A Methodology for Constructing Geometric Priors and Likelihoods for Deformable Shape Models

Derek Merck, Gregg Tracton, Stephen Pizer, and Sarang Joshi

Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill

Abstract. Deformable shape models require correspondence across the training population in order to generate a statistical model for use as a future geometric prior. Traditional methods use fixed sampling and assume correspondence, or attempt to induce correspondence by minimizing variance. In this paper, we define a training methodology for sampled medial deformable shape models (m-reps) which generates correspondence implicitly via a geometric prior. We present quantitative results of the method applied to real medical images.

1 Introduction

Automatic segmentation of medical images is a vital step in processing large populations for quantitative study of biological shape variability. In this paper, we present a methodology for statistically training the geometry of deformable model templates that can be used as geometric prior and basis for intensity training for automatic segmentation of gray images. Our method frames the problem as a special case of the general segmentation problem. Given a data set of human or otherwise expertly segmented training cases, we fit models to the labeled data, and then create our template by statistical analysis of the fit population. The fitting is an optimization over model parameters in a Bayesian framework, searching for the model with the highest posterior probability of fitting the data. Our posterior is decomposed into data likelihood and geometric prior terms. The data likelihood accounts for both image match and optional landmark match. The geometric prior encourages models to stay in a legal shape-space. We describe an implementation of the method using m-reps and present results showing that the method is accurate and yields models suitable for statistical analysis.

Deformable Models We desire a statistical model for a population of training data. Probabilistic deformable models describe shape variability via a probability distribution on the shape-space. Under the Gaussian model, the distribution of the training data can be modeled by the mean, a point in the space, and several eigenmodes of deformation. This model describes all possible shapes in the training data, and by extension, estimates the actual ambient shape-space from which the training data is drawn. This statistical model can then be used as

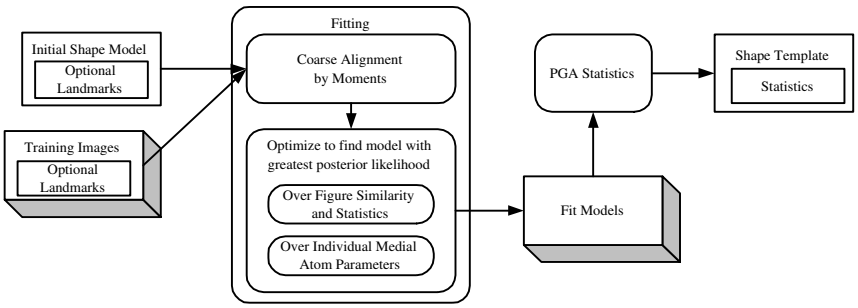

Fig. 1. The method starts with an initial model and n images from the training population. It results in a geometry trained model which can be used as a geometric prior for data segmentation and to train intensity likelihoods.

a geometric prior probability when searching for an object from this shape-space while segmenting new data sets.

For point distribution models (PDMs), the shape variance is classically described by principal component analysis (PCA) of the feature space [1]. However, this method requires Euclidian parametrization of the model features and is only as effective as its ability to identify and maintain correspondence between sampled features in the training data sets. Our training method uses sampled medial representations (m-reps), a symmetric space shape parameterization that allows us to model local bending and twisting as well as provides us with an object-centric coordinate system. Principal component analysis has been extended to non-Euclidean domains like that of m-reps as principal geodesic analysis [3]. The deformable template becomes the Fréchet mean of the training population and a set of principal modes of variance which may include coefficients of position, rotation, and scale for each sample.

An m-rep figure is parameterized as samples on an object's medial manifold. The medial samples or 'atoms' have eight parameters: three position, four pose (two spoke directions), and scale. An implied boundary surface must be computed from the medial samples; our implementation uses a variant of Catmull-Clark subdivision with a nearly interpolating surface and normal constraints [4]. M-reps are a hierarchical shape domain; collections of m-rep samples form figures, and collections of figures may be taken together to form more complex objects and object groups. Here, because we have restricted ourselves to single object binary labeled images and shapes which can be adequately modeled with a single medial manifold, we need only work at the atom and figure levels.

