DEFORMABLE M-REP SEGMENTATION OF OBJECT COMPLEXES

P. Thomas Fletcher, Stephen M. Pizer, A. Graham Gash, Sarang Joshi

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

ABSTRACT

While we have previously reported on multiscale segmentation of single-figure anatomic objects from medical images by deformable m-rep models, here we report on a method of segmentation of multiple geometrically related single-figure objects from a medical image. Inter-object geometric effects within segmentation of structures in the male pelvis from CT in the planning of radiation therapy of the prostate will be illustrated. Given a model in which each object is formed from a mesh of medial atoms, the method adds an object-complex similarity transformation stage ahead of the figural, medial atom, and boundary deformation stages of model transformation. The final three stages reward geometric interrelations among objects that occur in the model, according to the object-intrinsic coordinate systems provided by m-reps.

1. INTRODUCTION & REVIEW OF DEFORMABLE M-REP SEGMENTATION OF SINGLE-FIGURE OBJECTS

In previous papers [3, 5] we have described a representation for 3D objects called m-reps. An m-rep for a generic *figure* is a 2-manifold of medial atoms (Fig. 1). An interior medial atom is a medial position at which two vectors (called *port* and *starboard sails*) of equal length r share a tail, and a mesh-edge medial atom has an additional sail that bisects the other two sails and is of length greater than or equal to r. In the representation used in this paper the medial manifold is sampled into a quad-mesh. This medial representation implies (as opposed to being implied by) a boundary that is incident to and orthogonal to the sail tips for each sail and at the bisector vector tip. The locus of bisector tips for sheet-edge medial atoms forms the crest cycling the boundary of the figure.

The two coordinates (u,v) along and across the manifold provide an object-intrinsic along-figural parametrization, and the along-sail directions provide a corresponding cross-figural coordinate τ, measuring the signed fractional distance from the boundary to the medial locus by taking it to be zero at the medially implied boundary and negative inside. All of these coordinates measure distance in multiples of the figural width, r. The around-crest figural
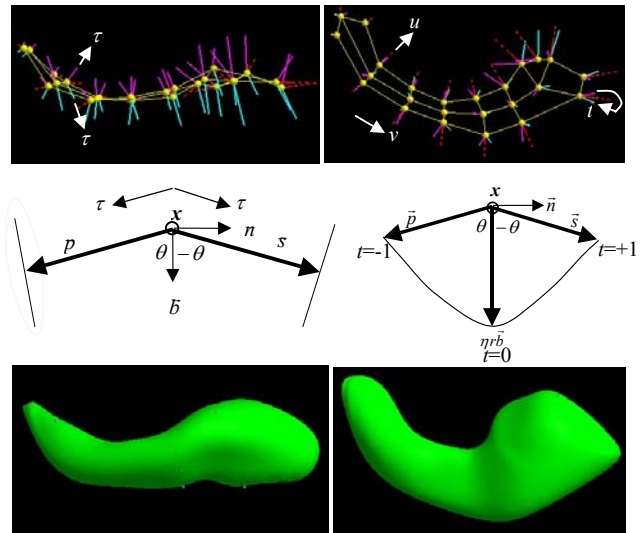

Fig. 1. Top: an m-rep for a hippocampus, viewed from two directions, with figural coordinates indicated. Each ball with two line segment sails forms a medial atom. Shown is a quad-mesh of samples of the continuous 2-manifold of medial atoms. Center: an internal medial atom and a sheet-edge medial atom, each with their implied boundaries and with figural coordinates indicated. Bottom: The boundary implied by the m-rep, viewed from the two directions.

coordinate t designates, by its value ± 1, on which side of the medial locus a point is. Around the crest as the point moves from one medial side to the other, t passes continuously from $+1$, through 0 at the crest, to -1. Thus any point in 3-space can be indicated by its object-intrinsic (*figural*) coordinates (u,v,t,τ).

In these same papers we have shown how an effective multiscale method of segmentation by deformable models has been constructed based on m-rep models. In this method at each of successively smaller object-related scale levels an objective function summing geometric typicality and geometry-to-image match is optimized. At each scale level the geometric typicality for a particular geometric primitive at that scale level is the sum of two terms. The first term measures, by τ^2 averaged over the medially implied boundary, the relationship between that primitive and its value provided by the next larger scale level. The second term measures the relationships between that geometric primitive and its value as predicted by the geometric primitives that immediately neighbor it

at that scale. The geometry-to-image match is measured at positions affected by that geometric primitive and with apertures that are proportional to the local figural width. The positions at which the match is measured are in object-intrinsic coordinates chosen at modeling time, typically in a collar about the medially implied boundary designated by a threshold on $|\tau|$.

