

# Automatic organ segmentation on torso CT images by using content-based image retrieval

Xiangrong Zhou\*<sup>a</sup>, Atsuto Watanabe<sup>a</sup>, Xinin Zhou<sup>b</sup>, Takeshi Hara<sup>a</sup>, Ryujiro Yokoyama<sup>c</sup>,  
Masayuki Kanematsu<sup>c,d</sup>, and Hiroshi Fujita<sup>a</sup>

<sup>a</sup> Department of Intelligent Image Information, Division of Regeneration and Advanced Medical Sciences, Graduate School of Medicine, Gifu University, Gifu-shi, 501-1194 Japan

<sup>b</sup> School of Information Culture, Nagoya Bunri University 365 Maeda, Inazawa-cho, Inazawa-shi, 492-8520 Japan

<sup>c</sup> Radiology Services, Gifu University Hospital, Gifu-shi, 501-1194 Japan

<sup>d</sup> Department of Radiology, Gifu University Hospital, Gifu-shi, 501-1194 Japan

## ABSTRACT

This paper presents a fast and robust segmentation scheme that automatically identifies and extracts a massive-organ region on torso CT images. In contrast to the conventional algorithms that are designed empirically for segmenting a specific organ based on traditional image processing techniques, the proposed scheme uses a fully data-driven approach to accomplish a universal solution for segmenting the different massive-organ regions on CT images. Our scheme includes three processing steps: machine-learning-based organ localization, content-based image (reference) retrieval, and atlas-based organ segmentation techniques. We applied this scheme to automatic segmentations of heart, liver, spleen, left and right kidney regions on non-contrast CT images respectively, which are still difficult tasks for traditional segmentation algorithms. The segmentation results of these organs are compared with the ground truth that manually identified by a medical expert. The Jaccard similarity coefficient between the ground truth and automated segmentation result centered on 67% for heart, 81% for liver, 78% for spleen, 75% for left kidney, and 77% for right kidney. The usefulness of our proposed scheme was confirmed.

**Keywords:** Torso CT images, Content-based image retrieval, Phase-correlation registration, Atlas-based segmentation.

## 1. INTRODUCTION

X-ray CT images have been routinely used in clinical diagnosis for lesion detection and classification. A modern CT scanner can generate a large number of 2D CT slices to show the details of the human body within several seconds. However, CT image interpretation (viewing a large number of CT slices for each patient manually in front of a monitor or as films) requires a lot of time, energy and experience. Therefore, computer-aided diagnosis (CAD) systems that can support CT image interpretations are strongly anticipated. As the fundamental part of the CAD systems, segmentation of the inner organs on CT images is required. However, due to image noise, artifacts, gray level inhomogeneities, and the similar image appearance of the adjacent anatomical structures, accurate and robust segmentation of the inner organ regions on CT images still remains a challenge.

The traditional approaches for developing automatic segmentation algorithms focus on how to generate a model to show the organ appearances on CT images by the human designers. A model that shows the anatomical structures of a special organ should be constructed firstly and then used as the prior knowledge for supporting the segmentation task. However, generating such a model to present all the possible anatomical structures in both normal and abnormal CT cases is difficult and sometimes unrealistic, especially in the case that only a limit number of CT images are known during the development. Therefore, the approach that can learn and update the knowledge of the model directly from database and solve the different organ segmentation problems simply and straightforwardly is expected.

This research proposes a universal approach that can be used to automatically segment the different massive organ regions on CT images. The proposed approach simplifies the organ segmentation process by finding its location in CT images, searching the image patterns that are similar to the inputted image in a database, and transferring the anatomical

structures in the selected image patterns directly to the inputted image as the references to guide the segmentation. This approach is fully based on machine-learning and data-driven methods that use more image data instead of complex algorithms to enhance the robustness and accuracy of the organ segmentation process.

