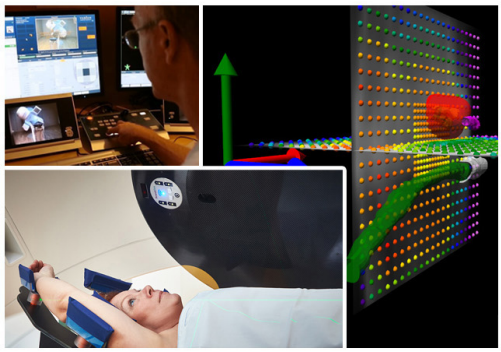
Medical Physics
The International Journal of Medical Physics Research and Practice

Automatic thoracic anatomy segmentation on CT images using hierarchical fuzzy models and registration

Kaiqiong Sun, Jayaram K. Udupa, Dewey Odhner, Yubing Tong, Liming Zhao, and Drew A. Torigian

Citation: Medical Physics **43**, 1487 (2016); doi: 10.1118/1.4942486
View online: http://dx.doi.org/10.1118/1.4942486
View Table of Contents: http://scitation.aip.org/content/aapm/journal/medphys/43/3?ver=pdfcov
Published by the American Association of Physicists in Medicine



Automatic thoracic anatomy segmentation on CT images using hierarchical fuzzy models and registration

Kaiqiong Sun[a)]
School of Mathematics and Computer Science, Wuhan Polytechnic University, Wuhan 430023, China

Jayaram K. Udupa, Dewey Odhner, and Yubing Tong
Medical Image Processing Group, Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Liming Zhao
Medical Image Processing Group, Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104 and Research Center of Intelligent System and Robotics, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Drew A. Torigian
Medical Image Processing Group, Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

(Received 1 July 2015; revised 31 January 2016; accepted for publication 7 February 2016; published 25 February 2016)

Purpose: In an attempt to overcome several hurdles that exist in organ segmentation approaches, the authors previously described a general automatic anatomy recognition (AAR) methodology for segmenting all major organs in multiple body regions body-wide [J. K. Udupa *et al.*, "Body-wide hierarchical fuzzy modeling, recognition, and delineation of anatomy in medical images," Med. Image Anal. **18**(5), 752–771 (2014)]. That approach utilized fuzzy modeling strategies, a hierarchical organization of organs, and divided the segmentation task into a recognition step to localize organs which was then followed by a delineation step to demarcate the boundary of organs. It achieved speed and accuracy without employing image/object registration which is commonly utilized in many reported methods, particularly atlas-based. In this paper, our aim is to study how registration may influence performance of the AAR approach. By tightly coupling the recognition and delineation steps, by performing registration in the hierarchical order of the organs, and through several object-specific refinements, the authors demonstrate that improved accuracy for recognition and delineation can be achieved by judicial use of image/object registration.

Methods: The presented approach consists of three processes: model building, hierarchical recognition, and delineation. Labeled binary images for each organ are registered and aligned into a 3D fuzzy set representing the fuzzy shape model for the organ. The hierarchical relation and mean location relation between different organs are captured in the model. The gray intensity distributions of the corresponding regions of the organ in the original image are also recorded in the model. Following the hierarchical structure and location relation, the fuzzy shape model of different organs is registered to the given target image to achieve object recognition. A fuzzy connectedness delineation method is then employed to obtain the final segmentation result of organs with seed points provided by recognition. The authors assess the performance of this method for both nonsparse (compact blob-like) and sparse (thin tubular) objects in the thorax.

Results: The results of eight thoracic organs on 30 real images are presented. Overall, the delineation accuracy in terms of mean false positive and false negative volume fractions is 0.34% and 4.02%, respectively, for nonsparse objects, and 0.16% and 12.6%, respectively, for sparse objects. The two object groups achieve mean boundary distance relative to ground truth of 1.31 and 2.28 mm, respectively.

Conclusions: The hierarchical structure and location relation integrated into the model provide the initial pose for registration and make the recognition process efficient and robust. The 3D fuzzy model combined with hierarchical affine registration ensures that accurate recognition can be obtained for both nonsparse and sparse organs. Tailoring the registration process for each organ by specialized similarity criteria and updating the organ intensity properties based on refined recognition improve the overall segmentation process. [http://dx.doi.org/10.1118/1.4942486]

Key words: shape modeling, fuzzy models, object recognition, fuzzy connectedness, segmentation, registration

1. INTRODUCTION

Computerized automatic anatomy recognition (AAR) refers to the process of identifying and delineating objects in medical images.[1] AAR becomes essential to make quantitative radiology a reality in routine radiological practice. Quantitative radiology can lead to numerous clinical advances. Due to the imaging process, acquired images often have various shortcomings. Different organs often have overlapping image intensity ranges, and significant gray value variations exist within individual organs and among organs from different patients. Automatic, efficient, and robust segmentation of multitudes of objects on radiological images is challenging.

To overcome the difficulty in segmentation stemming from the above mentioned problems, prior knowledge about individual objects is widely used in current segmentation methods. Anatomy recognition methods, whether or not using explicit object models, have made use of different types of anatomic knowledge and taken different strategies. Major and reliable structures were segmented in Ref. 2, and a region of interest was then determined based on their spatial relationships. A region of interest for the heart was decided by fuzzy constraints expressing anatomical knowledge.[3] The identification of the carina in the trachea was used to detect a set of landmarks for initializing a shape model in Ref. 4. Ribs were detected to initialize a shape model to segment lungs in Ref. 5. Other methods determine the pose of an object model in the image with the intent of subsequently delineating the object.[6–9]

Besides using the object model of a single organ for segmentation,[10,11] multi-object strategies have been studied to improve segmentation, as better constraints can be imposed on the search process by the simultaneous consideration of multiple objects.[12] A competition is set up among objects for delineating their regions/boundaries.[13,14] The inter-relationships among objects are included in the model to influence their localization and delineation.[15,16] Multi-object strategies try to strengthen segmentability by incorporating relevant information in model building, object recognition/localization, and subsequently delineation.[17–23] The dual relationship between objects is integrated in the segmentation algorithm in Ref. 24. Methods have also been proposed that focus on locating objects or their anatomic features in image volumes rather than delineating them.[25–27]

The anatomical knowledge about the shape and spatial layout of organs can also be expressed via labeled images, often called an atlas, which can be subsequently registered to a target image to locate objects in the image. This is a commonly adopted method for organ segmentation. A hierarchical segmentation procedure is formulated using the constructed multi-organ statistical hierarchical atlases.[28] The approach consists of hierarchical recursive processes of initial region extraction using probabilistic atlases and subsequent refinement using multi-level/multi-organ statistical shape models. An atlas-based segmentation approach based on a combination of multiple registration operations was presented in Ref. 29. Labels of the atlas images are propagated to target organ and the propagated labels are combined by spatially varying decision fusion weights. The dependence of different label errors is further considered.[30] In Ref. 31, a general method for multi-organ segmentation of abdominal computed tomography (CT) scans was presented based on a hierarchical atlas registration and weighting scheme that generates target specific priors from an atlas database by combining aspects from multi-atlas registration and patch-based segmentation. The final segmentation is obtained by applying an automatically learned intensity model in a graph-cuts segmentation step, incorporating high-level spatial knowledge.[31]

Owing to the fact that registration cannot handle topological changes in the organ boundaries, there is still a gap in the atlas approaches regarding the relationship between the deformed atlas after registration and the real objects. To overcome this nonconformity, a selection or fusion strategy has been used, although with an increase in the computational load of the recognition process.[29,31] Atlas-based registration and delineation processes are separated in Ref. 32, where atlas-to-subject registration is used to roughly recognize the organ followed by a refined process with fuzzy-connectedness (FC) delineation in its 2D form.

Most of the focus in the literature has been on specific organs/organ systems and not on the same general framework operating on a multitude of organs body-wide, especially for the strategies that determine the pose of an object. Previously, we presented a hierarchical body-wide fuzzy object modeling strategy for AAR.[1] It encoded object relationships in an anatomic hierarchy of objects and demonstrated that the same general method can be used to perform AAR in different body regions and even in different modalities with very good results. In this paper, we combine some of the object-oriented ideas embodied in that approach and bring in registration in the hierarchical setting to achieve improved recognition and delineation. We focus on the thoracic body region. Nomenclature lists the nine target organs and two composite objects. The composite objects are RS (respiratory system = left lung + right lung + trachea and bronchi) and IMS (internal mediastinum = esophagus + pericardium + arterial system + venous system), which are a union of other basic organs considered. The notion of composite objects is useful in combining objects of similar characteristics at a higher level of the hierarchy, which often makes object recognition (and delineation) more effective. For a precise anatomic definition of these organs, refer to Ref. 1 (the Appendix, Table A1). These objects are grouped into a hierarchical tree-like structure,[1] shown in Fig. 1. An object is called parent organ if there is a tree/subtree that is rooted at it. The derivate nodes of the tree are offspring organs.

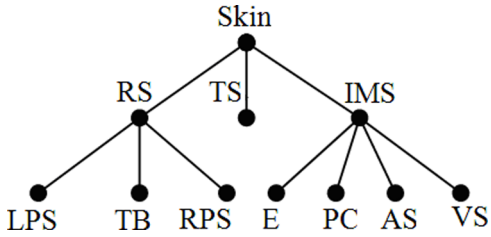


FIG. 1. Hierarchical arrangement of the objects in the thorax considered in this study.

Thin tubular objects will be called sparse objects and include: TB, E, AS, and VS. Compact, blob-like objects will be referred to as nonsparse objects and include: skin, LPS, RPS, and PC. TS is a hybrid between these two types of objects, as it has both types of features.

