

# Automatic anatomy recognition via multiobject oriented active shape models

Xinjian Chen

*Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Room 1C515, Building 10, Bethesda, Maryland 20892-1182*

Jayaram K. Udupa<sup>a)</sup>

*Medical Imaging Processing Group, Department of Radiology, UPENN, Philadelphia, Pennsylvania 19104*

Abass Alavi and Drew A. Torigian

*Hospital of the UPENN, Department of Radiology, UPENN, Philadelphia, Pennsylvania 19104*

(Received 18 June 2010; revised 15 October 2010; accepted for publication 25 October 2010; published 23 November 2010)

**Purpose:** This paper studies the feasibility of developing an automatic anatomy recognition (AAR) system in clinical radiology and demonstrates its operation on clinical 2D images.

**Methods:** The anatomy recognition method described here consists of two main components: (a) multiobject generalization of OASM and (b) object recognition strategies. The OASM algorithm is generalized to multiple objects by including a model for each object and assigning a cost structure specific to each object in the spirit of live wire. The delineation of multiobject boundaries is done in MOASM via a three level dynamic programming algorithm, wherein the first level is at pixel level which aims to find optimal oriented boundary segments between successive landmarks, the second level is at landmark level which aims to find optimal location for the landmarks, and the third level is at the object level which aims to find optimal arrangement of object boundaries over all objects. The object recognition strategy attempts to find that pose vector (consisting of translation, rotation, and scale component) for the multiobject model that yields the smallest total boundary cost for all objects. The delineation and recognition accuracies were evaluated separately utilizing routine clinical chest CT, abdominal CT, and foot MRI data sets. The delineation accuracy was evaluated in terms of true and false positive volume fractions (TPVF and FPVF). The recognition accuracy was assessed (1) in terms of the size of the space of the pose vectors for the model assembly that yielded high delineation accuracy, (2) as a function of the number of objects and objects' distribution and size in the model, (3) in terms of the interdependence between delineation and recognition, and (4) in terms of the closeness of the optimum recognition result to the global optimum.

**Results:** When multiple objects are included in the model, the delineation accuracy in terms of TPVF can be improved to 97%–98% with a low FPVF of 0.1%–0.2%. Typically, a recognition accuracy of  $\geq 90\%$  yielded a TPVF  $\geq 95\%$  and FPVF  $\leq 0.5\%$ . Over the three data sets and over all tested objects, in 97% of the cases, the optimal solutions found by the proposed method constituted the true global optimum.

**Conclusions:** The experimental results showed the feasibility and efficacy of the proposed automatic anatomy recognition system. Increasing the number of objects in the model can significantly improve both recognition and delineation accuracy. More spread out arrangement of objects in the model can lead to improved recognition and delineation accuracy. Including larger objects in the model also improved recognition and delineation. The proposed method almost always finds globally optimum solutions. © 2010 American Association of Physicists in Medicine.

[DOI: [10.1118/1.3515751](https://doi.org/10.1118/1.3515751)]

Key words: object recognition, image segmentation, active shape models, live wire

## I. INTRODUCTION

With medical imaging becoming increasingly functional and quantitative, clinical radiology will likely lay increasingly higher emphasis on quantification in routine practice. To facilitate quantitative radiology, computerized recognition, labeling, and delineation of anatomic organs, tissue regions, and suborgans, and guided by these, the delineation and

quantification of abnormalities will become important in clinical radiology. Once organs have been recognized and delineated, an important component of this task will also be to automatically report certain fundamental morphological, physiological, architectural, and functional information, pertaining to the organs and tissues, derived from single/multiple modality images. For the purpose of this paper, such an assistive process of recognizing, delineating, and quanti-

fying organs and tissue regions, occurring automatically during clinical image interpretation in the Radiology reading room, will be called *automatic anatomy recognition* (AAR). As a first step toward this larger goal, the purpose of this paper is to demonstrate in 2D images the feasibility of developing such a system in clinical radiology.

AAR is the process of identifying and delineating objects in medical images, in other words, the complete process of image segmentation. The whole segmentation operation can be thought of as consisting of two related processes: Recognition and delineation. *Recognition* is the high-level process of determining roughly the whereabouts of an object of interest and distinguishing it from other objects in the image. *Delineation* is the low-level process of determining the precise spatial extent of the object in the image. The efficient incorporation of high-level recognition help together with accurate low-level delineation has remained a challenge in image segmentation.

Historically, purely image-based segmentation methods, such as those employing thresholding, region growing, clustering, active contours, level sets, live wire, watershed, fuzzy connectedness, graph cut, and Markov random fields predate most methods employing object population shape and appearance prior models such as atlases, statistical shape and appearance models, and fuzzy geographic models. The relative merits of and the synergy that exists between these two approaches—purely image-based and model-based strategies—is clearly emerging in the segmentation field. As such hybrid methods that form a combination of the two approaches are emerging as powerful segmentation tools<sup>1–10</sup> where their superior performance and robustness over each of the component methods has been well demonstrated. Many model-based<sup>11–18</sup> as well as some purely image-based techniques,<sup>19–26</sup> specially tailored for specific body regions and image modalities, have been developed. However, in the context of AAR, it is better to have a general approach that is applicable to any (or most) body regions, image modalities, and protocols, and that do not depend heavily on the characteristics of fixed shape families and image modalities. While perhaps some of the above techniques can be generalized in this spirit of AAR, the same methods demonstrated to work in this general setting are few.

In this paper, we propose a general automatic anatomy recognition method which can be applied to different body regions and different image protocols. We propose an AAR approach by extending the 2D OASM method<sup>6</sup> to multiobject OASM (MOASM) and demonstrate the generality of the method on three different body regions and two image protocols. The AAR methodology described here consists of two main novel components: (a) multiobject generalization of OASM and (b) object recognition strategies. The OASM algorithm is generalized to multiple objects by including a model for each object and assigning a cost structure specific to each object in the spirit of live wire.<sup>27</sup> The delineation of multiobject boundaries is done in MOASM via a three level dynamic programming algorithm, wherein the first level is at pixel level which aims to find optimal oriented boundary segments between successive landmarks, the second level is

at landmark level which aims to find optimal location for the landmarks, and the third level is at the object level which aims to find optimal arrangement of object boundaries over all objects. The object recognition strategy attempts to find that pose vector (consisting of translation, rotation, and scale component) for the multiobject model that yields the smallest total boundary cost over all objects. The proposed method was evaluated on 2D images drawn from three routine clinical data sets from three different body regions: Chest CT, abdominal CT, and foot MRI data sets. An overall delineation accuracy of true positive volume fraction<sup>30</sup> (TPVF)  $>97\%$ , false positive volume fraction<sup>30</sup> (FPVF)  $<0.2\%$  was achieved, suggesting the feasibility of AAR and high accuracy. The results indicate that increasing the number of objects in the model improves both recognition and delineation accuracy in clinical images. More spread out arrangement of objects in the model can also lead to improved recognition and delineation accuracy. Including larger objects in the model also improved recognition and delineation.

This paper is organized as follows. In Sec. II, the complete methodology of our approach and the object recognition strategies are described. In Sect. III, we describe a detailed evaluation of this method in terms of its recognition and delineation accuracy and efficiency on three different data sets. In Sec. IV, we summarize our conclusions. A preliminary version of this paper has appeared in the Conference Proceedings of the *SPIE 2009 Medical Imaging Symposium*.<sup>28</sup>

## II. MULTIOBJECT ORIENTED ACTIVE SHAPE MODELS

### II.A. Overview of the approach

The proposed MOASM is a multiobject generalization of the OASM method.<sup>6</sup> Compared to one object in OASM, there are multiple objects in the model in MOASM. Let  $O_1, O_2, \dots, O_m$  be the physical objects of interest in a given body region  $B$ , such as the liver, lungs, heart, etc., in the thoracic region. For segmenting the boundaries of these objects in an image of a particular manifestation of them, ASM captures the statistical variations in the boundaries of these objects within these objects' family via a statistical shape model  $M$ . MOASM, like OASM, considers each object boundary to be oriented and determines a cost structure  $K$  associated with  $M$  via the principles underlying the live wire method. As per this cost structure, every shape instance  $\mathbf{x}_O$  of every object  $O$  in  $B$  generated by  $M$  is assigned a total cost  $K_O(\mathbf{x}_O)$  in a given image  $I$ . This cost is determined from the live wire segments generated in  $I$  between all pairs of successive landmarks of the shape instance  $\mathbf{x}_O$ . MOASM seeks that set of oriented boundaries in  $I$  corresponding to  $O_1, \dots, O_m$ , each of which is a sequence of live wire segments between successive pairs of landmarks of a shape instance  $\mathbf{x}_O$ , such that  $\mathbf{x}_O$  satisfies the constraints of  $M$ , and the total cost  $\sum_O K_O(\mathbf{x}_O)$  over all objects in  $I$  considered in the model is the smallest possible. The main steps involved in MOASM are listed below. Each step is described in detail in

each subsection of Sec. II. A comparison is also given between the proposed MOASM and the OASM method in each subsection.