The paper is organized as the follows. In Section 2, we introduce the outline of the proposed approach and show the details of three core parts: automatic organ localization, content-based image retrieval, and atlas-based segmentation respectively. And then, we apply this approach to the segmentations of 5 kinds of massive organ on non-contrast CT images and show the experimental results in Section 3. The performance the approach is evaluated and discussed in Section 4. Finally, we give a conclusion in Section 5.

## 2. METHODS

### 2.1 Outline

As described in the previous section, the proposed approach simplifies the organ segmentation process as a content-based image retrieval and anatomical structure transformation problem. In order to measure the image similarity between the same kinds of organ regions in different CT cases, the location (bounding box) of the target organ on the CT images should be detected firstly. Furthermore, the post-processing is also needed to compensate the individual variation between the query target (inputted CT image) and query results (image patterns in the DB). Therefore, the approach includes three processing steps: (1) automated target organ localization, (2) content-based image retrieval and atlas construction, and (3) atlas-based organ segmentation. Two techniques have been used in this approach; one is fast object localization based on machine-learning, and the other is image retrieval by using a phase-correlation registration based on the fast Fourier transform. The outline of the proposed approach is shown in Figure 1.

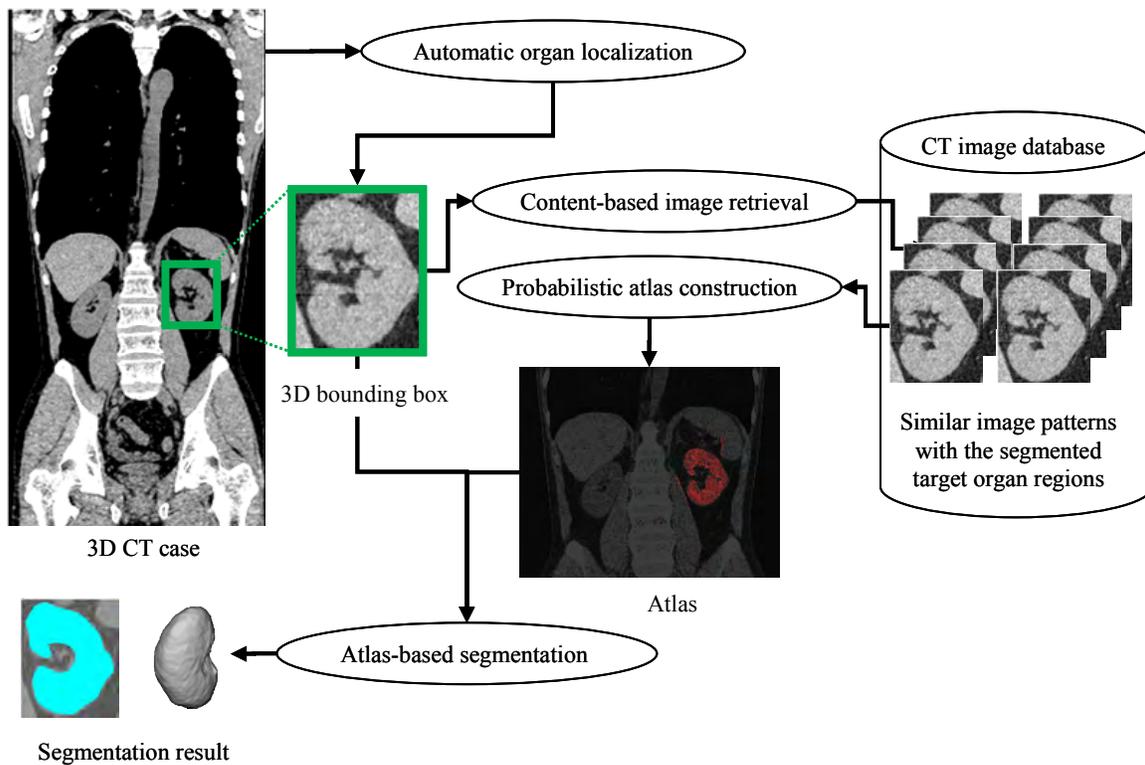


Fig. 1 Outline of the proposed approach for organ segmentation.

