MEDIAL AXIS BASED STATISTICAL SHAPE MODEL (MASSM): APPLICATIONS TO 3D PROSTATE SEGMENTATION ON MRI

Rob Toth, Rachel Sparks, Anant Madabhushi

Rutgers University
Department of Biomedical Engineering
Piscataway, New Jersey, 08854

ABSTRACT

In this paper, we present a novel methodology for computing statistical shape models (SSM's) by leveraging the medial axis model to determine shape variations between objects. Landmark based SSM's (LSSM's) are a popular approach to describing valid shape variation in an object of interest by applying principal component analysis to a set of landmarks on the surface of the object. However, defining landmarks which capture important shape variations can be difficult. Additionally, establishing landmark correspondences across different shapes is a challenging problem. In this work we utilize the medial axis to define the shape of the object, thereby enabling superior characterization of the underlying shape variations compared to the landmark based approach. Locations on the medial axis (medial atoms) are utilized to generate a SSM, one that we refer to as a medial axis based SSM (MASSM). The aim of the MASSM is to capture variations in the local symmetry of an object across different studies. We show analytically that reconstructing a shape using medial atoms yields a lower average error compared to reconstructing a shape using triangulations of landmarks on the boundary of a 2D object. We experimentally validate the ability of the MASSM to better reconstruct a 3D prostate volume, and to better segment that prostate compared to an LSSM on 34 3D T2-weighted endorectal Magnetic Resonance images. The accuracy of the LSSM in reconstructing the prostate was highly dependent on the number of landmarks N, while the accuracy of the MASSM was quite robust to the number of medial atoms N. In addition, in a segmentation test, the results showed an average Dice overlap of 0.93 for the MASSM while the LSSM showed an average Dice overlap of 0.88.

Index Terms— Prostate Capsule, Active Shape Model, Medial Axis, Segmentation, Statistical Shape Modeling

1. INTRODUCTION

Traditional statistical shape models (SSM's) use a set of anatomical landmarks to describe an object's shape [1]. The landmark based SSM (LSSM) is constructed by placing a set of landmarks on the surface of the object of interest, and principal component analysis (PCA) is used to capture the variations in the Cartesian coordinates of these landmarks [1]. However, there are several problems with using LSSM's.

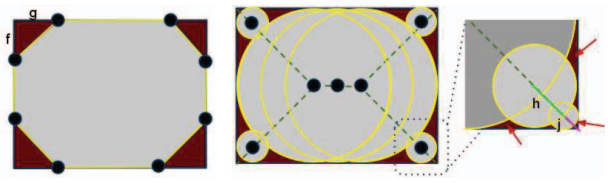
1. *Need for multiple landmarks to accurately capture shape variations.* Especially in 3D, the number of landmarks N to capture all the variation on the surface of the object is extremely high, leading to computational challenges in terms of performing PCA and in triangulating the object's surface.
2. *Need for accurate landmark alignment.* Each of the N landmarks must represent the same anatomical location across all training images, which can be infeasible, especially in 3D [2].
3. *Need for landmark triangulation for volume reconstruction.* In 2D, the landmarks (assuming they are ordered) can define a polygon. In 3D, however, triangulation of the point cloud must be done to reconstruct the volume. In both of these cases, there will inherently be some error in the surface reconstruction [3], especially near areas of high curvature.

In this paper, we aim to use points ("atoms") along the medial axis [4], to define a shape model. The medial axis is defined as all locations inside an object which are equidistant 2 or more surface points, and can be thought of as defining the local symetry of an object. Medial axis shape models (MASM's) are a popular approach for shape modeling in biomedical imagery due to their flexibility in describing object morphology. MASM typically compare differences in corresponding atoms of the MASM to quantify differences in shape between objects [5, 6]. Additionally, the inverse medial axis transform is able to reconstruct objects given a MASM, and in this paper we show analytically (in 2D) that reconstructing a shape using medial atoms yields a lower error than reconstructing a shape using a set of landmarks. In this paper, we introduce the medial axis statistical shape model (MASSM) which combines the MASM within the SSM framework to accurately describe shape variations. The closest related system [7] uses a MASM to segment medical imagery using a Bayesian approach, but this paper aims to use a MASM to characterize the underlying shape variation for a SSM.

