Performance Issues in Shape Classification

Samson J. Timoner[1], Pollina Golland[1], Ron Kikinis[2], Martha E. Shenton[3], W. Eric L. Grimson[1], and William M. Wells III[1,2]

[1] MIT AI Laboratory, Cambridge MA, USA
{samson,polina,welg,sw}@ai.mit.edu
[2] Brigham and Women's Hospital, Harvard Medical School, Boston MA, USA
kikinis@bwh.harvard.edu
[3] Laboratory of Neuroscience, Clinical Neuroscience Division, Department of Psychiatry, VAMC-Brockton, Harvard Medical School, Brockton, MA.
mshenton@warren.med.harvard.edu

Abstract. Shape comparisons of two groups of objects often have two goals: to produce a classifier to separate the two groups and to provide information to show differences between the groups. We examine issues that are important for shape analysis in a study comparing schizophrenic patients to normal subjects. We show that for this study, non-linear classifiers provide large accuracy gains over linear ones. Using volume information directly in the classifier provides gains over a classifier that normalizes the data for volume. Alignment methods also make small differences in classifier results. We compare two different representations of shape: displacement fields and distance maps. We show that for this study, the classifier based on displacement fields outperforms the one based on distance maps. We also show that displacement fields provide more information in visualizing shape differences than distance maps.

1 Introduction

Statistical studies of shape frequently attempt to compare the shape of an organ selected from two different groups. They are used to form connections between shape and the presence or absence of disease [1, 2], testing hypotheses in the differences between men and women [3, 1, 4], as well as examining biological processes. The goals of such studies are to classify new examples of an organ into either group and to inform a doctor what makes a shape more like either group.

This paper examines a set of issues in classifying organs and presenting information to doctors. Is it more important to use a good representation or a particular classification method? Is the choice of alignment technique critical or not? We examine these issues in one application, a data set of thirty amygdala-hippocampus Complexes [5]. Fifteen of these case are from normal subjects and fifteen suffer from schizophrenia.

1.1 Classification Methods

Until very recently, most groups doing shaped-based classification have used linear classifiers to separate two groups. The technology has generally been motivated by the desire to create deformable models of shape as the basis of automatic segmenters [6]. One creates a deformable model of shape using Principal Component Analysis (PCA) on representations of example organs. A generative model can be made by allowing the shape to deform along the principle modes of variation. To compare two groups, one forms such a model for each group and separates the features determined by PCA for each group using a hyperplane. Visualization of what makes a shape more or less like each group can be examined by moving perpendicular to the hyperplane.

Golland et. al [2] introduced more complex, non-linear classification techniques. That work showed that one can take derivatives of the classifier to determine what makes an organ more like either group.

1.2 Representation

There are numerous attractive representations of shape for classification. Most representations implicitly determine the points on the surface. For example, surfaces can be parameterized in a series of spherical harmonics [1]. Medial representations [7] are parameterizations of shapes based on a chain or sheet of atoms that project a surface. Distance maps embed a surface by labeling the voxels of a 3D-volume with the distance from the surface. Each of these parameterized models avoid establishing direct correspondences between surfaces.

Other representations, conversely, use explicit representations of correspondences. One can represent the surfaces of organs by a triangular mesh where the vertices of the mesh are at corresponding points on the different organs. One can also represent shape using deformation fields that establish correspondences not only between surface points, but also between interior points of the organ.

Correspondence-based representations have the potential to yield more information than implicit representations. Deformation fields can show not only whether surfaces moved in or out, but can also show local rotation or compression of an organ. For this reason, we consider deformation fields in this paper.

Unfortunately, finding a correspondances between organs is challenging. For example, medical organs typically have large smooth surfaces. It is not intuitive where points in one smooth surface should lie on a second surface. One typically overcomes the challenge by matching two shapes while minimizing an additional constraint. For example, one can match organs by treating them as viscous fluids [8], though many have argued that this type of matching can allow unrealistic correspondences. Finding corresponding surfaces that form a minimum description length of a dataset is a promising idea, though it is difficult to find in three spatial dimensions [9]. There have also been various types of surface matching by matching points of one surface to the closest points in another [10, 11].

We choose to treat organs like pieces of rubber and align them using a linear elastic model. Intuitively, high curvature regions in one object should match high curvature regions in the second object. Matching by minimizing an elastic energy accomplishes this feat. It is lower energy for a sharp portion of one surface to match against a sharp portion of another surface, rather than to flatten and match against an less sharp adjacent region.

