

Geometrically Proper Models in Statistical Training

Qiong Han, Derek Merck, Josh Levy, Christina Villarruel,
James N. Damon, Edward L. Chaney and Stephen M. Pizer

Medical Image Display & Analysis Group,
University of North Carolina, Chapel Hill, North Carolina 27599, USA.
<http://midag.cs.unc.edu>

Abstract. In deformable model segmentation, the geometric training process plays a crucial role in providing shape statistical priors and appearance statistics that are used as likelihoods. Also, the geometric training process plays a crucial role in providing shape probability distributions in methods finding significant differences between classes. The quality of the training seriously affects the final results of segmentation or of significant difference finding between classes. However, the lack of shape priors in the training stage itself makes it difficult to enforce shape legality, i.e., making the model free of local self-intersection or creases. Shape legality not only yields proper shape statistics but also increases the consistency of parameterization of the object volume and thus proper appearance statistics. In this paper we propose a method incorporating explicit legality constraints in training process. The method is mathematically sound and has proved in practice to lead to shape probability distributions over only proper objects and most importantly to better segmentation results.

1 Introduction

Image segmentation, the main target problem of this paper, is an important task upon which many other medical image processing applications are based. The goal is to delineate image regions corresponding to certain anatomical structures from the background. Deformable model based methods tackle the segmentation problem by representing anatomical objects with geometric models and deform them into images via object shape and object-relative image intensity information. Shape and intensity statistics have been used respectively as priors and likelihoods in segmentations and have become a standard component in deformable model methods. In order to get trained statistics, we need to extract deformable models from a set of training images and calculate shape statistics and related appearance statistics from the trained models. This process of extracting models and calculating prior and likelihood statistics is called training.

Given a set of segmented images, a typical training step includes converting those segmentations into deformable models in a specific representation. The conversion is often realized by fitting a deformable template into a batch of binary characteristic images typically produced manually from the training greyscale images. Therefore the training step is also a fitting process but without any shape prior statistics. The lack of priors in training can cause illegal shapes, with local self-intersection or creases, which

will yield unrealistic shape statistics and then defected appearance statistics relative to objects coordinates. Even in an iterative bootstrapping process [4] the results from the first pass training affects the final ones after multiple passes. Existing solutions to this challenge require humans to directly enforce smoothness on extracted deformable models [3]. This approach is time consuming and defeats the purpose of having an automatic training step given the binary images. Furthermore, human interferences might bring undesirable biases in the trained shapes and thus in the trained statistics.

In our framework, the deformable model is medial based. It has been argued that medial based models have strong descriptive power because they capture not only local positional information but also local orientational and size information [11]. Medial methods represent and parameterize both object surfaces and interior volumes. A particular strength of this representation is that powerful mathematics exists on the relation of the medial geometry and the geometry of the implied boundary and interior of objects [5]. In this paper we use this mathematics to provide a set of geometric constraints that can be explicitly applied to our deformable models in order to maintain the legality of the surfaces and interiors of the trained shapes.

The proposition this paper is designed to establish is that attention to model legality in fitting the models used in forming proper statistics on object geometry and on object-relative image intensity patterns is decisive for the quality of applications of the probability densities derived from the training process. Recall that achieving such legality by automatic means in Point Distribution Models (PDMs) of object boundaries required mathematics and serious computation [1]. The mathematics now available on medial representations lowers the computation necessary to assure legality from that required for PDMs.

The rest of the paper is organized as the follows. Sec. 2 covers background on the deformable method using a medial representation and medial mathematics. Sec. 3 details the proposed method for achieving medial legality and proper statistics in training. Sec. 4 shows how we generate synthetic test data and the results from application of our method to both synthetic and real world data. Sec. 5 concludes the paper.

2 Background

We begin with a brief review on deformable model based methods and medial representations. We then address some mathematical background necessary for the detailed description of the proposed method.

2.1 Medially Represented Deformable Model

Deformable models are probabilistic shape descriptions. Under the Gaussian model, the distribution of the training data is modeled by several modes of deformation about a point in the shape space. This distribution describes all shapes in the training data and moreover, for a sufficiently large training set, estimates the full ambient shape space from which the training data are drawn. This statistical framework has been shown [12] to provide a powerful basis for studies finding shape differences between classes of shape or, together with probabilistic models on image appearance via object-relative

intensity patterns, to provide a powerful means of segmentation by posterior optimization.

The classical model representation is landmarks or its extension to surfaces, the point distribution model (PDM) with shape variance described by principal component analysis (PCA) of the feature space [3]. PDMs assume feature correspondence by fixed sampling, or attempt to induce correspondence post hoc by minimizing variability in the parameterization [7]. While it is possible to provide local legality constraints with such boundary representations, the literature shows little attention to this issue, perhaps because legality checks require extracting information about surface normals. Another surface representation uses spherical harmonic (SPHARM) basis functions, which guarantees local shape legality at serious computational expense but handles surface locality with difficulty. Legal PDMs can, however, be derived from the SPHARM fits to training images [13].

