International Journal of Computer Vision 55(2/3), 85–106, 2003

Deformable *M-Reps* for 3D Medical Image Segmentation

STEPHEN M. PIZER, P. THOMAS FLETCHER, SARANG JOSHI, ANDREW THALL, JAMES Z. CHEN, YONATAN FRIDMAN, DANIEL S. FRITSCH, A. GRAHAM GASH, JOHN M. GLOTZER, MICHAEL R. JIROUTEK, CONGLIN LU, KEITH E. MULLER, GREGG TRACTON, PAUL YUSHKEVICH AND EDWARD L. CHANEY
Medical Image Display & Analysis Group, University of North Carolina, Chapel Hill

Received October 26, 2001; Revised September 13, 2002; Accepted June 27, 2003

Abstract. *M-reps* (formerly called DSLs) are a multiscale medial means for modeling and rendering 3D solid geometry. They are particularly well suited to model anatomic objects and in particular to capture prior geometric information effectively in deformable models segmentation approaches. The representation is based on *figural models*, which define objects at coarse scale by a hierarchy of figures—each figure generally a slab representing a solid region and its boundary simultaneously. This paper focuses on the use of single figure models to segment objects of relatively simple structure.

A single figure is a sheet of medial atoms, which is interpolated from the model formed by a net, i.e., a mesh or chain, of medial atoms (hence the name *m-reps*), each atom modeling a solid region via not only a position and a width but also a local figural frame giving figural directions and an object angle between opposing, corresponding positions on the boundary implied by the *m-rep*. The special capability of an *m-rep* is to provide spatial and orientational correspondence between an object in two different states of deformation. This ability is central to effective measurement of both geometric typicality and geometry to image match, the two terms of the objective function optimized in segmentation by deformable models. The other ability of *m-reps* central to effective segmentation is their ability to support segmentation at multiple levels of scale, with successively finer precision. Objects modeled by single figures are segmented first by a similarity transform augmented by object elongation, then by adjustment of each medial atom, and finally by displacing a dense sampling of the *m-rep* implied boundary. While these models and approaches also exist in 2D, we focus on 3D objects.

The segmentation of the kidney from CT and the hippocampus from MRI serve as the major examples in this paper. The accuracy of segmentation as compared to manual, slice-by-slice segmentation is reported.

Keywords: segmentation, medial, deformable model, object, shape, medical image

1. Introduction

Segmentation via deformable models has shown the advantage of allowing the expected geometric conformation of objects to be expressed (Cootes, 1993; Staib, 1996; Delingette, 1999; among others, also see McInerny, 1996 for a survey of active surfaces methods). The basic formulation is to represent an object by a set of geometric primitives and to deform the object by changing the values of the primitives to optimize an objective function including a match of the deformed object to the image data. Either the objective function also includes a term reflecting the geometric typicality of the deformed object, or the deformation is constrained to objects with adequate geometric typicality. In our work the objective function is the sum of a geometric typicality term and a geometry to image match term.

The most common geometric representation in the literature of segmentation by deformable models has been a mesh of boundary locations. The hypothesis described and tested in this paper is that improved

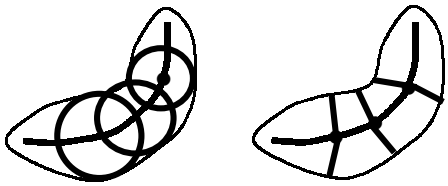

Figure 1. A 2D illustration of (left) the traditional view of the medial locus of an object as a sheet of disks (spheres in 3D) bitangent to the object boundary and our equivalent view (right) as an *m-rep*: a curve (sheet in 3D) of hubs at the sphere center and equal length spokes normal to object boundary. The locus of the spoke ends forms the medially implied boundary.

segmentations will result from using a representation that is at multiple levels of scale and that at all but the finest levels of scale is made from meshes of medial atoms. We will show that this geometric representation, which we call *m-reps*, has advantages in measuring both the geometric typicality and the geometry to image match, in providing the efficiency advantages of segmentation at multiple scales, and in characterizing the object as an easily deformable solid.

Many authors, in image analysis, geometry, human vision, computer graphics, and mechanical modeling, have come to the understanding, originally promulgated by Blum (1967), that the medial relationship[1] between points on opposite sides of a figure (Fig. 1) is an important factor in the object's geometric description. Biederman (1987), Marr (1978), Burbeck (1996), Leyton (1992), Lee (1995), and others have produced psychophysical and neurophysiological evidence for the importance of medial relationships (in 2D projection) in human vision. The relation has also been explored in 3D by Nackman (1985), Vermeer (1994), and Siddiqi (1999), and medial axis modeling techniques have been applied by many researchers, including Bloomenthal (1991), Wyvill (1986), Singh (1998), Amenta (1998), Bittar (1995), Igarashi (1999) and Markosian (1999). Of these, Bloomenthal and Wyvill provided skeletal-based soft-objects; Singh provided medial (wire-based) deformations; Amenta and Bittar worked on medially based reconstruction; Igarashi used a medial spine in 2D to generate 3D surfaces from sketched outlines; and Markosian used implicit surfaces generated by skeletal polyhedra.

One of the advantages of a medial representation is that it allows one to distinguish object deformations into along-object deviations, namely elongations and bendings, and across-object deviations, namely bulgings and attachment of protrusions or indentations. An additional advantage is that distances, and thus spatial tolerances, can be expressed as a fraction of medial width. These properties allow positions and orientations to be followed through deformations of elongation, widening, or bending. Because geometric typicality requires comparison of corresponding positions of an object before and after deformation and because geometry to image match requires comparison of intensities at corresponding positions, this ability to provide what we call a *figural coordinate system* is advantageous in segmentation by deformable models.

Medial representations divide a multi-object complex into objects and objects into *figures*, i.e., slabs with an unbranching medial locus (see Fig. 1). In the following we also show how they naturally divide figures into figural sections, and how by implying a boundary they aid in dividing the boundaries of these figural sections into smaller boundary tiles. This natural subdivision into the very units of medical interest provides the opportunity for segmentation at multiple levels of scale, from large scale to small, that provides at each scale a segmentation that is of smaller tolerance than the previous, just larger scale. Such a hierarchical approach was promulgated by Grenander (1981). Such a multi-scale-level approach is required for a segmentation that operates in time linear in the number of the smallest scale geometric elements, here the boundary tiles. The fact that at each level the units are geometrically related to the units of relatively uniform tissue properties yields effective and efficient segmentations.

Our *m-reps* representation described in Pizer (1999) and Joshi (2001) (in the first reference called DSLs) reverses the notion of medial relations descended from Blum (1967) from a boundary implying a medial description to a mesh of medial atoms implying boundaries, i.e., from an unstable to a stable relation. The radius-proportional ruler and the need to have locality at the scale of the figural section require it to use a width-proportional sampling of the medial surface in place of a continuous medial sheet.[2] These latter properties follow from the desire directly to represent shape, i.e., object geometry of some locality that is similarity transform invariant. The specifics are given later in this section.

M-reps also extend the medial description to the inclusion of a width-proportional tolerance, providing opportunities for stages of the representation with successively smaller tolerances. Representations with large tolerance can ignore detail and focus on gross

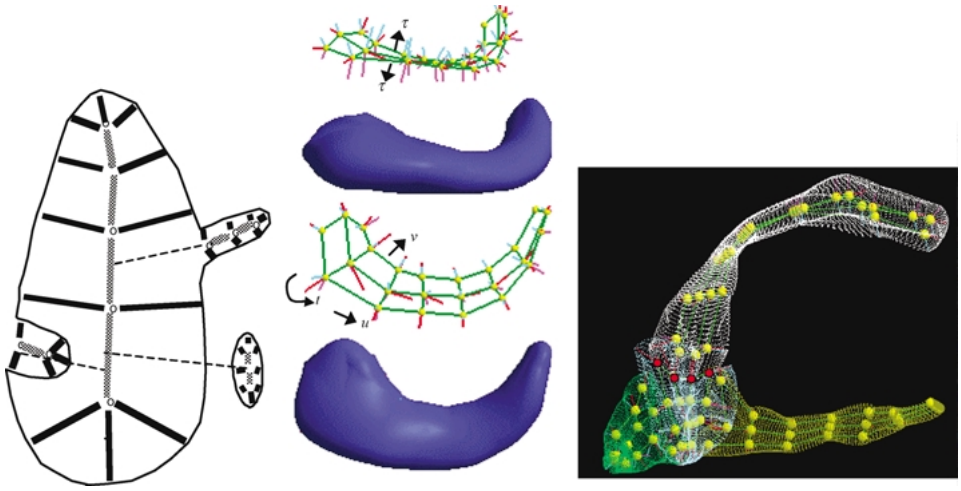

Figure 2. *M-reps*: In the 2D example (left) there are 4 figures: a main figure, a protrusion, an indentation, and a separate object. Each figure is represented by a chain of medial atoms. Certain medial atoms in a subfigure are interfigurally linked (dashed lines on the left) to their parent figures. In the 3D example of a hippocampus (middle) there is one figure, represented by a mesh of medial atoms. Each hub with two line segment spokes forms a medial atom (Fig. 3). The mesh is viewed from two directions, and the renderings below show the boundary implied by the mesh. The example on the right shows a 4-figure *m-rep* for a cerebral ventricle.

shape, and in these large-tolerance stages discrete samplings can be coarse, resulting in considerable efficiency of manipulation and presentation. Smaller-tolerance stages can focus on refinements of the larger-tolerance stages and thus more local aspects.

As described in Pizer (1999) and Joshi (2001), as a result of the aforementioned requirements an *m-rep* model of an object is a representation (data structure) consisting of a hierarchy of linked *m-rep* models for single figures (Fig. 2). A model for a single figure is made from a net (mesh or chain) of medial atoms (hence the name *m-reps*), each atom (Fig. 3) designating not only a medial position $\boldsymbol{x}$ and width r, but also a local figural frame $\underline{\boldsymbol{F}}$ implying figural directions, and the object's local narrowing rate, given by an object angle θ between opposing, corresponding positions on the implied boundary. In addition, width proportionality constants indicate net link length, boundary tolerance, boundary curvature limits, and, for measuring the fit of the atom to a 3D image, an image interrogation aperture. As detailed in later papers, a multifigure model of an object consists of a directed acyclic graph of figure nets, with interfigural links capturing information about subfigural location along the parent figure's medially implied boundary, figural width relative to the parent figure, and subfigural orientation relative to the parent figure. The elements of the figural graph also contain boundary displacement maps that can be used to give fine scale to the model.

Sometimes one wishes to represent and then segment multiple disconnected objects at the same time. An example is the cerebral ventricles, hippocampus, and caudate in which the structures are related but one

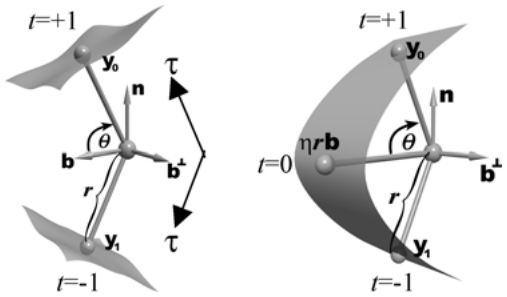

Figure 3. Medial atoms, made from a position $\boldsymbol{x}$ and two equal length boundary-pointing arrows $\vec{p}$ and $\vec{s}$ (for "port" and "starboard"), which we call "spokes". The atom on the left is for an internal mesh position, implying two boundary sections. The atom on the right is for a mesh edge position, implying a section of boundary crest. The atoms are shown in the "atom-plane" containing $\boldsymbol{x}$, $\vec{p}$ and $\vec{s}$. An atom is represented by the medial hub position $\boldsymbol{x}$; the length r of the boundary-pointing arrows; a frame made from the unit-length bisector $\vec{b}$ of $\vec{p}$ and $\vec{s}$, the $\vec{b}$-orthogonal unit vector $\vec{n}$ in the atom plane, and the complementary unit vector $\vec{b}^{\perp}$; and the "object angle" θ between $\vec{b}$ and each spoke. For a slab-like section of figure, $\vec{p}$ and $\vec{s}$ provide links between the medial point and the implied boundary (shown as a narrow curve), giving approximations, with tolerance, to both its position and its normal. The implied figure section is slab-like and centered on the head of the atom's spokes, i.e., it is extended in the $\vec{b}^{\perp}$ direction just as it is illustrated to do in the atom-plane directions perpendicular to its spoke.

is not a protrusion or an indentation on another. Another example are the pair of kidneys and the liver. In our system these can be connected by one or more connections between the representations of the respective objects, allowing the position of one figure to predict boundary positions of the other. This matter is left to a paper covering the segmentation of multi-object complexes (Fletcher, 2002).

