Abdominal Multi-organ CT Segmentation Using Organ Correlation Graph and Prediction-Based Shape and Location Priors

Toshiyuki Okada[1], Marius George Linguraru[2], Masatoshi Hori[1], Ronald M. Summers[3], Noriyuki Tomiyama[1], and Yoshinobu Sato[1]

[1] Department of Radiology, Graduate School of Medicine Osaka University, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan
toshi@image.med.osaka-u.ac.jp
[2] Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Medical Center, Washington DC, USA
[3] National Institutes of Health, Clinical Center, Radiology and Imaging Sciences, 10 Center Drive Bethesda, MD 20892, USA

Abstract. The paper addresses the automated segmentation of multiple organs in upper abdominal CT data. We propose a framework of multi-organ segmentation which is adaptable to any imaging conditions without using intensity information in manually traced training data. The features of the framework are as follows: (1) the organ correlation graph (OCG) is introduced, which encodes the spatial correlations among organs inherent in human anatomy; (2) the patient-specific organ shape and location priors obtained using OCG enable the estimation of intensity priors from only target data and optionally a number of untraced CT data of the same imaging condition as the target data. The proposed methods were evaluated through segmentation of eight abdominal organs (liver, spleen, left and right kidney, pancreas, gallbladder, aorta, and inferior vena cava) from 86 CT data obtained by four imaging conditions at two hospitals. The performance was comparable to the state-of-the-art method using intensity priors constructed from manually traced data.

Keywords: shape prediction, intensity model, partial least squares.

1 Introduction

Several general frameworks for abdominal multi-organ CT segmentation have been proposed. Recent approaches utilize a number of pairs of CT data and their manual traces, called atlases, as training data. Wolz et al. proposed a novel hierarchical method based on nonrigid registration and weighted label fusion using a large dataset of atlases [1]. The method uses sum of squared differences between target data and CT data in atlases, which means that the target data will be required to have similar contrast patterns as CT data in the atlases. Therefore, applicability to different imaging conditions (ICs) may be limited. In addition, nonrigid registration involves high computational cost and difficulty

K. Mori et al. (Eds.): MICCAI 2013, Part III, LNCS 8151, pp. 275–282, 2013.

in dealing with discontinuous deformations among abdominal organs. Machine-learning based approaches have also been studied [2,3]. Although these methods were validated for the various organs, they were not evaluated for organs having large inter-subject variations on shape and location, such as the pancreas and gallbladder. Since the features of the classifiers include the original CT value, the performance for unknown ICs is not clear.

Some multi-organ segmentation methods utilize prior models constructed from atlases [4,5]. Linguraru et al. used 4D graph incorporating multi-organ relations and two phases of intensity priors [4]. Okada et al. used patient-specific statistical atlases based on shape prediction from pre-segmented organ regions [5], which is also regarded as a generalization of some methods specialized to pancreas segmentation [6]. These methods used intensity priors based on intensity information obtained from traced CT data (hereafter we call "supervised intensity information"), and thus again the application to different ICs will be limited. Finally, Freiman et al. proposed a kidney segmentation method [7], in which the estimation of shape and intensity priors was performed by EM algorithm. Although intensity-based non-rigid registration between target and CT data in atlases was needed for initializing shape and intensity priors, only target data was used to estimate intensity prior without the atlases.

In this paper, we propose a framework for multi-organ segmentation, which is adaptable to any ICs without supervised intensity information. The features of the framework are as follows: (1) the organ correlation graph (OCG) is introduced, which encodes the spatial correlations among organs inherent in human anatomy. The concept of OCG is similar to [5], however, we provide fully automated methods for the construction and utilization of OCG. Then (2) the patient-specific organ shape and location priors obtained using OCG enable the estimation of intensity priors without using supervised intensity information and nonrigid registration. We evaluated the proposed method using 86 abdominal CT datasets obtained by four ICs at two hospitals.

