A Large-to-Fine-Scale Shape Prior for Probabilistic Segmentations Using A Deformable M-rep

Xiaoxiao Liu, Ja-Yeon Jeong, Joshua H. Levy,
Rohit R. Saboo, Edward L. Chaney, and Stephen M. Pizer

Medical Image Display & Analysis Group (MIDAG)
University of North Carolina, Chapel Hill, NC 27599, USA
sharonxx@cs.unc.edu

Abstract

Training a shape prior has been potent scheme for anatomical object segmentations, especially for images with noisy or weak intensity patterns. When the shape representation lives in a high dimensional space, principal component analysis is often used to calculate a low dimensional variation subspace from frequently limited number of training samples. However, the eigenmodes of the subspace tend to keep the large-scale variation of the shape only, losing the detailed localized variability which is crucial to accurate segmentations. In this paper, we propose a large-to-fine-scale shape prior for probabilistic segmentation to enable local refinement, using a deformable medial representation, called the m-rep.

Tests on the goodness of the shape prior are carried out on large simulated data sets of a) 1000 deformed ellipsoids with mixed global deformations and local perturbation; b) 500 simulated hippocampus models. The predictability of the shape priors are evaluated and compared by a squared correlations metric and the volume overlap measurement against different training sample sizes. The improved robustness achieved by the large-to-fine-scale strategy is demonstrated, especially for low sample size applications. Finally, posterior 3D segmentations of the bladder from CT images from multiple patients in day-to-day adaptive radiation therapy demonstrate that the local residual statistics introduced by this method improve the segmentation accuracy.

1. Introduction

Automatic 3D medical image segmentation via deformable models is challenging; adoption occurs only if its accuracy is competitive with manual segmentations. Deformable models with various representations have been proven effective in capturing variation across a population of certain geometric entities [10]. It is important to collect the statistics about the shape space via proper training in order to constrain the deformation field, especially for processing noisy image patterns. Given a sizable training sample, an off-line learning process can build up the probability distribution of the shape variation [2, 13], namely the shape prior.

With high dimensional data representations, principal-component-analysis (PCA)-based methods are often used to extract a robust subspace with a fairly small number of dimensions as the shape prior. However, as also pointed out in [12], the trained shape variation space often cannot reflect local perturbations of the object; oversmooth segmentations often result. Further, because the number of training samples is often smaller than the number of geometric features, a condition known as high dimension, low sample size (HDLSS), the extracted shape space is not robust in terms of limited predictability.

Considering the various degrees of localities for geometric features in many shape representations, an efficient way of representing the shape space is to decompose them into different scales.

Multiscale shape models such as hierarchical ASMs [4] and multiscale spherical wavelets [12] successfully address this issue via wavelet basis functions with shape localities. Here we propose a more general approach to "decompose" the variation space by training the residual statistics hierarchically, which tends to be applicable to various representations or parametrizations of deformable modelsl with notions of locality.

A recent study shows the medial representation has superior efficiency over the boundary representation [9]. We use a discrete medial representation called the *m-rep* (Fig.1) [18]. M-reps have been successfully applied to various medical image segmentation tasks [17] and statistical shape

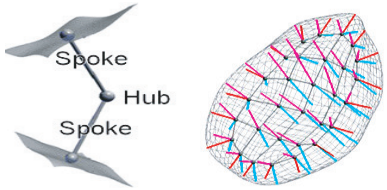

Figure 1. An m-rep model: An interior medial atom (left) with two spokes and an object (right). The object is composed of multiple medial atoms. Interior atoms have two spokes each. Exterior atoms (on the crest region) have three spokes each.

analysis [20]. In terms of average closest point distance and Dice similarity coefficient, the m-rep segmentation is among one of the best in literature for segmentation of male pelvic organs from CT [3, 15, 6, 14, 19]. However, the global m-rep segmentation results are not accurate enough due to the oversmooth boundaries, which often are several voxels off the manually segmented boundaries in local regions. Joshi *et al.* [8] developed a multiscale m-rep segmentation where the global scale stage optimizes the posterior probability of the shape with trained statistics and where the local scale stage is a deterministic optimization on the full space of atom deformations. However, without shape prior information, the local scale deformation can be unpredictable when the object boundary is in a low-contrast image region, and the ad hoc weighting penalties are tedious to optimize.

The m-rep models of objects with simple geometry usually contain a sparse grid of atoms; the atoms naturally provide the local scale. A two-scale (object-scale and atom-scale) probabilistic segmentation scheme via the m-rep is developed in this paper as an important step towards the ideal large-to-fine-scale strategy, in which more careful studies on the number of scales and the size of each scale are needed.

The rest of the paper is organized as follows: Section 2 begins by introducing the large-to-fine-scale shape prior via the m-rep. It ends by testing the robustness of the statistics against training sample size on 1000 simulated warped ellipsoids with both global and local deformations and 500 hippocampus m-rep models sampled from 51 trained variation modes. Section 3 presents the posterior segmentation methodology for applying shape statistics. In Section 4, bladder segmentation results from 79 clinical CT images are discussed in detail. Section 5 summarizes and discusses the method and results.