In the model deformation

1) the largest scale level is the object, in those papers limited to a single figure (so in this paper it is called "the figural stage"), and the figural transformation at that scale level is limited to a similarity transform composed with a figural elongation.

2) at the next smaller scale level each medial atom, in succession, is optimally transformed in its 8 first-order parameters: medial position, 3D orientation, figural width, and figural object angle; and neighbors are the up to 4 adjacent medial atoms. The necessary neighbor prediction can take place since each atom can be described, according to the model, in the intrinsic coordinates of a neighboring atom.

3) at the smallest scale level, finely spaced boundary tiling vertices implied by the m-rep are transformed by an optimal displacement along the medially implied normal. Neighbor boundary vertices predict the vertex in question to have the same displacement as they had. The boundary tiling is produced using the subdivision method of Thall [7].

This paper describes the extension of this method to a larger object-related scale level, that of the multi-object complex, for single figure objects without overlap. We show how the m-rep representation allows the location of one figure in the object-intrinsic coordinate system of a neighboring figure. This can be used to give predictions of an object to be transformed by a neighboring object as well as to prevent inter-object penetration when an object is being transformed at whichever scale level.

The remainder of this paper begins with a section describing the segmentation strategy for multi-object complexes followed by a section on the representation of inter-object geometry for single figure objects. These are followed by a section on results to date in regard to inter-object geometric effects within multi-object segmentation by deformable m-reps. We close with a discussion.

2. SEGMENTATION STRATEGY FOR MULTI-OBJECT COMPLEXES

The segmentation strategy described above needs to be augmented in two ways. First, we add a multi-object scale level 0 at the next larger scale above the three scale levels above. At this scale level a similarity transform for the whole object complex is chosen by optimization of an objective function that is the same as that used at the single-figure scale level, except that the boundaries and boundary collars used in the geometric typicality and geometry-to-image match terms, respectively, are computed from all objects' boundaries in the complex and the scale of the image interrogation is appropriately increased. Second, the figural scale level followed by the medial atom scale level is applied for the objects in succession. Each object is placed in an optimal position relative to that predicted by nearby, already placed objects, as measured by the value $shift^2$ described below. In the subsequent optimization of the object at the figural and medial atom stages, as at the smaller scale levels the geometric typicality function for each object contains a term, also described below, reflecting the relation with neighboring objects. The order in which the objects are transformed is designated during model building, in decreasing order of the strength (contrast, low variability) of the image data portraying the objects.

For each object to be transformed (OT), any of the objects that have already been transformed can provide geometric information relevant to the proper placement of the OT. Normally these are taken to be the nearby, or even abutting, objects. Any such object is designated at model building time as a neighbor object of the OT. The desire is that the OT stay in close to the same relationship in inter-object coordinates to each neighbor as it is in the model and that it not interpenetrate the neighbor, or at least not interpenetrate by an amount that is larger than can be corrected at the following smaller scale levels. As described in section 3, the figural coordinate systems allow the definition of the measure for disagreement, $shift^2$, from each neighbor's prediction; this measure is averaged over all predicting objects. The figural coordinate system also allows the definition of a measure for interpenetration, and this measure is added to the prediction disagreement measure to produce the inter-object neighbors term used in the geometric typicality first when the OT is first similarity transformed from its neighbor-predicted position and then when its medial atoms are transformed.

We will not show results of the boundary scale level in this paper, but suffice it to say, interpenetration must also be avoided at this stage.

3. INTER-OBJECT GEOMETRY

The inter-object neighbors term (Fig. 2), is the average, over the neighbor objects, $N_k(OT)$, of $-Diff(OT, N_k(OT))$, which measures the geometric relation between the OT and that neighbor object. $Diff(OT, N_k(OT))$ is computed (Fig. 2) as the sum of a term measuring shifts of the OT boundary from positions predicted by the neighbor and an interpenetration term $f(\tau)$ that is near zero if positions on the medially implied boundary of the OT fall into positions of the neighbor that have positive τ in the neighbor's figural coordinates, i.e., are outside its implied boundary, and rises to a very large value to the extent that the any position on the OT's boundary has negative τ in the
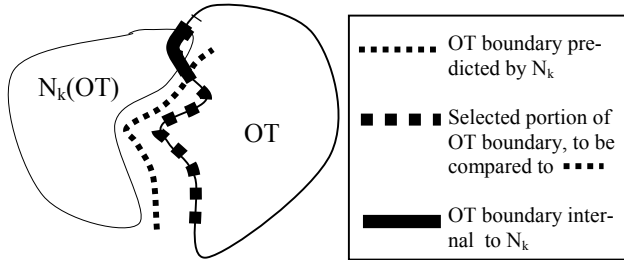
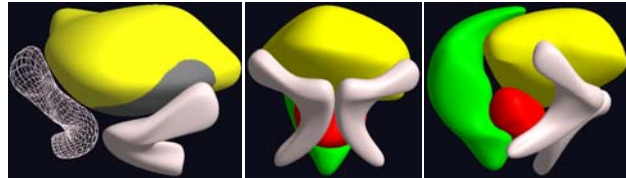Fig. 2. The neighbor-to-object to be transformed difference

