A universal approach for automatic organ segmentations on 3D CT images based on organ localization and 3D GrabCut

Xiangrong Zhou*[a], Takaaki Ito[a], Xinxin Zhou[b], Huayue Chen[c], Takeshi Hara[a], Ryujiro Yokoyama[a], Masayuki Kanematsu[d,e], Hiroaki Hoshi[f] and Hiroshi Fujita[a]

[a] Department of Intelligent Image Information, Division of Regeneration and Advanced Medical Sciences, Graduate School of Medicine, Gifu University, Gifu-shi, 501-1194 Japan
[b] School of Information Culture, Nagoya Bunri University, 365 Maeda, Inazawa-cho, Inazawa-shi, 492-8520 Japan
[c] Department of Anatomy, Division of Disease Control, Graduate School of Medicine, Gifu University, Gifu-shi, 501-1194 Japan
[d] Radiology Services, Gifu University Hospital, Gifu-shi, 501-1194 Japan
[e] Department of Radiology, Gifu University Hospital, Gifu-shi, 501-1194 Japan
[f] Department of Radiology, Graduate School of Medicine, Gifu University, Gifu-shi, 501-1194 Japan

ABSTRACT

This paper describes a universal approach to automatic segmentation of different internal organ and tissue regions in three-dimensional (3D) computerized tomography (CT) scans. The proposed approach combines object localization, a probabilistic atlas, and 3D GrabCut techniques to achieve automatic and quick segmentation. The proposed method first detects a tight 3D bounding box that contains the target organ region in CT images and then estimates the prior of each pixel inside the bounding box belonging to the organ region or background based on a dynamically generated probabilistic atlas. Finally, the target organ region is separated from the background by using an improved 3D GrabCut algorithm. A machine-learning method is used to train a detector to localize the 3D bounding box of the target organ using template matching on a selected feature space. A content-based image retrieval method is used for online generation of a patient-specific probabilistic atlas for the target organ based on a database. A 3D GrabCut algorithm is used for final organ segmentation by iteratively estimating the CT number distributions of the target organ and backgrounds using a graph-cuts algorithm. We applied this approach to localize and segment twelve major organ and tissue regions independently based on a database that includes 1300 torso CT scans. In our experiments, we randomly selected numerous CT scans and manually input nine principal types of inner organ regions for performance evaluation. Preliminary results showed the feasibility and efficiency of the proposed approach for addressing automatic organ segmentation issues on CT images.

Keywords: Torso CT images, automatic organ segmentation, organ localization, 3D GrabCut.

1. INTRODUCTION

Medical imaging devices such as computerized tomography (CT) and magnetic resonance (MR) scanners have been widely used in clinical medicine to support diagnosis, surgery, and therapy. Such devices can generate substantial image data quickly and enable visualization of detailed three-dimensional (3D) information regarding the interior of the human body. However, doctors need time and experience with the interpretation of such image data using a screen or monitor, effective image analysis algorithms, and software tools to increase efficiency and accuracy substantially and to reduce tedium and oversights during CT image interpretation. Automatic segmentation of principal inner organs and tissues is necessary as a basic component of medical imaging systems. Although many research studies have been proposed for extracting different target organ regions in 3D medical scans, automatic organ segmentations are still challenging owing to image noise, artifacts, gray-level inhomogeneity, and wide variation in organ appearance.

A universal process of automatic segmentation for a wide variety of organs can be considered as a combination of "object detection" and "foreground/background separation." Object detection recognizes the target location on the image

Medical Imaging 2014: Computer-Aided Diagnosis, edited by Stephen Aylward, Lubomir M. Hadjiiski,
Proc. of SPIE Vol. 9035, 90352V · © 2014 SPIE · CCC code: 1605-7422/14/$18 · doi: 10.1117/12.2043507