In the remainder of this paper we first (Section 2.1) detail the *m-reps* data structure and geometry, then (Section 2.2) describe how a continuous boundary and medial sheets are interpolated from the sampled sheet directly represented in an *m-reps* figure, and then (Section 2.3) detail the way in which the *m-rep* provides positional and orientational correspondences between models and deformed models. After brief discussions in Sections 2.4 of model building from segmented objects that serve for model training, in Section 3 we discuss the method of segmentation by deformable *m-reps*. In Section 4 we give results of segmentation of kidneys, hippocampi, and a horn of the cerebral ventricle by these methods, and in Section 5 we conclude with a comparative discussion of our method and indications of future directions in which our segmentation method is being developed.

2. Representation of Objects by M-reps

2.1. M-reps Geometry

Intuitively a figure is a main component of an object or a protrusion, an indentation, a hole, or an associated nearby or internally contained object. In Pizer (1999) we carefully define a figure, making it clear that the notion is centered on the association of opposing points on the figure called by Blum (1967) "medial involutes." Whereas Blum conceived of starting from a boundary representation and deriving the medial involutes, our idea is to start with a representation giving medial information and thus widths, and imply sections of figure bounded by involutional regions. As illustrated in Fig. 3, in order for a medial atom $\underline{\boldsymbol{m}}$ by itself to imply two opposing sections of boundary, as well as the solid region between them, we define $\underline{\boldsymbol{m}} = \{\boldsymbol{x}, r, \underline{\boldsymbol{F}}, \theta\}$ to consist of

(1) a position, $\boldsymbol{x}$, the skeletal, or "hub," position (this requires 3 scalars for a 3D atom). $\boldsymbol{x}$ gives the central location of the solid section of figure that is being represented by the atom.
(2) a width, r, the distance from the skeletal position to two or more implied boundary positions (1 scalar) and thus the length of both $\vec{p}$ and $\vec{s}$. r gives the scale of the solid section of figure that is being represented by the atom. That is, it provides a local ruler for the object.
(3) a frame $\underline{\boldsymbol{F}} = (\vec{n}, \vec{b}, \vec{b}^{\perp})$, implying the tangent plane to the skeleton, via its normal $\vec{n}$, and $\vec{b}$, the particular unit vector in that tangent plane that is along the direction of fastest narrowing between the implied boundary sections. The medial position $\boldsymbol{x}$ and the two boundary-pointing arrows, as illustrated in Fig. 3, are in the $(\vec{n}, \vec{b})$ plane. The continuous medial surface implied by the set of $\underline{\boldsymbol{m}}$ must pass through each $\boldsymbol{x}$ and have a normal of $\vec{n}$ there. $\underline{\boldsymbol{F}}$ requires 3 scalars for a 3D atom. $\underline{\boldsymbol{F}}$ gives the orientation of the solid section of figure that is being represented by the atom. That is, it provides a local figural compass for the object. The frame is given by first derivatives of $\boldsymbol{x}$ and r with respect to distance along the tangent plane to $\boldsymbol{x}$.
(4) an "object angle" θ that determines the angulation of the implied sections of boundary relative to $\vec{b}$ $\vec{b}$ is rotated by $\pm\theta$ towards $\vec{n}$ to produce normals $\vec{p}/r$ and $\vec{s}/$ to the implied boundary. θ is normally between $\pi/3$ and $\pi/2$, the angle corresponding to parallel implied boundaries.

In 3D, figures are generically slabs, though *m-reps* can also represent tubes. A slab figure is a 2-sheet of medial atoms satisfying the constraint that the implied boundary folds nowhere. The constraint expresses the relation that the Jacobian of the mapping between the medial surface and the boundary is everywhere positive. A more complete, mathematical presentation of the constraints of representations of which legal *m-reps* are a subset can be found in Damon (2002).

In our representation a discrete *m-rep* for a slab is a mesh $\{\underline{\boldsymbol{m}}_{\mathrm{ij}},\ 1 \le i \le m,\ 1 \le j \le n\}$ of medial atoms that sample a 2-sheet of medial atoms (Fig. 2). We presently use rectangular meshes of atoms (quad-meshes), even though creating and displaying *m-reps* based on meshes of triangles (tri-meshes) have advantages because of their relation to simplices. In slab figures the net is a two-dimensional mesh and internal nodes of the mesh have a pair of boundary-pointing vectors pointing from $\boldsymbol{x}$ to $\boldsymbol{x} + \vec{p}$ and $\boldsymbol{x} + \vec{s}$.

Implied by the atoms in the mesh are

(1) a tolerance $\tau = \beta\lambda r$ of boundary position normal to the boundary,

(2) the length of links to other primitives, approximately of length $\gamma\lambda r$, with $\gamma\lambda$ significant fraction of 1.0, and
(3) a constraint $\delta\lambda r$ on the radius of curvature of the boundary.
(4) a size $\iota\lambda r$ of the image region whose intensity values directly affect the atom when measuring the match of the geometric model to the image.

The constant λ specifies the scale of the figural representation, which will vary from stage to stage in an algorithm working among coarse and fine levels of representation. The proportionality constants β, γ, δ, and ι are presently set by experience (Burbeck, 1996; Fritsch, 1997; McAuliffe, 1996), and, to maintain magnification invariance, the constants $\beta\lambda$, $\gamma\lambda$, $\delta\lambda$, and $\iota\lambda$ decrease in proportion as the scale decreases.

The successive refinement, coarse-to-fine, of a medial mesh can provide a successive correction to the medially implied object by interpolating atoms at the finer spacing from those at the coarser spacing and then optimizing the finer mesh (Yushkevich, 2001), although we have not implemented this feature yet in the *m-rep* models used in segmentation. This refinement brings with it a decrease in tolerance of the implied boundary and radius of curvature constraint, the addition of patches of medial deformations relative to a figural (u, v) space (see Section 2.2) to handle heavily bent sections of slab, as well as proportionately smaller constant of radius proportionality, λ.

We call a figure represented via *m-reps* an *m-figure*. The net of medial atoms contains internal nodes and end nodes, as well. The end nodes for a slab are linked together to form the boundary of the mesh. For an object made from a single figure, the end nodes need to capture how the boundary of the slab or tube is closed by what is called a *crest* in differential geometry (Koenderink, 1990). For example, a pancake is closed at its sides by such a crest. Whereas the internal nodes for a slab-like segment have two boundary-pointing vectors, end nodes for slab-like segments have three boundary-pointing vectors, with the additional vector pointing from $\boldsymbol{x}$ in the $\vec{b}$ direction to the crest. Thus $\vec{b}$ must cycle as one moves around the figural crest. Internal nodes for tubes have a circle of boundary-pointing vectors, obtained by adding to $\boldsymbol{x}$ the full circle of rotations in a full circle of $\vec{p}$ about $\vec{b}$.

For slabs a sequence of edge atoms forms a curve of a crest or a curve of a corner closing the slab. As illustrated in Fig. 3, these segment closed ends may be rounded with any level of elongation η: the vertex is taken at $\boldsymbol{x} + \eta r\vec{b}$, and the end section in the principal direction across the implied crest is described by an interpolation using the position and tangent there and at the two points $\boldsymbol{x} + \vec{p}$ and $\boldsymbol{x} + \vec{s}$ and applying an interpolating function to produce a boundary crest with the desired extent, tangency, and curvature properties. We use this formulation for ends of end atoms represented as $\{\boldsymbol{x}, r, \underline{\boldsymbol{F}}, \theta, \eta\}$, instead of the Blum formulation, in which $\eta = 1$, in order to stabilize the image match at ends as well as to allow corners, i.e., crests of infinite curvature, to have a finitely sampled representation. Although corners do not normally appear in medical images, they are needed to model manufactured objects. Corner atoms have their vertex at $\boldsymbol{x} + r(1/\cos(\theta))\vec{b}$.

2.2. *Interpolated Medial Sheets and Figural Rendering*

As stated above, an *m-rep* mesh of medial atoms for a single figure should be thought of as a representation for a continuous sheet of atoms and a corresponding continuous implied boundary. The sheet extends to the space curve of points osculating the crest of the implied figural boundary. We interpolate the *m-rep* mesh into this sheet parameterized by $(u, v) \in [(j, j+1) \times (k, k+1)]$ for the mesh element with the jkth atom at its lower left corner. The interpolation is obtained by a process that has local support on these mesh elements and on the edge elements that are bounded by edge atoms only. If the medial atoms are separated by constant r-proportional distances, this parametrization satisfies the objective of representing shape locally.

The interpolation is achieved by applying a variant of subdivision surface methods (Catmull, 1978) to the mesh of implied boundary positions and normals given at the spoke ends (including the crest spokes). The variant, described in detail in Thall (2003), makes the subdivision surface match the position and the normal of the spoke ends to within their tolerance. This boundary surface (see Fig. 4) is C^2 smooth everywhere but the isolated points corresponding to the atoms at the corners of the mesh. From this surface Thall's method allows the calculation of interpolated medial atoms.

As it stands, Thall's method is unlikely to produce folded boundaries but is not guaranteed against folds. In our segmentation we penalize against the high curvatures that are precursors to a deformation that yields folds. The full avoidance of medial atoms that imply

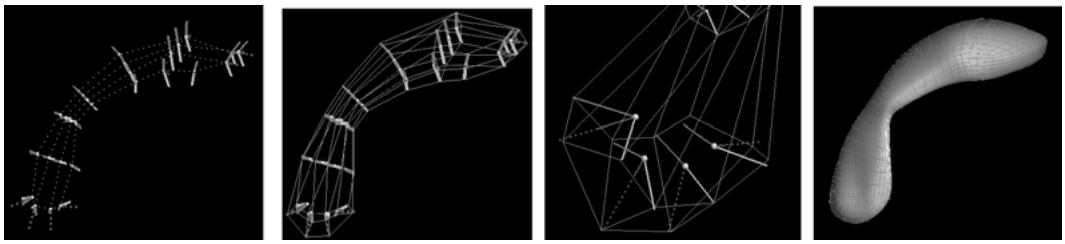

Figure 4. Left: A single-figure *m-rep*. Left middle: Coarse mesh of atom boundary positions for a figure. Right middle: Atom ends vs. interpolated boundary. Right: interpolated boundary mesh at voxel spacing.

a folded boundary will be achievable using the mathematical results found in Damon (2002).

The result of Thall's interpolation is that with each boundary position we can associate a *boundary figural coordinate* (Pizer, 2002) (Fig. 2), the figure number together with a side parameter t ($= -1$ for port, $= +1$ for starboard, with $t \in (-1, 1)$ around the crest) and the parameters (u, v), describing which interpolated atom's spoke touches the boundary there. For each figure we interpolate the atoms sufficiently finely that a set of voxel-size triangular tiles represent the boundary. The method computes the boundary position and associated normal and r value for an arbitrary boundary figural coordinate (u, v, t).

Points (x, y, z) in space can also be given a figural coordinate by appending an r-proportional distance τ (Fig. 1) to the figural coordinates of the closest medially implied boundary point. To allow the distinction by the sign of the distance of the inside and the outside of the figure, we take the distance to be relative to the medially implied boundary and to be negative on the interior of the figure. A procedure mapping arbitrary spatial positions (x, y, z) into figural coordinates (u, v, t, τ) has been written. Also, an arbitrary fine triangular tiling of the medially implied boundary can be computed. Rendering can be based on these triangular tiles or on implicit rendering using the τ function. As well, the correspondence under figural deformation given by figural coordinates is critically useful in computing the objective function used in segmentation.

