# Multi-scale 3-D Deformable Model Segmentation Based on Medial Description

Sarang Joshi, Stephen Pizer, P. Thomas Fletcher, Andrew Thall, and Gregg Tracton

Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill, Chapel Hill NC 27514
joshi@radonc.unc.edu

**Abstract.** This paper presents a Bayesian multi-scale three dimensional deformable template approach based on a medial representation for the segmentation and shape characterization of anatomical objects in medical imagery. Prior information about the geometry and shape of the anatomical objects under study is incorporated via the construction of exemplary templates. The anatomical variability is accommodated in the Bayesian framework by defining probabilistic transformations on these templates. The modeling approach taken in this paper for building exemplary templates and associated transformations is based on a multi-scale medial representation. The transformations defined in this framework are parameterized directly in terms of natural shape operations, such as *thickening* and *bending*, and their location. Quantitative validation results are presented on the automatic segmentation procedure developed for the extraction of the kidney parenchyma-including the renal pelvis-in subjects undergoing radiation treatment for cancer. We show that the segmentation procedure developed in this paper is efficient and accurate to within the voxel resolution of the imaging modality.

## 1 Introduction

Modern anatomic imaging technologies are enabling extremely detailed study of anatomy, while the development of functional imaging modalities are providing detailed *in vivo* associated information regarding the physiological function. While modern imaging modalities provide exquisite imagery of the anatomy and its function, automatic segmentation of these images and the precise quantitative study of the biological variability exhibited in these images continues to pose a challenge. In this paper we present a multi-scale medial framework based on deformable templates[8],[5],[19] for the automatic extraction and analysis of the shape of anatomical objects from the brain and abdomen, imaged respectively via MRI and CT. The multi-scale deformable template approach is based on the medial axis representation of objects first proposed by Blum [3] for studying shape. The approach presented herein is an extension of the early work by Pizer[14] and Firtsch[7] in 2D on deformable medial representation of objects.

We adopt a Bayesian approach of incorporating prior knowledge of the anatomical variations and the variation of the imaging modalities. Following the deformable templates paradigm, we incorporate prior information about the geometry and shape of the anatomical objects under study via the construction of exemplary templates. The infinite anatomical variability is accommodated in the Bayesian framework by defining probabilistic transformations on these templates[8]. The segmentation problem in this paradigm is that of finding the transformation $S$ of the template, that maximizes the posterior,

$$P(S|data) \propto P(data|S)P(S) \ ,$$

where $P(S)$ is the prior probability function capturing prior knowledge of the anatomy and its variability, and $P(data|S)$ is the data likelihood function capturing the image data-to-geometry relationship. For efficiency of implementation we equivalently maximize the log-posterior given by

$$LogP(S|data) = LogP(data|S) + LogP(S|data) \ , \text{up to an additive constant.}$$

The modeling approach taken in this paper for building exemplary templates and associated transformations is based on a multi-scale medial representation. The transformations defined in this framework are parameterized directly in terms of natural shape operations, such as *thickening* and *bending*, and their location.

This multi-scale approach has many stages of scale, at each of which the geometric primitives are intuitive for that scale and have the property that their spacing is comparable to the linear measure of the size of space (modeling aperture) that they summarize. This leads to a spatial tolerance that successively decreases with scale level. A Markov Random Field approach, described in detail in [15] is used to defining the energetics of the log probabilities needed for the posterior. The log probabilities at a given scale are not only conditioned on a neighborhood at that scale, but conditioned on the result of the next larger scale. The posterior at each scale can then be separately optimized successively decreasing the scale. This approach not only has benefits in providing an intuitive description of the shape and its deformation but also yields a method that, theoretically compared to most other deformable models approaches, can be expected to be distinctly more efficient for any level of accuracy, by being O(N) in the number of geometric primitives at the smallest scale (typically voxels or voxel-sized boundary tiles).

The multi-scale nature of our approach allows for the investigation of these properties at various scales from the coarse scale of entire body sections to the fine scale on the order of the resolution of the imaging modality. The intuitiveness derives from the ability to have many of the levels of scale describe medial properties. In addition, the size properties derived from medial description allow the creation of natural levels of scale each suited for shape description at that scale level. The next two sections discuss the medial representation of objects. Section 3 discusses the deformation of models to fit image data and the geometric measures used in the log prior term measuring geometric typicality. Section 4

discusses the log likelihood term measuring the match of a deformed model to a target image, and Section 5 gives segmentation results using this method.

## 2 Medial Representation of Objects.

Many authors in image analysis, geometry, human vision, computer graphics, and mechanical modeling have come to the understanding that the medial relationship between points on opposite sides of a figure is an important factor in the objects shape description. Biederman [1], Marr [12], Burbeck [4], Leyton [10], and others have produced psychophysical and neurophysiological evidence for the importance of medial relationships (in 2D projection) in human vision. The medial geometry has also been explored in 3D by Nackman [13], and Siddiqi [16], and medial axis modeling techniques have been applied by many researchers, including Bloomenthal [2], Wyvill [18],Thall[17], Igarashi [9] and Markosian [11]. Of these, Bloomenthal and Wyvill provided skeletal-based soft-objects; Igarashi used a medial spine in 2D to generate 3D surfaces from sketched outlines; and Markosian used implicit surfaces generated by skeletal polyhedra.