Our segmentation and shape analysis methodology uses the multi-scale discrete m-rep representation proposed in [11]. Medial representations provide a model-centric volumetric coordinate system for the object interior and hence, a framework for volumetric correspondence. As described in Sec. 2.2, the mathematics of m-reps based on [5] provides the means for legality guarantees over the whole object interior.

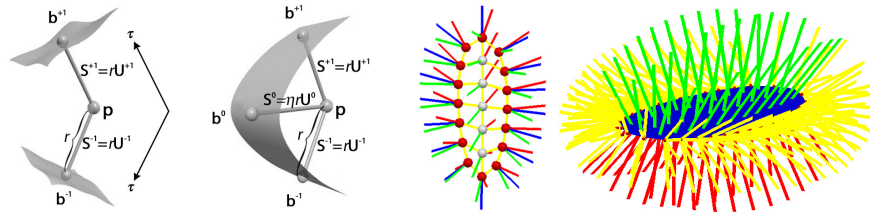


Fig. 1. From left to right: an internal atom of two spokes S^{+1} and S^{-1} , with τ to parameterize the object interior along the spokes; an end atom with an extra bisector spoke S^0 ; a discrete m-rep as a mesh of internal atoms (with white hubs) and end atoms (with red hubs); an interpolated spoke field on a smooth medial sheet (in dark blue).

An m-rep figure is a mesh of samples of an object’s medial axis, as shown in Fig. 1. Each internal medial sample, or medial atom $\mathbf{m} = (\mathbf{p}, r, \mathbf{U}^{+1}, \mathbf{U}^{-1}) \in \mathbf{M} = \mathbb{R}^3 \times \mathbb{R}^+ \times \mathcal{S}^2 \times \mathcal{S}^2$ has eight parameters: 3 in hub position \mathbf{p} , 4 in two medial spoke directions $\mathbf{U}^{+1/-1}$, and 1 in scale r .

In order to have more stability in the representation, each end atom can be considered as the compound of an internal atom $\{\mathbf{p}, r, \mathbf{U}^{+1}, \mathbf{U}^{-1}\}$ plus the bisector spoke corresponding to the crest line on the object boundary. The hub position and two regular spokes $S^{+1/-1}$ of each end atom are treated as one internal atom. An interpolation method can be applied to an m-rep to generate a continuous medial representation as a smooth double-sided spoke field [10], and the interpolated hub positions forms its own smooth surface called *medial sheet*, shown in dark blue in Fig. 1-rightmost.

Given that each atom lies on a Riemannian manifold [9], the distance between two atoms, $d(\mathbf{m}, \mathbf{o})$, is derived by a local flattening *Log*-map defined to map points on the

manifold to the local tangent plane. This leads to a Riemannian distance $d(\mathbf{m}, \mathbf{o})$. The squared distance between two m-rep models \mathbf{M} and \mathbf{O} with samples \mathbf{m}_i and \mathbf{o}_i respectively is defined as the sum of the squared distances between corresponding samples $d^2(\mathbf{M}, \mathbf{O}) = \sum_i d^2(\mathbf{m}_i, \mathbf{o}_i)$. These metrics allow for the extension of PCA into non-Euclidean domains such as m-reps [9].

The medial atom is the basic building block in m-rep representation, and the derivatives of the medial sheet and the two spoke directions in each medial atom contain sufficient information to tell whether any illegal shape exists or is about to appear [5]. The next subsection reviews the mathematics that enable us to put legality constraints on m-reps, detailed in Sec. 3.4.

2.2 Radial Shape Operator \mathbf{S}_{rad} and $r\mathbf{S}_{rad}$

Differential geometry reveals how a surface normal curves in its local surface tangent plane. The derivatives of surface normals can be summarized by a 2×2 symmetric matrix called the *shape operator*. Analogously, the *radial shape operator* [5] tells how a medial spoke changes while walking on the medial sheet. Since each m-rep atom has a spoke on both sides of the medial sheet, the spoke field on the medial sheet should be considered as double valued. Thus there are two radial shape operators defined for each atom: one for each side of the medial sheet.