2. Large-to-fine-scale shape space analysis

2.1. Locality of m-reps

In single-figure m-rep models, an object is a sheet of medial atoms represented by a quadrilateral mesh (Fig.1), which carries geometric properties such as widening, bending and tapering, with the locality scale given according to the grid spacing. Each atom controls a local boundary region implied by its spokes' ending points. The object as a whole gives these properties in a way reflecting global interrelations of atoms, whereas each atom, and its relation to its immediate neighbors, provides more local features.

For simple biological objects represented by a sparse grid of medial atoms, typically less than 50 atoms, object and atom scales are sufficient to describe the shape variation globally and locally. For geometrical structures that are characterized by hundreds of atoms, it might be necessary to define more locality scales by grouping n-nearest neighboring atoms, where n controls the size of each scale.

2.2. Large-to-fine-scale shape prior

Let $\mathbf{m}$ denote a m-rep model. When the i^{th} medial atom $\mathbf{A}_i$ of $\mathbf{m}$ lies on the interior of its medial grid, $\mathbf{A}_i$ is described by the following parameters: its hub position $\mathbf{p}_i \in \mathbf{R}^3$, the common length $r_i \in \mathbb{R}^+$ of its spokes, and a pair of unit spoke directions $\mathbf{U}_i^{+1}, \mathbf{U}_i^{-1} \in \mathbf{S}^2$, such that the points $\mathbf{p}_i + r_i\mathbf{U}_i^{+1}$ and $\mathbf{p}_i + r_i\mathbf{U}_i^{-1}$ lie on the boundary of the object and the outward normal directions to the surface at those points are $\mathbf{U}_i^{+1}$ and $\mathbf{U}_i^{-1}$, respectively. Each interior medial atom is understood to lie in the curved space $\mathcal{S} = \mathbb{R}^3 \times \mathbb{R}^+ \times \mathbb{S}^2 \times \mathbb{S}^2$. Each exterior medial atom requires an additional parameter from $\mathbb{R}^+$ that controls the distance to a third point of contact with the boundary along the vector $\left(\mathbf{U}_i^{+1} + \mathbf{U}_i^{-1}\right)/2$. To simplify notation, we will ignore this elongation factor and act as if each $\mathbf{A}_i \in \mathcal{S}$. Thus when $\mathbf{m}$ is composed of k medial atoms, it can be assumed to lie on the curved space formed by the direct product $\mathcal{S}^k$.

Fletcher *et al.* [5] developed metrics on $\mathcal{S}$ and $\mathcal{S}^k$ that allow these curved spaces to be understood as Riemannian manifolds. For a medial atom $\mathbf{A} \in \mathcal{S}$ and an m-rep object $\mathbf{m} \in \mathcal{S}^k$, they defined Riemannian Log maps from these manifolds to their tangent planes: $\text{Log}_{\mathbf{A}} : \mathcal{S} \to T_{\mathbf{A}}\mathcal{S}$ and $\text{Log}_{\mathbf{m}} : \mathcal{S}^k \to T_{\mathbf{m}}\mathcal{S}^k$ and Riemannian Exp maps from the tangent plane back to the manifold: $\text{Exp}_{\mathbf{A}} : T_{\mathbf{A}}\mathcal{S} \to \mathcal{S}$ and $\text{Exp}_{\mathbf{m}} : T_{\mathbf{m}}\mathcal{S}^k \to \mathcal{S}^k$.

Fletcher *et al.* use these metrics and maps to apply principal geodesic analysis (PGA), their generalized version of PCA for manifold data, to m-rep objects to train a shape prior. PGA applies PCA to the projection of manifold data by the Riemannian Log map into the tangent plane at their Fréchet mean. These eigenmodes of variation in the tangent plane are mapped onto geodesics on the manifold by the Riemannian Exp map. Our large-to-fine-scale shape prior is built by PGA first at the whole object scale and then at an individual atom scale. The atom-scale prior is built upon the residual space of the object-scale prior.

2.2.1 Object-scale prior

For the object scale, the whole grid of atoms is taken by the Riemannian Log map at the Fréchet mean object to a vector in the high dimensional shape space. For example, a model with 21 atoms in a 3×7 grid maps onto a 184 dimensional vector.

The first several eigenmodes with the largest eigenvalues produced by PGA usually cover most of the variation from the training samples. By only using the subspace composed of these few eigenmodes, we successfully reduce the dimension of the shape space from several hundred to a number often less than 10. We avoid eigenmodes with little variance that are not robustly estimated in the typical situation of a limited number of training cases and a high-dimensional object representation. The remaining variability is captured by training residual statistics at a local scale.

2.2.2 Atom residual prior

If we regard the object-scale prior as the *main* shape space, we can think of the difference between *main* shape space and the actual space as the *residual* shape space. The *residual* shape space contains high frequency signals of the shape details that are missing at the object scale. We again use PGA to analyze this *residual* shape space, but locally on each atom to match the "signal" with the "resolution".

