

Medical Image Segmentation Using Analysis of Isolable-Contour Maps

Smadar Shiffman*, Geoffrey D. Rubin, and Sandy Napel, *Member, IEEE*

Abstract—A common challenge for automated segmentation techniques is differentiation between images of close objects that have similar intensities, whose boundaries are often blurred due to partial-volume effects. We propose a novel approach to segmentation of two-dimensional images, which addresses this challenge. Our method, which we call intrinsic shape for segmentation (ISeg), analyzes isolabel-contour maps to identify coherent regions that correspond to major objects. ISeg generates an isolabel-contour map for an image by multilevel thresholding with a fine partition of the intensity range. ISeg detects object boundaries by comparing the shape of neighboring isolabel contours from the map. ISeg requires only little effort from users; it does not require construction of shape models of target objects. In a formal validation with computed-tomography angiography data, we showed that ISeg was more robust than conventional thresholding, and that ISeg's results were comparable to results of manual tracing.

Index Terms—Computed tomography angiography, segmentation, contour matching, shape analysis.

I. INTRODUCTION

SEGMENTATION is an important step in most applications that use medical image data. For example, segmentation is a prerequisite for quantification of morphological disease manifestations and for radiation treatment planning [1], [2], for construction of anatomical models [3], for definitions of flight paths in virtual endoscopy [4], for content-based retrieval by structure [5], and for volume visualization of individual objects [2]. The application that provided the incentive for our work was visualization of vasculature from computed-tomography angiography (CTA) data.

Our goal was to provide a segmentation tool that would allow experts to edit CTA volumes with little effort and good reproducibility. We observed that two characteristics of CTA data were likely to challenge existing segmentation methods. The first characteristic is the proximity of vasculature and bone structures—both of which appear as bright regions in CTA data—in several areas, for example, in the abdomen, where the aorta may contact lumbar vertebral osteophytes.

Manuscript received October 12, 1999; revised August 16, 2000. This work was supported by The National Library of Medicine under Grants LM-07033 and LM05305, by the National Institutes of Health (NIH) under Grants R01HL50305 and P41-RR09784, and by the Siemens Medical Systems, and the Lucas Foundation. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was C. Roux. *Asterisk indicates corresponding author.*

*S. Shiffman is with the departments of Psychiatry and Radiology, Stanford University, Stanford, CA 94305 USA.

G. D. Rubin and S. Napel are with the Department of Radiology, Stanford University, Stanford, CA 94305 USA.

Publisher Item Identifier S 0278-0062(00)09739-1.

Vessels and bones that touch each other are difficult to resolve by observation of intensity discontinuity alone because partial-volume effects (PVE) cause the discontinuity to weaken where the structures touch. The second characteristic is the profound variation in both normal and pathological structures across patients. This variation makes it difficult to use *a priori* shape information drawn from training data. Based on these characteristics we determined that a segmentation procedure for our application required the ability to resolve objects with similar intensities that are close, or that touch, throughout the volume, without requiring construction of global shape models of the target objects.

This paper describes our work toward a new segmentation method that allows experts to edit two-dimensional image data with ease and with good reproducibility, and that addresses the challenges associated with segmentation of CTA data. The organization of the paper follows. Section II outlines existing segmentation methods that we considered. Section III presents the theory that underlies our method, and explains how we applied this theory to perform segmentation. Section IV describes experiments that we conducted to validate our method. Section V presents a discussion of our work, and Section VI draws conclusions.

II. BACKGROUND

Several visualization and image-processing systems include interactive tools to expedite manual editing—for example, tools that elicit manual traces on a single image and then propagate the trace throughout a slab, or a sequence of images [1], [2], [6], [7]. Although these tools facilitate segmentation, they still require considerable user effort when they are used to edit target structures that are surrounded by structures with a similar intensity range. Many algorithms for segmentation of medical images use *a priori* knowledge in the form of probabilistic or geometric models to guide the segmentation process [3], [8]–[12]. Typically, these models are derived through an automated supervised-learning process [13] that produces a general model from training samples that an expert defines. In the past decade, researchers introduced a variety of deformable models [14]–[17], which minimize energy-functions to find the boundaries of objects. Several groups demonstrated that deformable models produced good segmentation results even when the initial conditions were data driven [9], [18], [19]. However, data-driven initial conditions could depict touching structures that have similar intensities as a single structure from the onset of the energy-minimization process, and thus might lead to an incorrect segmentation.

There are several semiautomated data-driven segmentation methods that are based on spatial connectivity of voxels [7], [20], [21]. The use of intensity thresholds to guide connectivity for segmentation is appealing because intensity thresholds are easy to implement. However, because objects that appear in medical images often have overlapping intensity distributions, thresholding is likely to result in misclassification of voxels. When the prior distributions for tissue types are known, we can find the thresholds that minimize the probability of voxel misclassification [22]. When elaborate prior distributions that account for imaging artifacts are not known, users must select thresholds through trial and error to produce results that portray the structures of interest adequately. Typically, users guess an initial threshold based on the expected intensity of the object of interest, and then they modify the threshold as they observe the results, until they find a satisfactory threshold. Users distinguish between good and inadequate thresholds by examining the extent to which the shape of the region of interest reflects the users' knowledge of anatomy and by the extent to which the region is separated from its surroundings.

We believed that the existing methods would not be adequate for segmentation of structures from CTA data because we expected that the methods would require considerable user input. In Section III, we present an alternative method that exploits shape information that is implicit within isointensity contours in medical images. Our method mimics the way users refine a depiction of an object's boundary by adjusting an intensity threshold and noting abrupt deviations from the shape of the boundary.

III. METHODS

Our method is based on an automated analysis of the shape of contours that result from applying a fine partition to the intensity range of raw images. The following two sections discuss imaging principles that constitute the foundations of our segmentation method, and the algorithms that we incorporated into our method.

A. Foundations

In this section, we explain how thresholding the intensity range of an image with a fine partition results in an isolabel-contour map. Then, we explore characteristics of objects' appearance in isolabel-contour maps, and we suggest that we can compare the shape of successive isolabel contours to detect approximate object boundaries.

Multilevel thresholding partitions the intensity range of an image into bins, and assigns a distinct label to voxels associated with each intensity bin. The resulting image includes sets of connected voxels, such that the voxels in each set have a uniform label, and each voxel belongs to exactly one set. The voxels that lie on the boundary of such a set form an **isolabel contour**. In formal notation, the partition of the intensity range I results in a set of ordered distinct intensity subranges $s_j = [i_j, k_j)$, where $\cup s_j = I$. We assign numeric labels l_j to voxels whose intensity is in the range s_j , where l_j reflects the relative intensity of s_j .

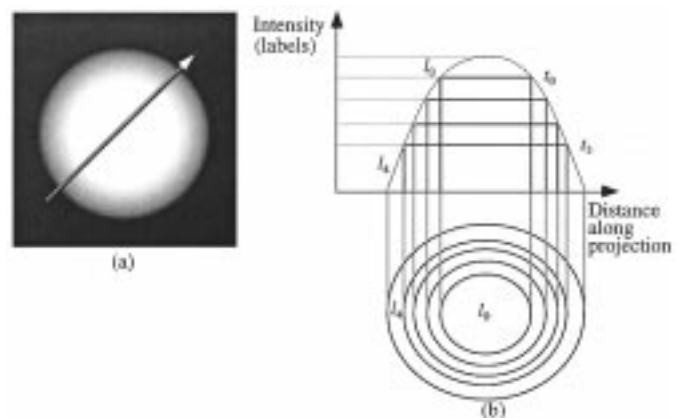


Fig. 1. A label pattern formed by thresholding of the intensity range with a fine partition. (a) A synthetic image of a circular object with a gradual intensity fall across the boundary and a line across the image. (b) The intensity profile of voxels along the line in part a, and the projections of voxels whose intensities are greater than or equal to the threshold intensities. Here, the l_i denote labels (l_0 corresponds to the highest range of intensities) and the t_i denote thresholds. Note the pattern of circular regions with decreasing labels.