space and outputs a minimum bounding box that tightly covers the organ region. Some ideas [1, 2] that had been applied successfully to face detection in computer vision research have been improved and used for organ localizations in CT images [3-8]. That progress indicated the possibility that locations of major organs can be detected correctly in 3D CT images. In contrast, foreground/background separation aims to determine the contour of the target organ region. Although the organ contours appear directly on images as local edge information, the traditional edge detection method was not sufficient for performing the foreground/background separation task on medical images owing to poor image contrast and ambiguity between the target organ and surrounding tissue. Some global information, such as the organ shape, inside/outside intensity distributions, is needed to assist in foreground/background separation. In order to balance local edge and global information, a cost function that includes both local and global information was defined and used to determine the best solution for contour detection. The graph cuts algorithm is a typical algorithm for performing this type of task [9]. In order to reduce the number of manual interactions (showing some foreground and background landmarks required by graph cuts), an iterated graph cuts algorithm, GrabCut, that requires only a bounding box of the target as the initialization is presented in [10]. GrabCut is a semi-automatic segmentation method, but it has the potential to become a fully automatic process by incorporating the object-detection result as a pre-processing step. As far as we know, most GrabCut applications have focused on segmentation in two-dimensional (2D) nature images, with no previous work reported for organ segmentation in 3D CT images.

In this paper, we propose an approach that can automatically segment inner organ and tissue regions in torso CT images. We aimed to use a simple method to deal with different kinds of inner organ segmentation issues generally and adapt the method to real clinical images generated by different CT scanners, including non-contrast and contrast-enhanced, for normal and abnormal cases. The basic idea of this approach is to combine object detection and the improved GrabCut algorithm based on a probabilistic organ atlas that we developed for supporting organ segmentation in our previous work [11].

2. METHODS

2.1 Outline

The process flow of the proposed approach is shown in Fig. 1. As described in the previous section, we divided the automatic organ segmentation process into two parts – (1) organ localization and (2) organ/background separation – and used these parts to achieve success and accuracy, respectively, in segmentation. The two parts were designed separately and independently of the target organs, and they were executed sequentially during the segmentation process. The details of each part are described in the following sections.

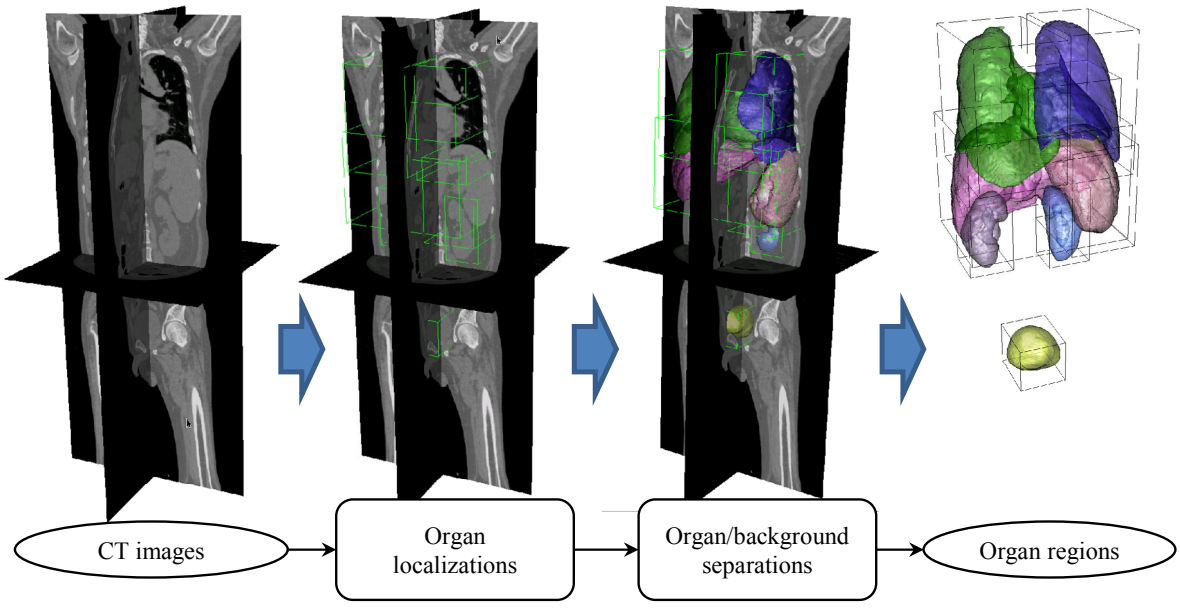

Fig.1. Process flow of the proposed approach for segmenting inner organs in a 3D CT scan.