The gray intensity similarities of different organs, especially if they happen to be adjacent organs, make a recognition process necessary before delineation to construct a complete segmentation process.[33] In this paper, the segmentation operation is divided into recognition and delineation processes. Following the hierarchical structure and location relation of thoracic anatomy defined in Ref. 1, the fuzzy shape model of different organs is registered to the given target image to achieve object recognition. A fuzzy-connected delineation method is then used to obtain the final segmentation result of organs with seed points provided by recognition. An overview of the proposed method is presented in Fig. 2.

The main contributions of the paper are as follows: The hierarchical recognition process is implemented in the registration framework compared to our previous work[1] which did not use registration. Hierarchical registration, which provides initial location of an offspring organ, takes advantage of the gray grouping property of organs and achieves robustness in recognition. The spatial relationship knowledge between different organs, exploited in the recognition process, is easily incorporated in the registration framework through defining exclusion region in search. The shape of an organ varies among different samples, and the degree of variation among different organs is different. To make recognition more tolerant to the variations and improve recognition performance, a similarity measure for registration is defined between the fuzzy model (FM) of the object and target image. Moreover, different weight strategies are used for different organs to strengthen recognition. Both the gray statistics of organs obtained in the modeling process and the gray statistics of organs estimated with the recognized result are made use of to enhance the accuracy of delineation. A brief description of some of the contents of this paper appeared in the SPIE Medical Imaging conference proceedings in 2014.[34] The main enhancements in this paper over the conference paper are more detailed and complete description of all major steps, more extensive evaluation, additional objects considered, and improved delineation results due to better parameter estimation for affinity functions.

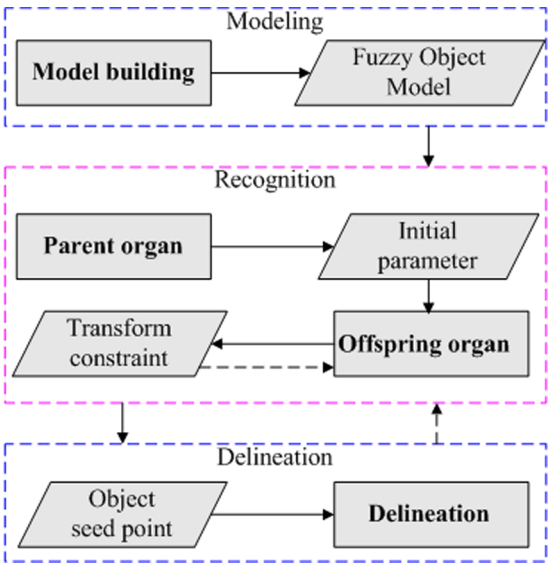


FIG. 2. Overview of the proposed hierarchical registration based segmentation method. The three boxes show the main steps involved in the method, namely, model generation, recognition or object localization, and delineation of the localized objects. Note the interaction between recognition and delineation steps.

The rest of this paper is laid out as follows. The fuzzy model idea[1] is summarized and some refinement is presented in Sec. 2. In Sec. 3, we delineate methods for automatically recognizing objects in target images that employ the hierarchical models. We present fuzzy-connectedness object delineation techniques in Sec. 4. In Sec. 5, experimental results and evaluation of the proposed method are presented, as well as a comparison to methods from the recent literature. Our conclusions are summarized in Sec. 6.

2. MODEL BUILDING

The hierarchical AAR approach[1] consists of the following steps: (1) Collecting image data for a specific population group and delineating objects for the purpose of model building. (2) Building fuzzy anatomy model of objects over the body region. (3) Recognizing objects in a given image I of the thorax. (4) Delineating the recognized objects. In step (2), we explicitly encode object size and positional relationships into the hierarchy and subsequently exploit this information in object recognition in step (3) and delineation in step (4). In the new hierarchical AAR approach, steps (3) and (4) are not separate and independent but are intertwined, with the delineation of an organ higher up in the hierarchical order a prerequisite for recognition of organs lower down in the hierarchical order, as shown with dotted arrow in Fig. 2.

2.A. Collecting image data and delineating objects for modeling

This retrospective study was conducted following approval from the Institutional Review Board at the Hospital of the University of Pennsylvania along with a Health Insurance Portability and Accountability Act waiver. Our goal is to build models of normal thoracic anatomy. Thoracic contrast-enhanced CT images that are radiologically near normal (with the exception of occasional minimal small clinically insignificant abnormalities) were obtained via a search of our hospital patient image database. The nine organs in Nomenclature (excluding RS and IMS) in training sample sets were delineated by using a combination of methods including live wire, iterative live wire,[35] thresholding, and manual painting, tracing and correction. To minimize human labor and to maximize precision and accuracy, algorithms in terms of a proper combination of these methods and the order in which objects are delineated are devised first, all of which operate under human supervision and verification.

2.B. Building fuzzy models of objects

In previous work, we proposed the concept of a fuzzy object model (FOM).[36] The FOM for a body region R and subject group G is a quintuple, $\mathrm{FOM}(R) = [H,M,\rho,\lambda,\eta]$. H here is a hierarchy, represented as a tree, of the objects $O_1,\ldots,O_L$ in R, where L represents the number of objects which are considered for inclusion in model building. Specifically, the hierarchy chosen for thorax is shown in Fig. 1. $M = \{\mathrm{FM}(O_l): 1 \le l \le L\}$ is a set of FM, one model per object. ρ describes the parent-to-offspring relationship in H over G. $\rho = \{\rho_{l,k}: O_l$ is a parent of $O_k,\ 1 \le l,\ k \le L\}$. λ is a set of scale factor ranges $\lambda = \{\lambda_l = [\lambda_l^b,\lambda_l^h] : 1 \le l \le L\}$ indicating the size variation of each object O_l over G. η represents a set of measurements pertaining to the objects in R.

To adapt this approach to the proposed registration framework, the above model is modified in several ways in this paper. In the AAR method,[36] $\mathrm{FM}(O_l)$ is created by aligning all samples of O_l to their mean geometric center, mean scale, and mean orientation (inertia axes) with an exact transformation without requiring searching or optimization, and subsequently averaging the distance transforms and mapping the averaged distance values to a fuzzy membership value through a sigmoid function (see Ref. 1 for details). (In this paper, when we refer to the *location* of an object, whether binary or fuzzy, it refers to the geometric center of the object which is obtained by averaging the coordinates of all its voxels.) Starting from such $\mathrm{FM}(O_l)$, we refine the fuzzy models by seven-parameter affine registration of each training sample to the model and regenerating the model. Figure 3 shows one slice of the refined fuzzy models of several organs created in this manner. At each voxel, the gray intensity denotes the membership of belonging to the organ. When a voxel is covered by all aligned samples, its gray value is K, or maximum membership or gray intensity of image. When a voxel is not covered by any aligned samples, its grayness is zero.

Moreover, the original CT image gray value statistics (mean m_ϕ, standard deviation σ_ϕ, and homogeneity σ_ψ) of each organ are also recorded in the model. These parameters are related to object recognition (Sec. 3) and delineation (Sec. 4). The values of these parameters are considered part of the description of η in the FOM quintuple for thorax. When used in seed point specification and delineation, the above mean gray intensity associated with some organs will be estimated again or refined, as described in Sec. 4, from the target image based on the recognition result. Besides the hierarchical relationship from parent to offspring in Fig. 1, some relative location relationship between different organs is also recorded in the model. TB is located between RPS and LPS in a left-to-right orientation. PC is located between RPS and LPS more anteriorly in a left-to-right orientation. These relations can be used to narrow the search range in recognizing organs, provided that related organs have been recognized, easily.

The fuzzy object model $\mathrm{FOM}(R)$ output at the end of the model building step is used in performing AAR on any target image I as described in Secs. 3 and 4.

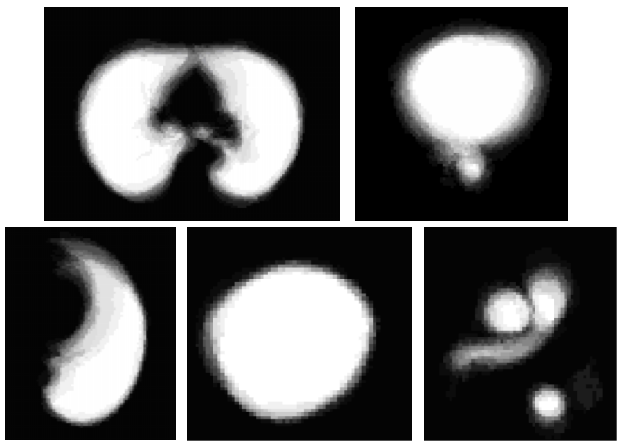


FIG. 3. Axial slices through different objects from the refined fuzzy model for: respiratory system, internal mediastinum (first row), left lung, pericardial region, and arterial system (second row). The gray intensity denotes the fuzzy membership of belonging to the organ, where brighter intensity means a higher membership.

3. HIERARCHICAL RECOGNITION

The goal of recognition in AAR is to output the pose of $\mathrm{FM}(O_l)$, or equivalently the pose-adjusted fuzzy model $\mathrm{FM}^T(O_l)$, for each O_l in a given target image I. In particular, the recognition step places the shape model of an organ, such as shown in Fig. 3, at the proper location in the target image after making a proper scale and rotation transformation. In the framework of this paper, the recognition step is realized through registration. The result can be represented as a transformation T^* with translation t, rotation θ, and scale s. The registration optimization process is given by

$$T^* = \underset{\{T(t,\theta,s)\}}{\arg\max} \left\{\mathrm{SI}\left(\mathrm{FM}^T(O_l), Id_l\right)\right\}, \quad \text{constrained by } i(T),\ c(T). \tag{1}$$

Here SI denotes the similarity between the transformed fuzzy object model, $\mathrm{FM}^T(O_l)$, and a transformed version Id_l (see below) of the target image I. Seven-parameter affine transform T (translation, rotation, and scale) is used for registration. Variables $i(.)$ and $c(.)$ are the initial condition and constraints for the transform variables. The algorithm of unconstrained optimization by quadratic approximation[37] is used for finding the optimal solution. Without proper constraints and initial condition, such a search for the solution is difficult, error-prone, and time-consuming for a large search space and due to the existence of many local optima. The hierarchical recognition method is aimed at rapidly locating the parent organ by making use of the gray grouping information of some organs and to provide initial parameters for the search of the offspring organ. Also, previously recognized and delineated organs provide constraints for the other organs by making use of the spatial relation. In the rest of this section, the transformed image Id_l and the similarity function SI in Eq. (1) are defined followed by the description of the hierarchical recognition process.