We tested our MASSM for segmenting Magnetic Resonance (MR) images of the prostate. Prostate MR segmentation is a necessary pre-requisite in (a) developing and validating computer aided diagnosis systems for detecting prostate cancer *in vivo*, (b) for prostate volume estimation, and (c) for guiding surgical treatments. In this work we evaluate our MASSM on a set of 34 T2-weighted, 3.0 Tesla prostate MR images, and compare it to the traditional LSSM in terms of reconstruction and segmentation accuracy.

Our paper is presented as follows. In Section 2, we provide an analytical justification for using the medial atoms to reconstruct an object. In Section 3, we present the mathematical framework for defining and calculating the MASSM. We describe and present results for several experiments with T2-weighted endorectal prostate MR images in Section 4. Finally, we offer concluding remarks and future directions in Section 5.

This work was made possible via grants from the Wallace H. Coulter Foundation, New Jersey Commission on Cancer Research, National Cancer Institute (R01CA136535-01, R01CA140772 01, R21CA127186 01, and R03CA143991-01), The Cancer Institute of New Jersey, and Bioimagene Inc. The authors would also like to thank Drs. Nicholas Bloch, Elizabeth Genega, Neil Rofsky, Robert Lenkinski, and Mark Rosen for imagery and annotations.



(a) (b)

Fig. 1: R is modeled by (a) LSSM L (landmarks shown as black nodes) and (b) MASM S (medial atoms shown as black nodes and radii displayed in yellow). The inset in (b) shows a magnified view of an area near p_s. The reconstruction errors for both the LSSM and MASM are shown in red, and it can be seen that the MASM (S) has less reconstruction error than the landmark method (L).

2. ANALYTICAL DESCRIPTION OF MEDIAL AXIS RECONSTRUCTION ERROR

Figure 1 shows the reconstruction error (red) of a MASM and LSSM in the context of a 2D rectangle R, which we will describe below.

Definition 1. For a d-dimensional object $\mathcal{O}$ ($d \geq 2$), a LSSM L is defined by N landmarks $p_n : n \in \{1, \ldots, N\}, p_n \in \mathbb{R}^d = [x_n, y_n, \ldots]$ on $Q(\mathcal{O})$, where $Q(\mathcal{O})$ represents the surface of $\mathcal{O}$.
Definition 2. We define the reconstruction $\mathcal{O}_L$ of object $\mathcal{O}$ from L as a linear interpolation (e.g. triangulation if $d = 3$) between $\forall p_n \in L$.
Definition 3. For any object $\mathcal{O}$, a MASM S is defined by N medial atoms $c_n : n \in \{1, \ldots, N\}, c_n \in \mathbb{R}^{d+1} = [x_n, y_n, \ldots, r_n]$ on the medial axis [4, 5] of R, where r_n represents the radius of atom n and c_n is equidistance to at least 2 points on $Q(\mathcal{O})$.
Definition 4. The reconstruction $\mathcal{O}_S$ of object $\mathcal{O}$ from S, is defined as all locations within r_n of each atom $c_n \in S$, where,

$$\mathcal{O}_S = \{d | \exists n : \| d - c_n \|_2 \leq r_n, c_n \in S\}. \quad (1)$$

Definition 5. Landmarks $p_1 \in \mathbb{R}^d$ and $p_2 \in \mathbb{R}^d$, are adjacent if $\nexists p_n : (\| p_n - p_1 \|_2 < \| p_1 - p_2 \|_2$ and $\| p_n - p_2 \|_2 < \| p_1 - p_2 \|_2)$.
Definition 6. A saliency point p_s on $Q(\mathcal{O})$ is defined as a point of high curvature (e.g. a corner of a rectangle).
Definition 7. Landmark p_μ is unique if $\nexists p_n : \| p_\mu - p_n \|_2 < \epsilon$.

Proposition 1. *For rectangle R (where $d = 2$) with length l, width w, and equally spaced landmarks L, (i) the distance f of the landmark closest to p_s is on the interval $0 \leq f \leq 2(l + w)/N$, and (ii) the difference in area between R_L and R, near p_s, is given by,*

$$e_1 = \tfrac{1}{2} f \cdot (f - 2(l + w)/N). \quad (2)$$

Proof. To prove (i) assume that the two closest points to p_s are p_1 and p_2, which are adjacent, and have distances f and g, respectively, from p_s. Since all landmarks are equally spaced on the surface of R, $f + g = 2(l + w)/N$. If $p_1 = p_s$ then $f = 0$ and $g = 2(l + w)/N$. If $p_2 = p_s$ then $g = 0$ and $f = 2(l + w)/N$. Hence, $0 \leq f \leq 2(l + w)/N$.

