

Refined Segmentation in Statistical Multiscale Framework

Xiaoxiao Liu, Robert E. Broadhurst, Qiong Han, Joshua H. Levy,
Joshua V. Stough, Rohit Saboo, Ja-Yeon Jeong,
Edward L. Chaney, Stephen M. Pizer

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

Abstract. Based on an effective statistical segmentation methodology using a deformable medial model, a local scale deformation approach is developed to refine the global scale segmentation results within a multi-scale framework. In the local scale segmentation, the probabilistic variations of locally aligned shape residuals from the global scale are learned from proper training followed by a posterior probability optimization in local regions. The resulting finer scale deformation improves the accuracy of the segmentation results, shown by experimental study on 3D CT images of the male pelvic area in day-to-day adaptive radiotherapy.

1 Introduction

3D medical image segmentation via deformable objects is challenging; adoption occurs only if its accuracy is competitive with manual segmentations. Deformable models with various representations have been proven to be effective in capturing a population of geometric entities [1]. An off-line learning process can build up the constrained variation space for deformation [2] [3] [4] [5] [6]. Considering the various degrees of locality for geometric features, an efficient way of representing the shape variance is to decompose them into different scales of localities [6] [7] [8] [9]. In this paper, we describe an approach to produce stable estimates of probability distributions for geometry based on medial representations and for geometry-relative intensity patterns based on regional intensity quantile descriptors.

We focus on the local refinement of the segmentation results from a global-scale deformation as an extension of the methods developed by Pizer et al [10] [11] [6]. Our methodology is based on the Bayesian segmentation method using a medial representation called *m-reps* [2], which explicitly describes both interior and boundary localities. An m-rep, made up of a grid of medial atom primitives, gives geometric properties, such as the widening, bending and tapering, with a locality given according to the grid spacing. The m-rep as a whole gives these properties in a way reflecting global interrelations of primitives, whereas each

primitive, and its relation to its immediate neighbors, gives more local information. Based on the multiscale shape description ability, coarse-to-fine strategy provides efficient segmentation [6].

We have shown the effectiveness of statistical m-rep segmentation in the global scale [12], which provides a good starting point for the refinement stage. Then the more local region-based scale segmentation will guide the model boundaries to deform further from the globally segmented result towards the “correct” locations. Similar to the global segmentation, in the finer scale, we apply a posterior optimization to improve the segmentation results iteration by iteration. But different from global scale, the shape priors in the local scale are residuals from global scale, and image match likelihoods are calculated on the local regions.

Section 2 reviews methods using a multiscale framework (§2.1 and §2.2), and the background for the global scale segmentations using m-reps (§ 2.3). Methodology for the local segmentation is discussed in §3; §3.1 describes the Principal Geodesic Analysis (PGA) training of the residual shape prior, and §3.2 describes the posterior probability optimization. Section 4 gives experimental results on 5 sets of clinical CT image data of the male pelvic area.

2 Background

2.1 Multiscale framework

The multiscale idea has been widely adopted in various segmentation methods [6] [7] [8] [9] [13]. Several of them apply scales to image intensity features; others apply them to model features. A coarse-to-fine multi-resolution protocol on landmarks is used by Christensen et al [13] for diffeomorphic mapping of brain anatomy. Krissian et al [7] developed multiscale detection of the vessel centerlines based on a cylindrical model, with the radius of the cross sections providing the scaling feature. The Active Shape Model (ASM) developed by Cootes et al [9] built multi-resolution image profiles around each landmark for boundary refinement. In m-rep segmentation the scales are localities of the model representation, as shown in Fig. 1, and a template model, which is the geodesic mean calculated from the training models, deforms in a series coarse-to-fine scales.

2.2 Probabilistic multiscale framework

The multiscale methods, ASM and the global scale of m-rep segmentation, take advantage of the statistics from learning. In [6] deterministic, geometrical penalties rather than log probability penalties are used at each scale, and the optimization is carried out over the whole space of m-rep atom transformations, including translation of the hub position, rotation of the spokes and scaling for the spoke length. Thus the optimization is not efficient by explicitly searching in all dimensions and not robust due to the possibility of being trapped at local minima. Lu et al [14] first proposed the probabilistic atom scale as one step of multiscale framework for multi-object segmentation theoretically, based on



Fig. 1. A single figure M-rep model: A medial atom (left) and a figure (right). The figure is composed of multiple atoms.