2.2 Organ localization

The location of an inner organ was defined by a ground truth 3D minimum bounding rectangle (MBR) that covered all of the voxels in the target organ region and that was aligned with the x-, y-, and z-axes; i.e., its six faces were parallel to the x-y, y-z, and z-x planes. An automatic organ localization method was proposed in our previous work [12]. This method uses a sliding window to scan all of the positions in a CT scan and determined the MBR of a target organ on the basis of template-matching techniques in a selected feature space, as shown in Fig. 2. The feature space used for distinguishing the target organ from the background was constructed using an ensemble-learning approach [8]. In fact, a large number of Haar-like features [2] and local binary patterns [13] were used as candidates, and the AdaBoosting algorithm [14] was use to select distinctive features and generate an ad-hoc detector for each target organ,.

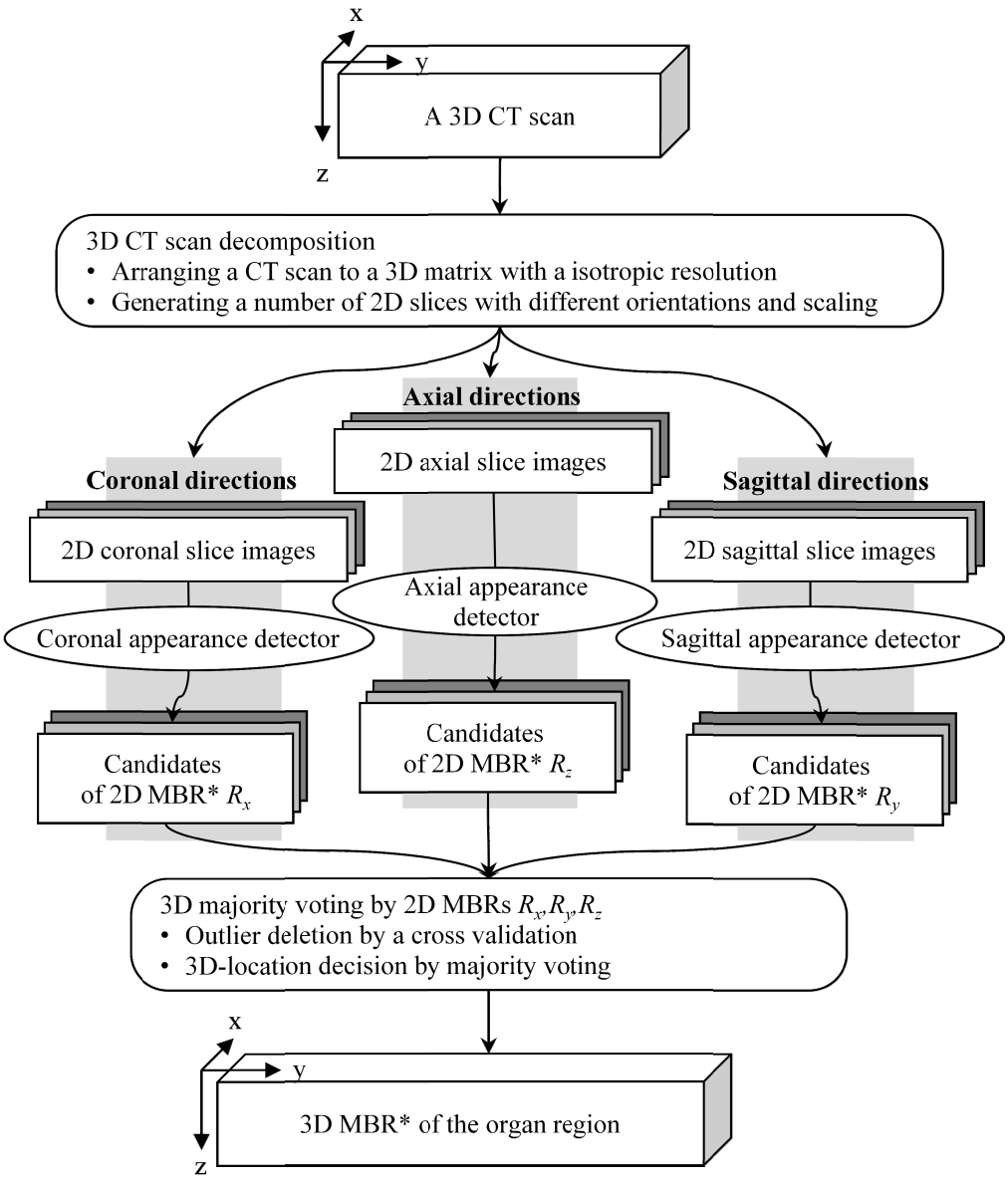

Fig. 2. Processing flow for organ localization in 3D CT images.