Let us only consider one side of the medial sheet since the same method applies to the other. Assume that there is a continuous spoke field $\mathbf{S}(\mathbf{u})$ with unit length spoke direction $\mathbf{U}(\mathbf{u})$ and spoke length $r(\mathbf{u})$ on the continuous medial sheet $\mathbf{p}(\mathbf{u})$, where $\mathbf{u} = (v_1, v_2)$ parameterizes the 2D medial sheet and the spoke field. $\mathbf{S}(\mathbf{u}) = r(\mathbf{u}) \cdot \mathbf{U}(\mathbf{u})$. The derivatives of the unit length spoke direction $\mathbf{U}(\mathbf{u})$ wrt (v_1, v_2) are calculated as follows, with \mathbf{U} and \mathbf{p}_{v_1/v_2} being 1×3 row vectors.

$$\frac{\partial \mathbf{U}}{\partial v_i} = a_{0,i} \mathbf{U} - a_{i,1} \mathbf{p}_{v_1} - a_{i,2} \mathbf{p}_{v_2}, \text{ where } i = 1, 2, \quad (1)$$

or rewriting in matrix form,

$$\frac{\partial \mathbf{U}}{\partial \mathbf{u}} = \begin{pmatrix} a_{0,1} \\ a_{0,2} \end{pmatrix} \mathbf{U} - \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \begin{pmatrix} \mathbf{p}_{v_1} \\ \mathbf{p}_{v_2} \end{pmatrix} \quad (2)$$

where $\frac{\partial \mathbf{U}}{\partial \mathbf{u}}$ is a 2×3 matrix with row i as the vector $\frac{\partial \mathbf{U}}{\partial v_i}$ and \mathbf{p}_{v_1} and \mathbf{p}_{v_2} are the derivatives of the medial sheet \mathbf{p} wrt parameters v_1 and v_2 . In these equations, the derivative of \mathbf{U} is decomposed by a generally non-orthogonal projection along the spoke direction \mathbf{U} to the tangent plane of the medial sheet spanned by \mathbf{p}_{v_1} and \mathbf{p}_{v_2} .

Let $\mathbf{A}_{\mathbf{u}} = \begin{pmatrix} a_{0,1} \\ a_{0,2} \end{pmatrix}$ and $\mathbf{S}_{rad} = \begin{pmatrix} a_{1,1} & a_{2,1} \\ a_{1,2} & a_{2,2} \end{pmatrix}$. \mathbf{S}_{rad} is called the radial shape operator.

The radial shape operator is a 2×2 matrix and in general not self-adjoint.

Then (2) \Rightarrow

$$\frac{\partial \mathbf{U}}{\partial \mathbf{u}} = \mathbf{A}_{\mathbf{u}} \mathbf{U} - \mathbf{S}_{rad}^T \begin{pmatrix} \mathbf{p}_{v_1} \\ \mathbf{p}_{v_2} \end{pmatrix} \quad (3)$$

$\mathbf{U}(\mathbf{u})$ is of unit length, $\frac{\partial \mathbf{U}}{\partial \mathbf{u}} \cdot \mathbf{U}^T = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\mathbf{U} \cdot \mathbf{U}^T = 1$, so by [5]-Sec.2

$$\mathbf{A}_u = \mathbf{S}_{rad}^T \begin{pmatrix} \mathbf{p}_{v_1} \\ \mathbf{p}_{v_2} \end{pmatrix} \mathbf{U}^T \quad (4)$$

Substituting (4) into (3) yields the means of computing \mathbf{S}_{rad} given $\frac{\partial \mathbf{U}}{\partial \mathbf{u}}$, \mathbf{U} and $\begin{pmatrix} \mathbf{p}_{v_1} \\ \mathbf{p}_{v_2} \end{pmatrix}$. That is, \mathbf{S}_{rad} depends on the spoke direction \mathbf{U} , and the derivatives of \mathbf{U} and \mathbf{p} .

Furthermore, the derivative of $\mathbf{S} = r\mathbf{U}$ can be expressed using $r\mathbf{S}_{rad}$ and elementary linear algebra [12]-Ch.3. An explicit matrix expression for $r\mathbf{S}_{rad}$ is obtained as follows.

$$\frac{\partial \mathbf{S}}{\partial \mathbf{u}} = \frac{\partial(r\mathbf{U})}{\partial \mathbf{u}} = r \frac{\partial \mathbf{U}}{\partial \mathbf{u}} + \begin{pmatrix} r_{v_1} \\ r_{v_2} \end{pmatrix} \mathbf{U} \quad (5)$$

Substituting (3) and (4) into (5) yields

$$\frac{\partial \mathbf{S}}{\partial \mathbf{u}} = r\mathbf{S}_{rad}^T \begin{pmatrix} \mathbf{p}_{v_1} \\ \mathbf{p}_{v_2} \end{pmatrix} (\mathbf{U}^T \mathbf{U} - \mathbf{I}) + \begin{pmatrix} r_{v_1} \\ r_{v_2} \end{pmatrix} \mathbf{U} \quad (6)$$