3.A. Likelihood map of organ

For each organ O_l, the similarity in Eq. (1) is defined between the transformed shape model $\mathrm{FM}^T(O_l)$ and the

likelihood image Id_l of O_l, instead of the original target image I. By making use of the mean and standard deviation, m_ϕ^l and σ_ϕ^l, of O_l in the original CT image, the likelihood image $Id_l(x)$, where x denotes a voxel, is constructed from $I(x)$ as

$$Id_l(x) = \begin{cases} K, & \text{if } I(x) \in \left[m_\phi^l - \alpha \cdot \sigma_\phi^l, m_\phi^l + \alpha \cdot \sigma_\phi^l\right] \\ K \cdot e^{-(I(x)-m_\phi^l)^2/2(\sigma_\phi^l)^2}, & \text{otherwise} \end{cases}, \quad (2)$$

where $\alpha = 2$ was used for all organs except $\alpha = 0$ for E. A small value of α is used for E to avoid the search algorithm running into a local optimum solution. K is the maximum gray intensity in the resulting likelihood image. The recognition step is thus implemented by registering the fuzzy model $\mathrm{FM}(O_l)$ to Id_l for each organ O_l. In other words, the fuzzy model is matched to the best evidence presented by I in terms of intensity for O_l. Figure 4 shows one slice of a target image I and an example of Id_l for RS, IMS, LPS, PC, and AS. It can be seen that the corresponding object region has high gray intensity in the likelihood image. However, there is also high gray intensity elsewhere due to gray-level overlap of different organs, which hampers the registration process for recognition.

3.B. Similarity measure for registration

The object region of an organ has high intensity in the likelihood image, but it does not have homogeneous gray intensity. On the other hand, the fuzzy shape model of an object (Fig. 3) integrates the shape of the object from many samples in the training set. It also has inhomogeneous gray intensity. From the modeling process, the highest intensity region in the shape model represents the most stable shape part of the organ, which should take greater role in recognition than other parts. Therefore, a similarity metric for registration is defined, with the consideration of the above observations, as the negative weighted sum of squared difference (WSSD) between Id_l and K,

$$\mathrm{SI}\left(\mathrm{FM}^T(O_l), Id_l\right) = -\sum_x w\left(\mathrm{FM}^T(O_l)(x)\right)[Id_l(x) - K]^2, \quad (3)$$

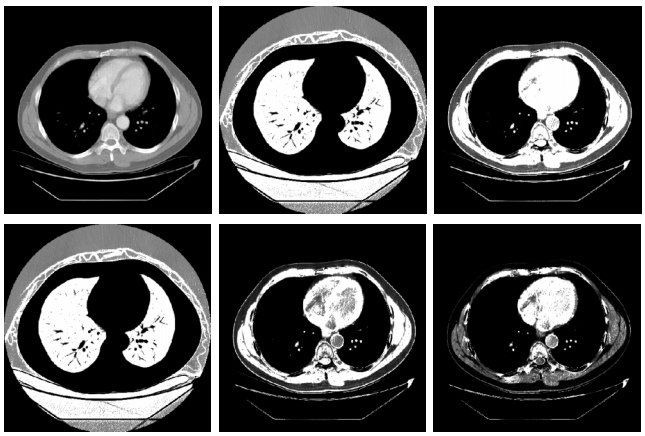


FIG. 4. One axial slice of a target CT image in the thorax (first row, upper left) and the corresponding likelihood images for respiratory system, internal mediastinum (first row from left to right), left lung, pericardial region, and arterial system (second row from left to right). The gray intensity denotes the likelihood of the pixel belonging to the organ.

where the weight w chosen is a function of the transformed (adjusted) model membership value $\mathrm{FM}^T(O_l)$. The manner in which w is chosen for each organ is described at the end of this section. Note that when similarity is maximal, the sum of weighted square of difference $Id_l(x) - K$ of all voxels in the domain of the transformed shape model $\mathrm{FM}^T(O_l)$ becomes minimum. It means that the transformed shape model and the corresponding data likelihood region appear most similar.

3.C. Hierarchical recognition process

A composite object can be more easily and robustly recognized than the individual objects. This is because the gray value range of the composite object is larger and this has far less gray value overlap with other nearby structures compared to single objects. The initial solution for the offspring organ from the recognized result of its parent will be very near to the desired solution sought by registration. Furthermore, the constraints or forbidden regions from previously recognized or delineated organs can narrow the solution space for other organs yet to be recognized and delineated.

With the location and scale of a parent organ known in the hierarchy, an estimation of the locations of all descendant objects in the tree is obtained through the object relationship stored in ρ in the hierarchical model as described in Ref. 1,

$$L_o^c = L_p^c + \left(L_o^m - L_p^m\right) \cdot S_p^c / S_p^m, \quad (4)$$

where L_p^c and S_p^c denote the already computed location and scale of parent organ in the target image. L_p^m and S_p^m denote the mean location and scale of parent organ saved in the model. L_o^m and L_o^c are the mean location in the model and the location to be estimated in the target image for the offspring organ.

In our hierarchical recognition process, the skin object is first segmented by thresholding combined with a morphology filter. With the geometric center and scale computed from the segmented skin, an estimation of the locations of all descendant objects is obtained through Eq. (4). Based on the initial location estimate from skin, and orientation and scale described below, objects RS, IMS, and TS are then recognized and subsequently also their offspring following the hierarchy.

With the skin object segmented, the intensity of the complementary (background) region within skin in the likelihood image Id of RS is set to 0. That is, these regions within skin have low belongingness to RS. The initial orientation and scale of RS are just those computed from segmented skin. The similarity for registration of RS is the negative sum of squared differences (SSD) defined between $\mathrm{FM}^T(O)$ of RS and the modified data likelihood of RS. With the above setting, RS is recognized. IMS is then recognized with the initial location estimated for skin and initial scale and orientation inherited from recognized RS. The similarity for registration to recognize IMS is negative WSSD defined in Eq. (3). It means that only a portion of voxels in the model $\mathrm{FM}^T(O)$ of IMS contributes to the computation of similarity, whereas the SSD used for RS is not appropriate for the completely different background in likelihood image and the $\mathrm{FM}(O)$ of IMS.

For the recognition of LPS and RPS, the initial location is estimated by Eq. (4) with the recognized result of RS. The initial scale and orientation are those inherited from RS. The similarity is negative SSD as for RS.

Before the recognition of TB, objects LPS and RPS are fully delineated using the method described in Sec. 4. With the segmented result of LPS and RPS, their regions are set as forbidden regions for TB, by setting the corresponding voxels in Id of TB to a very high value compared to K. The similarity for the registration to recognize TB is negative WSSD [Eq. (3)]. Also, the search range is confined between LPS and RPS, in the anatomic left-to-right direction.

PC and AS are recognized in the same way after the recognition of IMS. The initial locations are estimated by Eq. (4) with the recognized result of IMS. Their initial scales and orientations are inherited from those of IMS. The similarity for PC and AS is negative WSSD. The confined region that has smaller coordinate in the anterior–posterior direction than RS is imposed on the location search for PC.

Before the recognition of E, TB is delineated. Subsequently a threshold is applied to the target image and a possible air region is obtained by subtracting the union of LPS, RPS, and TB from the thresholded image. The data values of the voxels in the possible air region are set to K, which means highest membership in E. The recognition of E starts from the initial pose provided by the recognition result of IMS. The similarity is negative WSSD.

When negative WSSD is used as the similarity function for registration, the weight function $w(.)$ in Eq. (3) is specified as follows: For TB and E, $w(.)$ is $\mathrm{FM}^T(O)/K$. For IMS, PC, and AS, $w(.)$ is $\mathrm{FM}^T(O)/K$, if $\mathrm{FM}^T(O)$ is greater than $K/2$, and $w(.)$ is 0 otherwise. For TS, $w(.)$ is $\mathrm{FM}^T(O)/K$, if $\mathrm{FM}^T(O)$ is greater than $2K/3$, and $w(.)$ is 0 otherwise. The value of $\mathrm{FM}^T(O)$ at any point is between 0 and 1 mathematically. But in the implementation, its range is from 0 to K, which leads to $\mathrm{FM}^T(O)/K$ in [0, 1]. For all registration tasks that use negative WSSD as similarity, there is a scale range constraint if an initial scale is provided. This constraint confines the result of scale in 0.85–1.15 times the initial scale as used in Ref. 1. The interval is a relative scale, multiple of the mean scale size, in which nearly all samples fall into this range of mean scale. Even if some organ in the target image has a scale out of this range, it will not influence the result significantly. This is due to the fact that based on the recognition result there are also strict gray intensity requirements for the seed points as described later.

From the above settings for $w(.)$, it can be seen that the nonzero/zero region of this weight function comes from a threshold of $\mathrm{FM}^T(O_l)$. It leads to only a portion of the model contributing to the computation of the similarity and different weights are associated with them. Thus, in the modeling process, if the shape of one or several samples deviates from other samples greatly, they will contribute less to the shape model of the organ, and consequently will have less effect on recognition. The definition of similarity has more tolerance to the variation of organs in the target image and is used in the recognition of some organs which have high shape variation among different subjects. Specifically, for the highly sparse nature of the TS object, a high threshold on $\mathrm{FM}^T(O)$ of TS is used to compute the weight $w(.)$ aiming to recognize the vertebral part of the TS object.