Fletcher *et al*. have shown that $\mathcal{S}$ the space of medial atoms forms a symmetric space, and that the metric on $\mathcal{S}$ is left-invariant with respect to the symmetric space action: $\circ : \mathcal{S} \times \mathcal{S} \rightarrow \mathcal{S}$. For a precise definition of $\circ$, we refer the reader to [5].

Given two medial atoms $\mathbf{A}_i$, $\mathbf{A}'_i \in \mathcal{S}$, the residual $\Delta\mathbf{A}_i \in \mathcal{S}$, from $\mathbf{A}'_i$ to $\mathbf{A}_i$, can be understood with respect to the action $\circ$ and its inverse.

$$\begin{aligned}\Delta\mathbf{A}_i &\doteq (\mathbf{A}'_i)^{-1} \circ \mathbf{A}_i \\ &\doteq (\Delta\mathbf{p}_i, \Delta r_i, \Delta\mathbf{U}_i^{+1}, \Delta\mathbf{U}_i^{-1}). \\ &\doteq (\mathbf{p}_i - \mathbf{p}'_i, r_i/r'_i, \mathbf{R}_{\mathbf{U}_i^{+1\prime}}(\mathbf{U}_i^{+1}), \mathbf{R}_{\mathbf{U}_i^{-1\prime}}(\mathbf{U}_i^{-1})).\end{aligned} \tag{1}$$

$\mathbf{R}_\mathbf{w}$ represents the rotation along the geodesics in $\mathbb{S}^2$ that moves a point $\mathbf{w} \in \mathbb{S}^2$ to the North Pole (0, 0, 1) $\in \mathbb{S}^2$. The atom can be recovered from the residual as follows.

$$\begin{aligned}\mathbf{A}_i &\doteq \mathbf{A}'_i \circ \Delta\mathbf{A}_i \\ &\doteq (\mathbf{p}'_i + \Delta\mathbf{p}_i, r'_i \cdot \Delta r_i, \mathbf{R}^{-1}_{\mathbf{U}_i^{+1\prime}}(\Delta\mathbf{U}_i^{+1}), \\ &\qquad \mathbf{R}^{-1}_{\mathbf{U}_i^{-1\prime}}(\Delta\mathbf{U}_i^{-1})).\end{aligned} \tag{2}$$

The shape residuals that we use to train the atom prior are of the form that follows.

$$\Delta\mathbf{A}_i = (\hat{\mathbf{A}}_i^{obj})^{-1} \circ \mathbf{A}_i^{train}, \tag{3}$$

where $\mathbf{A}_i^{train}$ is the i^{th} atom of a training m-rep and $\hat{\mathbf{A}}_i^{obj}$, formally defined below, is the i^{th} atom in the projection of training model onto the main shape space. Let μ_i^{train} denote the the Fréchet mean of this atom from the set of training objects. The dot product of the model deviation vector in the tangent plane at μ_i^{train} with the K^{obj} eigenvectors learned via object-scale-PGA produces the coefficients for the projection model:

$$\begin{aligned}\mathbf{x}_p &= < \underset{\mu_i^{train}}{\text{Log}}\, \mathbf{A}_i^{train}, \mathbf{v}_p^{obj} > \\ \hat{\mathbf{A}}_i^{obj} &= \underset{\mu_i^{train}}{\text{Exp}} \left(\sum_{p=1}^{K^{obj}} \mathbf{x}_p \mathbf{v}_p^{obj} \right).\end{aligned} \tag{4}$$

Based on the projections onto the main shape space, the training samples for computing the atom prior therefore have to be the same with the object-scale prior. We apply PGA on $\left(\hat{\mathbf{A}}_i^{obj}\right)^{-1} \circ \mathbf{A}_i^{train}$ to learn the principal modes of residual variation for each atom. The atom parametrization is 8- or 9- dimensional, depending on whether it is an interior or an exterior atom (Fig.1). The PGA results on a single atom typically yields four non-neglible eigenmodes. Similar to the object-scale projection, we project the residual into its trained residual shape space. Together with the object-scale projection, we calculate the multiscale projection as follows:

$$\begin{aligned}\mathbf{x}_q &= < \underset{\hat{\mathbf{A}}_i^{obj}}{\text{Log}}\, \mathbf{A}_i^{train}, \mathbf{v}_q^{atom} > \\ \hat{\mathbf{A}}_i^{Multiscale} &= \underset{\hat{\mathbf{A}}_i^{obj}}{\text{Exp}} \left(\sum_{q=1}^{K^{atom}} \mathbf{x}_q \mathbf{v}_q^{atom} \right),\end{aligned} \tag{5}$$

with $K^{atom} \leq 8$.

