

Segmentation of Single-Figure Objects by Deformable M-reps

Stephen M. Pizer, Sarang Joshi, P. Thomas Fletcher, Martin Styner,
Gregg Tracton, James Z. Chen

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

Abstract. This paper describes the basis and behavior of segmentation of single figures in 3D by deformable m-reps models. Results are given for the segmentation of kidneys from CT and of hippocampi from MR images. Special focus is made on multi-scale-level stages of segmentation, on intrinsic correspondences under deformation that are provided by m-reps, and on the match against model-relative templates provided by both theoretical edge strength templates and templates derived from training images.

1. Introduction

A variety of authors have described methods of segmentation of anatomic objects from medical images by the single scale deformation of boundary models [Casseles 1997, Cootes 1999, Montagnat & Delingette 1998, Kelemen 1999, Staib 1996] In [Joshi 2001] we have described a method of multi-scale segmentation by deformation of models using an m-reps representation. The focus of this paper is not the method itself but its basis and results. In this section we sketch the method, for the restricted case where the object to be found can be well modeled by a single medial mesh, i.e., as a single *figure*, leaving the details to [Joshi 2001]. In section 2 we discuss the theoretical advantages of m-reps models for deformable model segmentation. Sections 3 and 4 cover the areas of emphasis of this paper, with section 3 focusing on one particular advantage of m-reps, that of providing an object-intrinsic coordinate system to give positional, orientational, and scale correspondence between deformed versions of an object. We explain how this correspondence is used in both the geometric typicality term and geometry to image match term in the objective function being optimized to accomplish the model deformation. Section 4 discusses the results of this segmentation approach on the segmentation of two single-figure objects: the union of the kidney parenchyma and renal pelvis as it appears in CT images and the hippocampus as it appears in MR images. Quantitative comparisons between the results and those of manual segmentation are given for the kidney. Because this segmentation method is only a step along the way to the final method that we anticipate, section 5 discusses directions in which anticipated further developments will be made.

The method. The deformable m-reps method operates from large to small scale levels, at each level deforming the represented object \underline{m} by optimizing an objective function $F(\underline{m}, I_{\text{target}})$ over the set of geometric transformations available at that scale level. As with many deformable model based segmentation methods, the objective function F is the sum of two terms, one measuring the geometric typicality of \underline{m} and

the other measuring the match of $\underline{\mathbf{m}}$ to the target image I_{target} . The algorithm for single figure objects is as follows. At each stage of this algorithm the geometric typicality measures deviation from the deformed model that is the result of the previous stage.

Algorithm

- 1a. Manually place the model in the 3D image, thereby choosing a similarity transform
- 1b. Find and apply the similarity transform which optimizes $F(\underline{\mathbf{m}}, I_{\text{target}})$
2. Until convergence, do
 - {For each medial atom in $\underline{\mathbf{m}}$ {Transform the atom to optimize $F(\underline{\mathbf{m}}, I_{\text{target}})$ }
3. For each boundary tile implied by $\underline{\mathbf{m}}$
 - {Shift the position of the tile along the tile's normal to optimize $F(\underline{\mathbf{m}}, I_{\text{target}})$ }

The initial models $\underline{\mathbf{m}}_0$ used in this work (Figs. 1 & 4) were developed in one of two ways:

By analysis of the geometry of a training set of hand segmented instances of the object over a variety of patients. This automatic analysis uses a method described in [Styner 2001].

By manual construction on a single training image according to set of rules determined by the mathematics of medial geometry [Pizer et al 2001].

Space limitations do not permit us to detail these model building methods here.

2. Theoretical Advantages of M-reps Based Segmentation

We desire successful segmentation performance that is linear in the number of the smallest scale geometric primitives, for example the boundary tiles defining the segmented object's surface or the voxels making up the object. We argue elsewhere that a) such behavior is achievable only by multi-scale-level segmentation with rather closely spaced scale levels, and b) at each level the diameter of the area or volume summarized by the geometric primitives (*atoms*) at that level and the distance of communication used at each geometric transformation of an atom or group of atoms are comparable. Either of these two distances can be taken as the measure of the scale at that level.