2.3. *Correspondence Through Deformation*

As previously described, figures are designed to provide a natural coordinate system, giving, first a position on the medial sheet, second a figural side, and finally a figural distance in figural width relative terms along the appropriate medial spoke from a specified position. As detailed in Section 3.3, this figural coordinate system is used in our segmentation method to follow boundary locations and other spatial locations through the model deformation process.

In particular, as illustrated in Fig. 5, a boundary point after deformation, identified by its figural coordinate can be compared to the corresponding point before deformation, and the magnitude of the r-proportional distance between these points can be used to measure the local deformation. Also, the intensity in the target image, at a figural coordinate relative to a putatively deformed model can be compared to the intensity in a training image (or training images) at the corresponding figural coordinate relative to the undeformed model.

Figure 5. Correspondence over deformation via figural correspondence.

Figure 6. *M-reps* models. Heavy dots show hubs of medial atoms. Lines are atoms' spokes. The mesh connecting the medial atoms is shown as dotted curves. Implied boundaries are rendered with shading. Hippocampus: see Fig. 2. Left: kidney parenchyma + renal pelvis. Middle: lateral horn of cerebral ventricle. Right: multiple single-figure objects in male pelvis: rectum, prostate, bladder, and pubic bones (one bone is occluded in this view).

2.4. *M-reps Model Building for Anatomic Objects from Training Images*

Model-building must specify of which figures an object or multi-object complex is made up, the size of the mesh of each figure, and the way the figures are related, and it must also specify each medial atom. In this paper, focused on single-figure objects, only the mesh size and its medial atoms must be specified. Illustrated in the panels of Fig. 6 are single-figure *m-rep* models of a variety of anatomic structures that we have built.

Because an *m-rep* is intended to allow the representation of a whole population of an anatomic object across patients, it is best to build it based on a significant sample of instances of the segmented object. Styner (2001) describes a tool for stably producing models from such samples. The hippocampus model shown in Fig. 2 was built using this tool. We also have a design tool for building *m-rep* models from a single training 3D-intensity data set and a b-rep from a previous segmentation, e.g., a manual segmentation, of the object in that image. The kidney model shown in Fig. 6 was built using this tool.

It is obvious that effective segmentation depends on building a model that can easily deform into any instance of the object that can appear in a target image. However, the production of models is not the subject of this paper, so we assume in the following that a satisfactory model can be produced and test this fact via the success of segmentations.

3. Segmentation by Deformable *M-reps*

3.1. *Visualizations*

3.1.1. Viewing an M-rep. To allow appreciation of the object represented by an *m-rep*, the *m-rep* and the implied boundary must be viewable. To judge if an *m-rep* adequately matches the associated image, capabilities described in Sections 3.2 and 3.3 are needed to visualize the *m-rep* in 3D relative to the associated image. If the match is not good, the user needs a tool to modify the *m-rep*, either as a whole or atom by chosen atom, while in real time seeing the change in the implied boundary and the relation of the whole *m-rep* or modified atom to the image. After this seldom required manual modification the *m-rep* or atom may then be attracted by the image data.

As seen in Figs. 2, 4, and 6, we view an *m-rep* as a connected mesh of balls, with each ball attached to a pair of spokes and, for end atoms, to the crest-pointing $\eta r \vec{b}$ vector. The inter-atom lines, the spokes, and the crest-pointing vectors can be optionally turned off. In addition, the implied boundary can be viewed as a dot cloud, a mesh of choosable density, or a rendered surface.

3.1.2. Visualization of the M-rep vs. a Target Image. Visualization of greyscale image data must be in 2D; only in 2D image cuts can the human understand the interaction of a geometric entity with the image data. The implied boundary of an *m-rep* can be visualized in 3D versus one or more of the cardinal tri-orthogonal planes (x, y; y, z; or x, z), with the choice of these planes dynamically changeable (Fig. 11). Other image planes in which it is useful to visualize the fit of the *m-rep* to the image data are the atom-based planes described in the next paragraph. Showing the curve of intersection of the *m-rep* implied boundary with any of these image planes is useful.

The desired relationship of a single medial atom to the image data is that at each spoke end there is in the image a boundary orthogonal to the spoke. This relation is normally not viewable in a cardinal plane. Instead one needs to visualize and edit the atoms in cross sections of the object that are normal to the boundary.

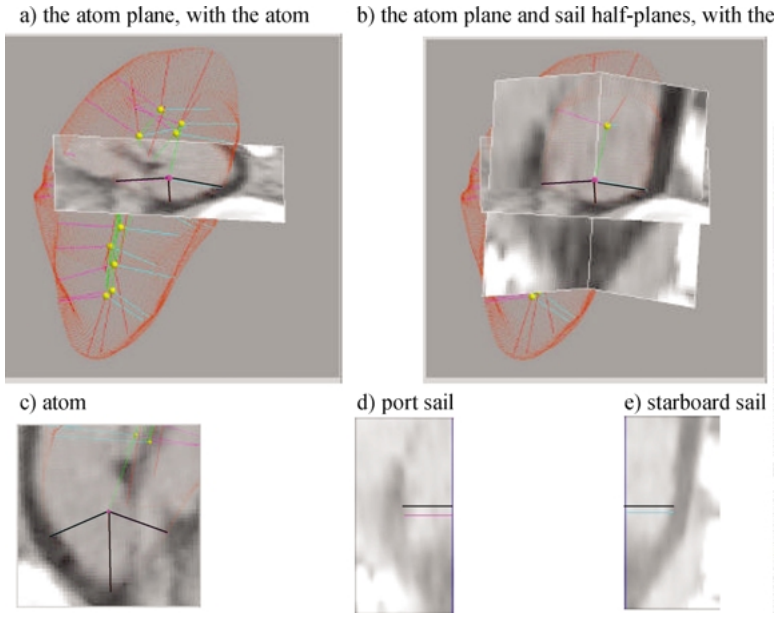

Figure 7. The viewing planes of interest for a medial atom: Top: 3D views. Bottom: 2D views.

The spokes of a medial atom, if they have been correctly placed relative to the image, are normal to the perceived boundary, so they should be contained in the cross section. We conclude that a natural view plane should contain both of the boundary pointing arrows of the medial atom, that is the plane passing through $\boldsymbol{x}$ of the atom and spanned by $(\vec{n}, \vec{b})$ of the atom (if the object angle is other than $\pi/2$). We call this the *atom plane*. We can superimpose the medial atom, as well as the medially implied boundary slice on the image in this plane (Figs. 7(a), (b), and (c)) and visualize in-plane changes in the atom and the boundary there.

On the other hand, the atom may be misplaced orthogonal to the atom plane or misrotated out of the atom plane, so it needs to be viewed in an image plane orthogonal to the atom plane. There is no such plane that contains both spokes, but if we could live with having only one spoke in the atom-orthogonal plane, we could view across the boundary to which the spoke should be orthogonal. The desired plane is that spanned by the atom's $\vec{b}^{\perp}$ and the chosen spoke of the atom. But we would only need the half plane ending at $\boldsymbol{x}$, the "port" or "starboard half-plane." Thus, we could draw simultaneously the two adjoining half planes for the respective spokes (Fig. 7).

When an atom is selected, it implies visualization planes, which can been chosen from among the atom plane, the port spoke half-plane, and the starboard spoke half-plane. The image data in the chosen plane(s) can be viewed in the 3D viewing window, with the image data texture rendered onto these planes (Fig. 7(a) and (b)), and in respective 2D windows for the chosen plane(s) (Figs. 7(c), (d) and (e)).

Manual editing of *m-reps*, should it be necessary, can be done using these visualizations. The viewer can point to positions in the three 2D viewing windows where the spoke ends should be moved, and the atom can be modified to as closely as possible meet these requirements, based on the object being locally a linear slab. Also, tools for manually rotating, translating, scaling, and changing the object angle and elongation (for end atoms) are easily provided.

3.2. *Multi-Scale-Level Model Deformation Strategy for Segmentation from Target Images*

Our method for deforming *m-reps* into image data allows model-directed segmentation of objects in volume data. The deformation begins with a manually chosen initial similarity transform of the model. To meet the efficiency requirements of accurate segmentation, the segmentation process then follows a number of stages of segmentation at successively smaller levels of scale (see Table 1). At each scale level the model is the result of the next larger scale level and we optimize

Table 1. Geometry by scale level.

Scale level k	Geometric entity	Transformation S^k	Primitive $\underline{z}_i^k$	Transformation parameters $\underline{\omega}_i^k$	Neighbors $\underline{N}(\underline{z}_i^k)$
1	Object ensemble	Similarity	Object ensemble pose	7: 3D sim transf params	None
2	Object	Similarity	Object pose	7: 3D sim transf params	Adjacent objects
3	Main figure	Similarity plus elongation	Figure pose	8: 3D sim transf params, 1 elongation param	Adjacent figures
3	Subfigure (attached to a host to represent a protrusion or indentation)	Similarity in figural coordinates of its host's figural boundary, plus hinging and elongation	Figural pose in host's cords and elongation	6: 4 2D sim transf params, 1 hinging param, 1 elongation param	Adjacent figures, possibly attached to same host
4	Through section of figure (medial atom)	Medial atom change	Medial atom value	8 (or 9): medial atom params ($+\eta$ for external atoms)	2–4 adjacent medial atoms
5	Boundary vertex	Displacement along medially implied normal	Boundary vertex position	1: displacement param	Adjacent boundary vertices

an objective function of the same form: the sum of a geometric typicality metric (detailed later in this section) and a geometry to image match metric (detailed in Section 3.3). At each scale level there is a type of geometric transformation chosen appropriate to that scale and having only at most 9 parameters.

The deformation strategy, from a model to a candidate obtained by geometrically transforming the model, follows two basic geometric principles, according to the conceptual structure presented in Sections 1 and 2.

(1) In both the geometric typicality and the model to image match metrics all geometry is in figurally related terms. Thus

- model-relative and candidate-relative positions correspond when they have common figural coordinates, and
- all distances are r-proportional.

(2) Calculating geometric typicality at any scale level is done in terms of the relations relevant to that scale, i.e., relative to its values predicted by the previous, next larger, scale and by its neighbors at its scale. The neighborhood of a medial atom is made up of its immediately adjacent atoms, and the neighborhood of a boundary tile vertex is made up of the adjacent boundary tile vertices.

To describe the algorithm in detail, we make a number of definitions.

The process begins with a model $\underline{\mathbf{z}}$ that is manually translated, rotated, and uniformly scaled into the image data by the user to produce an initialized model $\underline{\mathbf{z}}^0$. $\underline{\mathbf{z}}^0$ is successively transformed through a number of scale levels into deformed models $\underline{\mathbf{z}}^k$ until $\underline{\mathbf{z}}^5$ is the final segmentation. The details and descriptions of the primitives, their neighbor relations, and the associated transformations at each scale level are given in Table 1.

Let $\underline{\mathbf{z}}^k$ be the geometric representation at scale level k. Let $\underline{z}_i^k$ be the representation of the ith primitive at scale level k. At all scale levels $k \leq 4$, each $\underline{z}_i^k$ is represented as a collection of medial atoms, and a geometric transformation on $\underline{z}_i^k$ is computed by applying that transformation to each medial atom in its representation. Each primitive $\underline{z}_i^k$ for $k > 1$ has a small set of neighbors $\underline{\mathrm{N}}(\underline{z}_i^k)$ at scale level k and a geometric entity at the next larger scale $(k-1)$ that contains $\underline{z}_i^k$. We call this containing entity the parent primitive $\mathrm{P}(\underline{z}_i^k)$. While $\mathrm{P}(\underline{z}_i^k)$ is at scale level $k-1$, it is of the same type as $\underline{z}_i^k$. That is, for $k \leq 4$ $\mathrm{P}(\underline{z}_i^k)$ is represented as a superset of the set representing $\underline{z}_i^k$, and for $k = 5$ the parent of a boundary vertex is the corresponding vertex on the medially implied surface with zero displacement. Also associated with scale level k is a type of transformation S^k such that $\underline{z}_i^k = S^k\,\mathrm{P}(\underline{z}_i^k)$. Let the parameters $\underline{\omega}_i^k$ be the parameters of the particular transformation S^k applied to $\mathrm{P}(\underline{z}_i^k)$ at scale level $k-1$ to produce $\underline{z}_i^k$ at scale level k.

