Image Segmentation via Adaptive K-Mean Clustering and Knowledge-Based Morphological Operations with Biomedical Applications

Chang Wen Chen, *Senior Member, IEEE,* Jiebo Luo, *Member, IEEE,* and Kevin J. Parker, *Fellow, IEEE*

***Abstract*—Image segmentation remains one of the major challenges in image analysis, since image analysis tasks are often constrained by how well previous segmentation is accomplished. In particular, many existing image segmentation algorithms fail to provide satisfactory results when the boundaries of the desired objects are not clearly defined by the image intensity information. In medical applications, skilled operators are usually employed to extract the desired regions that may be anatomically separate but statistically indistinguishable. Such manual processing is subject to operator errors and biases, is extremely time consuming, and has poor reproducibility. We propose a robust algorithm for the segmentation of three-dimensional (3-D) image data based on a novel combination of adaptive K-mean clustering and knowledge-based morphological operations. The proposed adaptive K-mean clustering algorithm is capable of segmenting the regions of smoothly varying intensity distributions. Spatial constraints are incorporated in the clustering algorithm through the modeling of the regions by Gibbs random fields. Knowledge-based morphological operations are then applied to the segmented regions to identify the desired regions according to the *a priori* anatomical knowledge of the region-of-interest. This proposed technique has been successfully applied to a sequence of cardiac CT volumetric images to generate the volumes of left ventricle chambers at 16 consecutive temporal frames. Our final segmentation results compare favorably with the results obtained using manual outlining. Extensions of this approach to other applications can be readily made when *a priori* knowledge of a given object is available.**

***Index Terms*—Cardiac imaging, clustering, Gibbs random field, image segmentation, K-mean, morphological operations.**

I. Introduction

THREE-DIMENSIONAL (3-D) image segmentation has attracted considerable attention for the last few years, due to the advances in multidimensional image acquisition techniques. In general, 3-D image segmentation algorithms are derived directly from two-dimensional (2-D) image segmentation algorithms. Many existing image segmentation algorithms are suited for certain classes of boundary extraction when *a priori* knowledge of the object in the image is available. These segmentation approaches can be divided into two major categories: optimization in parameter space and optimization in image space. In the case of boundary finding through optimization in parameter space, a class of parameterized templates are used to model the objects [1]–[3]. All these approaches have been implemented through fitting the model to the image data by searching the parameter space for the best fit. Recently, Staib and Duncan [4] proposed an algorithm for boundary finding with probabilistic deformable models in order to increase the flexibility of enforcing the constraints in parameter space. In the case of boundary finding through the optimization in image space, the measure of fit is represented by certain image-related quantities [5], [6]. Among them, an elegant approach proposed by Kass, Witkin, and Terzopoulos [7] uses the boundary model primitives called *snakes* to form an energy minimization problem. The flexibility of *snakes* enables its applications to many interactive image segmentation schemes when the skilled operator is employed to supervise the segmentation process. As opposed to the parameter space based boundary representation, the image space based representation and optimization are usually not efficient in the incorporation of the *a priori* shape information to the algorithm, especially those specific algebraic properties of the object boundary other than the smoothness constraints.

In biomedical-oriented image analysis research, the development of robust 3-D image segmentation techniques is an imperative task for the processing of the huge amount of volumetric biomedical images or image sequences produced by various medical imaging modalities. In general, the intensity of the same anatomical structure within a given set of volumetric image produced by medical imaging device is spatially varying because of the inevitable inhomogeneity in the process of image acquisition. In addition, a biomedical image analysis task often needs to extract the desired regions that may be separate in anatomy but indistinguishable in intensity. To segment such biomedical structures from the 3-D volumetric images, labor-intensive and operator-dependent outlining of region-of-interest has traditionally been employed [8]–[10]. Recently, there have been some attempts working toward a robust and operator-independent segmentation of desired biomedical structures from given volumetric images [11]–[14]. However, these algorithms would require either initial input or interactive interface of a human operator for

Manuscript received February 13, 1994; revised July 28, 1997. This work was supported by the National Science Foundation under Grant EEC-92-09615 and the 1996 Whitaker Foundation Biomedical Engineering Research Award. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Patrick A. Kelly.

C. W. Chen is with the Department of Electrical Engineering, University of Missouri, Columbia, MO 65211 USA (e-mail: cchen@ece.missouri.edu).

J. Luo was with the Department of Electrical Engineering, University of Rochester, Rochester, NY 14627-0231 USA. He is now with the Imaging Science Technology Laboratory, Eastman Kodak Company, Rochester, NY 14650 USA (e-mail: luo@image.kodak.com).

K. J. Parker is with the Department of Electrical Engineering, University of Rochester, Rochester, NY 14627-0231 USA.

Publisher Item Identifier S 1057-7149(98)08726-0.

every set of images to overcome the difficulties in segmenting the spatially varying intensity as well as in identifying the ambiguous anatomical structures. Such segmentation processes are able to produce better results than the simple and time-consuming manual extraction, however, the results obtained are still subject to operator bias.

We describe here a novel 3-D image segmentation technique capable of robust segmentation using K-mean clustering and knowledge-based morphological operations. This technique aims at solving the problems encountered in the segmentation of images consisting of regions that may be separate in anatomy but indistinguishable in intensity. The proposed adaptive K-mean clustering algorithm is adopted from [15] and is capable of initial segmentation of the structures characterized by spatially varying intensity distributions. Spatial constraints are incorporated in the form of Gibbs random fields in our adaptive clustering algorithm to enforce the neighborhood configuration to overcome the noise in the given image. Simple morphological operations are then applied to clear the results obtained from K-mean clustering to form the initial segmented regions. Although we are able to overcome the difficulties originating from spatially varying intensity distributions and image acquisition noise with adaptive K-mean clustering and simple morphological operations, resolving the anatomical ambiguity presented in many biomedical image segmentations is still a challenging task. It is this challenge that motivated the development of the algorithm for knowledge-based morphological operations which determines desired final segmentation according to the *a priori* anatomical knowledge of the region-of-interest. To illustrate the effectiveness of this proposed algorithm, we have successfully implemented a robust segmentation on a sequence of cardiac CT volumetric images to extract time-varying chamber of left ventricle. The volumes of left ventricle extracted using this approach compare favorably with the volumes obtained using operator manual outlining. However, such knowledge-based segmentation is fast, reproducible, and without operator bias.

The application of the proposed algorithm to other applications of image segmentation can be easily adopted as long as the *a priori* knowledge of the structure-of-interest is available. In many biomedical image segmentation tasks, such knowledge is usually available since we often study certain biomedical structures with known anatomical information. The anatomical information can be used in the design of K-mean clustering when it is necessary to set the value K and to incorporate the spatial characteristics of each class. Such information is crucial in the design of knowledge-based morphological operations since it is the only way of intelligently identifying the anatomical structures from the possibly ambiguous segmentations obtained through adaptive K-mean clustering. It is true that a particular implementation scheme of 3-D image segmentation would depend on individual applications. However, the principles of this knowledge-based approach will provide, without doubt, the methodology in the design of an individualized 3-D image segmentation algorithms.

Section II describes the given CT volumetric data used in this paper to implement a knowledge-based approach to image segmentation. It also interprets how the cardiac anatomy would appear in a given volumetric image. Section III presents the adaptive K-mean clustering algorithm and the implementation of such algorithm on the given 3-D image data. Section IV explains the necessity for initial post processing after K-mean clustering and introduces the morphological operations based approach to accomplish such processing. Section V addresses the difficulty in identifying the structures from the given images when they are separate in anatomy but indistinguishable in intensity. It then describes a knowledge-based approach to overcome such difficulty with the proposed constrained morphological operations. Section VI presents the segmentation results obtained using this proposed approach. Section VII concludes this paper with discussion on future extensions to this approach and potential integration of the image segmentation with subsequent image analysis tasks.