(*MBR: minimum bounding rectangle)

In this work, 3D organ localization is regarded as the detection of several 2D crossed slices of the target organ and voting to select the final organ location on the basis of the detected slices. The processing flow of organ localization is shown in Fig. 2. An inner organ region in a 3D CT scan is constructed using a series of consecutive 2D slices along a given direction (x, y, z or other orientation). The appearance of an organ in each 2D slice is highly correlated with and similar to its appearance in neighboring slices. Our assumption was that the partial appearances of an inner organ in 2D slices along the same direction are similar and can be recognized by a single 2D detector. Here, we only required a "weak" 2D detector, which may have optimal balance between the false positive (FP) and true positive (TP) rates, enhancing both efficiency and quality. This is the advantage of the traditional ensemble-learning approach. The later major voting step is used to further reduce FP rate and make a correct decision.

After 2D detection in a CT scan, we obtained a number of 2D bounding boxes (TP) that tightly enveloped the target organ from different directions with many FPs lacking any relationship to the target organ regions. Based on the tendency that the distribution of FPs centers inherently on one location but the distribution of FPs does not, we used a majority-voting technique to distinguish and delete FPs in 2D detection results. Finally, the underlying 3D MDR was estimated using the corners of the detected 2D bounding rectangles [8]. Compared to our previous work [12], this work increased the number of directions for 2D slice detection from three to twelve, thereby achieving higher robustness and accuracy in the final 3D localization results in the voting stage.

2.3 Organ and background separation using 3D GrabCut

Target organ segmentation can be achieved using foreground/background separation within the bounding box. We followed the GrabCut procedure, which had been successfully used for foreground/background separation on nature images [10]. To adapt this procedure to the organ segmentation tasks in 3D CT images, we expanded the Grab-Cut algorithm from 2D to 3D and improved the original method by combining it with the probabilistic atlas that had been presented in our previous work [11] to improve segmentation performance.

The original GrabCut was an interactive foreground extraction method based on iterated graph cuts. Compared to graph cuts segmentation [9], GrabCut required only a bounding box of the foreground of the image for initialization and greatly reduced the need for user interaction. Because the bounding box of a target organ can be detected automatically from CT images in this work, we combined organ localization with 3D GrabCut to achieve target-free automatic segmentation in CT images.

The graph cuts approach (the basis for GrabCut) addresses foreground separation in an image I by minimizing a pre-defined energy function E. The global minimum of the defined energy function E can be expected to yield a good segmentation result. Generally, a good segmentation result S is considered to be one in which foreground and background image intensity histograms (shown by a model with parameter P) are coherent and reflect a tendency toward object solidity. This is captured by an energy function of the form

$$E(s,p,i) = U(s,p,i) + V(p,i) \qquad \text{---- (1)}$$

Here, U is the approximate accuracy of the result S relative to the image I with model parameters P, and V is a smoothness value that encourages coherence in foreground or background regions with similar gray levels. GrabCut uses a Gaussian mixture model (GMM) that shows gray-level foreground and background histograms of an image. Intuitively, U reflects the likelihood of the gray level on each pixel belonging to foreground or background, and V relates to the image gradients. Minimizing E leads to a separation between foreground and background that satisfies the conditions that the gray-level histogram can be represented well by the GMM and that the smoothness was preserved in both foreground and background regions. Details are provided in [10].

The drawback of the original GrabCut method for organ segmentation in CT images is that the difference in the CT number distributions between organ and background is very small. For example, abdominal organs such as liver, spleen, kidney, and muscles have similar CT numbers, and most such organs are connected without an explicit contour separating them completely in CT images. This fact causes difficulty for GMM learning and for predicting the likelihood of distinguishing the organ region from the background in CT images.

This work improved the GrabCut method by adding a prior to show the target organ region within the bounding box during the separation process. Instead of the likelihood U in Eq. 1, we introduce a posterior probability U' based on Bayes' theorem, as shown in the following formula:

$$U'(s,p,i) = U(s,p,i) \times Pr(s,i) \qquad \text{---- (2)}$$

The prior *Pr*(*s, i*) shows the spatial frequency of background and foreground on each voxel within the bounding box. We call this prior, which is generated by an accumulation of the labeled foregrounds on a number of training samples stored in a database, a "probabilistic atlas." The details for probabilistic atlas construction can be found in our previous work [11, 12].