The spokes being orthogonal to the boundary formed by their envelope requires [5] that $r_{v_1/v_2} = -\mathbf{p}_{v_1/v_2} \mathbf{U}^T$. Let $\mathbf{Q} = \begin{pmatrix} \mathbf{p}_{v_1} \\ \mathbf{p}_{v_2} \end{pmatrix} (\mathbf{U}^T \mathbf{U} - \mathbf{I})$ be a 2×3 matrix. Then

$$r\mathbf{S}_{rad} = \left(\left(\frac{\partial \mathbf{S}}{\partial \mathbf{u}} + \begin{pmatrix} \mathbf{p}_{v_1} \mathbf{U}^T \\ \mathbf{p}_{v_2} \mathbf{U}^T \end{pmatrix} \mathbf{U} \right) \mathbf{Q}^T (\mathbf{Q}\mathbf{Q}^T)^{-1} \right)^T \quad (7)$$

(7) shows how to compute $r\mathbf{S}_{rad}$ given the derivatives of \mathbf{p} , \mathbf{U} and \mathbf{S} wrt (v_1, v_2) .

Analogously with calling the eigenvalues and eigenvectors of the shape operator at a surface point the *principal curvatures* and *principal directions*, Damon named the eigenvalues of the radial shape operator \mathbf{S}_{rad} the *principal radial curvatures* $\kappa_{ri}, i = 1, 2$, and he named the eigenvectors the *principal radial directions* [5].

Considering a local radial flow from the medial sheet \mathbf{p} along one of the two spokes \mathbf{S} to the implied boundary as $\varphi(\mathbf{p}, t) = \mathbf{p} + t\mathbf{S}, t \in [0, 1]$. φ can be generalized to a global radial flow via the doubled-sided spoke field on the medial sheet. The spoke field is legal if and only if the Jacobian matrix of the global radial flow φ is never singular. This implies that for a legal spoke field, i.e., one free of any intersections among the spokes, it has to fulfill a legality condition [5]:

$$\lambda_{ri} < 1, \text{ where } \lambda_{ri} = r\kappa_{ri}, \text{ for all positive real eigenvalues } \lambda_{ri, i=1,2} \text{ of } r\mathbf{S}_{rad}. \quad (8)$$

This relatively simple legality condition can be converted into a geometric constraint in our training process, which allows direct control on the legality of the model interior and implied boundary. We will talk about our training process in general and then come back to how to use this legality condition explicitly in the training.

3 Method

3.1 Binary Training

As described in the introduction, the binary training starts with a fitting process. The task for binary fitting is to find the best member of the shape space for each binary

training image. Members of the shape space are parametric models, \mathbf{M} , with implied surfaces Ω . Training images, \mathbf{I} , are expert binary segmentations of 3D patient data, each with boundary voxels \mathbf{B} . We want to find the best \mathbf{M} for a given \mathbf{I} : $\arg \min_{\mathbf{M}} F_{obj}(\mathbf{M}|\mathbf{I})$. The objective function $F_{obj}(\mathbf{M}|\mathbf{I})$ is the sum of two parts: the data match measures the fit of \mathbf{M} to data including the image \mathbf{I} and image landmarks, and the geometric penalty measures the geometric appropriateness of \mathbf{M} . The data match has two terms, the image match term F_{img} and the landmark match term F_{ilm} .

3.2 Data Match

Image Match F_{img} enforces that the surface implied by model \mathbf{M} is in accordance with the boundary voxels \mathbf{B} , in the binary image. We define the image match as an integral over \mathbf{B} of the minimum distance to the implied model surface Ω as the follows.

$$F_{img}(\mathbf{M}, \mathbf{I}) \propto \sum_{b_i \in \mathbf{B}(\mathbf{I})} \min_{\Omega} (d^2(b_i, \Omega(M))) \quad (9)$$

In our current implementation, Ω is generated via a modified Catmull-Clark algorithm with additional normal constraints [14]. Ideally, we want to measure the distance of the label boundary surface from the model, $d^2(\mathbf{B}, \Omega)$. However, this is computationally expensive given finely sampled subdivision surfaces required for accurate matches and the large number of candidate surfaces generated during optimization.

Furthermore, we note that when \mathbf{B} and Ω are very close, the distance function is nearly symmetric, $|d(\mathbf{B}, \Omega) - d(\Omega, \mathbf{B})| < \epsilon$ except when in a neighborhood the normals of \mathbf{B} and Ω strongly differ. So we simplify by approximating our ideal function with the more tractable $d(\Omega, \mathbf{B})$.

In implementation, we generate a single space filling lookup table for distance from the label boundary by Danielsson’s algorithm [6]. Trilinear interpolation gives a very fast measure of the distance at any point in space to the closest boundary point on \mathbf{B} . Then we let $d(\omega_i, \mathbf{B})$ be the lookup of the position of ω_i in the distance map and $d(\Omega, \mathbf{B}) = \sum_{\omega_i \in \Omega} d(\omega_i, \mathbf{B})$.