4. DELINEATION OF ORGANS

Once the recognition process is completed and the adjusted model $\mathrm{FM}^T(O)$ is output for a given image I for an organ O, organ delineation is performed on I using the method and object order presented below.

In our system, object delineation is carried out via an adaptation of the iterative relative fuzzy connectedness (IRFC) algorithm.[38] Adaptation consists of using the fuzzy model at recognition to devise ways of automatically finding seed voxels in the object and the background tissue components and forming affinity functions (needed by IRFC) by taking into account the model constraints.

The FC framework[38,39] is graph-based. An ordered graph (C, α) is associated with the given image $I = (C, f)$, where C is the set of all voxels of I, $f(x)$ is the image intensity at voxel x, and α is an adjacency relation on C such as 6-, 18-, or 26-adjacency of voxels. Each ordered pair (c, d) of adjacent voxels in α is assigned an affinity value $\kappa(c,d)$ which constitutes the weight assigned to arc (c, d) in the graph. To each path π in the graph (or equivalently in I) in the set of all possible paths $\Pi_{a,b}$ between two voxels a and b of C, a strength of connectedness $Q(\pi)$ is determined, which is the minimum of the affinities along the path. The connectivity measure $Q^*(a,b)$ between a and b is then defined to be $Q^*(a,b) = \max\{Q(\pi) : \pi \in \Pi_{a,b}\}$. The notion of connectivity measure can be generalized to the case of "between a set A and voxel b" by a slight modification: $Q^*(A,b) = \max\{Q(\pi)\colon \pi \in \Pi_{a,b}$ and $a \in A\}$.

By using a fast algorithm to compute $Q^*(A,b)$, the machinery of FC allows a variety of approaches to define and compute "objects" in images by specifying appropriate affinity functions and seed sets. In particular, in IRFC, two seed sets A_O and A_B are indicated for an object O and its background B, respectively. Then the object indicated by A_O is separated *optimally* from the background indicated by A_B by an iterative competition in connectivity measure between A_O and every voxel $c \in C$ and between A_B and c. In IRFC, A_O and A_B are specified usually with human interaction. Below, we address the affinity function and seed specification based on the recognition result from Sec. 3.

4.A. Affinity function

In this paper, affinities $\kappa_O(c,d)$ and $\kappa_B(c,d)$ for O and B are designed separately. Each of κ_O and κ_B has three components. The description below is for κ_O. The same applies to κ_B,

$$\kappa_O(c,d) = \omega_1\psi_O(c,d) + \omega_2\phi_O(c,d) + \omega_3\gamma_O(c,d). \quad (5)$$

Here, $\psi_O(c,d)$ represents a homogeneity component of affinity, which maps the gray intensity difference between c and d into an affinity value in [0, 1]. It means that the more similar image intensities $f(c)$ and $f(d)$ are at c and d, the greater is this component of affinity between c and d. As commonly done

in the FC literature, we set

$$\psi_O(c,d) = \exp\left[-\left((f(c)-f(d))/\sigma_\psi^o\right)^2/2\right]. \quad (6)$$

$\phi_O(c,d)$, the object feature component, on the other hand, describes the "degree of nearness" of the intensities at c and d to the intensity m_ϕ^O expected for the object O under consideration. This nearness is expressed by

$$\phi_O(c,d) = \exp\left[-\left(\left(\max\{f(c),f(d)\} - m_\phi^O\right)/\sigma_\phi^O\right)^2/2\right]. \quad (7)$$

The third component $\gamma_O(c,d)$ incorporates fuzzy model information into affinity depending on the distance of c or d to the support domain of the transformed fuzzy model of the object and the two fuzzy model membership values $\mu_O(c)$ and $\mu_O(d)$ at c and d for the object. This component has high value when c and d both have high membership or both have low membership in the registration aligned object model for the corresponding foreground tissue (far away from the edge of the support domain of the transformed fuzzy model of object).

Usually there are several organs forming the background for an organ. Thus, for each organ in the background, its object-based affinity component can be computed by the mean and standard deviation with Eq. (7). A combined single object-based affinity for background is constructed by taking the maximum over all background objects.

The weights for the three components in Eq. (5) are chosen differently for each organ (they add up to 1), as described in Sec. 5. The homogeneity parameter is set equal for object and background $\left(\sigma_\psi^O = \sigma_\psi^B\right)$ and estimated from uniform regions in the training images (after leaving out high gradient regions), as commonly done in the FC literature.[14] The remaining parameters (m_ϕ^O, σ_ϕ^O, m_ϕ^B, σ_ϕ^B) are estimated automatically from the training data sets from knowledge of O and B regions for each object. All these parameters are estimated during the model building process and recorded in the element η of the fuzzy object model.

4.B. Parameter estimation for affinity computation

Since significant gray value variations exist within individual organs and amongst organs from different patients, some parameters obtained from the training set in the modeling process may deviate significantly from the real gray characters of the organ in the target image, especially for organs that have small gray intensity range. We estimate the parameters m_ϕ and σ_ϕ for some organs based on the recognition result.

The updated parameters m'_ϕ and σ'_ϕ of an organ O_l are computed from the gray intensities in the target image on the voxels covered by a threshold result of $\mathrm{FM}^T(O_l)$. For different shapes of organs, the threshold result of $\mathrm{FM}^T(O_l)$ may have different overlap in the expected object. Thus, different strategies are adopted to update the gray value statistics of organs. Specifically, for PC, the sample set of gray intensities is fitted with the sum of two Gaussian distributions and two sets of mean and standard deviation are obtained to substitute the previous ones from training samples. Thus two object-based affinities are computed and the maximum of the two affinities is used as the object-based affinity κ_O for PC. For AS, the sample set of gray intensities is fitted with the sum of two Gaussian distributions. The maximum of the means of the two distributions is used to substitute the mean from training samples. For E, we model the sample set of gray intensities as one Gaussian distribution and the mean of the distribution is used to substitute the mean from training samples.

These updated gray value statistics of organs are used in parameter definition required to compute the object-based affinity in Eq. (7) as well as in seed point specification described below.

4.C. Seed specification

Seed sets A_O and A_B are found by a joint criterion of a threshold for image intensity and for model membership for each of O and B. The threshold interval Th_O for O is a small interval around m_ϕ^O. This will be denoted by Th_l for object O_l. The threshold interval Th_B for background is a union of similar threshold intervals for the background objects. (In principle, all objects other than O can be considered to be background objects of O; however, in practice, only the anatomic neighbors of O matter.) The only new parameters are Th_O^M and Th_B^M used as model thresholds for indicating A_O and A_B, respectively. Thus,

$$A_O = \{v \in C : f(v) \in Th_O \text{ and } \mu_O(v) \in Th_O^M\}, \quad (8)$$

$$A_B = \{v \in C : f(v) \in Th_B \text{ and } \mu_B(v) \in Th_B^M\}. \quad (9)$$

In our implementation, Th_O^M is fixed at $[K/2, K]$ and Th_B^M is set to [0, 0]. For some organs, the definition of the seed for background (based on gray criterion for both target image and model) can be substituted by a threshold result on the distance function of the support domain of the model, i.e., an erosion result of the support domain of the model.

In summary, the main parameters needed for each object are as follows. For recognition, besides the gray value mean and standard deviation obtained from model building in Sec. 2, the parameter α is needed in Eq. (2) to compute the likelihood map of organ and the weight function $w(.)$ is needed in Eq. (3) to compute the similarity. For the optimization problem in Eq. (1), the search range and step are needed to be set. For delineation, the weights for the three components in Eq. (5) are needed to be set and the threshold intervals for gray value and fuzzy model in Eqs. (8) and (9) are needed. To segment skin, threshold and scale of filter are needed, as described in Sec. 5.

5. RESULTS

The image data sets (same as those used in Ref. 1) are selected from our hospital patient image database by a board-certified radiologist (Torigian). They were verified to be of acceptable quality and radiologically normal, with exception of minimal incidental focal abnormalities. They consist of contrast-enhanced breath-hold chest CT studies acquired during full inspiration from fifty subjects aged from 50 to 60 yr. The images were of size $512 \times 512 \times 51$–69 with 12 bits of

gray-level resolution and voxel size of $0.9\times0.9\times5$ mm^3. We have tested the proposed method on organs listed in Nomenclature by using 20 samples for building the models and 30 for testing. In all data sets, any extra slices falling outside the body region as per definition are removed manually first. Ground truth segmentations of all organs in all images were created as previously explained.

A composite object is recognized only to provide initial parameters for the recognition of its offspring organs. Thus, delineation methods described in Sec. 4 are not needed and have not been implemented for composite objects. Since VS has wide image intensity variation and strong overlap with AS and TS, and since it is adjacent to the latter, acceptable recognition could not be obtained. Hence recognition and delineation results are not reported for VS. We are not aware of any published paper that successfully addressed this issue except our own recent work using a different strategy expressly designed to handle sparse objects.[40]

Thorax region is considered to extend axially from 5 mm below the base of the lungs to 15 mm above the apex of the lungs (arms are not included in this study).[1] Similarly, each object included in the thorax is defined precisely irrespective of whether it is open-ended, because it straddles body regions (for example, esophagus), or closed and contained within the thorax but is contiguous with other objects.[1]

The recognition result is the transformed fuzzy object model $\mathrm{FM}^T(O_l)$. A special threshold $K/2$ is used on it and a binary image is obtained, the location and orientation of which are used to compute the recognition error. The location error is the distance between the centers of the above binary object and the ground truth object, the ground truth object obtained by the same manual method described in the section of model building. To characterize delineation accuracy, the following two independent measures are defined: false positive volume fraction (FPVF) and false negative volume fraction (FNVF). In addition, we report mean Hausdorff distance (HD) between the true and delineated boundary surfaces. The HD measure is defined as the mean over all test subjects of the median of the distances of the points on the delineated object boundary surface from the true object boundary surface.