To prove (ii), $\forall p_n$ are connected by straight lines (from Definition 2). In R, corner p_s makes a right angle, and hence f and g define the lengths of sides of a right triangle (shown as a red triangle in Figure 1(a)). Since $f + g = 2(l + w)/N$, Equation 2 follows as the area of the triangle. □

For a given number of landmarks N on $Q(R)$, Proposition 1 shows that the reconstruction error is entirely dependent on distance f from the saliency point to the closest landmark.

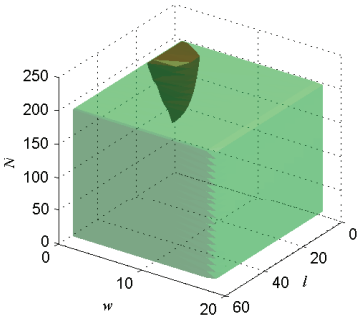


Figure 2: The solid brown region represents all values of l, w, and N for which $\overline{e_1} < \overline{e_2}$, while the translucent green region represents $\overline{e_2} < \overline{e_1}$. In a majority of cases, medial atoms had a lower reconstruction error ($\overline{e_2}$) than landmarks ($\overline{e_1}$).

Proposition 2. *For rectangle R (where $d = 2$) with length l, width w, and medial atoms S equally spaced on the medial axis, (i) the distance j of c_n closest to p_s is on the interval $0 \leq j \leq \left(l + \frac{(\sqrt{2}-2)}{2} \cdot w\right)/N$, and (ii) the difference in area between R_S and R near p_s is given by,*

$$e_2 = r_M^2(1 - \tfrac{\pi}{4}) + \sum_{m=1}^{M-1} r_m^2(1 - \tfrac{\pi}{4}) - r_{m+1}^2(1 - \tfrac{5\pi}{4}) + \cdot I_{m+1}, \quad (3)$$

where M is the number of atoms on the axis segment near p_s.

Proof. To prove (i) for R there are 5 segments of S, shown as 5 green dotted lines in Figure 1(b). The first segment is parallel to the edge of R and has a length $l - w$. The other 4 segments of S correspond to each p_s. S bisects each corner at p_s to form a 45° angle with R and has a length of $\sqrt{2}/2 \cdot w$. The total length of S is $d_S = \left(l + (\sqrt{2} - 2)/2 \cdot w\right)$. As there are N equally spaced atoms, the distance h between adjacent atoms is $h = d_S/N$. Assume the atom closest to p_s is c_1, where $\| p_s - c_1 \|_2 = j$. If $j = h + \epsilon, 0 < \epsilon < h$, then there can exist an atom c_2, where $\| p_s - c_2 \|_2 = \epsilon$, which would imply that $\| c_1 - c_2 \|_2 = h$, and hence c_2 would be closer to p_s. Therefore, h is a upper bound on j.

To prove (ii), the reconstruction error e_2 is constrained to regions near p_s. The number of atoms M along the segment of S near p_s is defined as $M = \lfloor \sqrt{2}/2 \cdot n \cdot w/(2\sqrt{2} \cdot w + l - w) \rfloor$ where $\lfloor \cdot \rfloor$ indicates the floor operator. The radius r_m of each atom m is the distance to the closest edge of R, and is given as $r_m = \sqrt{2}/2 \cdot (j + (M - m) \cdot h)$. The error e_2 is now defined as the difference between the area near p_s and the area of overlapping c_m given by Equation 3 where I_{m+1} is the area of intersection between atoms c_m and c_{m+1}. □

For a given number of medial atoms N, Proposition 2 shows that the reconstruction error is entirely dependent on distance j from the saliency point p_s to the closest atom c_1.

Figure 2 shows the results for $1 < l < 60, 1 < w < 20, 10 < N < 200$. $\overline{e_1}$ is the average e_1 over f, and $\overline{e_2}$ is the average e_2 over j. The green translucent region represents values in which $\overline{e_2} < \overline{e_1}$ and the brown solid region represents $\overline{e_1} < \overline{e_2}$. The green region is approximately 28 times larger than the brown region, which suggests that unless a large number of p_n are chosen relative to the size of the shape, the medial axis will yield a more accurate reconstruction.

Proposition 3. *Given $\mathcal{O}$ (where $d = 2$) with perimeter length u, k saliency points, and N unique p_n randomly distributed on $Q(\mathcal{O})$, the probability P that $\forall s \in \{1, \ldots, k\}$, $\exists p_n : \| p_s - p_n \|_2 < \epsilon$ is given by the hypergeometric distribution,*

$$P = \frac{\binom{k}{k}\binom{u/\epsilon - k}{N - k}}{\binom{u/\epsilon}{N}}, \quad (4)$$

where u/ϵ defines the number of unique locations on $Q(\mathcal{O})$.