### II.A.1. Algorithm MOASM

#### II.A.1.a. Model building /training phase.

- (T1) Specify landmarks on boundaries of objects  $O_1, \dots, O_m$  in the training images provided for body region  $B$ .
- (T2) Construct a shape model  $M$  for the objects in  $B$  from the landmarks and training images.
- (T3) Create boundary cost function  $K$ .

#### II.A.1.b. Segmentation phase.

- (S1) Initialization/recognition: Determine, in the given image  $I$  of  $B$  of a patient, the pose at which  $M$  should be set in  $I$  so that the model boundaries are close to the real object boundaries in  $I$ . Let the shape instance of the multiple object assembly corresponding to the recognized site be  $\mathbf{x}=(\mathbf{x}^{O_1}, \dots, \mathbf{x}^{O_m})$ .
- (S2) Delineation: For the shape instances  $\mathbf{x}$  of the multiple object assembly, determine the best oriented boundaries in  $I$  as per the MOASM method.
- (S3) If the convergence criterion is satisfied, output the best oriented boundaries found in S2 and stop. Else subject  $\mathbf{x}$  to constraints of model  $M$  and go to step S2.

In this procedure, T1–T3 constitute training or model creation steps, and S1–S3 represent initialization/recognition and delineation steps. These steps are described in the following sections in detail.

### II.B. T1: Specifying landmarks

Similar to the OASM method, MOASM specifies landmarks for multiple objects. Suppose each object  $O_i$  considered for inclusion in the model has  $n_i$  landmarks,  $1 \leq i \leq m$ . Then the vector representation for planar shapes would be

$$\begin{aligned} \mathbf{x} &= (\mathbf{x}^{O_1}, \mathbf{x}^{O_2}, \dots, \mathbf{x}^{O_m}) \\ &= (x_1^{O_1}, x_2^{O_1}, \dots, x_{n_1}^{O_1}, x_1^{O_2}, \dots, x_{n_2}^{O_2}, \dots, x_{n_m}^{O_m}) \\ &= (x_1^{O_1}, y_1^{O_1}, x_2^{O_1}, y_2^{O_1}, \dots, x_{n_1}^{O_1}, y_{n_1}^{O_1}, x_1^{O_2}, y_1^{O_2}, \dots, x_{n_2}^{O_2}, y_{n_2}^{O_2}, \dots, \\ &\quad x_{n_m}^{O_m}, y_{n_m}^{O_m}). \end{aligned} \quad (1)$$

In many ASM studies, a manual procedure is used to label the landmarks in a training set, although automatic methods are also available for this purpose. For our approach, any such method will work, although we have used the manual method in producing all presented results.

### II.C. T2: Building model $M$

To obtain a true shape representation of the family of an object  $O_i$ , location, scale, and rotation effects within the family need to be filtered out. This is done by aligning shapes within the family of  $O_i$  (in the training set) to each other by changing the pose parameters (scale, rotation, and

translation).<sup>29</sup> For multiple objects, the object assemblies are aligned. Principal component analysis (PCA) is then applied to the aligned  $N$  training shape vectors  $\mathbf{x}_i$ ,  $i=1, \dots, N$ . The model  $M$  is then constructed following the ASM procedure.<sup>29</sup>

### II.D. T3: Creating boundary cost function $K$

Similar to OASM, the live wire method of Ref. 25 is used in MOASM because of its boundary orientedness properties and the distinct graph model it uses. Unlike OASM, MOASM trains the boundary cost functions  $K_1, K_2, \dots, K_m$  for each object  $O_1, O_2, \dots, O_m$  separately, where  $K_i$  is the boundary cost structure for  $O_i$ . Each cost structure  $K_i$  is specifically optimally tailored for each object  $O_i$ , as briefly explained below (see Ref. 25 for details).

A *boundary element*, *bel* for short, of  $I$  is an ordered pair  $(d, e)$  of four-adjacent pixels  $d$  and  $e$ . It represents the oriented edge between pixels  $d$  and  $e$ ,  $(d, e)$ , and  $(e, d)$  representing its two possible orientations. To every *bel* of  $I$ , we assign a set of features. The features are intended to express the likelihood of the *bel* belonging to the boundary (of a particular object  $O_i$ ) that we are seeking in  $I$ . The cost  $c(b)$  associated with each *bel*  $b$  of  $I$  is a linear combination of the costs assigned to its features,

$$c(b) = \frac{\sum_i w_i c_f(f_i(g))}{\sum_i w_i}, \quad (2)$$

where  $w_i$  is a positive constant indicating the emphasis given to feature function  $f_i$  and  $c_f$  is the function to convert feature values  $f_i(g)$  to cost values  $c_f(f_i(g))$ . In live wire,<sup>25</sup>  $f_i$ 's represent features such as intensity on the immediate interior of the boundary, intensity on the immediate exterior of the boundary, and gradient magnitude at the center of the *bel*.  $c_f$  is an inverted Gaussian function, and here, identical weights  $w_i$  are used for all selected features.

For the purpose of MOASM, we shall utilize the feature of live wire to define the best oriented path between any two points as a sequence of *bels* with minimum total cost. The only deviation in this case is that the two points will be taken to be any two successive landmarks employed in the model, and the landmarks are assumed to correspond to pixel vertices. With this facility, we assign a cost to every pair of successive landmarks of any shape instance  $\mathbf{x}^{O_i}$ , which represents the total cost of the *bels* in the best oriented path  $\langle b_1, b_2, \dots, b_h \rangle$  from landmark  $x_k^{O_i}$  to landmark  $x_{k+1}^{O_i}$ . That is,

$$\kappa(x_k^{O_i}, x_{k+1}^{O_i}) = \sum_{j=1}^h c(b_j). \quad (3)$$

For any shape instance  $\mathbf{x}^{O_i}=(x_1^{O_i}, x_2^{O_i}, \dots, x_{n_i}^{O_i})$  of  $O_i$ , the cost structure  $K_i=\kappa(\mathbf{x}^{O_i})$  may now be defined as

$$K_i(\mathbf{x}^{O_i}) = \sum_{k=1}^{n_i} D(x_k^{O_i}) \kappa(x_k^{O_i}, x_{k+1}^{O_i}) D(x_{k+1}^{O_i}), \quad (4)$$

where we assume that  $x_{n_i+1}^{O_i}=x_1^{O_i}$  and  $D(x_k^{O_i})$  is the Mahalanobis distance for the intensity profile at  $x_k^{O_i}$  for object  $O_i$  as defined below,

$$D(\mathbf{x}_k^{O_i}) = (g_k - \bar{g}_k)^T (S_k^g)^{-1} (g_k - \bar{g}_k), \tag{5}$$

where  $g_k$  is the intensity profile at landmark  $\mathbf{x}_k^{O_i}$  (we sample  $2l+1$  intensities along the direction perpendicular to the boundary at this landmark) and  $\bar{g}_k$  and  $S_k^g$  are the mean and covariance matrix of the intensity profile at this landmark, respectively.

That is,  $K_i(\mathbf{x}^{O_i})$  is the weighted sum of the costs associated with the best oriented paths between all  $n_i$  pairs of successive landmarks of shape instance  $\mathbf{x}^{O_i}$ . Thus, once the best cost function is determined via training (as described in Ref. 25),  $K_i$  is also determined automatically by Eq. (4). Then, for MOASM, the total boundary cost function can be created as  $K(\mathbf{x}) = \sum_{i=1}^m K_i(\mathbf{x}^{O_i})$ . The multiobject model  $M$  together with the multi object cost structure  $K$  constitutes our MOASM:  $M_{\text{MOASM}} = (M, K)$ .

**II.E. S1: Automatic initialization/recognition**

The goal of automatic initialization, or object recognition, is to find an initial pose for  $M$  in the given image  $I$ . Compared to OASM, automatic initialization (recognition) via MOASM is more robust and efficient. As the number of objects in the model increases and if the objects are selected strategically, the ways in which objects can be fitted in the given image to be segmented become severely constrained, and therefore, the search space becomes substantially smaller. So, even when there is only one object of interest, MOASM uses multiple objects in the model just to make the recognition of the object of interest, and hence its delineation, more accurate, robust, and effective.