5.A. Segmentation of skin and location estimation and recognition of its offspring organs RS, IMS, and TS

For all target images, binarization with a threshold of 700 (over the CT image gray scale of 0–4095 or −300 when expressed in Hounsfield Units) produces the rough skin object, shown on the left in Fig. 5. Then morphological opening with structuring element of size 7×7 is implemented on each slice to remove the small noisy components. A morphological closing with a linear structuring element of length 180 pixels followed by reconstruction is used to fill the holes in the rough skin object. This structuring element is chosen to be comprised of the background region of skin, and not be comprised of the region of LPS and RPS. Subsequently morphology opening with a disk structuring element of radius 7 pixels followed by reconstruction is used to delete the noisy segments disconnected to the main part in the image. The middle of Fig. 5 shows an example of final skin object delineation where the HD distance error is 0.43 mm. For the accuracy of segmentation for skin and other objects on all the 30 test samples, see Table III.

With the segmented skin, its volume and location can be obtained. In the model, the normalized volume of skin and location relationship between the offspring organs and skin is recorded. Thus an estimation of the locations of all descendant objects in the hierarchy is obtained through aligning the location of skin in the model to the segmented skin in the target image. The resize scale is decided by the ratio of the current segmented skin volume to the normalized volume of skin in the model. The location and scale alignment is expressed by Eq. (4). The mean location estimation errors at this stage for RS, IMS, and TS over all 30 target images are shown in Fig. 6(a), with mean being 10, 14.6, and 9.3 mm respectively, or roughly 2–3 voxels [noting that the voxel size is limited by its greatest dimension (5 mm) and typically errors occur in the direction orthogonal to the slice planes].

The maximum image gray value K is usually 4095 which is used to define the similarity in Eq. (3) and the threshold for seed specification. To implement the optimization solution of Eq. (1) for registration, the relative search range for the transform parameters is set to 20 voxels and the search

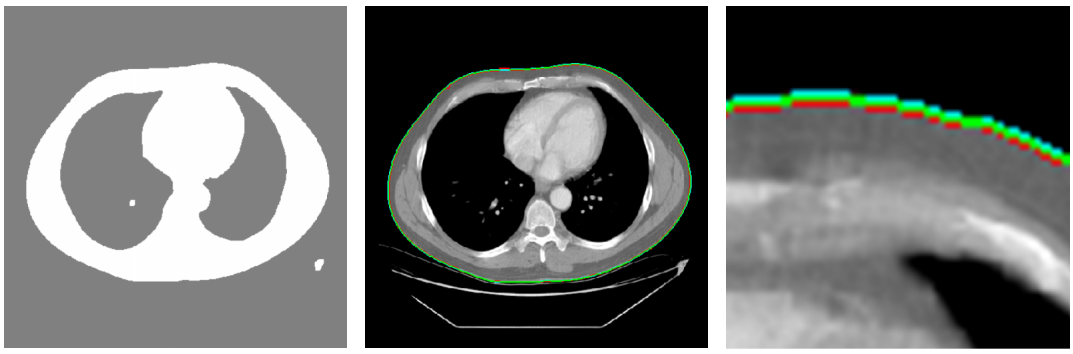


FIG. 5. One axial slice showing rough skin (left), skin results overlaid on target CT image in the thorax (middle), and an enlarged patch of the middle image selected from its top–right portion. In the middle and right images, the ground truth is shown in red, and delineation result is shown in blue. Green color denotes the overlapping of blue and red. Very small amount of red and blue points in the image denotes the strong agreement between ground truth and segmented result (see color version online).

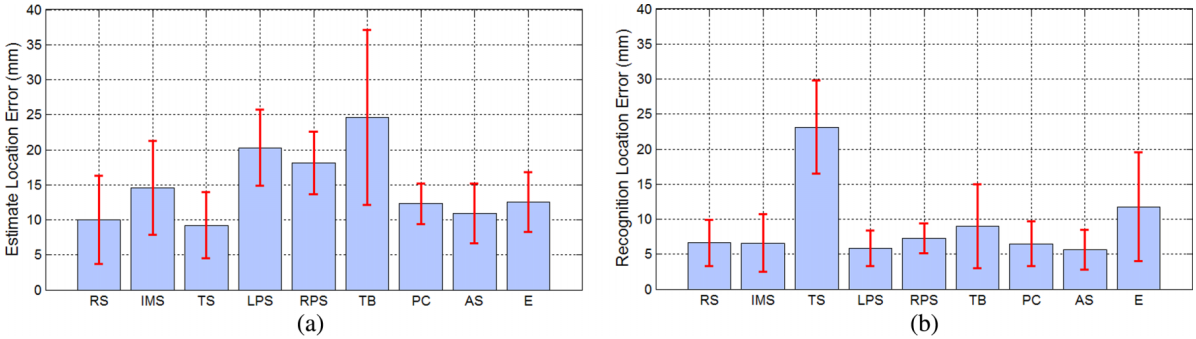


FIG. 6. (a) Location error stemming from placement of objects directly relative to their parent after the parent has been recognized/delineated. (b) Final location error.

step is 1, 1, and 0.01 for translation, rotation, and scale, respectively. Since good initial solution is provided based on hierarchical recognition, multi-resolution strategy is not needed in registration, which decreases the computational cost considerably.

The mean recognition errors over all 30 target images after registration are shown in Fig. 6(b). It shows a distinct decrease of location error compared to the error stemming from locations estimated from the skin object, as shown in Fig. 6(a). Since the recognition of TS is aimed at the vertebral parts instead of the whole TS, slightly increased recognition error in (b) over (a) is obtained for TS. However, TS recognition is stable and its accuracy is enough to produce object seeds for TS and hence delineation correctly.

Figure 7 illustrates the recognition result for RS and IMS of a sample with location error being 3.99 and 6.9 mm, respectively, on two slices, where it can be seen that the offspring organs of RS and IMS are placed very close to the corresponding organ in the target image, providing excellent initial location for their recognition. The first row of Fig. 8 illustrates sample recognition results for TS on two slices. Note how the vertebral positions are accurate but the ribs are not since the weight in WSSD for recognition specified in the end of Sec. 3.C leads to the domination of vertebral part in the recognition.

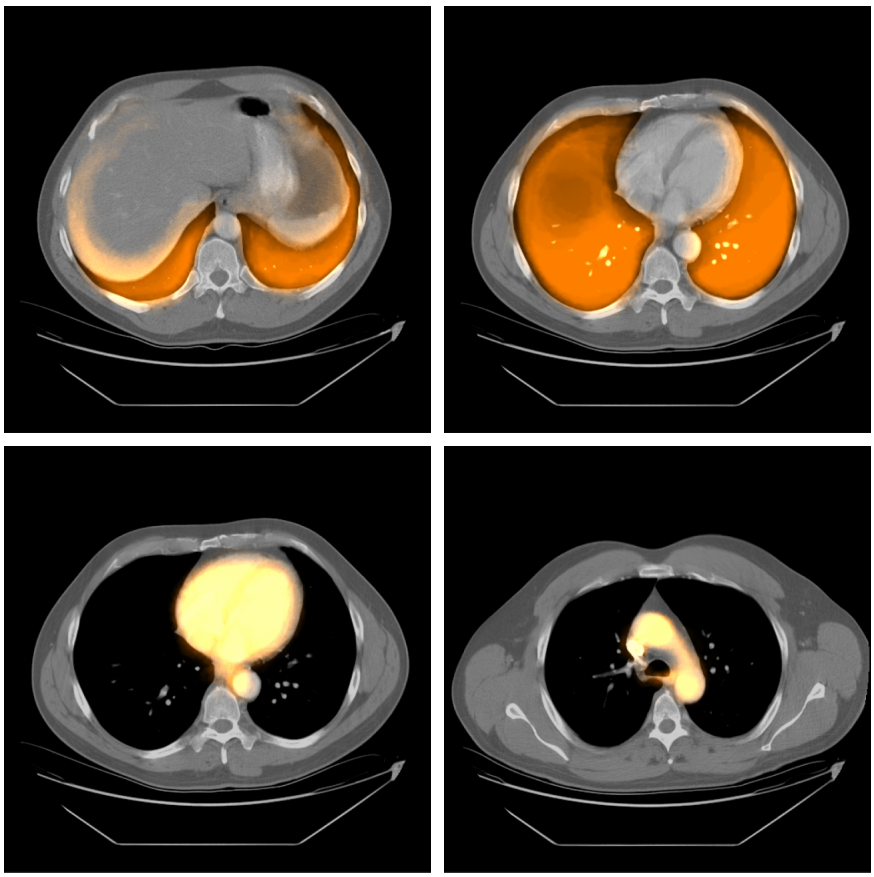


FIG. 7. Two axial slices of recognized results of object RS (respiratory system) (first row) and IMS (internal mediastinum) (second row) overlaid on an original CT image.