Proposition 3 shows that to have a high probability of capturing all k saliency points, the number of landmarks N must approach u/ϵ. For example, reasonable values of u and ϵ for an MR prostate slice are 250 mm and 0.25 mm respectively. Therefore, $u/\epsilon = 1000$, and the hypergeometric distribution yields an extremely low probability of hitting all k saliency points (regardless of the value of k) unless the number of landmarks N is very close to 1000.

3. MEDIAL AXIS STATISTICAL SHAPE MODEL (MASSM)

3.1. Constructing the MASSM

Step 1. N medial atoms for training image $i \in \{1, \ldots, I\}$ are equally sampled on the medial axis, and defined as S_i.
Step 2. All $S_i, i \in \{2, \ldots, I\}$ are aligned to S_1 using an orthogonal Procrustes technique.
Step 3. A matrix $\mathbf{M} \in \mathbb{R}^{4N \times I}$ is defined, where each column represents the N medial atoms of a single image $i \in \{1, \ldots, I\}$.
Step 4. The first $e = 10$ Eigenvectors and Eigenvalues of the covariance of $\mathbf{M}$ are calculated, defined as $\Phi \in \mathbb{R}^{4N \times e}$ and $\lambda \in \mathbb{R}^e$ respectively. We therefore define a MASSM as,

$$S_{\mathbf{b}} = \overline{S} + \Phi \cdot \left(\mathbf{b} * \sqrt{\lambda}\right) \tag{5}$$

where $\overline{S} \in \mathbb{R}^{4N}$ represents the mean set of medial atoms, $\cdot$ represents matrix multiplication, $*$ represents element-wise multiplication, $\mathbf{b} \in \mathbb{R}^e$ represents a set of e parameters defining the shape, and $S_{\mathbf{b}} \in \mathbb{R}^{4N}$ represents the medial atoms given the parameters $\mathbf{b}$. The reconstructed volume $V_{\mathbf{b}}$ is then defined using Equation 1.

3.2. Using MASSM to Segment an Image

Step 1. A classifier is used to determine which voxels belong within the object, given as W. For example, Cosio [8] used a Bayesian classifier comprised of a sum of Gaussians to calculate W.
Step 2. Parameter $\mathbf{b}$ is used to calculate $S_{\mathbf{b}}$ and subsequently $V_{\mathbf{b}}$ according to Equations 5 and 1 respectively. Initially, $\mathbf{b} = 0$.
Step 3. The number of true positive and true negative voxels T between $V_{\mathbf{b}}$ and W are calculated as $T = |V_{\mathbf{b}} \cap W| + |W| - |V_{\mathbf{b}} \cup W|$.
Step 4. Modify $\mathbf{b}$ and repeat Step 2 until the number of voxels in Step 3 is maximized, so that $\mathbf{b}^* = \text{argmax}_{\mathbf{b}}\, T$.
The final segmentation is given as $V_{\mathbf{b}^*}$, and the calculation of $\mathbf{b}^*$ was done using a simulated annealing technique [9]. In this paper, we assume that *Step 1* has already been completed.

4. EXPERIMENTAL RESULTS AND DISCUSSION

4.1. Data and Experimental Design

The goals of the experiments were (1) to show that the MASSM can better reconstruct the prostate volume than the LSSM, and (2) to show that the MASSM can better segment the prostate compared to the LSSM, given a W. Given a set of landmarks L from the LSSM, the reconstructed volume V_L is obtained via triangulation of L. We tested our MASSM on 34 T2-weighted, 3.0 Tesla, 3D endorectal *in vivo* prostate MR images.

4.2. MASSM Based Shape Variations of the Prostate

Figure 3 shows qualitatively an example of $V_{\mathbf{b}}$. Changing the first principal component ($\mathbf{b}_1$) primarily changes the effect of the endorectal coil, where $+3$ standard deviations represents the endorectal coil pushing against the prostate, while -3 standard deviations represents minimal effect of the endorectal coil.

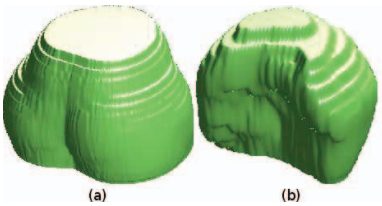


Figure 3: $V_{\mathbf{b}}$ is shown for the prostate MASSM. $\mathbf{b}_1 = -3$ is shown in (a) and $\mathbf{b}_1 = +3$ is shown in (b).