In our method the scale levels are indicated by the numbered steps in the algorithm above: [1] the figural scale levels, [2] the medial atom or figural section scale level, [3] the boundary atom scale level. In steps 1 and 2 the object is represented by medial atoms, and at step 3 it is represented by boundary atoms. A representation of a figure by medial atoms is called an *m-rep*.

An m-rep for a generic figure in our system is a quad-mesh of medial atoms (Fig. 1), where an interior medial atom is a medial position at which two vectors (called *port* and *starboard sails*) of equal length r share a sail and a mesh-edge medial atom in addition is equipped with a bisector vector of the two sails of length greater or equal to the common sail length. To allow shape representation and thus magnification invariance not only globally but locally, the spacing of the atoms in an m-rep is approximately proportional to the sail lengths r of the atom.

An m-rep represents a continuous 2-manifold of medial atoms. Thall and Yushkevich have developed two methods for interpolation of this manifold from the m-rep that are consistent with first order medial geometry [Fletcher et al. 2001]. The method of Thall is used in the results described here. The manifold of medial atoms defines a continuous m-rep implied boundary obtained from the union of all the sail-end positions, together with an interpolation of the crest between the sail tips and through the bisector vector tips of each mesh-edge atom. At each such boundary position the corresponding medial atom sail is normal to the implied boundary.

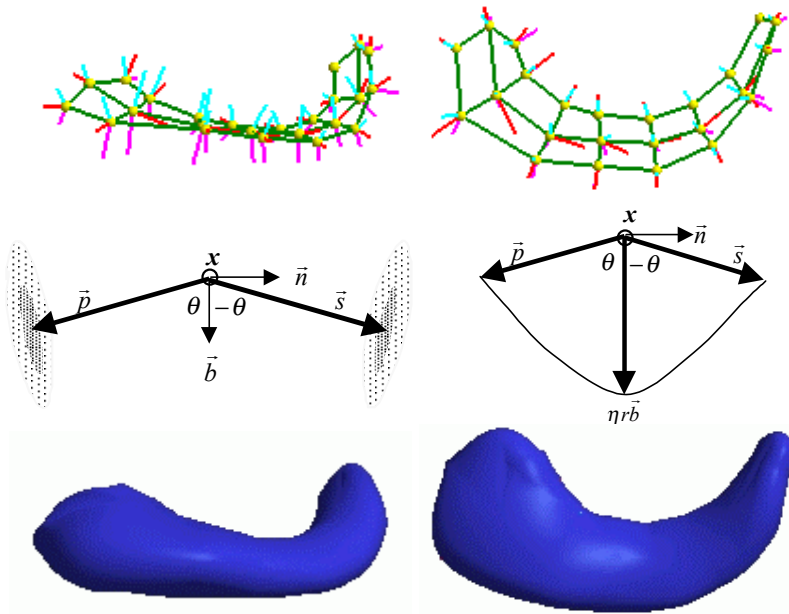


Figure 1. Top: an m-rep for a hippocampus, viewed from two directions. Each ball with two line segment sails forms a medial atom. Center: an internal medial atom and a mesh-edge medial atom, each with their implied boundaries. Bottom: The boundary implied by the m-rep, viewed from two directions.

M-reps have four advantages for representing objects to be segmented by deformable models.

1. Each medial atom represents an interior section of a figure, leading to a special capability for deformation of the interior.
2. Since figures typically have anatomic names and each medial atom in the m-rep corresponds to an interior slab of the figure bounded by its immediately neighboring medial atoms, the geometric transformations involved in the deformation can be described with medical relevance and with appropriate locality.
3. An m-rep lends itself directly to representation at multiple scale levels. There are additional important scale levels besides those already mentioned (figural, figural section, boundary). For objects consisting of more than one figure or for sets of objects, situations not described in this paper, there are the larger scale multi-

object and object scale levels. Also described elsewhere [Yushkevich 2001], there are opportunities for multiple levels of meshing of each figure.