2 Methods

We analyze eight organs, that is, the liver, spleen, left and right kidneys, gallbladder, aorta, inferior vena cava (IVC), and pancreas. The fundamental idea is to incorporate inter-organ spatial correlations to attain stable and accurate segmentation of the target organs. To do so, we first perform the segmentation of relatively stable organ in their position and shape (hereafter called "predictor" organs), and then segment other less stable organs whose positions and shapes are expected to be well-predicted by pre-segmented organs. In order to realize the above concept, we begin with a single organ segmentation method described in 2.1. In 2.2, we have a basic segmentation module to analyze multiple organs by using shape and location priors incorporating prediction from predictor organs. In 2.3, we obtain intensity priors from untraced CT data using the prediction-based priors in 2.2. Finally, we introduce OCG to assemble the single segmentation modules to establish a multi-organ segmentation system.

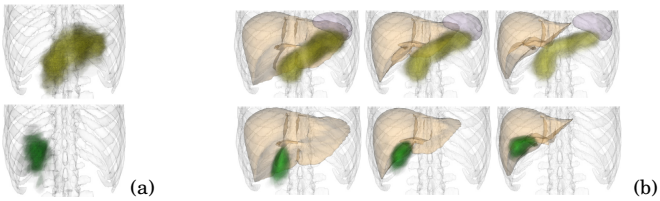


Fig. 1. Conventional and proposed probabilistic atlas (PA). (a) Conventional PA. (b) Proposed prediction-based PA. Upper: Pancreas. Lower: Gallbladder.

2.1 Single Organ Segmentation Method

As priors of the target organs, we utilize a statistical shape model (SSM) and a probabilistic atlas (PA) as shape and location priors [8], and an intensity model (IM) represented as Gaussian mixture model of the intensity distribution as intensity prior. After CT data is spatially transformed to the normalized space defined by the liver dome top and the rectangle circumscribing the bone tissue regions, voxel-wise MAP segmentation using PA and IM is performed, followed by SSM fitting with IM [8]. Finally graph-cut refinement is performed [9].

2.2 Prediction-Based Shape and Location Priors

We assume that $n-1$ predictor organs are available for the target organ. Let a set of predictor organ shapes be $\mathbf{s}$ and the target organ shape be $\mathbf{v}$, and let S and V be training data of the predictor and target organ shapes, respectively. We denote the function of PLSR (partial least squares regression) trained using S and V by $PLSR(\mathbf{s}; S, V)$ [10]. Given the pre-segmented predictor organ shapes $\mathbf{s}^*$, the prediction equation is given by $\mathbf{v} = \mathbf{v}^* + \mathbf{r}$, where $\mathbf{v}^* = PLSR(\mathbf{s}^*; S, V)$, $\mathbf{v}$ is the true shape, and $\mathbf{r}$ denotes the residual after the prediction. We represent $\mathbf{r}$ using PA and SSM [5].

To obtain PA and SSM of $\mathbf{r}$, we use the average shape of V as reference shape $\mathbf{v}'_0$. The true shape $\mathbf{v}$ is transformed using the 3D deformation field generated by correspondences from the predicted shape $\mathbf{v}^*$ to reference $\mathbf{v}'_0$, and then the transformed true shape $\mathbf{v}'$ is obtained. Now, we have $\mathbf{r}' = \mathbf{v}' - \mathbf{v}'_0$ of all training data, which are the residuals in the reference space, and then PA and SSM are generated. When we use PA and SSM for segmentation, the inverse of the 3D deformation field is used to bring them to the patient space. Figure 1 shows the prediction-based PA for the pancreas and gallbladder (and the comparison with the conventional PA). Segmentation is performed using the method described in 2.1 by replacing the conventional PA and SSM by the above-described prediction-based ones.

2.3 Estimating Intensity Model from Untraced CT Data

Two types of intensity models are constructed: (1) target-data specific intensity model (TD-IM) where only one CT dataset, which is the segmentation target is

used; (2) imaging-condition specific intensity model (IC-IM), where a sufficient number of CT data acquired using the same IC as the target are used.

