

# Autosegmentation of Images in Radiation Oncology

Edward L. Chaney, PhD, Stephen M. Pizer, PhD

It is likely that segmentation for medical purposes is performed more often for radiation therapy than for all other medical applications combined. *Segmentation* here refers to creating 3-D geometric models of anatomic objects from volume images, primarily computed tomographic and magnetic resonance images, that are needed to guide critical treatment planning and delivery decisions in radiation therapy. Manual slice-by-slice contouring methods currently used in clinical practice are extremely time consuming and require highly skilled personnel. Also, contoured structures demonstrate significant interuser and intrauser variability that can adversely affect clinical decisions and contribute to nonuniform practice. Goals of automatic segmentation include the reduction of human effort, expertlike performance, more uniform clinical practice through the reduction of interuser and intrauser variability, and ultimately, real-time computation to support online applications such as beam reshaping and replanning during treatment delivery.

Two general approaches show strong promise for radiation therapy applications: the mapping of voxels and deformable shape models (DSMs). Voxel mapping assumes that the spatial distribution of voxels in a target image is a small-scale rearrangement of the voxels in an atlas image that has been presegmented by an expert [1]. So-called diffeomorphic methods construct a smooth and invertible high-dimensional registration transformation at the spatial scale of a voxel by solv-

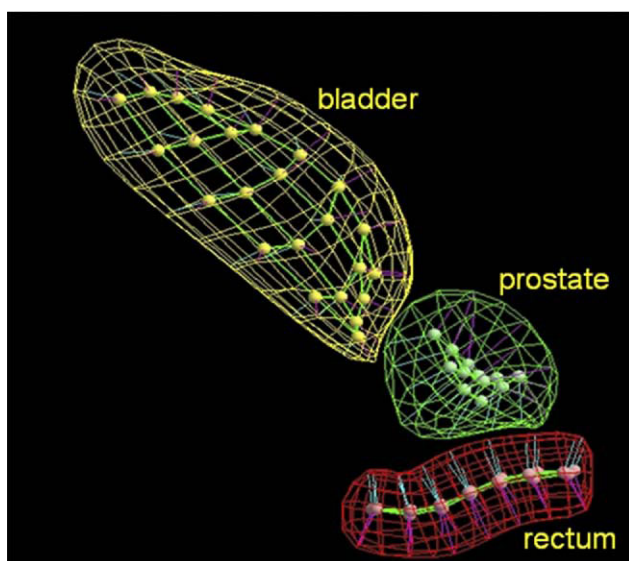
ing differential equations describing the “flow” of intensity values in the atlas image onto corresponding intensity values in the target image [2]. The transformation maps the segmentation labels to automatically segment the target image. Voxel-mapping methods have not been shown to work robustly for anatomic sites other than the brain, mainly because target images are large-scale rearrangements of the atlas images. For some sites, such as the head and neck, multiple atlases are created, and the user selects the atlas that most closely matches the target image (IKOE, Houston, Texas [<http://www.ikoemed.com>]). For other sites, such as the male pelvis, challenges posed by large shape and intensity differences between target and atlas images have not yet been overcome. Because the shape of an anatomic object is implied by its segmentation in the atlas, voxel-mapping methods do not take into account statistical knowledge about shape and shape variability of anatomic objects.

Deformable shape models involve creating an explicit mathematical representation (the model) of a target object. For segmentation, the “starting” DSM (eg, a mean model) is initialized in the target image and deformed by optimizing an objective function to closely match the shape of the target object. Nontrainable DSMs do not use prior knowledge about the shape and shape variability of a target object or knowledge about image intensity patterns (the “appearance”) in and around an object. Statistically trainable DSMs attempt to overcome this shortcoming by us-

ing prior shape and appearance information. The geometric prior is the mean shape of an object together with the principal modes of shape variation [3]. The appearance prior captures information about intensity patterns and their variations relative to the object of interest [4,5].

Boundary representations (b-reps) are statistically trainable DSMs composed of directly recorded boundary locations [6,7]. Philips Medical Systems (Andover, Massachusetts) [8] offers b-rep organ models for the Pinnacle RTP system based on the work of Pekar et al [9]. Varian Medical Systems (Palo Alto, California) offers a hybrid model approach for the Eclipse RTP system that includes b-reps [10]. Boundary representations have shortcomings that derive mainly from their relatively simple form. For example, b-reps are shells instead of solids and lack a “natural” coordinate system to facilitate representation of the appearance prior and interobject relationships and the establishment of tissue-voxel correspondence needed for dose accumulation across multiple treatment images of the same patient.