- Figures yield an intrinsic coordinate system with space varying frame and distance metric. Indeed, multi-figure complexes also yield an intrinsic coordinate system, not described in this paper due to lack of space. As described in section 3, this leads to correspondences under deformation that are important with deformable models.

3. Intrinsic Figural Coordinates and Their Use in Geometric Typicality and in Geometry to Image Match

Geometric typicality functions measure the closeness of a deformed model for a figure to a mean or most typical form of the figure. Since sensing of the figure is typically at the boundary, it is natural for this geometric typicality measure to involve the distances of corresponding boundary points between the deformed state of the figure and the typical state (Fig. 2). In a medial geometry, allowing shape to be characterized at all levels of scale (locality), the distances must be taken in multiples of r (the length of the relevant medial sail). In our method as it presently stands, the geometric typicality is measured by the mean squared r -proportional offset of the boundary. This mean is taken over the section of boundary appropriate for the present level of scale, and it is measured between the version of the deformed model produced at the next larger level of scale and the newly deformed candidate $\underline{\mathbf{m}}$.

As is common, our method measures the match of $\underline{\mathbf{m}}$ to the target image in a region near the object boundary that we call a boundary *collar* (Fig. 2). A fruitful way of looking at matching $\underline{\mathbf{m}}$ to I_{target} is that at corresponding positions in the collar, before and after deformation of the model into $\underline{\mathbf{m}}$, a template defined on the model must match as closely as possible a template image defined with respect to the model. The template can be an ideal image, e.g., defined by directional derivatives of a Gaussian, the training image from which the model was built, or the statistics of a set of training images from which the model was built.

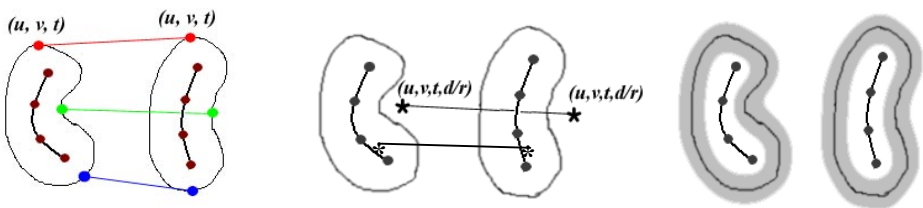


Fig. 2. Medially implied correspondences between a typical figure and a deformed figure for figural boundary positions (leftmost), for positions interior and exterior to the boundary (center), and for the boundary collar (right).

The boundary and collar position correspondences required by geometric typicality and the image match, respectively, are well provided by the medially based intrinsic coordinate system for a figure, as follows. Let the two dimensions of the quad mesh

forming the m -rep be called u and v . Let the atom positions be taken as integer values of u and v . Let the interpolated medial sheet be parametrized by its (u,v) "figural" coordinate system in which distances are r -proportional along the medial manifold.

The medially implied boundary is parametrized via the medial coordinates, but in addition a parameter t is needed to select the side of the medial sheet. We let $t=1$ for boundary points touched by port sails and let $t=-1$ for boundary points touched by starboard sails. At the crest, where the boundary switches from the port side to the starboard side, t varies smoothly from $+1$ to -1 such that $t=0$ at the crest. Then every boundary point is parametrized by its figural coordinates (u,v,t) . Each boundary point thereby carries a normal in the direction of the sail abutting there and a ruler $r(u,v)$. The (u,v,t) values are used to produce correspondences for boundary points to measure the geometric typicality between the state before and that after deformation and are also used in the boundary displacement final stage of the segmentation.

Points inside the figure and outside it but inside the caustic surface can also be put into correspondence in a figurally relative manner. Correspondences outside the caustic surface have also been defined [Crouch 2001], but this is beyond the scope of this paper. In the medial framework distances are measured in an r -proportional fashion along the sails, i.e., along boundary normals. Thus if d is the Euclidean distance from the medially implied boundary to a point in space, with points interior to the figure having negative distances and points exterior to the figure having positive distances, $(u,v,t,d/r)$ provides a shape-respecting figural coordinate for a point. Correspondences between the collars of a figure and its deformed version used in computing geometry to image match are then done according to equal values of these figural coordinates for space. More precisely, at model building time a boundary sampling defined by equally spaced samples of u , v , and t is determined and an equally spaced sampling of d/r between $-k$ and $+k$ is specified ($k = 0.33$ is a typical value). These sample positions for the part of the boundary that can be shifted at the respective level of scale are used in producing the geometry to image match measure.