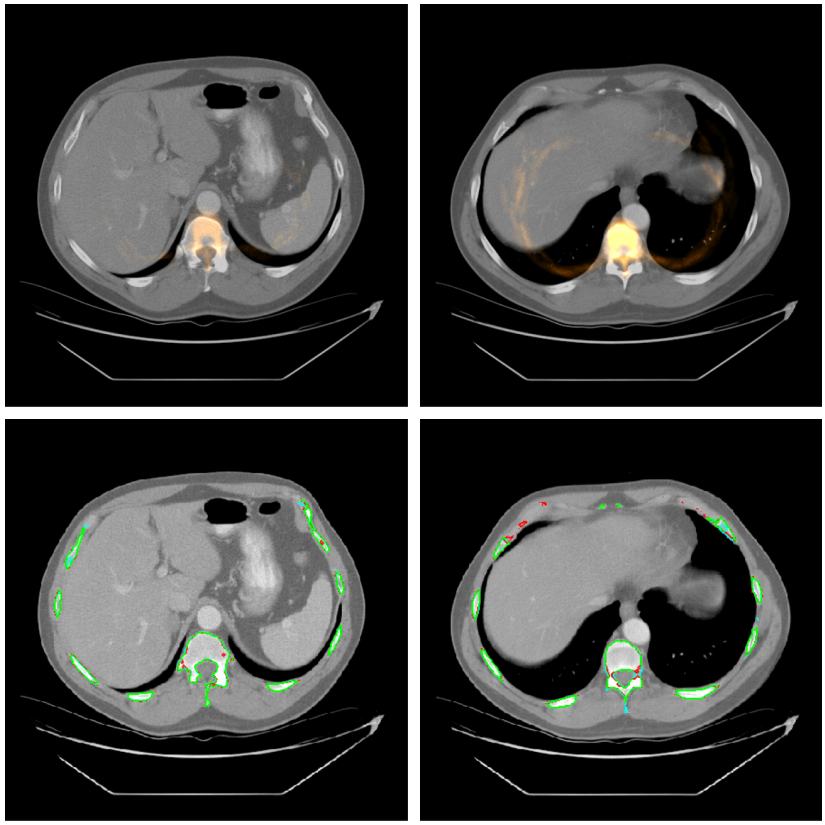


FIG. 8. Two axial slices illustrating recognition (top) and delineation (bottom) results for thoracic skeleton on CT images.

5.B. Recognition of offspring organs of RS and IMS

The mean location errors for the offspring organs of RS and IMS over the 30 target images are shown in Fig. 6(a), with mean being roughly 2–5 voxels. The mean location recognition error for these organs is shown in Fig. 6(b). Overall, the mean of initial location error and final recognition error of all listed organs over all test samples is 14.7 and 7.38 mm (recognition error excludes TS), respectively. The first row of Figs. 9 and 10 shows sample recognition results in two slices overlaid on original image for LPS, RPS, TB and PC, AS, and E, respectively.

Since E does not have any distinct shape difference along the z-axis, it is extremely difficult to recognize its accurate

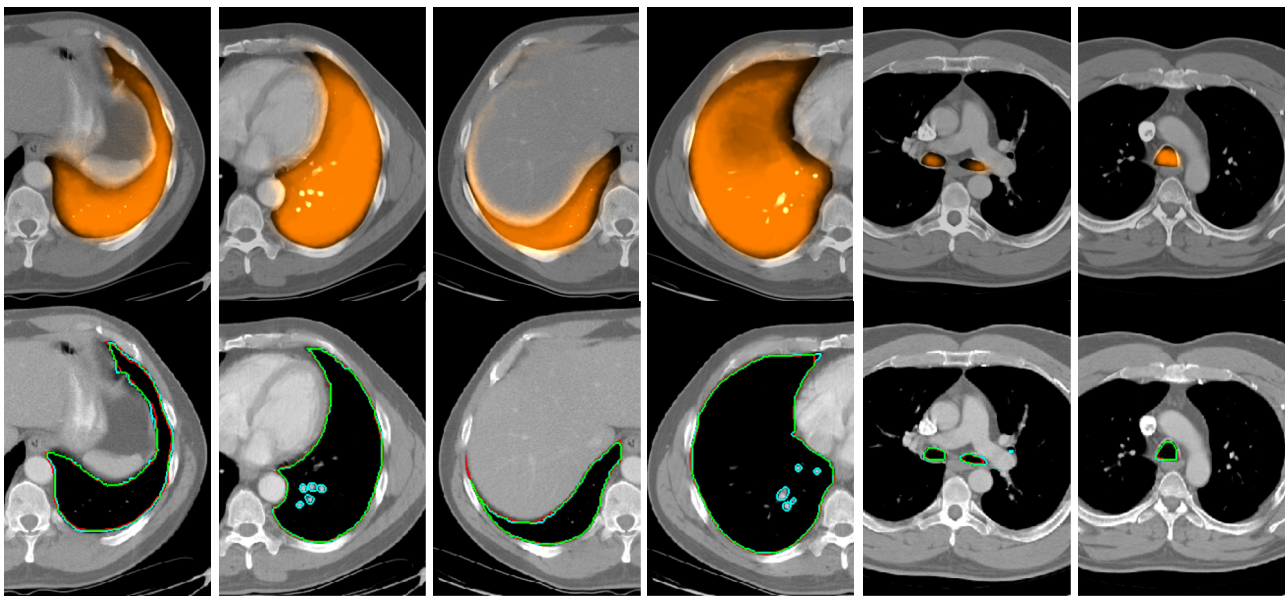


FIG. 9. Two axial slices showing recognition (first row) and delineation (second row) results for left lung (first two columns), right lung (middle two columns), and trachea and bronchi (last two columns) on thoracic CT images.

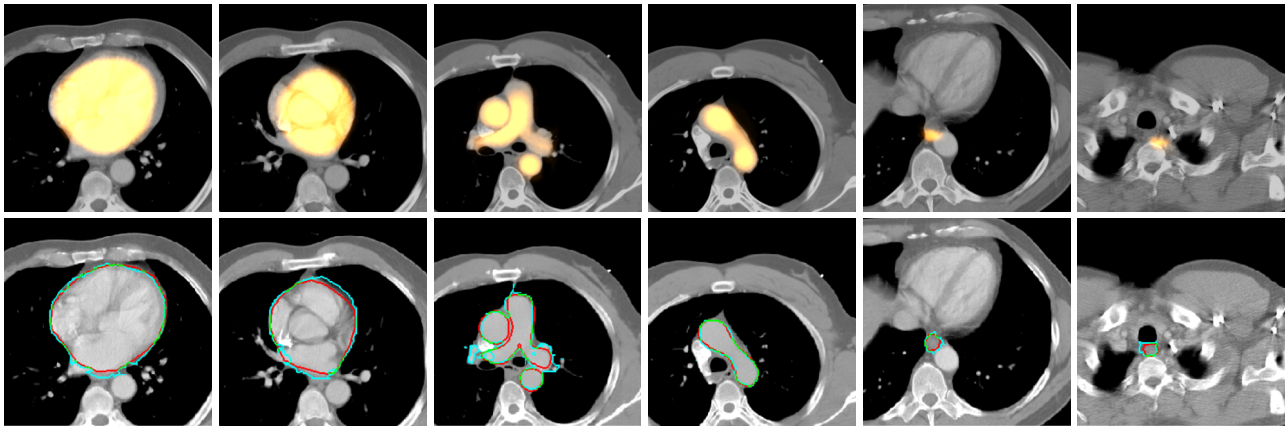


FIG. 10. Two axial slices showing recognition (first row) and delineation (second row) results for pericardial region (first two columns), arterial system (middle two columns), and esophagus (last two columns) on thoracic CT images.

location and orientation. A mean recognition error of 11.7 mm is obtained, see Fig. 6(b), which we believe is outstanding and is due to the intricate method of making use of anatomic information in various stages in a hierarchical manner. No effective methods have been reported in the literature for localizing and delineating esophagus on CT images. When the location error is computed only based on the error along x- and y-axis orientation, the mean error reduces to 4.03 mm with standard deviation 2.71 mm. Thus, the recognition result may provide proper seeds on only part of the slices with the methods in Sec. 4.C. In this paper, the seed specification method is implemented only in slices within a certain interval for all test images.

5.C. Delineation of offspring organs of RS and IMS

PC and AS are spatially contiguous and have strong overlap in gray value distributions and therefore it is difficult to segment them only based on object and homogeneity based affinity. Model based affinity brings prior information into the delineation process after effective recognition. Specifically, for AS, model based affinity is defined only on the estimated region of PC from the recognition result of PC, and in this manner, IRFC sets up a competition in terms of connectivity strength between PC and AS. There is strong overlap of gray intensity between E and its surrounding structures. Thus model based affinity is also needed for E. The delineation of other organs does not need the model component since the hierarchical recognition itself handles effectively the recognition of foreground and background tissues for effective seed placement.

TABLE I. Weight parameter setting for computation of affinity for different objects.

Object	Homogeneity ω_1	Object ω_2	Model ω_3
LPS, RPS, TB, TS	0.5	0.5	0
PC	0.25	0.25	0.5
AS	0.8 (0.5)	0.2 (0)	0 (0.5)
E	0.9	0	0.1

Table I summarizes the weights for homogeneity, object, and model based components in Eq. (5) for each organ. Two groups of weight parameters are used for AS, with and without model based affinity.

Table II lists the background organs for each organ. From the gray statistics of the organs, the object based component of affinity for object and background is computed with Eq. (7). A prime is added to the variable for some organs to indicate that the gray statistics for them are the updated values estimated in Sec. 4.B. Others are from the modeling process. Thoracic adipose tissue (TAT) is not an organ listed in Nomenclature. However, its gray statistics are needed, and hence used, in delineation.

The mean intensity of TB from the training set is lower than that of LPS and RPS, although in some it is greater. Therefore, its statistics are used in the computation of object-based component of affinity for the background of TB to prevent leakage into LPS or RPS. The model based affinity can also be used to prevent leakage.

Delineation results over the tested 30 subjects in terms of FNVF, FPVF, and HD for the eight objects are listed in Table III.

Because of the strong intensity overlap between E and the surrounding structures, pure intensity based affinity cannot separate E from the background adequately. If only homogeneity based affinity is used, the three measures reduce to 0.10, 0.003, and 10.3 mm. Large FNVF leads to large distance errors. Therefore, model based affinity is adopted to further confine the object and background. Since localization along

TABLE II. Anatomic background objects considered for different objects. TAT denotes thoracic adipose tissue and is used only as a background object.