We employ multilevel thresholding to generate pseudoconcentric contours. Because occurrences of PVE yield a gradual intensity fall across the boundaries of objects, multilevel thresholding with an intensity partition that is sufficiently fine results in labeled images whose isolabel contours form conspicuous patterns (Fig. 1). Because isolabel-contour patterns resemble isoelevation contours on topographical maps we refer to the labeled images as *isolabel-contour maps*. If we observe an area within an isolabel-contour map that extends from an object's center toward its boundary, we see a distinct pattern. Where the intensity gradient is monotonic in the raw image, the pattern of labels in the isolabel-contour map is monotonic as well. We observe dense contour patterns in areas of abrupt intensity gradients, and widespread contour patterns in areas of gradual intensity gradients. Fig. 2 shows examples of isolabel-contour maps from CTA images.

The shape of an isolabel contour that encloses a set of voxels depends on the shape of the object of which the voxels are an image, on noise, on intensity variations within the image, and on manifestations of partial volume. We observed that, in general, isolabel contours that enclose voxels from only a single object have shapes that are similar to the shape of the object boundary, and thus, that these contours have similar shapes; similarly, isolabel contours that enclose voxels from different objects often have different shapes. Isolabel contours that enclose voxels from the same object often have similar orientation, but their size and position (center of mass) may vary, depending on the rate and uniformity of the intensity fall across the boundary of the object. Partial-volume effects may result in saddle-like patterns for tortuous structures whose plane of curvature is perpendicular to the image plane or for vessels that branch in the through-plane direction. Fig. 2 shows examples of saddle-like patterns.

We can describe the entire pattern of contours in an isolabel-contour map as a relationship of enclosure. This description allows us to manipulate easily voxels that related contours enclose. For any two isolabel contours in an

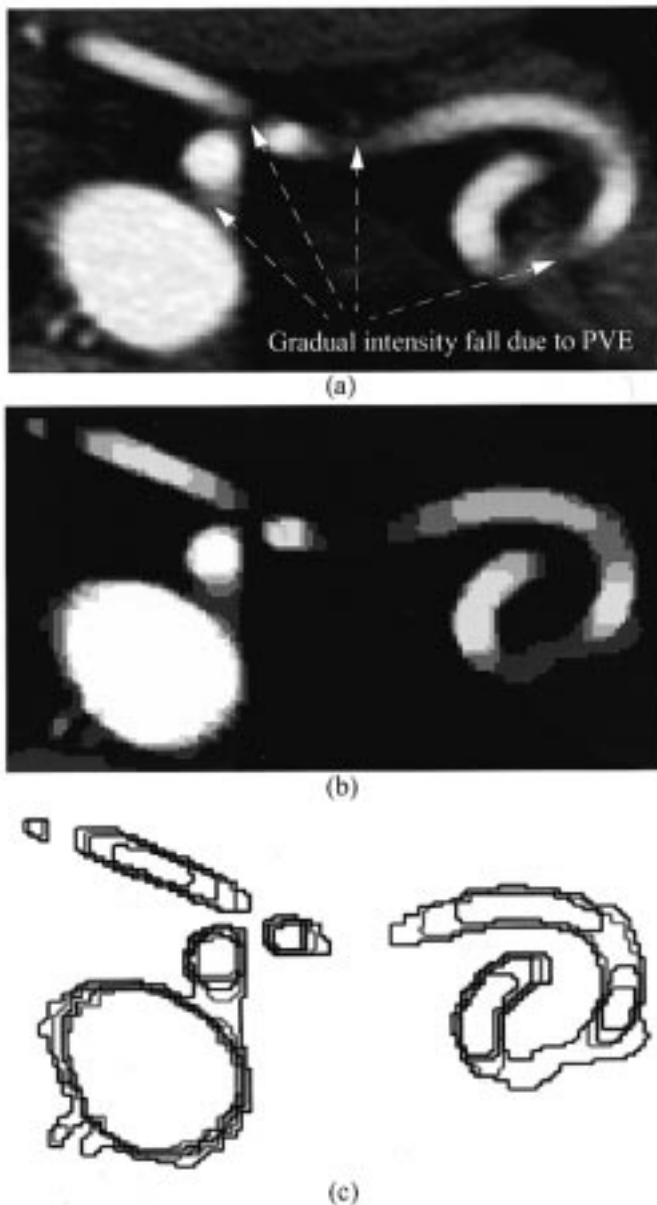


Fig. 2. A CTA section and its isolabel-contour map. (a) The raw section. Note the gradual intensity fall due to extensive PVE where the vessel is oriented at a small angle with respect to the image plane. (b) The isolabel-contour map. The map resulted from a seven-threshold partition of the intensity range. Note the saddle-like patterns of isolabel contours for the branching and tortuous vessel. (c) Contour patterns for the major structures in (a). For clarity, the figure shows only the contours that correspond to the six highest thresholds, and displays the contours in a variety of grayscale values. Note that the contour patterns are dense along sharp intensity gradients and are more widespread where intensity falls gradually.

isolabel-contour map, c_i, c_j , with sets of connected voxels that they encompass r_i, r_j , the following statement holds:

$$\langle r_i \cap r_j = \emptyset \rangle \vee \langle r_i \cap r_j = r_i \rangle \vee \langle r_i \cap r_j = r_j \rangle.$$

The statement is true because its negation implies that there are voxels with more than one label assignment, and that implication is inconsistent with generation of the isolabel-contour map via multilevel thresholding. We define an **enclosure** relation R_e

over the image. If r_{i+1} is the set of voxels encompassed by contour c_{i+1} , and r_i is the set of voxels encompassed by contour c_i , then $(c_{i+1}, c_i) \in R_e$, or c_{i+1} encloses c_i , if and only if $r_{i+1} \supseteq r_i$. The enclosure relation allows us to produce hierarchical descriptions of contours and of the regions that they circumscribe. In this descriptions, inner-regions, or regions that do not circumscribe any other regions, are depicted as leaves of a tree structure. Fig. 3 illustrates a labeled image and a corresponding region-based tree structure.

The characteristics of isolabel contours suggest that we can analyze isolabel-contour maps to determine the approximate boundary of an object. This analysis is based on matches of pairs of successive isolabel contours, or contours that result from successive intensity thresholds. For example, in Fig. 3, contours that enclose regions r_1 and r_2 are successive contours. Note that successive isolabel contours may consist of disjoint point sets, or may share a subset of points. We use the following definition in our analysis. Given an ordered set of successive contours $C = c_1, \dots, c_{n-1}, c_n$, and a distance measure $D(c_i, c_j)$ that assigns scores to pairs of successive contours c_i, c_j according to how well the contours match, C is a pattern of contours with similar shape if and only if, for every successive pair of contours $c_{i-1}, c_i \in C, D(c_{i-1}, c_i) < T_m$, where T_m is a predefined distance value. We observed that large shape changes in sequences of successive contours are typically associated with a transition from within an object into the background. Thus, we regard deviations from a pattern of contours with similar shape as a cue that suggests the presence of an object boundary in the vicinity of where the deviation occurred. Fig. 4 shows a set of successive isolabel contours derived from a CTA section.