Markov Random Field (MRF) assumptions. To obtain both mathematical correctness and tight, thus stably trainable probability distributions, we provide the atom scale with a local alignment of the shape residuals.

Since the m-rep and thus the shape residual lives in a curved space (cf. §3.1), the alignment issue becomes crucial. In this paper, in order to focus on the fine scale optimization, we only discuss single object cases.

2.3 Global scale segmentation

In m-rep models a figure is a sheet of medial atoms represented by a quadrilateral mesh. Here we call the global segmentation the figure stage, followed by an atom stage, which is at a local scale. In the figure stage the entire sheet of medial atoms deform together, restricted by the variations learned from the multiple sheets of atoms of the training samples. PGA [15] is the method for learning. We compute a mean model from the training m-rep models and use the first several principal modes, which are able to cover more than 90% total variations of the training samples, as the deformation modes during optimization. PGA largely reduces the dimension of the variation space from nearly three hundred to just 8 dimensions or less. While deforming in the constrained space, an image match measurement computed by Regional Intensity Quantile Functions (RIQFs) on the tuple of subregions forming the region near the implied boundary of the m-rep [16] evaluates how well the model fits into the grayscale image, which leads the optimization process into convergence. Together, the figure stage deforms the model into a global fitness with the image. However, not all of the small variance modes are noise; some of them surely contribute to small deviations from the ideal segmentation in specific regions. The overall quality of the figure stage can be seen from Fig. 2 and Fig. 4. To improve the segmentation results as the needs of clinical application, more local deformations should be enabled in the next finer scale.

3 Method

In the atom stage each medial atom deforms individually within a residual variance space learned from training that is individual for each atom. The residual here is the atom change needed to improve the figure stage result to the ground truth. Similar to the figure stage, after alignment appropriate to the relevant scale we use PGA to describe the residual shape space. RIQFs are used for the image match measurement, but locally on regions about the spokes of the individual atoms.

3.1 Training for Atom Stage Prior

The trained residual deformation space for an atom and the prior is captured by PGA at each atom from its shape residuals from the figure stage. The atom resides in the symmetric space $\mathbf{G} = \mathbb{R}^3 \times \mathbb{R}^+ \times \mathbf{S}^2 \times \mathbf{S}^2$. Given two medial atoms $\mathbf{A}_i, \mathbf{A}'_i \in \mathbf{G}$, $\Delta\mathbf{A}_i$, the residual from \mathbf{A}'_i to \mathbf{A}_i , is calculated by the subtraction operator .

$$\ominus : \mathbf{G} \times \mathbf{G} \rightarrow \mathbf{G}, \Delta\mathbf{A}_i \doteq \mathbf{A}_i \ominus \mathbf{A}'_i \doteq (\mathbf{p}_i - \mathbf{p}'_i, \mathbf{r}_i/\mathbf{r}'_i, \mathbf{R}_{\mathbf{u}'_i}(\mathbf{u}_i), \mathbf{R}_{\mathbf{v}'_i}(\mathbf{v}_i)). \quad (1)$$

where $\mathbf{A}_i = (\mathbf{p}_i, \mathbf{r}_i, \mathbf{u}_i, \mathbf{v}_i)$ with the hub position $\mathbf{p}_i \in \mathbf{R}^3$, the spoke length $\mathbf{r}_i \in \mathbf{R}^+$, and two unit spoke directions $\mathbf{u}, \mathbf{v} \in \mathbf{S}^2$. $\mathbf{R}_{\mathbf{w}}$ represents the rotation along the geodesics in \mathbf{S}^2 that moves a point $\mathbf{w} \in \mathbf{S}^2$ to the North Pole $(0, 0, 1) \in \mathbf{S}^2$. The addition operator is correspondingly defined as

$$\oplus : \mathbf{G} \times \mathbf{G} \rightarrow \mathbf{G}, \mathbf{A}'_i \oplus \Delta\mathbf{A}_i \doteq (\mathbf{p}'_i + \Delta\mathbf{p}_i, \mathbf{r}'_i \cdot \Delta\mathbf{r}_i, \mathbf{R}_{\mathbf{u}'_i}^{-1}(\Delta\mathbf{u}_i), \mathbf{R}_{\mathbf{v}'_i}^{-1}(\Delta\mathbf{v}_i)). \quad (2)$$