Suppose  $p$  denotes the pose vector for the object assembly, which includes a location component  $(tx, ty)$ , scale component  $(s)$ , and orientation component  $(\theta)$ . The goal of recognition is to find the best initial pose in  $I$  for  $M_{\text{MOASM}}$ . Our experiments indicate that, because of the globally optimal delineations and the orientedness nature of  $M_{\text{MOASM}}$ , the small variations in orientation observed in strict protocol guided clinical image acquisitions can be automatically handled in the delineation step, and thus,  $\theta$  can be ignored for the purpose of recognition. Thus,  $p$  becomes three-dimensional  $(tx, ty, s)$ . The recognition task is then to find

$$p^* = \arg \min_p BC(p, I, K_1, \dots, K_m), \tag{6}$$

where  $BC(p, I, K_1, \dots, K_m) = \sum_{i=1}^m K_i(\mathbf{x}^{O_i})$ , which denotes the sum of the cost (as per cost structures  $K_1, \dots, K_m$ ) of the boundaries of all  $m$  objects delineated starting with the model at  $p$ . As  $m$  increases, the constraints imposed become quite severe and the search for  $p^*$  becomes remarkably easier, and accurate recognition can be achieved by discretizing the space of  $p$  and by performing exhaustive search within a restricted space of the  $p$  vectors. The automatic initialization method proposed here is an essential underpinning of the MOASM method. It relies on the fact that, at a position of a shape instance of MOASM that is close to the correct boundaries of  $O_1, \dots, O_m$  in  $I$ , the total cost BC of the oriented boundaries of all objects included in MOASM is likely to be sharply smaller than the cost of the oriented

boundary assembly found at other locations in  $I$ . In our implementation, the recognition process consists of the following steps. Let the given image to be segmented be  $I$  and the mean pose vector  $\bar{p} = (\bar{tx}, \bar{ty}, \bar{s})^t$  represent the mean location  $(\bar{tx}, \bar{ty})$  and scale  $(\bar{s})$  observed over all training images for the objects included in the model. Let  $SS$  be a search space—a set of pose vectors—centered around  $\bar{p}$ .

**Object recognition algorithm**

Begin

1. For each  $p \in SS$ , perform steps 2 and 3.
  2. Place  $M$  at  $p$ . Deform each object as per standard ASM searching method.
  3. Evaluate  $BC(p, I, K_1, \dots, K_m)$ .
  4. Find  $p^*$  corresponding to the minimum BC over all tested  $p \in SS$ .
  5. Output the delineated object shapes  $\mathbf{x}_p = (\mathbf{x}_p^{O_1}, \dots, \mathbf{x}_p^{O_m})$  corresponding to  $p^*$ .
- end

It is important to note that  $SS$  is a small set of pose vectors uniformly sampled in a region around  $\bar{p}$  in the implementation. Typically, we have used a search region of  $100 \times 100 \times 0.5$  (in  $x, y, s$  units, where  $x$  and  $y$  units are in terms of the number of pixels) which is sampled at every 20–25 pixels along  $x$  and  $y$  and 0.1–0.15 in  $s$ . Thus,  $SS$  consists of roughly 50–125 vectors.

**II.F. S2: Delineation**

This step assumes that the recognized (initialized) shape assembly instance  $\mathbf{x}_p$  of the objects in  $M$  derived from step S1 is sufficiently close to the actual boundaries of  $O_i$  (for all  $1 \leq i \leq m$ ) in  $I$ . It then determines what the new position of the landmarks of the objects represented in  $\mathbf{x}_p$  should be such that the sum of the costs of the oriented paths between all pairs of successive landmarks is the smallest, where the cost takes into account not only the boundary cost  $K_i$  of each object  $O_i$  [Eq. (4)] but also the relationship among objects. This is accomplished through a three-level dynamic programming (3LDP) algorithm. Compared to 2LDP in the OASM method, the 3LDP algorithm in MOASM not only

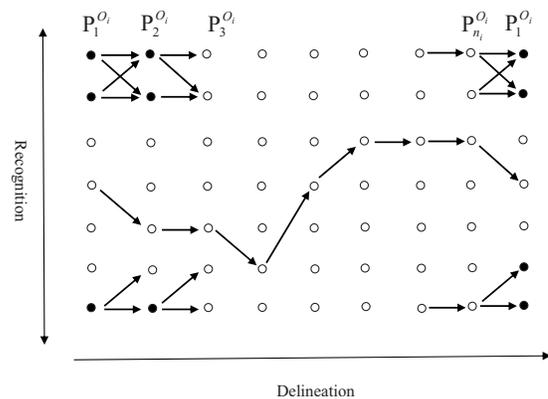


FIG. 1. The weighted graph used in the second level dynamic programming. The nodes in each column represent the set of points  $P_k^{O_i}$  that are sampled along a line orthogonal to the boundary at each landmark  $\mathbf{x}_k^{O_i}$ .

attempts to find an optimal path for each object but also aims at delineating all objects in a consistent and globally optimal way.

Let the shape for object  $O_i$  to be modified be  $\mathbf{x}^{O_i} = (x_1^{O_i}, x_2^{O_i}, \dots, x_{n_i}^{O_i})$ . At each landmark  $x_k^{O_i}$ ,  $L=2q+1$  ( $q > l$ , where  $l$  is the number of points selected on each side of  $x_k^{O_i}$  in the appearance aspect of the model during training) points are selected, including  $x_k^{O_i}$ , with  $q$  points on either side of  $x_k^{O_i}$  along a line perpendicular to the shape boundary at  $x_k^{O_i}$ . Let the set of these points be denoted by  $P_k^{O_i}$ .

From each point in  $P_k^{O_i}$ , there exists a minimum cost oriented path to each point  $P_{k+1}^{O_i}$  in  $I$ , which can be determined via a first level DP as in the live wire method. We are interested in that set of minimum cost paths, one selected between each pair  $(P_k^{O_i}, P_{k+1}^{O_i})$ , such that, the resulting closed boundary is continuous, and its total cost is the smallest possible. This problem can be solved via a second level of DP as illustrated in the graph of Fig. 1. In this graph, the set of nodes is  $P_1^{O_i} \cup P_2^{O_i} \cup \dots \cup P_{n_i}^{O_i} \cup P_1^{O_i}$ , and the set of directed arcs is  $(P_1^{O_i} \times P_2^{O_i}) \cup (P_2^{O_i} \times P_3^{O_i}) \cup \dots \cup (P_{n_i-1}^{O_i} \times P_{n_i}^{O_i}) \cup (P_1^{O_i} \times P_1^{O_i})$ . Each arc  $(u, v)$  in this graph has a cost which is simply the cost  $\kappa(u, v)$  [Eq. (3)] of the minimum cost oriented (live wire) path between  $u$  and  $v$  in  $I$ . Thus, each directed arc such as  $(u, v)$  also represents an oriented boundary segment from  $u$  to  $v$  as a sequence of *bels*. Note that a directed path, such as the one shown in the middle of Fig. 1, starting from some node  $u$  in  $P_1^{O_i}$  in the first column and ending in the same node in the last column constitutes a closed, oriented boundary as a sequence of *bels*. Our aim is to find the best of all such directed paths (closed, oriented boundaries), each starting from some node in the first column and ending in the same node in the last column. This problem can be solved by a second level of DP. Since there are  $L$  nodes in the first (and every) column, in OASM, the second level DP is applied  $L$  times, and a directed closed path among the  $L$  optimal paths which yields the lowest total cost is considered to be the best and the desired delineation of  $O_i$ .

In the MOASM method, instead of choosing the best among the  $L$  optimal paths for each object (as done in OASM), one more level (object level) optimization is carried out. This is done by combining the cost of the optimal boundaries with the cost associated with object relationships to ensure propriety of relationships as described below. Suppose we determine a linear order for the objects, denoted by  $O_1, \dots, O_m$ , and fix this order once for all (how to choose the order is discussed later). Let the  $L$  optimal boundaries of  $O_i$  found above be  $B_\ell^{O_i}$ ,  $1 \leq \ell \leq L$ . We quantify the relationship between successive objects  $(O_i, O_{i+1})$ ,  $1 \leq i \leq m$ , in the sequence via three factors, for  $1 \leq i \leq m$  and  $1 \leq \ell, k \leq L$ : Relative position  $RP_{\ell,k}^i = \bar{x}_\ell^{O_i} - \bar{x}_k^{O_{i+1}}$ , where  $\bar{x}_\ell^{O_i}$  is the geometric center of  $B_\ell^{O_i}$ ; relative size  $RS_{\ell,k}^i = A_\ell^{O_i} / A_k^{O_{i+1}}$ , where  $A_\ell^i$  is the  $B_\ell^{O_i}$  area enclosed by  $B_\ell^{O_i}$ ; and boundary distance  $BD_{\ell,k}^i$ , which is the Hausdorff distance between  $B_\ell^{O_i}$  and  $B_k^{O_{i+1}}$ . The relationship  $R_\ell^i$  between  $O_i$  and  $O_{i+1}$  is quantitatively expressed by a weighted combination of the three factors as

$$R_{\ell,k}^i = \sum_{x \in \{RP, RS, BD\}} w_x G_x(\mu_x, \sigma_x), \quad (7)$$

where  $G_x(\mu_x, \sigma_x)$  denotes a Gaussian function with mean  $\mu_x$  and standard deviation  $\sigma_x$ .  $G_{RP}$  is a Gaussian function of  $RP_{\ell,k}^i$ . Its parameters  $\mu_{RP}$  and  $\sigma_{RP}$  are estimated from the training data sets.  $G_{RS}$  and  $G_{BD}$  and their parameters are defined in an analogous manner.  $w_x$  are weights chosen such that  $\sum_x w_x = 1$ . (In all experiments, we used  $w_{RP} = w_{RS} = 0.3$ .)