The similarity transform S consisting of translation by $\underline{t}$, rotation O and uniform scaling α applied to a medial atom $\underline{m} = \{\boldsymbol{x}, r, \underline{\boldsymbol{F}}, \theta\}$ produces $S \circ \underline{m} = \{\alpha O x + \underline{t}, \alpha r, O \circ F, \theta\}$. Figural elongation by ν leaves

fixed the medial atoms at a specified atom row i (the hinge end for subfigures) and successively produces translations and rotations of the remaining atoms in terms of the atoms in the previously treated, adjacent row i^-, as follows:

$$S^3(v) \circ \underline{m}_{ij} = \left\{x_{i^-j} + v(x_{ij} - x_{i^-j}),\, r_{ij},\, \left(\underline{F}_{ij}\underline{F}_{i^-j}^{-1}\right)^v \circ \underline{F}_{i^-j},\, \theta_{ij}\right\}$$

The subfigure transformation applies a similarity transform to each of the atoms in the hinge. This transformation, however, is not in Euclidean coordinates but in the figural coordinates of the boundary of the parent. That transformation is not used in this paper, so its details are left to Liu (2002). The medial atom transformation S^4 translation by $\underline{t}$, rotation O, r scaling α, and object angle change $\Delta\theta$ applied to a medial atom $\underline{m} = \{\boldsymbol{x}, r, \underline{\boldsymbol{F}}, \theta\}$ produces $S^4(\underline{t}, O, \alpha, \Delta\theta) \circ \underline{m} = \{x + \underline{t}, \alpha r, O \circ \underline{F}, \theta + \Delta\theta\}$. The boundary displacement transformation τ applied to a boundary vertex with position $\boldsymbol{y}$, medial radial width r, and medially implied normal $\vec{n}$ yields the position $y + \tau r\vec{n}$.

The algorithm for segmentation successively modifies $\underline{\mathbf{z}}^{k-1}$ to produce $\underline{\mathbf{z}}^k$. In doing so it passes through the various primitives $\underline{z}_i^k$ in $\underline{\mathbf{z}}^k$ and for each i optimizes an objective function $\mathrm{H}(\underline{\mathbf{z}}^k, \underline{\mathbf{z}}^{k-1}, \mathrm{I}) = \mathrm{w}^k(\text{-Geomdiff}(\underline{\mathbf{z}}^k, \underline{\mathbf{z}}^{k-1})) + \mathrm{Match}(\underline{\mathbf{z}}^k, \mathrm{I})$. $\mathrm{Geomdiff}(\underline{\mathbf{z}}^k, \underline{\mathbf{z}}^{k-1})$ measures the geometric difference between $\underline{\mathbf{z}}^k$ and $\underline{\mathbf{z}}^{k-1}$, and thereby $\text{-Geomdiff}(\underline{\mathbf{z}}^k, \underline{\mathbf{z}}^{k-1})$ measures the geometric typicality of $\underline{\mathbf{z}}^k$ at scale level k. $\mathrm{Match}(\underline{\mathbf{z}}^k, \mathrm{I})$ measures the match between the geometric description $\underline{\mathbf{z}}^k$ and the target image I. Both $\mathrm{Geomdiff}(\underline{\mathbf{z}}^k, \underline{\mathbf{z}}^{k-1})$, and $\mathrm{Match}(\underline{\mathbf{z}}^k, \mathrm{I})$ are measured in reference to the object boundaries B^k and B^{k-1}, respectively implied by $\underline{\mathbf{z}}^k$ and $\underline{\mathbf{z}}^{k-1}$. The weight w^k of the geometric typicality is chosen by the user.

For any medial representation $\underline{\mathbf{z}}$, the boundary B is computed as a mesh of quadrilateral tiles as follows, with each boundary tile vertex being known both with regards to its figural coordinates $\underline{u}$ and its Euclidean coordinates $\underline{y}$. For a particular figure, $\underline{u} = (u, v, t)$, as described in Section 2.2. When one figure is an attached subfigure of a host figure, with the attachment along the v coordinate of the subfigure, there is a blend region whose boundary has coordinates $\underline{u} = (u, w, t)$, where u and t are the figural coordinates of the subfigure and $w \in [-1, 1]$ moves along the blend from the curve on the subfigure terminating the blend ($w = -1$) to the curve on the host figure terminating the blend ($w = +1$). This blending procedure is detailed in Liu (2002).

As mentioned in Section 2.2, the computation of B is accomplished by a variation of Catmull-Clark subdivision (Catmull, 1978) of the mesh of quadrilateral tiles (or, in general, tiles formed by any polygon) formed from the two (or three, spoke ends of the medial atoms in $\underline{\mathbf{z}}$. Thall's variation (2003) produces a limit surface that iteratively approaches a surface interpolating in position to spoke ends and with a normal interpolating the respective spoke vectors. That surface is a B-spline at all but finitely many points on the surface. The program gives control of the number of iterations and of a tolerance on the normal and thus of the closeness of the interpolations. A method for extending this approach to the blend region between two subfigures is presently under evaluation.

$\mathrm{Geomdiff}(\underline{\mathbf{z}}^k, \underline{\mathbf{z}}^{k-1})$ is computed as the sum of two terms, one term measuring the difference between the boundary implied by $\underline{\mathbf{z}}^k$ and the boundary implied by $\underline{\mathbf{z}}^{k-1}$, and, in situations when $\underline{\mathrm{N}}(\underline{z}_i^k)$ is not empty, another term measuring the difference between boundary implied by $\underline{\mathbf{z}}^k$ and that implied by $\underline{\mathbf{z}}^k$ with $\underline{z}_i^k$ replaced by its prediction from its neighbors, with the prediction based on neighbor relations in $\mathrm{P}(\underline{z}_i^k)$. The second term enforces a local shape consistency with the model and depends on the fact that figural geometry allows a geometric primitive to be known in the coordinate system of a neighboring primitive. The weight between the neighbor term and the parent term in the geometrical typicality measure is set by the user. In the tests described in Section 4, the neighbor term weight was 0.0 in the medial atom stage and 1.0 in the boundary displacement stage.

The prediction of the value of one geometric primitive $\underline{z}_j^k$ in an *m-rep* from another $\underline{z}_i^k$ at the same scale level using the transformation S^k is defined as follows. Choose the parameters of S^k such that S^k applied to the $\underline{\mathbf{z}}^k$ subset of $\underline{\mathbf{z}}^{k-1}$ is close as possible to $\underline{\mathbf{z}}^k$ in the vicinity of $\underline{z}_j^k$. Apply that $S^k$to $\underline{\mathbf{z}}^k$ to give predictions $(S^k\ \underline{\mathbf{z}}^k)_j$. Those predictions depend on the prediction of one medial atom by another. Medial atom $\underline{z}_j^4 = \{\boldsymbol{x}_j, r_j, \underline{\boldsymbol{F}}_j, \theta_j\}$ predicts medial atom $\underline{z}_i^4 = \{\boldsymbol{x}_i, r_i, \underline{\boldsymbol{F}}_i, \theta_i\}$ by recording $T = \{(\boldsymbol{x}_j - \boldsymbol{x}_i)/r_j, (r_j - r_j)/r_j, \underline{\boldsymbol{F}}_j\underline{\boldsymbol{F}}_i^{-1}, \theta_j - \theta_i\}$, where $\underline{\boldsymbol{F}}_j\underline{\boldsymbol{F}}_i^{-1}$ is the rotation that takes frame $\underline{\boldsymbol{F}}_i$ into $\underline{\boldsymbol{F}}_j$. T takes $\underline{z}_i^4$ into $\underline{z}_j^4$ and when applied to a modified $\underline{z}_i^4$ produces a predicted $\underline{z}_j^4$.

The boundary difference $\mathrm{Bdiff}(\underline{\mathbf{z}}_1, \underline{\mathbf{z}}_2)$ between two *m-reps* $\underline{\mathbf{z}}_1$ and $\underline{\mathbf{z}}_2$ is given by the following average r-proportional distance between boundary points that correspond according to their figural coordinates, although it could involve points with common figural

coordinates other than at the boundary and it will in the future involve probabilistic rather than geometric distance measures.

$$\text{Bdiff}(\underline{\mathbf{z}}_1, \underline{\mathbf{z}}_2) = \left[-\int_{B_2} \frac{\|\underline{y}_1 - \underline{y}_2\|^2}{2(\sigma r(\underline{y}_2))^2} \, d\underline{y} \right] \Big/ \, area(B_2).$$

The r-value is that given by the model at the present scale level, i.e., the parent of the primitive being transformed. The normalization of distance by medial radius r makes the comparison invariant to uniform scaling of both the model and the deformed model for the local geometric component being adjusted at that scale level.

Finally, the geometry to image match measure Match ($\underline{\mathbf{z}}^k$, I) between the geometric description $\underline{\mathbf{z}}^k$ and the target image I is given by $\int_{-\tau_{\max}}^{\tau_{\max}} \int_{B^k} G(\tau) \, I_{\text{template}}(y, \tau) \hat{I}(y', \tau) \, d\underline{y} d\tau$ where y and y' are boundary points in $B(\underline{\mathbf{z}}^k)$ and $B(\underline{\mathbf{z}}^k_{\text{template}})$ that agree in figural coordinates, $G(\tau)$ is a Gaussian in τ, $\hat{I}$ is the target image I rms-normalized with Gaussian weighting in the boundary-centered collar $\tau \in [-\tau_{\max}, \tau_{\max}]$ for the deformed model candidate (see Fig. 8), and the template image I_{template} and the associated model $\underline{\mathbf{z}}_{\text{template}}$ are discussed in Section 3.3.1.

In summary, for a full segmentation of a multi-object complex, there is first a similarity transformation of the whole complex, then a similarity transform of each object, then for each of the figures in turn (with parent figures optimized before subfigures) first a similarity-like transform that for protrusion and indentation figures respects their being on the surface of their parent, then modification of all parameters of each medial atom. After all of these transformations are complete, there is finally the optimization of the dense boundary vertices implied by the medial stages. Since in this paper we describe only the segmentation of single figure objects, there are three stages beyond the initialization: the figural stage, the medial atom (figural section) stage, and the boundary displacement stage.

For all of the stages with multiple primitives (in the case tested in this paper, the medial atom stage and the boundary stage), we follow the strategy of iterative conditional modes, so the algorithm cycles among the atoms in the figure or boundary in random order until the group converges. The geometric transformation of a boundary vertex modifies only its position along its normal [1 parameter]; the normal direction changes as a result of the shift, thus affecting the next iteration of the boundary transformation.

3.3. *The Optimization Method and Objective Function*

Multiscale segmentation by deformable models requires many applications of optimization of the objective function. The optimization must be done at many scale levels and for increasingly many geometric primitives as the scale becomes smaller. Efficient optimization is thus necessary. We have tried both evolutionary approaches and a conjugate gradient approach to optimization. The significant speed advantages of the conjugate gradient method are utilizable if one can make the objective function void of nonglobal optima for the range of the parameters being adjusted that is guaranteed by the previous scale level. We have thus designed our objective functions to have as broad optima as possible and chosen the fineness of our scale levels and intra-level stages to guarantee that each stage or level

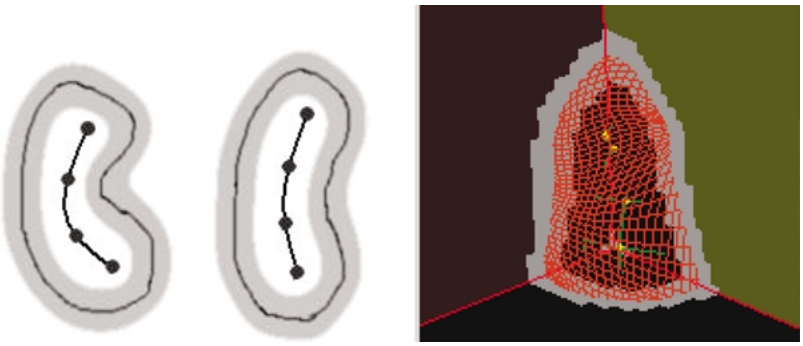

Figure 8. The collar forming the mask for measuring geometry to image match. Left: in 2D, both before and after deformation. Right: in 3D, showing the boundary as a mesh and showing three cross-sections of the collar.

produces a result within the bump-free breadth of the main optimum of the next stage or level.