II. 3-D Image Data and Cardiac Anatomy

The 3-D image data used in this research are a sequence of CT volumetric images obtained from dynamic spatial reconstructor (DSR), a unique ultra-fast multislice scanning CT system built and managed by the Mayo Foundation [16]. According to Ritman and his colleagues [8], [10], the DSR synchronously scans multiple, parallel, transaxial sections within 1/100 s. These scans are repeated 60 times/s. Up to 120 images of transaxial sections were reconstructed for each 1/60-s scan sequence, and post scan manipulation and interpolation of the reconstructed images were used to generate cubic voxels of such volumetric image sequences. With such rapid, extensive data collection, high-resolution volumetric images, largely free of motion blur, can be generated for moving organs, such as the heart. In a typical DSR experiment, 16 volumes are reconstructed within a cardiac cycle, with each volume representing one time instant. Each reconstructed volumetric image usually contains roughly 120 slices of size 128 $\times$ 128, however, this research is based on a sequence of 16 volumes, each containing 95 90 $\times$ 90 slices. Each slice of a reconstructed volume represents an approximately 0.9 mm thick transverse cross section of the scanned anatomy, and each of the volumetric elements, or voxels, represents an (0.9 mm^3 cube) of tissue. To bring out the left ventricle chamber as a bright object, a Roentgen contrast agent is injected into the right atrium several seconds prior to the scanning of the heart. A few cross sections of the original DSR data from a canine heart scan are presented in Fig. 1.

In a typical volume of such images, the left ventricle is included in a high intensity region which would also includes the left atrial chamber and aorta. Although there are valves separating the left ventricle chamber from left atrial chamber and aorta, the valves of this canine heart, are at the order of 1 mm thick and their visibility in the volume is diminished by the partial volume effect and the resolution limitation of DSR scanner [17]. Furthermore, the valves open and close alternatively during a cardiac cycle so that whether or not they appear in an acquired image would also depend on the timing of the image acquisition. Therefore, the left ventricle chamber often appears connected with the left atrial chamber and aorta in the acquired volumetric images.

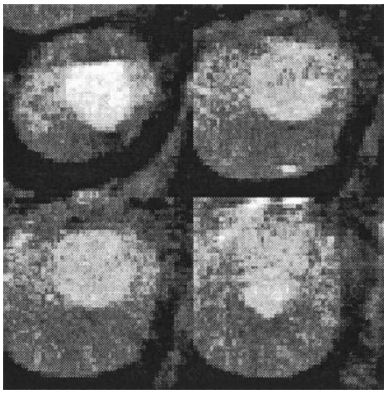


Fig. 1. A few typical cross sections of the original DSR data.

Overall, the intensity of the left ventricle is much brighter than the myocardium. However, within the volume, the intensity distribution of left ventricle chamber is not uniform due to the uneven distribution of the contrast agent. The nonuniform distribution of the contrast agent is difficult to model and compensate, and therefore requires the image segmentation algorithm to be adaptive to local properties of the intensity distribution. In addition to the nonuniform distribution of contrast agent, the noise from the errors in scanning and volume reconstruction causes the blurring of the structural borders and hence introduces the ambiguity in boundary classifications. According to Higgins [10], the left ventricle chamber appearing in the CT volumetric images is a large, bright, smooth, solid region, varying in size and shape over time, approximately attached to the left atrial chamber and aorta through the valves, and separated from the myocardium by a strong, but blurred and noisy, interface.

These are two types of ambiguity present in the given CT volumetric data: imaging related ambiguity and anatomy related ambiguity. Imaging related ambiguity can be resolved by designing an adaptive clustering algorithm while the anatomy related ambiguity cannot be resolved using image intensity information alone. It is necessary for us to utilize the *a priori* knowledge of the cardiac anatomy in order to identify the separations between the left ventricle and the left atrium and aorta. The simultaneous presence of imaging related ambiguity and anatomical ambiguity is the major challenge in this research and our successful solution to this challenging problem constitutes the main contributions of our work.

III. Segmentation Using Adaptive K-Mean Clustering

Traditional statistical image segmentation algorithms, as simple as thresholding or as complicated as K-mean [18] and even fuzzy K-mean clustering [19], all classify the pixels into clusters based only on their intensity values. Each cluster is usually characterized by a constant intensity and no spatial constraint is imposed. In practice, images are usually noise contaminated versions of the reflected density function, and the image intensity of the same class may change over space due to some physical constraints of the imaging system as we have discussed in Section II. In many biomedical applications, even though the relative intensity is evident for different clusters within a small neighborhood, different clusters at different locations may have similar intensity appearance due to the inhomogeneous nature of the imaging media. Therefore, a single global threshold is usually inapplicable to such images even within the same 2-D cross-section. The ability of being adaptive to the local intensity distribution is generally required for a robust image clustering algorithm to obtain the correct clustering results. In addition, certain spatial constraints are needed to prevent the algorithm from misclustering caused by the impulse noise introduced in the process of image acquisition and reconstruction. Such spatial constraint is based on the assumption that a pixel generally tends to belong to same cluster as most of its neighbors unless it is on the edge of a sharp region transition.

With the successful application of Markov random field in image segmentation [21]–[25], several extensions to the traditional K-mean clustering algorithm based on Gibbs random fields have recently been proposed [26]–[28]. These extensions have included the spatial constraints through the modeling of the spatial distribution of the clusters as Gibbs random fields. Such modeling of spatial distribution indeed imposes the spatial continuity in the process of clustering and produces more robust results than traditional K-mean algorithm. However, they all assume that the intensity or its related parameters would be constant within each clustered region. The adaptive K-mean algorithm we develop is based on the segmentation algorithm proposed recently by Pappas [15]. His algorithm includes not only the 2-D spatial constraints characterized by Gibbs random fields, but also the adaptive capability specified through iterative estimation of local means of each region. We have extended Pappas' algorithm through the development of 3-D spatial constraints to suit the volumetric nature of the image data and an enhanced adaptive capability to account for the varying characteristics of the cluster means as well as cluster variances.

K-mean clustering is often suitable for biomedical image segmentation since the number of clusters (K) is usually known for images of particular regions of human anatomy. In biomedical applications, the spatially varying intensity change of a biomedical structure is usually caused by inhomogeneity in the process of image acquisition, such as the inhomogeneous distribution of the contrast agent in CT imaging or inhomogeneous distribution of the magnetic field gradient in MR imaging. It has been shown that [29] such intensity change can be updated locally during the segmentation process through a maximum *a posteriori* (MAP) estimation scheme. In this research, the update of intensity change is represented by the estimated local cluster means and local cluster variances. If we denote a given image by y and a segmentation of this image by x, according to Bayes' rule, the *a posteriori* probability can be expressed as

$$p(x|y) \propto p(y|x)\, p(x) \tag{1}$$

where $p(x)$ is the *a priori* probability of the segmentation, and $p(y|x)$ represents the conditional probability of the image data given the segmentation. The Gibbs random field can be characterized by a neighborhood system and a potential function. A Gibbs random field constrained image segmentation is

accomplished by assigning labels to each pixel in the given image. A label $x_s = i$ implies that the pixel s belongs to the ith class of the K classes. Therefore, we have

$$p(x_s|x_t, \quad \forall t \neq s) = p(x_s|x_t, \quad t \in N_s) \tag{2}$$

where N_s represents the defined neighborhood for pixel s. Associated with each neighborhood system are cliques and their potentials. A clique C is a set of sites where all elements are neighbors. If we consider that a 2-D image is defined on the Cartesian grid and the neighborhood of a pixel is represented by its four nearest pixels [30], then the two-point clique potentials are defined as

$$V_C(x) = \begin{cases} \beta, & \text{if } x_s = x_t \text{ and } s, t \in C \\ -\beta, & \text{if } x_s \neq x_t \text{ and } s, t \in C. \end{cases} \tag{3}$$