Given  $L$  optimal shapes for each  $O_i$ ,  $1 \leq i \leq m$ , the best shape for each  $O_i$  and the arrangement of objects in  $I$  is determined via a third level DP (3LDP) carried out on the graph shown in Fig. 2. Each of the  $L$  optimal shapes  $B_\ell^{O_i}$  for each  $O_i$  constitutes a node in this graph and a pair such as  $(B_\ell^{O_i}, B_k^{O_{i+1}})$ ,  $1 \leq i \leq m$ ,  $1 \leq \ell, k \leq L$ , constitutes a directed arc. The integrated cost assigned to this arc is

$$IC(B_\ell^{O_i}, B_k^{O_{i+1}}) = K_i(B_\ell^{O_i}) \cdot K_{i+1}(B_k^{O_{i+1}}) \cdot R_{\ell,k}^i. \quad (8)$$

As in Fig. 1, we observe that there is an optimum path in the graph of Fig. 2, starting from each node in the first column and ending on some node in the last column. The MOASM method finds each such optimum path (which represents one arrangement of the object assembly) and the best among  $L$  such arrangements via 3LDP.

The 3LDP algorithm may be summarized as follows. Given a shape assembly  $\mathbf{x}_p = (x_p^{O_1}, x_p^{O_2}, \dots, x_p^{O_m}) = (x_1^{O_1}, x_2^{O_1}, \dots, x_{n_1}^{O_1}, x_1^{O_2}, \dots, x_{n_2}^{O_2}, \dots, x_{n_m}^{O_m})$  and  $I$ , it outputs a new shape assembly  $\mathbf{x}_o$  and the associated optimum oriented boundaries as a sequence of *bels*.

#### Algorithm 3LDP

---

*Input:* Initialized shape assembly  $\mathbf{x}_p$   
*Output:* Resulting shape assembly  $\mathbf{x}_o$  and the associated oriented boundaries.

```

begin
1. For each object  $O_i$  ( $1 \leq i \leq m$ ), do steps 2 to 4.
2. Determine  $P_1^{O_i}, P_2^{O_i}, \dots, P_{n_i}^{O_i}$  sets of points for  $O_i$  in  $I$ .
3. Determine costs  $\kappa(u, v)$  via first level DP for all directed arcs in Fig. 1.
4. For each point  $u$  in  $P_1^{O_i}$ , determine the best directed path from  $u$  back to  $u$  via the second level DP, the corresponding shape  $\mathbf{x}^{O_i}$ , and its total cost  $K_i(\mathbf{x}^{O_i})$ .
5. Compute arc costs  $IC(B_\ell^{O_i}, B_k^{O_i})$  in Fig. 2. Find the best path  $\mathbf{x}_o$  which has minimal total cost in the graph of Fig. 2 via third level DP.
6. Output the optimum shape assembly  $\mathbf{x}_o$  and the corresponding oriented boundaries.
end

```

## II.G. S3: Applying model constraints

The convergence criterion used here is a measure of the distance between two shapes encountered in two consecutive executions of step S2. This measure is simply the maximum of the distance between corresponding landmarks in the two shape assemblies of all objects. If this distance is greater than 0.5 pixel unit, the optimum shape assembly found in step S2 is subjected to the constraints of model  $M$ . Then the iterative process is continued by going back to step S2. Otherwise, the

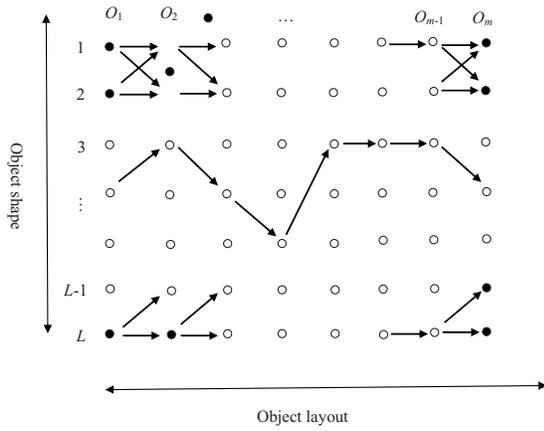


FIG. 2. The graph used in the third level of optimization. There are  $m$  objects to be delineated and  $L$  possible optimal shapes for each object. The horizontal direction represents different objects. The vertical direction represents the different object shapes.

MOASM process is considered to have converged and it stops with an output of the optimum shape assembly and the optimum oriented boundaries found in the process.

**III. EXPERIMENTAL RESULTS**

In this section, we demonstrate both qualitatively, through image display, and quantitatively, through evaluation experiments, the extent of effectiveness of the MOASM strategy. Three clinical image data sets which include abdominal CT, chest CT, and foot MRI have been considered. Our method of evaluation, based on the framework of reference,<sup>30</sup> will focus on the analysis of accuracy and efficiency of MOASM. We will also present an evaluation of the proposed object recognition strategies. We will consider manual segmentation performed in these different data sets to constitute a surrogate of true segmentation for assessing the delineation and recognition accuracy of the methods.

**III.A. Image data sets**

The image data sets and objects employed in the experiments are described in Table I and Fig. 3. Both CT data sets

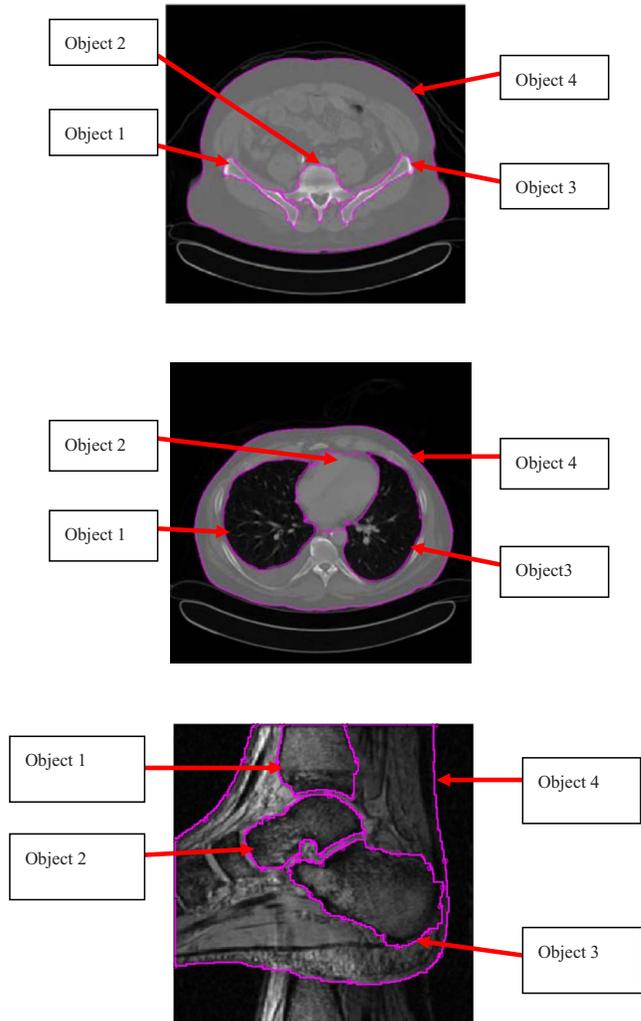


FIG. 3. The four objects selected in the abdominal CT, chest CT, and foot MRI data sets.

constitute slices selected from 3D CT studies on a Siemens Sensation 16 CT scanner with a slice spacing of 5 mm, image size of  $512 \times 512$ , and pixel size of  $0.78 \times 0.78$  mm<sup>2</sup>. The MRI Foot data set constitutes slices selected from 3D studies on a GE 1.5T MRI scanner with a slice spacing of 1.3

TABLE I. Description of the image data sets used in the three AAR experiments.