To date we have assumed that the probability distribution of local residuals $\Delta\mathbf{A}_i$ on atom i is independent of neighboring atoms. We are thereby assuming that the correlations between the atoms are contained in the object-scale PGA and that the residual variations of each atom are independent of the other atoms. Let $\hat{\mathbf{A}}_i^{obj}$ and $\hat{\mathbf{A}}_i^{Multiscale}$ denote the projection into the object and atom scale shape spaces described above of a particular atom of an object instance. Let $\hat{\mathbf{m}}^{obj} = \cup_i \hat{\mathbf{A}}_i^{obj}$ and $\hat{\mathbf{m}}^{Multiscale} = \cup_i \hat{\mathbf{A}}_i^{Multiscale}$ denote the projections of that object instance into the shape spaces. The multiscale prior is thus

$$\begin{aligned}p\left(\hat{\mathbf{m}}^{Multiscale}\right) &= p\left(\hat{\mathbf{m}}^{Multiscale} | \hat{\mathbf{m}}^{obj}\right) p\left(\hat{\mathbf{m}}^{obj}\right) \\ &= \left(\prod_i p\left(\left(\hat{\mathbf{A}}_i^{obj}\right)^{-1} \circ \hat{\mathbf{A}}_i^{Multiscale} \right) \right) \times \\ &\qquad p\left(\hat{\mathbf{m}}^{obj}\right) \quad .\end{aligned} \tag{6}$$

2.3. Robustness measurement of the shape prior

According to Muller [11] the robustness of a probability distribution estimation method can be measured by the fit of test populations to estimated probability distributions, over a variety of training sample sizes. The fit is evaluated by the squared correlation ρ^2 between test cases and their projections in the space estimated from training cases, which reduces to the following formula:

$$\rho^2 = \frac{\sum_{i=1}^{N} d^2(\hat{\mathbf{m}}_{i,test}, \overline{\mathbf{m}}_{train})}{\sum_{i=1}^{N} d^2(\mathbf{m}_{i,test}, \overline{\mathbf{m}}_{train})}, \quad (7)$$

where $\hat{\mathbf{m}}_{i,test}$ is the projection of the test model $m_{i,test}$ onto the shape space and $\overline{\mathbf{m}}_{train}$ is the Fréchet mean of the training sample. The distance is the Euclidean distance for Cartesian space, and it is the geodesic distance for the manifold in our m-rep case. The projected model indicates the closest model to the test model in the trained shape space, according to geodesic distances d in the feature space. The closer ρ^2 is to 1, the better the predictability provided by the shape prior.

2.3.1 Toy experiment on deformed ellipsoids

We evaluated the robustness of the statistics first on a simulated data set of 1000 warped ellipsoids. M-rep models of warped ellipsoids were produced by successively applying random, independent, global bending, tapering, and twisting to a base ellipsoid, followed by a random, small, local perturbation of the hub position for a fixed atom to create a visible protrusion or indentation at that local region of the surface (Fig. 2).

Each transformation was sampled from a zero-mean Gaussian with its respective variance. The accumulated percentages of variation of object eigenmodes are demonstrated in Fig. 2, which shows the distribution of the shape variation among the trained eigenmodes from 20 randomly selected m-rep models. We found that the first three eigenmodes obtained by PGA each captured one of the three global shape changes, namely bending, twisting, and tapering, that were introduced by our simulation. The local surface perturbation that we simulated was captured by the 4th and 5th eigenmodes, as were small variations at other random regions. We attribute these other local variations to the limited number of training samples.

We used the ρ^2 measure to compare two shape priors. We trained one shape prior by using our multiscale approach and retaining three object-scale eigenmodes as well as four atom scale eigenmodes. The other shape prior was trained at the object scale only, and retained three eigenmodes. Twenty tests were run for each training sample size, and there was no overlap between the training and testing populations. and the stars in the figures are the outliers. The

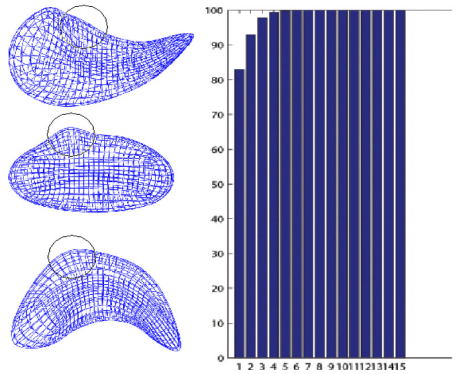

Figure 2. Left: 3 typical training m-reps that were generated by a random mix of three global variations (bending, twisting and tapering) and one local variation (in the circled areas). Right: vs. k, the cumulative percentage of variation explained by the first k principal geodesic eigenmodes of 20 randomly selected m-reps.

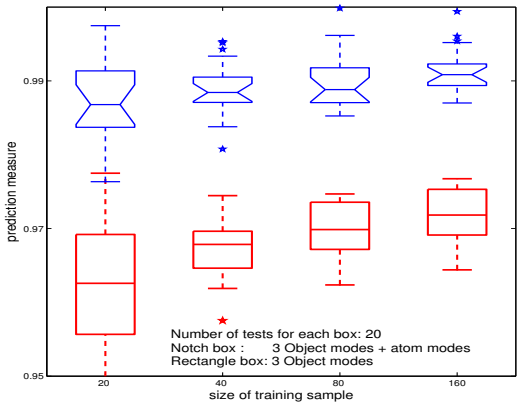

Figure 3. The distribution of the prediction measurement ρ^2 from twenty random training and test populations as a function of training sample sizes are shown for multiscale (notched boxes) and object-scale (rectangular boxes) shape priors.

distribution of ρ^2 as a function of training sample size from these tests is shown in Fig. 3. We can clearly see that our multiscale shape prior offers superior predictability than the object-scale shape prior alone.