At a boundary location where the surface normal differs from the distance gradient, i.e. the normal to Ω , by more than a certain threshold, $d(\omega_i, \mathbf{B})$ is replaced by the distance along the surface normal to the nearest binary boundary location on \mathbf{B} .

Thin object regions also pose a challenge for the image match term. An advantage of the m-rep parameterization is that the medial skeleton can be thought of as the limiting case of a morphological erosion. This allows us to segment very thin object, the image of which is less than a voxel in thickness. We fit an initially dilated model to a dilation of the initial binary labels in the training image, and then we contract the model surface by the same amount via an inverse scaling of the spoke length (radius) parameter.

Landmark Match F_{ilm} allows for identified explicit feature correspondences. An expert may identify a few anatomically important and easily identifiable landmarks in the training image population, and we penalize the distance from a medially specified point on the model to these landmarks images. Each image landmark \mathbf{LI}_i identified in \mathbf{I} has a

real tolerance value ϕ_i that reciprocally weights the associated distance to the specified model point.

In detail, landmarks \mathbf{LM} in an m-rep model \mathbf{M} are identified as spoke ends of medial atoms. Corresponding image landmarks in the image data \mathbf{I} are identified as points in image volumetric space. In our implementation, the Euclidean squared distance from the corresponding spoke ends to the image points is computed and summed, weighted by individual confidence factors $1/\phi_i^2$:

$$F_{itm}(\mathbf{M}, \mathbf{I}) \propto \sum_{\mathbf{LM}_i \in \mathbf{LM}} \frac{1}{\phi_i^2} d^2(\mathbf{LM}_i, \mathbf{LI}_i) \quad (10)$$

3.3 Geometry Penalty

Besides the measure of the dissimilarity between the model and the image data, we also put geometric constraint on the shapes to measure the appropriateness of \mathbf{M} .

The geometric penalty typically penalizes parameterizations that lead to irregular coverage of the boundary implied by \mathbf{M} , via the irregularity penalty F_{reg} . We argue that it should also include a term to penalize illegality of the interior and boundary implied by \mathbf{M} , which will be detailed in Sec. 3.4.

Irregularity Penalty The irregularity penalty F_{reg} of a discrete structure, such as a discrete m-rep may reasonably measure the deviation of every atom from the average of its symmetrically placed neighbors. This term penalizes non-uniform spacing and changes in spoke length and direction of medial atoms. It contributes to proper object geometry and to correspondence across the training cases. For each medial atom \mathbf{m}_i , the regularity is calculated as the squared Riemannian distance between \mathbf{m}_i and the Fréchet mean of its neighboring atoms $N(\mathbf{m}_i)$, where the Fréchet mean is defined as $FMean(\{\mathbf{o}_i\}) = \arg \min_{\mathbf{m}} \sum_i d^2(\mathbf{m}, \mathbf{o}_i)$ [9]. The penalties are then accumulated for all the medial atoms of the object:

$$F_{reg}(\mathbf{M}) \propto \sum_{i=1}^n d^2(\mathbf{m}_i, FMean(N(\mathbf{m}_i))) \quad (11)$$

3.4 Incorporation of $r\mathbf{S}_{rad}$ into F_{obj} as a Geometric Penalty

As described in Sec. 2.2, a fulfilled legality condition guarantees there is no local self-intersection or creases in a medial spoke field. In order to use the condition as a geometric penalty on a discrete m-rep, we need to calculate the $r\mathbf{S}_{rad}$ matrix and its eigenvalues $\lambda_{ri}, i = 1, 2$ first, and design a function $F_{leg}(\lambda_{ri})$ as the illegality penalty.

Calculating $r\mathbf{S}_{rad}$ According to (7), $r\mathbf{S}_{rad}$ can be calculated from the derivatives of the medial sheet \mathbf{p} and spoke \mathbf{S} .

We calculate the derivatives of the medial sheet and spoke by the finite differences between neighboring atoms. $r\mathbf{S}_{rad}$ is then calculated by (7). Eigen-decomposition is

applied to the $r\mathbf{S}_{rad}$ to get both the eigenvalues λ_{ri} . Recall that the spoke field is double-sided on the medial sheet, and thus there are two $r\mathbf{S}_{rad}$ matrices and corresponding four eigenvalues for each medial atom.