4.3. Experiment 1: Reconstruction Accuracy

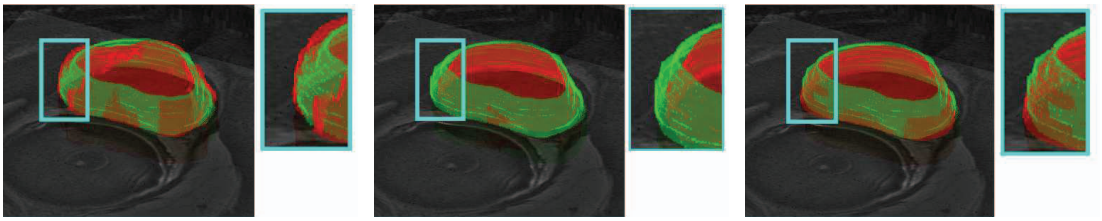
N medial atoms were extracted from each study, and the original volume was reconstructed. In addition, N landmarks were extracted from each study and the original volume was triangulated. The accuracy of these reconstructions were given in terms of both Dice [10] overlap, and Mean Absolute Distance (MAD).

The results of the reconstruction accuracy for both MASSM and LSSM are shown in Figures 5(a) and 5(b) for different values of N. The mean values over $I = 34$ studies are shown as the points on the graph, the standard deviations are shown as vertical bars. The results demonstrate that if there are a sufficient number of landmarks to capture variations on the surface of the shape (approximately $N \geq 500$), then the triangulated landmarks are able to accurately reconstruct the volume. The LSSM's Dice values were slightly higher than the MASSM's (indicating higher accuracy), but the MAD values were also slightly higher (indicating lower accuracy). However, once N begins to decrease below a certain threshold, it can be seen that the MASSM is still able to accurately reconstruct the shape, while triangulating the landmarks yields a significantly decreased reconstruction accuracy, in terms of both Dice and MAD. This suggests that the MASSM is more robust to variations in N than the LSSM.

4.4. Experiment 2: Segmentation Accuracy

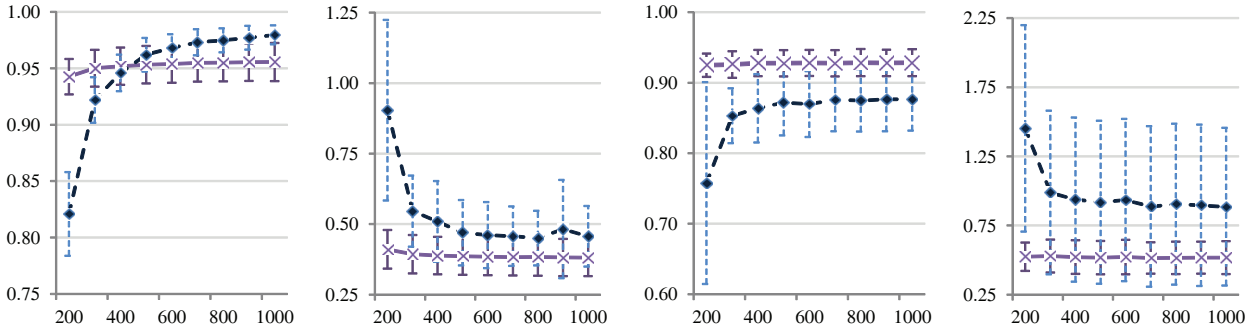
Experiment 2 evaluated the MASSM in a segmentation scheme, and determined how well the MASSM would constrain the segmentation to only valid shapes, given a classification result W. To obtain W for this experiment, the surface of the ground truth segmentation on each image was perturbed by random distances both inwards and outwards. An example of W is shown in Figure 4(a) in red. We tested whether the MASSM would constrain the segmentation to only learned shapes better than the LSSM. The accuracy of the segmentation result with respect to the ground truth was determined. For the MASSM the accuracy of $V_{\mathbf{b}^*}$ was determined. The LSSM was also fit to W in the manner described in Section 3.2, yielding a reconstructed prostate volume V_L, for which the accuracy was determined.