From this toy experiment with carefully designed artificial variation, predictability of the two-scale shape prior is shown. It suggests that the atom scale residual modes are able to capture the fairly small amount of local region variation ignored by the object scale modes .

2.3.2 Experiments on simulated hippocampus models

To further investigate the advantages of the large-to-fine-scale prior on anatomical structures, we simulated a large

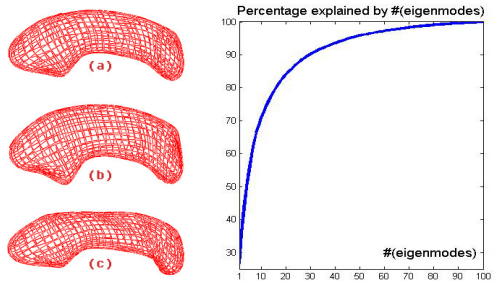

Figure 4. Left: (a) The Fréchet mean of the 290 m-reps. (b) and (c) are randomly selected from the 500 samples using 51 eigenmodes. Right: The accumulated percentages of the variances explained by the eigenmodes trained from the 290 m-reps.

group of hippocampus m-rep models based on 290 hand-segmented MR images. First, the m-rep models were fitted to images were obtained by a semi-automated process [7]. Global PGA was then used to form a Gaussian probability distribution from these models. 51, 33, and 17 eigenmodes respectively capture 95%, 90% and 80% total variation of the 290 models (see Fig. 4). By a Gaussian random sampling on the 51 modes, 500 hippocampus models were generated. Example simulated models are shown in Fig. 4.

Shape priors with different number of eigenmodes were evaluated. We first used ρ^2 to compare the 33 object-modes prior with the multiscale prior composed of 10 object modes and 9 atom modes (Fig. 5, top). For training sample size N, there are at most $N-1$ modes actually available. So in our tests with 20 samples, 19 object modes are used rather than 33 object modes we used in the tests on larger sample sizes. We also calculated Dice coefficients to measure volume overlap between the test models and the corresponding projected models (Fig. 5, bottom). The two measurements consistently demonstrate that adding the atom scale improves the robustness of the estimation.

In medical image segmentation, a 33-dimensional space is too large to search for the optimum shape since robustness is challenged by local minima during optimization. In this example, using only 10 object modes followed by atom modes shows much better predicability. More importantly, much better performance was seen in tests with small training sample size, indicating the superiority of this large-to-fine-scale strategy in HDLSS situations.

More comparisons are shown in Fig. 6. Six shape priors with different combination of modes were tested. To make the comparison easier, the median ρ^2 values among 20 tests are shown. The two-scale priors are much more robust than

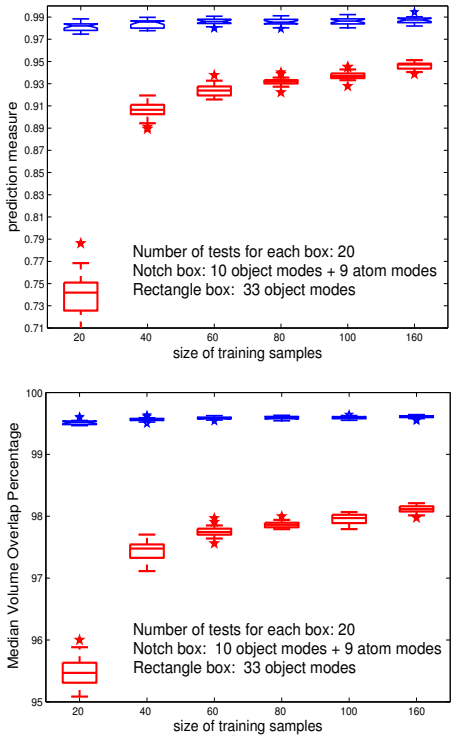

Figure 5. Comparisons on the robustness of the shape priors between using 33 object modes with using 10 object modes followed by atom modes. Top: ρ^2 measurement; Bottom: Volume overlap between the test models and the projected models.

the single-scale object-scale priors overall. As the training sample size increases, the curves converge to different values. For the object-scale priors (in red), predictability increases with the number of modes used. Two-scale priors (in blue) converge much faster as the training sample increases and show their strength with training sample sizes less than 60. No measurements are calcuated at training sample size 20 and 30 for the priors with more eigenmodes. The improvements from the object-scale prior are most noticeable at smaller training sample sizes.

3. Large-to-fine-scale posterior optimization

A large-scale deformation (object stage) followed by a fine-scale deformation (atom stage) can optimize the model hierarchically, as shown in Fig. 7. In each stage, a Bayesian posterior optimization scheme is applied using the off-line trained statistics at the corresponding scale.

In the object stage the entire sheet of medial atoms deform together, restricted by the variation learned from the multiple sheets of atoms of the training samples. In the atom stage, each atom is constrained in its own variation space trained from corresponding atoms of the training samples.