Our representation, described in [Pizer 1999], expands the notion of medial relations from that of a medial atom implying boundaries by including of a width-proportional tolerance and by using a width-proportional sampling of the medial manifold in place of a continuous representation. The advantages, relative to the ideas of medial axis descended from Blum [1967], are in representational and computational efficiency and in stability with respect to boundary perturbation. Associating a tolerance with the boundary position provides opportunities for stages of the representation with successively smaller tolerance. Representations with large tolerance can ignore detail and focus on gross shape, and in these large-tolerance stages, discrete sampling can be coarse, resulting in considerable efficiency of manipulation and presentation. Smaller-tolerance stages can focus on retirements of the larger-tolerance stages and thus more local aspects.

The medial representation used in this paper called m-rep, is based on a hierarchical representation of linked figural models, defined at coarse scale by a hierarchy of figures protrusions, indentations, neighboring figures, and included figures which represent solid regions and their boundaries simultaneously. The linked collection of figural components imply a fuzzy, i.e., probabilistically described boundary position with a width-proportional tolerance. At small scale these figural boundaries are made precise by displacing a dense sampling of the m-rep implied boundary. A model for a single figure is made from a net, (a mesh or a chain) of medial atoms; each atom describing not only a position and width, but also a local figural frame giving figural directions, and an object angle between opposing, corresponding positions (medial involutes) on the implied boundary. A figure can be expressed as a sequence over scale of nets, implying successively refined (smaller tolerance) versions of the figural boundary.

### 2.1 Single figure description via m-rep.

We now describe the representation of single figural forms. Detailed motivation and description can be found in Fletcher [6]. Our representation is based on the notion of medial involutes of Blum [1967] and starts with a parameterization of a medial atom $\mathbf{m}$ that locally implies opposing figural boundaries as illustrated in Fig. 1. The medial atom $\mathbf{m}$ by itself not only implies two opposing sections of boundary, but as well the solid region between them. Medial atoms on the interior of the medial manifold are defined as a four tuple $\mathbf{m} = \{\mathbf{x}, r, \boldsymbol{F}, \theta\}$, consisting of:

1. $\mathbf{x} \in I\!R^3$, the skeletal position,
2. $r \in I\!R^+$, the local width defined as the distance from the skeletal position of two or more implied boundary positions,
3. $\mathbf{F} \in SO(3)$ the local frame parameterized by $(\mathbf{n}, \mathbf{b}, \mathbf{b}^\perp)$, where $\mathbf{n}$ is the normal to the medial manifold , $\mathbf{b}$ is the direction in the tangent plane of the fastest narrowing of the implied boundary sections,
4. $\theta \in [0, \frac{\pi}{2}]$ the object angle determining the angulation of the implied sections of boundary relative to $\mathbf{b}$.

The two opposing boundary points implied by the medial atom are given by $\mathbf{y} = \mathbf{x} + \mathbf{p}$ and $\mathbf{y} = \mathbf{x} + \mathbf{s}$. The vectors $\mathbf{p}$ and $\mathbf{s}$ are given by

$$\mathbf{p} = rR_{(\mathbf{b},\mathbf{n})}(\theta)\mathbf{b} \quad , \quad \mathbf{s} = rR_{(\mathbf{b},\mathbf{n})}(-\theta)\mathbf{b} \ ,$$

where $R_{(\mathbf{b},\mathbf{n})}(\theta)$ is a rotation by $\theta$ in the $(\mathbf{b}, \mathbf{n})$ plane.

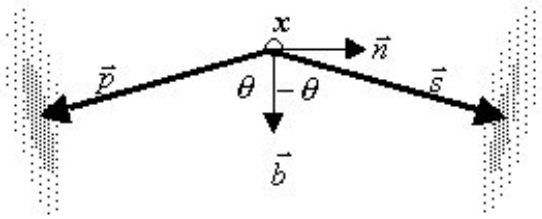


**Fig. 1.** A medial atom defined by the 4-tuple $\{\mathbf{x}, r, \mathbf{F}, \theta\}$ with involutes $\mathbf{P}$ and $\mathbf{S}$ perpendicular to the implied surface.

For stability at the ends in image matching, medial atoms on the boundary of the medial manifold also include an extra parameter $\eta$ that captures the elongation of the edge away from a spherical end cap. The end section of the medially implied boundary is as a parametric curve form one involute to the

other passing through the point $\mathbf{x} + \eta r\mathbf{b}$ and orthogonal to $\mathbf{b}$. The curve $\mathbf{c}(t)$ parametrized by $t \in [-1, 1]$ is defined by

$$\mathbf{c}(t) = \mathbf{x} + r\eta(t)R_{(\mathbf{b},\mathbf{n})}((1-t)\theta))\mathbf{p} \ , \text{ where}$$
$$\eta(t) = (\cos(t\pi) + 1)\frac{(\eta - 1)}{2} + 1 \ ,$$

with $\theta$ being the object angle.