When the target image is noisy and the object contrast is low, the interstep fineness requirement just laid out requires multiple substages of image blurring within a scale level. That is, at the first substage the target image must be first blurred before being used in the geometry to image match term. At later substages the blurring that is used decreases.

At present the largest scale level involved in a segmentation requires a single user-selected weight, between the geometric typicality term and the geometry to image match term. All smaller scale stages require two user-selected weights, the one just mentioned plus a weight between the parent-to-candidate distance and the neighbor-predictions-to-candidate distance. However, we intend in the future that our objective function be a log posterior probability. When this comes to pass, both terms in the objective function will be probabilistic, as determined by a set of training images. These terms then would be a log prior for the geometric typicality term and a log likelihood for the geometry to image match term. In this situation there is no issue weighting the geometric typicality and geometry to image match terms. However, at present our geometric typicality term is measured in r-proportional squared distances from model-predicted positions and the geometry to image match term is measured in rms-proportional intensity squared units resulting from the correlation of a template image and the target image, normalized by local variability in these image intensities. While this strategy allows the objective function to change little with image intensity scaling or with geometric scaling, it leaves the necessity of setting the relative weight between the geometric typicality term and the geometry to image match term.

The remainder of this section consists of a subsection detailing the geometry-to-image match term of the objective function, followed by a section detailing the boundary displacement stage of the optimization.

3.3.1. The Geometry-to-Image Match Measure. It is useful to compute the match between geometry and the image based on a model template. Such a match is enabled by comparing the template image I_{template} and the target image data I at corresponding positions in figural coordinates, at figural coordinates determined in the model. The template is presently determined from a single training image I_{template}, in which the model $\underline{\mathbf{z}}$ has been deformed to produce $\underline{\mathbf{z}}_{\text{template}}$ by applying the *m-reps* deformation method through the medial atom scale level (level 4) on the characteristic image corresponding to a user-approved segmentation. In our implementation the template is defined only in a mask region defined by a set of figural coordinates, each with a weight of a Gaussian in its figural distance-to-boundary, τ, about the model-implied boundary. The standard deviation of the Gaussian used for the results in this paper is $^1/_2$ of the half-width of the collar. The mask is choosable as a collar symmetrically placed about the boundary up to a user-chosen multiple of r from the boundary (Fig. 8) or as the union of the object interior with the collar, a possibility especially easily allowed by a medial representation. In the results reported here we use a boundary collar mask. The mask is chosen by subdividing the boundary positions affected by the transformation with a fixed mesh of figural coordinates (u, v) and then choosing spatial positions to be spaced along each medial spoke (implied boundary normal) at that (u, v). These along-spoke positions are equally spaced in the figural distance coordinate τ up to a plus or minus a fixed cutoff value $\tau_{\max}$ chosen at modeling time. For the kidney results reported in Section 4, this cutoff value was 0.3, so the standard deviation of the weighting Gaussian in the intensity correlation is 0.15.

The template to image match measure is choosable in our tool from among a normalized correlation measure, with weights, and a mutual information measure, with weights, but for all the examples here the correlation measure has been used and the weight in all mask voxels is unity. The correlation measure that we use is an average, over the boundary sample points, of the along spoke intensity profile correlations at these sample points. For the geometry to correspond to the volume integral of these point-to-corresponding-point correlations, each profile must be weighted by the boundary surface area between it and its neighboring sample points, and the profile must be weighted by its r-proportional length. In addition, as indicated above, we weight each product in the correlation by a Gaussian in τ from the boundary. Also, to make the intensity profiles insensitive to offsets and linear compression in the intensity scale, the template is offset to a mean of zero and both the template and the target image are rms-normalized. The template's rms value is computed within the mask in the training image, and the target image's rms value is computed for a region corresponding to a blurred version of the mask after the manual placement of the model.

In our segmentation program the template is choosable from among a derivative of Gaussian, described more precisely below, and the intensity values in the training image in the region, described in more detail in Section 4.2. In each case the template is normalized by being offset by the mean intensity in the mask and normalized in rms value.

The derivative of Gaussian template for model to image match is built in figural coordinates in the space of the model, i.e., the space of the training image. That is, each along-spoke template profile, after the Gaussian mask weighting, is a derivative of a Gaussian with a fixed standard deviation in the figural coordinate τ, or equivalently an r-proportional standard deviation in Euclidean distance. We choose 0.1 as the value of the standard deviation in τ. Since this template is associated with the target image via common figural coordinates, in effect the template in the target image space is not a derivative of 3D Gaussian but a warped derivative of 3D Gaussian, with the template's standard deviation in spatial terms increases with the figural width.

3.3.2. Boundary Displacement Optimization. The boundary deformation stage is much like active surfaces, except that the geometric typicality term consists not only of a term measuring the closeness of each boundary displacement to that at each of the neighboring boundary positions but also a term measuring the log probability of these displacements in the medially based prior. Since the tolerance of the medially implied boundary is r-proportional, the log Gaussian medially based prior, conditional on the medial estimate, is proportional to the negative square of the r-normalized distance to the medially implied boundary (Chen, 1999). The method of Joshi (2001), with which we complete the segmentation, uses this combined geometric typicality measure, and its boundary to image match measure is a log probability based on the object and its background each having normal intensity distributions.

4. Segmentation Results: Deformed Models

4.1. Segmenting the Kidney from CT; Segmentation Accuracy

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, but a single result using a Gaussian derivative template is shown in Fig. 9.

Extracting the kidney from CT images is challenging under the conditions of the work reported here for radiation therapy treatment planning (RTP). The kidney sits in a crowded soft tissue environment where parts of its boundary have good contrast resolution against surrounding structures but other parts have poor contrast resolution. Also, not too far away are ribs and vertebrae, appearing very light with very high contrast (Fig. 10). The typical CT protocol for RTP involves non-gated slice-based imaging, without breath holding and without injecting the patient with a "dye" to enhance the contrast of the kidney. During the time interval between slice acquisition the kidneys are displaced by respiratory motion, resulting in significantly jagged contours in planes tilted relative to the slice plane (Fig. 10). A combination of partial volume and motion artifacts causes the poles to be poorly visualized or spuriously extended to adjacent slices. Motion and partial volume artifacts degrade the already poor contrast between the kidney and adrenal gland, which sits on top of the kidney.

The single figure *m-rep* used here includes part of the pelvis along with the kidney parenchyma, mimicking segmentation as performed for RTP. The complex architecture of the renal pelvis acts as structure noise for this *m-rep*. When using a Gaussian template that is designed to give increased response at boundaries next

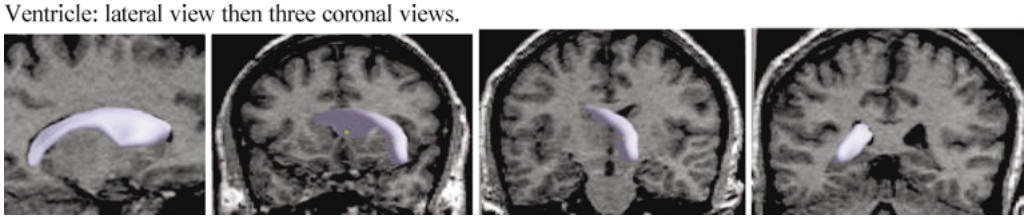

Figure 9. Segmentation results of the lateral horn of a cerebral ventricle at the *m-rep* level of scale (i.e., before boundary displacement) from MRI using a single figure model.

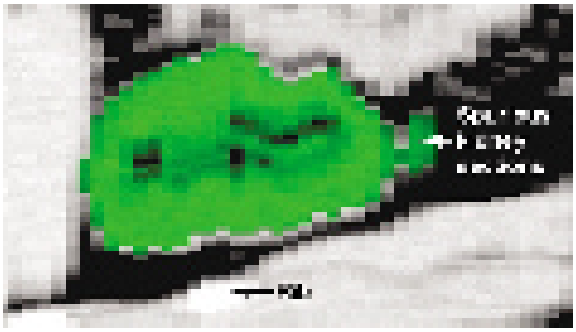

Figure 10. Sagittal plane through a CT of the kidney, used in this study, demonstrating significant partial volume and breathing artifacts. A human segmentation is shown as a green tint. Note the scalloped boundary and spurious sections of the kidney, which were segmented by one of two human raters but excluded by *m-rep* segmentation. Note also the nearby high-contrast rib that can create a repulsive force when a Gaussian derivative template is used.

to which are non-narrow strips of object with intensity lighter than its background, the following behaviors are noted. (1) If the geometric penalty weight is low, the kidney *m-rep* can move inside a sequence of vertebral bodies because the high contrast on only a portion of the model results in a high model to image match value. This is easily prevented by an adequately high weight for geometric typicality. (2) A portion of the implied boundary of the kidney *m-rep* can move to include a rib, as a result of the high contrast of the rib. (3) A portion of the implied boundary of the kidney *m-rep* can move to include part of the muscle. (4) The boundary at the liver, appearing with at most texture contrast, does not attract the implied *m-rep* boundary, with the result that the geometric typicality term makes the kidney not quite follow the kidney/liver boundary. In other organs, parts of the edge with high contrast of opposite polarity to other parts would repel the *m-rep* boundary. Avoiding some of these difficulties necessitated replacing the first post-initialization similarity transform by a similarity transform augmented by an elongation for the main figure. Despite these challenges, segmentation using a Gaussian derivative template, built on a single right kidney, is successful when compared against human performance.

An example deformation sequence is shown in Fig. 11, showing the improved segmentation at each stage. Results of a typical kidney segmentation are visualized in Fig. 12. Comparisons between *m-rep* segmentation (rater C) and human segmentation

Table 2. Comparison of *m-reps* segmentation to manual segmentation for six examples selected from 24 kidneys (twelve kidney pairs). Distances are in cm. The examples span the range, from best to worst, of human-*m-rep* volume overlap.

Kidney code	Volumes compared	Vol. Overlap	1st-Qtl*	2nd-Qtl*	3rd-Qtl*	⟨Dist⟩	Haussdorf distance
639 R	A–B	0.947	0.000	0.200	0.200	0.117	0.693
	A–C	0.940	0.000	0.000	0.200	0.124	1.342
	B–C	0.940	0.000	0.200	0.200	0.123	1.183
646 L	A–B	0.935	0.000	0.200	0.200	0.133	1.149
	A–C	0.939	0.000	0.000	0.200	0.108	0.917
	B–C	0.940	0.000	0.000	0.200	0.103	1.077
634 L	A–B	0.953	0.000	0.000	0.200	0.103	0.600
	A–C	0.903	0.000	0.200	0.200	0.182	0.894
	B–C	0.899	0.200	0.200	0.283	0.191	0.849
633 L	A–B	0.951	0.000	0.000	0.200	0.106	0.721
	A–C	0.909	0.000	0.200	0.200	0.162	1.217
	B–C	0.896	0.000	0.200	0.283	0.189	1.131
637 R	A–B	0.957	0.000	0.000	0.200	0.079	0.566
	A–C	0.843	0.000	0.200	0.400	0.272	1.789
	B–C	0.840	0.000	0.200	0.400	0.280	1.897
635 L	A–B	0.950	0.000	0.000	0.200	0.097	0.800
	A–C	0.807	0.000	0.200	0.400	0.307	2.209
	B–C	0.808	0.000	0.200	0.400	0.314	2.272

*Quartile columns give the surface separation associated with each quartile, e.g., an entry of.200 in the 2nd-Qtl column means that 50% of all voxels on the surfaces of the compared segmentations are separated by no more than .200 cm. (.200 cm is the smallest unit of measurement). ⟨Dist⟩ is the median distance between the surfaces of the compared segmentations.