The processing flow of the improved 3D GrabCut method is shown in Fig. 3. The input is a pair of CT scans with a bounding box to indicate the target organ location. The output is a binary mask showing the target organ region in the CT scan. The details of each processing step are as follows.

1. Using the location of the bounding box and the CT image within the bounding box, a patient-specific atlas *Pr(s, i)* is generated using fusion of labels selected from a CT database. The bounding box is used for normalizing organ shapes and locations for different patient cases, and the CT image within the bounding box is used for similar case selection in atlas construction [12].

2. Two GMMs are estimated for representing the number of CT histograms within/outside of the bounding box or separated organ/background masks using an expectation–maximization (EM) algorithm. Based on the learned parameters of the GMMs, two likelihood maps for organ and background, respectively, are predicted.

3. The posteriors for the organ region and background are calculated using Eq. 2 based on the probabilistic atlas (spatial information) and likelihood (intensity information) obtained in steps 1 and 2, respectively.

4. A 3D graph based on posteriors and image gradients is constructed and a min-cut algorithm [9] is used to separate the organ and the background in 3D.

5. Steps 2 through 4 are repeated until convergence.

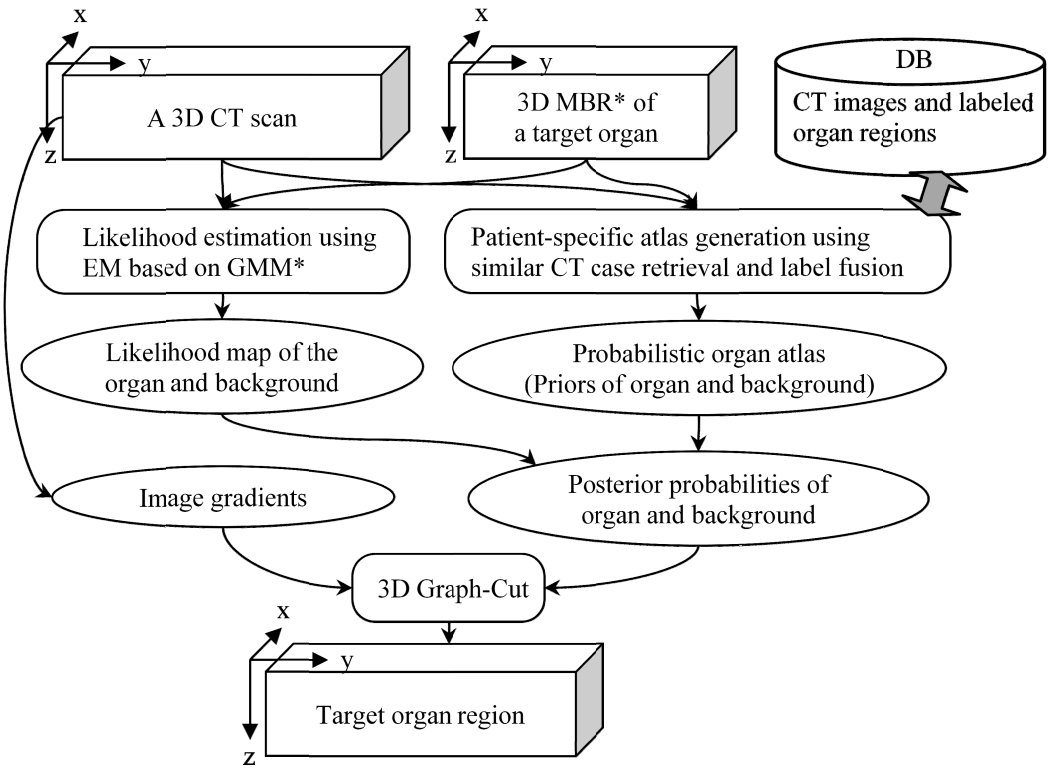

Fig. 3. Processing flow of 3D GrabCut method for organ segmentation in 3D CT scan using a probabilistic atlas.

