Automatic Anatomy Recognition of Sparse Objects

Liming Zhao[a,b], Jayaram K. Udupa[a], Dewey Odhner[a], Huiqian Wang[a], Yubing Tong[a],

Drew A. Torigian[a]

[a]Medical Image Processing Group, Department of Radiology, University of Pennsylvania, 423 Guardian Drive, Blockley Hall, 4th Floor, Philadelphia, PA 19104.
[b]Research Center of Intelligent System and Robotics, Chongqing University of Posts and Telecommunications, Chongqing, 400065, P. R. China.

Abstract: A general body-wide automatic anatomy recognition (AAR) methodology was proposed in our previous work based on hierarchical fuzzy models of multitudes of objects which was not tied to any specific organ system, body region, or image modality. That work revealed the challenges encountered in modeling, recognizing, and delineating sparse objects throughout the body (compared to their non-sparse counterparts) if the models are based on the object's exact geometric representations. The challenges stem mainly from the variation in sparse objects in their shape, topology, geographic layout, and relationship to other objects. That led to the idea of modeling sparse objects not from the precise geometric representations of their samples but by using a properly designed optimal super form. This paper presents the underlying improved methodology which includes 5 steps: (a) Collecting image data from a specific population group G and body region B and delineating in these images the objects in B to be modeled; (b) Building a super form, S-form, for each object O in B; (c) Refining the S-form of O to construct an optimal (minimal) super form, S^*-form, which constitutes the (fuzzy) model of O; (d) Recognizing objects in B using the S^*-form; (e) Defining confounding and background objects in each S^*-form for each object and performing optimal delineation. Our evaluations based on 50 3D computed tomography (CT) image sets in the thorax on four sparse objects indicate that substantially improved performance (FPVF~2%, FNVF~10%, and success where the previous approach failed) can be achieved using the new approach.

Key words: sparse anatomic objects, object modeling, fuzzy models, image segmentation, fuzzy connectedness

1. INTRODUCTION

Image segmentation in radiology is the analytic underpinning of all efforts that seek quantitative information from medical images. It is required in all image analysis processes, including research efforts in other related disciplines such as image registration, interpolation, filtering, and artifact suppression, and clinical applications such as the study, diagnosis, prognosis, radiomics, image-based therapy planning, and response assessment of diseases. Most past developments in image segmentation focused on specific objects and organ systems that were tailored to specific imaging modalities and protocols and hence are difficult to generalize to a different object, organ system, or different modalities. Demonstrated segmentation methods that are independent of objects, body-region, and imaging modality and protocol are at present rare. With body-wide imaging becoming common in many systemic diseases, increasing need for body-region-wide and body-wide quantitative information, and emerging trends in and their implications to medicine of Big Data to Knowledge programs, segmentation methods that generalize to the whole body and different imaging modalities are needed at present.

Motivated by the above needs, we have been developing a general computerized body-wide automatic anatomy recognition (AAR) methodology [1-4] based on hierarchical fuzzy models of multitudes of objects in several body regions. The goal was to build a general AAR system that is not tied to any specific organ system, body region, or image modality. Our work revealed that, while our AAR methodology was effective in segmenting spatially compact blob-like objects, which we refer to as *non-sparse* objects, thin, tubular, less space-filling objects, called *sparse objects*, posed special challenges in effective model building, and therefore in recognition and delineation. The main difficulty arose from the fact that if we built an object model based on the object's exact geometric representation, then the sparse nature of the object caused very little spatial overlap among its different samples used for model building. The difficulty stems mainly from the variation in sparse objects (compared to their non-sparse counterparts) in their shape, topology, geographic layout, and relationship to other objects. For example, the descending portion of the thoracic aorta is often straight and directed vertically inferiorly while in some subjects it may be oblique oriented, curved, or even tortuous, whereas other portions of the thoracic aorta, especially the aortic arch, do not vary much. While many segmentation methods have been reported in the literature specific to each sparse object, such as the main arterial trunks in the thorax, abdomen, and pelvis, esophagus, spinal cord, and ribs [5-9], they are not generalizable in the above spirit of the AAR system.

The above observation of the variations in sparse objects led us to the idea of modeling sparse objects not from the precise geometric representations of their samples but through use of properly designed optimal super forms. This paper presents the underlying concepts and demonstrates the significant improvement afforded by the new strategy to localize or recognize these objects and subsequently delineate them within the AAR fuzzy modeling framework.

