

Deformation Analysis for Shape Based Classification

Polina Golland¹, W. Eric L. Grimson¹, Martha E. Shenton², and Ron Kikinis³

¹ Artificial Intelligence Laboratory,
Massachusetts Institute of Technology, Cambridge, MA.

{polina,welg}@ai.mit.edu

² Laboratory of Neuroscience, Clinical Neuroscience Division, Department of
Psychiatry, VAMC-Brockton, Harvard Medical School, Brockton, MA.
mshenton@warren.med.harvard.edu

³ Surgical Planning Laboratory, Brigham and Women's Hospital,
Harvard Medical School, Boston, MA.

kikinis@bwh.harvard.edu

Abstract. Statistical analysis of anatomical shape differences between two different populations can be reduced to a classification problem, i.e., learning a classifier function for assigning new examples to one of the two groups while making as few mistakes as possible. In this framework, feature vectors representing the shape of the organ are extracted from the input images and are passed to the learning algorithm. The resulting classifier then has to be interpreted in terms of shape differences between the two groups back in the image domain. We propose and demonstrate a general approach for such interpretation using deformations of outline meshes to represent shape differences. Given a classifier function in the feature space, we derive a deformation that corresponds to the differences between the two classes while ignoring shape variability within each class. The algorithm essentially estimates the gradient of the classification function with respect to node displacements in the outline mesh and constructs the deformation of the mesh that corresponds to moving along the gradient vector. The advantages of the presented algorithm include its generality (we derive it for a wide class of non-linear classifiers) as well as its flexibility in the choice of shape features used for classification. It provides a link from the classifier in the feature space back to the natural representation of the original shapes as surface meshes. We demonstrate the algorithm on artificial examples, as well as a real data set of the hippocampus-amygdala complex in schizophrenia patients and normal controls.

1 Introduction

Statistical studies of anatomical shape in different populations are important in understanding anatomical effects of diseases (comparing patients *vs.* normal controls) or biological processes (for example, comparing different age groups). Volume and area measurements were originally used for such studies, but recently more sophisticated shape based techniques have been used to identify

statistical differences in the shape of a particular organ in different subject groups [1,6,7,8,9,11,13]. One of the important steps in the analysis is the interpretation of the mathematical description of shape variation in terms of image related quantities. It allows us to argue about the shape differences in biologically meaningful terms of organ development and deformation.

Most work in constructing models of shape and its variation has been motivated by and used for segmentation tasks, where a generative model of shape variation guides the search over the space of possible deformations of the “representative” shape template, with the goal of matching it to a new input image. Typically, this approach uses Principal Component Analysis (PCA) to build a linear model of variation and compress it to a few principal modes of deformation [5,10,12,18]. In this case, the shape variation within the population is represented using these principal modes of deformation. In our previous work, we extended this approach to linear discriminant analysis by visualizing the shapes resulting from moving the input feature vectors along the vector normal to the best separating hyperplane [8]. In this paper, we present a generalized approach to deriving the deformation that represents the differences between two classes, where that deformation is deduced from a non-linear classifier learned from training examples. Note that unlike pattern recognition, where the learned classifier is used to label new examples, we are much more interested in the classifier’s spatial structure, from which we infer the shape differences in terms of organ deformations.

The paper is organized as follows. In the next section, we discuss the problem of statistical shape analysis and briefly outline our approach to learning the shape differences that is described in detail elsewhere [9]. It is followed by the derivation of the deformation that corresponds to differences in the shape represented by the classifier function. The algorithm is then demonstrated on artificial examples, as well as a real data set of the hippocampus-amygdala complex in schizophrenia patients compared to a normal control group.

2 Background on Shape Based Classification

Shape Representation. The first step in shape based statistical analysis is extraction of the shape parameters, or a shape descriptor. The wealth of shape descriptors used in medical image analysis and computer vision include parametric models: such as Fourier descriptors [17,18] or spherical harmonic functions [2,10], as well as a variety of non-parametric models: landmark based descriptors [1,5], distance transforms [9,12], medial axes [8,14] and deformation fields obtained by warping a template to each of the input shapes [6,7,11,13].

We choose to use distance transforms as shape descriptors. A distance transform, or a distance map, is a function that for any point in the image is equal to the distance from the point to the closest point on an outline. By first segmenting example images into structures of interest, we can then convert each example into a distance map, based on the boundary of the segmented object. We then sample points in the distance transform to form a feature vector, which captures

the shape of the segmented object. Since the values of the distance transform at neighboring voxels are highly correlated, using it as a shape descriptor provides the learning algorithm with redundant information on the structure of the feature space. Moreover, the local nature of the distance transform's sensitivity to noise in the object's pose and segmentation makes it an attractive shape representation for statistical analysis [9,12]. The shape information extraction is an inherently noisy process, consisting of imaging, segmentation and feature extraction, with every step introducing errors. The vectors in the training set are therefore computed with uncertainty, and it is important that small errors in any of the steps do not cause significant displacements of the resulting feature vectors.