(*MBR: minimum bounding rectangle; *GMM: Gaussian mixture model)

3. EXPERIMENT AND RESULTS

A dataset consisting of 1300 3D volumetric CT scans was used in this experiment. These CT scans were collected at Gifu University Hospital using two types of multi-slice CT scanners (LightSpeed Ultra16 of GE Healthcare and Brilliance 64 of Philips Medical Systems). Each CT scan used a common protocol (120 kV/Auto mA) and covered the entire human torso region. Each 3D CT scan has approximately 800 to 1200 axial CT slices with isotopic spatial resolution of approximately 0.6 to 0.7 mm and density (CT number) resolution of 12 bits. A portion of the CT cases was scanned with contrast enhancements. The patients ranged in age from 25 to 92 years. All CT images were taken for patients with real or suspected abnormalities.

Eighteen major organs and tissues (including heart, liver, gallbladder, spleen, pancreas, stomach, bladder, left-lung, right-lung, left-kidney, right-kidney, left-femur-head, right-femur-head, left-psoas-major-muscle, right-psoas-major-muscle, uterus and rectum, abdominal rectus muscle, and inferior vena cava (IVC) with ventral aorta) were selected as targets for evaluating the performance of the organ localization method. 3D locations of the targets in 300 3D CT scans were marked manually by the authors to train the 2D detectors. This led to approximately 4000 to 10000 positive 2D training samples and 40000 to 100000 negative 2D training samples, which were then used to train the 2D detectors for organs along the sagittal, coronal, axial and other directions. The detectors for each organ were trained separately and independently. Each 2D location detector consisted of 15 to 20 cascades, and each stage in the cascades was a classifier created by combining (boosting) five to 100 weak classifiers. The trained detectors were applied to localizing the 18 target organs in the other 1000 3D CT scans. We selected nine types of solid organ and tissue regions as targets for testing the performance of the segmentation method. The number of GMM mixture components was fixed empirically at three for foreground and five for background.

Accuracy of the bounding box and segmentation results was measured, first using a subjective evaluation by the authors and then by randomly selecting a number of samples for quantitative evaluations. The ground truth of each target organ region was manually labeled by the authors, and the Jaccard index between the segmented result and the ground truth was used as the evaluation measure. An example of detected organ locations in a 3D CT scan is shown in Fig. 4. The segmentation results of nine types of organ or tissue regions are shown in Fig. 5. Typical computing time using a general-purpose computer equipped with an Intel Core2Duo 2.23-GHz CPU was 45 seconds (localization: 15 seconds; segmentation: 30 seconds) per torso CT scan.

4. DISCUSSION AND CONCLUSION

In this study, the detected location was considered to be correct if most (two thirds of the volume) of the detected 3D rectangle and the ground truth MBR overlapped with each other. Our subjective evaluation based on 1000 CT test cases showed that the success rates for 14 organs (left-lung, right-lung, heart, liver, left-kidney, right-kidney, left-psoas-major-muscle, right-psoas-major-muscle, IVC with ventral aorta, left-femur-head, right-femur-head, bladder, abdominal rectus muscles, uterus and rectum) were in the range of 95% to 100%. The success rates for stomach and spleen were 94% and