### 2.2 Automatic organ localization based on machine-learning

The location of an inner organ is defined by a ground-truth 3D minimum bounding rectangle (MBR) that covers all the voxels in the target organ region. We handle the organ location by detecting the MBR of the target inner organs in 3D

CT images. Our method is to treat 3D organ localization in a 3D CT image as training several independent 2D object detectors for a series of 2D image slices [1]. For an unseen 3D CT case, our method applies different 2D detectors to each voxel independently to detect a number of 2D target candidates and votes those 2D candidates back to the 3D space. Finally, we judge the existence of the target by checking the mutual consent of the responses from all 2D detectors, selecting the majority of the related 2D candidates in the 3D voting space as the target location (Figure 2). This solution aims to enhance the robustness of the training results by reducing the feature dimension (from 3D to 2D) and increasing the number of training samples (one 3D training sample consists of a large number of 2D training samples) during the machine learning. More details for training the 2D detectors and voting process are described in the following paragraphs.

A solid organ region in a 3D CT case is constructed by a series of consecutive 2D slices along a given direction (sagittal, coronal, or axial direction). The appearance of an organ in each 2D slice is highly correlated and similar to its neighbor slices. Our basic assumption is that the appearances of a solid organ in 2D slices along the same direction are similar and could be recognized by a single 2D detector. We take the 2D slices from the 3D training CT cases (with manually labeled ground-truth 3D MBRs) for training of the 2D organ-location detectors. Specifically, the slices along the sagittal, coronal, and axial directions are used for training the detectors to find candidates for 2D MBRs  $R_x$ ,  $R_y$ , and  $R_z$ , respectively [2]. Without loss of generality, in the following we focus on describing the training algorithm for finding the candidates for 2D MBR  $R_z$ . We collect the slices of the 3D training images along the axial body direction. If a slice intersects the ground-truth 3D MBR, we further check the 2D bounding rectangle resulting from this intersection. If the corseted target organ in this slice is representative, i.e., the target organ pixels account for a high percentage of the area of the 2D bounding rectangle, we crop this slice by this 2D bounding rectangle and then take the cropped slice as a positive 2D training sample. We randomly select a set of training slices cropped by rectangles that have no overlap with the ground-truth MBR as the negative 2D training samples. We then apply a cascaded AdaBoosting algorithm [3], using 2D Haar-like features [4], to train the 2D target organ-location detector, which can be applied to other axial-direction CT slices to find the candidates for 2D MBR  $R_z$ . In the same way, we can train 2D detectors for finding the candidates for 2D MBRs  $R_x$  and  $R_y$  using slices along the coronal and sagittal directions.