We want to describe $p(\Delta\mathbf{A}_i | \Delta\mathbf{A}_{\mathbf{M}\{i\}})$, where $\mathbf{A}_{\mathbf{M}\{i\}}$ contains all the other atoms in the medial sheet. We make an assumption here that the large-range dependency among all the atoms on the medial sheet has been explicitly covered by the figure stage, so that after alignment relative to neighbor atoms, the atom stage deformations are localized and can be expressed by the following MRF model [14]

$$p(\Delta\mathbf{A}_i | \Delta\mathbf{A}_{\mathbf{M}\{i\}}) = p(\Delta\mathbf{A}_i | \Delta\mathbf{A}_{\mathbf{N}\{i\}}), i \in \mathbf{M}. \quad (3)$$

where $\mathbf{A}_{\mathbf{N}\{i\}}$ contains the neighbor atoms for atom i , \mathbf{M} is the collection of all atoms within the object. We simplify this conditional probability by decomposing it into two parts: the probability of self driven deformation $p(\Delta\mathbf{A}_i)$, and the neighbor driven deformation $p(\Delta\mathbf{A}_i^{ngbr})$. Further, $p(\Delta\mathbf{A}_i^{ngbr})$ can be approximated by the residual of the interpolated atom, which is the geodesic mean of its neighbor atoms. In our male pelvic organ segmentations, the residuals of the interpolated atom is shown small enough to be ignored. Then the final residual probability becomes $p(\Delta\mathbf{A}_i)$, as shown in (4)

$$p(\Delta\mathbf{A}_i | \Delta\mathbf{A}_{\mathbf{N}\{i\}}) \approx p(\Delta\mathbf{A}_i \oplus \Delta\mathbf{A}_i^{ngbr}) \approx p(\Delta\mathbf{A}_i). \quad (4)$$

For the residual term $\Delta\mathbf{A}_i$, atom \mathbf{A}_i is in the m-rep model fitted to the hand-segmented image, and atom \mathbf{A}'_i is in the figure stage resulting model, after both have been locally aligned according to its neighbors' configurations. We use PGA to train the shape prior $p(\Delta\mathbf{A}_i)$ for each atom from the sampling of ground truth models and their corresponding figure stage results. For each atom, our feature space has 8 dimensions, only the first several 2-5 principal modes which cover the most variance are used in the atom stage deformation.

It is important to align the residuals across cases by their neighbors before we do PGA on them, in order to guarantee mathematical correctness and the tightness of the statistics. To align the residuals, the subtrahend and minuend atom sets should be aligned first. One possibility is to align both the operands into an atlas model before the subtraction in symmetric space. The alternative is to align the operands to each other. And for each segmentation iteration, the current target atom needs to be aligned to the same model you used for its training. Either possibility would use a Procrustes alignment based on geodesic distances between m-reps. In this High Dimensional Low Sample Size (HDLSS) situation, we found that alignments at multiple times introduce noise into the prior, and considering the sensitivity of the residual statistics to alignments, we align the atom in ground truth model to corresponding figure stage model and then do the subtraction to get $\Delta\mathbf{A}_i$. Therefore, in each iteration during the segmentation, we need also align our resulting atom to the figure stage model via its neighbors to be able to use the PGA statistics.

3.2 Atom stage posterior optimization

In the Bayesian optimization process of the atom stage deformation, we compute $\arg \max_{\Delta\mathbf{A}_i} (\log p((\mathbf{A}_i \oplus \Delta\mathbf{A}_i) | \mathbf{I}_i) = \arg \max_{\Delta\mathbf{A}_i} [\log p(\mathbf{I}_i | \mathbf{A}_i \oplus \Delta\mathbf{A}_i) + \log p(\Delta\mathbf{A}_i)]$, where \mathbf{I}_i is the intensity distribution in the local region around the atom and the target model is updated by "adding" up shape residuals. The first term, image likelihood term, is evaluated by RIQF principal mode coefficients on geometric local regions corresponding to the spoke ends of the atom, which probabilistically represents the appearance of a region in an image. The details of the image statistical method can be found in [16]. Given the figure stage results, we randomly loop over all the atoms and update each atom by adding the residual deformation which gives the best log posterior probability value. To be consistent with our training, before each addition with the residuals from the residual PGA space, we locally align the atom to the figure stage result by the configuration of its neighbors. Instead of transforming the current atom first by alignment and then inverse alignment, we apply the inverse alignment transformation to the residual primitive and directly add it to the current atom. Usually it takes 1-4 iterations for the optimization to converge. In each iteration, if one atom gets updated, all its neighbors will be updated later due to the new neighbor relationship.