For a 3-D image, a straightforward extension of 2-D neighborhood system concludes that the neighborhood of a voxel can be represented by its six nearest neighbors [29]. A Gibbs distribution can then be defined as

$$p(x) \propto \exp\left\{-\sum_C V_C(x)\right\} \tag{4}$$

where V_C is certain clique potential for clique C. For a $4 \times 4 \times 4$ 3-D lattice, there will be 24 cliques within each 4×4 cross section, and 16 cliques between two cross sections. The total number of two-point cliques for such 3-D lattice is therefore 144. If we model the conditional density as a Gaussian process with mean μ_s and variance σ_s at a pixel location s, then it can be written as a spatially varying density function with respect to pixel location s, as follows:

$$p(y|x) \propto \exp\left\{-\sum_s \frac{1}{2\sigma_s^2}(y_s - \mu_s)^2\right\}. \tag{5}$$

Then, the overall probability function will be

$$p(x|y) \propto \exp\left\{-\sum_s \frac{1}{2\sigma_s^2}(y_s - \mu_s)^2 - \sum_C V_C(x)\right\}. \tag{6}$$

There are two components in the overall probability function. One corresponds to the adaptive capability that force the segmentation to be consistent with local image distribution with locally estimated mean μ_s and variance σ_s. The other corresponds to the spatial continuity constraint characterized by the clique potentials within a given 3-D lattice.

MAP estimation based adaptive K-mean clustering can be implemented using various optimization techniques depending on the specific applications. In particular, when the sampling lattice of the given 3-D images are not uniformly structured, the clique potentials need to be carefully modified according to the relative length of each lattice unit. A nonuniformly structured sampling lattice is a common practice in medical imaging since resolution within a cross section is usually different from the resolution between cross sections. In the case of CT volumetric data used in this research, we assign the same β to the clique potentials both within a cross section and between cross sections, since the sampling lattice of the CT volumetric data is uniformly structured. The proposed adaptive clustering algorithm applied to the CT volumetric data is implemented using the method of *iterative conditional modes* [31]. First, an initial segmentation x is acquired through the simple K-mean algorithm. The value K is chosen to be four in the case of CT volumetric image data according to the available *a priori* knowledge of the cardiac structure. Then, overall probability function is maximized on a point-by-point basis, with the mean μ_s and the variance σ_s of each cluster s being updated after each iteration. Therefore, the optimization is accomplished through alternating between MAP estimation of the clustered regions and iterative update of the cluster means and variances. Such alternating processing is repeated until no pixels change classes. The result is the desired segmentation of the given 3-D images.

In the case of segmenting the left ventricle from CT images, the adaptive clustering algorithm is applied to obtain four clusters with the brightest cluster corresponding to potential left ventricle chamber. Even though only the brightest cluster is used for subsequent extraction of left ventricle chamber, the multicluster segmentation ($K = 4$) is still necessary. Because the given volumetric images indeed consist of four clusters of intensity corresponding to four types of biomedical structures, the binary clustering ($K = 2$), or thresholding, would produce incorrect segmentation results.

There exist a number of differences between this algorithm and that of Pappas [15]. One important difference is the introduction of iterative estimation of cluster variances in the process of optimization. The assumption of the changing variance and the implementation of estimation scheme allow us to account for the noise levels to change from one local area to another, and from one cluster to another. This additional feature of the proposed scheme enhances the flexibility of the adaptive K-mean clustering algorithm, since, in practice, the variances of different clusters are generally different and the variance of a specific cluster also changes with location. This varying nature in both intensity and variance is due to the inhomogeneous distribution of density in various medical images. The specific characteristics of CT images used in this research have already been discussed in Section II. The computational complexity of estimating variances is more complicated than the mean estimation as described in [15]. The second difference is the choice of the number of classes. The value of K used in our approach corresponds to the actual number of classes present in the given image. Therefore, the choice of K is based on the available knowledge, whereas the choice of K in [15] is not based upon any available knowledge. Instead, it is stated that the best choice $K = 4$ appears to be related to the famous four-color theorem. The third difference is the choice of the parameter β of the Gibbs random field. It is evident that the parameter β is related to the image contents as well as imaging conditions. According to biomedical structure and known imaging condition, we have chosen the parameter β such that the spatial constraint is strong enough to smooth out the noise while still preserving the structural details.

IV. Initial Morphological Operations

For many biomedical image analysis tasks, one usually focuses on a particular structure contained in a given image.

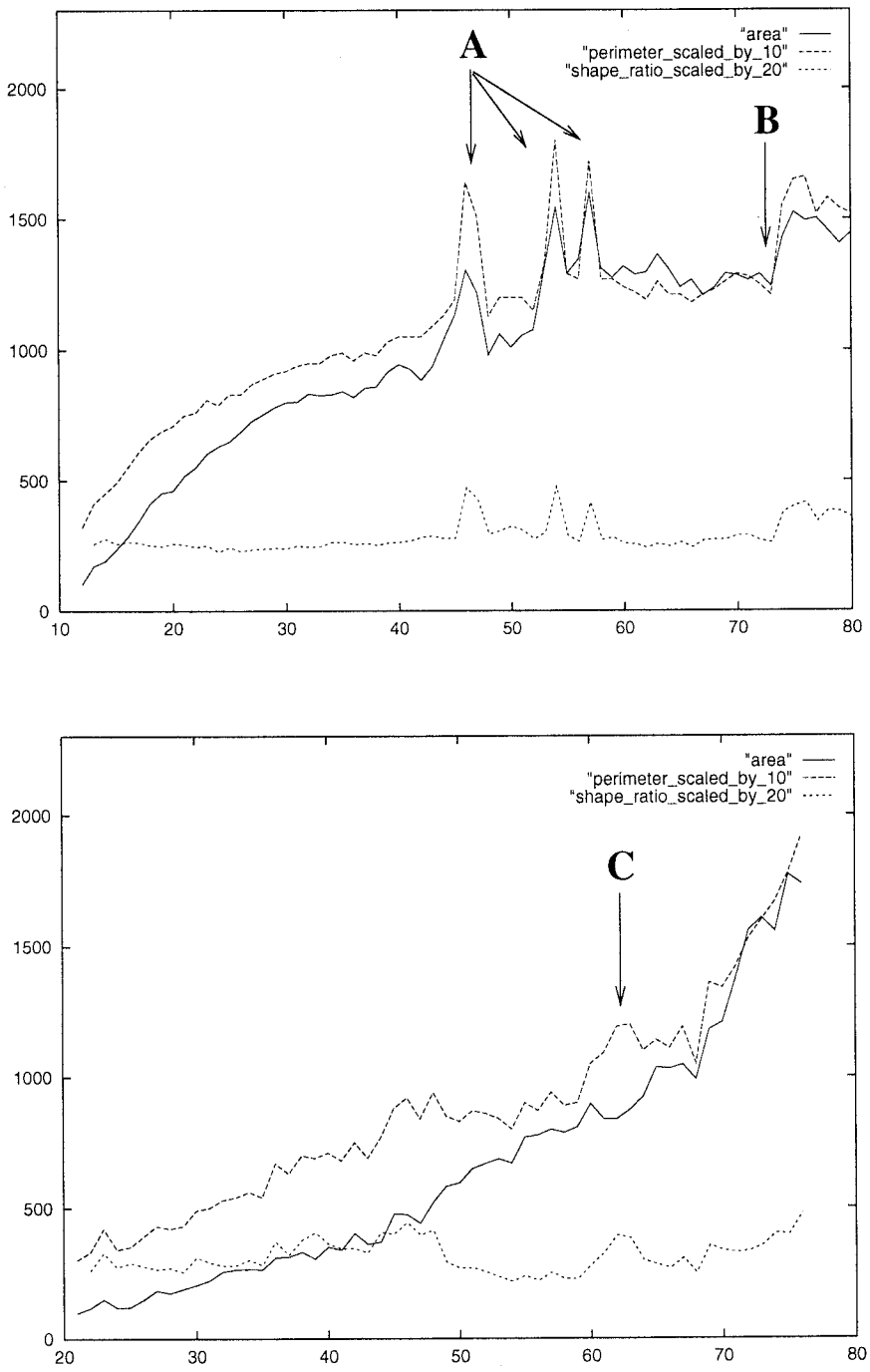


Fig. 2. Typical action curves.