In the above representation $\mathbf{x}$ gives the central location of the solid section of figure that is being represented by the atom $\mathbf{m}$. The scalar $r$ gives the local scale and size of the solid section of figure that is being represented by the atom. The object angle $\theta$ and the direction $\mathbf{b}$ also define the gradient of the scalar field $r$ via

$$\nabla r = -\mathbf{b}\cos\theta \ .$$

The scalar field $r$ also provides a local ruler for the precise statistical analysis of the object.

There are three basic types of medially defined figural segments with corresponding medial manifolds $\mathcal{M}$ of dimension $0, 1, 2$ respectively. Figural segments with two dimensional medial manifolds represent slab-like segments, tube-like segments, where the medial manifold is an one dimensional space curve, and spherical segments, where the medial manifold consists of a single point. Shown in Fig. 2 are examples of slab like and tubular figures. In this paper we will focus on slab-like segments having 2-dimensional medial manifolds discretized into a net of medial atoms. For easy of implementation we have been using a quadrilateral mesh of discretized medial atoms $\mathbf{m}^k_{i,j} \in \mathcal{M}$ , $(i, j) \in [1, N] \times [1, M]$ for approximating the continuous medial manifold at particular scale $k$ with tolerance and the level of discretization inversely proportional to scale with the final scale having tolerance on the order of the resolution of the imaging modality. We define a medial scale space by a sequence of successive refinement of medial nets defined via offsets from a spline interpolation of medial atoms from the scale above.

### 2.2 Spline Interpolation of medial atoms.

Given a quadrilateral mesh of medial atoms $\mathbf{m}_{i,j}, (i, j) \in [1, \cdots, N] \times [1, \cdots, M]$ we define a continuous medial surface via a Bézier interpolation of the discretely sampled medial atoms. The medial position $\mathbf{x}(u, v), u \in [i, i + 1], v \in [j, j + 1]$ is defined via a bicubic polynomial interpolation of the form

$$\mathbf{x}(u, v) = \sum_{m,n=0}^{3} \mathbf{d}_{m,n} u^m v^n$$

with $\mathbf{d}_{m,n}$ are chosen to satisfy the known normal/tangency and continuity conditions at the sample points $\mathbf{x}_{i,j}$.

Given the interpolation of the medial positions the radius function $r(u, v)$ is also interpolated as a bicubic scalar field on the above interpolated medial

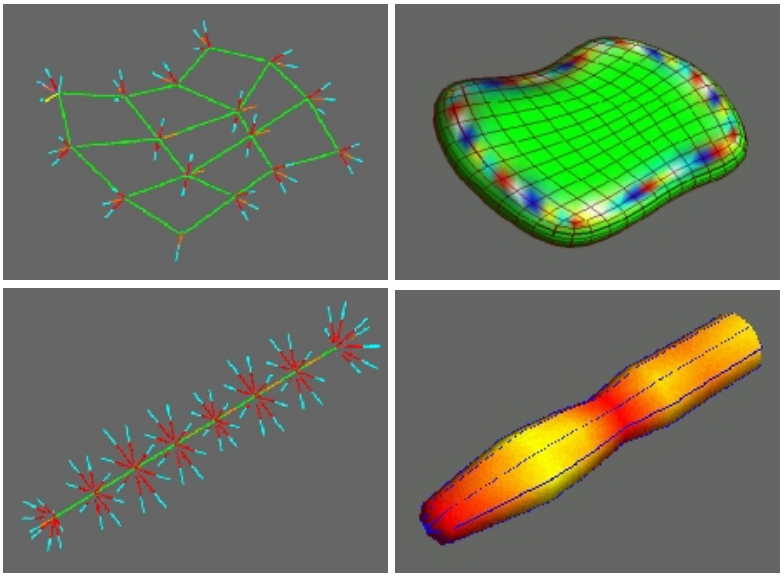

**Fig. 2.** Top rows shows an example of a slab like figure with 2 dimensional medial manifold. Shown in the bottom row is tubular figure with 1 dimensional medial manifold.

manifold given $r$ and $\nabla r$ at the mesh points points $\mathbf{x}_{i,j}$. Having interpolated $r$ and its gradient, the frame $\mathbf{F}$ and the object angle $\theta$ are defined via the relation ship $\nabla r = -\mathbf{b}\cos\theta$

### 2.3 Figural coordinate system.

The prior (geometric typicality) measure requires geometrically consistent correspondence between boundary points in the model and those in a deformed model. The likelihood (deformed model to target image match) measure requires correspondence between template intensities at positions in 3-space relative to the model and target image intensities at positions in 3-space relative to the deformed model. Both of these correspondences are made via the medial geometry, in a way detailed in [6] but summarized here.