2 Method

The general procedure is shown in figure 1. The method begins with n hand segmented images, in which every voxel has been labeled as inside or outside

the object of interest and anatomic landmarks have optionally been identified. Then, given a starting model with approximately the correct sampling rate for the object's complexity, we find the most likely deformation of the model into each training image by optimization. Finally, we gather geometric statistics of the fit shapes and produce the trained template model.

2.1 Definitions

Data: Let I be a binary labeled image. Let $LI = \{li_1..., li_m\}$ be a set of landmarks identified in the image, each with confidence σ_i.
Model: Let M be a medial model with samples $\{m_1, ..., m_n\}$. Let Ω_M^s be the discrete approximation of the implied surface of model M at subdivision level s. Let LM be a set of landmarks corresponding to LI identified in M.
Distances: Let $d(x, y)$ be the Riemannian distance between two points[3]. For points in 3-space, this is the regular Euclidean distance; for medial samples, this is a symmetric space distance.

2.2 Designing a Posterior

We use a Bayesian framework and seek the model that maximizes $P(M|Data)$ and so has the greatest conditional probability given the data. By Bayes Rule $P(M|Data) \propto P(Data|M)P(M)$, so we describe $P(M|Data)$ as two terms: a data likelihood and a geometric prior. Our data likelihood term must reflect our desires that the model's surface be accurate to the data's label boundary and, if landmarks have been identified in the images, that the model's corresponding landmarks must interpolate those points. Our geometric prior term must reflect our desires that the fit model be stable and implicitly maintain correspondence, which we define as proximate in shape to the starting model and preserving topology (no bending, folding). As log is a monotonic function, we can maximize the log probability to the same result as maximizing the original term. Therefore, we actually describe the elements of a log posterior in the next sections.

2.3 Data Likelihood

Our data likelihood term is jointly conditioned on the model's fit to the image data, and the model's fit to the landmark data. Assuming these factors are independent, the joint probability is the product $P(Data|M)) = P(I|M)P(LI|M)$.

Image Likelihood Our image likelihood term is specific to the problem of matching binary labeled data from an existing by-hand segmentation. In this case, we have the correct answer at hand: the binary image labeling. We compute the boundary of the label and consider it as a surface, B, then using a Gaussian model, we define the log image likelihood as a function of the distance of the label boundary to the model surface:

$$log(P(I|M)) \propto -\frac{1}{\alpha} \sum_{\omega_i \epsilon \Omega_M^s} d^2(B, \omega_i) \tag{1}$$

with α a weighting scalar.

Landmark Likelihood We introduce a landmark term to create model-to-image correspondences at anatomically important points and also to add limited points of explicit model-to-model correspondence across the training population [5] [6]. An expert identifies landmarks in the training image population, then we constrain the fit models to always interpolate these points at the same object-coordinates. Similar to the image likelihood term, we define the landmark likelihood as a normal probability over the positions of the landmarks LM identified in M, with mean at the landmarks LI identified in the image, and with σ_i equal to a tolerance or confidence assigned to each pairing lm_i to li_i.

$$log(P(LI|M)) \propto -\sum_{li_i \epsilon LI} \frac{d^2(lm_i, li_i)}{\sigma_i} \tag{2}$$

2.4 Geometric Prior

Geometric Reference Likelihood Our aim is to construct a geometric prior before we have a statistical estimate of the shape space. If we had a statistical model available, we would define a geometric prior as the exponentiated Mahalanobis distance of a model from the mean shape. However, before we have a statistical model, we make two simplifying assumptions: first that our starting model is representative of the mean of this shape space, and second, that the shape space is isotropic with an identity covariance. In this case, the Mahalanobis distance to the mean is simplified to the sum of the atom-to-atom symmetric space distances between the candidate model and the starting model. Let R be the starting model and β be a scalar weighting factor, then the log probability of the shape is as follows:

$$log(P_{Ref}(M)) \propto -\frac{1}{\beta} \sum_{M} d^2(m_i, r_i) \tag{3}$$

This term is invariant to the subdivision level of the surface and does not assume point-to-point correspondences on the model's surface.