We have presented the theoretical underpinning of our method for identifying boundaries of coherent regions. Section III-B describes the procedures that we implemented for identifying approximate boundaries through extraction and matching of isolabel contours.

B. Algorithms

The goal of our method is to portray objects of interest as coherent regions that a user can select easily by clicking with a mouse. Following a preprocessing task that removes noise from raw images with an anisotropic diffusion filter [23], the method generates isolabel-contour maps by thresholding the image intensity range with a fine partition (see Fig. 2). Then, the method extracts isolabel contours from the maps and compares their shapes to detect object boundaries (see Fig. 4). The detected boundaries are estimates of the true boundaries. The probability of observing only small estimation errors increases with an increase in the number of thresholds. The method designates regions that the detected boundaries circumscribe by assigning grayscale codes to voxels enclosed within these boundaries. We refer to the resulting regions as **salient regions**. We implemented our method within a program that we call intrinsic shape for segmentation (ISeg). We now explain the details of processing the contours within the program.

1) *Extraction of Isolabel Contours*: ISeg's method for extracting isolabel contours is based on the following observation. When we follow the thresholds in an intensity-range partition in monotonic order, and follow the corresponding projections that

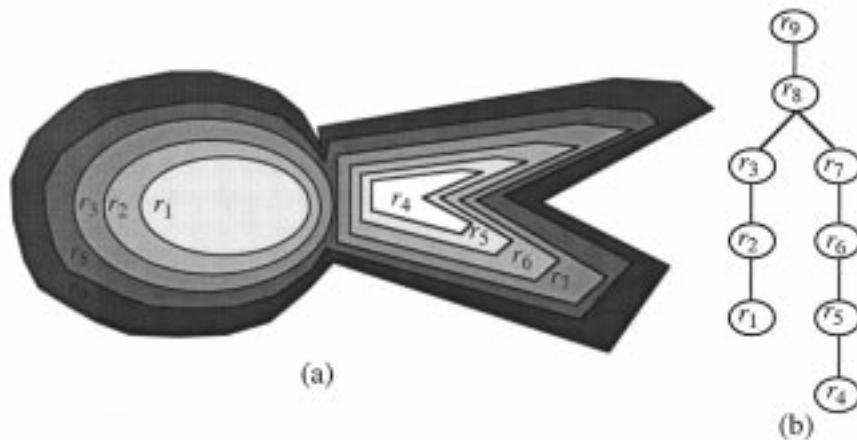


Fig. 3. A hierarchical description of regions and their corresponding contours in a labeled image. (a) A hypothetical labeled image for a vessel and a bone that are close to each other and that have a similar intensity range (b) The corresponding region tree. The vessel is centered at region r_1 ; the bone is centered at region r_4 . Because of the proximity of the vessel and the bone structures, there are isolabel contours that circumscribe both objects, and those are regarded as part of the background. The two objects are represented by two separate branches in the tree structure.

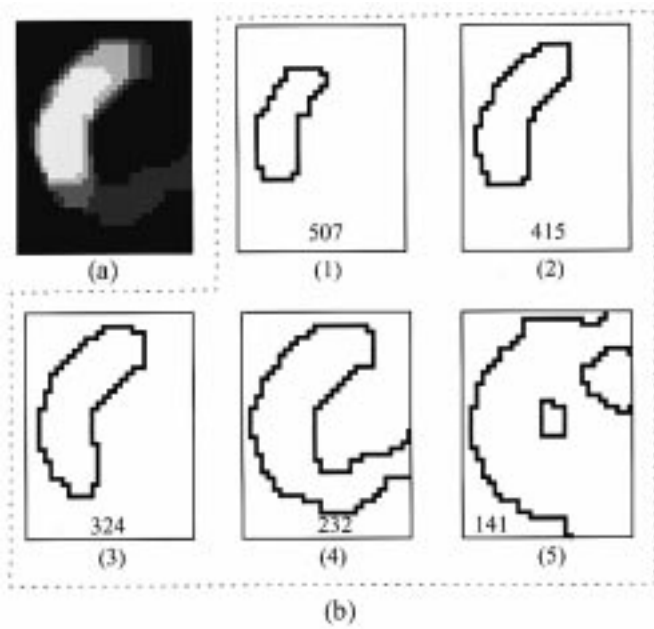


Fig. 4. Isolabel contours from the CTA section in Fig. 2. (a) An excerpt from the rightmost third of the isolabel-contour map. (b) (1-5) Several contours along the perimeter of part of the tortuous vessel. The intensity threshold in HU is indicated for each isolabel contour. Note the shape similarity of contours 1 through 4, and the shape difference between contours 4 and 5. The shape difference suggests a transition from within the vessel into the background.

the thresholds form on the image plane, we observe that the projections extend in a specific direction with respect to the object (see Fig. 1). Similarly, for CTA sections where the target objects are bright, if we examine labels from an isolabel-contour map in descending order of associated intensity we encounter isolabel contours that extend from within the objects to their surroundings. ISeg examines isolabel contours in order of their associated intensities, and thus guarantees that the program considers contours along the boundary of an object. For CTA data, ISeg considers labels in decreasing order, extracts the respective contours, and then matches each pair of successive contours. The

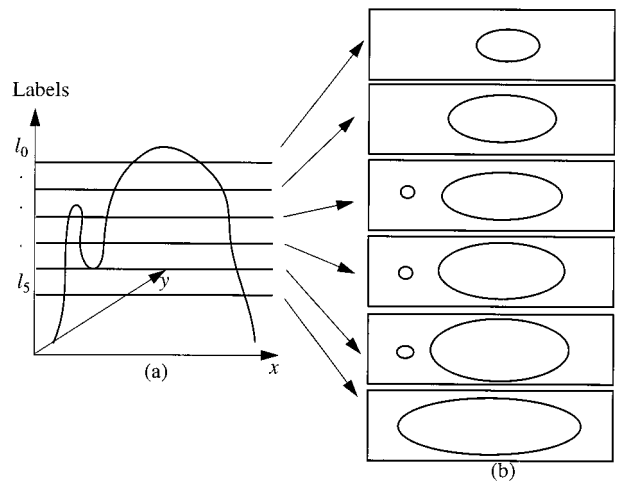


Fig. 5. Contours that ISeg extracts for a hypothetical object with a nonmonotonic intensity profile. (a) The intensity profile along a section that is parallel to the x-intensity plane, and five intensity thresholds. (b) Contours that the program extracts.

effective matching order for every object is from the innermost contour to the outermost contour.

We can view the order in which ISeg considers isolabel contours as a bottom-up, breadth-first traversal of the hierarchical representation of the contours (Fig. 5). The program maintains a separate data-structure instance for each path of successive contours within the representation tree. Each instance includes the list of coordinates for the most recent contour that the program considered for that path. The following procedure illustrates the course in which ISeg considers contours for an arbitrary iteration j .

- 1) Given a current label l_j , find all regions with that label that have not been visited. The regions designate entrances into new subtrees. Establish a new path structure for the new subtrees.
- 2) For each existing path, find the contour that is associated with label l_{j+1} , and compare the contour to the contour associated with label l_j for that path.