The continuous medial manifold of a figure, defined via the spline interpolation describe above, is parameterized by $(u, v)$, with $u$ and $v$ taking the atom index numbers at the discreet mesh positions. A parameter $t \in \{-1, 1\}$ designates the side of the medial manifold on which an implied boundary point lies. As described in section 2.1, $t$ varies continually between $-1$ and 1 as the implied boundary point moves around the crest of the object from one side of the medial axis to another. For single figures boundary correspondences are defined via the common parameterization $(u, v, t)$.

Positions in the image in the neighborhood of the implied boundary are indexed by $(u, v, t, \hat{d})$, where $(u, v, t)$ is the parameterization of the closest point on

the medially implied boundary and $\hat{d}$ is the signed distance (interior = negative, exterior = positive) from the boundary in multiples of the local radius $r$ of the medial point at $(u, v)$.

### 2.4 Connecting m-reps figures into objects.

As illustrated in Fig. 3, protrusion and indentation figures combine into objects in a hierarchical fashion, with the same Boolean operators of union and difference as with Constructive Solid Geometry models, but here recognizing the tolerance of the figures. A figure may be separated from all other figures, or it may be the parent of one or more attached sub-figures: protrusion and/or indentation. A sub-figure on a slab or tube or sphere may be a slab or tube. The interior of a protrusion sub-figure is combined with the parent by union of their interiors with the modification that the boundaries may smoothly blend. An indentation subfigure subtracts its interior from its parent, in the set theoretic sense, again with smooth blending. As illustrated in Fig. 3, a slab protrusion or indentation on a figure has a segment of its medial meshs end atoms that are at the open end of the figure and on the implied boundary of the parent, where the subfigure attaches to its parent. If the subfigure is a tube, it has a single open-end atom where the tube is attached to its parent, and a closed end atom at the other end. We call these the hinge atoms. The remaining end atoms form the closure of that figure. We intersect the subfigures interpolated medial mesh with the implied boundary of the parent figure. In what is presented herein we will concentrate on single figure objects.

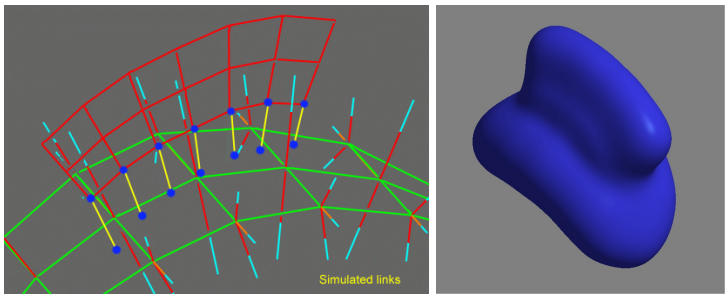


**Fig. 3.** Fig. showing the medial mesh of protrusion sub figure with hinge atoms and the resulting blended implied surface.

### 2.5 Construction of m-rep figures.

Described in Fletcher [6] is an extensive modeling package used for the manual modeling of the anatomical objects using m-reps. Using the visualization and computer aided design techniques developed, we have built numerous models of

anatomical objects. In this paper we focus on the automatic segmentation of the kidney as imaged in CT for radiation treatment for cancer. Shown in Fig. 4 is the template m-rep model of the kidney built from a CT of the abdomen.

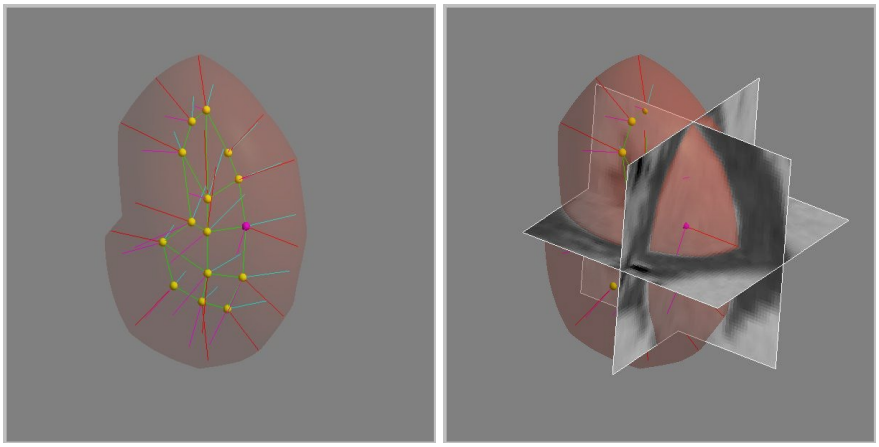

**Fig. 4.** Fig. showing the m-rep model of the template kidney. The left panel shown the medial atoms and the implied surface. The right panel shows the model overlaid on the associated CT imagery.