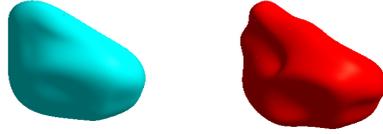


Fig. 2. Comparison of figure stage result (left) and the atom stage result (right): Rendered surfaces of a particular bladder case. Notice the local surface variations of the atom stage results with the smoothness property kept.

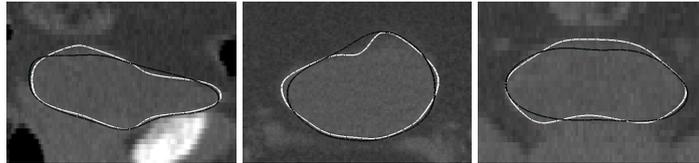


Fig. 3. Comparison of figure stage result (black) and the atom stage result (white): Sagittal(left), Axial (Middle) and Coronal (Right) slices through bladder CT data set. Notice the improvements on the higher contrast regions.

4 Experimental Results and Discussion

We tested the multiscale methodology on five patient image sets, each of on-average 16 daily CT scans of the male pelvic area taken during the radiotherapy. Each image has an in-plane resolution of 512×512 with voxel dimension of $0.98 \text{ mm} \times 0.98 \text{ mm}$ and an inter-slice distance of 3 mm . Expert manual segmentation was provided. For each patient dataset, we carried out a leave-one-day-out study; namely, we had training done on all other days (approximately 15) when segmenting the target day. As the evaluation metric we use volume overlap, which is the intersection over average with the hand segmented binary image, and average surface distance, which is the average distance between two surfaces.

An example of typical bladder segmentation results are shown to demonstrate how the atom stage performs. Fig. 2 shows the 3D surface of the m-rep models of the figure stage result (left) and the atom stage result (right), and Fig. 3 shows the three orthogonal planes through the models with the corresponding slices of gray scale CT image as the background. In this case the average surface distance is reduced from 1.54 mm at the figure stage to 1.15 mm after the atom stage, and volume overlap is increased from 91.3% to 93.8% .

Fig. 4 depicts the bladder segmentation results of 80 images from 5 different patients. We pool all the cases together by sorting the evaluation metrics to show the trend of the improvement regardless of its specificity of a certain patient or day, without strict correspondence of the figure stage and the atom stage for the same case. As we can see from the figure, the overall curve is improved towards a better fit. Even for some unsatisfying figure stage results, with volume overlap

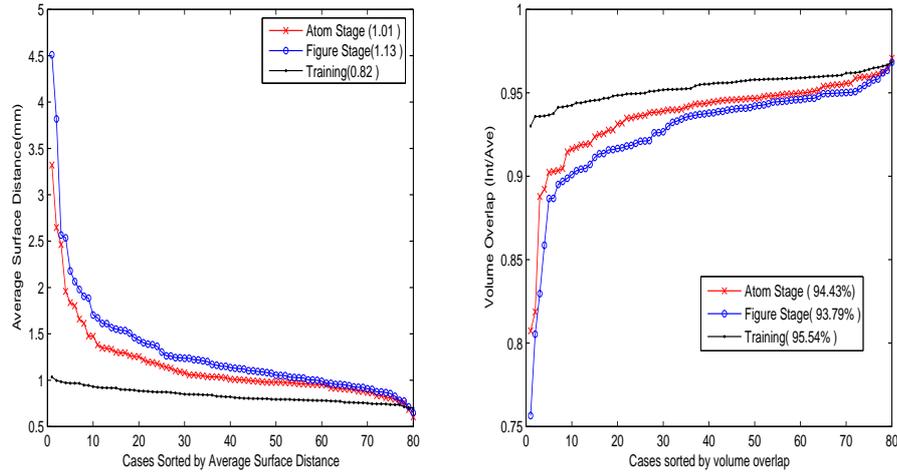


Fig. 4. Bladder segmentation results for the 80 pooled CT images, measured by average surface distance and volume overlap. The curves give the trend of the overall performance. The curve labeled training measures the fit of the training m-rep to the binary image given the “correct” answer and thus provides the best possible segmentation that could be achieved by a m-rep. Notice the improvement from the figure stage to the atom stage.

under 90% and average surface distance above 2 mm, the atom stage is able to make some improvements.

