methods

SPHARM is a parametric boundary shape description that can represent spherical topology objects. Truncating the basis functions, which are spherical harmonics, results in object models at different levels of detail. By means of a mapping of the basis functions to the sphere, a PDM can be obtained [7]. The mapping function is designed to get a mesh with an equal area distribution and minimum angular distortion. 

The presence of noise in the landmark locations of the SPHARM-PDM may obscure the statistical significance of the hypothesis testing of differences. Several noise sources may be present, among them manual segmentation errors or a non-accurate landmark correspondence. One way to reduce noise variance is by means of using linear expansions that simultaneously provide dimensionality reduction and filtering [8]. Noise variance is reduced when a subset of the basis functions are used in the expansion. The lower the dimensionality is used, the greater the noise variance reduction will be obtained. However, this noise reduction is inherently related to a reduction of the

---

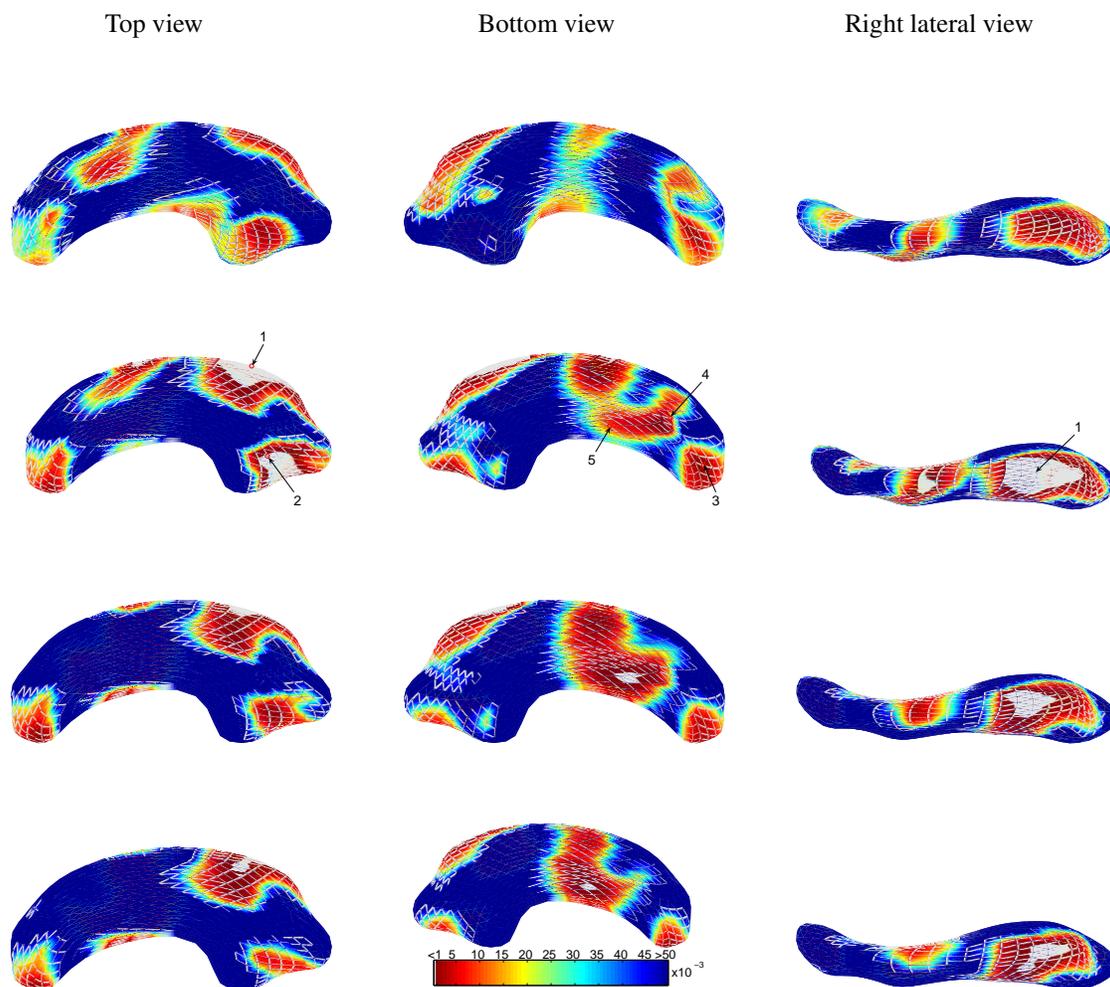
\*olmos@unizar.es, Aragon Institute of Engineering Research, Mara de Luna 1, 50018-Zaragoza (Spain)

shape space dimensionality, which may introduce bias in the reconstructed shape. Quantification of this effect was studied in detail in for electrocardiogram compression, where several noise distributions where analyzed [8].