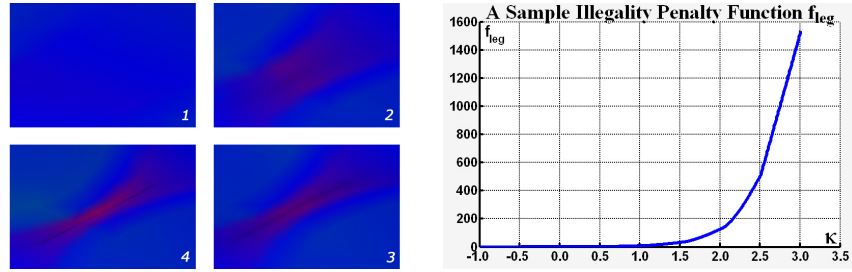


Fig. 2. Left: gradual formation of a self-intersection on a surface portion, rendered by the maximal λ_r of corresponding spokes to surface points. Blue/red means legal ($\lambda_r \leq 0$)/illegal ($\lambda_r \geq 1$) and any intermediate color shows the tendency for corresponding surface points to become illegal; **Right:** a sample illegality function f_{leg} as a cubic Hermite curve.

Fig. 2-Left shows a visualization of λ_r on the implied boundary of an m-rep. λ_r serves as a distinct indicator of local shape illegalities. Next we show the design of a penalty function on λ_r that will be used in the illegality penalty.

Illegality Penalty Recall that in (8), legality holds *iff* $\lambda_{ri} < 1$ for all real eigenvalues of $r\mathbf{S}_{rad}$. Theoretically it is a binary condition that the sign of $\lambda_{ri} - 1$ determines the legality of the implied boundary. However, in practice smaller λ_{ri} desirably leads to smoother surface and binary fit. In order to incorporate this into the objective function $F_{obj}(\mathbf{M}|\mathbf{I})$ as a penalty term, we define a smooth function $f_{leg}(\lambda_{ri})$ that has the following properties.

1. Strictly monotonically increasing;
2. Does not penalize negative λ_{ri} ;
3. Increasingly penalizes when $\lambda_{ri} > 0$ approaches or passes 1.

Ideally f_{leg} should approach ∞ when λ_{ri} approaches 1. In practice we choose f_{leg} to be smooth at $\lambda_{ri} \in [1, \infty]$ in order to simplify the gradient-based optimizations on the objective function. We use a cubic Hermite curve to define f_{leg} with the luxury of freely picking the function slopes at desired control points, which in our case are $\lambda_{ri} = 0, 1$. A cubic Hermite curve is C^2 continuous.

A sample f_{leg} is shown in Fig. 2-Right. The penalties of all n atoms' λ_{ri} in an m-rep M are then summed up and combined into the objective function $F_{obj}(\mathbf{M}|\mathbf{I})$ as an explicit illegality penalty F_{leg} for training.

$$F_{leg}(\mathbf{M}) \propto \sum_{i \in [1,4]; j \in [1,n]} f_{leg}(\lambda(\mathbf{m}_j)_{ri}) \quad (12)$$

The objective function for binary training is then defined as

$$F_{obj}(\mathbf{M}|\mathbf{I}) = \alpha F_{img}(\mathbf{M}, \mathbf{I}) + \beta F_{ilm}(\mathbf{M}, \mathbf{I}) + \gamma F_{reg}(\mathbf{M}) + \delta F_{leg}(\mathbf{M}) \quad (13)$$

where $\alpha, \beta, \gamma, \delta$ are the parameters controlling the relative weights of the penalty terms.

As a normalization, the binary image \mathbf{I} is uniformly scaled into a unit cube before the fitting process, and thus the calculated values of the four penalty terms are unitless. The binary fitting starts with a gross alignment step typically via a similarity transformation implied by the moments of the model and binary volumetric image, followed by an optimization on the objective function over m-rep parameters.

After fitting the deformable template into all binary images, the shape statistics are calculated by the Principal Geodesic Analysis (PGA) [9] on the fit models. The fit models are mapped back to their corresponding grayscale images, which are divided into regions using model coordinates provided by m-reps. Intensity histograms are then collected and converted into Regional Intensity Quantile Functions (RIQFs) [2]. PCA is applied to the RIQFs to form the appearance statistics. The PGA shape statistics and the RIQFs appearance statistics form the trained statistics that are used in applications such as segmentation or shape discrimination.

This paper focuses on image segmentation. We show some training and segmentation results based on the methods we proposed, using both synthetic and real world data. The synthetic binary images are generated by warping a standard ellipsoid by diffeomorphic deformations, and the real world data are male pelvic CT images. The next section starts by describing the generation of the synthetic data and ends with results of training on both synthetic and real world data and segmentations on male pelvis CT images of prostates and bladders.

4 Results

Adding an illegality penalty has led to better model fits hence to better statistics on geometry and intensity patterns. Our experience is that this led to considerably better segmentation results, indeed ones so good that on bladder and prostate segmentation from CT images the computer results compared to the manual segmentations by the human who did the training were closer than another human's manual segmentations were to the trainer's.