2 Methods

We explore several issues that are important to shabed-based classification. We start by forming two different representations of the shapes. We then compare the results of linear and non-linear classifiers. We explore whether normalizing the data for volume is harmful or useful. Finally, we consider the effects of different alignment methods. Every combination is examined, though we report only a subset of the results.

The data consists of segmented amygdala-hippocampus complexes, fifteen from schizophrenic patients and fifteen from normal patients [5]. The objects are represented using both signed distance maps and displacement fields.

Signed distance maps are formed [2] by labeling each voxel with its distance from the surface. The resulting representation is the vector of labels of the voxels.

To form the displacement field representations of the left complexes, one left amygdala-hippocampus complex is chosen randomly as a basis. It is meshed with tetrahedra to facilitate the matching process. The mesh is then treated as a linear elastic material and deformed to match the amygdala hippocampus complexes as described in [12]. A similar procedure is carried out for the right complexes starting with meshing a complex and then matching to the rest of the data. For each match, the displacement of the nodes of the tetrahedra form a roughly uniform sampling of the displacement field. The resulting representation vector is the concatenation of all the displacements of the nodes of the tetrahedra.

Except where stated, the initial segmentations were aligned using second order moments. When both sides are considered together, the vectors of each side were simply concatenated. Each section indicates whether data has been normalized to one volume, or not scaled at all.

Let $\boldsymbol{x}$ be the representation vector of the complexes for either representation. The squared distance between two amygdalas, $||\boldsymbol{x}-\boldsymbol{x'}||^2$, is defined to be $(\boldsymbol{x}-\boldsymbol{x'})^T(\boldsymbol{x}-\boldsymbol{x'})$. For displacement fields, this distance is simply the square of the length of the displacement field between each complex. For distance maps, there is no simple interpretation of distance between shapes.

We train linear and non-linear classifiers of the data. The non-linear classifier is a support vector machine (SVM), described in [2]. We used a Gaussian kernel function $K(\boldsymbol{x},\boldsymbol{x}_k) = -e^{||\boldsymbol{x}-\boldsymbol{x}_k||^2/\gamma}$ in the SVM where γ is proportional to the square of the width of the kernel. We pick γ to optimize the training accuracy of the classifier, tested by "leave one out" training.

Our goal is not only to form the classifier, but to determine what makes an amygdala more or less schizophrenic. Differentiating a pre-thresholded, SVM classifying function with respect to shape would seem to yield the answer. The derivative at $\boldsymbol{x}$ is $\sum_k \frac{2}{\gamma}\alpha_k y_k(\boldsymbol{x_k}-\boldsymbol{x})e^{-||\boldsymbol{x}-\boldsymbol{x}_k||/\gamma}$, where the $\{\alpha_k\}$ are constants determined by the SVM and $\{y_k\}$ are -1 for on group and 1 for the other. Using distance maps this answer is not sufficient. A small change to a distance map does not yield another distance map. Therefore, one must project a derivative back onto the manifold of distance maps [2]. However, displacement fields form a vector space; a small change in a displacement field yields another displacement field. Thus for this case, differentiating the classifier is sufficient for our goals.

3 Results

When forming correspondences between shapes, it is important to verify that correspondences between objects are meaningful. We formed displacement fields between complexes using a linear elastic model [12]. Figure 1 shows a number of points found to correspond. The hand segmented complexes have notably different structure. The typical member of the data set has a nearly horizontal “head” like the rightmost two complexes; a few have the head at an angle (leftmost), or practically no such structure at all (second from the left). Even with the shape differences, in all examples, Point 1 stays slightly above the tip of the head and Point 2 stays on the side of the head. Examining the base of the complexes, some bases are nearly flat and level while others are curved and angled. A review of points near the base also shows that points are approximately in the same position relative to major structures.