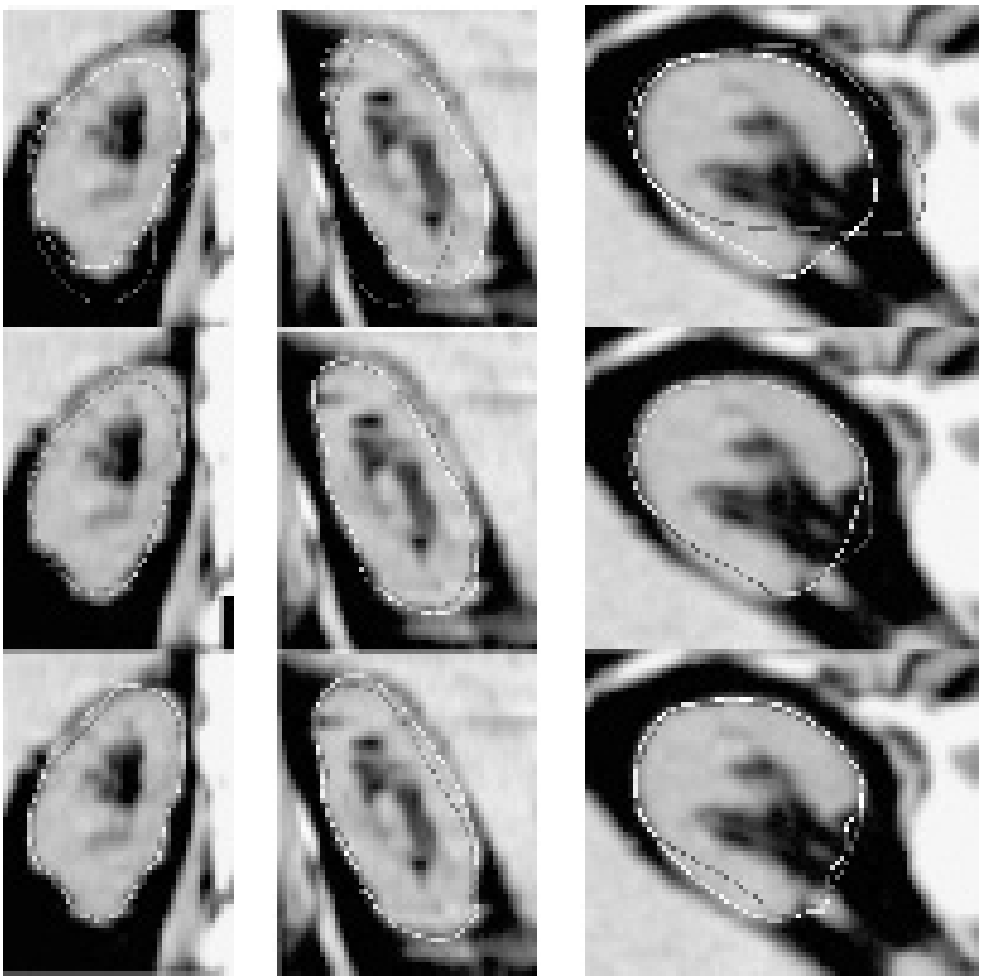

Figure 11. Stage by stage progress: all rows, from left to right, show results on Coronal, Sagittal and Axial CT slices. Each row compares progress through consecutive stages via overlaid grey curves to show the kidney segmentation after stage N vs. white curves after stage $N+1$. Top row: stages are the initial position of the kidney model vs. the figural similarity transform plus elongation. Middle: the similarity transform plus elongation vs. medial atom transformations. Bottom: medial atom transformations vs. 3D boundary displacements.

(raters A and B) using our evaluation system Valmet (Gerig, 2001) are given in Table 2 and Figs. 13–15. Comparisons are given for 12 kidney pairs (12 right kidneys and corresponding left kidneys. Manual segmentation by A and B was performed slice-by-slice using the program, Mask (Tracton, 1994). Within-slice pixel size was approximately 1 mm, and slice thickness varied image to image between 3 mm and 8 mm. Images were resampled for *m-rep* segmentation to yield isotropic 2 mm voxels. At the comparison stage using Valmet the segmented volumes, originally represented by sets of contours for humans and as 3D surfaces for *m-reps*, were scan converted to produce voxelized (2 mm voxels) representations.

The median volume overlap for human segmentations, as measured by the overlap volume divided by the union of the two volumes being compared, is 94% ($\sigma = 1.7\%$, min = 90%, max = 96%). The mean surface separation, averaged over all kidneys, is 1.1 mm ($\sigma = 0.3$ mm); the mean surface separation for a given kidney is defined in terms of closest points, i.e., as

$$\frac{1}{2}\left[\frac{1}{N_1}\left(\sum_{i=1}^{N_1}\min_{j=1,\ldots,N_2}\left|\underline{y}_i^1-\underline{y}_j^2\right|\right)+\frac{1}{N_2}\left(j\sum_{i=1}^{N_2}\min_{i=1,\ldots,N_1}\left|\underline{y}_i^1-\underline{y}_j^2\right|\right)\right],$$

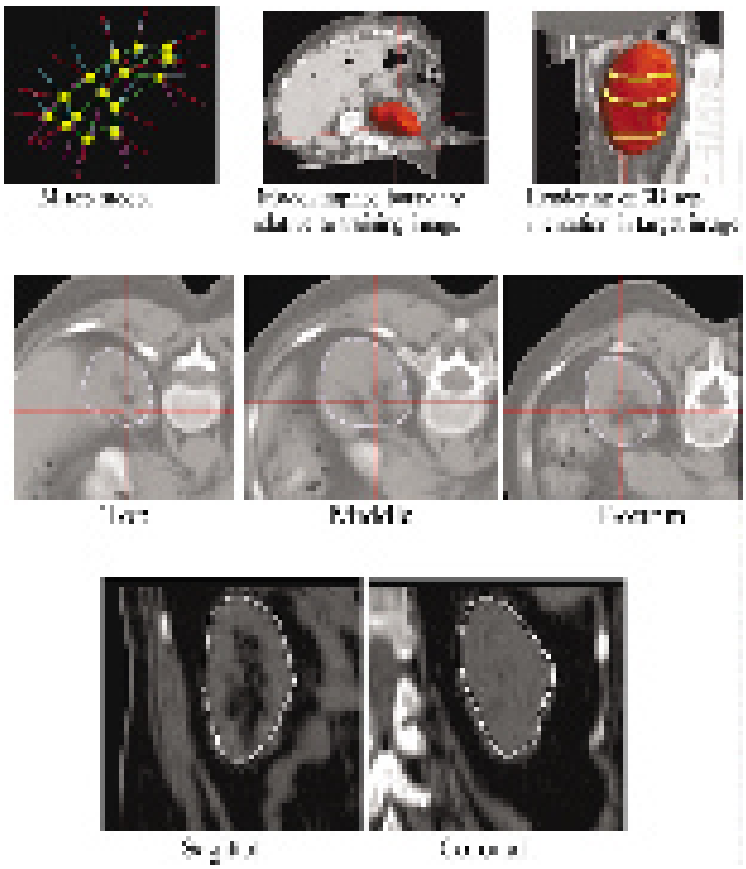

Figure 12. Kidney model and segmentation results. (Segmentation results at the *m-rep* level of scale (i.e., before boundary displacement) on kidneys in CT using a single figure model. The three light curves on the rendered *m-rep* implied boundary in the 3D view above right show the location of the 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).

where N_1 and N_2 are the respective numbers of boundary voxels in the two kidneys being compared and $\underline{y}_i^1$ and $\underline{y}_j^2$ are the coordinates of the boundary voxel centers of the respective kidneys. The median volume

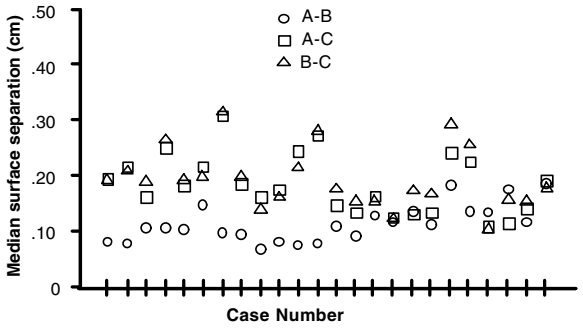

Figure 13. Scattergram of median surface separations for all kidneys.

overlap between human and *m-rep* segmentations is 89% ($\sigma = 3.4\%$, min $= 81\%$, max $= 94\%$), and the mean surface separation, averaged over all kidneys, is 1.9 mm ($\sigma = 0.5$ mm). The distinction between *m-rep* to human comparison and human to human comparison is statistically significant for average distance and maximum distance metrics, though from a clinical point of view the distance differences are small. The average surface separations (over each boundary vertex point on the reference segmentation and, for each, the closest point on the segmentation being evaluated) for human-human and human-*m-rep* comparisons correlate well with the pixel dimensions at segmentation with MASK and *m-reps* respectively. Because Valmet measures offsets and overlaps only to the closest voxel (Table 2), the image resampling and scan conversion steps introduced a bias against *m-reps* which is thought to account for most of the difference between human-human

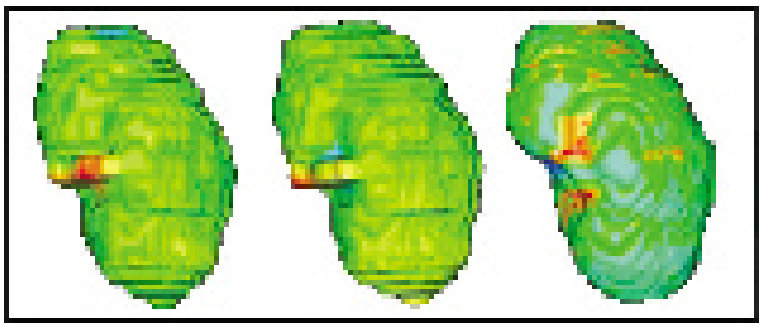

Figure 14. Valmet pairwise comparisons for a left kidney. The comparison result is color-coded on a reference surface selected from human (A or B) and *m-rep* segmentations (see Fig. 15). Green represents a subvoxel surface correspondence between the two compared segmentations. Red represents a section where the surface of the reference segmentation is outside the compared surface. Blue represents a section where the surface of the reference segmentation is inside compared surface. Left: Reference shape from human B, color coding from human A. Middle: Reference shape from human B, color coding from *m-reps*. Right: Reference shape from *m-reps*, color coding from human B. In this case the volume overlap for A and B was 93.5% and the *m-rep* overlap was 94.0% with both A and B.

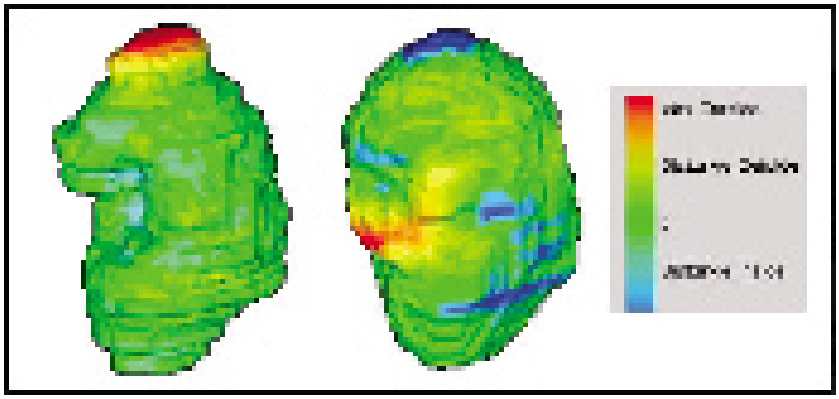

Figure 15. Valmet comparisons for a kidney with significant motion artifacts (see Fig. 10), reflecting human segmentations' preservation of artifactual scalloping vs. the *m-rep* segmentations' yielding a smooth surface. Left: Reference shape from human A, color coding from human B. In this case both A and B contoured spurious sections at the top of the kidney, but rater A contoured one additional slice. Center: Reference shape from *m-reps*, color coding from human A. Right: Color coding scheme.

and human-*m-rep* metrics. Also, consistencies in manual segmentation also favorably bias human-human comparison. For example 2D contouring tends to preserve jagged edges (Fig. 10) caused by motion artifacts while 3D *m-rep* surfaces tend to smooth the edges (Figs. 14–15).