2. METHODS

The proposed AAR approach for sparse objects consists of the following steps:

(a) Collecting image data from a specific population group G and body region B and delineating in these images the objects in B to be modeled.

(b) Building a super form, S-form, for each object O in B.

(c) Refining the S-form to construct an optimal super form S^*-form for each object, from which a hierarchical model of the object assembly is built.

(d) Recognizing objects in B by using their models in a hierarchical order.

(e) Defining confounding and background objects in each S^*-form for each object and performing optimal delineation.

These steps are described below.

(a) *Collecting image data and delineating objects*

We will follow the notations defined in [4]. G: the population group under consideration. B: the body region of focus. $O_1, \ldots, O_L$: L objects considered in B. $I = \{I_1, \ldots, I_N\}$: the set of images of B for G from N subjects used for model building and training the parameters of the AAR algorithms. $I_{n,\ell}$: the binary image representing the delineation of object O_ℓ in the image $I_n \in I$. The image set I is obtained from our patient image database by searching for images that are radiologically near normal in B. Fifty contrast-enhanced breath-hold CT image data sets of the thorax from 50 male

patients in the age range of 50-60 were utilized. These were collected from our hospital patient image database and all images were verified to be of acceptable quality and radiologically near normal with exception of minimal incidental focal abnormalities. Images were of routine clinical resolution with a voxel size of 0.7 mm x 0.7 mm x 5 mm. One half of the data set was used for model building and the remaining 25 image data sets were used for testing the methods. 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 (HIPAA).

We utilize the data sets created in the work of [4] and follow the definitions given in [4] for each body region and object. That is, the images for the body region and each object were consistently created following the definitions in Table A1 in [4]. In this paper, we focus on the thorax and the following sparse objects in that body region: thoracic skeleton (TSk), arterial system (AS), venous system (VS), and esophagus (E). The remaining (non-sparse and sparse) objects are as in our earlier work: Tskn – thoracic skin, LPS & RPS – left & right pleural spaces, TB – trachea and bronchi as a single object, RS – respiratory system (= LPS+RPS+TB), PC – pericardial region, IMS – internal mediastinum (= E+PC+AS+VS). The delineations were carefully created for all data sets with expert supervision and verification as explained in [4].

(b) *Building a rough super form, S-form, for each object*

As in our previous AAR work [4], we define the *Fuzzy Anatomy Model FAM*(B) of the body region B (for *G*) to be a quintuple FAM(B) = (H, $\mathcal{M}$, ρ, λ, η). H here is a hierarchy, represented as a tree, of the objects in B. We use the same hierarchy as in [4] with about 10 objects in Thorax. $\mathcal{M} = \{FM_\ell : 1 \leq \ell \leq L\}$ is a set of fuzzy models, one model per object in B. ρ describes the parent-to-offspring relationship in H over G. λ is a set of scale factor ranges indicating the size variation of each object O_ℓ over G. η represents a set of measurements pertaining to the objects in B. The hierarchy for thoracic objects is depicted in Figure 1.
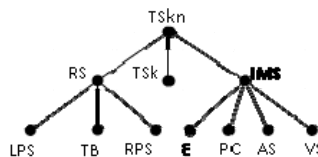

Figure 1. Hierarchy used for building the fuzzy anatomy model of the objects of the thorax.

The models for the non-sparse objects are built in the same manner as in [4]. Briefly, the process involves first aligning all binary images $I_{n,\ell}$ representing samples of object O_ℓ to the mean geometric center of all samples, then scaling each binary image to the mean scale factor, and finally estimating a fuzzy membership value at each voxel based on distance transform and a sigmoid function to convert averaged distance values to fuzzy membership values as explained in [4].

The models of sparse objects are created differently as described below. First, all N samples $I_{n,\ell}$ of O_ℓ are aligned as for non-sparse objects and a union of the aligned binary volumes is formed. This is followed by morphological closing operations to create a volume with minimal small inconsequential 3D cavities. The result is called the *S-form* of object O_ℓ.