There are several cases where the atom stage gives no improvements, partly because the figure stage has already done a good job, with volume overlap larger than 94% and average surface distance less than 1 mm, leaving little room for the refinement. When part of the figure stage result is some cm from correct, typically due to a poor initialization, the atom stage makes the results little better and sometimes even a little worse, with volume overlap under 85% and average surface distance above 2.5 mm. We are working to improve these results by developing better initialization methods.

The improvements in bladder segmentation are most noticeable at parts of the boundary where the contrast is high. Here the atom stage has moved the segmentation from rather good to completely acceptable for clinical use.

Our preliminary atom stage segmentations on the prostate did not show improvements from the figure stage. There are several reasons. First, almost all around the prostate boundary the contrast is low. Second, the atom alignment in the prostate model is less accurate than the bladder because its medial sheet is typically heavily bent. Additionally, the figure stage results are in many cases more satisfying than for the bladder, leaving little room for the refinement.

Further study on more accurate alignments will be carried out to increase the accuracy of the atom stage prostate segmentations.

5 Conclusion

We have provided a method to refine segmentations after the global scale in a probabilistic multiscale framework. Priors generated via local alignments on the shape residuals and image likelihoods computed from local RIQFs are applied to a posterior probability optimization for the local scale segmentation. Experiments on 80 CT images have demonstrated that more accurate fits to the objects are achieved by the small scale deformations.

Our preliminary results suggest that the residual statistics are effective to drive the local primitives to deform towards better posterior estimates. A robust figure stage is desired to produce stable statistics for the atom stage, especially for the HDLSS problem in our radiotherapy datasets. Also, accurate alignment is important to obtain correct and tight residual statistics for segmentation.

From the point of view of clinical use on our adaptive radiotherapy target problem, our multiscale method for inter-day segmentation within a patient must be tested with variability training from inter-day variations on other patients, a strategy that has been shown to work at the figure stage [17]. Also, a more reliable alignment method should lead to better local scale segmentation. Finally, the refinement method has only been tested on intra-patient datasets. It will be interesting to try it on inter-patient data sets as well. Larger training samples will produce more robust statistics but these will be needed to handle the larger inter-patient variations.

References

1. McInerney, T., Terzopoulos, D.: Deformable models in medical image analysis: A survey. *Medical Image Analysis* **1**(2) (1996) 91–108
2. (Anon)
3. Cootes, T., Taylor, C. J. and Cooper, D.H., Graham, J.: Active shape models-their training and application. *CVGIP: Image Understand* **61**(1) (1994) 38–59
4. Kervrann, C., Heitz, F.: Statistical deformable model-based segmentation of image motion. *IEEE transaction on image processing* **8**(4) (1999)
5. Nikou, C., Heitz, F., Armspach, J., Bueno, G., Vernon, D.: A physically-based statistical deformable model for brain image analysis. *Proceedings of the 6th European Conference on Computer Vision-Part II* **1843** (2000) 528–542
6. (anon)
7. Krissian, K., Ellsmere, J., Vosburgh, K., Kikinis, R., Westin, C.F.: Multiscale segmentation of the aorta in 3d ultrasound images. *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)* (2003) 638–641
8. Park, J.Y., McInerney, T., Terzopoulos, D., Kim, M.H.: A multiscale deformable model for extracting complex surfaces from volume images. *Seventh Pacific Conference on Computer Graphics and Applications (PG'99)* (1999) 208

9. Cootes, T.F., Taylor, C., Lanitis, A.: Active shape models: Evaluation of a multiresolution method for improving image search. 5 th British Machine Vision Conference (1994) 327–336
10. (Anon)
11. Christensen, G.E., Rabbitt, R.D., Miller, M.I.: Deformable templates using large deformation kinematics. IEEE Transactions on Image Processing **5**(10) (1996) 1435–1447
12. (Anon)
13. Christensen, G., Joshi, S., Miller, M.: Volumetric transformation of brain anatomy. IEEE Transactions on Medical Imaging **16**(6) (1997) 864–877
14. (Anon)
15. (Anon)
16. (Anon)
17. (anon)