Object	Background
LPS	RPS, IMS, TS, TAT
RPS	LPS, IMS, TS, TAT
TB	TB, IMS, TAT
TS	PC′, AS′, RS, TAT
PC	AS′, RS, TAT
AS	AS′, RS, TS, TAT

TABLE III. Mean ± standard deviation of delineation accuracy for the different objects.

Object	FNVF	FPVF	HD (mm)
Skin	0.00 ± 0.00	0.01 ± 0.01	0.7 ± 0.47
LPS	0.04 ± 0.02	0.00 ± 0.00	0.8 ± 0.39
RPS	0.03 ± 0.01	0.00 ± 0.00	0.85 ± 0.39
PC	0.09 ± 0.08	0.00 ± 0.00	2.89 ± 0.82
TB	0.06 ± 0.07	0.00 ± 0.00	0.85 ± 0.37
AS	0.12 ± 0.06	0.00 ± 0.00	2.88 ± 1.86
E	0.2 ± 0.07	0.00 ± 0.00	3.1 ± 0.88
TS	0.09 ± 0.06	0.01 ± 0.01	1.57 ± 0.67

z-axis is less accurate, a very small weight is used for the model component. With the same consideration, the object and background seed strategy in Sec. 4 is implemented only in the middle one third slices, which is sufficient for delineation based on the property of IRFC.

The second row of Fig. 9 shows the delineation results in two slices overlaid on original image for LPS, RPS, and TB, respectively. The HD error for them is 1.3, 1.1, and 0.38 mm, respectively. The second row of Fig. 10 shows the delineation results in two slices for PC, AS, and E, respectively. The distance error of them is 3.36, 2.04, and 2.57 mm, respectively. The second row of Fig. 8 shows the delineation result for TS of distance error 1.3. The delineation results shown in the second row of Figs. 8–10 are all corresponding to the recognition results shown in the first row.

The computation time for the segmentation process has been estimated on Dell precision t3600 with 4-core Intel Xeon 3.6 GHz CPU with 8 GB RAM. The mean time for one organ over the 30 samples is reported in Table IV. The time for the computation of data likelihood is omitted since it is insignificant. It can be seen that the mean time for all organs for one subject is less than 2 min. The delineation of skin takes most of the time. However, if skin is not the segmentation object, other organs can still be recognized with the same framework. We have tested the recognition of RS and IMS without the initial condition from skin. There is no distinct difference in recognition accuracy and compute time. This is because the gray intensity grouping property makes the recognition robust.

TABLE IV. Mean computational time in seconds per subject for different objects and organ operations.

Object	Recognition	Delineation
Skin	—	48.40
RS	1.37	—
IMS	0.90	—
LPS	0.97	5.97
RPS	1.00	6.67
TB	0.10	6.53
PC	0.33	6.57
AS	0.97	8.30
E	0.33	8.03
TS	2.47	19.00
All	8.43	109.00

5.D. Comparison to results reported in the literature

Compared to the AAR strategy of Ref. 1 that does not use hierarchical registration, these results are generally better for both recognition and delineation. An overall improvement in mean FPVF and FNVF of nonsparse organs (skin, LPS, RPS, and PC) from 1.25% and 5.5% in Ref. 1 to 0.34% and 4.02% and sparse organs (TB, AS, and E) from 1% and 27% in Ref. 1 to 0.16% and 12.6% has been achieved by the new strategy. Note that these comparisons are on exactly the same data sets for the two methods. For the sparse organs, the decrease of both FPVF and FNVF is substantial, which led to the substantial decrease of distance error from 4.5 to 2.28 mm. This is mainly because of adaptive parameter estimation for the affinity computation used for some organs and more consideration for the object-based affinity computation and weight grouping for different weights. There is also improvement in the computation time for recognition by the new strategy. However, computation time has increased for the new strategy for object delineation. This is because of the new strategy for skin, which is time-consuming but achieves higher accuracy. In the new strategy, the search neighborhood in IRFC (adjacency relation in graph) is 18-adjacency for TS and 6-adjacency for others. Thus, in delineation TS is more time-consuming than others.

The performance of the proposed method is comparable to other state-of-the-art organ segmentation methods.[4,5,41–43] All methods, except our previous work,[1] focus on single or very few organs. Sofka *et al.*[4] proposed a method for shape model initialized by using landmarks and refined using a freeform refinement. The detection of a set of landmarks near the lung borders is based on the pose parameters of left and right lungs, which is predicted by a hierarchical detection network after identifying the carina in the trachea. The method was tested in 68 volume images. The symmetric point-to-mesh distance of detection results and annotations for both lungs is about 2.0 mm. Sun *et al.*[5] developed a lung segmentation method with a shape model and graph-cut approach. The shape model was trained using 41 scans with segmented lungs and initialized in a new scan with detected ribs. The method was tested on 30 scans and a mean Dice coefficient of 0.975 was obtained. The mean Dice coefficient for LPS and RPS on the 30 testing samples with our proposed method is 0.9737 and 0.9757, respectively.

In Ref. 41, a shape-based human airway segmentation scheme was proposed to suppress the leakage into surrounding areas which is based on fuzzy connectivity method. A mathematical shape optimization approach was embedded into the fuzzy connectivity algorithm. The method was tested on six real image datasets with a FNVF and FPVF of 0.1482 and 0.0119. Another method[44] for airway extraction adopted other standards such as total tree length and number of branches since it is difficult to manually and accurately trace and mark the 3D airway trees depicted on CT images and use these as a reference standard for evaluation purposes.

A model based 3D level set esophagus segmentation method initialized with a principal curve tracing algorithm is reported in Ref. 42 to solve the esophagus centerline detection problem. The result on thoracic CT scans (resolution

$0.98\times0.98\times3.75$ mm^3) from eight subjects was reported with a point-wise mean distance error of 2.1 mm. A multi-step method is proposed in Ref. 43 for esophagus segmentation which integrated a detector that is trained to learn a discriminative model of the appearance, prior shape expressed using a Markov chain model, and a nonrigid deformation to better fit the boundary of the organ. It achieved a mean boundary distance error of 1.80 mm. The segmentations typically ranged from the thyroid gland down to a level below the left atrium. Note that these methods are specially designed for esophagus and are difficult to generalize to other organs. Our method for esophagus is general and is within a general framework of AAR for other major objects within the thorax.

6. CONCLUSIONS

A general recognition method and a more specific delineation method for segmentation of thoracic organs are presented in this paper. Without loss of the core strategy, all organs are treated in the same way with parameter tuning for different organs. The parameters for recognition and delineation of each organ are kept constant among all test samples. Thus the segmentation method is automatic and general. Region based fuzzy modeling[1] is easy to implement in the optimization process and can be adapted to varying object topology. Hierarchical registration combined with fuzzy modeling can achieve robust recognition of parent organs and speed up the recognition of offspring organs. Weight differences in similarity metric provide tolerance for the fuzziness in data and emphasize the dominant parts of objects. Since there is a great deal of variation in the manner in which objects vary over a population and their geographic context, parameters will have to be set up differently for different classes of objects. We have demonstrated that still this can be achieved within a general framework. Model based affinity in fuzzy connectedness can regulate the delineation of contiguous objects having similar gray intensity.

Due to the existence of strong gray overlap between some organs and their surrounding structures, shape constraints based on the recognition result are used as complementary conditions to pure intensity constraints. However, due to the shape variability of organs among different samples, especially of the sparse organs, shape constraints introduced from the aligned fuzzy model may produce adverse side effects. Other local image information about gray intensity or shape is possible to be integrated into the IRFC algorithm to improve the delineation result. A global intensity parameter is adopted in the delineation for one organ in the target image. However, since there is large intensity variation at different locations of the same organ, such as in AS, such a global parameter may be inadequate to compute the affinity for all voxels. This observation is also desirable to be considered in the affinity computation in IRFC.

ACKNOWLEDGMENTS

This work was carried out while the first author was visiting the Medical Image Processing section of the University of Pennsylvania with support from a Grant No. 61163046 from the National Natural Science Foundation of China.

NOMENCLATURE

Skin	Skin boundary
RS	Respiratory system
LPS	Left lung
RPS	Right lung
TB	Trachea and bronchi
TS	Thoracic skeleton
IMS	Internal mediastinum
PC	Pericardial region
E	Esophagus
AS	Arterial system
VS	Venous system

[a)]Author to whom correspondence should be addressed. Electronic mail: kaiqiong@163.com