Medial representations (m-reps), proprietary IP of Morphormics, Inc ([www.morphormics.com](http://www.morphormics.com)) represent objects in terms of their centers and widths [11]. Medial representations are solid models of both interior and surface geometries. Continuous m-reps are a class of so-called skeletal structures that have been extensively studied and reported in the literature [11]. This strong mathematical foundation provides a basis for developing a



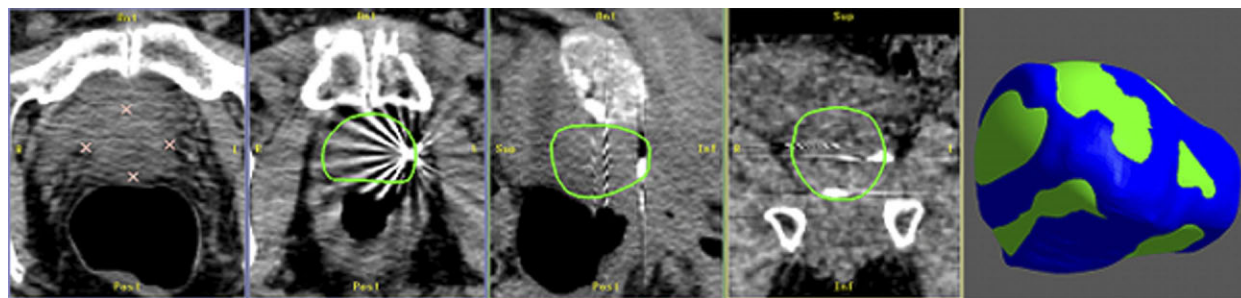
**Fig 1.** Medial representation models for the prostate, bladder, and rectum.

principled approach for shape representation and shape analysis via m-reps. In medial terms, a simple slablike object such as a bladder can be represented as a skeletal framework comprising a 2-D curved sheet, called the medial sheet, lying midway between opposing surfaces of the object. At all points on both sides of the sheet, equal-length spokes extend to opposite sides to the object boundary. A set of spokes together with their common origin, called a hub, is called a medial

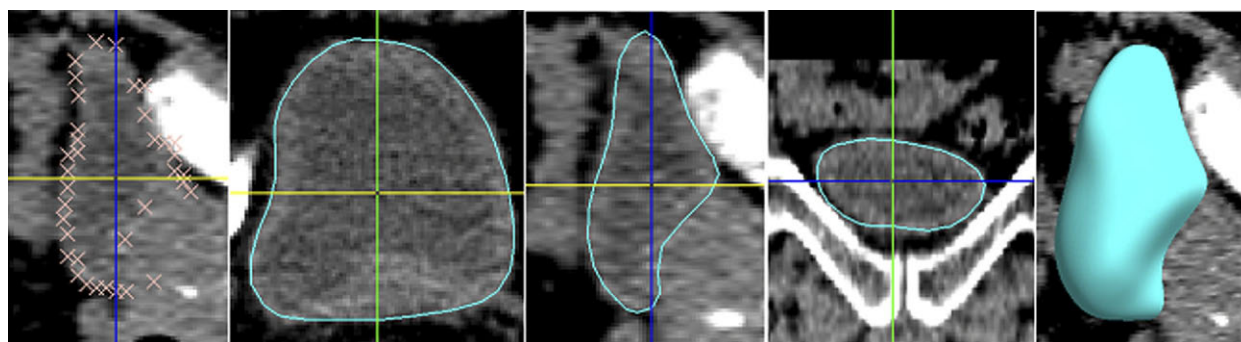
atom. In practice, a sampled grid with up to a few tens of atoms, each with 2 or 3 spokes, is used to represent the medial sheet (Figure 1). Atoms at intermediate points on the sheet can be interpolated. Tubelike objects, such as the rectum, are represented by a quasi-tube m-rep comprising a single medial chain of atoms, in which each atom can have several spokes of unequal lengths to accurately represent noncircular cross-sections. The full surface for a slablike or tubelike object can be

recreated from the spoke ends. Every medial atom can be understood as a simple transformation of any other. The m-rep architecture and the tightly coupled interatom relationships allow m-reps to richly capture key geometric properties of shape and shape variability.