Many segmentation systems use directional derivatives of a Gaussian at some scale as a measure of contrast, which is expected to be high at a boundary. This is equivalent to correlation with a derivative of Gaussian template in the normal direction at each boundary position of the object. This suggests that a correlation method with a template related by figural correspondence to the image(s) on which the model is based is advisable. Other measures of template match, such as normalized mutual information are possible [Willis 2001].

A different template is preferable for images in which the object appears at low contrast at some boundary positions, the polarity or other form of the contrast changes along the boundary (Fig. 3), or either object or the background region in some portions of the boundary is quite thin. In these cases, which are more common than not, a template made from a training image or the mean of a set of training images is quite attractive. Such a template can avoid having the deformation be attracted by a high contrast nearby boundary in a region along which the object sought bounds an object with similar intensity and thus is known to provide no contrast.

We have implemented such a training image template for collars of half-width $r/3$. For certain kidneys they provide improved segmentation, and for the hippocampus they are essential.

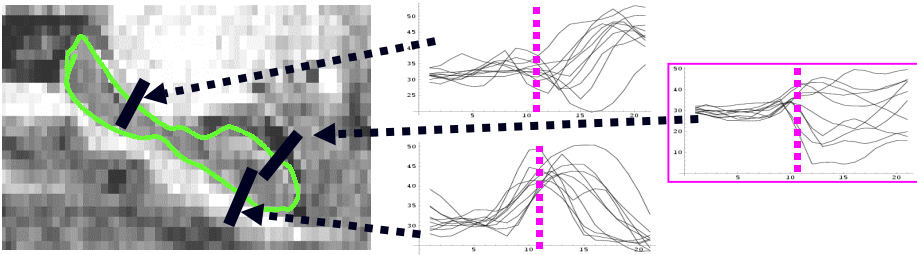


Fig 3. Intensity profiles for a set of patient's normal to the boundary at three separate locations. Compliments of G. Gerig.

4. Results on Single Figure Objects

We have tested this method for the extraction of three anatomic objects well modeled by a single figure: the lateral cerebral ventricle, the kidney parenchyma + pelvis, and the hippocampus. Extracting the lateral ventricle from MR images is not very challenging because the ventricle appears with high contrast, so our successful results are not shown. Extracting the kidney from CT has some challenge because in certain sections of the kidney, where it abuts the liver, there is essentially no contrast and the high contrast spine is nearby. Results of a kidney segmentation are visualized in Figs. 4 & 5. As laid out in Table 1, for 30 kidneys (18 right kidneys and 12 left kidneys, with the model having been built on a right kidney) we were able to extract the kidney with an median accuracy of boundary position of 1 voxel (2.0 mm) as compared to a human manual segmentation. All of the measurements in Table 1 are made relative to a human segmentation that classified each voxel as in or not in the object and did not take anatomic understanding well into account when segmenting the renal pelvis from the ureter and other background. Moreover, the measurement tool measures offsets and overlaps only to the closest voxel. Therefore, our real median boundary accuracy is subvoxel, and the overlap percentatges are understated. Human to human agreement is of the same order.