In this research, the segmentation results are intended for the study of the dynamics of the left ventricle chamber. The subsequent analysis of motion and deformations is usually based on the left ventricle volumes extracted from the given CT volumetric data [32]. Ideally, upon K-mean clustering, the cluster corresponding to the brightest regions would represent the left ventricle chamber and can be used for left ventricle dynamics analysis. The representation can be obtained by converting the clustering result into a binary image in which each pixel is labeled as either belonging or not belonging to the desired region representing left ventricle chamber. However, the region so obtained may not be fully connected due to the noise and may contain left atrium and aorta due to the anatomical ambiguities. Two types of morphological operations are applied to the binary images in order to clean the noise and to clarify the anatomical ambiguities. They are initial processing morphological operations described in this section and knowledge-based morphological operations described in Section V.

Initial morphological operations are necessary since the spatial smoothness can be implemented through such operations. In addition, the morphological operations can also be used to correct certain large misclustered regions that are inevitably present in segmentation results when using

the adaptive K-mean algorithm since it is still an intensity-based approach. The spatial constraints used in the MAP estimation have indeed largely overcome the difficulties in segmenting the noise contaminated images. However, certain large misclustered regions are the consequence of trade-off between preserving the structural detail and suppressing the contaminating noise. In general, those misclustered regions due to intensity ambiguity can be removed using simple morphological operations which can also be used to smooth the segmented volume to obtain a better representation of the left ventricle boundary.

Such misclustered regions due to intensity ambiguity often appear in a cross section as small bright regions disconnected from main bright region, or small cavities inside the main bright region. Binary morphological operations have been proved capable of deleting small disconnected regions, filling cavities and smoothing the region-of-interest [33] and [34]. Digital morphological operations, opening followed by closing, are applied to the binary image to eliminate the small isolated regions outside the potential left ventricle volume and small cavities inside such volume. These operations are defined as an ordered combinations of fundamental operations, dilation and erosion, which can be written as

$$O(\mathbf{A}, \mathbf{B}) = D[E(\mathbf{A}, \mathbf{B}), \quad \mathbf{B}] \tag{7}$$

$$C(\mathbf{A}, \mathbf{B}) = E[O(\mathbf{A}, -\mathbf{B}), -\mathbf{B}] \tag{8}$$

where O and C denote the operators opening and closing, respectively, and D and E denote the operators, erosion, and dilation, respectively. If the translation operation is defined as

$$\mathbf{A} + x = \{a + x: \quad a \in \mathbf{A}\} \tag{9}$$

then the operators erosion and dilation can be written as

$$E(\mathbf{A}, \mathbf{B}) = \mathbf{A} \ominus \mathbf{B} = \{x: \quad -\mathbf{B} + x \subset \mathbf{A}\} \tag{10}$$

$$D(\mathbf{A}, \mathbf{B}) = \mathbf{A} \oplus \mathbf{B} = \{x: \quad (-\mathbf{B} + x) \cap \mathbf{A} = \emptyset\} \tag{11}$$

where $\mathbf{A}$ is the image and $\mathbf{B}$ is the structuring element. Note that $-\mathbf{B}$ denotes the reflection of $\mathbf{B}$ with regard to the origin, x denotes a point in space, and a is a point in the image $\mathbf{A}$. Here we use circular element with radius r which can be determined by a compromise between the noise suppression performance and preservation of details. The larger the element is, the more noise is suppressed and the more details are lost. Fortunately, in many applications, we know in advance the rough size and shape of the object we are looking for in the segmented images. Such knowledge can be used in this initial postsegmentation processing in which we try to eliminate small regions that are considered not part of the left ventricle, and cavities inside the left ventricle chamber region.

The opening and closing operations are inherently idempotent, that is, repeatedly applying the same combination of opening and closing will not change the image after the first application. Thus we choose a larger element in the closing operation in order to compensate some possible detriment that a certain opening operation may cause, such as opened hollow areas at or inside the boundary. Such intuitive analysis has also been confirmed by test images from cardiac CT sequences. This asymmetric combination of opening and closing can be carefully balanced without significantly altering the shape since the left ventricle chamber is fairly large and smooth. The largest region in each cross section is chosen as the candidate region belonging to the left ventricle chamber. The stack of all regions forms the volumetric representation of extracted left ventricle chamber presumably free of ambiguities caused by inhomogeneous and noise corrupted intensity distributions. However, the anatomical ambiguity caused by the lack of clear boundary between left ventricle and left atrium and aorta would still be present after the initial morphological operations. The elimination of such ambiguity needs to utilize the knowledge of the left ventricle anatomy and the implementation of such knowledge-based algorithm is described in the next Section.

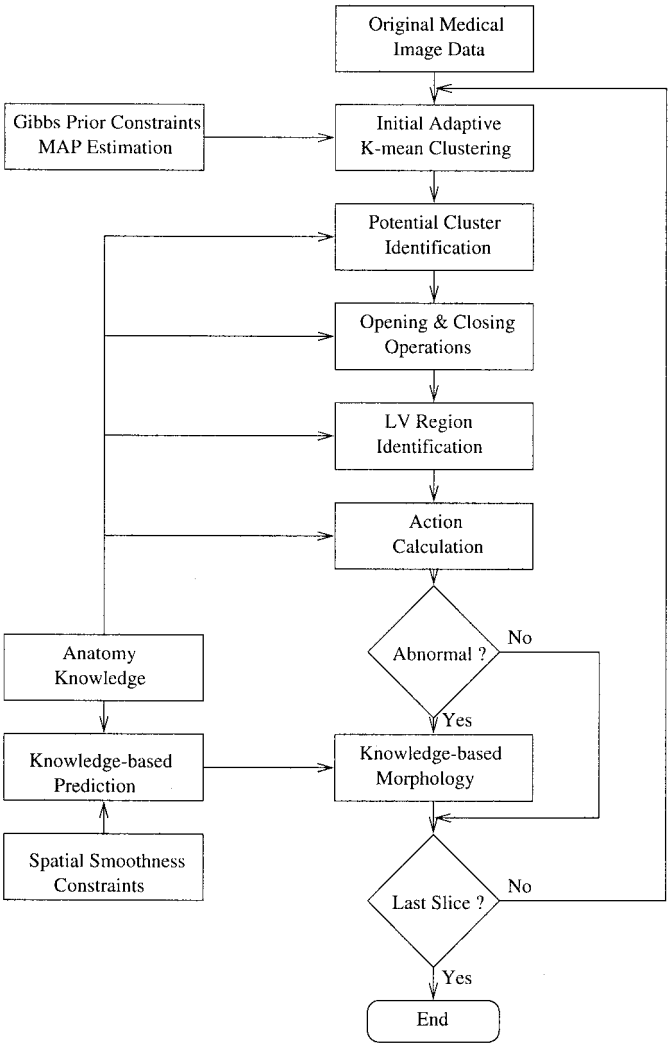


Fig. 3. Block diagram of the overall 3-D segmentation scheme.