The procedure is adequate for objects with intensity profiles that have a single peak and for objects with intensity profiles that have multiple peaks. For example, consider the object in Fig. 5. The intensity profile for the object has more than one peak. ISeg maintains a single path at l_0 and l_1 ; then, it establishes a new path at l_2 . The program maintains two paths at l_3 and l_4 ; it merges the two paths into one parent path at l_5 .

2) *Comparison of the Shape of Contours*: To compare the shape of two contours at a given iteration, ISeg extracts the contours, represents the contours as turning-angle sequences [24], and computes the difference between sequences. We now describe the details of these subtasks. At each iteration ISeg matches the current contour, which represents the state that the program reached in the previous iteration, to the next contour, which the program extracts from the contour map (Fig. 6). ISeg extracts the next contour by growing connected voxels with the specified label from voxels that are adjacent to the current isolabel contour from the outside. The voxels on the boundary of the enlarged region define the next isolabel contour. ISeg extracts these voxels by tracking eight-neighbor-connected¹ voxels with the specified label that have an eight-neighbor with a different label.

ISeg follows several steps to generate turning-angle sequences of a contour. The program decomposes the contour into functions $x(t), y(t)$, in parametric space with parameter t , and smooths the contour by fitting a cubic spline to a sample of equispaced domain points. Smoothing the contours reduces the effects of noise and geometric distortions (for small objects) on the results of contour matching. However, the degree of smoothing must be such that it does not eliminate object shape characteristics. Note that the splines affect only the matching process and do not define coherent regions directly. When ISeg finds that two contours are similar, it includes in the current region the voxels to which it fit the outer spline.

ISeg computes a spline in the following way. For a contour of n points, the program samples $\lceil n/4 \rceil$ control points from the contour, fits a spline to the resulting control-point sequence, and samples $n + 1$ points from the continuous spline. To reduce the sensitivity of contour matching to the selection of control points, ISeg computes four splines for each contour, each spline with a different set of control points, and then averages the resulting coordinates for spline points that correspond in the original contour's parameter space.

Finally, ISeg represents the shape of the final spline by computing a sequence of turning angles [24], where each turning angle represents the slope of the line connecting two successive points in the average spline. The turning-angle representation of a contour depicts direction changes of a hypothetical walk along the contour, and thus it implicitly encompasses the shape of the contour. Only the contour itself is necessary to derive this representation.

ISeg assigns a similarity score to each pair of successive contours by performing a nonlinear elastic match between the respective sequences of turning angles [25]. The matching method is based on the notion of a *mapping* between two sequences that

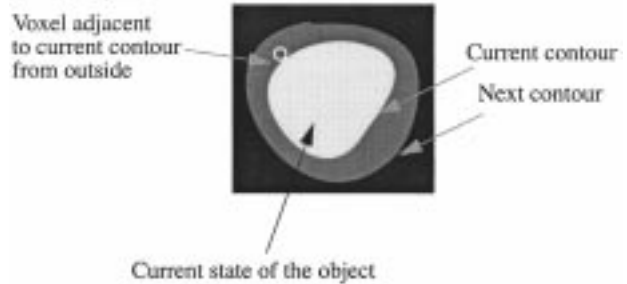


Fig. 6. The relative position of the current contour and the next contour at a given contour-matching iteration. ISeg identifies the next contour by applying connectivity to voxels that are adjacent to the current contour from its outside.

assigns all elements in one of the sequences to an element in the other sequence. Because an element in one sequence can be mapped to more than one mode element in the other sequence, there exist multiple mappings between the two sequences. The matching method examines all possible mappings, and for each mapping, computes a mapping score—the average absolute difference between corresponding sequence elements. The similarity score $D(x, y)$ is the minimum score over all possible element mappings of sequences x and y . The matching algorithm uses dynamic programming to find the mapping with the minimal score [26]. Fig. 7 shows the turning-angle sequences for two successive contours and portrays the mapping that produces the minimum score (0.04 radians).

The hierarchical description of isolabel contours allows for efficient matching of contours in an isolabel-contour map. ISeg tracks parent-child relations between regions, and terminates matching of contours for paths that reach a shared parent that has been processed already. Thus, the program processes each contour at most twice—once as the first contour in a pair and once as the second contour—and not once for every path in the tree that includes the contour. To understand the point at which the hierarchical representation saves computation, we analyze the performance of the contour matching process. We assume that there are n thresholds and, therefore, $n - 1$ labels, and that each contour in the isolabel-contour map encloses directly² exactly q contours. Under this assumption there are q^{n-2} terminal contours, which do not enclose any contours within. If we compare the shape of successive contours in an exhaustive manner, the total number of contour comparisons is

$$(n - 2) \cdot q^{n-2}. \quad (1)$$

If we exploit the hierarchical representation of the contours to avoid unnecessary comparisons, the number of comparisons is equal to the number of branches in the tree

$$\sum_{i=1}^{n-2} q^i. \quad (2)$$

The expressions in (1) and (2) can both be viewed as sums of $n - 2$ powers of q . All powers of q in (2) are smaller than those

¹Eight-neighbor voxels are oriented with respect to each other like the eight major directions on a compass.

²Contour a directly encloses contour b if there is no other contour c that encloses b where a also encloses c .

$$A = \{6.20, 6.25, 0.18, 0.50, 0.88, 1.18, 1.41, 1.59, 1.82, 2.12, 2.54, 3.01, 3.44, 3.78, 4.03, 4.21, 4.37, 4.57, 4.85, 5.24, 5.61, 5.90, 6.09, 6.19\}$$

$$B = \{6.24, 0.02, 0.21, 0.43, 0.64, 0.89, 1.22, 1.58, 1.87, 2.08, 2.30, 2.61, 3.08, 3.55, 3.87, 4.10, 4.28, 4.42, 4.55, 5.72, 5.02, 5.44, 5.80, 6.03, 6.15\}$$

(a)

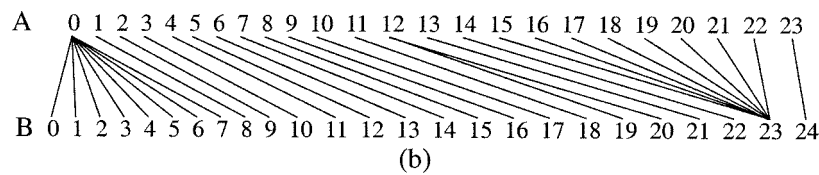


Fig. 7. Turning-angle sequences and their optimal mapping. (a) Angle sequences for contour A and that for contour B in radians. (b) A mapping of ordinal elements from the two sequences that produces the minimal matching score. Note that because the minimal matching score considers the differences between corresponding elements throughout the entire sequences, the difference between a given element in one sequence and its matching element from the other sequence may not be the minimal difference possible for that element.

in (1), except for the final one, which is equal to the powers of q in (1). Therefore

$$\sum_{i=1}^{n-2} q^i < (n-2) \cdot q^{n-2} \quad (3)$$

and we conclude that exploiting the contour hierarchy allows us to reduce the computation associated with comparison of successive isolabel contours. From (3), we observe that we obtain the maximum reduction in computation

$$(n-2) \cdot q^{n-2} - \sum_{i=1}^{n-2} q^i \quad (4)$$

from a tree-structure representation when the branching factor is greater than one and the tree is balanced. In practice, however, the branching factor varies throughout the tree, and the difference between computation with and without the tree structure is smaller than the difference indicated by (4).