[1]J. K. Udupa, D. Odhner, L. Zhao, Y. Tong, M. M. S. Matsumoto, K. C. Ciesielski, A. X. Falcao, P. Vaideeswaran, V. Ciesielski, B. Saboury, S. Mohammadianrasanani, S. Sin, R. Arens, and D. A. Torigian, "Body-wide hierarchical fuzzy modeling, recognition, and delineation of anatomy in medical images," Med. Image Anal. **18**(5), 752–771 (2014).
[2]O. Camara, O. Colliot, and I. Bloch, "Computational modeling of thoracic and abdominal anatomy using spatial relationships for image segmentation," Real-Time Imaging **10**, 263–273 (2004).
[3]A. Moreno, C. M. Takemura, O. Colliot, O. Camara, and I. Bloch, "Using anatomical knowledge expressed as fuzzy constraints to segment the heart in CT images," Pattern Recognit. **41**, 2525–2540 (2008).
[4]M. Sofka, J. Wetzl, N. Birkbeck, J. Zhang, T. Kohlberger, J. Kaftan, J. Declerck, and S. K. Zhou, "Multi-stage learning for robust lung segmentation in challenging CT volumes," in *14th International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer Berlin Heidelberg, 2011), pp. 667–674.
[5]S. Sun, C. Bauer, and R. Beichel, "Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach," IEEE Trans. Med. Imaging **31**, 449–460 (2012).
[6]A. Criminisi, J. Shotton, and S. Bucciarelli, "Decision forests with long-range spatial context for organ localization in CT volumes," *MICCAI Workshop on Probabilistic Models for Medical Image Analysis (MICCAI-PMMIA), 2009*.
[7]X. Zhou, S. Yamaguchi, X. Zhou, H. Chen, T. Hara, R. Yokoyama, M. Kanematsu, and H. Fujita, "Automatic organ localizations on 3D CT images by using majority-voting of multiple 2D detections based on local binary patterns and Haar-like features," Proc. SPIE **8670**, 86703A (2013).
[8]*Recent Advances in Biomedical Image Processing and Analysis*, edited by T. M. Deserno (Springer, Hiedelberg, Germany, 2011), Chap. 5.
[9]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**, 725902–725908 (2009).
[10]A. Shimizu, T. Kimoto, H. Kobatake, S. Nawano, and K. Shinozaki, "Automated pancreas segmentation from three-dimensional contrastenhanced computed tomography," Int. J. Comput. Assisted Radiol. Surg. **5**, 85–98 (2010).
[11]T. Okada, R. Shimada, Y. Sato, M. Hori, K. Yokota, M. Nakamoto, Y.-W. Chen, H. Nakamura, and S. Tamura, *Automated Segmentation of the Liver from 3-D CT Images Using Probabilistic Atlas and Multi-Level Statistical Shape Model* (Springer-Verlag, Berlin, Germany, 2007), pp. 86–93.
[12]X. Chen and J. K. Udupa, "Automatic anatomy recognition via multi-object oriented active shape models," Med. Phys. **37**(12), 6390–6401 (2010).
[13]J. A. Bogovic, J. L. Prince, and P. L. Bazin, "A multiple object geometric deformable model for image segmentation," Comput. Vision Image Understanding **117**, 145–157 (2013).
[14]P. K. Saha and J. K. Udupa, "Relative fuzzy connectedness among multiple objects: Theory, algorithms, and applications in image segmentation," Comput. Vision Image Understanding **82**, 42–56 (2001).

[15]J. J. Cerrolaza, A. Villanueva, and R. Cabeza, "Hierarchical statistical shape models of multiobject anatomical structures: Application to brain MRI," IEEE Trans. Med. Imaging **31**, 713–724 (2012).
[16]N. Duta and M. Sonka, "Segmentation and interpretation of MR brain images: An improved active shape model," IEEE Trans. Med. Imaging **17**, 1049–1062 (1998).
[17]X. J. Chen, J. K. Udupa, A. Alavi, and D. A. Torigian, "GC-ASM: Synergistic integration of graph-cut and active shape model strategies for medical image segmentation," Comput. Vision Image Understanding **117**, 513–524 (2013).
[18]C. Chu, M. Oda, T. Kitasaka, K. Misawa, M. Fujiwara, Y. Hayashi, R. Wolz, D. Rueckert, and K. Mori, "Multi-organ segmentation from 3D abdominal CT images using patient-specific weighted-probabilistic atlas," Proc. SPIE **8669**, 86693Y–86697Y (2013).
[19]M. G. Linguraru, J. A. Pura, V. Pamulapati, and R. M. Summers, "Statistical 4D graphs for multi-organ abdominal segmentation from multiphase CT," Med. Image Anal. **16**, 904–914 (2012).
[20]C. Lu, Y. Zheng, N. Birkbeck, J. Zhang, T. Kohlberger, C. Tietjen, T. Boettger, J. S. Duncan, and S. K. Zhou, "Precise segmentation of multiple organs in CT volumes using learning-based approach and information theory," Med. Image Comput. Comput. Assisted Intervention **15**, 462–469 (2012).
[21]C. Meyer, J. Peters, and J. Weese, "Fully automatic segmentation of complex organ systems: Example of trachea, esophagus and heart segmentation in CT images," Proc. SPIE **7962**, 796216 (2011).
[22]T. A. Shen, H. S. Li, and X. L. Huang, "Active volume models for medical image segmentation," IEEE Trans. Med. Imaging **30**, 774–791 (2011).
[23]G. Tsechpenakis and S. P. Chatzis, "Deformable probability maps: Probabilistic shape and appearance-based object segmentation," Comput. Vision Image Understanding **115**, 1157–1169 (2011).
[24]J. Ulén, P. Strandmark, and F. Kahl, "An efficient optimization framework for multi-region segmentation based on lagrangian duality," IEEE Trans. Med. Imaging **32**(2), 178–188 (2013).
[25]A. Criminisi, D. Robertson, E. Konukoglu, J. Shotton, S. Pathak, S. White, and K. Siddiqui, "Regression forests for efficient anatomy detection and localization in computed tomography scans," Med. Image Anal. **17**, 1293–1303 (2013).
[26]J. Zhou and J. C. Rajapakse, "Segmentation of subcortical brain structures using fuzzy templates," NeuroImage **28**, 915–924 (2005).
[27]X. Zhou, S. Wang, H. Chen, T. Hara, R. Yokoyama, M. Kanematsue, and H. Fujita, "Automatic localization of solid organs on 3D CT images by a collaborative majority voting decision based on ensemble learning," Computerized Medical Imaging and Graphics **36**(4), 304–313 (2012).
[28]T. Okada, K. Yokota, M. Hori, M. Nakamoto, H. Nakamura, and Y. Sato, "Construction of hierarchical multi-organ statistical atlases and their application to multi-organ segmentation from CT images," in *Medical Image Computing and Computer-Assisted Intervention* (Springer Berlin Heidelberg, 2008), pp. 502–509.
[29]I. Isgum, M. Staring, A. Rutten, M. Prokop, M. A. Viergever, and B. Ginneken, "Multi-atlas-based segmentation with local decision fusion—Application to cardiac and aortic segmentation in CT scans," IEEE Trans. Med. Imaging **28**(7), 1000–1010 (2009).
[30]H. Wang, J. W. Suh, S. R. Das, J. B. Pluta, C. Craige, and P. A. Yushkevich, "Multi-atlas segmentation with joint label fusion," IEEE Trans. Pattern Anal. Mach. Intell. **35**(3), 611–623 (2013).
[31]R. Wolz, C. Chu, K. Misawa, K. Mori, and D. Rueckert, "Multiorgan abdominal CT segmentation using hierarchically weighted subjectspecific atlases," in *Medical Image Computing and Computer-Assisted Intervention*, Lecture Notes in Computer Science Vol. 7510, edited by N. Ayache, H. Delingette, P. Golland, and K. Mori (Springer, Berlin, Heidelberg, 2012), pp. 10–17.
[32]Y. Zhou and J. Bai, "Multiple abdominal organ segmentation: An atlas-based fuzzy connectedness approach," IEEE Trans. Inf. Technol. Biomed. **11**(3), 348–352 (2007).
[33]J. Liu and J. K. Udupa, "Oriented active shape models," IEEE Trans. Med. Imaging **28**, 571–584 (2009).
[34]K. Sun, J. K. Udupa, D. Odhner, Y. Tong, and D. A. Torigian, "Automatic thoracic anatomy segmentation on CT images using hierarchical fuzzy models and registration," Proc. SPIE **9036**, 90361P-1–90361P-8 (2014).
[35]A. Souza and J. K. Udupa, "Iterative live wire and live snake: New user-steered 3D image segmentation paradigms," Proc. SPIE **61443N**, 1159–1165 (2006).
[36]J. K. Udupa, D. Odhner, A. X. Falcao, K. C. Ciesielski, P. A. V. Miranda, P. Vaideeswaran, S. Mishra, G. J. Grevera, B. Saboury, and D. A. Torigian, "Fuzzy object modeling," Proc. SPIE **7964**, 79640B-1–79640B-10 (2011).
[37]M. J. D. Powell, "UOBYQA: Unconstrained optimization by quadratic approximation," Math. Program., Ser. B **92**, 555–582 (2002).
[38]K. C. Ciesielski, J. K. Udupa, A. X. Falcao, and P. A. V. Miranda, "Fuzzy connectedness image segmentation in graph cut formulation: A linear time algorithm and a comparative analysis," J. Math. Imaging Vision **44**, 375–398 (2012).
[39]J. K. Udupa and S. Samarasekera, "Fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation," Graphical Models Image Process. **58**, 246–261 (1996).
[40]L. Zhao, J. K. Udupa, D. Odhner, H. Wang, Y. Tong, and D. A. Torigian, "Automatic anatomy recognition of sparse objects," Proc. SPIE **9413**, 94133N-1–94133N-8 (2015).
[41]F. Y. Rizi, A. Ahmadian, N. Rezaie, and S. A. Iranmanesh, "Leakage suppression in human airway tree segmentation using shape optimization based on fuzzy connectivity method," Int. J. Imaging Syst. Technol. **23**(1), 71–84 (2013).
[42]S. Kurugol, E. Bas, D. Erdogmus, J. G. Dy, G. C. Sharp, and D. H. Brooks, "Centerline extraction with principal curve tracing to improve 3D level set esophagus segmentation in CT images," *33rd Annual International Conference of the IEEE EMBS*, *Boston, MA*, *August 30–September 3, 2011*.
[43]J. Feulner, S. K. Zhou, M. Hammon, S. Seifert, M. Huber, D. Comaniciu, J. Hornegger, and A. Cavallaro, "A probabilistic model for automatic segmentation of the esophagus in 3-D CT scans," IEEE Trans. Med. Imaging **30**(6), 1252–1264 (2011).
[44]J. Pu, C. Fuhrman, W. F. Good, F. C. Sciurba, and D. Gur, "A differential geometric approach to automated segmentation of human airway tree," IEEE Trans. Med. Imaging **30**, 266–278 (2011).