The TD-IM is constructed as follows. Let I_0 be the target CT data. First, PA is binarized using the threshold value determined by the 1-percentile in the probability distribution of PA. The binarized region R is the initial volume of interest (VOI). The histogram H_0 is calculated from the intensity distribution of I_0 within R. IM is calculated by fitting a Gaussian mixture model (GMM) to H_0. MAP segmentation is applied to I_0 using PA and IM, and R is updated by the extracted region. The processes calculating IM with R and updating R by MAP are repeated sufficient times.

The IC-IM is constructed as follows. Let $I = \{I_i\}(i = 1, \ldots, N)$ be N optional CT data acquired using the same IC as I_0. Note that manual traces are unavailable for these CT data. TD-IMs are calculated for each $I_i(i = 0, 1, \ldots, N)$. Let $H = \{H_i\}(i = 0, 1, \ldots, N)$ be the calculated histograms. The average histogram $\bar{H} = \frac{1}{N+1} \sum_{i=0}^{N} H_i$ is calculated. IM is calculated by fitting GMM to $\bar{H}$.

Segmentation is performed using the method described in 2.1 by replacing conventional IM by the above-described TD-IM or IC-IM.

2.4 Multi-organ Segmentation Based on Organ Correlation Graph

The OCG is defined as a set of nodes and directed edges. Each node corresponds to an organ, and each directed edge denotes the correlation from one organ to the other. We consider a simple OCG defined by three nodes and two edges (Fig. 2 (a)). Two nodes having out-edges are regarded as the predictor organs, whose segmentation is assumed to be stable enough as predictors, and the other having in-edge as the target organ or response.

Among the eight organs, liver segmentation has been studied intensively, and is now sufficiently accurate and stable by itself [1,4,5,8]. The spleen and kidneys have also been studied by a sufficient number of works, which showed relatively good accuracy [1,4,5]. Finally, the segmentation of other organs has not been well-studied nor shown to be so accurate. Based on the above observations, we classified the eight organs into the following three types.

Type 1, $V_{C1} = \{Liver\}$: Only out-edges are defined (no in-edges). Segmentation is performed unconditionally.

Type 2, $V_{C2} = \{Spleen,\ Right\ kidney,\ Left\ kidney\}$: Both in- and out-edges are defined. Segmentation is performed under the condition that the segmentation is completed for at least one node connected by an in-edge. In addition, the organ can be a predictor for a node connected by an out-edge.

Type 3, $V_{C3} = \{Pancreas,\ Gallbladder,\ Aorta,\ IVC\}$: Only in-edges are defined (no out-edges). Segmentation is performed conditionally like for V_{C2}. This type cannot be a predictor.

Given the constraints on the above three types of nodes, the edges connections representing organ correlations are automatically defined based on shape predictability by PLSR [10]. Let $V_p = V_{C1} \cup V_{C2}$ be a set of nodes for prediction and $V_r = V_{C2} \cup V_{C3}$ be a set of nodes for response. Let V_{all} be all possible

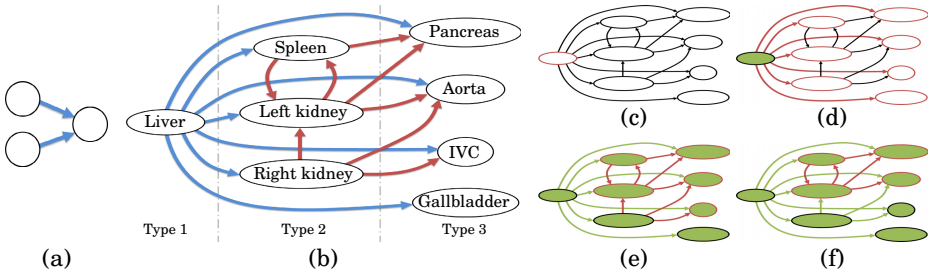


Fig. 2. Multi-organ segmentation based on organ correlation graph (OCG). (a, b) OCG. Blue and red edges indicate the directed edges from a node in Types 1 and 2, respectively. (c-f) Sequential segmentation steps based on OCG. Red border indicates the nodes to be segmented. Green node indicates segmented nodes.