4.1 Synthetic Data

We applied a diffeomorphic deformation to the ambient space \mathbb{R}^3 in order to generate the synthetic ellipsoid models. Given the implicit form for the original ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1$, the ambient diffeomorphic deformation is defined as: $x' = x; y' = (y \cos(\rho_{twist}x) - z \sin(\rho_{twist}x))e^{\rho_{taper}x}; z' = (y \sin(\rho_{twist}x) + z \cos(\rho_{twist}x))e^{\rho_{taper}x} + \rho_{bend}x^2$, where $\rho_{bend}, \rho_{twist}, \rho_{taper}$ are three independent normally-distributed variables that control the three types of deformations: bending, twisting and tapering, as shown in Fig. 3-Left. The standard ellipsoid gives an initial binary volumetric image

to start with. By sampling three normal distributions we get a set of the three values. We then apply the deformation based on each set of $\{\rho_{bend}, \rho_{twist}, \rho_{taper}\}$ to the initial binary image and get a warped binary image. By this means we get a set of 150 sample binary images of warped ellipsoids.

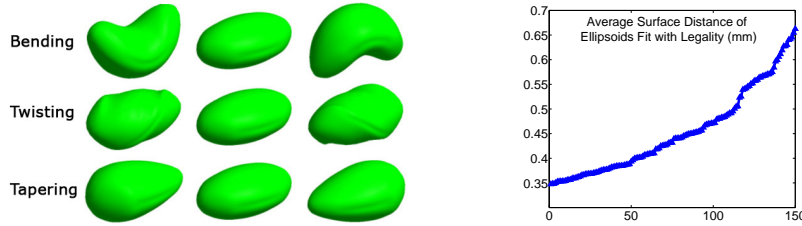


Fig. 3. Left: warped ellipsoids by three deformations of bending, twisting and tapering. Each deformation is shown as $-2\lambda/0/+2\lambda$ away from the mean; **Right:** fitting results of 150 sample images shown as the average surface distances from model to binary boundary.

For each sampled binary image, landmarks are also picked automatically as the four extreme points on the middle section of the ellipsoid boundary, plus two vertices at the two warped tips. A template model was generated by sampling the medial axis of the standard ellipsoid. The template was then fit into all the sample binary images using our binary fitting with the geometric illegality penalty. In this study we only measure the quality of the binary fit in term of the average surface distance between model and binary boundaries. Since the synthetic binary images are generated independently from our medial representation, this study serves as a consistency test on our fitting method. The fit results in Fig. 3-Right show that the m-rep surface is on the average, over the boundary, closer than one voxel from the binary boundaries as the truth, while the three axes (a, b, c) of the original ellipsoid are approximately (50, 30, 23) voxels in the binary image. Furthermore in some test cases, our proper training shows more robustness by providing good fit shapes that we failed to get without the illegality penalty. Next on real world data, we will show both training and segmentation results.

4.2 Real World Data

We used CT images ($1 \times 1 \times 3mm$) of prostates and bladders from 5 patients (80 images) as the real world test data. Each patient has from 13-18 images from multiple days. For each patient, we successively left each day out, trained on all remaining days using both our proper training method and the training without legality constraints, and segmented the left-out day image using the trained shape and appearance statistics from both training methods from all other days. The results show that the trained shapes not only have smoother surfaces (Fig. 4-Left) but also fit better (Fig. 4-Right) into the binary images. The robustness that our proper training process provides allows us to get good fits even from the images that we failed to fit without the legality constraint.

Sorted statistics over all 80 images are given in Fig. 4-Right. For comparison, m-rep fits to humans average 93%, and the average agreement between two humans' segmentations of 16 prostates is 81% volume overlap and 1.9mm average closest point surface separation. For the majority of cases, the segmented results for both the bladder and prostate were judged qualitatively good.