(c) *Constructing an optimal super form, S -form, for each object*

In this step, for each object O_ℓ, its optimal super form, denoted S^*-*form*, is created from its S-form by applying an erosion operation t times to the S-form, where t is determined by minimizing the sum $\text{FPVF}^2+\text{FNVF}^2$ over all training samples of O_ℓ. Here FPVF and FNVF denote the false positive and false negative volume fractions, respectively, over all training samples yielded by the eroded S-form with respect to the aligned true delineations $I_{n,\ell}$ of O_ℓ. The underlying idea is that the S-form is generally too large as a super form, and for employing it as a model, its most compact version is preferable. The resulting S^*-form is taken to be the fuzzy model FM_ℓ of O_ℓ. For estimating all other parameters of the Fuzzy Anatomy Model FAM(B), namely $\mathcal{M}$, ρ, λ, and η, the S^*-form is used, following the methods described in [4] for fuzzy models.

(d) *Recognizing objects in* B

To localize an object in a given test image, we employ the optimal threshold-based strategy of [4] with a modification. This strategy involves finding an optimal threshold interval Th_ℓ for object O_ℓ by rehearsing a recognition trial over the training images in I. The trial consists of finding a single threshold interval which when applied to the training images yields the smallest sum FPVF+FNVF with respect to the model FM_ℓ over all training images. In the case of a sparse object, we take its S^*-form to be its fuzzy model FM_ℓ. (This training step is actually part of the model building process itself.) To actually recognize a sparse object in a given test image, the learned optimal threshold Th_ℓ is applied to the test image and the optimal pose for FM_ℓ that minimizes FPVF+FNVF is found. The recognition process proceeds hierarchically making essential use of the relationship ρ and the size range λ to make the search process accurate and very efficient [4].

(e) *Defining co-objects and performing optimal delineation*: The recognized objects are finally delineated by using the Iterative Relative Fuzzy Connectedness (IRFC) algorithm [10-13]. From knowledge of the co-objects for each object (this is anatomic knowledge that is also part of FAM(B)), the models associated with the object and its co-objects, and the intensity characteristics of the objects, we automatically determine the seeds needed for the object and its co-objects within the mask indicated by the object model (S^*-form). Again using the intensity characteristics of the object and its co-objects, optimal delineation of the object is performed via the IRFC mechanism.
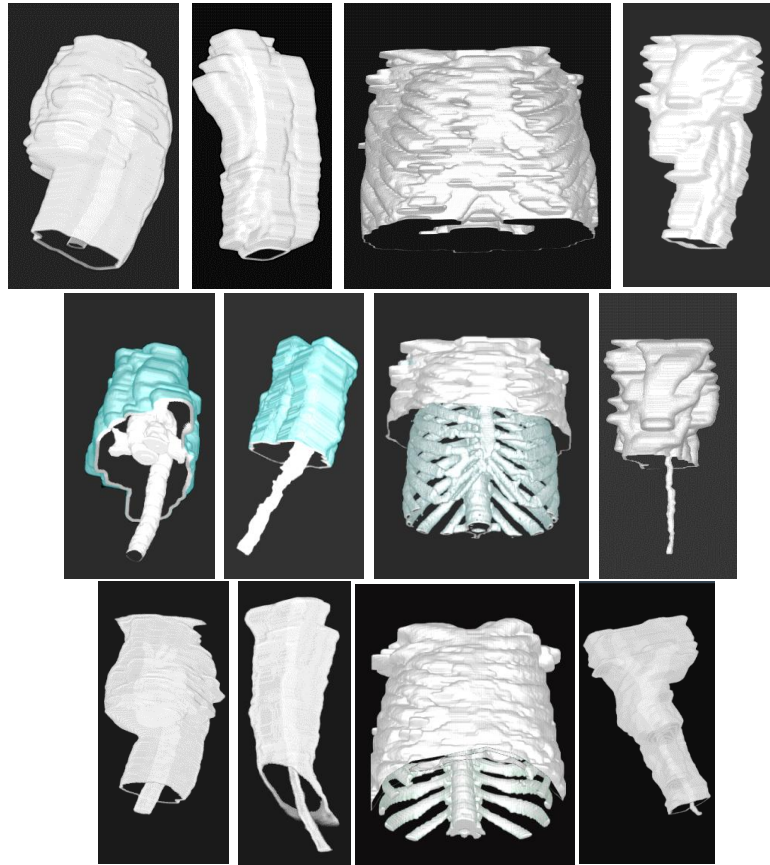
3. RESULTS
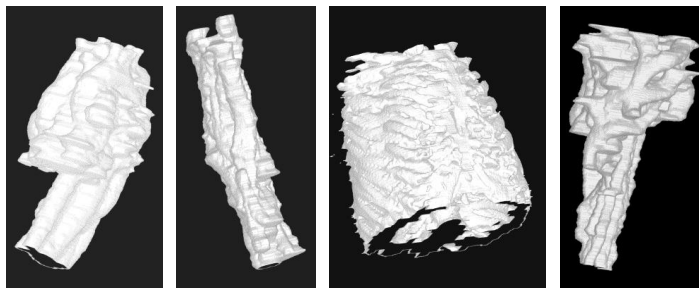
Figure 2. Objects (L to R): AS, E, TSk, and VS. 1st row shows the *S*-form before morphological closing. 2nd row shows the resulting *S*-form. 3rd row indicates the *S*-form (partially shown) at recognition superimposed on a true object to indicate coverage by the model. 4th row similarly shows the S^*-form.