## 3 Transformation of m-reps figures.

Having defined the construction of typical anatomical objects via m-rep figures, anatomical variability is accommodated by defining a cascade of transformations $S^k, k = 0, \cdots, N$ increasing in dimensionality. These transformations are applied globally to the entire object as well as locally to individual atoms at various scales. Each transformation is applied at its own level of locality to each of the primitives appearing at that level. At each level of locality by the Markov random field framework the primitive is related only to immediately neighboring primitives at that level. Each level's result provides both a initial value and a prior for the primitives at the next smaller scale level. The transformation at the last (smallest) scale level is finally a dense displacement field applied to the boundary of the figure on the scale of the voxel resolution of the imaging modality.

### 3.1 Object-level similarity transformation.

To begin with, a similarity transformation $S^0 = (\alpha, \mathbf{O}, \mathbf{t}) \in [(I\!R^+ \times SO(3)) \ltimes I\!R^3]$ is defined on the scale of the entire object and is applied to the whole medial

manifold $\mathcal{M}$. The similarity transformation $S^0$ scales, translates and rotates equally all the medial atoms of the object, that is

$$\mathbf{m}^1_{i,j} = S^0 \circ \mathbf{m}_{i,j} = \{\alpha \mathbf{O}\mathbf{x}_{i,j} + \mathbf{t}, \alpha r, \mathbf{O} \circ \boldsymbol{F}, \theta\} \ .$$

Notice that the similarity transformation does not affect the object angle. As the medial representation is invariant under the similarity transformation, this is equivalent to applying the similarity transformation $S^0$ to the implied boundary $\mathcal{B}$ of the medial mesh to yield the transformed boundary $\mathcal{B}^1$.

A prior is induced on the above defined transformation based on the displacement of the implied boundary of the objects. Throughout, an independent Gaussian prior on boundary displacement is used with variance proportional to the local radius $r$. For the whole object similarity transformation $S^0$ the log-prior becomes

$$LogP(S^0) = \left[ - \int_{\mathcal{B}} \frac{||\mathbf{y} - S^0 \circ \mathbf{y}||^2}{2(\sigma r(\mathbf{y}))^2} d\mathbf{y} \right] \ .$$

### 3.2 Atom level transformation.

Having accomplished the gross placement of the figure, attention is now focused on the sub-sections of the figure defined by each of the medial atoms. At this stage local similarity transformations as well as rotations of the local angulation, $S^1_{i,j} = (\alpha, \mathbf{O}, \mathbf{t}, \beta)_{i,j} \in [(I\!R^+ \times SO(3)) \ltimes I\!R^3] \times [-\frac{\pi}{2}, \frac{\pi}{2}]$ are applied to the medial atom, that is,

$$\mathbf{m}^2_{i,j} = S^1_{i,j} \circ \mathbf{m}^1_{i,j} = (\alpha_{i,j} \mathbf{O}_{i,j} \mathbf{x}^1_{i,j} + \mathbf{t}_{i,j}, \alpha_{i,j} r^1_{i,j}, \mathbf{O}_{i,j} \circ \mathbf{F}^1_{i,j}, \theta^1_{i,j} + \beta_{i,j}) \ . \quad (1)$$

The resulting implied boundary is defined as $\mathcal{B}^2$. A prior on the local atom transformations $S^1_{i,j}$ is also induced based on the displacement of the implied boundary with an additional Markov random field prior on the translations, guaranteeing the smoothness of the medial manifold. In keeping with the level of locality Let $\mathcal{B}^1_{ij}$ be the portion of the implied boundary affected by the atom $\mathbf{m}^1_{i,j}$. The prior energy on the local transformation $S^1_{i,j}$ of the atom $\mathbf{m}^1_{i,j}$ becomes

$$LogP(S^1) = \left[ - \int_{\mathcal{B}^1_{i,j}} \frac{||\mathbf{y} - \mathbf{y}'||^2}{(\sigma r(\mathbf{y}))^2} d\mathbf{y} - \sum_{i,j} \sum_{n,m=-1}^{n,m=1} \frac{||\mathbf{t}_{i,j} - \mathbf{t}_{i+n,j+m}||^2}{||\mathbf{x}^1_{i,j} - \mathbf{x}^1_{i+n,j+m}||} \right] \ ,$$

where $\mathbf{y}$ is the corresponding position on the figural boundary implied by the transformed atom $\mathbf{m}^2$, and $\mathbf{t}_{i,j}$ is the translation component of the local transformation $S^1_{i,j}$. Good association between points on the boundary $\mathbf{y}$ and the deformed boundary $\mathbf{y}'$ is made using the figural coordinate system describe in section 2.3. The point $\mathbf{y}'$ is the point on the deformed model having the same $(u, v, t)$ coordinates as that of the original point $\mathbf{y}$. The integral in the above prior is implemented as a discrete sum over a set of boundary points by defining a sampling of the $(u, v, t)$ coordinate space and calculating the associated implied boundary before and after an atom deformation.

### 3.3 Dense boundary displacement field transformation.

At the final stage the implied boundary of the figure is displaced in the normal direction using a dense displacement field defined on the implied boundary $\mathcal{B}^2$,

$$\mathbf{y}' \in \mathcal{B}^3 = \mathbf{y} + \mathbf{n}(\mathbf{y})d(\mathbf{y}), \mathbf{y} \in \mathcal{B}^2 \ ,$$

where $\mathbf{n}(\mathbf{y})$ is the normal to the implied boundary at $\mathbf{y} \in \mathcal{B}^2$.