Data set	Objects	Number of landmarks used	Image domain	No. of images
Abdominal CT	(1) Right pelvic bone	10	512 × 512	40
	(2) Vertebra	12		
	(3) Left pelvic bone	10		
	(4) Skin boundary	12		
Chest CT	(1) Right lung	10	512 × 512	40
	(2) Heart	9		
	(3) Left lung	10		
	(4) Skin boundary	12		
Foot MRI	(1) Tibia	7	256 × 256	40
	(2) Talus	14		
	(3) Calcaneous	15		
	(4) Skin boundary	11		

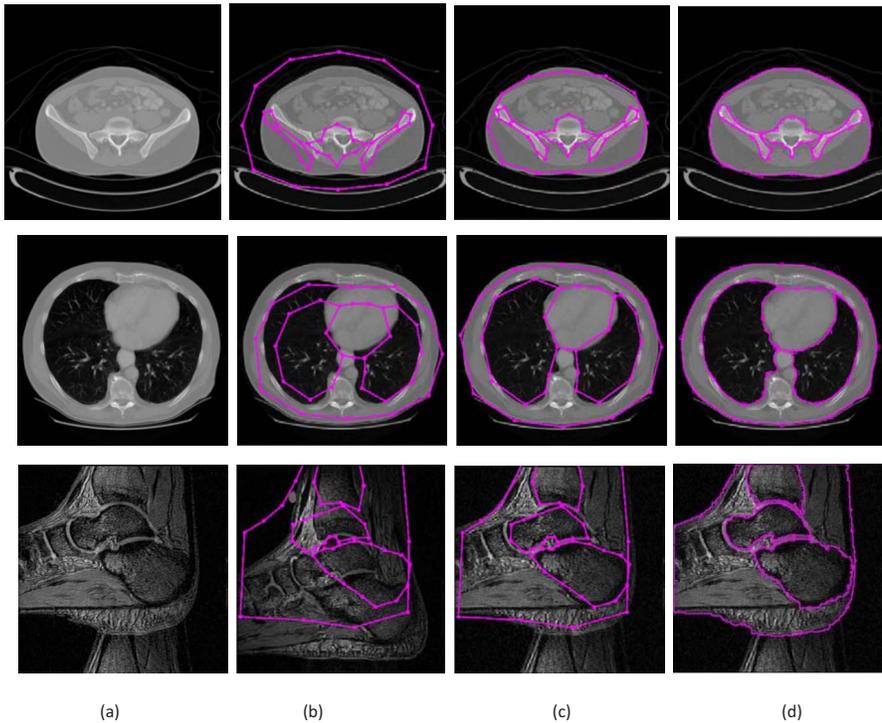


FIG. 4. One example from each of the three data sets: (a) original image, (b) default initial model pose, (c) automatic recognition result, and (d) delineation result by MOASM.

mm, image size of  $256 \times 256$ , and pixel size of  $0.55 \times 0.55 \text{ mm}^2$ . In each data set, 40 slices selected from full 3D images, acquired from 15 different subjects are used. These slices are approximately at the same anatomic location in the body, so that, for each object, the 40 2D images in each set can be considered to represent images of a family of objects of same shape. Two to three slices are taken on average from the same subject's data, either from the same 3D image or from different 3D images. Among the 40 images, 25 are randomly selected for training, and the rest are used as test images. MOASM models were built by manually selecting landmarks on the segmented boundaries.

### III.B. Qualitative analysis

A subjective inspection revealed that, in all experiments, the MOASM results matched the perceived boundary very well. Some examples are displayed for the three data sets in Fig. 4. Automatic initialization based on location and scale search worked well in all cases in the sense that initialized shapes were found close to the true boundary. In the figure, the original image with the mean shape assembly in its default pose is displayed in Fig. 4(b). In Fig. 4(c), the model

shapes resulting from the recognition process, which form an input to the MOASM algorithm, are displayed. The final delineation results are shown in Fig. 4(d).

### III.C. Quantitative analysis of segmentation

Here, we will focus on the analysis of accuracy and efficiency of MOASM from the perspective of the final delineation results and compare them with those of the classical ASM method. Accuracy here relates to how well the segmentation results agree with the true delineation of the objects. Efficiency indicates the practical viability of the method, which is determined by the amount of time required for performing computations and for providing any user help needed in segmentation. The measures that are used under each of these groups and their definitions are given below. In these analyses, all test data sets from all three groups of data were utilized.

*Accuracy.* The following measures, called TPVF and (FPVF) (Ref. 30) are used to assess the accuracy of the methods. The ground truth for delineation (and recognition) is provided by manually delineating the shapes. In all applications, all data sets have been previously manually segmented

TABLE II. Mean and standard deviation of TPVF and FPVF for ASM and MOASM.

Data set	TPVF		FPVF	
	ASM	MOASM	ASM	MOASM
Abdominal CT	$88.56\% \pm 2.2\%$	$97.3\% \pm 1.1\%$	$0.66\% \pm 0.01\%$	$0.13\% \pm 0.01\%$
Chest CT	$87.89\% \pm 0.15\%$	$98.3\% \pm 0.2\%$	$0.99\% \pm 0.02\%$	$0.17\% \pm 0.01\%$
Foot MRI	$90.01\% \pm 0.34\%$	$97.2\% \pm 0.3\%$	$0.41\% \pm 0.01\%$	$0.11\% \pm 0.01\%$

TABLE III. Mean operator time  $T_o$  and computational time  $T_c$  (in seconds) in all experiments for ASM and MOASM. The number of landmarks used is indicated by  $n\ell$ .

Data Set	$T_o$		$T_c$	
	ASM	MOASM	ASM	MOASM
Abdominal CT	160( $n\ell=132$ )	60( $n\ell=44$ )	18	30
Chest CT	130( $n\ell=123$ )	55( $n\ell=41$ )	16	28
Foot MRI	170( $n\ell=141$ )	65( $n\ell=47$ )	12	20

in the domain. Table II lists the mean and standard deviation values of TPVF and FPVF achieved by using ASM and MOASM methods. TRVF/FPVF are calculated as per Ref. 30 for each object separately. Then the average TRVF/FPVF over all objects is used as the performance. It shows that MOASM produces considerably more accurate final segmentations than the ASM method. It is important to note that only a small set of landmarks is needed for each object in MOASM as shown in Table I due to the efficiency of live wire for boundary delineation. However, with more landmarks, the method will have even superior performance.

*Efficiency.* Both methods are implemented on an Intel Pentium IV PC with a 3.4 GHz CPU using MATLAB programming. In determining the efficiency of a segmentation method, two aspects should be considered—the computation time ( $T_c$ ) and the human operator time ( $T_o$ ). The mean  $T_c$  and  $T_o$  per data set estimated over the three data sets for each experiment are listed in Table III.  $T_o$  measured here is the operator time required in the training step (for selecting landmarks). Table III shows that the operator time (training) required in MOASM is much less than that of ASM since far fewer landmarks are needed in MOASM. The computation time required in MOASM is a little more than that of ASM because of the 3LDP algorithm.

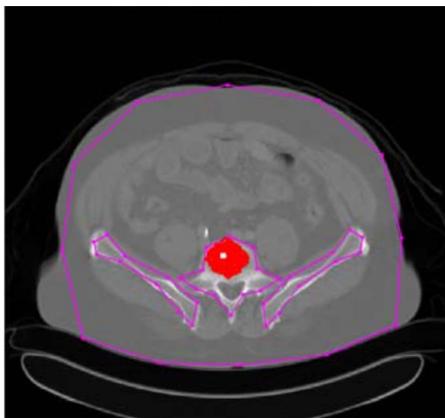


FIG. 5. Illustration of robust region RR for an image from the abdominal data set. The white point is the geometric center of the recognized shape assembly. In this case recognition is perfect. The darker region around the white point located within the vertebral body is the robust region RR considering just the  $x$  and  $y$  location components of  $p$ .

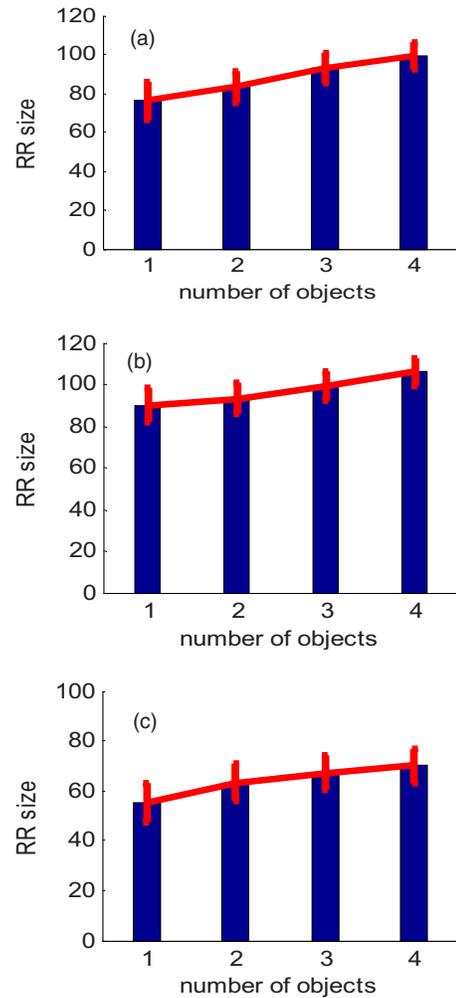


FIG. 6. The volume of the robust region in the  $(x, y, s)$  space for different number of objects. (a) Abdominal CT data set, (b) chest CT data set, and (c) MRI data set. The mean value and standard deviation over all test data sets are shown.