The results in Figure 5 show that the segmentation accuracies decrease as N decreases. However, consistent with Experiment 1, the LSSM accuracy decreases much more rapidly than the MASSM accuracy. In addition, even with a large N ($N = 1000$), the accuracy of the LSSM (Dice = 0.88, MAD = 0.89 mm) was still less than the accuracy of the MASSM (Dice = 0.93, MAD = 0.42 mm), and the LSSM had a higher standard deviations. For the reconstruction, the LSSM had a higher Dice value compared to the MASSM with $N = 1000$, but a lower Dice in the segmentation which suggests that the MASSM is better able to capture the underlying shape variations and is able to constrain the appearance model to valid shapes.



(a) Perturbed Mask (W) (b) LSSM (V_L) (c) MASSM ($V_{\mathbf{b}^*}$)

Fig. 4: Qualitative results are shown with the ground truth in green. (a) shows W in red, (b) shows V_L in red, and (c) shows $V_{\mathbf{b}^*}$ in red. The inset of each image highlights a region in which the MASSM was better able to constrain the segmentation to a valid shape.



(a) Reconstruction Dice (b) Reconstruction MAD (c) Segmentation Dice (d) Segmentation MAD

Fig. 5: The reconstruction ((a) and (b)) and segmentation ((c) and (d)) results are shown above in terms of Dice ((a) and (c)) and MAD in mm ((b) and (d)). The solid purple line represents the medial atoms and the dotted blue line represents the landmarks. The error bars represent the standard deviations over 34 studies and the X-axis represents N, where $200 \leq N \leq 1000$.

5. CONCLUDING REMARKS

In this paper, we presented a statistical shape model which leverages the medial axis to capture shape variations. Analytically, we demonstrated in 2D that a MASSM will more accurately reconstruct the shape of an object compared to a landmark based SSM (LSSM). We demonstrated experimentally that the MASSM can accurately reconstruct and segment the prostate on 34 T2-w MRI studies. When reconstructing the prostate volume, with a limited number of landmarks ($N = 200$), the LSSM achieved a mean MAD value of 0.90 mm. With $N = 200$ atoms, the MASSM was able to achieve a mean MAD of 0.41 mm. Only for large values of N ($N \geq 500$) was the LSSM able to achieve a high reconstruction accuracy (mean MAD less than 0.50 mm). However, even when using a large number of landmarks ($N = 1000$), the MASSM still outperformed the LSSM in a segmentation task. The MASSM was able to achieve a mean MAD of 0.52 mm for $N = 1000$, while the LSSM only achieved a mean MAD of 0.89 mm for $N = 1000$. This demonstrates the ability of the MASSM to better model the shape variation of the prostate than the LSSM, even when a large number of landmarks were used. Future work will entail testing on a larger cohort of data, and exploring accurate appearance models to combine with the MASSM.

6. REFERENCES

[1] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham, "Active shape models - their training and application," *CVIU*, vol. 61, no. 1, pp. 38–59, Jan 1995.

[2] M. Styner, K. Rajamani, L.P. Nolte, G. Zsemlye, G. Székely, C. Taylor, and R. Davies, "Evaluation of 3D correspondence methods for model building," in *IPMI 2732*, 2003, pp. 63–75.

[3] D. D. Morris and T. Kanade, "Image-consistent surface triangulation," *CVPR*, vol. 1, pp. 1332, 2000.

[4] H. Blum, "A transformation for extracting new descriptors of shape," in *Models for Perception of Speech and Visual Form*. 1967, pp. 362–380, Weiant Wathen-Dun.

[5] P.T. Fletcher, C. Lu, S.M. Pizer, and S. Joshi, "Principal geodesic analysis for the study of nonlinear statistics of shape," *IEEE TMI*, vol. 2, no. 8, pp. 995–1005, 2000.

[6] R.E. Sparks, R. Toth, J. Chappelow, G. Xiao, and A Madabhushi, "An integrated framework for analyzing three-dimensional shape differences: Evaluating prostate morphometry," in *IEEE ISBI*, 2010, pp. 1081–1084.

[7] S. Joshi, S. Pizer, P.T. Fletcher, P. Yushkevich, A. Thall, and J.S. Marron, "Multiscale deformable model segmentation and statistical shape analysis using medial descriptions," *IEEE TMI*, vol. 21, pp. 538–550, 2002.

[8] F.A. Cosio, "Automatic initialization of an active shape model of the prostate," *MedIA*, vol. 12, no. 4, pp. 469–483, Aug 2008.

[9] R.M. Lewis and V. Torczon, "Pattern search algorithms for bound constrained minimization," *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 1082–1099, 1999.

[10] L.R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 263, pp. 297–302, 1945.