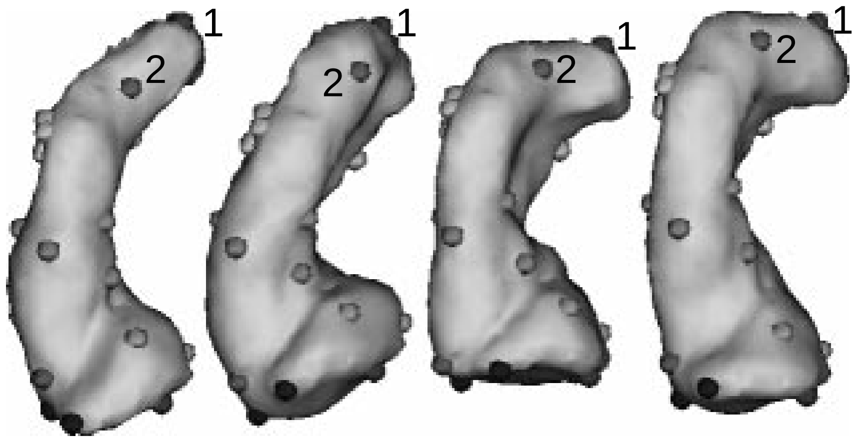

Fig. 1. Surfaces of four matched left complexes. To show correspondence across the shapes, points were randomly selected and represented by small spheres on each surface. Points 1 and 2 are referenced in the text.

3.1 Classifier Comparison

We compare the cross validation accuracy of linear and SVM-based [2] classifiers. Table 1 shows the classification accuracy using data normalized to remove relative volume. The table shows that support vector machines generally performed better than linear classifiers by 10 to 20 percentage points. We found this to be the case in all trials, with one exception. When deformation fields were aligned using absolute orientation [13], linear classifiers improved their classifying ability to as high as 70%. (See Section 3.3.)

For individual sides, deformation field-based classifiers perform slightly better than distance map-based classifiers. When considering both sides together, that performance improvement becomes much larger. For comparison to other methods, Table 1 shows results from Gerig et al. [1] who made a classifier comparing the two sides in each subject using the same data. That classifier's accuracy is in between the two we tested.

Normalized Structure	Cross Validation Accuracy using Linear Classifier		
	Deformation Field	Distance Map	
Left Complex	$60 \pm 18\%$	$57 \pm 18\%$	
Right Complex	$53 \pm 18\%$	$53 \pm 18\%$	
Both Complexes	$57 \pm 18\%$	$53 \pm 18\%$	

Normalized Structure	Cross Validation Accuracy using SVM		
	Deformation Field	Distance Map	Gerig et al.
Left Complex	$67 \pm 18\%$	$70 \pm 17\%$	
Right Complex	$73 \pm 17\%$	$77 \pm 17\%$	
Both Complexes	$80 \pm 16\%$	$67 \pm 18\%$	$73 \pm 17\%$

Table 1. A comparison of cross-validation accuracy for SVMs versus linear classifiers. The data is normalized to the same volume. The range is the 95% confidence interval.

Structure	Cross Validation Accuracy using SVM		
	Deformation Field	Distance Map	Gerig et al.
Left Complex	$77 \pm 17\%$	$73 \pm 17\%$	
Right Complex	$77 \pm 17\%$	$70 \pm 17\%$	
Both Complexes	$87 \pm 16\%$	$70 \pm 17\%$	$87 \pm 16\%$

Table 2. Cross-validation accuracy for the different representations using SVMs. For deformation fields and distance maps, the data is not normalized by volume. For the third column, volume was added separately to shape data. [1].

3.2 Including Volume

Table 2 examines the effects of not normalizing for volume in the data, using an SVM. Comparing Tables 1 and 2, volume normalization generates improvements for the displacement field based classifier of between 3 and 10 percentage points. For the classifier based on distance maps, volume improves or hurts classifier, but the effect is roughly 3 percentage points each way. Gerig et al. [1] include volume as a separate variable in their classifier; doing so improves the performance of their classifier to the same as the deformation field-based classifier.

3.3 Alignment

Variations in patient position during imaging causes misalignment between subjects. We align shapes so that classifiers are not confused by such rigid transformations. Unfortunately, there are often several reasonable registration methods for a given representation. One can align displacement fields, for example, by the moments of the voxels of the initial segmentations, the moments of the tetrahedral mesh, the moments of the nodes of the mesh, or by absolute orientation [13], removing global translations and rotations from the deformation field. One can align both amygdalas with the same rigid transformation, or do each separately. We test all of these alignment methods, repeating each test with and without volume normalization. Choosing different alignment methods causes classifiers to have a range of between 3 and 10 percentage points, typically closer to 3%. The worst results on any complex is shown in Table 1, 67% accuracy. The best results are 87% achieved in three different ways. Most results are between 70 and 80%. Aligning both amygdalas with the same transformation produced lower-accuracy classifiers than classifiers based on aligning each side separately. Examining visualization of the differences between the classes as in section 3.4, the different alignments had small impacts on the differences found between groups.