combinations of the nodes in V_p. For each set in V_{all}, the prediction of node $v_r \in V_r$ is performed by applying PLSR and the prediction error is calculated [10]. Then a set of nodes $V_{min}(v_r)$, which has minimum prediction error, is selected. Let $E(v_r) = \{\{v, v_r\} | v \in V_{min}(v_r)\}$ be a set of directed edges to v_r. OCG $G =< V, E >$ is defined as $V = V_{C1} \cup V_{C2} \cup V_{C3}$ and $E = \bigcup_{v_r \in V_r} E(v_r)$. Figure 2 (b) shows OCG constructed from the eight organs.

Let v be a node to be segmented, $V_{out}(v)$ be a set of nodes having the edge directed from v, and $V_{in}(v)$ be a set of nodes having the edge directed to v. v is segmented using prediction-based priors constructed from the nodes in $V_{in}(v)$. After the segmentation of v, all nodes in $V_{out}(v)$ become ready for segmentation at the next stage. Let V^j_{target} be a set of nodes ready for segmentation at Stage j, and $V_{extracted}$ a set of segmented nodes. Let $Segmentation(v; V_{predictor})$ denote the segmentation of v based on prediction from $V_{predictor}$. The process of multi-organ segmentation based on OCG is formulated as follows:

1. $j \leftarrow 0$, $V^j_{target} \leftarrow V_{C1}$, $V_{extracted} \leftarrow \emptyset$.
2. For each $v \in V^j_{target}$, apply $Segmentation(v; V_{in}(v) \cap V_{extracted})$.
3. $V^{j+1}_{target} \leftarrow \bigcup_{v \in V^j_{target}} V_{out}(v)$, $V_{extracted} \leftarrow V_{extracted} \cup V^j_{target}$, $j \leftarrow j + 1$.
4. Repeat steps 2 and 3 a fixed number of times.

Fig. 2 (c) to (f) show the process of multi-organ segmentation based on OCG.

3 Experimental Results

We tested the proposed methods using 86 abdominal CT data obtained by four different ICs at two hospitals: Osaka University Hospital and National Institutes of Health (NIH). Table 1 shows the details of the four conditions. Figure 3 shows examples of CT data. In all data, the eight organs were manually segmented. No intensity information combined with manual traces was used in the proposed TD-IM and IC-IM.

Table 1. Detatils of CT scans used in experiment

Dataset	Institution	Phase	# of cases	Slice thickness [mm]
A	Osaka Univ. Hospital	late arterial phase	10	2.5
B	Osaka Univ. Hospital	late arterial phase	39	0.625
C	NIH	portal venous phase	25	0.68 - 1.25
D	NIH	non-contrast	12	1.0

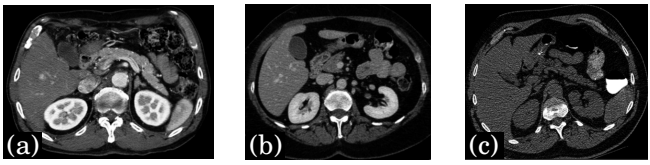
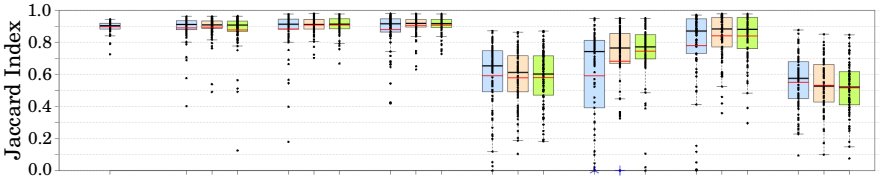


Fig. 3. Examples of CT cross-sectional images of our data. (a) Late arterial phase. (b) Portal venous phase. (c) Non (intra-venous) contrast but artifact due to oral contrast.