In the basis functions are selected are the dominant eigenvectors of the covariance matrix, as in PCA, a low dimensionality shape space is obtained where most of the shape variance is represented. The low dimensionality provides an important noise variance reduction, which hopefully improve statistical significance of the hypothesis testing of shape differences.

### 3 Results

Statistical shape analysis was performed in a hippocampus dataset of schizophrenic patients and healthy controls described elsewhere in more detail [7, 9]. We investigated the shape of the hippocampus structure in the left and right brain hemisphere in schizophrenic patients (56 cases) and healthy controls (26 cases). The subjects in this study are all made gender and same handedness. The two populations are matched for age and ethnicity. The hippocampi were segmented from IR SPGR MRI datasets ( $0.93 \times 0.93 \times 1.5$  mm) using a manual outlining procedure [9]. The objects were normalized via a rigid-body Procrustes alignment and scaled to unit volume. The SPHARM-PDM was computed using a sampling of 1002 points along the boundary [7]. PCA was performed on the whole set of 82 aligned shapes.



**Figure 1.** Statistical maps of the local shape analysis of right hippocampus, with several numbers of eigenmodes for PCA: 7,10,20 and 81 from top to bottom. Larger regions with stronger significance are found on reconstructed hippocampus with 10 modes. White color represents regions with p-values lower than 0.001.

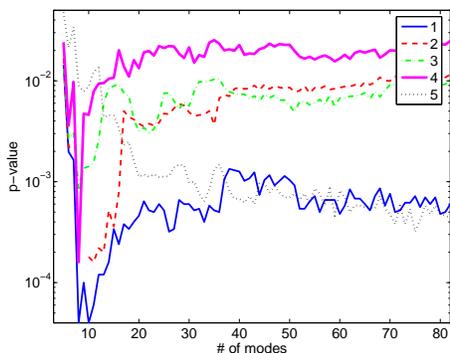
The template of healthy controls and schizophrenic patients was computed as the average of each landmark location. For each landmark, a multivariate Hotelling  $T^2$  test including x,y,z coordinates was computed as mean difference statistical test. A non-parametric permutation test was used in order to estimate the p-value with 50000

permutations.

Statistical shape analysis was performed on reconstructed shapes with a varying number of PCA modes. Fig. 1 shows the p-value maps of right hippocampus for the following number of modes: 7, 10, 20 and 81. The last value corresponds to the case of no-filtering. No big regions with significant differences were found on the left hippocampus. In this paper we only show results on the right hippocampus.

According to the results shown in Fig. 1, PCA filtering with 10 eigenmodes provides p-value maps with larger regions with significant differences in the mean templates. In addition, the statistical significance is notably increased.

In order to illustrate this effect, we selected five landmarks pointed with arrows in the second row of Fig. 1. Each landmark is located at the biggest and most significant regions at any number of modes. The p-values at these locations are shown in Fig. 2. In general, each landmark attains the minimum p-value at different number of modes, but for most of them the minimum is achieved for dimensions in the interval 8-10. Landmarks 1-4 showed a fast and steep improvement of statistical significance at very low PCA dimensionalities (below 10), while for larger dimensions a poorer significance was obtained due to the increase of noise variance. On the other hand landmark 5 did not experience a degradation of the statistical significance at high dimensionalities. Landmark 5 was located in a smaller region and with larger p-value.



**Figure 2.** Statistical significance at specific landmark locations specified in the second row of Fig. 1 for varying number of eigenmodes.

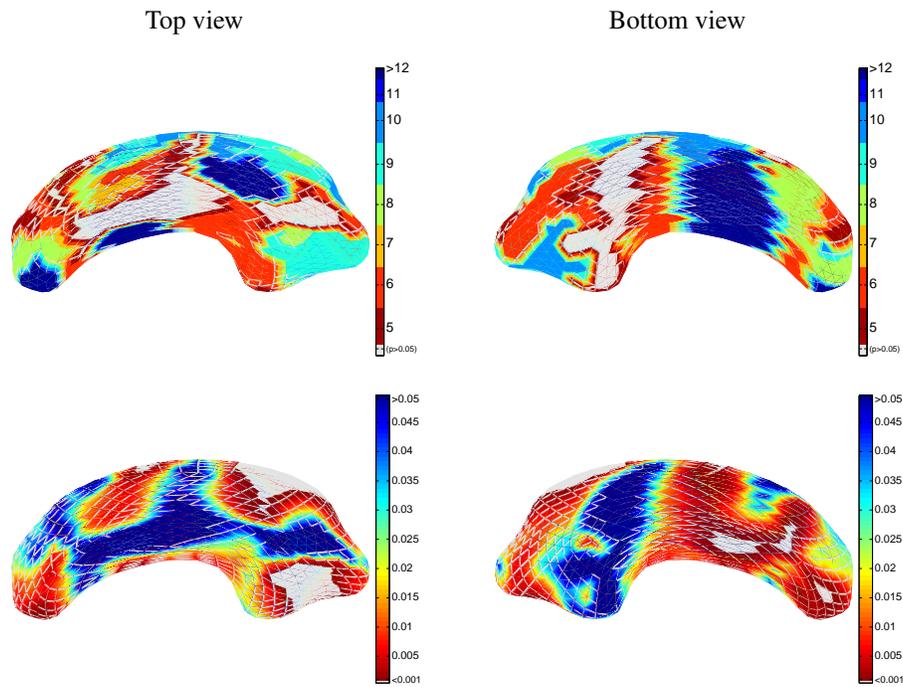
Statistical maps shown in Fig. 1 shows the sensitivity of the dimensionality parameter. In order to make the selection of this parameter more easy, a new map was built with the PCA dimensionality that minimized the p-value at each landmark. This map is shown at the top of Fig. 3. Non-significant regions are represented as white color. The map with the minimum p-values is shown at the bottom row of Fig. 3.