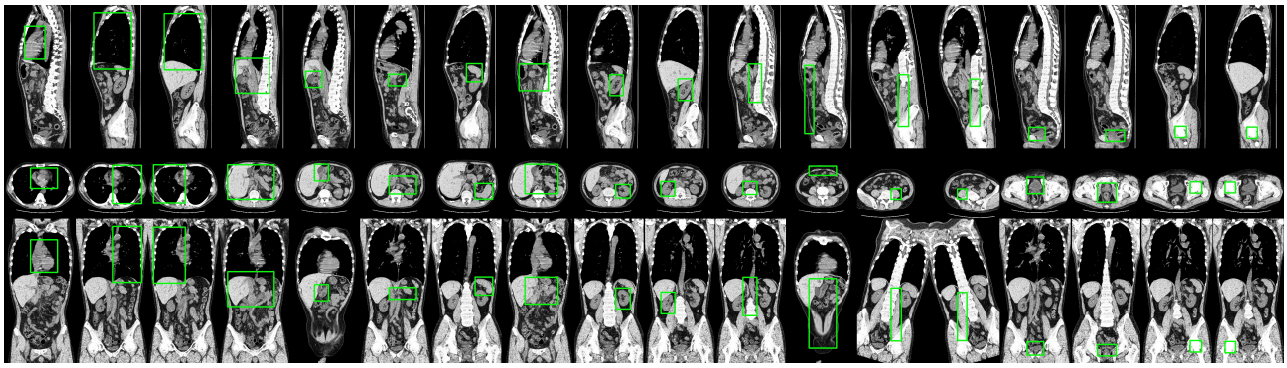

Fig.4. An example of the localization results for 18 kinds of major organ and tissues in a 3D CT scan. Three slices that pass through the detected center position of the target organ are shown. The rectangle indicates the detected organ location (bounding rectangle of the heart, left-lung, right-lung, liver, gallbladder, pancreas, spleen, stomach, left-kidney, right-kidney, IVC & ventral aorta, abdominal rectus muscles, left-psoas-major-muscle, right-psoas-major-muscle, bladder, uterus & rectum, left-femur-head, and right-femur-head, from the left to right sides).

90%, respectively, and the rates for pancreas and gallbladder were 82% and 84%, respectively. These results show that the proposed approach can perform localization tasks successfully in over 90% of CT cases, except for two types of organs. The failure of the localization for pancreas and gallbladder was due to poor performance of the trained 2D detectors. This problem was caused by the smaller sample number (smaller volume in 3D) and larger variances in those two organs in contrast to the other organs during the training stage. For the quantitative evaluation, the distance from the detected center position of the target organs was less than 50 voxels in most CT cases, with modes ranging from 10 to 20 voxels. The histograms of the difference in volume for the most organs were distributed mostly between -20% and 20%, with the modes near zero except for the bladder and liver. This evaluation result showed that the proposed approach can find the center position approximately and determine the size of the bounding rectangle of the target organ regions. However, for cases where the target organ had a large variance in the volume (bladder) or an asymmetrical shape (liver), the accuracy of the detected bounding rectangle needs to be improved.

The average values of the Jaccard indices for the segmentation results of each organ were distributed from 0.8 to 0.9. Preliminary results showed the usefulness and efficiency of the proposed method for segmenting massive organs in CT images. Compared to the original GrabCut method, which produced Jaccard indices ranging from 0.6 to 0.7, the proposed method improved the accuracy of the segmentation results for every type of organ. This improvement resulted from the probabilistic atlas, which introduced the spatial prior of the organs for supporting the segmentation process. On the other hand, the proposed method was unsuccessful in segmenting some digestive organs, such as stomach, pancreas, and rectum. This problem was caused by the large variances of CT number distributions and shapes for organ regions that cannot be handled by GMM learning and probabilistic atlas. Therefore, the cost function defined in Eq. 1 must be changed or expanded to adapt to digestive organ segmentations in non-contrast CT images.

In conclusion, we proposed a universal approach that can be used to segment major organs and tissues automatically in 3D CT scans. This approach was applied to nine types of organ regions, and its efficiency and accuracy were validated using real clinical CT scans.

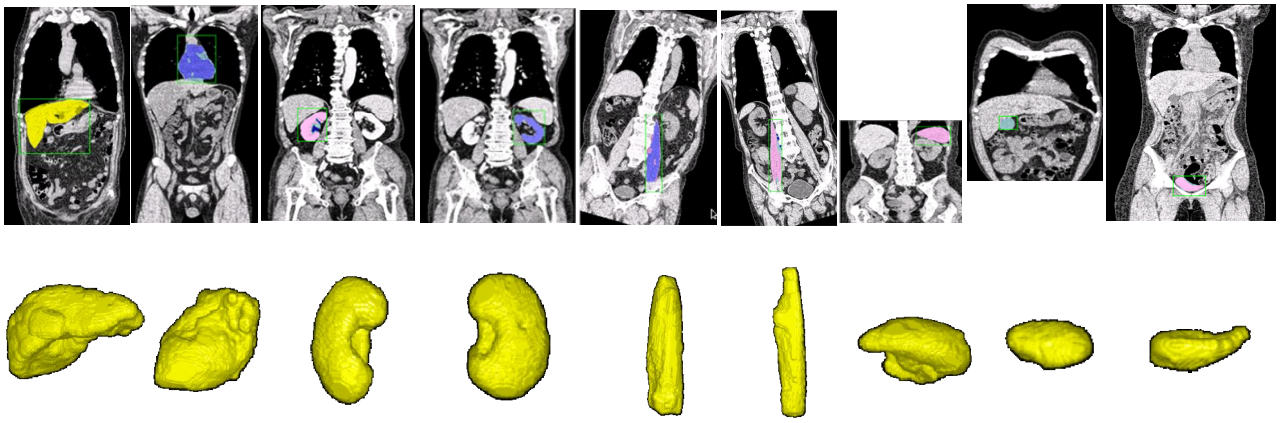

Fig. 5. Example of segmentation results for nine types of major organ and tissue in 3D CT scans. One coronal slice that passed through the detected center position of the target organ is shown. The segmented organ regions of liver, heart, right-kidney, left-kidney, left-psoas-major-muscle, right-psoas-major-muscle, spleen, gallbladder, and bladder are shown in color from left to right. 3D contours of each segmented organ were visualized using surface rendering and are shown under the images.