As an example, a randomly sampled CTA case with 158 sections contained up to seven branches for a single branching point, and a total of 4662 branching points. Without the hierarchical representations, the program performed a total of 1 003 332 contour comparisons. With the hierarchical representation, the program performed 223 182 contour comparisons. Thus, the hierarchical representation yielded a reduction in the number of contour comparisons by a factor of 4.5.

3) *Results of contour matching:* The contour matching process results in an image with one or more salient regions. For example, for the hypothetical objects in Fig. 3, the program notes two large shape changes. The large shape change between the contours that circumscribe regions r_3 and r_8 designates a transition from within the vessel into the background, and the large shape change between the contours that circumscribe regions r_7 and r_8 designates a transition from within the bone into the background. Thus, the salient region for the vessel includes all the voxels that r_3 encloses, and the salient region for the bone includes all the voxels that r_7 encloses. Fig. 8 shows salient regions for the section in Fig. 2.

Each salient region is surrounded by a zero background (black in the figure) and consists of one or more sets of

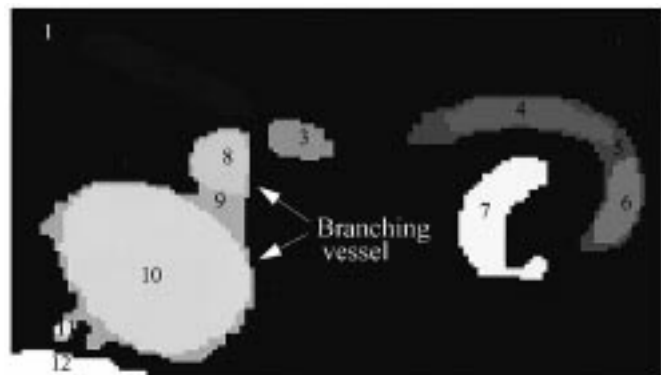


Fig. 8. Demonstration of salient regions that resulted from a contour map generated by 14 thresholds for the image in Fig. 2(a). The numeric labels indicate distinct grayscale codes.

connected voxels, where each set is labeled with a distinct grayscale code. Different grayscale codes for a single salient region result from multiple occurrences of contours with a significant shape difference; the multiple codes may suggest the presence of substructures within a larger object—for example, branches, within a vascular tree. However, multiple grayscale codes could also result for a single structure, when the program computes a shape difference just above the shape-difference threshold for contours that humans might perceive as similar. In this case, we would observe sets of connected voxels with different gray scale codes such that one set encloses the other.

ISeg maintains a hierarchical description of salient regions to support efficient automated post-processing of salient regions—for example, selection of all subregions that are included within a region that the user selected with a mouse.

IV. EVALUATION

We evaluated the performance of ISeg with three separate experiments. First, we performed a sensitivity analysis to determine the robustness of ISeg to small changes in parameter values. Second, we compared the robustness of ISeg to robustness of conventional thresholding. Finally, we compared the results of ISeg to those of expert manual editing.

TABLE I
F VALUES FOR PARAMETERS AND PARAMETER INTERACTIONS

Parameter (or interaction)	F value	Pr > F	Parameter (or interaction)	F value	Pr > F
LI	1.01	0.3156	LI*SDT	0.83	0.3636
UI	0.78	0.3777	UI*SDT	1.87	0.1722
LI*UI	0.35	0.5552	LI*UI*SDT	0.84	0.3605
NT	1.34	0.2471	NT*SDT	1.06	0.3030
LI*NT	1.48	0.2237	LI*NT*SDT	1.08	0.2996
UI*NT	1.25	0.2639	UI*NT*SDT	1.23	0.2668
LI*UI*NT	0.99	0.3189	LI*UI*NT*SDT	0.82	0.3654
SDT	1.56	0.2123			

Parameter interactions are designated by '*'. Pr>F denotes the area under the curve of the F distribution that is to the right of the F value; Pr>F is the probability of rejecting the null hypothesis when it is true. The unit of analysis is a region. The dependent variable is NVM.

A. Sensitivity Analysis

The goal of the first experiment was to evaluate the sensitivity of ISeg to 5% changes in parameter settings. The parameters that we examined were lower intensity (LI) and upper intensity (UI) that we expected of the target object, number of intensity thresholds (NT), and shape-difference threshold (SDT). Through experimentation with four CTA cases, we determined settings for these parameters that produced good results according to our subjective judgment: LI = 49 HU, UI = 598 HU, NT = 170, SDT = 0.048. We wished to determine whether small changes in these parameters produced results that were significantly different from the results that we generated with the original parameter setting. In the sensitivity analysis, we investigated the effect of 5% changes. Investigation of additional change rates would yield a more elaborate analysis of the robustness of our method.

We selected a random set of 11 cases from an archive of abdominal-pelvic CTA cases that Stanford University acquired with a General Electric Helical CT scanner in 1997. The section spacing was 2 mm and a section thickness was 3 mm. The in-plane resolution of the images ranged from 0.39 mm to 0.7 mm. The number of sections in each case ranged from 152 to 198. We sampled three sections from each case, for a total of 33 sections. We activated ISeg to identify salient regions in the collection of images for the original parameter settings; then we activated the program for 16 combinations of the four parameters at two levels each, where, for each parameter, each level constituted a 5% change, either above or below the original setting. We considered the results that ISeg produced for the original parameter setting as the set of **baseline images**, and the results that ISeg produced for the 16 parameter-level combinations as 16 sets of **treatment images**. We compared pairs of corresponding regions from baseline and treatment images. We assumed that significant sensitivity of ISeg to parameter changes would reflect in notable differences of segmentation results between the paired regions.

We compared artery regions from each of the 16 treatment sets to corresponding regions in the baseline set, a total of 1244 pairs. For each pair of regions we computed the number of

voxel mismatches. To obtain a measure of voxel mismatch relative to the size of the regions we divided the number of mismatches by the number of voxels in the baseline region, and thus obtained the **normalized voxel mismatch (NVM)**. For salient regions that included multiple grayscale codes—this situation could result from multiple contour pairs for which the shape difference exceeded the SDT—we recorded the smallest NVM value, which represented the best-case scenario. We performed an analysis of variance for a four-way classification to determine whether any of the four parameters or parameter interactions explained mismatches between regions in treatment images and corresponding regions in baseline images. The null hypothesis was that *none of the parameters (LI, UI, NT, SDT), or any combination of these parameters, explained the differences in corresponding regions from the baseline images and the treatment images.*

The results of the analysis of variance showed that there was not sufficient evidence to reject the null hypothesis at $p < 0.05$, for all parameter combinations or for LI, UI, NT, and SDT alone. Table I shows the **F value** and **p value** for each parameter or parameter combination. The table shows the extent to which the parameter or parameter combinations explain the differences observed between the groups. From the table we see that, if we rejected the null hypothesis for each parameter or parameter combination—we assumed that the parameter or parameter-level combinations explained the differences in the NVM values for the 16 treatments—then we would risk making errors at the rates specified in the third and sixth columns. Because we did not have sufficient confidence to reject the null hypothesis, we accepted it. Power calculations showed that, at $\alpha = 0.05$, $1 - \beta$ was greater than 80%. We concluded that ISeg was not sensitive to 5% changes in the parameters LI, UI, SDT, and NT. These findings suggested that, for CTA data, the initial setting of LI, UI, SDT, and NT was likely to be adequate.