V. Knowledge-Based Morphological Operations

The necessity of further identification of the region-of-interest lies in the fact that those two problems we have pointed out are still unsolved. First of all, due to the anatomy related ambiguity, part of the left atrium and aorta are classified connected with the upper part of the left ventricle. As we have pointed out earlier, such ambiguity can not be differentiated

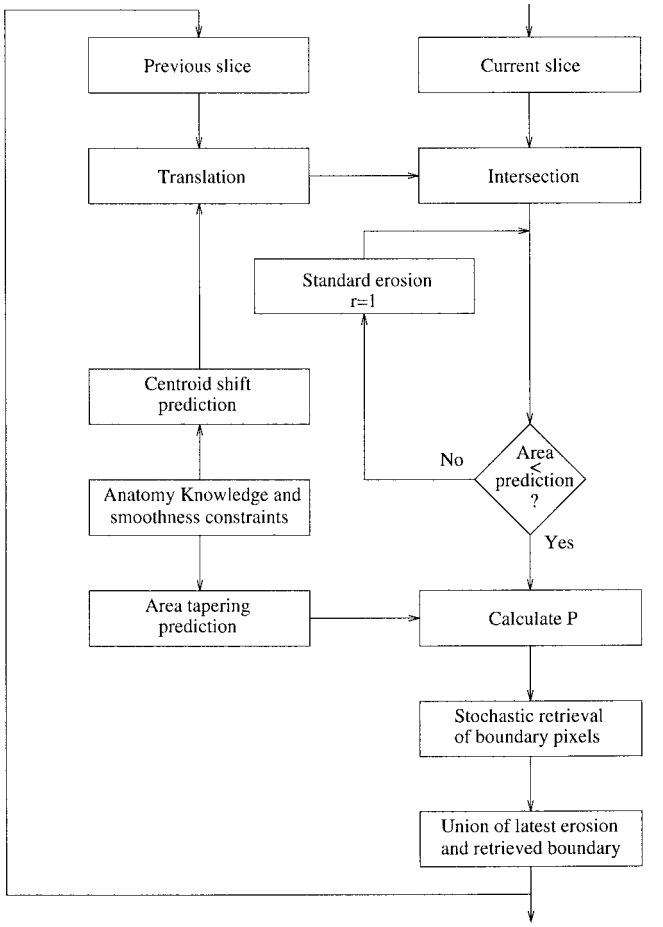


Fig. 4. Block diagram of the composite knowledge-based morphological operations.

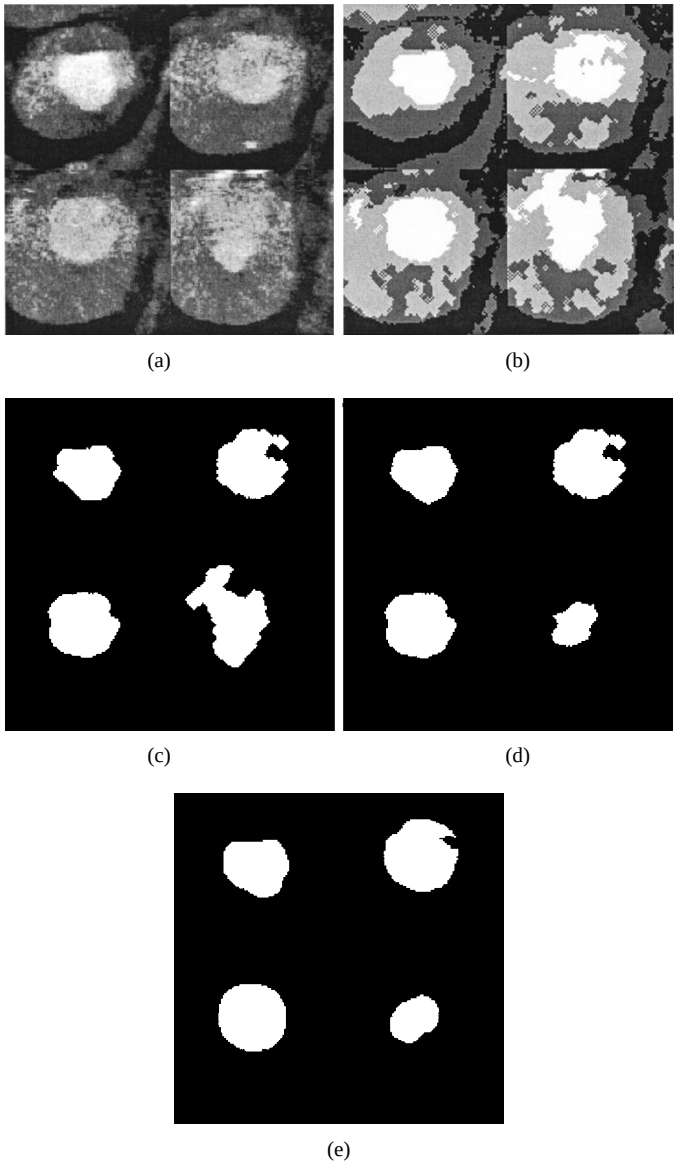


Fig. 5. Comparison of the segmentation results. (a) Original CT image. (b) K-mean segmentation. (c) Initial ROI. (d) Final ROI. (e) Manual segmentation.

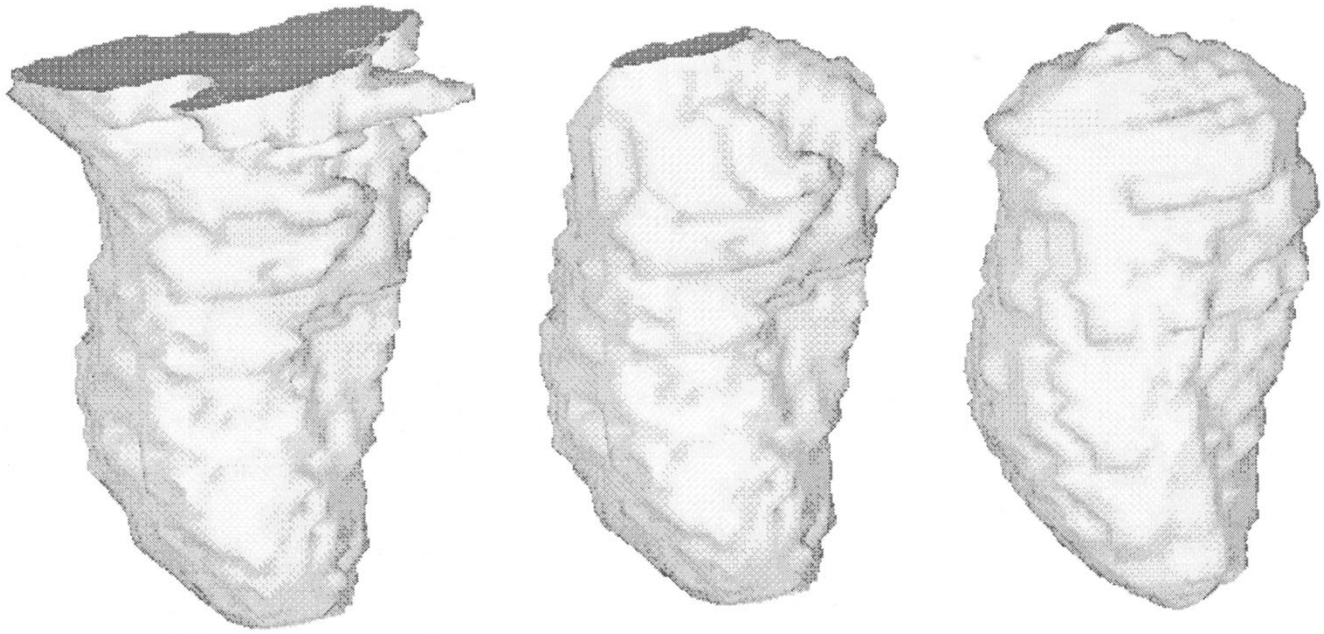
solely by the intensity. Furthermore, sometimes the local noise may become so remarkable that large nonuniformly can occur in the object or background region. The consequence may be a large area attached to or hollow space inside the left ventricle chamber even after the initial morphological operations. Existing schemes dealing with such problem all introduced human interferences one way or another, either outlining the valves by a human operator, or applying additional constraints interactively to the location where the ambiguity is present [11] and [14].