Therefore the real median boundary accuracy is subvoxel and the human-*m-rep* overlap percentages are understated. We judge the best *m-rep* segmentations to be at least as good as the most careful manual slice by slice segmentation done for RTP in routine clinical practice. Suboptimal performance is associated primarily with insufficient training and shortcomings of the Gaussian derivative template (Fig. 16).

4.2. Templates Made from a Training Image, and Segmenting the Hippocampus

Objects that appear without high contrast over most of their boundary cannot be extracted using a Gaussian derivative template. For these the pattern of image intensity profiles in the boundary-centered mask region as a function of (figurally defined) boundary position are frequently characteristic. As a result an intensity template as function of figural position drawn from one or more training images can be used as the correlate for the target image in producing a geometry to image match measurement. Our *m-reps* based segmenter can use such training intensity templates. In what we report

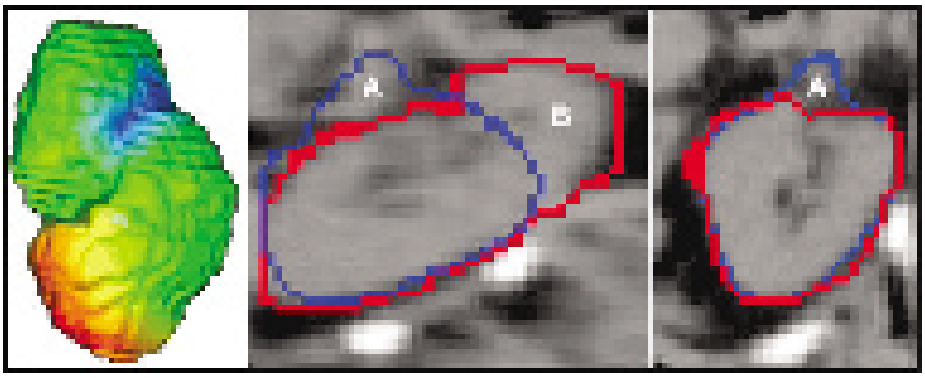

Figure 16. Correctable *m-rep* failure mistakenly included in our analysis (worst case in Table 1). Left: Valmet comparison with reference shape from human B, color coding from *m-reps*. Center: Sagittal plane showing *m-rep* (blue) and human (red) surfaces. Two problems mentioned in the text are illustrated. In the region labeled "A" the *m-rep* model deformed into structures related to the kidney pelvis that were poorly differentiated from the kidney parenchyma. In the region labeled "B" the *m-rep* model did not elongate fully during the first transformation stage. Right: Transverse plane illustrating the deformation of the *m-rep* model into peri-pelvic structures in region A. Even in this case there is close correspondence between human and *m-rep* contours excluding regions A and B. After more careful user-guided initialization a successful *m-rep* segmentation was obtained for this kidney, but those results were omitted in the analysis.

below, we use a template drawn from a single training image with the *m-rep* fitted by the approach of this paper into a slightly blurred version of the human's binary hand segmentation for that training image. This *m-rep* fitted into the training segmentation also formed the model used in these studies on four other cases.

We first report on using the image intensity profile method to fit a hippocampus *m-rep* from the training case into the slightly blurred versions (Gaussian blurring with std. dev. = 1 voxel width) of binary segmentations for the four other hippocampi. This process succeeded in all four cases (Fig. 17), suggesting two conclusions. First, one can use such fitting of *m-reps* to already segmented cases to provide statistical description of the geometry via the statistics of the fitted *m-reps* (Styner, 2002). Second, the method of 3D deformable *m-rep* optimization via intensity templates is geometrically sound.

We move on to the extraction of the left hippocampus from MRI, an extraction that is very challenging for humans and has great variability across human segmenters. The pattern of intensities across the boundary of the left hippocampus is characteristic, to the extent that the structures abutting or near the hippocampus follow a predictable pattern, with each structure having its typical intensity in MRI. However, because of the variability of this structural pattern and thus the variability of the intensity pattern, and because of the indistinct contrast at both the tip and the tail of the hippocampus, hippocampal segmentation from MRI provides a major challenge for automatic segmentation of the hippocampus as a single object. The poor contrast makes it necessary to omit the boundary displacement stage from the segmentation and so to stop after the medial atom stage. Our method beginning from a manually placed hippocampus model produced a reasonable, hippocampus-shaped segmentation overlapping with the human segmentation in 3 of the 4 cases (Fig. 17), and in the remaining case, the structural pattern around the hippocampus was so different as to make the method fail. Moreover, in only 1 of the 3 semi-successful cases was the segmentation credible to a human rater. However, in each of those 3 cases when the same segmentation was begun from the *m-rep* fitted to the human segmentation, it produced a successful segmentation with a geometry to image match measure higher than that achieved when starting from the manually placed model. This result suggests that when the contrast in the images is weak the optimization technique needs to be changed to avoid local optima or that a better measure for geometry-to-image match must be found or that multiple object geometry must be used in the process. We are investigating all three of these paths. In regard to geometry-to-image match, we judge that the intensity template must really be statistical, reflecting the range of patterns in a family of training images, and that is indeed a direction that, long since, we have intended our method development to go. Our figural coordinate system and correlation method is exactly what is needed for such segmentation to be based on principal components analysis, in a way analagous

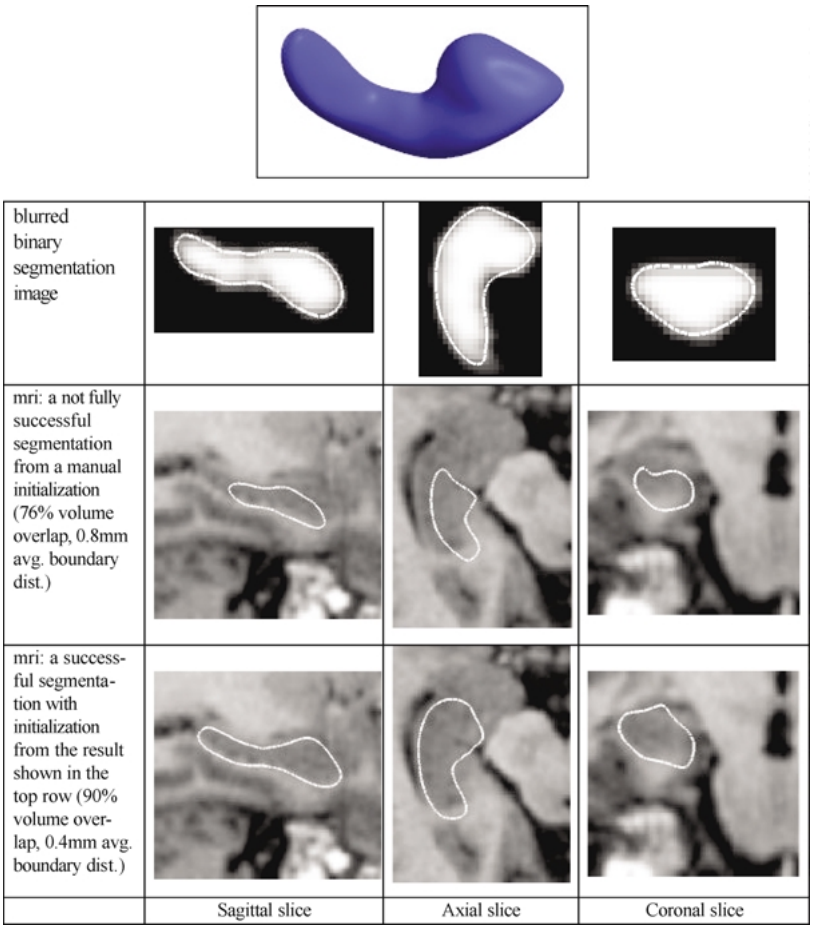

Figure 17. Hippocampus results using training intensity matches, for one of the three target images with typical results. The top image shows the m-rep-segmented hippocampus from the blurred binary segmentation image. Each row in the table shows in three intersecting triorthogonal planes the target image overlaid with the implied boundary of the *m-rep* hippocampus segmentation using the image template match. The top row shows the segmentation from the blurred binary image produced from a human segmentation. The middle row shows the corresponding segmentation using the mri image as the target and an intialization from a manual placement of the model determined from training image. The bottom row shows the segmentation of the same target mri with both the initialization and the model being the segmentation result on the blurred binary image for that case.

to the boundary point based segmentation methods of Cootes and Taylor (1993).

We have found that the figural stage of segmentation using a training image template is better when the training image intensity template as well as the target image is rather blurred with a standard deviation of a significant fraction of the average r of the object. This causes small differences between the template image and the target image in the relative positioning of the object sought to surrounding structures to have only a small effect. Thus when we apply the technique to kidney segmentation, the fact that, for example, in the training image the kidney nearly abuts the spinal column but is somewhat separated from it in the target image causes that portion of the kidney segmented from the target image to move toward spine, but only to a small degree if blurred images are used. With the hippocampus, this same blurring approach was used, but since the hippocampus is much narrower than the kidney, the level of blurring used was correspondingly smaller.

4.3. *Speed of Computation*

The speed of a 3D segmentation is 2–4 minutes for a 15-atom kidney model on a Pentium 4, 1.8 GHz Dell

Inspiron 8200 laptop computer, with

- preprocessing computations taking less than 1 second,
- the object similarity transform plus elongation stage taking under 1 second per iteration and on the average requiring 8 iterations for a total time of 2 seconds to determine the object similarity transform,
- the atom transformations taking on the average approximately 45 seconds per iteration through all of the atoms for the kidney, with the time per iteration roughly proportional to the surface area of the object, and the number of iterations required for the kidney being 2 to 3, and
- the boundary deformation stage taking approximately 3 seconds.

While the method's speed has already benefited strongly from moving much of the computation from the deformation stage to the model building stage, there is still much room for speedup by more medial levels of coarse to fine and just by more careful coding.

5. Discussion and Conclusions

The main contribution of this paper is the detailing of a method for 3D image segmentation that uses the medial model representation called *m-reps* both to capture prior knowledge of object geometry and as the basis of measurement of model to image match. There are number of other contributions. We have laid out how to base the geometry in 3D of a figure on a new form of medial atom that allows a spatially sampled representation while carrying a local ruler and compass. We have described a way of deriving the continuous implied boundary of a single-figure object from that representation. We have described a means of calculating a parametrization of space in terms of medially relative positions and signed r-proportional distance from this implied boundary and the usefulness of this parameterization in computing geometric typicality via correspondence at the object boundary and computing geometry to image match via this correspondence and the correspondence of distances from the boundary in r-proportional terms. This space parametrization also allows the interior of the object to be distinguished as associated with a particular region or in the blend region. An additional contribution is having a set of object and figure based scale levels and the means of optimal deformation at multiple scale levels by considering the model information at each relevant scale. The initial successes of the segmentation method provide evidence of the usefulness of these ideas. M-rep based visualizations of an object in both 3D and versus the 3D image relative to a medial atom were also described.

Our quantitative validation on kidney segmentation is encouraging: robust and accurate enough for clinical use. A recent study on segmenting the kidney from CT suggests that the result is very stable to variations in the manual initialization. However, achieving full human accuracy remains a goal. Controlled, quantitated validations on other objects are needed. The results of such a study will be reported in a future paper. We find the results so far, together with the theoretical strengths of the deformable *m-reps* method, encouraging evidence that this approach will increase the maximum robustness achievable in 3D segmentation of anatomic objects or that it can do such segmentation at a given level of robustness faster than alternative methods.