Fig. 3. Left: Shown in grey on the lighter bladder OT is the boundary section of the OT that is governed by the left pubic bone, shown in white. The right pubic bone, which also governs the bladder, is shown in wireframe. Right pair: Model of male pelvis in front and side views: pubic bones (white), bladder (light), rectum (mid-grey), prostate (darkest).

neighbor's figural coordinates, i.e., is inside its implied boundary. The neighbor prediction terms involves the shift d of each selected OT boundary point from its position predicted from $N_k(OT)$ and r, the medial radius associated with that boundary point in the model of $N_k(OT)$. Then the neighbor prediction term is $(d/r)^2$ averaged over a section of the OT boundary selected at model building time.

This model building includes, for each OT and for each governing neighbor $N_k(OT)$, selecting a boundary section of OT to be governed. These points, whose positions are taken to determine the inter-object geometry, are chosen as all OT boundary points within a user-chosen $|\tau|$ threshold from $N_k(OT)$ (Fig.3), with the region being chosen indicated on an interactive display.

For each of the selected OT boundary points, that point's figural coordinates $(u,v,t)_{OT}$ in the OT coordinate system and that point's figural coordinates $(u,v,t,\tau)_k$ in the governing (neighbor) object's coordinate system are recorded; these latter coordinates will be used as the predictor of the boundary point $(u,v,t)_{OT}$ after object k has been transformed. These predictors can be expected to be good only within about $2r$ from the surface of OT and within its external focal surfaces.

During the transformation of the OT during segmentation, for each selected boundary point the Euclidean coordinates (ξ, η, ζ) are extracted for each boundary figural coordinate $(u,v,t)_{OT}$. For each such point we find the predicted position (x,y,z) corresponding to the neighbor's $(u,v,t,\tau)_k$, and we measure the disagreement between the predicted and candidate position in r-proportional units: $shift^2 = [(\xi-x)^2 + (\eta-y)^2 + (\zeta-z)^2)]/r(u,v)^2$. Also, we find the new τ_k of (ξ, η, ζ), compute the interpenetration penalty $f(\tau_k)$, and add this to $shift^2$. We average the resulting inter-object penalty term over all the boundary points and over all neighboring objects k.

4. MULTI-OBJECT SEGMENTATION BY DEFORMABLE M-REPS

The implementation uses a conjugate gradient algorithm to optimize the objective function at each scale level. A legality term is added to the geometry term during the atom deformation level to prevent self-intersections of the boundary. This term penalizes high curvature, i.e., the average over the boundary of the square of the dot product between adjacent triangles' normals.

An m-rep model of the male pelvis (Fig. 3) was built to fit a single CT training image. The model's objects follow: right pubic bone: 3×5 atoms; left pubic bone: 3×5 atoms, with the right bone as its neighbor; bladder: 4×4 atoms, with both bones as neighbors; rectum: 3×5 atoms, with the bladder as its neighbor; prostate: 3×3 atoms, with all 4 other objects as neighbors. We model only sections of the rectum and the bones, namely, the sections that are near other objects of interest.

The model was then deformed using the above method into two different target patients. The target images used for the bones were windowed CT data. Because the goal was to test the effectiveness of using the inter-object geometric relationships and not the image match, the target images for the other organs were label images from hand segmentations of CT data.

The number of iterations required was approximately 20 at each object stage and approximately 5 passes through all the atoms at the medial atom stage. The total time to segment all five objects was about 20 minutes on a 1.7 GHz Pentium 4.