In order to reduce potential errors and eliminate operator biases associated with human interferences, a knowledge-based intelligent approach is adopted. This is motivated by the success of shape modeling the left ventricle by superquadrics with tapering and bending deformations in the CT volumetric data based shape and motion analysis of the left ventricle [32]. According to the anatomy of the left ventricle, the shape of its cross sections is of similar shape within their neighborhood while the area of these cross sections decreases from the middle to the two ends. In addition, the curved nature of its long axis results in a continuous shift of the centroid from one cross-section to another. Therefore, the principle of smooth transition in cross-section shape can be applied to decide whether a discontinuity due to misclassification has taken place. To avoid an over constrained operation, we make no assumption of a specific parametric model of the global shape. Instead, we predict, in first or second order, the tapering in area and the shift of centroid within a small neighborhood of a given cross section and use such predictions to detect possible misclassifications due to image or anatomy related ambiguities.

For many correctly classified cross sections, no action is taken in the knowledge-based morphological operation illustrated in Fig. 4. Whether or not the knowledge-based morphological operation is applied to a particular slice is dependent on an action criterion defined, for current slice K, as

$$ACTION_k = \frac{\triangle_k \text{area}}{T1} + \frac{\triangle_k \text{peri}}{T2} + \frac{\triangle_k \text{ratio}}{T3} \qquad (12)$$

where $\triangle_k$ denotes the difference operator and T_i's are preset thresholds. The action defined in (12) has taken into consideration the variations in area, perimeter as well as shape ratio (perimeter^2/area). Two cases would trigger the start of



(a) (b) (c)

Fig. 6. Comparison of the segmentation in 3-D volumes. (a) Segmentation result without knowledge-based morphology. (b) Final segmentation result using the proposed approach. (c) Result of manual segmentation.

knowledge-based morphological operations. Fig. 2 shows two typical action curves in which the horizontal axis denotes the slices and the vertical axis denotes the number of voxels. The shape ratio should be in the range of 4–6, and should change relatively smoothly. In the case that the value of the action criterion is larger than 3.0, which means either all the three measures are abnormal or some are extraordinarily abnormal, the knowledge-based morphological operation is activated to remove the extra areas attached to the desired region (e.g., type A in Fig. 2). In another case both perimeter and the shape-ratio change abruptly, even though the area change is within tolerated variation. This case corresponds to one or more hollow areas inside the region-of-interest, since the perimeter is computed by counting all boundary pixels, including those inside the region-of-interest. The appropriate closing operation with an element of larger radius is applied to fill in the hollow areas inside the desired region that were not filled using initial morphological operation described in Section IV (e.g., type B in Fig. 2). To ensure a robust processing, this procedure starts from the middle of the volume where the segmentation and initial morphological operation have produced good result, and proceeds in opposite directions to the two ends. Experiments have shown that our algorithm is insensitive to the choice of the starting slice as long as it is chosen to be in the middle of the whole volume.

When the knowledge-based morphological operation is activated, the following steps are undertaken as illustrated in Fig. 4, the block diagram of the knowledge-based morphological operations. First, a template is generated using the predicted tapering parameter based on several of its previous slices. This template is translated using the predicted shift of centroid and compared with the given cross section. Depending on the comparison result, dilation or erosion is applied to obtain a region whose area is sufficiently close to predicted template. Notice that the standard operations defined in (7), (9), and (10) are implemented using a structure element with radius of one. This discretized operation would result in a change of cross-section area that equals to the total number of pixels on the boundary at each iteration. It is evident that such standard operations are not applicable to the cases when the desired increase/decrease of the area equals to only a fraction of total increase/decrease caused by a standard operation.

To meet the necessity of increasing or decreasing the area at a rate equal to a fraction of area caused by standard operation, a nonstandard operation with stochastic property is introduced. Such a nonstandard operation is equivalent to a standard operation with a structure element of subpixel radius. A stochastic erosion $\hat{E}[A, B, p_0]$ is defined as a conditional erosion in that whether a boundary pixel is removed or not is determined by a random probability generator with uniform distribution

$$P[A, p_0] = \{x_i \in A | p_i > p_0,\ p_i \in UNIF[0, 1],\ \forall\, x_i \in A\}. \quad (13)$$

The implementation of such stochastic erosion can be written as

$$\hat{E}[A, B, p_0] = E[A, B] \bigcup P[(A - E[A, B]), p_0] \quad (14)$$

where $UNIF[0, 1]$ denotes the uniform distribution in $[0, 1]$, and $P[A, p_0]$ is a random selecting process. The parameter p_0 is determined by the ratio of the number of points needed to be eliminated and the total number of points of the current boundary. Note that this stochastic erosion is a logical way of implementing conditional erosion to determine whether a boundary pixel should be eliminated. A conditional dilation can be defined with similar operations. The nature of these subpixel operations would result in a more accurate processing