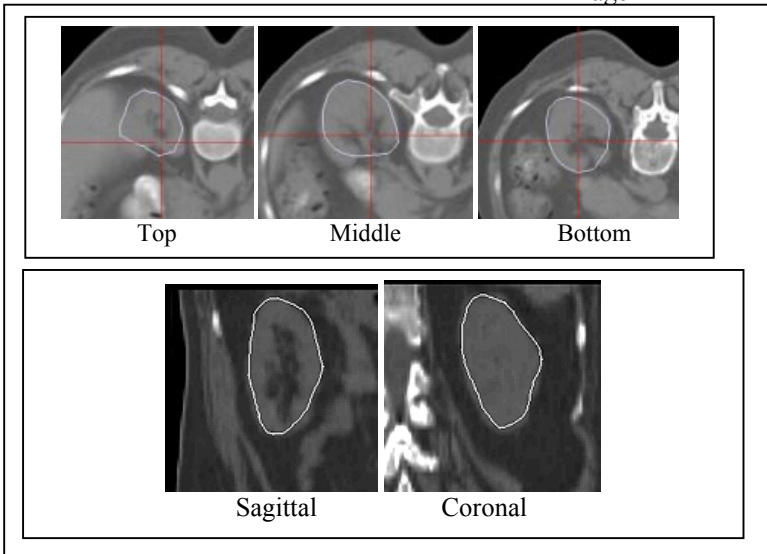
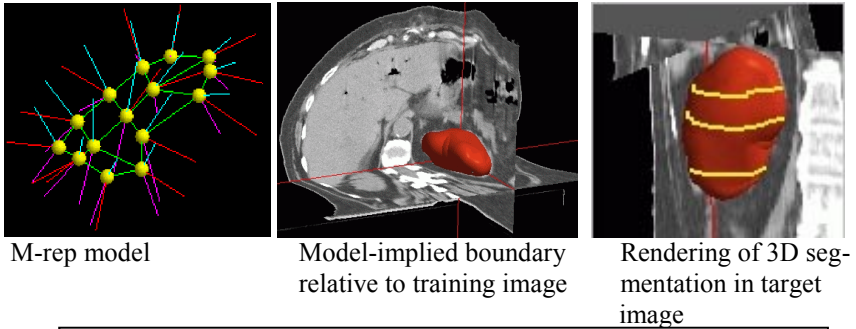
5. Conclusions

Extracting the hippocampus from MRI is very challenging for humans and has great variability across human segmenters. This provides a major challenge for automatic segmentation, but so far we have achieved 8 successful, reproducible segmentations. Fig. 6 shows comparison for our deformable m-reps segmentation based on a training template to human manual segmentation.

| | | |
|---|------------------------------|---------------------------------------|
| Median boundary offset: | Median case: 2.0* mm. | Worst case: 2.0* mm |
| 3rd quartile boundary offset: | Median case: 2.0* mm. | Worst case: 3.4[†] mm |
| Percent volume overlap: | Median case: 89% | Worst case: 81% |

Table 1. Comparison of deformable m-reps segmentation to manual segmentation of four kidneys from CT, using a Gaussian derivative template. *distance between face-adjacent voxels, [†]distance between corner-adjacent voxels

Fig. 4. Kidney model and segmentation results.



Segmentation results at the m-rep level of scale (i.e., before boundary displacement) on kidneys in CT using a single figure model. The three yellow curves on the m-rep implied boundary rendered in red in the 3D view above, right show the location of the slices shown in the center row. On these slices the curve shows the intersection of the m-rep implied boundary with the slices. The slices in the lower row are the sagittal and coronal slices shown in the 3D view.

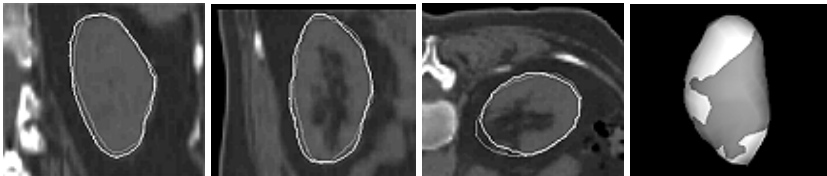


Fig 5. The result of 3D boundary displacements on the kidney. Grey curves: before displacement. White curves: after displacement. Left three: on three orthogonal slices. Right: surface rendered.