Thus, our feature extraction procedure consists of the following steps:

- Example images are segmented into relevant structures;
- These images are aligned by computing a distance map, then computing the transformation that aligns the moments of the maps;
- The aligned images are clipped to a common size;
- The aligned and clipped 3D distance transforms are used as feature vectors, with individual voxels being the vector components.

2.1 Statistical Analysis

Once the feature vectors are extracted from the images, they are used to study the statistical properties of two groups and to compare two populations, for example, a population of normals and a population of schizophrenics. As mentioned in the introduction, statistical techniques have been extensively used to build generative models of shape variation within a single population by computing the principal modes of deformation [5,10,12,14,18]. This approach has been previously used to identify and visualize the differences between two populations by comparing generative models derived separately for each population, for example by studying the first few principal modes of deformation in the two groups [7,13]. In addition, several techniques have been demonstrated that are based on learning the differences between the classes directly from the training examples by either comparing individual landmarks [1,7] or by actually training a classifier to label new examples [6,11]¹. Our technique [8,9] belongs to this last group and is based on the premise that in order to automatically detect statistical shape differences between two populations, one should try to build the best possible classifier for labeling new examples. In the optimal case, the classifier function will exactly represent the important differences between the two classes, while ignoring the intra-class variability.

We use the Support Vector Machines (SVM) learning algorithm to arrive at a classification function. In addition to the sound theoretical foundations of

¹ In both papers, a classifier was learned from the training data. However, it was not used to demonstrate the shape differences, but only to establish statistical significance.

SVMs, they have been demonstrated empirically to be quite robust and seemingly free of the over-fitting problems to which other learning algorithms, such as neural networks, are subject. Below, we state without proof the main facts on SVMs that are necessary for derivation of our results. The reader is referred to the original publications on SVMs [3,20] for more details on the method.

The Support Vector Machines algorithm exploits kernel functions for non-linear classification. A kernel is a function $K : R^n \times R^n \mapsto R$, such that for some mapping function $\Psi : R^n \mapsto R^m$ from the original space to a higher dimensional space ($n \leq m$), the value of the dot product in R^m can be computed by applying K to vectors in R^n :

$$K(\mathbf{u}, \mathbf{v}) = \langle \Psi(\mathbf{u}) \cdot \Psi(\mathbf{v}) \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in R^n. \tag{1}$$

Thus, while mapping the points from the original space R^n to the higher dimensional space R^m and performing computations explicitly in R^m can be prohibitive because of the dimensionality of the space, we can compute certain quantities in R^m without ever computing the mapping if the answer depends only on dot products of the vectors in R^m .

Given a training set of l pairs $\{(\mathbf{x}_k, y_k)\}_1^l$, where $\mathbf{x}_k \in R^n$ are observations and $y_k \in \{-1, 1\}$ are corresponding labels, the SVM learning algorithm searches for the optimal classification function

$$f_K(\mathbf{x}) = \sum_{k=1}^l \alpha_k y_k \langle \Psi(\mathbf{x}) \cdot \Psi(\mathbf{x}_k) \rangle + b = \sum_{k=1}^l \alpha_k y_k K(\mathbf{x}, \mathbf{x}_k) + b \tag{2}$$

that given a new example \mathbf{x} assigns it a label $\hat{y}(\mathbf{x}) = \text{sign}(f(\mathbf{x}))$. In the higher dimensional space defined by the mapping Ψ , the separation boundary is a hyperplane whose normal is a linear combination of $\Psi(\mathbf{x}_k)$'s

$$\mathbf{w} = \sum_{k=1}^l \alpha_k y_k \Psi(\mathbf{x}_k), \tag{3}$$

but it can be an arbitrary complex surface in the original space. Training vectors with non-zero α 's are called support vectors, as they define, or "support", the hyperplane. The optimal classifier (i.e., the optimal set of coefficients α_k 's and the threshold b) is found by maximizing the margin between the classes with respect to the separating boundary, which can be reduced to a constrained quadratic optimization problem defined in terms of the pairwise dot products of $\Psi(\mathbf{x}_k)$'s, and can therefore be solved without explicitly computing the mapping into the higher dimensional space.

For a linear kernel $K(\mathbf{u}, \mathbf{v}) = \langle \mathbf{u} \cdot \mathbf{v} \rangle$, the mapping Ψ is the identity, and the classifier is a linear function in the original space. Several non-linear kernels have been also proposed. In this work, we use a family of Gaussian kernel functions

$$K(\mathbf{u}, \mathbf{v}) = e^{-\|\mathbf{u}-\mathbf{v}\|^2/\gamma},$$

where the parameter γ determines the width of the kernel. One of the important properties of this family of classifiers is its locality: moving a support vector

slightly affects the separation boundary close to the vector, but does not change it in regions distant from the vector. Following the discussion in the previous section, this is a desirable property in the presence of noise in the training examples.