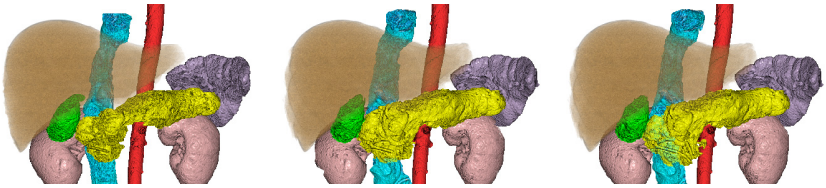
The segmentation accuracy using the proposed TD-IM and IC-IM was compared with the intensity model constructed from manual traces (hereafter called "supervised IC-IM"). Leave-one-out cross validation was performed for the evaluation of segmentation accuracy. The supervised IC-IM was constructed by leave-one-out method for each IC. Jaccard index (JI) and Dice coefficient (DC) were used for evaluation. The multi-organ segmentation based on OCG was performed until Stage 4. For liver segmentation, only TD-IM was used.



	Liver	Spleen	Right kidney	Left kidney	Pancreas	Gallbladder	Aorta	IVC
Proposed TD-IM	0.899 (0.946)	0.891 (0.940)	0.884 (0.932)	0.880 (0.932)	0.592 (0.718)	0.592 (0.682)	0.781 (0.843)	0.551 (0.696)
Proposed IC-IM	-	0.895 (0.943)	0.909 (0.951)	0.906 (0.950)	0.579 (0.718)	0.683 (0.770)	0.841 (0.907)	0.532 (0.680)
Supervised IC-IM	-	0.878 (0.929)	0.909 (0.951)	0.908 (0.951)	0.580 (0.718)	0.745 (0.837)	0.840 (0.906)	0.517 (0.667)

*: $p < 0.05$ **: $p < 0.01$

Fig. 4. Summary of accuracy evaluation for three intensity models. Average JIs of 86 cases using proposed TD-IM (blue box plot), proposed IC-IM (orange box plot), and supervised IC-IM (green box plot) for each of eight organs were plotted. Red and black lines in box indicate average and median JIs, respectively. In the liver, only the result using TD-IM was shown. Average JI and DC (shown in parenthesis) values and statistical significance are also shown below the plots.



(a) (b) (c)

	Liver	Spleen	Right kidney	Left kidney	Pancreas	Gallbladder	Aorta	IVC
Proposed IC-IM	0.919 (using TD-IM)	0.929	0.968	0.951	0.704	0.859	0.966	0.614
Supervised IC-IM		0.929	0.977	0.959	0.690	0.836	0.969	0.579

Fig. 5. Typical segmentation result. (a) Manual segmentation. (b) Proposed method using IC-IM. (c) Result using supervised IC-IM. Orange, purple, pink, yellow, green, red, and cyan volumes indicate the liver, spleen, kidneys, pancreas, gallbladder, aorta, and IVC, respectively. Jaccard index of each segmented organ is also shown. In the liver, JI using the proposed TD-IM is shown.

Figure 4 shows the segmentation accuracy for the three intensity models. Figure 5 shows typical segmentation results. Average JIs of the liver, spleen, right and left kidneys using the proposed intensity model were around 0.90 for all methods. These performances were comparable to state-of-the-art segmentation methods based on the availability of large training sets with manual tracings. In the pancreas and IVC, TD-IM achieved the best performance with average JI of 0.592 and 0.551, respectively. In the gallbladder (GB), supervised IC-IM achieved the best performance with average JI of 0.745. No significant accuracy improvement was observed between the proposed IC-IM and supervised IC-IM except for the GB.

The similarity between the proposed IC-IM and supervised IC-IM was calculated using normalized cross correlation (NCC). All NCCs were higher than 0.9 (average was 0.98 ± 0.03), except for GB where NCC was 0.64.