B. Comparison of ISeg and Intensity Thresholding

The goal of the second experiment was to compare the robustness of ISeg and the robustness of conventional intensity thresh-

olding. ISeg and conventional thresholding both use a threshold value as part of the segmentation process. However, conventional thresholding applies a criterion in intensity space and uses an intensity threshold to determine object membership of voxels, whereas ISeg applies a criterion in shape space and uses a shape-difference threshold to determine object membership of voxels. We wished to determine whether ISeg was more robust than conventional thresholding, when we examined the robustness of each method with respect to its individual threshold settings. The experiment used the same paradigm as the sensitivity-analysis experiment to assess robustness: comparison of baseline images, which ISeg generated with an initial set of parameters, to treatment images, which ISeg generated with alternate parameters. However, unlike the sensitivity analysis, in this experiment we compared the robustness of two different segmentation methods.

For this study, we used the same raw dataset and measures as in the sensitivity analysis (see Section IV-A). First, we ran ISeg 11 times, each time with a different value for SDT. The values for SDT ranged from 0.036 to 0.06, and represented 5% increments over a predefined center-point value of 0.048. The fixed parameter values were $LI = 49$ HU, $UI = 598$ HU, and $NT = 170$ —the same values we used for the baseline images in the sensitivity analysis. We regarded the results that ISeg produced for $SDT = 0.048$ as the baseline images, and the results that ISeg produced for all other SDT values as treatment images. We computed the NVM for corresponding artery regions in baseline and treatment images.

Second, we measured the sensitivity of conventional thresholding to the intensity threshold; we performed intensity thresholding for our dataset 11 times, each time with a different intensity threshold. The intensity thresholds ranged from 150 HU to 250 HU, and represented 5% increments over a predefined center-point intensity threshold of 200 HU. We regarded the results from the threshold of 200 HU as the baseline images, and the results from all other threshold values as treatment images. We computed the NVM for corresponding artery regions in baseline and treatment images. For each pair of corresponding baseline and treatment images we computed the mean NVM (MNVM) from the NVM values of all region pairs.

Third, we compared the results of the two methods for corresponding percent threshold increments; for each increment, we performed a paired t -test to test the hypothesis that the *MNVM values for ISeg were smaller than those for conventional thresholding*. The paired t -test compared the MNVM for values for each pair of corresponding images, one resulting from intensity thresholding and the other resulting from ISeg. We considered for the test only those regions that had overlapping counterparts in the baseline images (94% for conventional thresholding results, and 99% for ISeg's results).

To obtain a meaningful statistical analysis, we concentrated on only observation with a reasonable standard error, and we excluded 51 observations (out of 749) for which conventional thresholding resulted in notable leaks, or protrusions, into adjacent structures. We believed that because these observations indicated poor results on the part of conventional thresholding, excluding the observations would not bias the analysis in favor of ISeg. Table II shows the results of the paired t -test; Fig. 9

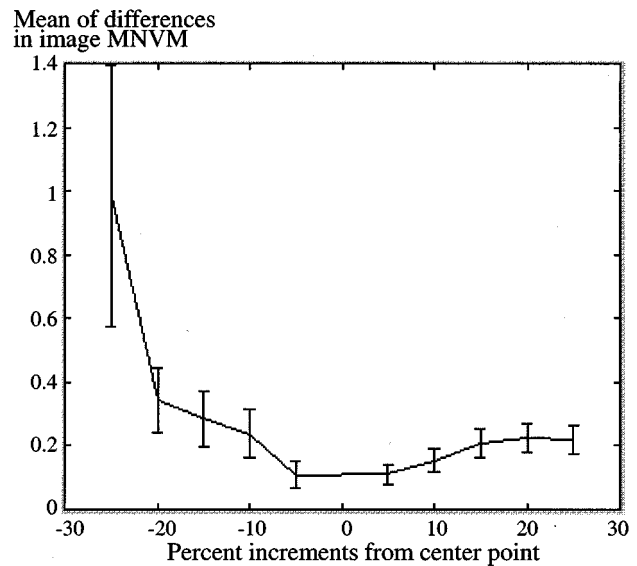


Fig. 9. Error-bar plot of mean difference in image MNVM for conventional thresholding and for ISeg. The positive differences in MNVM values indicate that MNVM values for conventional thresholding were larger than those for ISeg at all increments.

TABLE II
RESULTS OF PAIRED t -TEST FOR COMPARISON OF CONVENTIONAL THRESHOLDING AND ISeg

Percent increment	Mean of differences	Standard error	T value	$\text{Pr} > T $
-25	0.9838334	0.4080844	2.4108574	0.0218
-20	0.3413646	0.1000427	3.4121885	0.0018
-15	0.2854101	0.0871485	3.2749844	0.0025
-10	0.2389239	0.0742975	3.2157736	0.0030
-5	0.1093748	0.0438471	2.4944592	0.0180
5	0.1104296	0.0312476	3.5340196	0.0013
10	0.1541976	0.0377041	4.0896801	0.0003
15	0.2101106	0.0446525	4.7054570	0.0001
20	0.2240587	0.0444016	5.0461802	0.0001
25	0.2201652	0.0432196	5.0941033	0.0001

shows an error-bar plot for the results of the test. From the results we see that, in general, the mean differences and standard error of the mean for MNVM increased with an increase in the magnitude of the increment. We concluded that, where there were no leaks from objects into adjacent structures in conventional thresholding, ISeg's method for finding object boundaries in shape space was less sensitive to changes in threshold values compared to conventional thresholding in intensity space.

C. Comparison of Editing with ISeg to Manual Editing

The goal of the third experiment was to determine whether artery salient regions that ISeg produced for CTA data were comparable to experts' manual editing of arteries from these data. We recruited five radiology experts who had experience in editing medical image volumes using conventional editing tools. The editing tools—intensity thresholding, connectivity

TABLE III
RESULTS OF PAIRED t -TEST FOR COMPARISON OF EXPERT-EXPERT DISPARITIES TO COMPUTER-EXPERT DISPARITIES

	Set 1 Expert-expert MNVM	Set 1 Expert-expert NVM variance	Set 2 Computer- expert MNVM	Set 2 Computer- expert NVM variance	Pr> T
t -Test 1	0.29	0.19	0.5	6.56	0.114
t -Test 2	0.29	0.19	0.32	0.09	0.175

Note the reduction in variance of NVM for Set 2 after removal of three outlier regions from the set. After removal of the outliers, the mean disparity measures (MNVM) for Set 1 and Set 2 were similar.

tracking, and manual tracing—were available within the General Electric, Inc., Advantage Windows Workstation. In the following discussion, we refer to the use of these editing tools as **manual editing**. Because we observed considerable variability within results that different experts produced for the same regions (Table III), we could not use expert edits as a definitive gold standard for evaluating the accuracy of ISeg. Alternatively, we evaluated ISeg by comparing the disparities between results that ISeg produces and results that multiple experts produced for a set of images to disparities between results that different expert produced for the same set of images. We believed that if the disparities between ISeg's results and those of manual editing were similar to interexpert disparities, the results of both editing methods were comparable.

We selected randomly a set of 75 CTA sections from 16 patients, and arranged the section in six sets. To determine variability of expert manual editing, we had five experts independently edit arteries from the sections in all the sets. To reduce the effect of the order of editing on the results, the experts edited the sections by subsets in random order. To reduce the effect of the timing and the order of subset editing for a given day on the results, the experts edited at most one subset per day.