Interpretation of the Analysis in the Image Domain. Once the mathematical description of the shape variation is obtained, it needs to be mapped back into the image domain for visualization and anatomical interpretation. In the generative case, this is traditionally done by visualizing the first few principal modes of deformation derived from PCA [5,10,12,18]. Since the model is linear, this step is particularly simple and intuitive. This technique can also be used in linear discrimination by noting that for a linear classifier, the deformation that represents the differences between the two classes is fully determined by the normal to the separating hyperplane and can be used for visualization in the image domain [8]. In fact, this is a special case of the general solution presented in the next section.

Recently, a non-linear kernel based version of PCA was proposed in [19], following the approach of kernel based SVMs. It was utilized in [15] for constructing a non-linear model of shape variation of face contours for a wide range of the viewing angles. In this application, the generative linear model in the higher dimensional space had to be mapped back into the image domain for sampling new examples of face contours. The analysis presented in the next section derives an analogous result for the discriminative case.

3 From a Classifier to Shape Differences

In this section, we first analyze the classification function in the original feature space and derive the direction that corresponds to the maximal change in the classifier output while minimizing displacement within the class. Then we note that the shape descriptors used for classification are not necessarily the optimal representation for modeling shape differences in the image domain and show how to express the “optimal direction” in terms of deformations of the outline mesh of the segmented object, which is a more appropriate representation.

3.1 Shape Differences in the Feature Space

Eq. (2) and Eq. (3) imply that the classification function $f(\mathbf{x})$ is directly proportional to the signed distance from the input point to the separating boundary computed in the higher dimensional space defined by the mapping Ψ . It therefore depends only on the projection of the vector $\Psi(\mathbf{x})$ onto the normal to the hyperplane \mathbf{w} , while ignoring the component of $\Psi(\mathbf{x})$ that is perpendicular to \mathbf{w} . This suggests that in order to create a displacement of $\Psi(\mathbf{x})$ that corresponds to the difference between the two classes, one should change the vector’s projection onto \mathbf{w} while keeping its perpendicular component constant. Below, we derive the displacement of \mathbf{x} in the original space that achieves the desired displacement of $\Psi(\mathbf{x})$.

As we move from \mathbf{x} to $\mathbf{x} + d\mathbf{x}$ in the original space, the corresponding vector in the range of Ψ moves from $\mathbf{z} = \Psi(\mathbf{x})$ to $\mathbf{z} + d\mathbf{z} = \Psi(\mathbf{x} + d\mathbf{x})$. The deviation of the displacement vector $d\mathbf{z}$ from the direction of the classification vector \mathbf{w}

$$\|e\|^2 = \left\| d\mathbf{z} - \frac{\langle d\mathbf{z} \cdot \mathbf{w} \rangle}{\|\mathbf{w}\|} \right\|^2 = \|d\mathbf{z}\|^2 - \frac{\langle d\mathbf{z} \cdot \mathbf{w} \rangle^2}{\|\mathbf{w}\|^2} \tag{4}$$

can be expressed as a function of dot products in the range of Ψ as shown here:

$$\|\mathbf{w}\|^2 = \left\| \sum_k \alpha_k y_k \Psi(\mathbf{x}_k) \right\|^2 \tag{5}$$

$$= \sum_{k,m} \alpha_k \alpha_m y_k y_m \langle \Psi(\mathbf{x}_k) \cdot \Psi(\mathbf{x}_m) \rangle \tag{6}$$

$$= \sum_{k,m} \alpha_k \alpha_m y_k y_m K(\mathbf{x}_k, \mathbf{x}_m), \tag{7}$$

$$\langle d\mathbf{z} \cdot \mathbf{w} \rangle = \sum_k \alpha_k y_k \langle (\Psi(\mathbf{x} + d\mathbf{x}) - \Psi(\mathbf{x})) \cdot \Psi(\mathbf{x}_k) \rangle \tag{8}$$

$$= \sum_k \alpha_k y_k (K(\mathbf{x} + d\mathbf{x}, \mathbf{x}_k) - K(\mathbf{x}, \mathbf{x}_k)) \tag{9}$$

$$= \sum_k \alpha_k y_k \sum_i \left. \frac{\partial K(\mathbf{u}, \mathbf{v})}{\partial u_i} \right|_{(\mathbf{u}=\mathbf{x}, \mathbf{v}=\mathbf{x}_k)} dx_i \tag{10}$$