Figure 2 displays for the four sparse objects considered (AS, E, TSk, and VS) 3D renditions of their *S*-form and *S**-form. Here, the recognition process is also illustrated in the figure for both forms. We can see that the recognized objects are inside both super forms. For reference, the true objects are depicted to demonstrate their coverage by the super form models.
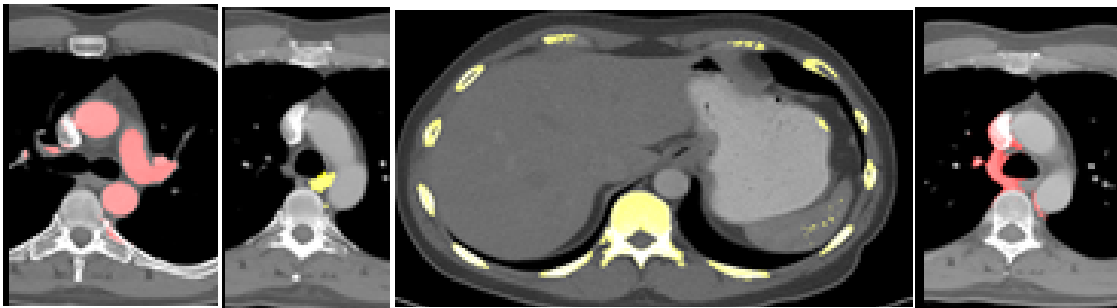

Figure 3. Delineation results. Objects from left to right are: AS, E, TSk, and VS.

Figure 3 shows some sample slices of the 3D segmentation result for the four sparse objects. We note that the segmentation of E, VS, and AS in these images is extremely challenging.

Some quantitative results of evaluation are listed in Table 1 for the four sparse objects where the delineation errors are expressed as FPVF and FNVF with respect to true object delineations. The mean and standard deviation of FPVF and FNVF over the 25 test cases are listed in the table. For comparison, the table also lists the two measures obtained for the same objects on the same test cases by using the previous non-sparse modeling technique of [4]. Note from the table that the delineation results have improved for all objects, in some cases significantly, compared to the previous AAR method which modeled the sparse and non-sparse objects in the same manner using exact object forms in the training set. The earlier method had failed at the recognition stage for VS (see [4] for details), and hence there was no delineation result for VS which is indicated by a "-" in the table. The new approach however was successful for VS with very good delineation results. The delineation of esophagus was quite poor in the previous method with nearly 50% FNVF. Sparse modeling however helped to bring this to under 10%.

	Method	AS	E	TSk	VS
FPVF	Sparse AAR	0.016 0.005	0.018, 0.003	0.017 0.014	0.034 0.006
	Prev AAR	0.01 0	0 0	0.19 0.05	-
FNVF	Sparse AAR	0.114 0.065	0.098 0.037	0.098 0.088	0.126 0.05
	Prev AAR	0.17 0.17	0.49 0.19	0.13 0.07	-

Table 1. FPVF and FNVF (mean and SD) over the 25 test image sets. For reference, these values for the previous AAR method [4] are also shown. The entry "-" means the delineation of the object (VS) was not tested by the previous method in [4] due to the fact that the object's recognition results were not acceptable.

4. CONCLUDING REMARKS

(a) The previous AAR approach [4] was a large effort to establish body-wide object recognition and delineation independent of body region, image modality, and organ system. This paper further advances that technology by devising new methods of modeling, recognizing, and delineating sparse objects. Sparse objects have posed challenges to the AAR methodology as well as to other modeling and atlas approaches to object segmentation in the literature.