To obtain ISeg's data, we selected all artery regions from the results that ISeg produced for the 75 sections. To compare salient regions that ISeg generated to manually edited sections, we generated two sets of observations. Set 1 included NVM figures for pairs of corresponding regions that different experts edited; these figures measured interexpert variability. Set 2 comprised NVM figures for pairs of corresponding regions, where an expert generated one region via manual editing, and we generated the other region with ISeg. For both Set 1 and Set 2 we computed the MNVM per image. To compare the disparity between manual-editing results and ISeg results to interexpert variability, we tested the hypothesis that sets 1 and 2 differed significantly with a paired t -test that compared MNVM values for 75 corresponding sections. The null hypothesis was that *MNVM values for sections in sets 1 and 2 did not differ significantly*.

We conducted the paired t -test twice: once with the original Set 2, and once after we excluded three outlier regions (0.5% of all regions) for which the differences between the expert's edit and the computer results were very large—as measured in NVM of 45, 35, 17. There were no outliers of similar magnitude in Set 1, as can be seen from Table III. The results of both tests indicated that, if we rejected the null hypothesis, we would risk falsely concluding that the MNVM values in sets 1 and 2

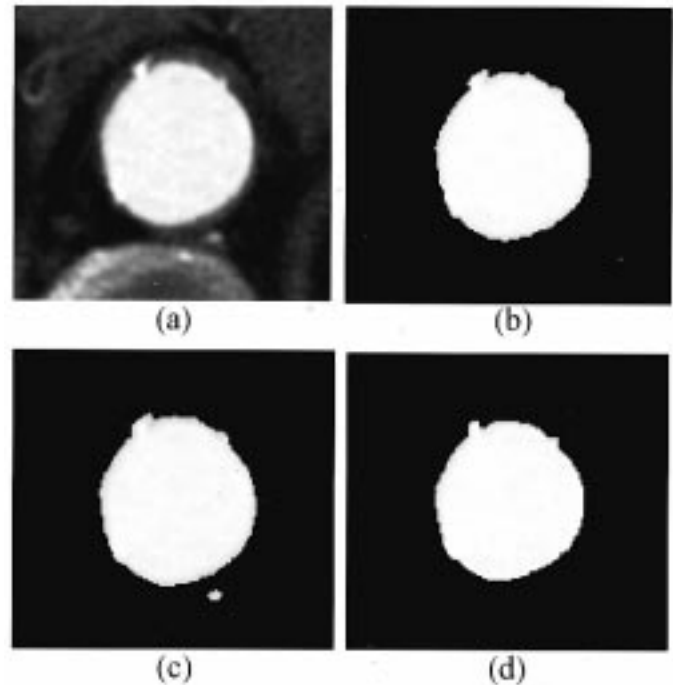


Fig. 10. Edited sections of the aorta that experts and ISeg generated. (a) Raw section (b)–(c) Sections that experts generated via manual editing. (d) A section that ISeg generated.

differed significantly, with $p > 0.114$ and $p > 0.175$, respectively. Therefore, we accepted the null hypothesis. Table III summarizes the results of the statistical analysis. With a sample size of 75 images, and for $\alpha = 0.05$, the power of the test was 0.99. We concluded that the salient regions that ISeg generated were comparable to those that experts generated via manual editing. Fig. 10 demonstrates edited sections that experts and ISeg generated.

V. DISCUSSION

We now turn to a discussion of the contributions and the limitations of our work. The contribution of ISeg is in ISeg's capability to automatically extract visual information that is necessary to support diagnosis, quantification, and therapy planning. ISeg mimics the way experts perform segmentation via intensity thresholding through perception of shape changes. ISeg takes

advantage of PVE, which is inherent in image acquisition regardless of imaging modality. Because of PVE, imposing a fine partitioning of the intensity range results in an isolabel-contour map in which contours that reside along the boundaries of objects carry the shape of the object and thus have a similar shape.

ISeg interprets successive isolabel contours that have similar shape as an indication for the shape of an object, and regards contours with different shape as designating transition from within an object into the background. With this intrinsic-shape criterion, the program can meet the challenge of distinguishing between objects of similar appearance that are close to each other, regardless of their specific intensities or shapes. The shape criterion allows ISeg to accommodate PVE, which causes intensity distortions along the boundaries of objects. ISeg fits splines to contours prior to matching the contours to reduce the effects of noise and shape distortion on the matching results. The hierarchical representation of isolabel contours allows the program to perform the contour-matching procedure efficiently.

To the best of our knowledge, comparing the shape of isolabel contours to detect transitions from within an object into the background is a novel idea. Our method relates to the segmentation method that Zheng *et al.*[27] developed in that both methods analyze features of regions that result from thresholding to determine the region that best represents the target object. Zheng *et al.* use three coarsely spaced thresholds and general features such as size growth and central-position shift. We use a finer partition of the intensity range and features more closely related to the boundaries of the objects. Our method also relates to the image registration method of Amit [28] in that both methods analyze contour patterns to perform image-analysis. Amit used local topography, or patterns of shape-similar isolabel contours, to identify landmarks for registering images to templates.

ISeg uses shape information that it extracts from isolabel contours in labeled images—it does not rely on external prior shape models to find object boundaries. ISeg requires only a fairly lax specification of the intensity range for target objects; the sensitivity analysis showed that segmentation results were insensitive to 5% changes in the upper and lower intensities that the user specified for the target objects. The sensitivity analysis and the comparison to conventional thresholding demonstrated that ISeg was robust for segmenting CTA data. ISeg’s robustness results in part from the fact that the program reduces the effects of noise by fitting smooth splines to isolabel contours before it compares the shape of contours. From the third evaluation experiment we concluded that the boundaries that ISeg found differed from boundaries that experts delineated using conventional editing tools as much as the latter boundaries differed from one another. A paired t-test showed that the computer-expert editing disparities were insignificantly different from expert-expert editing disparities.

Note that the results of the evaluation represent the best potential scenario for the evaluation datasets. Because salient regions may include multiple regions with different grayscale codes for a single object (for example, see Fig. 8), and because each of these regions could potentially designate the target object, ISeg requires a post-processing procedure to select the final target objects. Whether the best depiction of a target object is obtained

in practice depends on the post-processing procedure used. For example, a user can designate target structures by clicking with a mouse on specific grayscale regions. Selecting a known segmented structure with a mouse is much easier and more reproducible than tracing around the structure in a raw section. For a three-dimensional (3-D) object, the user could select a target structure from only several images, and a program could then automatically include other connected segments. In [29], we described a semiautomated region-selection method that used salient regions that ISeg generated to extract 3-D vessel trees from CTA data.

The limitations of the method that we developed arise from the underlying assumptions that characterize target datasets for which the method is likely to succeed. The first assumption is that there exist intensity thresholds that separate between target objects and structures that surround them. This assumption is less likely to hold as the ratio of noise standard deviation to the contrast between the true object and its background increases. Although noise-removal methods could improve the contrast-to-noise ratio, it is likely that a reduced level of noise will remain. For CTA images, ISeg was able to separate the aorta flow channel from surrounding structures—for example, the vena cava; the typical contrast in these images was about 200 HU, and the standard deviation of the noise in the arteries was about 25 HU. However, ISeg was not able to separate thrombi from the vena cava; the contrast was only about 20 HU, and the standard deviation of the noise within the thrombi was similar to that within the aorta.

The second assumption is that sections of target objects are at least 50 pixels in area. For small objects, whose shape is distorted due to discretization, the intensity fall that is associated with PVE is not as smooth as it is for large objects, and thus the shape of isolabel contours along the boundary of the object image does not reflect well the shape of the underlying object. Consequently, the contour-matching procedure is likely to detect large shape changes for contours that enclose voxels that belong to the same object. This false detection either would lead to elimination of a salient region for the object—in the event that ISeg failed to detect two successive contours with similar shape—or might compromise the coherence of ISeg’s depiction of the salient region.