### III.D. Evaluation of object recognition strategies

It is clear that the effectiveness of recognition depends very much on the integrated cost function  $IC$ , which in turn depends on the cumulative boundary cost function  $BC$ , and the cost structures  $K_1, \dots, K_m$ . This in turn implies that the recognition efficacy is influenced by the particular combination of objects included in MOASM. Here, we evaluate the relationship between object recognition and delineation accuracy and the number of objects considered in  $M$  and their actual spatial distribution.

*Robust region.* To characterize the robustness of the recognition method, we define a concept called *robust region* RR as

$$RR = \{\text{pose vector } p \mid \text{when the recognized object assembly is at } p, \text{ the resulting delineation accuracy is acceptably "high" (say, with TPVF} > 97\% \text{ and FPVF} < 3\%)\}. \quad (9)$$

RR is a useful factor to quantify the robustness of the recog-

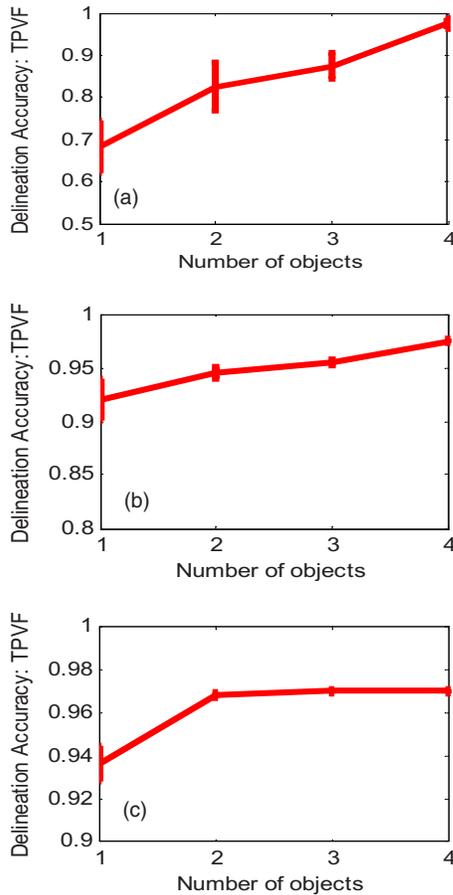


FIG. 7. The dependence of delineation accuracy on the number of objects in the three data sets. (a) Abdominal CT data set, (b) chest CT data set, and (c) foot MRI data set. For labels 1, 2, and 3, all combinations of 1, 2, and 3 objects were considered. The mean value over all test data sets and standard deviation are shown.

dition method and to compare among recognition strategies. The size of RR is an indicator of the robustness of the recognition strategy resulting from a particular choice of the objects included in MOASM. The larger the robust region is, the more robust the recognition method is. Figure 5 shows an illustration of the robust region RR (red region) for an image from the abdominal CT data set. The white point is the geometric center of the recognized shape. The red region is the robust region RR, considering only the  $x, y$  location component of  $p$ . In this case, as the white point is within the red region, it implies high delineation accuracy. The size of the robust region in the AAR method over all images in these three data sets is roughly 22 pixels  $\times$  23 pixels  $\times$  0.20, 23 pixels  $\times$  24 pixels  $\times$  0.20 and 20 pixels  $\times$  22 pixels  $\times$  0.16, respectively. Figure 6 illustrates how the size (volume) of RR ( $x$  range  $\times$   $y$  range  $\times$  scale range) varies for different numbers of objects in these three data sets. The results suggest that the MOASM recognition process becomes more robust as the number of objects in the model increases. This is true even in studying only one object.

*Accuracy dependence on the number of objects, objects' distribution, and object size.* It is interesting to note that the recognition accuracy improves with increasing number of

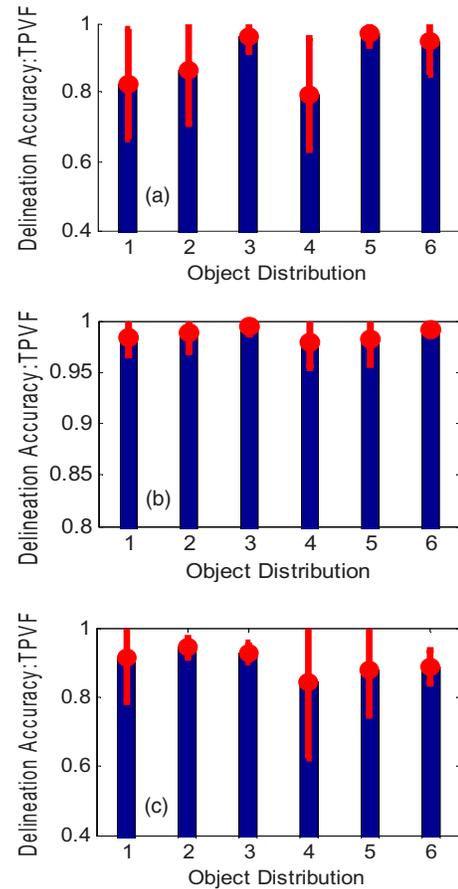


FIG. 8. The dependence of delineation accuracy on the distribution of objects in the three data sets. Distribution: 1: objects 1, 2; 2: objects 1, 3; 3: objects 1, 4; 4: objects 2, 3; 5: objects 2, 4; 6: objects 3, 4. (a) Abdominal CT data set; (b) chest CT data set; (c) foot MRI data set.

objects. Figure 7 shows the experimental results on the three data sets. In order to test how the objects' distribution may influence recognition accuracy, two objects were selected in different combinations for the three data sets. Figure 8 shows the experimental results on these three data sets. We observe that the recognition accuracy is higher if the objects are distributed more evenly. We also found that it is a good idea to include the encompassing object (such as skin boundary) as it provides a strong and robust constraint. Figure 9 demonstrates that delineation accuracy depends on the size of the object also. The figure shows how the delineation accuracy of object 1 in the three data sets improves as the object used for recognition is chosen to have increasingly larger size. On the horizontal axis, label 1 indicates that only object 1 is used in the model for recognizing and delineating object 1. Label 2 indicates that, in addition to object 1, object 2 is also employed in the model, and so on.

*Interdependence between recognition and delineation.* The effectiveness of AAR depends on the interaction between object recognition and object delineation; in other words, the pose at which the objects should be initialized influences delineation accuracy. Conversely (obviously), the recognition accuracy itself depends on delineation accuracy since  $IC$ ,  $BC$ , and  $K_i$  will depend on the delineated bound-

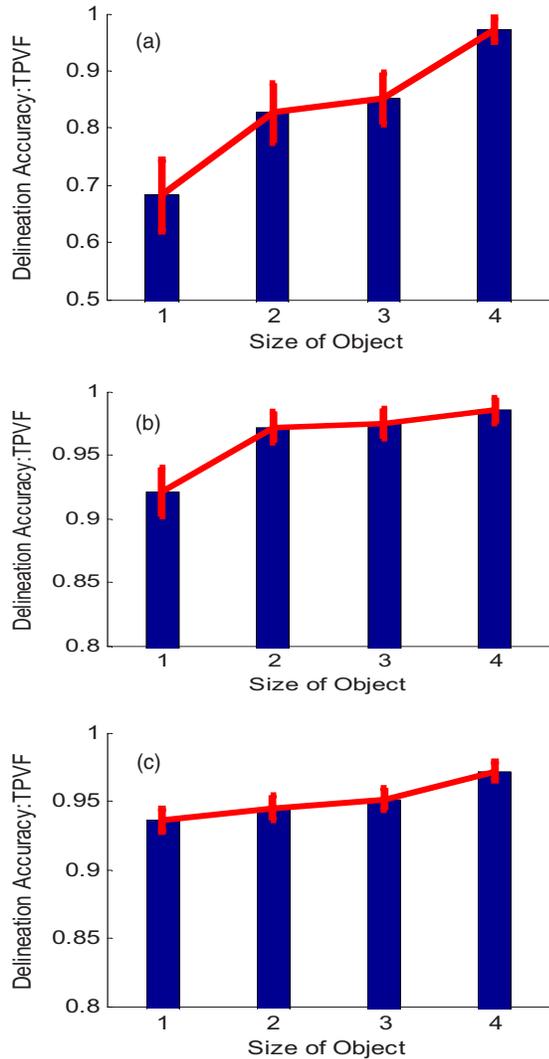


FIG. 9. Illustration of how better recognition via larger objects included in MOASM influences delineation accuracy. The delineation accuracy is shown for object 1 in all three data sets. The horizontal axis indicates the number of the object included in the model (with increasing size) in addition to object 1 for facilitating recognition. (a) Abdominal CT data set. (b) Chest CT data set. (c) Foot MRI data set.