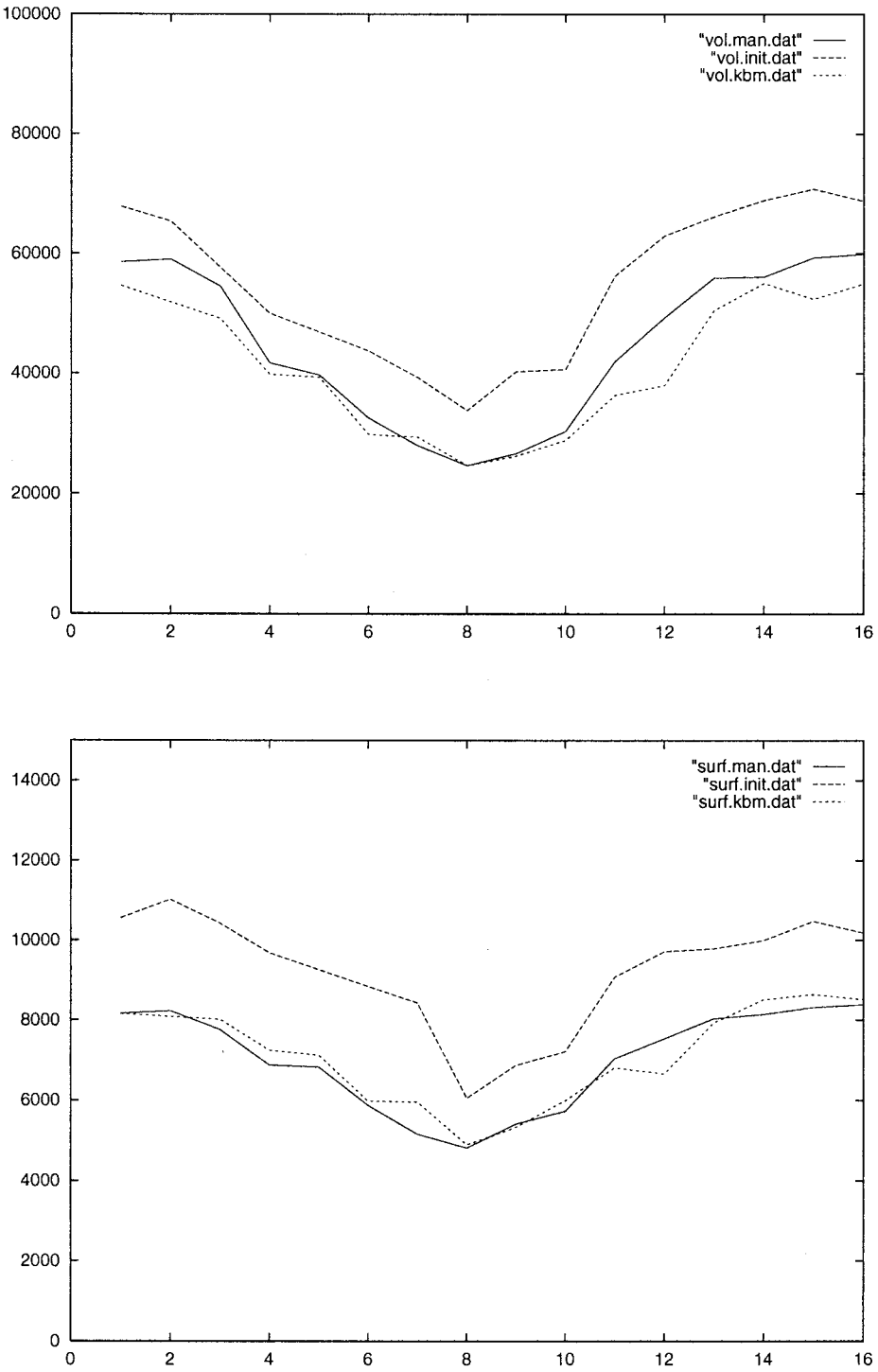


Fig. 7. Comparison of the volume (upper) and surface area (lower) over the cardiac cycle obtained using manual outlining, and segmentations with and without the knowledge-based morphological operation.

of the boundary by a random process driven dilation and erosion operations. With its stochastic characteristics, the subpixel operations are also able to produce visually pleasing boundaries, especially in the case of resolving anatomy related ambiguity (e.g., point C in Fig. 2).

VI. Segmentation Results

The final result of this knowledge-based approach is a segmented left ventricle chamber whose image related ambiguities (corresponding to point A, B in action curve) as well as anatomy related ambiguities (corresponding to point C in action curve) are successfully resolved. Results of a few typical 2-D cross sections are shown in Fig. 5. Based upon the extracted cross-sections, we can reconstruct the whole left ventricle volume. 3-D rendering of the left ventricle volume for visualization and animation of cardiac dynamics is accomplished using AVS, a software of advanced visual systems. 3-D reconstructed volumes are shown in Fig. 6. Our results

compare favorably with the results obtained using manual outlining. Quantitative comparisons are also shown in Fig. 7 in which the horizontal axis denotes the time frame and the vertical axis denotes the number of voxels. This knowledge-based approach yields results very close to those of the manual segmentation with an average percentage difference of 6.7% and 1.9%, in volume and surface area respectively. Notice that, according to [9], the intraobserver volume variability of manual segmentation is on the order of $\pm 5\%$. On the contrary, both the volume and surface area obtained without applying the knowledge-based morphological operation are generally much larger than the manual results due to the inclusion of left atrium and aorta (i.e., unresolved anatomical ambiguities), the average percentage difference is 27.1% for volume and 31.9% for surface area. Reproducibility of the knowledge-based approach is very good in that it is insensitive to the choice of the starting cross section as long as the starting slice is in the middle part of the volume. The segmentation results produced by the proposed approach compare favorably with those presented in [10]–[12]. In [12], neither volumes nor surface areas were compared with manual segmentation. However, from the display results, large errors are expected since the appearance of the extracted left ventricle chambers is quite different from the manual results. In [10], the anatomy-related ambiguity is resolved using an interactive manual delineation. Portions of some 2-D slices which contain the left atrium and aorta are deleted and the resultant empty regions are propagated through the volume so that no processing is applied to the ambiguous regions. The rest of the left ventricle boundary is obtained through adaptive thresholding. With proper starting slice, the extracted left ventricle chamber can be obtained. Only volumes are compared with the results of full manual segmentation and the results are presented in [10] as a plot of volume versus time frames. Since no numerical average difference has been provided, an estimate from the plot indicates that the average difference is between five to ten percent, which is compatible with the performance of this approach. In [11], a method of image segmentation applicable to general circumstances is described and the results of extracting left ventricle chambers are presented. The adaptive thresholding scheme in [10] has now been replaced by interactive relaxation labeling, however, the anatomy-related ambiguity is again resolved using manual delineation. The performance of the scheme presented in [11] is about the same as those presented in [10]. The major difference between the approach presented in this paper and the approaches presented in [10] and [11] is the way the anatomy-related ambiguity is resolved.

VII. Conclusion and Discussion

We have presented a knowledge-based approach for 3-D image segmentation. This approach is able to resolve the challenging issues corresponding to both image related and anatomy related ambiguities in the segmentation of 3-D medical images. The application in CT volumetric image segmentation of the left ventricle has shown that the knowledge-based approach can be very successful, especially in medical image analysis where *a priori* knowledge of the biomedical objects is available. The generalization of this approach to other similar application can be made readily. This segmentation scheme can be incorporated into an integrated approach in which image segmentation and shape analysis are mutually constrained to produce a consistent estimation of cardiac shape as well as cardiac dynamics [32].

Two extensions of this approach are currently under investigation. One extension addresses a scheme of true 3-D morphological operations. Such 3-D operation may be able to produce better result in the process of initial morphological operation so that image related ambiguity may be resolved completely. This will allow the easier implementation of the knowledge-based operations. The other extension involves the incorporation of temporal information into the segmentation procedure. The temporal changes of the shape of the biomedical object will undoubtedly provide useful information which can be used to resolve the ambiguity at a particular frame of given image data. The temporal information can also be incorporated into a integrated approach where both segmentation and subsequent analysis are studied.

References

[1] B. Widrow, "The 'rubber mask' technique—I and II," *Pattern Recognit.*, vol. 5, pp. 175–211, 1973.

[2] A. L. Yullie, D. S. Cohen, and P. W. Hallinan, "Feature extraction from faces using deformable templates," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, June 1989, pp. 104–109.

[3] U. Grenander, Y. Chow, and D. M. Keenan, *Hands: A Pattern Theoretic Study of Biological Shapes.* Berlin, Germany: Springer-Verlag, 1991.

[4] L. H. Staib and J. S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 1061–1075, 1992.

[5] C. W. K. Gritton and E. A. Parrish, Jr., "Boundary location from an initial plan: The bead chain algorithm," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 5, pp. 8–13, 1983.

[6] D. B. Cooper, "Maximum likelihood estimation of Markov-process blob boundaries in noisy images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-4, pp. 372–384, 1979.

[7] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, pp. 321–331, 1988.

[8] E. A. Hoffman and E. L. Ritman, "Shape and dimensions of cardiac chambers: Importance of CT section thickness and orientation," *Radiology*, vol. 115, pp. 739–744, 1985.

[9] T. Iwasaki *et al.*, "Mass of left ventricular myocardium estimated with dynamic spatial reconstructor," *Amer. J. Physiol.*, vol. 246, pp. H138–H142, 1984.