3.4 Visualization of Differences

The goal of this study is to visualize differences between the classes. Figure 2 shows those differences, found from derivatives of the classifier as described in Section 2. For each derivative examined, for both displacement field and distance map based classifiers, the two gradients shown are seen to be qualitatively similar. Comparing the left half of the grid to the right half, schizophrenic to normals, the visualizations are very close to opposites, that is the results suggest that making a schizophrenic more normal is the opposite of making a normal subject more schizophrenic; this result is not guaranteed in a non-linear classifier.

In the bottom of Figure 2, the derivatives of the classifier based on deformation fields are shown in the form of a vector field. (Because distance maps do not use correspondences, it is not possible to show motions tangential to the surface using distance maps.) The vector fields show that there is motion along the surface in several places. Most notably, there is a clear rotation of the "head" of the complex in the image. Conversely, most of the motion in the base is simply compression or expansion. There is also a rotation in the base of the left amygdala, though much smaller in magnitude than the rotation in the "head", and very difficult to see in the image.

4 Discussion

We examined which issues in shape classification have the largest impact on classification accuracy. It is clear from Table 1 that the non-linear SVM methods outperform linear classifiers by 10 to 20 percentage points. Volume information (Table 2) consistently improves results the deformation field-based classifier, by 4 to 10 percentage points, as well as the classifier of Gerig et al. [1]. Thus, for this case, including volume information was helpful, but not as helpful as using a non-linear classifier over a linear one. Alignment techniques generally had a smaller effect. Differences between alignment methods generally produced accuracy changes of a few percentage points. Though, in a couple case, different alignment methods produced a range of classification accuracies as large as 10 percentage points.

Displacement fields outperformed distance maps in this study. Classification rates were higher in almost every test performed. Interestingly, Tables 1 and 2 suggest that deformation field-based classifiers were able to find correlations between the deformations on different sides to improve the classification rate, while distance maps-based classifiers were not. Perhaps most importantly, displacement fields provided vector fields in visualizations which added an important tool for visualizing shape differences.

The issues that concern us are how well these conclusions generalize to other data sets. We believe that non-linear classification methods will almost always outperform linear classification methods. We also believe that the gains due to the inclusion of volume, or various alignment techniques exist, but will be much smaller. We feel that correspondence-based methods do provide an advantage over non-correspondence-based methods because they provide additional information for visualizing class differences.

5 Conclusion

We examined several issues that are important for performing shape comparison studies: complexity of the classifier, volume information, alignment method, and representation. For the shape differences between amygdala-hippocampus complexes, non-linear classifiers provide 10-20 percentage point accuracy gains over linear methods. For this study, not normalizing for volume provides a smaller gain, in the range of 4 to 10 percentage points. Using different alignment methods generally produces an even smaller impact on classification accuracy.

We have shown that for the cases examined, deformation field-based classifiers outperform distance maps as a measure of shape. Deformation fields form classifiers of higher accuracy and produce more information for the visualization of shape differences.

Acknowledgements: S.J. Timoner is supported by the Fannie and John Hertz Foundation and NSF ERC grant, J.H.U Agreement #8810274. W. Wells is supported by the same NSF grant and NIH grant 1P41RR13218. Dr. Shenton's data collection was supported by NIMH grants R01 50740 and K02 01110, and a Veterans Administration Merit Award.

References

1. G. Gerig, M. Styner, et al., "Shape versus size: Improved understanding of the morphology of brain structures," in *MICCAI*, (Utrecht), pp. 24–32, October 2001.
2. P. Golland, W. E. L. Grimson, M. E. Shenton, and R. Kikinis, "Deformation analysis for shape based classification," in *IPMI*, (Davis, CA), pp. 517–530, June 2001.
3. C. Davatzikos, et al., "A computerized approach for morphological analysis of the corpus callosum," *Journal Computer Assisted Tomography*, 20(1):88–97, 1996.
4. A. M. C. Machado and J. C. Gee, "Atlas warping for brain morphometry," in *SPIE Medical Imaging, Image Processing*, pp. 642–651, 1998.
5. M. E. Shenton et al., "Abnormalities in the left temporal lobe and thought disorder in schizophrenia: A quantitative magnetic resonance imaging study," *New England Journal of Medicine*, 327:604–612, 1992.
6. T. F. Cootes et al., "The use of active shape models for locating structures in medical images," *Image and Vision Computing*, 12(6):9–18, 1992.
7. S. Pizer et al., "Segmentation, registration and measurement of shape variation via image object shape," *IEEE Trans. on Medical Imaging*, 18(10):851–865, 1996.
8. G. Christensen, R. Rabbit, and M. Miller, "Deformable templates using large deformation kinematics," *Transactions on Image Processing*, 5(10):1435–1447, 1996.
9. R. H. Davies, T. F. Cootes, and C. J. Taylor, "A minimum description length approach to statistical shape modeling," in *IPMI*, (Davis), pp. 50–63, June 2001.
10. L. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-d and 3-d images," *PAMI*, 15:1131–1147, November 1993.
11. L. Staib and J. Duncan, "Boundary finding with parametrically deformable models," *PAMI*, 14:1061–1075, November 1992.
12. S. J. Timoner, W. E. L. Grimson, R. Kikinis, and W. M. Wells, "Fast linear elastic matching without landmarks," in *MICCAI*, (Utretcht), pp. 1358–60, October 2001.
13. B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society of America A*, vol. 4, pp. 629–642, 1987.