As with the local atom transformations the prior is induced on the dense displacement field using a Markov random field prior derived from energetics associated with thin elastic membranes to guarantee smoothness. The log-prior on the displacement field $d(y)$ becomes

$$LogP(d(x)) = \left[ - \int_{\mathcal{B}^2} \frac{|d(\mathbf{y})|^2}{(\sigma r(\mathbf{y}))^2} - \int_{\mathcal{B}^2} |\nabla d(\mathbf{y})|^2 d\mathbf{y} \right] \tag{2}$$

The above above prior is implements via a discrete implementation as follows. Let $\mathbf{y}_i \in \mathcal{B}_2, i = 1, \cdots, N$ be the set of discrete boundary points on the implied boundary $\mathcal{B}^2$. Let $\mathcal{N}(\mathbf{y}_i)$ be the set of neighbors of the point $\mathbf{y}_i$. The discrete approximation of equation 2 becomes

$$- \sum_{i=1}^{N} \frac{|d(\mathbf{y}_i)|^2}{(\sigma r(\mathbf{y}_i)} - \sum_{i=1}^{N} \sum_{j \in \mathcal{N}(\mathbf{y_i})} \frac{|d(y_j) - d(y_i)|^2}{||y_j - y_i||} \ .$$

## 4 Image Data Log-Likelihood

Having defined the transformation and the associated prior energetics, we now define the data likelihood function needed for defining the posterior. We have been defining the data likelihood functions, using the object centered coordinate system developed in section 2.3, by defining correlation functions between a predefine template image $I_{temp}$ and the data $I_{data}$ in the neighborhood of the boundary of the medially define object $\mathcal{B}$. Leting $\delta$ be the size of the collar around the object, in multiples of $r$ the local radius, the data log likelihood function becomes

$$\int_{-\delta}^{\delta} \int_{\mathcal{B}} I_{temp}(\mathbf{y}, \hat{d}) I_{data}(\mathbf{y}', \hat{d}) d\mathbf{y} d\hat{d} \ , \tag{3}$$

where $(\mathbf{y}, \hat{d}) \in I\!R^3$ is the point in the template image at distance $r\hat{d}$ away from the boundary point $\mathbf{y}$, and $(\mathbf{y}', \hat{d})$ is the point in the data image at distance $r\hat{d}$ away from the boundary point $\mathbf{y}'$ in the transformed object $\mathcal{B}'$. This association between points in the template image and the data image is made using the object coordinate system described in section 2.3. The image positions in the neighborhood of the implied boundary are indexed by $(u, v, t, \hat{d})$, where $(u, v, t)$ is the parameterization in the object centered coordinate system of the closest point on the medially implied boundary $\mathcal{B}$, and $\hat{d}$ is the signed distance (interior = negative, exterior = positive) from the boundary in multiples of the local

radius $r$ of the medial point at $(u, v)$. In implementing the correlation defined in Eqn. 3 care must be taken in implementing the surface integral by a discrete voxel summation. The template image needs to be normalized by the determinant of the Jacobian associated with the implied model surface $\mathcal{B}$. At model building time intensities in the template image $I_{temp}$ are associated with their positions' $(u, v, t, \hat{d})$ values. As the model deforms, a target image position is calculated for each template $(u, v, t, \hat{d})$ value, using the deformed model, and the intensity interpolated at that target image position is associated with the corresponding template intensity.

We have have been using two basic types of templates: an analytical template derived from the derivative of the Gaussian and an imperical template learned from an example image from which the template medial model was built.

Using the data likelihood defined above and the prior defined in previous section, the log posterior is defined as a weighted sum of the two terms with weights chosen by the user. For optimizing the log-posterior with respect to the global object similarity transformation and the local atom-by-atom transformation, we have been using a genetic optimization algorithm. Genetic algorithms have the advantages of not being susceptible to local minimum and not requiring the computation of the derivative of the posterior with respective to the transformation parameters. For optimizing the posterior with respect to the dense displacement field $d(bfy)$ we have been using a simple gradient decent algorithm.

## 5 Results

We have been using the automatic segmentation procedure for extracting the kidney parenchyma-including the renal pelvis-in subjects undergoing radiation treatment for cancer. Results from a series of three data sets are presented. Using a few seconds, the user rigidly place the template model in the subject data set using the the interactive tool described in Fletcher [6]. This initialization stage of the algorithm is followed by the hierarchical automatic segmentation which takes on the order of 5 minutes for convergence depending on the data set. At the first scale level, a object similarity transformation is estimated accommodating gross size and orientation differences between the template model kidney and the subject's kidney.

Fig. 5 compares the results of the similarity transformation to the clinical hand segmentation in the axial, coronal, and sagittal views through the kidney. The yellow contour of the resulting implied boundary is overlaid, for comparison with the clinical hand segmentation shown in red. Note that the clinical hand segmentation did not include the renal pelvis, while our single figure model of the kidney used in this study includes the renal pelvis.