The most significant regions (white color regions at the bottom row in Fig. 3) are mostly obtained with dimensionalities from 8 to 10.

## 4 Discussion and conclusions

In a previous study [7] significant shape differences between hippocampi from schizophrenic patients and healthy controls were mainly found at the tail of the right hippocampus. A univariate t-test was computed at each boundary point on the signed distance between templates. In this work a multivariate  $T^2$  test was used with the three spatial coordinates instead. The p-value maps at the bottom row of Fig. 1 correspond to the case of no PCA-filtering. The regions with statistically significant differences in the no-filtered shapes became larger and more significant when PCA-filtering, achieving the maximum significance when about 10 PCA modes were used. We hypothesize that the noise reduction achieved by PCA-filtering was the main reason for such improvement.

In this work we have shown that PCA-based filtering of SPHARM-PDM provides noise reduction. As a consequence, multivariate Hotelling  $T^2$  hypothesis testing on filtered hippocampi allowed to find new and more statistically significant regions with differences between healthy controls and schizophrenic patients. Further research should focus on the clinical implications of these findings.



**Figure 3.** Top row: map with the number of modes with largest statistical significance. In this case white color is used for non-significant ( $p \geq 0.05$ ) regions. Bottom row: corresponding minimum p-value. White color represent highly significant regions.

## Acknowledgements

We are thankful to Guido Gerig and Neuroimage analysis lab at UNC for providing original MR data and the hippocampus segmentation of schizophrenia study. This work was funded by grants PI04/1795 and PI04/0689 from Health Research Foundation (Spain). The work of M Bossa is supported by the Aragon Government under FPI grant B097/2004. The hippocampal schizophrenia study was funded by the Stanley Foundation.

## References

1. T. Cootes, C. Taylor, D. Cooper et al. "Active shape models-their training and application." *Comp. Vis. Graphica Image Proc.* **61**, pp. 38–59, 1995.
2. G. Gerig, M. Styner, M. Shenton et al. "Shape versus size: improved understanding of the morphology of brain structures." In *MICCAI*, pp. 24–42. 2001.
3. G. Gerig, M. Styner, D. Jones et al. "Shape analysis of brain ventricles using spharm." In *MMBIA, IEEE Press*, pp. 171–178. 2001.
4. L. Shen, J. Ford, F. Makedon et al. "Hippocampal shape analysis surface-based representation and clasification." In *SPIE-Medical Imaging*. 2003.
5. J. G. Csernansky, S. Joshi, L. Wang et al. "Hippocampal morphometry in schizophrenia by high dimensional mapping." In *Proc. Natl. Acad. Sci.*, volume 95, pp. 11406–11411. 1998.
6. J. G. Csernansky, L. Wang, J. Swank et al. "Preclinical detection of alzheimer's disease: hippocampal shape and volume predict dementia onset in the elderly." *Neuroimage* **25**, pp. 783–792, 2005.
7. M. Styner, J. A. Lieberman, D. Pantazis et al. "Boundary and medial shape analysis of the hippocampus in schizophrenia." *Medical Image Analysis* **8**, pp. 197–203, 2004.
8. S. Olmos, J. Garcia, R. Jané et al. "ECG signal compression plus noise filtering with truncated orthogonal expansions." *Signal Processing* **79(1)**, pp. 97–115, Oct 1999.
9. S. Schobel, M. Chakos, G. Gerig et al. "Duration and severity of illness and hippocampal volume in schizophrenia as assessed by 3d-manual segmentation." *Schizophrenia Research* **49(1-2)**, pp. 165, 2001.