[10] W. E. Higgins, N. Chung, and E. L. Ritman, "Extraction of left ventricle chamber from 3D CT images of the heart," *IEEE Trans. Med. Imag.*, vol. 9, pp. 384–395, 1990.

[11] W. E. Higgins, M. W. Hansen, and W. L. Sharp, "Interactive relaxation labeling for 3D cardiac image analysis," in *Biomedical Image Processing and Biomedical Visualization*, R. S. Acharya and D. B. Goldgof, Eds. San Jose, CA: SPIE, Feb. 1993, vol. 1905.

[12] H. K. Tu and D. B. Goldgof, "Left ventricle boundary detection from spatio-temporal volumetric CT images," in *Biomedical Image Processing and Biomedical Visualization*, R. S. Acharya and D. B. Goldgof, Eds. San Jose, CA: SPIE, Feb. 1993, vol. 1905.

[13] A. Singh, L. von Kurowski, and M. Y. Chiu, "Cardiac MR image segmentation using deformable models," in *Biomedical Image Processing and Biomedical Visualization*, R. S. Acharya and D. B. Goldgof, Eds. San Jose, CA: SPIE, Feb. 1993, vol. 1905.

[14] T. McInerney and D. Terzopoulos, "A finite element based deformable model for 3D biomedical image segmentation," in *Biomedical Image Processing and Biomedical Visualization*, R. S. Acharya and D. B. Goldgof, Eds. San Jose, CA: SPIE, Feb. 1993, vol. 1905.

[15] T. Pappas, "An adaptive clustering algorithm for image segmentation," in *IEEE Trans. Signal Processing*, 1992, vol. 40, pp. 901–914.

[16] E. L. Ritman, "Fast computed tomography for quantitative cardiac analysis—state of the art and future perspectives," in *Proc. Mayo Clinic*, vol. 65, pp. 1336–1349, 1990.

[17] R. A. Robb, Ed., *Three-Dimensional Biomedical Imaging.* Boca Raton, FL: CRC Press, 1985.
[18] J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles.* Reading, MA: Addison-Wesley, 1974.
[19] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms.* New York: Plenum, 1982.
[20] R. Chellappa, "Two-dimensional discrete Gaussian Markov random field models for image processing," in *Progress in Pattern Recognition—2,* K. Lanal and A. Rosenfield, Eds. Amsterdam, The Netherlands: Elsevier, 1985, pp. 79–112.
[21] J. Zhang and J. W. Modestino, "A model-fitting approach to cluster validation with application to stochastic model-based image segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 1009–1017, 1990.
[22] F. C. Jeng and J. W. Woods, "Compound Gauss-Markov random fields for image estimation," *IEEE Trans. Signal Processing,* vol. 39, pp. 683–697, 1991.
[23] R. Chellappa and A. K. Jain, *Markov Random Fields—Theory and Applications.* New York: Academic, 1993.
[24] J. Zhang, J. W. Modestino, and D. A. Langan, "Maximum-likelihood parameter estimation for unsupervised stochastic model-based image segmentation," *IEEE Trans. Image Processing,* vol. 4, pp. 404–420, 1994.
[25] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 721–741, 1984.
[26] H. Derin and H. Elliot, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 39–55, 1987.
[27] S. Lakshmanan and H. Derin, "Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 799–813, 1989.
[28] M. M. Chang, A. M. Tekalp, and M. I. Sezan, "Bayesian segmentation of MR images using 3D Gibbsian priors," in *Image and Video Processing,* M. Rabbani, M. I. Sezan, and A. M. Tekalp, Eds. San Jose, CA: SPIE, Feb. 1993, vol. 1903.
[29] J. Besag, "Spatial interaction and the statistical analysis of lattice system," *J. R. Stat. Soc.*, vol. 36, pp. 192–326, 1974.
[30] ———, "On the statistical analysis of dirty pictures," *J. R. Stat. Soc.*, vol. 48, pp. 259–302, 1986.
[31] C. W. Chen, J. Luo, K. J. Parker, and T. S. Huang, "CT volumetric data based left ventricle motion estimation: an integrated approach," *J. Comput. Med. Imag. Graph.*, vol. 19, pp. 85–100, 1995.
[32] C. R. Giardina and E. R. Dougherty, *Morphological Methods in Image and Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1988.
[33] J. Serra, *Image Analysis and Mathematical Morphology—Volume 2: Theoretical Advances.* New York: Academic, 1988.

Chang Wen Chen (S'86–M'92–SM'97) received the B.S. degree in electrical engineering from University of Science and Technology of China in 1983, the M.S.E.E. degree from University of Southern California, Los Angeles, in 1986, and Ph.D. degree in electrical engineering from University of Illinois at Urbana-Champaign in 1992.

In the summers of 1989 and 1990, he was employed at National Center for Supercomputing Applications, Champaign, IL, working with the Visualization Service and Development Group. From August 1992 to September 1996, he was on the faculty of the Department of Electrical Engineering, University of Rochester, Rochester, NY. In 1995, he spent his Summer at Kodak working with the Imaging Science Research Laboratories. Since September 1996, he has been an Assistant Professor of Electrical Engineering at the University of Missouri-Columbia. His current research interests include image and video coding, wireless communication, wavelets, biomedical image processing, graphics, and visualization.

Dr. Chen currently serves as an Associate Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. He is a senior author of the 1994 SPIE Best Student Paper for Visual Communication and Image Processing. He received a Whitaker Foundation Biomedical Engineering Research Award in 1996. He is a member of Tau Beta Pi and SPIE.

Jiebo Luo (S'92–M'96) was born in Kunming, China, in 1967. He received the B.S. and M.S. degrees in electrical engineering from the University of Science and Technology of China, Hefei, China, in 1989 and 1992, respectively. In 1995, he received the Ph.D. degree in electrical engineering from the University of Rochester, Rochester, NY.

He was a Research Assistant in the Image Processing Laboratory and the NSF Center for Electronic Imaging Systems, University of Rochester, from September 1992 to November 1995. In the summer of 1995, he was employed at the Joseph C. Wilson Center for Technology of Xerox Corporation, Webster, NY. In 1995, he became a Senior Research Scientist, and is currently a Research Associate in the Imaging Science Technology Laboratory, Imaging Research and Advanced Development, Eastman Kodak Company, Rochester, NY. His research interests include image enhancement and manipulation, digital photography, image and video coding, wavelets and applications, medical imaging, pattern recognition, and computer vision. He has published over 30 technical papers.

Dr. Luo holds a number of U.S. and European patents. He was the recipient of the 1994 SPIE Best Student Paper Award for Visual Communication and Image Processing. He is a member of the SPIE.

Kevin J. Parker (S'79–M'81–SM'87–F'95) received the B.S. degree in engineering science (summa cum laude), from the State University of New York, Buffalo, in 1976, and the M.S. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology in 1978 and 1981, respectively.

From 1981 to 1985, he was an Assistant Professor of electrical engineering and radiology. His research interests are in medical imaging, linear and nonlinear acoustics, and digital halftoning.

Dr. Parker has received awards from the National Institute of General Medical Sciences (1979), the Lilly Teaching Endowment (1982), the IBM Supercomputing Competition (1989), the World Federation of Ultrasound in Medicine and Biology (1991). He is a member of the IEEE Sonics and Ultrasonics Symposium Technical Committee, and serves as reviewer and consultant for a number of journals and institutions. He is a member of the Acoustical Society of America and the American Institute of Ultrasound in Medicine. He was named a Fellow in both the IEEE and the AIUM for his work in medical imaging. In addition, he was recently named to the Board of Governors of the AIUM.