$$= \nabla f(\mathbf{x}) d\mathbf{x}, \tag{11}$$

where u_i denotes the i -th component of vector \mathbf{u} , and

$$\|d\mathbf{z}\|^2 = \|\Psi(\mathbf{x} + d\mathbf{x}) - \Psi(\mathbf{x})\|^2 \tag{12}$$

$$= K(\mathbf{x} + d\mathbf{x}, \mathbf{x} + d\mathbf{x}) + K(\mathbf{x}, \mathbf{x}) - 2K(\mathbf{x} + d\mathbf{x}, \mathbf{x}) \tag{13}$$

$$= \sum_{i,j} \left[\left. \frac{1}{2} \frac{\partial^2 K(\mathbf{u}, \mathbf{u})}{\partial u_i \partial u_j} \right|_{\mathbf{u}=\mathbf{x}} - \left. \frac{\partial^2 K(\mathbf{u}, \mathbf{v})}{\partial u_i \partial u_j} \right|_{(\mathbf{u}=\mathbf{x}, \mathbf{v}=\mathbf{x})} \right] dx_i dx_j \tag{14}$$

$$= \sum_{i,j} \left. \frac{\partial^2 K(\mathbf{u}, \mathbf{v})}{\partial u_i \partial v_j} \right|_{(\mathbf{u}=\mathbf{x}, \mathbf{v}=\mathbf{x})} dx_i dx_j \tag{15}$$

$$= \sum_{i,j} H_{\mathbf{x}}(i, j) dx_i dx_j = d\mathbf{x}^T H_{\mathbf{x}} d\mathbf{x}, \tag{16}$$

where matrix $H_{\mathbf{x}}$ is the upper right quarter of the Hessian of the kernel function K evaluated at (\mathbf{x}, \mathbf{x}) .

Substituting into Eq. (4), we obtain

$$\|e\|^2 = d\mathbf{x}^T \left(H_{\mathbf{x}} - \|\mathbf{w}\|^{-2} \nabla f^T(\mathbf{x}) \nabla f(\mathbf{x}) \right) d\mathbf{x}. \tag{17}$$

This is a quadratic form that defines the deviation of the displacement $d\mathbf{z}$ from \mathbf{w} as a function of the displacement $d\mathbf{x}$ in the original feature space. The error is minimized when $d\mathbf{x}$ is aligned with the smallest eigenvector of matrix

$$Q_{\mathbf{x}} = H_{\mathbf{x}} - \|\mathbf{w}\|^{-2} \nabla f^T(\mathbf{x}) \nabla f(\mathbf{x}). \quad (18)$$

We call this direction *principal deformation*, in analogy to the principal modes of variation in the generative case. We can compute $Q_{\mathbf{x}}$ in the original space using Eq. (7), Eq. (10), and Eq. (15) to compute $\|\mathbf{w}\|$, $\nabla f(\mathbf{x})$ and $H_{\mathbf{x}}$ respectively. Note that since the mapping Ψ is non-linear, the principal deformation is generally not spatially uniform and has to be determined for every input feature vector \mathbf{x} separately.

Interestingly, an analytical solution of this problem exists for a wide family of kernel functions, including the linear kernel ($K(\mathbf{u}, \mathbf{v}) = \langle \mathbf{u} \cdot \mathbf{v} \rangle$), as well as all kernel functions that depend only on the distance between the two input vectors ($K(\mathbf{u}, \mathbf{v}) = k(\|\mathbf{u} - \mathbf{v}\|)$). For these functions, $H_{\mathbf{x}} = cI$, where c is a constant and I is an identity matrix, and the principal deformation vector is parallel to the largest eigenvector of the matrix $\|\mathbf{w}\|^{-2} \nabla f^T(\mathbf{x}) \nabla f(\mathbf{x})$, which is the gradient of the classification function $\nabla f^T(\mathbf{x})$. It is a well known fact that in order to achieve the fastest change in $f(\mathbf{x})$, one moves along the gradient of the function. In our case, the gradient of the function also corresponds to the smallest distortion of the feature vector along irrelevant dimensions. For the linear kernel, $H_{\mathbf{x}}$ is the identity matrix ($c = 1$), and the principal deformation is equal to the classification vector (the distortion is zero). For the Gaussian kernel, $c = \frac{2}{\gamma}$, and while the principal deformation is parallel to the function gradient, it does not achieve the classification direction precisely. One example of commonly used kernel functions for which $H_{\mathbf{x}}$ is not diagonal, and therefore the general eigenvalue problem has to be solved, is the polynomial kernel family:

$$K(\mathbf{u}, \mathbf{v}) = (1 + \langle \mathbf{u} \cdot \mathbf{v} \rangle)^d, \quad (19)$$

where d is the degree of the polynomial.