Fig. 5, shows the improvement in the segmentation as a result of the atom deformation process, thus accommodating more local object shape changes.

Fig. 6 shows the improvement due to the final stage of the deformation, as the dense displacement field accommodates the fine featured variation in the shapes of the kidney.

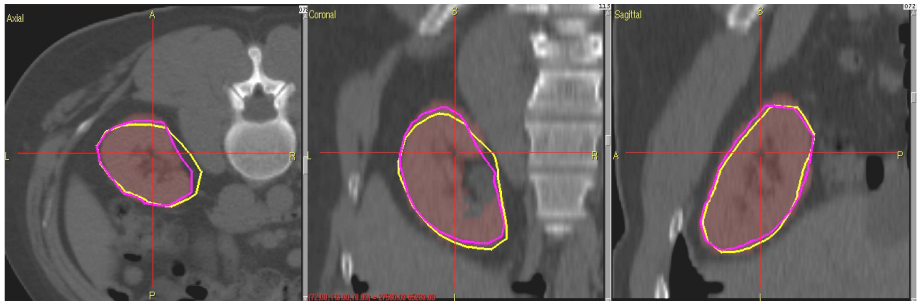


**Fig. 5.** Axial (left), coronal (middle) and sagittal (right) slices through the subject kidney CT data set with the hand segmentation overlaid in red. In yellow are the contours of the result of the object similarity transformation. The result of the atom deformation is shown in purple.


**Fig. 6.** The improvement in the segmentation of the kidney after the dense displacement field deformation. The purple contour shows the result of the atom transformation while the blue contour shows the result of the dense displacement field deformation. The hand segmentation is again overlaid in red.

For quantitative comparisons of the segmentations of the method with manual segmentations, we have used two metrics from a geometric scoring package developed by Guido Gerig and Matthieu Jomier called VALMET : relative overlap and mean surface distance. The relative overlap measure is defined as the ratio of the intersection of the two segmentations divided by the union. Although the relative overlap is commonly used in the literature for scoring segmentations it is sensitive to the size of the object and not very effective in characterizing shape differences between two segmentations. The symmetric, mean surface distance $D_s$ between the boundary of the two segmenations using Euclidean distance transforms of the segmentations is defined as follows. Let $y_i^1, i = 1, \cdots, N \in \mathcal{B}^1$ and $y_j^2, j = 1, \cdots, M \in \mathcal{B}^2$ be the boundary points of two segmenattions $\mathcal{B}^1, \mathcal{B}^2$; the mean surface distance then is

$$D_s(\mathcal{B}^1, \mathcal{B}^2) = \frac{1}{2}\left[\frac{1}{N}\sum_{i=1}^{N} \min_{j=1\cdots M} ||y_i^1 - y_j^2|| + \frac{1}{M}\sum_{j=1}^{M} \min_{i=1\cdots N} ||y_i^1 - y_j^2||\right] .$$

Shown in table 1 is the summary of the results from the study for the three data sets. The results shown above are typical of the three data sets and are form Data set 613. The segmentation improves at each stage of the algorithm for all three data sets. The accuracy of the segmentation as measured via the mean surface distance is on the order of the resolution of the data set and on average within one pixel of the hand segmentation.

| Data Set (cm) | Scale Level | Relative Overlap | Surface Distance (cm) |
|---|---|---|---|
| 613 $0.15 \times 0.15 \times 0.5$ | Similarity Transformation | 0.85 | 0.26 |
| | Atom deformation | 0.86 | 0.23 |
| | Field deformation | 0.90 | 0.16 |
| 608 $0.2 \times 0.2 \times 0.4$ | Similarity Transformation | 0.88 | 0.22 |
| | Atom deformation | 0.89 | 0.19 |
| | Field deformation | 0.93 | 0.14 |
| 1402 $0.15 \times 0.15 \times 0.3$ | Similarity Transformation | 0.77 | 0.65 |
| | Atom deformation | 0.86 | 0.38 |
| | Field deformation | 0.90 | 0.38 |

**Table 1.** Table showing the relative overlaps and the mean surface distance between the manual segmentations and the automatic segmentations at the different stages of the hierarchical procedure for the three data sets processed.

## 6 Discussion and Conclusion

It can be seen from the quantitative analysis of the segmentations that the accuracy of the automatic segmentation as measured via the average surface distance

is on the order of the resolution of the imaging modality. Although these results show that our current methodology can segment structures in the abdomen such as the kidney with high level of accuracy, improvement can be expected from the change in the image template used in the data likelihood. All the results shown in this paper were generated using a Gaussian derivative template for the data-likelihood. In the CT image the kidney sits in an environment where part of its boundary is distinctly darker, but part of its boundary region is formed by the liver. Not too far away are bony structures such as the ribs and vertebrae, appearing with very high contrast. When using a Gaussian template that is designed to give increased response at boundaries, the following behaviors are noted. The kidney m-rep can move inside a sequence of vertebral bodies because of the high contrast. Also 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. To avoid such behaviour careful initialization was required with very high prior weight. The boundary at the liver, appearing with at most texture contrast, does not strongly attract the implied m-rep boundary. With a high weight associated with the prior the automatic segmentation does not quite follow the kidney/liver boundary

We expect these problems to be overcome by the use of our already implemented but not yet tested training image template in place of the Gaussian derivative template that would allow a spatially varying template capturing the different gray scale characteristics of the kidney boundaries. This model to image match would be further improved a statistical model reflecting image intensity variations across a population of subjects.