Geometric Regularity Likelihood However, without a statistical model, a prior based on the distance from the starting model alone is not sufficient to induce shape legality. We simplify our enforcement of legality to simply looking at regularity of sampling on the medial sheet. Our regularity likelihood is motivated by Markov random field assumptions about a sample's dependence on neighboring parameters. Let $\mathcal{N}(m)$ be the neighborhood of atom m and δ be a scalar weighting factor, then:

$$log(P_{Reg}(M)) \propto -\frac{1}{\delta} \sum_{m_i \epsilon M} \frac{1}{|\mathcal{N}(m_i)|} \sum_{m_j \epsilon \mathcal{N}(m_i)} d^2(m_i, m_j) \tag{4}$$

Minimizing this expression as a function of m_i is equivalent to minimizing the symmetric space distance of each sample m to the Fréchet mean of its neighbors.

Finally, we describe our full $P(M)$ as an independent joint distribution of the geometry reference likelihood and regularity likelihood with joint density $P(M) \propto P_{Ref}(M)P_{Reg}(M)$.

2.5 Optimization

We combine terms to form an objective function of the full log posterior:

$$\begin{aligned} log(P(M|Data)) \propto & -\tfrac{1}{\alpha} \textstyle\sum_{\omega_i \epsilon \Omega_M^s} d^2(B, \omega_i) - \sum_{li_i \epsilon LI} \frac{d^2(lm_i, li_i)}{\sigma_i} \\ & -\tfrac{1}{\beta} \textstyle\sum_M d^2(m_i, r_i) - \tfrac{1}{\delta} \sum_{m_i \epsilon M} \frac{1}{|\mathcal{N}(m_i)|} \sum_{m_j \epsilon \mathcal{N}(m_i)} d^2(m_i, m_j) \end{aligned} \tag{5}$$

We initialize the optimization by coarse alignment of the model to the data via method of moments. We compute the model's centroid, volume, and covariance by integral on the model's surface according to Stokes' Theorem. We compute the image moments directly by integration over the voxel volume. The model moments are then aligned to the data by a similarity transformation applied to the model. Because of ambiguity in orientation, we chose the rotation with the smallest angle.

We then optimize by seeking a model parameterization that maximizes our log posterior. Because our shape models are organized hierarchically, we maximize our objective function over parameters appropriate to the working level of the model description. At the figure level, the objective function is optimized over a similarity transform. At the medial sample level, we optimize over each sample's similarity component (position, pose, size), as well as each sample's extended parameters (local bending, twisting). We maximize via a conjugate gradient search as described in Numerical Recipes [7]. To find the gradient direction at each step in the parameter space, we use numeric differentiation; we evaluate the model at several values for each parameter and take an initial step in the aggregate best direction. Because conjugate gradient search performs best given a relatively isotropic global minimum, some effort is expended tuning the parameter scalings of the algorithm.

Surface-to-Surface Distance Implementation Details Ideally we desire to measure the distance of the label boundary surface from the model, however, this is computationally exorbitant given finely sampled subdivision surfaces required for accurate matches and the large number of candidate surfaces generated for optimization. Furthermore, we note that near zero, the distance of the label boundary to the model is equal to the distance of the model to the label boundary, so we simplify by approximating one by the other.

Because the label boundary is static, it is quite fast to generate a lookup table for distance function from the model surface to the label boundary. We compute our lookup table inside and outside the label boundary by Danielsson's algorithm [8]. Six connected neighbors are used to find B. Trilinear interpolation of values out of the lookup table gives a very fast measure of the distance at any point in space to the closest boundary point on B. Then we let $d(B, \omega)$ be the lookup of the position of ω in the distance map.