3.2 From Shape Differences to Deformation

If the feature vectors extracted from the input images can be used successfully by a generative model of shape variation, the result derived above is sufficient: in order to visualize shape differences between the classes, one should estimate the principal deformation and then generate examples in the image domain that demonstrate the changes that the shape undergoes as its feature vector moves along the principal deformation direction. However, many shape representations that are well suited for classification do not lend themselves naturally to this approach. The distance transform is one such representations. Distance transforms do not form a linear vector space – a linear combination of two distance transforms is not always a valid distance transform – and therefore cannot be used directly for generating new examples of the shape. Some representations do not provide an easy way to model deformations locally, which could be a

desired property. In this section, we demonstrate how one can use two different representations for classification and visualization and still obtain the optimal reconstruction.

We choose to use surface meshes extracted from the segmented images to model shape deformation. By selectively changing locations of the mesh nodes, this representation can be used very effectively for modeling local deformations of the shape. Furthermore, if we form a feature vector by concatenating node positions in the image, we can change the values of the vector components arbitrarily and still represent a particular shape in the image domain. However, this representation is unsuitable for classification purposes, mainly because the feature vector length could be different for different example shapes, and moreover, it is difficult to establish correspondences between nodes in different meshes. A finite set of points, often manually selected landmarks, can be used as a shape descriptor [1,5], but we are interested in a dense representation that can be extracted fully automatically from the segmented images.

Once the classifier function has been learned in the original feature space (distance transforms in our example), we can perform the search for the principal deformation direction as described in the previous section. But instead, we transfer the search into the space of node displacements in the original surface mesh by computing a local Jacobian of the distance transform components with respect to the displacements of the mesh nodes. We restrict the node motion to the direction of the normal to the outline at each node, combining the displacements into a new feature vector \mathbf{s} . For the distance transform descriptors, each voxel is only affected by the closest node in the outline mesh². We can express infinitesimal changes in the distance transform feature vectors $d\mathbf{x}$ as a function of the displacements of the points on the outline of the shape $d\mathbf{s}$:

$$J_{\mathbf{x}}(i, j) = \frac{\partial x_i}{\partial s_j} = \begin{cases} 1, & \text{node } j \text{ is the closest node to voxel } i \\ 0, & \text{otherwise} \end{cases}, \quad (20)$$

or, in matrix notation, $d\mathbf{x} = J_{\mathbf{x}}d\mathbf{s}$. Substituting into Eq. (18), we can see that the principal deformation in the space of node displacements \mathbf{s} is the smallest eigenvector of the matrix

$$Q_{\mathbf{s}} = J_{\mathbf{x}}^T (H - \|\mathbf{w}\|^{-2} \nabla f^T(\mathbf{x}) \nabla f(\mathbf{x})) J_{\mathbf{x}}. \quad (21)$$

Note that $J_{\mathbf{x}}$ is very sparse, and in fact, only one element in every row is non-zero (or a few, for the skeleton voxels), and that the dimensionality of the new space is significantly lower than the dimensionality of the distance transform feature vectors, which simplifies the computation.

² For most voxels, there is only one such node, but for voxels that belong to a skeleton of a shape there are at least two closest nodes. However, the number of skeleton voxels is typically small (5-8% in our experiments) and they do not affect the results significantly. We ignore the skeleton voxels in the derivation, but our experiments demonstrated that the resulting deformations do not change almost at all if we incorporate the skeleton voxels into the analysis.

In the case of linear or Gaussian kernels, the principal deformation in the space of the mesh deformations can be computed explicitly:

$$\mathbf{s}' = J_{\mathbf{x}}^T \nabla f^T(\mathbf{x}), \quad (22)$$

or equivalently, for node i , the optimal displacement is

$$s'_j = \sum_i \frac{\partial f(\mathbf{x})}{\partial x_i} \frac{\partial x_i}{\partial s_j} = \sum_{\substack{i, \text{ such that node } j \text{ is the} \\ \text{closest node to voxel } i}} \frac{\partial f(\mathbf{x})}{\partial x_i}. \quad (23)$$

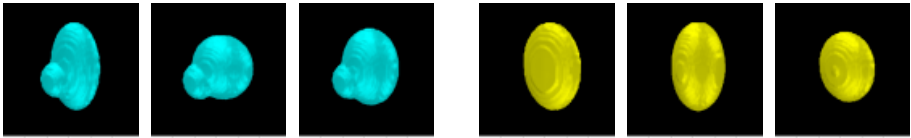
To summarize, we use deformations of the original shapes to demonstrate shape differences learned from example images. The shape descriptors used for classification are linked to the node displacements in the outline mesh through a linear model, which allows searching for the principal deformation in the space that is particularly well suited for deformation representation. This approach can be applied to a variety of shape descriptors, as long as one can establish local dependence of the descriptor components on the displacements of the mesh nodes.

4 Experimental Results

In this section, we demonstrate the method on two data sets. The first one is an artificially constructed example, for which the shape differences are known in advance. The second example is a real medical data set.