At the same time, the method sometimes fails to segment variants of an object that has significant variability in shape over the population. An example, is the prostate, which varies across individuals in the degree to which it bends to "saddlebag" around the rectum. In situations of large saddlebagging, the ends of the prostate are so distant from the model that the distance-based geometric penalty prevents convergence to the fully bent prostate at its ends. The solution, we guess, lies in extending the transformation at each scale level to include not only the transformation listed in Table 1 but also a first principal component transformation of the collection of medial atoms associated with that scale level, with the first principal component eigenmode being computed from a training set.

We have completed the programming for an early version of computing optimal subfigural geometric transformations, but have not put the program in the form we are ready to test. Each subfigural transformation iteration will require roughly the fraction of the time for the whole object iteration that the subfigure's boundary is of the whole object's boundary. Moreover, the number of iterations can be expected to be distinctly lower than for the whole object, since the intialization, based on the object's transformation, will typically be better than that for the object as a whole. Thus, each subfigure transformation will typically take a few tens of seconds.

We have also implemented a means of deforming multiple objects in a way that the position of one object's implied boundary in another object's figural coordinates allows rewarding its relationship (e.g., abutment) to that controlling object (Fletcher, 2002). We need to continue to test this method for the segmentation of multiple abutting objects and thus to make the segmentation method applicable to whole sections of the body. With that development our segmentations will form a means of elastic registration between an atlas and a patient's target image for the whole imaged body section. Both the multifigure object segmentation and multiobject segmentation methods and their validation will be reported in later papers.

A very pervasive improvement will lie in replacing both the geometric typicality measures and the geometry to image match measures by statistical (log probability) objective functions based on multiple training images. We expect improved performance by the use of a geometric typicality term that is based not simply on distances from a model but instead reflecting normal variability in a training set, and a geometry to image match term reflecting image intensity variations relative to the geometric model; moreover, that would make it unnecessary to interactively choose a weight for the geometric typicality term relative to the model to image match term. The fact that the geometry to image match module already computes the template correlations that are used in computing the dot product between an eigenmode and the target image within the masked region will make this step straightforward, once the training process is developed. Another direction for future development is the use of mechanical geometric typicality measures in place of statistical ones when the transformation is intrapatient rather than between an atlas model and a patient. Here the medial sampling provides promise for strides in efficiency of mechanical modeling.

Another matter for the future is the evaluation of the usefulness of multiple levels of sampling at the medial stages (3 and 4) of the segmentation. Coarse spacing will produce efficient operation but will not accommodate small radii of curvature relative to the object width. The accommodation of patches requiring these larger curvatures can be done efficiently at the finer scale medial substages, once the larger scale optimization has been accomplished. We will need to study when smaller scale wrinkles should be accommodated at a medial substage and when they can be better represented by boundary displacements.

Acknowledgments

We appreciate advice from Valen Johnson and Stephen Marron on statistical matters and help from Guido Gerig, Joshua Stough, Martin Styner, and Delphine Bull. This work was done under the partial support of NIH grants P01 CA47982 and EB02779, NSF SGER Grant CCR-9910419, and a fellowship from the Link Foundation. A gift from Intel Corp. provided computers on which some of this research was carried out.

Notes

1. The reader unfamiliar with the literature on the mathematics of medial loci can find the relevant definitions, overview, and references in Pizer (2003) in this issue.
2. This sheet has both a boundary and singular points in its interior, so in mathematical terminology it is a "stratified set". We will use the phrase "medial sheet" to refer to this stratified set.

References

Amenta, N., Bern, M., and Kamvysselis, M. 1998. A new Voronoi-based surface reconstruction algorithm. *Computer Graphics Proceedings*, Annual Conference Series, 1998, ACM SIGGRAPH: 415–422.

Attali, D., Sanniti di Baja, G., and Thiel, E. 1997. Skeleton simplification through non significant branch removal. *Image Processing and Communications*, 3(3/4):63–72.

Bittar, E., Tsingos, N., and Gascuel, M. 1995. Automatic reconstruction of unstructured 3D Data: Combining a medial axis and implicit surfaces. *Computer Graphics Forum (Eurographics '95)*, 14(3):457–468.

Blum, H. 1967. A transformation for extracting new descriptors of shape. In *Models for the Perception of Speech and Visual Form*. W. Wathen-Dunn, (Ed.). MIT Press: Cambridge MA, pp. 363–380.

Burbeck, C.A., Pizer, S.M., Morse, B.S., Ariely, D., Zauberman, G., and Rolland, J. 1996. Linking object boundaries at scale: A common mechanism for size and shape judgments. University of North Carolina, Computer Science Department, Technical Report TR94–041. *Vision Research*., 36(3):361–372.

Catmull, E. and Clark, J. 1978. Recursively generated B-spline surfaces on arbitrary topological meshes. *Computer Aided Design*, 10:183–188.

Chen, D.T., Pizer, S.M., and Whitted, J.T. 1999. Using multiscale medial models to guide volume visualization. Tech Report TR99–014, Dept. of Comp. Sci., Univ. of NC at Chapel Hill.

Cootes, T.F., Hill, A., Taylor, C.J., and Haslam, J. 1993. The use of active shape models for locating structures in medical images. In *Information Processing in Medical Imaging*, H.H. Barrett and A.F. Gmitro (Eds.). Lecture Notes in Computer Science, vol. 687.Springer-Verlag, Heidelberg, pp. 33–47.

Damon, J. 2002. Determining the geometry of boundaries of objects from medial data. Internal report, Dept. of Mathematics, Univ. of NC. Available at webpage http://midag.cs.unc.edu/pubs/papers/Damon_SkelStr_III.pdf.

Delingette, H. 1999. General object reconstruction based on simplex meshes. *International Journal of Computer Vision*, 32:111–146.

Fletcher, P.T., Pizer, S.M., Gash, A.G., and Joshi, S. 2002. Deformable *m-rep* segmentation of object complexes. Proc. Int. Symp. Biomed. Imaging 2002, a compact disc, IEEE. Also available at webpage http://midag.cs.unc.edu.

Fritsch, D.S., Pizer, S.M., Yu, L., Johnson, V., and Chaney, E.L. 1997. Localization and Segmentation of Medical Image Objects using Deformable Shape Loci. In *Information Processing in Medical Imaging (IPMI)*, J. Duncan and G. Gindi (Eds.), Lecture Notes in Computer Science, vol. 1230. Springer-Verlag, Heidelberg, pp. 127–140,

Furst, J.D. 1999. Height Ridges of Oriented Medialness. Ph.D. Thesis, Univ. of NC at Chapel Hill, 1999.

Gerig, G., Jomier, M., and Chakos, M. 2001. Valmet: A new validation tool for assessing and improving 3D object segmentation. In *Proc. MICCAI 2001*, LNCS 2208, Springer, pp. 516–523.

Grenander, U. 1981. *Regular Structures: Lectures in Pattern Theory, vol. III*, Springer-Verlag, 1981.

Igarashi, T., Matsuoka, S., and Tanaka, H. 1999. Teddy: A sketching interface for 3D freeform design. *Computer Graphics Proceedings*, Annual Conference Series, 1999, ACM SIGGRAPH, pp. 409–416.

Joshi, S., Pizer, S., Fletcher, P.T., Thall, A., and Tracton, G. 2001. Multi-scale 3-D deformable model segmentation based on medial description. *Information Processing in Medical Imaging 2001* (IPMI '01). Lecture Notes in Computer Science, vol. 2082, Springer, pp. 64–77.

Kelemen, A., Szekely, G., and Gerig, G. 1999. Three-dimensional model-based segmentation. *IEEE-TMI*, **18**(10):828–839.

Koenderink, J.J. 1990. *Solid Shape*. MIT Press: Cambridge, MA.

Lee, T.S., Mumford, D., and Schiller, P.H. 1995. Neuronal correlates of boundary and medial axis representations in primate striate cortex. *Investigative Ophthalmology and Visual Science* Annual Meeting, abstract #2205.

Leyton, M. 1992. *Symmetry, Causality, Mind*. MIT Press: Cambridge, MA.

Liu, G.S., Pizer, S.M., Joshi, S., Gash, A.G., Fletcher, P.T., and Han, Q. 2002. Representation and segmentation of multifigure objects via *m-reps*. University of North Carolina Computer Science Department technical report TR02–037, at webpage http://www.cs.unc.edu/Research/MIDAG/pubs/papers/.

Marr, D. and Nishihara, H.K. 1978. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc Royal Soc*, Series B, 200:269–294.

Markosian, L., Cohen, J.M., Crulli, T., and Hughes, J. 1999. Skin: A constructive approach to modeling free-form shapes. *Computer Graphics Proceedings*, Annual Conference Series, 1999, ACM, SIGGRAPH, pp. 393–400.

McAuliffe, T., Eberly, D., Fritsch, D.S., Chaney, E.L., and Pizer, S.M. 1996. Scale-space boundary evolution initialized by cores. In *Visualization in Biomedical Computing (VBC '96) Proceedings*, H. Hoehne and R. Kikinis (Eds.), Lecture Notes in Computer Science, vol. 1131. Springer-Verlag: Heidelberg, pp. 174–182.

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

Nackman, L.R. and Pizer, S.M. 1985. Three-dimensional shape description using the symmetric axis transform, I: Theory. *IEEE Trans. PAMI*, 7(2):187–202.

Pizer, S.M., Fritsch, D.S., Yushkevich, P., Johnson, V., and Chaney, E.L. 1999. Segmentation, registration, and measurement of shape variation via image object shape. *IEEE Transactions on Medical Imaging*, 18:851–865.

Pizer, S.M., Siddiqi, K., Székely, G., Damon, J.N., and Zucker, S.W. 2003. Multiscale medial loci and their properties. *Int. J. Comp. Vis.*, 55(2/3):155–179.

Pizer, S.M., Fletcher, P.T., Thall, A., Styner, M., Gerig Guido, and Joshi Sarang 2002. Object models in multiscale intrinsic coordinates via *m-reps*. *Image and Vision Computing* 21:5–15.

Siddiqi, K., Bouix, S., Tannenbaum, A., and Zucker, S. 1999. The Hamilton-Jacobi skeleton. *ICCV'99 Proceedings, IEEE*, 828–834.

Singh, K. and Fiume, E. 1998. Wires: A geometric deformation technique. In *Computer Graphics Proceedings*, Annual Conference Series, 1998, ACM SIGGRAPH, pp. 405–414.

Staib, L.H. and Duncan, J.S. 1996. Model-based deformable surface finding for medical images. *IEEE Trans. Med. Imaging*, 15(5):1–12.

Styner, M. and Gerig, G. 2001. Medial models incorporating object variability for 3D shape analysis. *Information Processing in Medical Imaging 2001 (IPMI '01)*. Lecture Notes in Computer Science, Springer, 2082:502–516.

Styner, M., Gerig, G., Lieberman, J., Jones, D., and Weinberger, D. 2002. Statistical shape analysis of neuroanatomical structures based on medial models *Medical Image Analysis*, to appear.

Thall, A. 2003. Deformable solid modeling via medial sampling and displacement subdivision. UNC Computer Science Dept., Dissertation to be defended August 2003, at webpage http://www.cs.unc.edu/Research/MIDAG/pubs/phd-thesis.

Tracton, G., Chaney, E.L., Rosenman, J.G., and Pizer, S.M. 1994. MASK: Combining 2D and 3D Segmentation methods to enhance functionality. Math. Methods Med. Imaging III, F.L. Bookstein et al. (Eds.), SPIE Proc. 2299:98–109.

Vermeer, P.J. 1994. Medial axis transform to boundary representation conversion, Ph.D. Thesis, Purdue University, May 1994.

Yushkevich, P., Pizer, S.M., and Culver, T. 1999. Statistical object shape via a medial representation. Tech Report TR00–002, Dept. of Comp. Sci., Univ. of NC at Chapel Hill.

Wyvill, G., McPheeters, C., and Wyvill, B. 1986. Data Structures for soft objects. *Visual Computer*, 2(4):227–234.