Results from the m-rep segmentation of one of the images are shown in Fig. 4. First the right bone transforms into place (Fig 4a). Next the left bone's initial position is predicted in relation to the right bone and it then is optimally transformed. Figs. 4b&c show the layout at this stage, where the bones have penetrated the bladder. Fig 4d shows the nicely predicted bladder's position from its relation to the bones. After the bladder and the rectum have been transformed, removing all interpenetrations, the prostate's position is predicted. Fig. 4e shows how nicely the prostate is predicted solely on its geometric relationship to the other objects. The final result (Fig. 4f) is attractive but shows how the inter-object relations can prevent segmentation of unpredicted sections of an object. In the second target image (not shown) the bone segmentations succeeded as before. The rectum label image was from a lower section of the rectum than was modeled. However, since the rectum is rewarded for maintaining its relationship with the bladder, the model is not drawn by
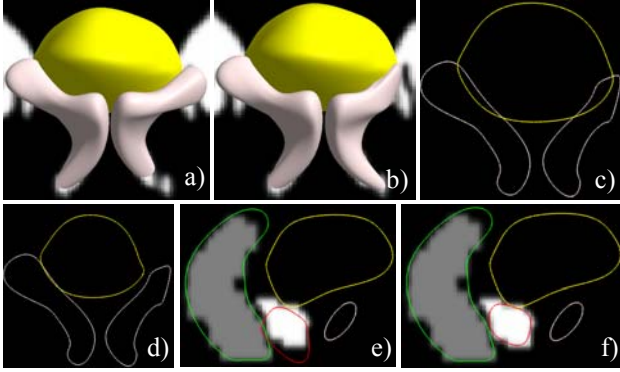
Fig. 4. A multi-object segmentation of male pelvic structures: successive scale levels and steps. a) After the right bone transformation, vs. the windowed image that provided the image data. b) After both bone transformations. c) Same layout as b, slice view. d) Same slice as c after bladder position has been predicted. e) After bladder and rectum segmentations, the prostate's position is predicted. f) Prostate in final position vs. image data.

the image data to the lower section. The other objects again predict the prostate well, and its segmentation is of similar quality as in Fig. 4.

5. DISCUSSION AND CONCLUSION

Characterizing one object in the figural coordinates of another appears to be effective in initially locating objects and in keeping them in the proper relation. We anticipate that beginning a segmentation at a multi-object scale level will for many segmentations make any manual initial placement of the model unnecessary, as the amount by which the multi-object model will have to move from a standard position in the image will be a fraction $\ll 0.5$ of the extent of the model. The resulting method for simultaneous extraction of multiple objects in a coarse-to-fine fashion while avoiding interpenetration has potentially improved abilities compared to other deformable model methods with rich geometric priors [1, 2, 4] that are either designed for single-figure objects or handle multi-object complexes at a single scale level and thus somewhat slowly. Results of actual use of our method in multi-object segmentation on real images will come at the presentation at the ISBI Conference, according to the research schedule set when this paper was invited. Validation vs. human segmentation is planned for after ISBI.

A direction for improvement is a common coordinate system for interstitial regions related to all nearby objects. Another important direction to improve segmentation using this approach is further to develop geometry-to-image match. Other developments that are anticipated are the straightforward extension to multifigure objects, the improvement of the boundary folding penalty, and having separate weights for an object to be transformed sliding along a neighboring object and pulling away or toward that neighbor.

Serious improvement in segmentation should derive from using multi-object models built from populations, extending the Styner method [6] that presently applies to single objects. The extension will yield not only object model medial sheets and samplings, but also probability distributions for the figures and medial atoms to use in place of the manually weighted distance functions presently used for geometric typicality, probability distributions of neighbor relations to use in place of the manually weighted distance functions presently used, and probability distributions of image intensities in an intrinsic multi-object coordinate system in place of the normalized intensity correlations presently used for geometry-to-image match. The ability to derive probability distributions of multi-object geometry will also lead to studies involving statistical geometric characterization of populations of multi-object complexes in normal vs. pathological classes.

REFERENCES
[1] Cootes, TF, CJ Taylor, DH Cooper, J Graham (1995). Active Shape Models - Their Training and Application. Computer Vision and Image Understanding , **61**(1):31-59.

[2] Cootes, T and Taylor A (1999). Unified Framework for Atlas Matching Using Active Appearance Models. IPMI 1999, **1613**: 322-333.

[3] Joshi, S, SM Pizer, PT Fletcher, A Thall, G Tracton (2001). Multi-Scale 3-D Deformable Model Segmentation Based on Medical Description. Proc. IPMI 2001, **2082**:64-77.

[4] Kelemen, A, G Szekely, G Gerig (1999). Elastic Model-Based Segmentation of 3D Neuroradiological Data Sets. *IEEE Transactions On Medical Imaging*, **18**: 828-839.

[5] Pizer, SM, JZ Chen, PT Fletcher, Y Fridman, D Fritsch, AG Gash, J Glotzer. MR Jiroutek, S Joshi, KE Muller, A Thall, G Tracton, P Yushkevich, EL Chaney (2001). Deformable M-Reps for 3D Medical Image Segmentation. Submitted, on web at http://midag.cs.unc.edu/pubs/papers/IJCV01-Pizer-mreps.pdf.

[6] Styner, M, G Gerig (2001). Medial Models Incorporating Object Variability for 3D Shape Analysis. Proc. IPMI 2001, **2082**: 502-516.

[7] Thall, A (2002). Fast C^2 Interpolating Subdivision Surfaces using Iterative Inversion of Stationary Subdivision Rules. University of North Carolina Computer Science Deptartment technical report TR02-001.