In the artificial example, we generated 30 examples of ellipsoidal shapes of varying sizes. The width, height and thickness of the shapes were sampled out of a ± 10 voxel variation range centered at 20, 30 and 40 voxels respectively. The 10 first examples were chosen to be the first class and had a spherical bump added to them. Fig. 1a shows example shapes from each class. We extracted the distance transforms (Fig. 1b), performed the classification and constructed the optimal deformation both in the distance transform space (Fig. 1c) and in the space of the mesh node displacements (Fig. 2). We can see that the classifier successfully identified the area of the bump as the area of the highest differences between the two groups. Fig. 3 shows several frames in the deformation process for two different input shapes.

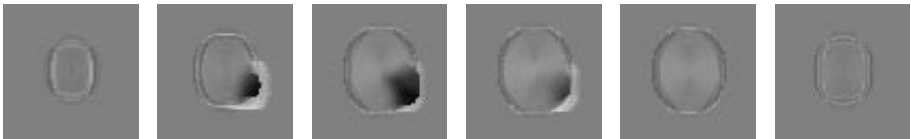
We also report the results of the method applied to a data set that contains MRI scans of 15 schizophrenia patients and 15 matched controls. In each scan, the hippocampus-amygdala complex was manually segmented. More details on the subject selection and data acquisition can be found in [16]. The same paper reports statistically significant differences in the left hippocampus based on relative volume measurements (the volume of the structure normalized by the total volume of intracranial cavity). We previously showed statistically significant differences between the groups based on SVM classification using distance transforms as feature vectors [9]. Here, we present the deformations that correspond to the learned differences. Fig. 4 and Fig. 5 show two examples of the



(a) Three example shapes from each class.



(b) Several slices of the distance transform for the top left shape.



(c) Several slices of the principal deformation for the top left shape. The values are linearly scaled so that dark regions correspond to negative values, bright areas correspond to positive values.

Fig. 1. Artificial data set. Example shapes are shown for each class, as well as the distance transform and the principal deformation for the top left shape.

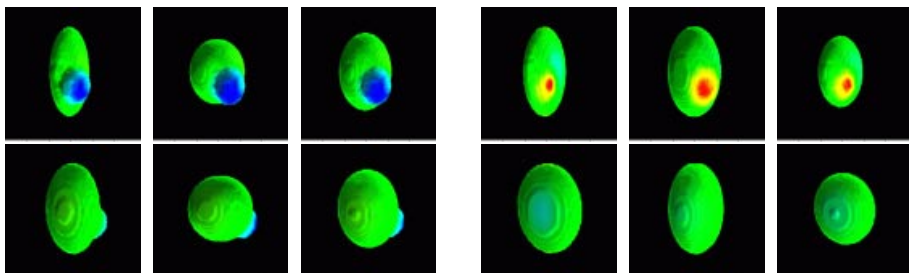
most prominent deformations identified for normal controls and schizophrenia patients respectively. These examples were among those selected as support vectors (a total of seven support vectors was selected, other vectors represent similar deformations to shown here) by the algorithm and each represents a different region in the feature space and the corresponding deformation. We can see that the deformation is well localized, and that the hippocampus in the schizophrenia group seems to be thinner and shorter than the normal one. While this result and its physiological implications need further validation, it demonstrates our approach. We are currently working on validating the technique on several other data sets.

5 Conclusions

In this paper, we have demonstrated a general technique for deriving the deformation representing statistical differences between two classes of anatomical shapes. We identify the differences by training a classifier in the space of feature vectors (distance transforms). Then the classifier is analyzed to arrive at the



(a) Principal deformation in the mesh space.



(b) Two more views of the principal deformation.

Fig. 2. Example shapes from Fig 1a with the principal deformation values painted on. Colors are used to indicate the direction and the magnitude of the deformation, varying from blue (moving inward) to green (zero deformation) to red (moving outwards).

surface based representation of the deformation. We use kernel based classification technique (SVMs), but this approach can be extended to other classification methods that express the classification function analytically in terms of the training data. In order to derive the deformation, we compute the derivative of our representation with respect to small displacements of the points on the outline and use it to express the changes in the classification function in terms of the deformations of the outline.

Surface based representations are excellent for expressing deformations, which could be easy and intuitive to interpret in anatomical terms. However, these representations are usually difficult to use as feature vectors, mostly because it is difficult to reliably establish correspondences between the points on the meshes that represent example shapes. Our approach overcomes this problem by performing classification using volume based feature vectors and then mapping them back to the surface based representation for every individual shape example.

The proposed technique can provide the currently missing step of detailed interpretation and visualization of a mathematical model of the statistical differences in the anatomical shape detected from example images.

Acknowledgments. Quadratic optimization was performed using PR_LOQO optimizer written by Alex Smola.