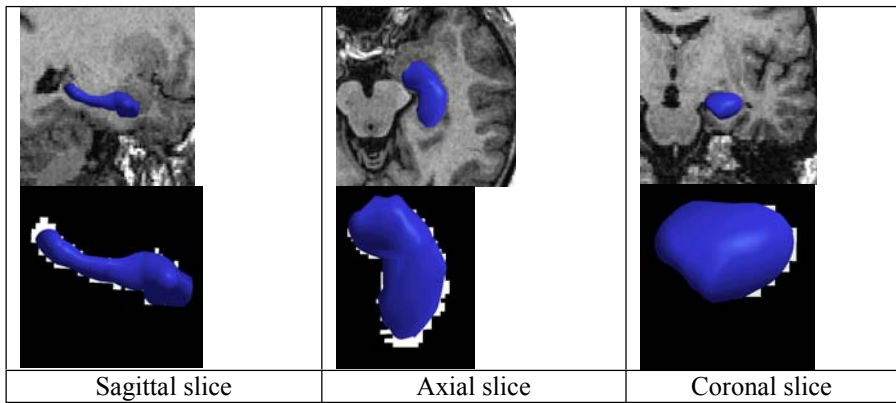


Fig. 6. Hippocampus results. In the left image of each pair the grayscale image is overlaid with the implied boundary of the m-rep hippocampus segmentation using the image template match. In the right image of each pair the match with a manual segmentation of a human expert is visualized in a zoomed version of the same slice as the grayscale image.

The time required for a segmentation is as follows. The manual initial placement, orientation, and magnification requires less than a half minute for a human. Steps 1b (the figural similarity transform) and 3 (the boundary displacement) each require a few seconds each; the total computing time is dominated by step 2, the atom by medial atom optimizations. For the conjugate gradient optimization algorithm that we use and the atom spacing of between $0.5r$ and r that we use, we find that between 2 and 5 passes of optimization of all the atoms are needed for convergence. The time for the optimization of an atom is roughly proportional to the area of implied boundary covered by that atom and its immediate neighbors. Put another way, a full pass optimizing all the atoms via 5 steps of the conjugate gradient process requires time proportional to the implied surface area of the whole figure, in a way only slightly related to the number of atoms. For the 3×5 atom kidney and the 3×8 hippocampus models, this requires time of about 40 seconds per full pass of the atoms, for a total time of 1.5-3 minutes for convergence on a 700MHz PIII PC with 128MB RAM. We have not optimized our code, nor have we taken advantage of the multiple processors available on our standard PCs.

Segmentation by deformable m-reps appears to work effectively and efficiently on single figure objects, even for challenging image situations of variable contrast. While no comparison to alternative techniques have been made and the reproducibility of these results against variable manual placement of the model has not yet been adequately studied, this method seems to have peculiar advantages. More results can be found at the website www.cs.unc.edu/Research/Image/MIDAG/defmreps.

We ascribe the good quality of these results to three factors: 1) the multiple scale level approach with geometric typicality at each stage depending on the result of the previous stage, 2) the ability of the m-rep representation to provide object-intrinsic correspondences at the figural and figural section stages, 3) the availability of the

training intensity template for the more difficult cases. While we have only a few cases of the kidney and all (but still only a few) cases of the hippocampus where we have used this training image template, we have found that success with the training image template in all cases if the match was successively made against a highly blurred version of the target image, then a moderately blurred version of the target image, and then the target image.

The method described here is by no means fully developed. The metrics, the segmentation algorithm, and the visualizations and user interface, and the program code have already been extended to deal with objects made up of multiple attached figures which must be kept in the correct geometric relations as they deform. Examples are the cerebral ventricle, the vertebra, and the kidney parenchyma. Extension has also been made to deal with multiple nonoverlapping figures which must be kept in the right geometric relations and to remain noninterpenetrating. Examples are the pubic bones, bladder, prostate, and rectum in the male pelvis and the full set of cerebral ventricles. Early, incomplete trials of the extended versions of the code suggest that m-reps have particular advantages also with multiple attached figures and multiple nonoverlapping figures.