We have also been working on extending this frame work to the deformation of objects with multiple attached sub-figures and multiple objects with priors induced on the transformations that reflect the knowledge of the associated relative typical geometry.

## 7 Acknowledgement

We thank Prof. Gerig and Matthieu Jomier for the use of their scoring tool for the comparison of segmentation as well as for the many insightful discussions and comments. We would like to also thank Dr. Zhi Chen for the generating the table comparing the segmentations. We also thank Prof. Ed. Chaney for providing us the data sets and invaluable insights. This work was supported by NIH Grants P01 CA47982 R01 CA67183 This research was carried out on computers donated by Intel.

## References

1. Irving Biederman. Recognition-by-Components: A Theory of Human Image Understanding. *Psychological Review*, 94(2):115–147, 1987.
2. Jules Bloomenthal and Chek Lim. Skeletal methods of shape manipulation. In *Proc. Shape Modeling and Applications*, pages 44–47. IEEE, 1999.

3. H. Blum. A transformation for extracting new descriptors of shape. In *Models for the Perception of Speech and Visual Form*. MIT Press, 1967.
4. A. C. Burbeck, S M Pizer, B. S. Morse, D. Ariely, G. Zauberman, and J. Rolland. Linking object boundaries at scale: a common mechanism for size and shape judgments. In *Computer Science Department technical report TR94-041*, page 361:372, Chapel Hill, 1996. University of North Carolina.
5. T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active shape models - their training and application. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 1(61):38–59, 1994.
6. P. T. Fletcher, S. M. Pizer, A. Thall, and A. G. Gash. Shape modeling and visualization in 3d with m-rep object models. In *Submited to Information Processing in Medical Imaging*, 2000.
7. D. Fritsch, S. Pizer, L. Yu, V. Johnson, and E. Chaney. Segmentation of Medical Image Objects using Deformable Shape Loci. In *International Conference on Information Processing in Medical Imaging*, pages 127–140, Berlin, Germany, 1997. Springer-Verlag.
8. U. Grenander. *General Pattern Theory*. Oxford Univ. Press, 1994.
9. Takeo Igarashi, Satoshi Matsuoka, and Hidehiko Tanaka. Teddy: A sketching interface for 3d freeform design. *Proceedings of SIGGRAPH 99*, pages 409–416, August 1999.
10. M. Leyton. *Symmetry, Causality, Mind*. MIT Press, Boston, 1992. 620 pages.
11. Lee Markosian, Jonathan M. Cohen, Thomas Crulli, and John F. Hughes. Skin: A constructive approach to modeling free-form shapes. *Proceedings of SIGGRAPH 99*, pages 393–400, August 1999.
12. David Marr and H. K. Nishihara. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. Roy. Soc. London Ser. B*, 200:269–294, 1978.
13. Lee R. Nackman. *Three-Dimensional Shape Description Using the Symmetric Axis Transform*. PhD thesis, UNC Chapel Hill, 1982. under the direction of Stephen M. Pizer.
14. S. Pizer, D. Fritsch, P. Yushkevich, V. Johnson, and E. Chaney. Segmentation, registration, and measurement of shape variation via image object shape. *IEEE Transactions on Medical Imaging*, 18:851–865, October 1999.
15. S.M. Pizer, T. Fletcher, Y. Fridman, D.S. Fritsch, A.G. Gash, J.M. Glotzer, S. Joshi, A. Thall, G Tracton, P. Yushkevich, and E.L. Chaney. *Deformable M-Reps for 3D Medical Image Segmentation*. In Review, ftp://ftp.cs.unc.edu/pub/users/nicole/defmrep3d.final.pdf, 2000.
16. Kaleem Siddiqi, Sylvain Bouix, Allen Tannenbaum, and Steven W. Zucker. The hamilton-jacobi skeleton. In *Proc. Computer Vision*, volume 2, pages 828–834. IEEE, 1999.
17. A. Thall, S. M. Pizer, and P. M. Fletcher. Deformable solid modeling using sampled medial surfaces: a multiscale approach. In *Computer Science Department technical report TR00-005*, Chapel Hill, 2000. University of North Carolina.
18. G. Wyvill, C. McPheeters, and B. Wyvill. Data structure for soft objects. *Visual Computer*, 2:227–234, 1986.
19. Alan Yuille and Peter Hallinan. *Active Vision*, chapter Deformable Templates. MIT Press, Cambridge, MA, 1992.