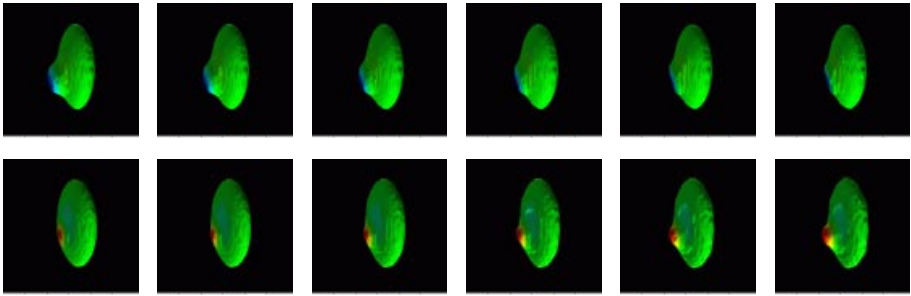
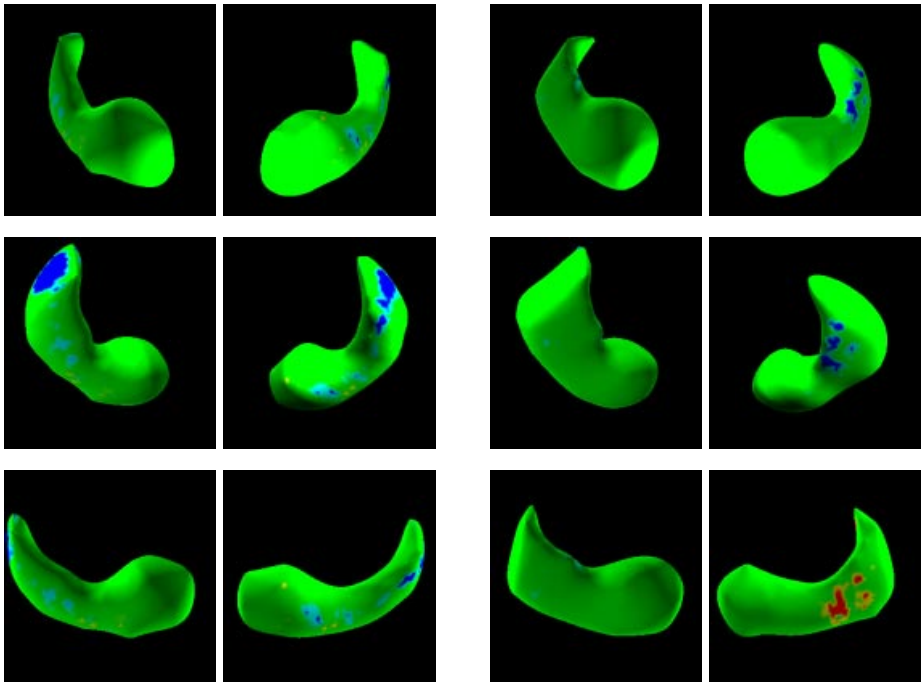


Fig. 3. Several frames in the deformation process for the first example shape in each group in Fig. 1a. Note how the shape of the protrusion changes as we vary the amount of deformation.

This research was supported in part by NSF IIS 9610249 grant. M. E. Shenton was supported by NIMH K02, MH 01110 and R01 MH 50747 grants, R. Kikinis was supported by NIH PO1 CA67165, R01RR11747, P41RR13218 and NSF ERC 9731748 grants.

References

1. F. L. Bookstein. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Medical Image Analysis*, 1(3):225-243, 1996.
2. C. Brechbühler, G. Gerig and O. Kübler. Parameterization of closed surfaces for 3-D shape description. *CVGIP: Image Understanding*, 61:154-170, 1995.
3. C. J. C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(2):121-167, 1998.
4. C. J. C. Burges. Geometry and Invariance in Kernel Based Methods. In *Advances in Kernel Methods: Support Vector Learning*, Eds. B. Schölkopf, C. J. C. Burges, A. J. Smola, MIT Press, 89-116, 1998.
5. T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Graham. Training Models of Shape from Sets of Examples. In *Proc. British Machine Vision Conference*, 9-18, Springer-Verlag, 1992.
6. J. G. Csernansky *et al.* Hippocampal morphometry in schizophrenia by high dimensional brain mapping. In *Proc. Nat. Acad. of Science*, 95(19):11406-11411, 1998.
7. C. Davatzikos, *et al.* A Computerized Method for Morphological Analysis of the Corpus Callosum. *J. Computer Assisted Tomography*, 20:88-97, 1996.
8. P. Golland, W. E. L. Grimson and R. Kikinis. Statistical Shape Analysis Using Fixed Topology Skeletons: Corpus Callosum Study. In *Proc. IPMI'99*, LNCS 1613:382-387, 1999.
9. P. Golland, W. E. L. Grimson, M. E. Shenton and R. Kikinis. Small Sample Size Learning for Shape Analysis of Anatomical Structures. In *Proc. MICCAI'2000*, LNCS 1935:72-82, 2000.
10. A. Kelemen, G. Székely, and G. Gerig. Three-dimensional Model-Based Segmentation. In *Proc. IEEE Intl. Workshop on Model Based 3D Image Analysis*, Bombay, India, 87-96, 1998.