VI. CONCLUSION

On-going technology innovations increase the ease with which we can obtain high-quality large volumetric image data sets to support medical care. We need accurate and efficient image-processing tools to realize the potential benefits embodied within image information. We described a promising method and a program that could help radiology experts to edit large volumetric datasets. Our evaluation experiments showed that ISeg was robust and superior to conventional thresholding for delineating arteries in CTA data, and that the program produced results that were comparable to results of manual tracing. Because our method does not depend on the content of images or on the acquisition process, we expect that the methods will generalize beyond the scope of CTA data.

ACKNOWLEDGMENT

The authors would like to thank M. Walker, C. H. Yan, S. Y. Yen, and D. Paik for useful discussions, and R. Altman and R. Motwani and B. Madore for reviewing this work. They would like to thank R. Shifrin, C. Beaulieu, S. Heiss, B. Daniel, and L. Logan for their editing of CTA cases for their evaluation experiments. Finally, they would also like to thank the reviewers for their helpful critique.

REFERENCES

- [1] E. Chaney and S. Pizer, "Defining anatomical structures from medical images," in *Seminars Radiat. Oncol.*, vol. 2, 1992, pp. 215–225.
- [2] G. Tracton, E. Chaney, J. Rosenman, and S. Pizer, "Mask: Combining 2-D and 3-D segmentation methods to enhance functionality," in *Proc. SPIE Conf. Medical Imaging*. Bellingham, WA, 1994, pp. 98–109.
- [3] J. F. Brinkley, "A flexible, generic model for anatomic shape: Application to interactive two-dimensional medical image segmentation and matching," *Comput. Biomed. Res.*, vol. 26, no. 2, pp. 121–142, 1993.
- [4] W. Lorensen, A. Ferenc, and R. Kikinis, "The exploration of cross-sectional data with a virtual endoscope," in *Interactive Technology and the New Paradigm for Health Care*, R. Mattheus and J. Christensen, Eds. Ohmsha, Japan: IOS Press, , 1995, pp. 221–230.
- [5] S. C. Orphanoudakis, C. Chronaki, and S. Kostomanolakis, "I^{sup} 2/C: A system for the indexing, storage, and retrieval of medical images by content," *Med. Informatics*, vol. 19, pp. 109–122, 1994.
- [6] R. A. Robb and C. Barillot, "Interactive display and analysis of 3-D medical images," *IEEE Trans. Med. Imag.*, vol. 8, pp. 217–226, June 1989.
- [7] D. Hentschel, J. Ezrielev, R. Fisher, C. Flanders, A. Bani-Hashemi, C.-C. Liang, S.-P. Liou, S. Samaddar, S. Singh, and D. Ney, "Techniques for editing and visualizing CT-angiographic data," in *Proc. Conf. Visualization in Biomedical Computing*, R. A. Robb, Ed., 1994, pp. 307–318.
- [8] J. Duncan, L. Staib, T. Birkholzer, R. Owen, P. Anandan, and I. Bozma, "Medical image analysis using model-based optimization," in *Proc. 1st Conf. Visualization in Biomedical Computing*. Atlanta, GA, 1990, pp. 370–377.
- [9] L. H. Staib and J. S. Duncan, "Model-based deformable surface finding for medical images," *IEEE Trans. Med. Imag.*, vol. 15, pp. 720–731, Oct. 1996.
- [10] M. Kamber, R. Shinghal, D. L. Collins, G. S. Francis, and A. C. Evans, "Model-based 3-D segmentation of multiple sclerosis lesions in magnetic resonance brain images," *IEEE Trans. Med. Imag.*, vol. 14, pp. 442–445, June 1995.
- [11] G. Christensen, S. Joshi, and M. Miller, "Volumetric transformation of brain anatomy," *IEEE Trans. Med. Imag.*, vol. 16, pp. 864–877, Dec. 1997.
- [12] A. Kelemen, G. Szekely, and G. Gerig, "Elastic model-based segmentation of 3-D neuroradiological data sets," *IEEE Trans. Med. Imag.*, vol. 18, pp. 828–839, Oct. 1999.
- [13] L. Clarke, R. Velthuizen, M. Camacho, J. Heine, M. Vaidyanathan, L. Hall, R. Thatcher, and M. Silbiger, "MRI segmentation: Methods and applications," *Magn. Reson. Imag.*, vol. 13, pp. 343–368, 1995.
- [14] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Computer Vision*, vol. 1, pp. 321–331, 1987.
- [15] L. D. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-D and 3-D images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1131–1147, Nov. 1993.
- [16] G. Szekely, A. Kelemen, C. Brechbuhler, and G. Gerig, "Segmentation of 3-D objects from MRI volume data using constrained elastic deformations of flexible Fourier surface models," in *Proc. Conf. Computer Vision, Virtual Reality and Robotics in Medicine*, Nice, France, 1995, pp. 495–505.
- [17] G.-H. Chuang and C.-C. Kuo, "Wavelet descriptor of planar curves: Theory and applications," *IEEE Trans. Image Processing*, vol. 5, pp. 56–70, Jan. 1996.
- [18] T. McInerney and D. Terzopoulos, "Medical image segmentation using topologically adaptable snakes," in *Proc. Conf. Computer Vision, Virtual Reality, and Robotics in Medicine*, Nice, France, 1995, pp. 92–101.
- [19] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 158–175, Feb. 1995.
- [20] H. Cline, W. Lorensen, R. Kikinis, and F. Jolesz, "Three-dimensional segmentation of MR images of the head using probability and connectivity," *J. Comput. Assist. Tomogr.*, vol. 14, pp. 1037–1045, 1990.
- [21] H. Cline, C. Dumoulin, W. Lorensen, S. Souza, and W. Adams, "Volume rendering and connectivity algorithms for MR angiography," *Magn. Reson. Med.*, vol. 18, pp. 384–394, 1991.
- [22] P. Sahoo, S. Soltani, A. Wong, and Y. Chen, "A survey of thresholding techniques," *Comput. Vision Graphics Image Processing*, vol. 41, pp. 233–260, 1988.
- [23] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 629–639, July 1990.
- [24] E. M. Arkin, L. P. Chew, D. P. Huttenlocher, K. Kedem, and J. S. B. Mitchell, "An efficiently computable metric for comparing polygonal shapes," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 209–211, Mar. 1991.
- [25] R. Fagin and L. Stockmeyer, "Relaxing the triangle inequality in pattern matching," *Int. J. Comput. Vision*, vol. 30, no. 3, pp. 219–231, 1998.
- [26] G. Cortelazzo, G. A. Mian, G. Vezzi, and P. Zamperoni, "Trademark shapes description by sting-matching techniques," *Pattern Recogn.*, vol. 27, no. 8, pp. 1005–1018, 1994.
- [27] R. Zheng, Y. H. Chang, and D. Gur, "Computerized detection of masses in digitized mammograms using single image segmentation and a multilayer topographic feature analysis," *Academic Radiol.*, vol. 2, pp. 959–966, 1995.
- [28] Y. Amit, "Graphical shape templates for automatic anatomy detection with application to MRI brain scans," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, pp. 28–40, Jan 1997.
- [29] S. Shiffman, "Segmentation of medical image volumes using intrinsic shape information," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1999.