This approximation may be enhanced by computing the true label boundary to model distance at a minimal number of points. We identify points where we would expect the label boundary to model distance to be different than the model to label boundary distance by checking the gradient of the distance map against the surface normal. If their dot product is below some threshold, we compute a new distance along the model's normal for that point.

2.6 Statistical Analysis

We then build a deformable shape model by computing the Fréchet mean and PGA eigenvectors of the fit models [3]. This model becomes a better estimator of the shape space than our untrained geometric prior. We now redefine our statistical geometric prior explicitly as proportional to an exponentiated Mahalanobis distance of the candidate model to the mean model. This covariance weighted distance is computed by projecting the model into the PGA space and then simply taking the norm of the eigenvalue scaled PGA coefficients. That is, the Mahalanobis distance squared $= \sum \frac{\alpha_i^2}{\lambda_i}$ where α are the coefficients of M expressed in PGA eigenvectors and λ are the corresponding eigenvalues.

3 Results

We fit two sample data sets with our implementation of the method: a left kidney trained from forty-two landmarked training images at 0.3mm voxel resolution, and a left hippocampus trained for forty-six unlandmarked training images at 0.5mm voxel resolution. We present results showing that the implied surfaces of the fit models are accurate to the training data, thus the statistical model is gathered over a population of correct models. We also show that the fit models create a stable statistical sample by leave-one-out testing.

Quality of Fit Both data sets were automatically fit by our method. There were no 'illegal' results, no shapes with broken topology, and all looked encouragingly like the correct shape. Quantitative results are summarized in table 1.

Stability We then calculated the Fréchet mean and PGA eigenmodes of our fit kidney models. Figure 2 shows the mean and +/- two standard deviations of deformation along the first four eigenmodes of the shape space. As shown in

	After Figure			After Figure and Atom		
Structure	μ Ave Dist	σ Ave Dist	Worst Dist	μ Ave Dist	σ Ave Dist	Worst Dist
Kidney	0.180mm	0.023mm	0.977mm	0.161mm	0.010mm	1.201mm
Hippocampus	0.536mm	0.058mm	3.556mm	0.263mm	0.013mm	2.880mm

Table 1. Model surface to label boundary distances after optimization stages

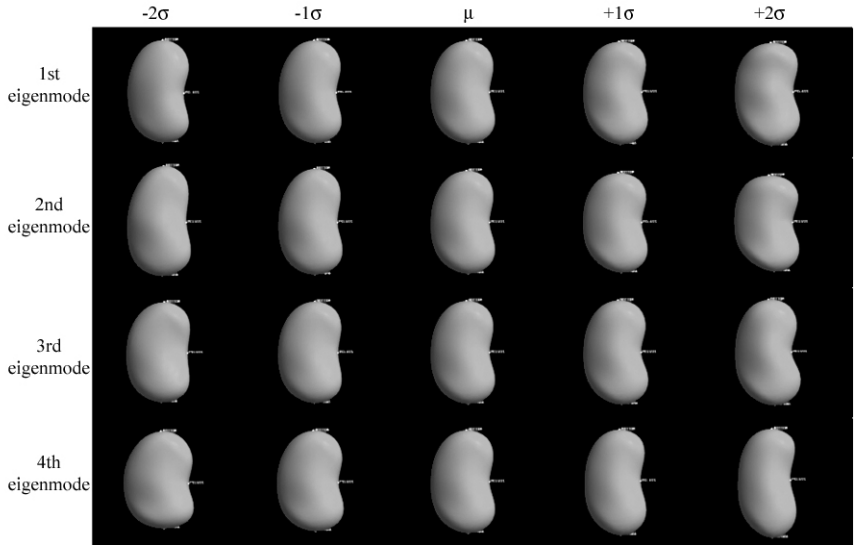

Fig. 2. Trained kidney template and two standard deviations of deformation in each of first four principal modes of shape variance

figure 3, the first four eigenmodes account for over 60% of the training data's shape variance.