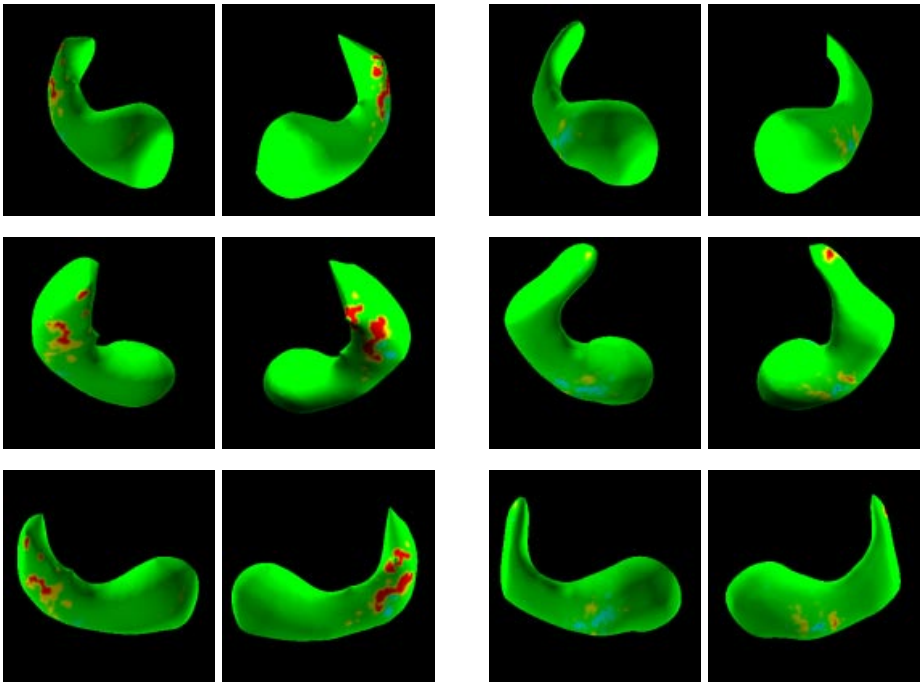


(a) Normal control #1.

(b) Normal control #2.

Fig. 4. Two examples (support vectors) from the normal control group with the principal deformation painted on the surface. In each row, the right and the left hippocampus are shown from a different viewpoint (anterior, posterior and lateral respectively). The color-coding scheme is identical to Fig. 2.

11. J. Martin, A. Pentland, and R. Kikinis. Shape Analysis of Brain Structures Using Physical and Experimental Models. *In Proc. CVPR'94*, 752-755, 1994.
12. M. E. Leventon, W. E. L. Grimson and O. Faugeras. Statistical Shape Influence in Geodesic Active Contours. *In Proc. CVPR'2000*, 316-323, 2000.
13. A. M. C. Machado and J. C. Gee. Atlas warping for brain morphometry. *In Proc. SPIE Medical Imaging 1998: Image Processing*, SPIE 3338:642-651, 1998.
14. S. M. Pizer, *et al.* Segmentation, Registration, and Measurement of Shape Variation via Image Object Shape. *IEEE Transactions on Medical Imaging*, 18(10): 851-865, 1996.
15. S. Romdhani, S. Gong and A. Psarrou. A Multi-View Nonlinear Active Shape Model Using Kernel PCA. *In Proc. BMVC'99*, 483-492, 1999.
16. M. E. Shenton, *et al.* Abnormalities in the left temporal lobe and thought disorder in schizophrenia: A quantitative magnetic resonance imaging study. *New England J. Medicine*, 327:604-612, 1992.
17. L. Staib and J. Duncan. Boundary finding with parametrically deformable models. *IEEE PAMI*, 14(11):1061-1075, 1992.



(a) Schizophrenia patient #1.

(b) Schizophrenia patient #2.

Fig. 5. Two examples (support vectors) from the schizophrenia group with the principal deformation painted on the surface. In each row, the right and the left hippocampus are shown from a different viewpoint (anterior, posterior and lateral respectively). The color-coding scheme is identical to Fig. 2.

18. G. Székely *et al.* Segmentation of 2D and 3D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models. *Medical Image Analysis*, 1(1):19-34, 1996.
19. B. Schölkopf, A. Smola, and K.-R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10:1299-1319, 1998.
20. V. N. Vapnik. *Statistical Learning Theory*. John Wiley & Sons, 1998.