Medial representations used for segmentation have both geometric and appearance priors. The geometric prior is trained by fitting m-rep models to truth contours and analyzing the resulting collection of m-reps for each organ using principal geodesic analysis [3]. Principal geodesic analysis is similar to principal component analysis but is better suited for the high-dimensional nonlinear shape space, called symmetric space, in which m-reps “live.” Like principal component analysis, principal geodesic analysis yields a mean m-rep, basic modes of variation, and variances associated with each mode. An important difference compared with principal component analysis, however, is that principal geodesic analysis on m-reps results in principal modes that are better aligned with natural shape changes, such as bending, twisting, elongation, widening, and deformation caused by external or internal pressure (as for the bladder



**Fig 2.** (Left panel) User-selected initialization points on an axial slice near the prostate base. Points also are selected on slices near the center and apex. (Center panels) Final segmentation results on midaxial, sagittal, and coronal slices. Note that the medial representation (m-rep) appearance prior is robust against image artifacts. (Right panel) Example comparison between an m-rep segmentation (green) of the prostate and the human expert (blue) who generated the initialization points. Interweaving of the surfaces demonstrates excellent agreement.



**Fig 3.** (Left panel) Automatically generated initialization points for the bladder in a mid-sagittal slice. Points also are automatically generated in midaxial and coronal slices. (Center panels) Final segmentation results in midaxial, sagittal, and coronal slices. (Right panel) Surface rendering of segmented bladder.

and rectum). The appearance component is trained by statistically analyzing regional and local image intensity patterns in and around fitted m-rep models. One strength of m-reps is that the appearance component is expressed in m-rep-relative coordinates, tightening statistics, making appearance more informative, and allowing one object to “know” about the appearance of neighbor objects. One form of the appearance prior, called a regional intensity quantile function, is a probability density function derived from exterior and interior regional intensity histograms [5].

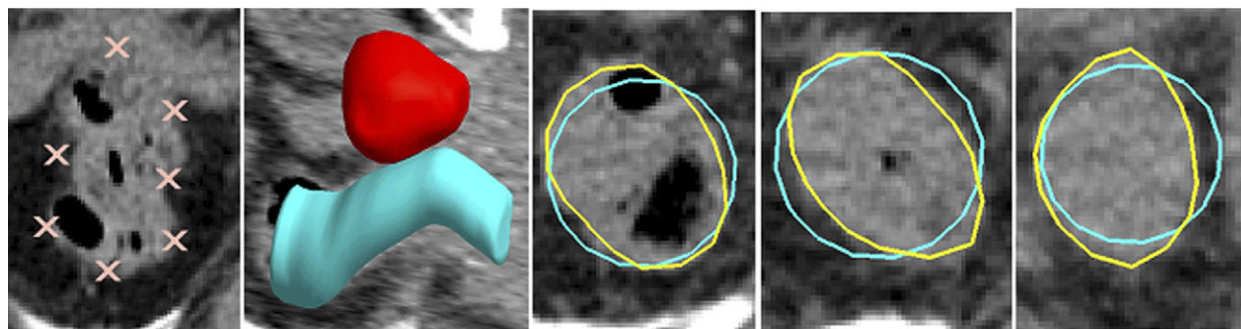
The segmentation of an organ involves initialization of the mean m-rep in the target image, followed

by deformation into the grayscale image data. Computationally, both steps are based on optimizing an objective function in a Bayesian-like framework summarized by

$$\begin{aligned} & \log(\text{probability of model}) \\ & \propto \log(\text{probability of model's} \\ & \text{geometry}) \\ & + \log(\text{probability of model} \\ & \text{given the image data}). \end{aligned}$$

For initialization, the starting m-rep is deformed into a few initialization points on the surface of the object that are selected by the user or automatically generated. Optimization of the objective

function yields the most probable m-rep geometry that is consistent with the initialization points and leads to a segmentation result that is less likely to require editing, because it is biased toward the user's initialization points. Example results for the prostate, bladder, and rectum are shown in Figures 2 to 4. In most cases, m-reps at their current state of development perform like an experienced human for these 3 organs. Models for the seminal vesicles and femoral heads will be available soon to round out the male pelvic structures. Other sites being investigated include subcortical brain structures, head and neck, abdomen, and 4-D lung.



**Fig 4.** (Left) Initialization points placed by point-and-click interaction on an axial slice through the rectum. Points are placed on 3 slices. (Second from left) Surface rendering of rectum medial representation (blue) deformed into initialization points. The autosegmented prostate is shown in red. (Right 3 panels) Comparison of final segmentation results (yellow) with initialization results on 3 axial slices.

Edward L. Chaney, PhD, and Stephen M. Pizer, PhD, are from The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; and Morphormics, Inc, Chapel Hill, North Carolina.

**Edward L. Chaney, PhD**, Morphormics, Inc, 6320 Quadrangle Dr, Suite 380, Chapel Hill, NC 27517; e-mail: [ed.chaney@morphormics.com](mailto:ed.chaney@morphormics.com).