With variance computed as the sum of the PGA feature space eigenvalues, the entire training set has a symmetric space variance of 0.0013 units. To demonstrate that our fitting method generates tightly clustered models suitable for our statistical model, we recompute Fréchet means on subsets of the training set, leaving out each one of the training cases in turn. This results in a set of forty-two different kidney templates which, taken together, define a new PGA feature space. This feature space has a variance of 7.3e-007 units, three orders of magnitude tighter distribution than the original training set.

4 Discussion

This methodology for training deformable shape templates is accurate and extends to a variety of shapes. Results in this paper were generated with our

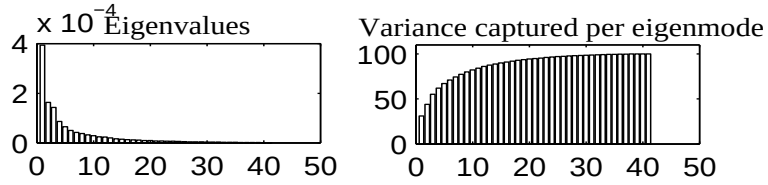

Fig. 3. Eigenvalues of the kidney population and variance captured per eigenmode

implementation of the method, *Binary Pablo*. Binary Pablo has a variety of visualization and m-rep modeling tools, and automatic fitting runs as a configurable batch process. It works quickly, producing the kidney and hippocampus models presented in our results in under two minutes per model, and population analysis can be easily parallelized over a network, scaling in speed with number of machines. Binary Pablo is currently being applied to a variety of shapes in several different labs. It is is available as a freely licensed download from our group and is distributed with a complete user's guide and example data.

Substantial continuing research in UNC MIDAG is focused on shape model training for shape analysis, primarily developing extensions for multi-figure models with more complex medial sheets [9] and for multi-object constructs (i.e., bladder, rectum, prostate in the male pelvis) [10]. Results of this paper also suggest the possibility of training by statistical bootstrapping; feeding the statistically trained deformable model back into the pipeline to increase training accuracy.

References

1. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models - their training and application. Computer Vision, Graphics, and Image Processing: Image Understanding **1** (1994) 38–59
2. Joshi, S., Pizer, S., Fletcher, P., Yushkevich, P., Thall, A., Marron, J.: Multiscale deformable model segmentation and statistical shape analysis using medial descriptions. IEEE Transactions on Medical Imaging **21** (2002) 538–550
3. Joshi, S., Pizer, S., Fletcher, P., Yushkevich, P., Thall, A., Marron, J.: Multiscale deformable model segmentation and statistical shape analysis using medial descriptions. IEEE Transactions on Medical Imaging **21** (2002) 538–550
4. Thall, A.: Deformable Solid Modeling via Medial Sampling and Displacement Subdivision. PhD thesis (2004)
5. Bookstein, F.: Principal warps: Thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence **1** (1989)
6. (Anonymous)
7. Press, W.H., Vetterling, W.T., Teukolsky, S.A., Flannery, B.P.: Numerical Recipes in C++: the art of scientific computing. (2002)
8. Danielsson, P.E.: Euclidean distance mapping. Computer Graphics and Image Processing (1980) 227–248

9. Han, Q., Pizer, S., Merck, D., Joshi, S., Jeong, J.Y.: Multi-figure anatomical objects for shape statistics. In: To appear, Information Processing in Medical Imaging (IPMI). (2005)
10. Chaney, E., Pizer, S., Joshi, S., Broadhurst, R., Fletcher, T., Gash, G., Han, Q., Jeong, J., Lu, C., Merck, D., Stough, J., Tracton, G., J. Bechtel, M., Rosenman, J., Chi, Y., Muller, K.: Automatic male pelvis segmentation from ct images via statistically trained multi-object deformable m-rep models. In: Presented at American Society for Therapeutic Radiology and Oncology (ASTRO). (2004)