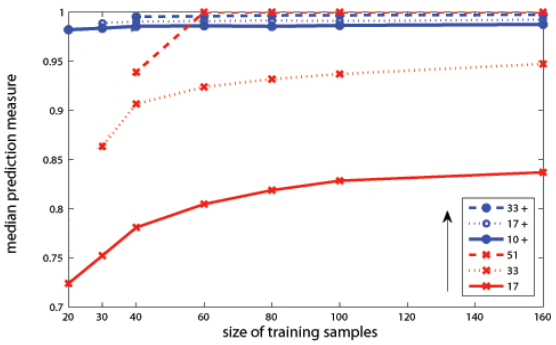

Figure 6. ρ^2 measurements from six different combinations of the number of object-scale modes and atom-scale modes: each marker shows the median ρ^2 from 20 tests. Red curves are the results of using 17, 33, and 51 object-scale modes only. Blue curves are the results of using 10, 17, and 33 object-scale modes followed by 9 atom-scale modes (tagged by"+").

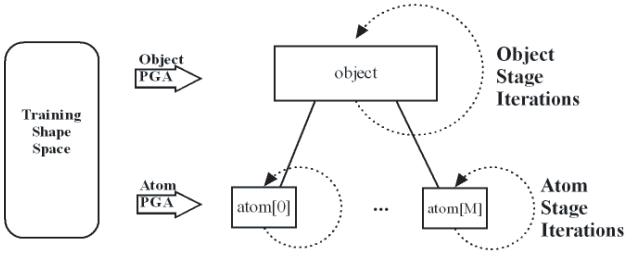

Figure 7. The probabilistic multiscale segmentation scheme for single object m-rep models: The object and atom stage shape statistics are collected from training samples and then used during the optimization process in the corresponding stages.

The object-scale segmentation results provides a good starting point for the local refinement stage, where the model deforms locally towards a more "correct" configuration.

Besides the shape prior training, the likelihood function on image intensity patterns $\mathbf{I}$ is also trained. This likelihood training uses PCA on regional intensity quantile functions (RIQFs). The image regions are defined locally corresponding to the spoke ends of the atoms. The details for the image appearance model can be found in [1]. With the training statistics, at each of the two scales the model deforms by optimizing the reciprocal-standard-deviation-weighted sum of two penalty terms: the log geometric prior $\log p(\mathbf{m})$ and the log likelihood of the m-rep $\log p(\mathbf{I}|\mathbf{m})$. The optimization is over the principal geodesic coefficients of the shape $\mathbf{m}$, thus restricting the result to the shape space spanned by the principal geodesic directions.

The local scale posterior optimization is described by the formula:

$$\begin{aligned}&\arg\max_{\Delta\mathbf{A}_i} \log p\left(\Delta\mathbf{A}_i \circ \mathbf{A}_i | \mathbf{I}_i\right)\\ &= \arg\max_{\Delta\mathbf{A}_i} \left(\log p\left(\mathbf{I}_i | \Delta\mathbf{A}_i \circ \mathbf{A}_i\right) + \log p\left(\Delta\mathbf{A}_i \circ \mathbf{A}_i\right)\right)\end{aligned} \tag{8}$$

where $\Delta\mathbf{A}_i = \mathbf{A}_i^{-1} \circ \mathrm{Exp}_{\mathbf{A}_i}\left(\sum_q \mathbf{x}_q \mathbf{v}_q^{atom}\right)$ is the atom scale variation at $\mathbf{A}_i$ (see equation 5) and is implied by the coefficient vector $\mathbf{x}$. The objective function at each atom, indexed by i, is optimized over the coefficient vector $\mathbf{x}$. Starting from the object-stage result, we randomly loop over all the atoms and update each atom by adding the residual deformation that maximizes the log posterior probability density.

4. CT image segmentation results

We tested the methodology on a data set of CT images of the male pelvis from 5 patients. In the average patient 16 daily CT scans were taken during image-guided radiation therapy. The images have in-plane resolution of 512×512 voxels with pixel dimension of 0.98 mm × 0.98 mm and an inter-slice distance of 3 mm. Manual segmentation by an expert was provided.

For each patient, we carried out two leave-one-day-out bladder segmentation experiments, one at object scale only and the other with our two-scale method. Training was done on all other days when segmenting the target day image. The results were compared with the manual segmentations and measured in terms of average closest surface point distance. Fig. 8, depicts the segmentation results via box plots. The training models are the m-rep models fitted to the binary images from the human segmentation, which can be regarded as the desired segmentation results. The initializations are done by optimizing the model over the object-scale shape space to minimize the surface distance to manually specified points on three designated slices of the image only. This can be regarded as a part of the object stage. After initialization, the model is pretty much correctly located in the image. The object stage then further optimizes the model into a large-scale shape that is a good starting point for the atom stage. The observed improvements in median average surface distance, on the order of $0.1mm$, have been shown to correspond to improvements on the order of $0.5mm$ in 90^{th} percentile (nearly worst) position distances. The decreased interquartile range of the two-scale method indicates improved robustness. The atom scale refinement on average brings the accuracy into the subvoxel level.