4 Discussion and Conclusion

We have presented a framework for multi-organ segmentation from abdominal CT which is adaptable to any contrast ICs without using supervised intensity information. The performance was comparable to the supervised intensity prior by using additional (untraced) CT data for intensity prior modeling. Results demonstrate that initialization of organ location and shape becomes accurate enough to estimate the intensity model in an unsupervised manner by using OCG and prediction-based priors.

Among the eight analyzed organs, GB was the only organ for which the supervised intensity information was significantly useful. Even for GB, however, if CT cholangiography data are used, in which the gallbladder contrast is enhanced, our method may be adaptable while the conventional method needs a number of manual traces to obtain additional supervised intensity information.

In the proposed method, failures of liver segmentation affect the subsequent prediction and segmentation of other organs. In non-contrast data, the prediction and segmentation of GB failed when the segmented liver region was leaked into GB. The low NCC of IM on GB was due to this leakage problem in addition to the small size of GB. Although this problem also occurred in IVC, its segmentation could be recovered using prediction-based priors.

Future work will include adding other abdominal organs, such as gastrointestinal tract, and the application to diseased organ segmentation.

Acknowledgments. This work is partly supported by KAKENHI No. 21103003.

References

1. Wolz, R., Chu, C., Misawa, K., Mori, K., Rueckert, D.: Multi-organ Abdominal CT Segmentation Using Hierarchically Weighted Subject-Specific Atlases. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 10–17. Springer, Heidelberg (2012)
2. Seifert, S., Barbu, A., Zhou, S.K., Liu, D., Feulner, J., Huber, M., Suehling, M., Cavallaro, A., Comaniciu, D.: Hierarchical parsing and semantic navigation of full body CT data. In: Pluim, J.P.W., Dawant, B.M. (eds.) Medical Imaging 2009: Image Proceedings, SPIE, vol. 7259, p. 725902 (2009)
3. Montillo, A., Shotton, J., Winn, J., Iglesias, J.E., Metaxas, D., Criminisi, A.: Entangled decision forests and their application for semantic segmentation of CT images. In: Székely, G., Hahn, H.K. (eds.) IPMI 2011. LNCS, vol. 6801, pp. 184–196. Springer, Heidelberg (2011)
4. Linguraru, M.G., Pura, J.A., Pamulapati, V., Summers, R.M.: Statistical 4D graphs for multi-organ abdominal segmentation from multiphase CT. Med. Image Anal. 16(4), 904–914 (2012)
5. Okada, T., Linguraru, M.G., Hori, M., Suzuki, Y., Summers, R.M., Tomiyama, N., Sato, Y.: Multi-organ segmentation in abdominal CT images. In: 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, pp. 3986–3989 (2012)
6. Shimizu, A., Kimoto, T., Kobatake, H., Nawano, S., Shinozaki, K.: Automated pancreas segmentation from three-dimensional contrast-enhanced computed tomography. Int. J. Comput. Assist. Radiol. Surg. 5(1), 85–98 (2010)
7. Freiman, M., Kronman, A., Esses, S.J., Joskowicz, L., Sosna, J.: Non-parametric iterative model constraint graph min-cut for automatic kidney segmentation. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part III. LNCS, vol. 6363, pp. 73–80. Springer, Heidelberg (2010)
8. Okada, T., Shimada, R., Hori, M., Nakamoto, M., Chen, Y.W., Nakamura, H., Sato, Y.: Automated segmentation of the liver from 3D CT images using probabilistic atlas and multi-level statistical shape model. Acad. Radiol. 15(11), 1390–1403 (2008)
9. Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In: Eighth IEEE International Conference on Computer Vision, pp. 105–112 (2001)
10. Rao, A., Aljabar, P., Rueckert, D.: Hierarchical statistical shape analysis and prediction of sub-cortical brain structures. Med. Image Anal. 12(1), 55–68 (2008)