Two important directions to improve deformable m-reps based segmentation are the following. Applying the segmentation at multiple levels of medial meshing, in coarse to fine order, is expected to speed the method for any level of effectiveness. This multiscale approach will overcome the question of what is the best level of meshing and replace it by the question of the spacing between the scale levels.

The replacement of the geometric distance measures for geometric typicality and average intensity correlation or mutual information for the geometry to image match measure by log probability measures [Cootes & Taylor 1999] has two important advantages. First, the probabilities reflect the modes of variability in the respective population. Second, the arbitrary, manually selected weight between geometric typicality and geometry to image match is no longer necessary. We will soon begin work on methods for measuring these probabilities from training sets, at each of the relevant scale levels based on a Markov random field model, and for using them in the model deformation process.

Acknowledgements

We are grateful to conceptual, geometric, algorithmic, or code contributions from Andrew Thall, Graham Gash, and Paul Yushkevich. We thank Edward Chaney and Guido Gerig for providing driving problems for this segmentation and images to segment. We appreciate the contributions of Daniel Fritsch and John Glotzer to earlier versions of this method. This work was done under the partial support of NCI Grant P01 CA47982 and NSF grant CCR SGER 9910419.

References

1. Casselles, V, R Kimmel and G Sapiro (1997). Geodesic Active Contours. *IJCV* **22**(1): 61-79.
2. Cootes, TF, C Beeston, GJ Edwards, CJ Taylor, A Unified Framework for Atlas Matching Using Active Appearance Models. *Information Processing in Medical Imaging* (IPMI 1999), A Kuba, M Samal, A Todd-Pokropek eds., Springer LNCS **1613**: 322-333.
3. Crouch, JR, SM Pizer, EL Chaney, M Zaider, Elastic registration of prostate images using the finite element method with m-rep models. Poster, 2nd Int. Conf. On Innovative Solutions for Prostate Cancer Care 2001. Also at website http://www.cs.unc.edu/Research/Image/MIDAG/pubs/presentations/prostate-mrep-Crouch2001_files/frame.htm.
4. Joshi, S, SM Pizer, PT Fletcher, A Thall, G Tracton, Multi-Scale 3-D Deformable Model Segmentation Based on Medical Description. *Information Processing in Medical Imaging* (IPMI 2001), MF Insana, RM Leahy, eds., Springer LNCS **2082**: 64-77.
5. Kelemen, A, G Szekely and G Gerig (1999). Elastic Model-Based Segmentation of 3D Neuroradiological Data Sets. *IEEE Transactions On Medical Imaging* **18**: 828-839.
6. Montagnat, J, H Delingette (1998). Globally Constrained Deformable Models for 3D Object Reconstruction. *Signal Processing* **71**: 173-186
7. Pizer, SM, PT Fletcher, Y Fridman, D Fritsch, AG Gash, J Glotzer, S Joshi, A Thall, G Tracton, P Yushkevich, EL Chaney, Deformable M-Reps for 3D Medical Image Segmentation. Internal report 2001 at website <ftp://ftp.cs.unc.edu/pub/users/MIDAG/defmrep3d.final.pdf>
8. Staib, LH, JS Duncan (1996). Model-based Deformable Surface Finding for Medical Images. *IEEE Trans. Med. Imaging* **15**(5): 1-12.
9. Styner, M, G Gerig, Medical Models Incorporating Object Variability for 3D Shape Analysis. *Information Processing in Medical Imaging* (IPMI 2001), MF Insana, RM Leahy, eds., Springer LNCS **2082**: 502-516.
10. Fletcher, PT, SM. Pizer, A Thall, P Yushkevich with S Joshi. Medial Geometry and Interpolation through Medial Correspondence. In preparation (2001).
11. Willis, L, A Geometric Description of Lung Shape During Respiration via M-reps and Normalized Mutual Information. MS thesis 2001, Dept. of Biomedical Engineering, UNC in preparation
12. Yushkevich, P, SM Pizer, S Joshi, JS Marron, Intuitive, Localized Analysis of Shape Variability. *Information Processing in Medical Imaging* (IPMI 2001), MF Insana, RM Leahy, eds., Springer LNCS **2082**: 402-408.