We also found that the improvements in the bladder segmentation are most noticeable at parts of the boundary where the contrast is high (Fig. 10). An example of typical (median surface distance) bladder segmentation re-

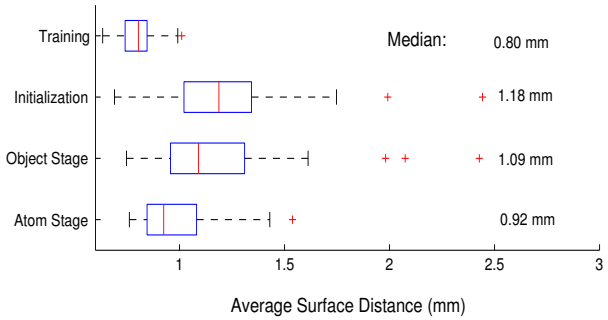

Figure 8. Leave-one-day-out bladder segmentation results for 79 images from 5 patients.

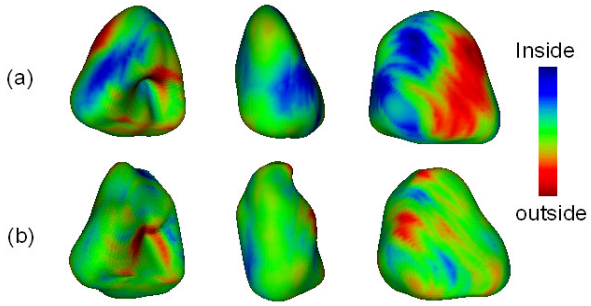

Figure 9. Surface distance errors shown in color from blue (5.5 mm inside) to red (5.8 mm outside) (a) An object-stage result with average surface distance of 1.15 mm. (b) An atom-stage result with average surface distance of 0.92 mm.

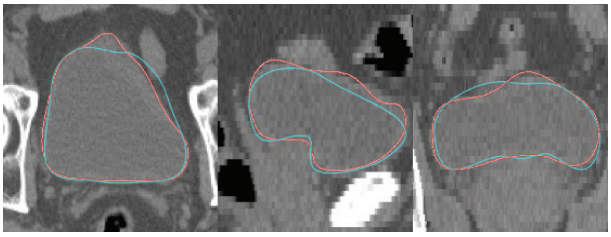

Figure 10. 2D contours comparison: object stage contours (in cyan) and the atom stage contours (in red) in 2D slices from axial (left), coronal (middle) and sagittal (right) views.

sults from the leave-one-day-out study is shown to demonstrate the difference made by the atom stage visually on the boundaries. Figure 9 shows surface distance to the reference segmentation as a color map on the model boundary. Figure 10 shows the model at both stages on three orthogonal gray scale CT image slices. The localized refinement is clearly shown in the three slices.

In addition, we applied the statistical atom stage to a leave-one-patient-out [16] study. For each of the 5 patients in succession, one was chosen as the target patient and the other 4 formed the training sample. We pooled the variation

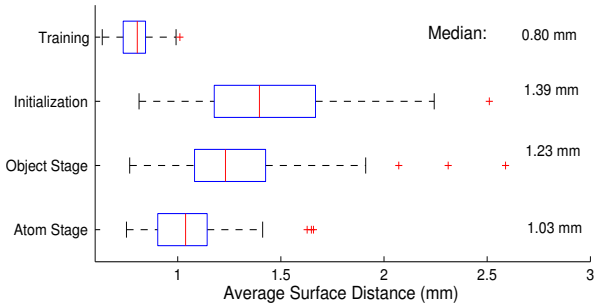

Figure 11. Leave-one-patient-out bladder segmentation results for 79 images from 5 patients.

of the training models relative to their patient-specific mean after a pelvic bone based alignment [16].

At segmentation time, we took the previous days' mean model as the initialization for the next day's image segmentation, with the first day's segmentation being provided beforehand. Compared to the leave-one-day-out study, although the training sample size is about 3 times larger, this experiment was a greater challenge due to its more varied shape space across different patients and the weaker correspondences for the local regions across patients. Nevertheless, as seen in Fig. 11 there is improvement resulting from the atom-scale refinement.

We also applied the strategy to caudate segmentation from MR images, and saw improved results from the two-stage approach. We conclude that the local residual statistics effectively enable the refinement from an object-scale shape. In low contrast regions where even humans have difficulties in locating the boundary, the atom stage tends to avoid deformations of the object-stage result.

5. Conclusions

In this paper, we presented a novel method to build up a large-to-fine-scale shape prior by discovering the fine residual statistics from the large scale hierarchically via m-rep. We showed improved robustness on two large simulated data sets, evaluated with the ρ^2 measure describing the predictiveness of an estimated probabilitiy distribution and volume overlap measurements against different training sample sizes. We applied the large-to-fine-scale shape priors within a multiscale posterior optimization segmentation framework to improve the segmentation accuracy. Our experiments suggest that the residual shape statistics can be used to restrict results to the trained shape space and to penalize the local deformation towards better posterior estimates. Segmentation experiments on 79 CT bladder images demonstrated improved segmentation accuracy by the local scale refinement.