ACKNOWLEDGEMENTS

Authors thank the members of Fujita Laboratory who assisted to establish the dataset. This research work was funded in part by a Grant-in-Aid for Scientific Research on Innovative Areas (21103004) and in part by Grant-in-Aid for Scientific Research (C23500118), MEXT, Japan.

REFERENCES

1. P. Viola and M. J. Jones: Rapid object detection using a boosted cascade of simple features, IEEE CVPR, 2001.

2. R. Lienhart and J. Maydt: An extended set of Haar-like features for rapid object detection, IEEE ICIP 2002, vol. 1, pp. 900-903, 2002.
3. Y. Zheng, B. Adrian, G. Bogdan, S. Michael, and C. Dorin: Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features, IEEE TMI, vol. 27, pp.1668-1681, 2008.
4. H. Ling, S. K. Zhou, Y. Zheng, B. Georgescu, M. Suehling, and C. Dorin: Hierarchical, learning-based automatic liver segmentation, IEEE CVPR, 2008.
5. M. Dikmen, Y. Zhan, and X. S. Zhou: Joint detection and localization of multiple anatomical landmarks through learning, SPIE Medical Imaging 2008, vol. 6915, 2008.
6. A. Criminisi, J. Shotton, and S. Bucciarelli: Decision forests with long-range spatial context for organ localization in CT volumes, in MICCAI workshop on Probabilistic Models for Medical Image Analysis (MICCAI-PMMIA), 2009.
7. X. Zhou, and H. Fujita: Automatic organ localization on X-ray CT images by using ensemble learning techniques, in Machine Learning in Computer-aided Diagnosis: Medical Imaging Intelligence and Analysis, ed. by K. Suzuki, IGI Global, USA, pp. 403-418, 2012.
8. X. Zhou, S. Wang, H. Chen, T. Hara, R. Yokoyama, M. Kanematsu, H. Hoshi, 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, vol. 36, 4, pp. 304-313, 2012.
9. Y. Boykov and M.-P. Jolly: Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In International Conference on Computer Vision, vol. I, pp. 105–112, July 2001.
10. C. Rother, V. Kolmogorov, and A. Blake: GrabCut: Interactive foreground extraction using iterated graph cuts, ACM Trans. Graph., vol. 23, pp. 309–314, 2004.
11. X. Zhou, T. Kitagawa, T. Hara, H. Fujita, X. Zhang, R. Yokoyama, H. Kondo, M. Kanematsu and H. Hoshi: Constructing a probabilistic model for automated liver region segmentation using non-contrast X-ray torso CT images, Proc. of 9th International Conference for Medical Image Computing and Computer-Assisted Intervention - MICCAI 2006, Part II, vol. 4191, 856-863, Springer Berlin/Heidelberg, 2006.
12. X. Zhou, A. Watanabea, X. Zhou, T. Hara, R. Yokoyama, M. Kanematsu, and H. Fujita: Automatic organ segmentation on torso CT images by using content-based image retrieval, Proc. of SPIE Medical Imaging, Vol.8314, 83143E-1 - 83143E-7, (2012).
13. X. Zhou, A. Yamaguti, X. Zhou, T. Hara, R. Yokoyam, 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. of SPIE Medical Imaging 2013: Computer-Aided Diagnosis, 8670, 86703A-1 - 86703A-7, (2013).
14. T. Ojala, M. Pietikäinen, and D. Harwood: Performance evaluation of texture measures with classification based on Kullback discrimination of distributions, Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR 1994), vol. 1, pp. 582 - 585, 1994.
15. Y. Freund, and R. E. Schapire: A Decision-theoretic generalization of on-line learning and an application to boosting. In Computational Learning Theory: Eurocolt 95, Springer-Verlag, pp. 23–37, 1995.