(b) The idea that sparse objects should not be modeled by their precise geometric specification but by an optimal super form is novel. It fits well within the fuzzy modeling framework and provides a tangible solution to this difficult problem of segmenting thin, subtle, and often poorly defined objects. This is a general concept which we demonstrated in this paper via fuzzy modeling techniques. The principle however may be applicable to other sparse (and non-sparse) objects not considered in this paper. The idea may be applicable also to other object modeling approaches.

(c) The idea of defining confounding objects in the delineation process, encoding that information within the model, and exploiting that information within the IRFC framework for optimal delineation is another novel component of this paper. This idea follows naturally from the super form idea. The super form will naturally contain other

objects than the sparse object it attempts to represent. By studying the anatomy closely, however, all confounding objects can be enumerated and this information can be exploited in the proper delineation of the sparse object.

(d) Optimal super form is an effective concept for modeling sparse anatomic objects. However, there may be other more effective ways of defining the optimality criterion instead of the function $FPVF^2+FNVF^2$ used in this paper, including criteria that consider image intensity.

(e) The embodiment of this idea presented in this paper as S^*-form yields very good object recognition and delineation results in our preliminary tests for some sparse objects in the thorax, with a mean FPVF of 2% and FNVF of 10% over all objects. Also, where the previous method failed for some objects, the new method was successful.

(f) The new method fits naturally within the previous AAR framework retaining its high computational efficiency.

REFERENCES

[1] Udupa J K, Odhner D, Falcao A X, et al. Fuzzy object modeling[C], SPIE Proceedings, Medical Imaging, 79640B-1-79640B-10(2011).

[2] Udupa J K, Odhner D, Falcão A X, et al. Automatic anatomy recognition via fuzzy object models[C], SPIE Proceedings, Medical Imaging, 831605-1-831605-8(2012).

[3] Udupa J K, Odhner D, Tong Y, et al. Fuzzy model-based body-wide anatomy recognition in medical images[C], SPIE Proceedings, Medical Imaging, 86712B-1-86712B-7(2013).

[4] Udupa J K, Odhner D, Zhao L, et al., Body-wide hierarchical fuzzy modeling, recognition, and delineation of anatomy in medical images[J], Medical Image Analysis, 18, 752-771(2014).

[5] Duquette A A, Jodoin P M, Bouchot O, et al., 3D segmentation of abdominal aorta from CT-scan and MR images[J]. Computerized Medical Imaging and Graphics, 36, 294-303(2012).

[6] Avila-Montes O C, Kukure U, Kakadiaris I A., Aorta segmentation in non-contrast cardiac CT images using an entropy-based cost function[C].SPIE Medical Imaging. International Society for Optics and Photonics, 76233J-76233J-8(2010).

[7] Ecabert O, Peters J, Schramm H, et al., Automatic model-based segmentation of the heart in CT images[J]. IEEE Transactions on Medical Imaging. 27, 1189-1201(2008).

[8] Grosgeorge D, Petitjean C, Dubray B, et al., Esophagus Segmentation from 3D CT Data Using Skeleton Prior-Based Graph Cut[J]. Computational and mathematical methods in medicine 2013,1-6(2013).

[9] McIntosh C, Hamarneh G, Toom M, et al., Spinal cord segmentation for volume estimation in healthy and multiple sclerosis subjects using crawlers and minimal paths[C].Healthcare Informatics, Imaging and Systems Biology (HISB).25-31(2011).

[10] Udupa J K, Saha P K, Lotufo R A. Relative fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation[J], IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(11): 1485-1500(2002).

[11] Ciesielski K C, Udupa, J K, Saha P K, Zhuge Y. Iterative relative fuzzy connectedness for multiple objects with multiple seeds[J], Computer Vision and Image Understanding, 107(3), 160-182(2007).

[12] Ciesielski K C, Udupa J K, Falcao, A X, deMiranda, P AV. Fuzzy connectedness image segmentation in graph cut formulation: A linear-time algorithm and a comparative analysis[J], Journal of Mathematical Imaging and Vision, 44, 375-398(2012).

[13] Ciesielski K C, de Miranda P A V, Falcao A X, Udupa J K. Joint graph-cut and relative fuzzy connectedness image segmentation algorithm[J], Medical Image Analysis, 17 (8), 1046-1057(2013).