aries. Experiments were also carried out to study the interdependence between recognition and delineation. We find that (Fig. 10) with the improvement of recognition, delineation also becomes more accurate. It means that good delineation can help in recognition and perfect recognition can make delineation most accurate. This is the spirit of the synergy established between ASM and live wire by MOASM. Delin-

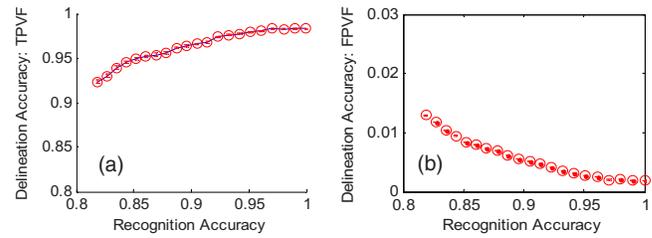


FIG. 10. Recognition accuracy versus delineation accuracy in the abdominal CT data set. (a) TPVF versus recognition accuracy; (b) FPVF versus recognition accuracy.

ation accuracy is shown over the same objects that are used in the model for recognition. Here, the recognition accuracy is defined as an inverse function of the distance of the geometric center *loc* of the model at the end of recognition from RR, as follows:

$$A_R = \exp(-\alpha \cdot \text{dist}(\text{loc}, \text{RR})), \quad (10)$$

where  $\alpha$  is a constant (here  $\alpha=0.1$ ). If the distance is 0,  $A_R = 1$  (perfect recognition).

### III.E. How far is MOASM solution from the global optimum?

As described in Ref. 6, the single-object OASM algorithm finds a globally optimal boundary if the recognition step brings the model to a pose, wherein the line profile drawn orthogonal to the model shape boundary at every landmark is intersected by the globally optimal boundary. This property is valid also for the MOASM algorithm. From Figs. 6–10, it is also clear that MOASM has better performance than (the single-object) OASM in terms of both delineation and recognition accuracy. To determine how often globally optimum solutions are actually found, we examined the MOASM recognition results for each of the four objects in all ( $15 \times 3 = 45$ ) test data sets to ascertain how often the above optimality property is satisfied. Table IV suggests that, overall, in 97% of the cases the globally optimal solution is actually found. We also computed the percent difference between the true optimum boundary cost (which can be determined since the true boundaries are known) and the optimum found by the MOASM method for the chest CT data set. These are listed in Table V, which shows that the differences are quite small, although for object 3 (heart) the cost function can be

TABLE IV. Results showing how often globally optimal boundaries are found for the four objects in the three data sets.

	Object 1	Object 2	Object 3	Object 4	Total
Abdominal CT	15/15=100%	14/15=93.3%	15/15=100%	14/15=93.3%	43/45=95.5%
Chest CT	15/15=100%	15/15=100%	14/15=93.3%	15/15=100%	44/45=97.7%
Foot MRI	15/15=100%	15/15=100%	15/15=100%	14/15=93.3%	44/45=97.7%
Total	45/45=100%	44/45=97.7%	44/45=97.7%	43/45=95.5%	131/135=97.0%

TABLE V. Difference between the true optimum cost and the cost of the optimum boundaries actually found for the different objects in the chest CT data set. Mean values over the 15 test images are shown.

Object 1	Object 2	Object 3	Object 4	Total
1.4%	2.3%	9.7%	0.04%	3.8%

perhaps improved. Better cost functions also lead to a larger RR and hence more efficient recognition.

All results reported thus far in this section were produced assuming the orders O1, O2, O3, and O4 for objects in all three data sets for devising the cost function  $IC$  referred to in Fig. 2, where the number assignment to objects is as depicted in Fig. 3. To test the dependence of delineation accuracy on this order, for the three data sets, we ran the entire segmentation experiment for five other object sequences. The results for all six sequences are summarized in Table VI. The results suggest that the delineation accuracy is more or less independent of the object sequence employed for  $IC$ .

#### IV. CONCLUDING REMARKS

This paper presented a generalization of the single-object oriented ASM methodology to multiple objects. The rationale for this generalization was, primarily, for improving the accuracy and robustness of object recognition, and secondarily, and consequently, also for improving delineation accuracy. The robustness of recognition also implies automated recognition. The paper introduced several new concepts: (1) using an assembly of objects even when there is only one object of interest, (2) separating recognition from delineation although they are interdependent, (3) the notion of robust region, (4) assessing recognition accuracy via RR independently of delineation, and (5) the dependence of recognition and delineation accuracy on the number and geographic distribution of objects included in the MOASM. These concepts were illustrated and validated via routine clinical chest CT, abdominal CT, and foot MRI data sets. The evaluated results indicate that (a) an overall delineation accuracy of TPVF  $>97\%$ , FPVF  $<0.2\%$  can be achieved, suggesting the feasibility of AAR. (b) Increasing the number of objects can considerably improve both recognition and delineation accuracy in clinical images. (c) More spread out arrangement of objects can lead to improved recognition and delineation accuracy. The larger the robust region is, the more robust the

recognition method is. It can be used to assess recognition accuracy and to compare among methods. It also helps to quantify the synergy in the method between recognition and delineation. Experimental verification indicates that the MOASM method almost always finds the global optimum and performs better than OASM.

Some ideas underlying MOASM can be further improved. First, the idea of creating an ordered sequence of the objects included in MOASM is somewhat artificial since there is no natural or meaningful sequential order for the objects. Our motivation for the idea was to achieve object level optimization within the existing computational framework of DP. Although, as illustrated in Table VI, there seems to be little variation in delineation accuracy when this order is changed, a more appropriate formulation would have been to define a complete graph of the objects for each  $1 \leq \ell \leq L$ , finding an optimum spanning tree in each, and finally choosing the best among the  $L$  optimal spanning trees. Second, in MOASM, we have made the cost function  $BC$  object specific which helps considerably in making recognition powerful via the integrated cost function  $IC$ . The latter can be possibly further improved by devising different cost functions for the section of the boundary between each pair of successive landmarks in each object. Since objects generally have interfaces with multiple neighboring objects, this strategy may match the cost function with the real variability in the characteristics commonly observed along object boundaries.

In this paper, the focus was on 2D images. Although we took 2D subsets of 3D images for illustration and evaluation, there are many 2D imaging modes and 2D type applications even in medical imaging where MOASM will be applicable. Generalizing the techniques to 3D images is certainly not trivial mainly from the view point of creating appropriate 3D models for OASM. Once the models are built and a 3D delineation algorithm is devised, then the above concepts are directly applicable to 3D images. One way is to segment the organ slice-by-slice, whose feasibility has been demonstrated for 3D organ initialization.<sup>7</sup> Alternatively, as demonstrated in Ref. 31 for OASM, multiple 2D object assembly models covering the body region can be employed to recognize and delineate 3D objects in 3D images.

#### ACKNOWLEDGMENTS

The authors' work is supported by NIH Grant No. EB004395.

TABLE VI. Delineation accuracy (mean and standard deviation) for six different object sequences considered in composing IC for the three data sets.