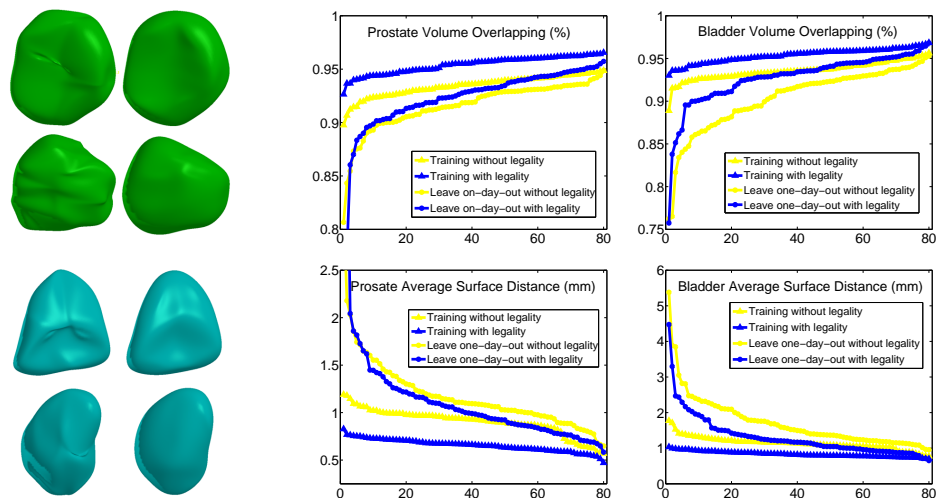


Fig. 4. **Left:** sample trained shapes from training without/with legality constraints, shown in the left/right column; **Right:** sorted measures (volume overlapping as *Dice similarity coefficients* [8] and average surface distance in *mm*) comparing m-rep segmentations (circle) and trained m-reps (triangle) to human segmentations, using training methods without (yellow) and with (blue) illegality penalty.

5 Conclusion and Future Work

The results in Fig. 4 suggest that the segmentations results based on our proper training process are improved from the results based on training without illegality penalty. Our training process incorporated with the legality constraint helps us to get proper shape and appearance statistics and lead to segmentation results that are as good as human segmentations. This shows that our method is not only mathematically sound but also proves in practice to improve segmentation results.

In the current implementation, we use adapted Catmull-Clark subdivision to generate the implied boundary used in the training, and legality constraints are applied to sample medial atom spokes. A medial spoke interpolation method [10] is also being adopted to generate both a smooth spoke field and implied boundary, which can further improve the legality of the interior and boundary of trained shapes. We are also working on extending the same framework to train models for anatomical objects with multiple

parts, such as livers and hearts. Application of the properly trained shape statistics in methods finding significant differences between classes is in progress.

6 Acknowledgement

The authors are grateful to all members in MIDAG group, especially Joshua Stough for the latest segmentation results and the help in creating some of the figures, and Graham Gash, Gregg Tracton, Eli Broadhurst, Ja-Yeon Jeong, Xiaoxiao Liu and Rohit Saboo for conceptual, algorithmic, code, and operations contributions. The work was done under the support of NIH grant P01 EB02779.

References

1. C. Brechbuhler, G. Gerig, and O. Kubler. Parametrization of closed surfaces for 3-d shape description. *Computer Vision and Image Understanding*, 61(2):154–170, 1995.
2. R. Broadhurst, J. Stough, S. Pizer, and E. Chaney. A statistical appearance model based on intensity quantiles. In *International Symposium on Biomedical Imaging (ISBI)*, pages 422–425, 2006.
3. T. Cootes, D. Cooper, C. Taylor, and J. Graham. Active shape models - their training and application. *Computer Vision and Image Understanding*, 61(1):18–59, 1995.
4. E. Dam, P. T. Fletcher, S. M. Pizer, G. Tracton, and J. Rosenman. Prostate shape modeling based on principal geodesic analysis bootstrapping. In *MICCAI*, (2):1008–1016, 2004.
5. J. N. Damon. Smoothness and geometry of boundaries associated to skeletal structures I: Sufficient conditions for smoothness. *Annales de Institut Fourier*, 53(6):1941–1985, 2003.
6. P. Danielsson. Euclidean distance mapping. *Computer Graphics and Image Processing*, pages 227–248, 1980.
7. R. Davies, C. Twining, T. Cootes, and C. Taylor. A minimum description length approach to statistical shape modelling. *IEEE Transactions on Medical Imaging*, 21(5):525–537 2002.
8. L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26:297–302, 1945.
9. P. T. Fletcher, C. Lu, S. M. Pizer, and S. Joshi. Principal geodesic analysis for the nonlinear study of shape. *Transactions on Medical Imaging*, 23(8):995–1005, 2004.
10. Q. Han, S. M. Pizer, and J. N. Damon. Interpolation in discrete single figure medial objects. In *IEEE CVPR Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 2006)*, page 85, 2006.
11. S. M. Pizer, T. Fletcher, Y. Fridman, D. S. Fritsch, A. G. Gash, J. M. Glotzer, S. Joshi, A. Thall, G. Tracton, P. Yushkevich, and E. L. Chaney. Deformable m-reps for 3d medical image segmentation. *International Journal of Computer Vision - Special UNC-MIDAG issue*, 55(2):85–106, 2003.
12. K. Siddiqi and S. M. Pizer, editors. *Medial Representations: Mathematics, Algorithms and Applications*. Springer. In press, 2007.
13. M. Styner, I. Oguz, S. Xu, C. Brechbuhler, D. Pantazis, J. Levitt, M. Shenton, and G. Gerig. Framework for the statistical shape analysis of brain structures using spharm-pdm. In *Open Science Workshop at MICCAI*, <http://hdl.handle.net/1926/215>, 2006.
14. A. Thall. Deformable solid modeling via medial sampling and displacement subdivision. *PdD thesis, Dept. of Comp. Sci., UNC @ Chapel Hill*, 2004.