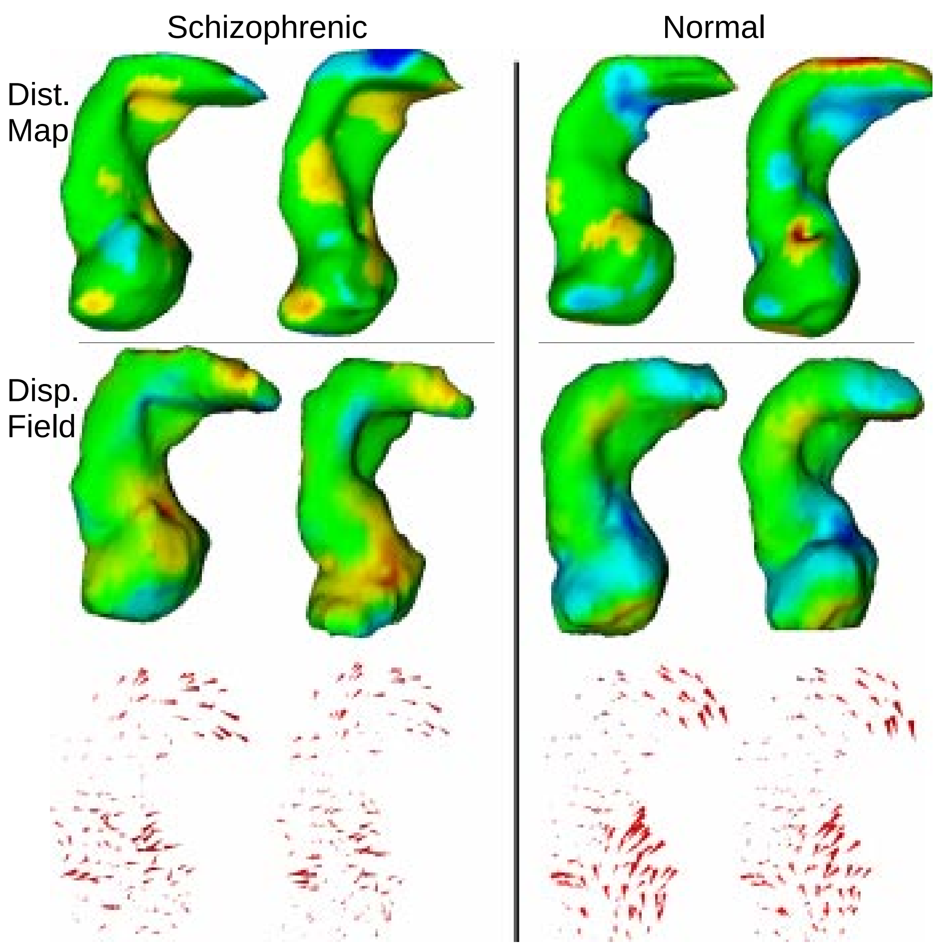

Fig. 2. The top four plots show the deformation of the surfaces relative to the surface normal for the left amygdala-hippocampus complex. For Schizophrenic subjects, "deformation" indicates changes to make the complexes more normal. For Normal subjects, "deformation" indicates changes to make the complexes more diseased. The 2x2 grid of surfaces shows deformations of Schizophrenic/Normal subjects using Distance Maps/Displacement Fields as representations. In each entry in the grid, the two largest deformations evaluated at the support vectors of the SVM classifier are shown; the larger one is on the left. The color coding is used to indicate the direction and magnitude of the deformation, changing from blue (inward) to green (no motion) to red (outward). The bottom two plots are the deformations fields used to generate the plots directly above them. Note that motion along the surface does not affect the colors in the surfaces.