Order of objects	TPVF			FPVF		
	Abdominal CT	Chest CT	Foot MRI	Abdominal CT	Chest CT	Foot MRI
O1-O2-O3-O4	97.36% $\pm$ 1.12%	98.33% $\pm$ 0.26%	97.25% $\pm$ 0.35%	0.13% $\pm$ 0.01%	0.17% $\pm$ 0.01%	0.11% $\pm$ 0.01%
O1-O3-O2-O4	97.05% $\pm$ 1.19%	98.21% $\pm$ 0.29%	97.15% $\pm$ 0.36%	0.15% $\pm$ 0.01%	0.18% $\pm$ 0.01%	0.13% $\pm$ 0.01%
O2-O1-O3-O4	97.12% $\pm$ 1.20%	98.15% $\pm$ 0.32%	97.17% $\pm$ 0.31%	0.14% $\pm$ 0.01%	0.19% $\pm$ 0.01%	0.13% $\pm$ 0.01%
O2-O3-O1-O4	96.96% $\pm$ 1.21%	98.19% $\pm$ 0.31%	97.12% $\pm$ 0.39%	0.16% $\pm$ 0.01%	0.18% $\pm$ 0.01%	0.12% $\pm$ 0.01%
O3-O1-O2-O4	96.93% $\pm$ 1.25%	98.17% $\pm$ 0.29%	97.11% $\pm$ 0.38%	0.15% $\pm$ 0.01%	0.17% $\pm$ 0.01%	0.14% $\pm$ 0.01%
O3-O2-O1-O4	97.31% $\pm$ 1.15%	98.29% $\pm$ 0.29%	97.26% $\pm$ 0.33%	0.14% $\pm$ 0.01%	0.16% $\pm$ 0.01%	0.12% $\pm$ 0.01%

- <sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: jay@mail.med.upenn.edu; Telephone: 215-662-6780; Fax: (215) 898-9145; URL: <http://www.mipg.upenn.edu>
- <sup>1</sup> Y. Zhou and J. Bai, "Atlas-based fuzzy connectedness segmentation and intensity nonuniformity correction applied to brain," *IEEE Trans. Biomed. Eng.* **54**, 122–129 (2007).
- <sup>2</sup> J. Hansegard, S. Urheim, K. Lunde, and S. I. Rabben, "Constrained active appearance models for segmentation of triplane echocardiograms," *IEEE Trans. Med. Imaging* **26**(10), 1391–1400 (2007).
- <sup>3</sup> M. A. Horsfield, R. Bakshi, M. Rovaris, M. A. Rocca, S. V. S. R. Dandamudi, P. Valsasina, E. Eudica, F. Lucchini, C. R. G. Guttmann, M. P. Sormani, and M. Filippi, "Incorporating domain knowledge into the fuzzy connectedness framework: Application to brain lesion volume estimation in multiple sclerosis," *IEEE Trans. Med. Imaging* **26**, 1670–1680 (2007).
- <sup>4</sup> J. Yao and D. Chen, "Live level set: A hybrid method of livewire and level set for medical image segmentation," *Med. Phys.* **35**, 4112–4120 (2008).
- <sup>5</sup> M. Rousson and N. Paragios, "Prior knowledge, level set representation and visual grouping," *Int. J. Comput. Vis.* **76**, 231–243 (2008).
- <sup>6</sup> J. Liu and J. K. Udupa, "Oriented active shape models," *IEEE Trans. Med. Imaging* **28**, 571–584 (2009).
- <sup>7</sup> X. Chen, J. Yao, Y. Zhuge, and U. Bagci, in *3D Automatic Image Segmentation Based on Graph Cut-Oriented Active Appearance Models*, Proceedings of 2010 IEEE 17th International Conference on Image Processing, Hong Kong, September 26–19, 2010, pp. 3653–3656.
- <sup>8</sup> A. Besbes, N. Komodakis, G. Lings, and N. Paragios, in *Shape Priors and Discrete MRFs for Knowledge-Based Segmentation*, Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, June 20–25, 2009, pp. 1295–1302.
- <sup>9</sup> D. Freedman and T. Zhang, in *Interactive Graph Cut Based Segmentation with Shape Priors*, Proceedings of 2005 IEEE Conference on Computer Vision and Pattern Recognition, San Diego, June 20–25, 2005, pp. 755–762.
- <sup>10</sup> P. Kohli, J. Rihan, M. Bray, and P. H. S. Torr, "Simultaneous segmentation and pose estimation of humans using dynamic graph cuts," *Int. J. Comput. Vis.* **79**, 285–298 (2008).
- <sup>11</sup> A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen, "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling," *IEEE Trans. Med. Imaging* **21**, 1151–1166 (2002).
- <sup>12</sup> H. Park, P. Bland, and C. Meyer, "Construction of an abdominal probabilistic atlas and its application in segmentation," *IEEE Trans. Med. Imaging* **22**, 483–492 (2003).
- <sup>13</sup> M. B. Cuadra, C. Pollo, A. Bardera, O. Cuisenaire, J.-G. Villemure, and J.-P. Thiran, "Atlas-based segmentation of pathological MR brain images using a model of lesion growth," *IEEE Trans. Med. Imaging* **23**, 1301–1314 (2004).
- <sup>14</sup> P. L. Bazin and D. L. Pham, "Homeomorphic brain image segmentation with topological and statistical atlases," *Med. Image Anal.* **12**, 616–625 (2008).
- <sup>15</sup> J. H. Noble and B. M. Dawant, "Automatic segmentation of the optic nerves and chiasm in CT and MR using the atlas-navigated optimal medial axis and deformable-model algorithm," *Proc. SPIE* **7259**, 725916 (2009).
- <sup>16</sup> M. Ehm, T. Klinder, R. Kneser, and C. Lorenz, "Automated vertebra identification in CT images," *Proc. SPIE* **7259**, 72590B–72590B (2009).
- <sup>17</sup> S. C. Mitchell, B. P. F. Lelieveldt, R. J. van der Geest, H. G. Bosch, J. H. C. Reiber, and M. Sonka, "Multistage hybrid active appearance model matching: Segmentation of left and right ventricles in cardiac MR images," *IEEE Trans. Med. Imaging* **20**, 415–423 (2001).
- <sup>18</sup> G. Christensen, R. Rabbitt, and M. Miller, "3-D brain mapping using a deformable neuroanatomy," *Phys. Med. Biol.* **39**, 609–618 (1994).
- <sup>19</sup> T. Kaneko, L. Gu, and H. Fujimoto, "Recognition of abdominal organs using 3D mathematical morphology," *Syst. Comput. Japan* **33**, 75–83 (2002).
- <sup>20</sup> C.-C. Lee, P.-C. Chung, and H.-M. Tsai, "Identifying multiple abdominal organs from CT image series using a multimodule contextual neural network and spatial fuzzy rules," *IEEE Trans. Inf. Technol. Biomed.* **7**, 208–217 (2003).
- <sup>21</sup> B. Haas, T. Coradi, M. Scholz, P. Kunz, M. Huber, U. Oppitz, L. Andre, V. Lengkeek, D. Huyskens, A. van Esch, and R. Reddick, "Automatic segmentation of thoracic and pelvic CT images for radiotherapy planning using implicit anatomical knowledge and organ-specific segmentation strategies," *Phys. Med. Biol.* **53**, 1751–1771 (2008).
- <sup>22</sup> P. Campadelli, E. Casiraghi, S. Pratisoli, and G. Lombardi, "Automatic abdominal organ segmentation from CT images," *Electronic Letters on Compute Vision and Image Analysis* **8**(1), 1–14 (2009).
- <sup>23</sup> S. Seifert, A. Barbu, S. K. Zhou, D. Liu, J. Feulner, M. Huber, M. Suehling, A. Cavallaro, and D. Comaniciu, "Hierarchical parsing and semantic navigation of full body CT data," *Proc. SPIE* **7259**, 725902 (2009).
- <sup>24</sup> X. Zhou, T. Hayashi, M. Han, H. Chen, T. Hara, H. Fujita, R. Yokoyama, M. Kanematsu, and H. Hoshi, "Automated segmentation and recognition of the bone structure in non-contrast torso CT images using implicit anatomical knowledge," *Proc. SPIE* **7259**, 72593S–72593S (2009).
- <sup>25</sup> H. Ling, S. K. Zhou, Y. Zheng, B. Georgescu, M. Suehling, and D. Comaniciu, in *Hierarchical, Learning-Based Automatic Liver Segmentation*, Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, June 23–28, 2008, pp. 1–8.
- <sup>26</sup> F. Liu, B. Zhao, P. K. Kijewski, L. Wang, and L. H. Schwartz, "Liver segmentation for CT images using GVF snake," *Med. Phys.* **32**, 3699–3706 (2005).
- <sup>27</sup> A. X. Falcao, J. K. Udupa, and F. K. Miyazawa, "An ultra-fast user-steered image segmentation paradigm: Live wire on the fly," *IEEE Trans. Med. Imaging* **19**, 55–62 (2000).
- <sup>28</sup> X. J. Chen, J. K. Udupa, X. F. Zheng, A. Alavi, and D. Torigian, "Automatic anatomy recognition via multi object oriented active shape models," *Proc. SPIE* **7259**, 72594P–72594P (2009).
- <sup>29</sup> T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models – Their training and application," *Comput. Vis. Image Underst.* **61**, 38–59 (1995).
- <sup>30</sup> J. K. Udupa, V. R. Leblanc, Y. Zhuge, C. Imielinska, H. Schmidt, L. M. Currie, B. E. Hirsch, and J. Woodburn, "A framework for evaluating image segmentation algorithms," *Comput. Med. Imaging Graph.* **30**, 75–87 (2006).
- <sup>31</sup> J. Liu, "Synergistic hybrid image segmentation: Combining image and model-based strategies," Ph.D. thesis, Department of Bioengineering, University of Pennsylvania, 2006.