A condition for building up the localized shape priors is the correspondence of the local regions across training samples. Better correspondence generates tighter statistics. Here in m-reps, atoms with the same index correspond to each other geometrically. In other representations without natural correspondence, resampling techniques can be used to set the correspondence at different scales.

For HDLSS problems, the strategy of multiple residual scales has been shown to improve the estimations of probability distributions of shapes. Computationally, optimizing in a high dimensional space is much more likely to yield local optima and less efficient than the successive optimization over far fewer dimensions at each stage provided by the large-to-fine-scale approach.

Besides m-reps, the usage of the local residual statistics is applicable to various representations with notions of localities. More studies need to be carried as to the number of scales and the size of each scale.

Acknowledgments: This work was supported by NIH grant P01 EB02779. We are thankful to Robert E. Broadhurst, Joshua V. Stough, Qiong Han, Gregg Tracton and Graham Gash from MIDAG for conceptual, algorithmic and code contributions.

References

[1] R. E. Broadhurst, J. V. Stough, S. M. Pizer, and E. Chaney. A statistical appearance model based on intensity quantiles. *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 422–425, 2006.

[2] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models-their training and application. *Computer Vision and Image Understanding*, 61:38–59, January 1995.

[3] M. Costa, H. Delingette, and N. Ayache. Automatic segmentation of the bladder using deformable models. *Proc. of IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 904–907, 2007.

[4] C. Davatzikos, X. Tao, and D. Shen. Hierarchical active shape models, using the wavelet transform. *IEEE Transactions on Medical Imaging*, 22(3):414–423, 2003.

[5] P. Fletcher, C. Lu, S. Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, 23:995–1005, 2004.

[6] D. Gibou, F nd levy. Partial differential equations based segmentation for radiotherapy treatment planning. *Mathematical Biosciences and Engineering*, 2(2):209–226, 2005.

[7] Q. Han, D. Merck, J. Levy, C. Villarruel, J. Damon, E. Chaney, and S. Pizer. Geometrically proper models in statistical training. In *Proceedings of Information Processing in Medical Imaging*, pages 751–762, 2007.

[8] S. Joshi, S. M. Pizer, P. T. Fletcher, P. Yushkevich, A. Thall, and J. S. Marron. Multiscale deformable model segmentation and statistical shape analysis using medial descriptions. *IEEE Transactions on Medical Imaging*, 21(5), 2002.

[9] K. Leonard. Efficient shape modeling: -entropy, adaptive coding, and blum's medial axis versus the boundary curve. *International Journal of Computer Vision*, 72(2):183–199, 2007.

[10] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: A survey. *Medical Image Analysis*, 1(2):91–108, 1996.

[11] K. E. Muller, S. Ray, and Y.-Y. Chi. Goodness of prediction for principal components, including high dimension, low sample size. *In preparation*, 2008.

[12] D. Nain, S. Haker, A. Bobick, and A. Tannenbaum. Multiscale 3d shape analysis using spherical wavelets. *Medical Image Computing and Computer-Assisted Intervention(MICCAI) 2005*, 3750:459–467, 2005.

[13] C. Nikou, F. Heitz, J. Armspach, G. Bueno, and D. Vernon. A physically-based statistical deformable model for brain image analysis. *Proceedings of the 6th European Conference on Computer Vision-Part II*, 1843:528–542, 2000.

[14] T. Pasquier, D.and Lacornerie, M. Vermandel, J. Rousseau, E. Lartigau, and N. Betrouni. Automatic segmentation of pelvic structures from magnetic resonance images for prostate cancer radiotherapy. *Int. J. Rad Onc. Biol. Phys.*, 68(2):592–600, 2007.

[15] V. Pekar, T. McNutt, and M. Kaus. Automated model-based organ delineation for radiotherapy planning in prostatic region. *Int. J. Rad Onc. Biol. Phys.*, 60(3):973–980, 2004.

[16] S. M. Pizer, R. E. Broadhurst, J.-Y. Jeong, Q. Han, R. Saboo, J. V. Stough, G. Tracton, and E. L. Chaney. Intra-patient anatomic statistical models for adaptive radiotherapy. *MICCAI Conference: Workshop on From Statistical Atlases to Personalized Models: Understanding Complex Diseases in Populations and Individuals*, pages 43–46, 2006.

[17] S. M. Pizer, E. L. Chaney, R. E. Broadhurst, X. Fang, J.-Y. Jeong, J. Stough, and G. Tracton. Segmentation of kidneys and pelvic organs from ct by posterior optimization of m-reps. *techreport*, 2006.

[18] S. M. Pizer, P. T. Fletcher, and S. C. e. a. Joshi. Deformable m-reps for 3d medical image segmentation. *International Journal of Computer Vision*, 55(2-3):85–106, 2003.

[19] M. Rousson, A. Khamene, and M. e. a. Diallo. Constrained surface evolutions for prostate and bladder segmentation in ct images. *CVBIA*, pages 251–260, 2005.

[20] M. Styner and G. Gerig. Three-dimensional medial shape representation incorporating object variability. *Computer Vision and Pattern Recognition*, pages 651–656, 2001.