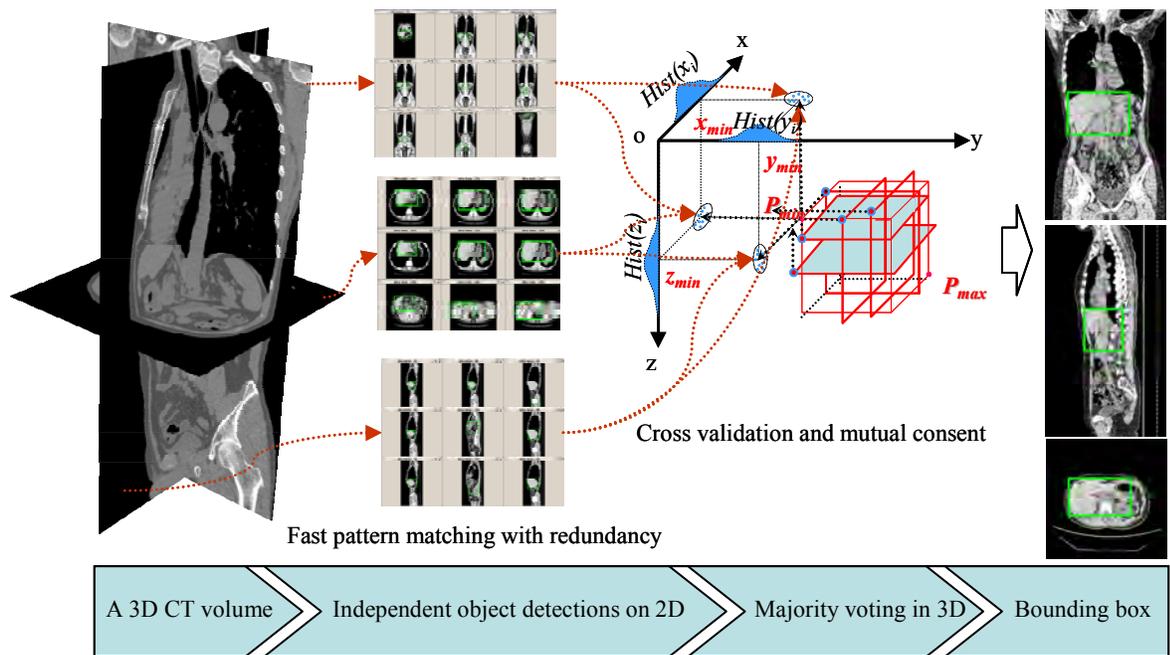


Fig.2 Processing flow of automatic organ localization.

Given an unseen 3D CT case, we first apply the three trained 2D location detectors to all the slices along the three directions to determine the possible existence of a target organ. Each location in the CT case will be checked three times, i.e., along the axial, coronal and sagittal directions, and only the locations that pass all three examinations are regarded as candidates for the target location. Ideally, the 2D bounding rectangles detected from different slices provide consistent

values of two corner coordinates ( $x_{min}, y_{min}, z_{min}$  and  $x_{max}, y_{max}, z_{max}$ ) of the bounding rectangles, from which we can derive the 3D MBR. In practice, however, the detected 2D rectangles may not lead to consistent values of  $x_{min}, y_{min}, z_{min}$  and  $x_{max}, y_{max}, z_{max}$  because of various noise and detection errors. We propose to use a majority-voting technique to achieve an optimal estimate of the values of  $x_{min}, y_{min}, z_{min}$  and  $x_{max}, y_{max}, z_{max}$  by selecting the mode of each coordinate value.

### 2.3 Content-based image retrieval based on phase correlation and probabilistic atlas construction

The aim of this processing step is to generate an atlas that predicts the prior probability for each voxel in the inputted CT case that belongs to the target organ region. This prior probability is calculated from a number additional CT cases that pre-stored in a database. Actually, this prior probability can be obtained simply by counting the frequency (number of the CT cases) that the voxel in current position of those CT cases is in the target organ region. As the precondition for counting the frequency in each voxel, the spatial coordinates of different the CT images should be unified and the appearances of the organ regions on query images and the CT images for atlas construction should be similar. The previous condition can be satisfied by image registration and the succeeding one can be accomplished by image similarity measuring. Those two processing modules are related together and can be concluded as a content-based similar image retrieval processing based on a CT image database.

The processing flow of the atlas construction for an inputted CT case is showed in Figure 3. A 3D-ROI that cropped from the inputted CT image based on the detected 3D bounding box (see Section 2.2) is used as the query image. The proposed scheme searches all the CT images in database and selects a number of CT cases that have the similar organ appearances to the query images. The basic idea of our method is to use the value of the cross correlation function as the similarity measure between two images and maximize this value to accomplish the spatial registration (rigid deformation) of those two images. The iteration of the image deformation and similarity measuring can be accomplished by using the Fast Fourier Transform (FFT) algorithm. The phase correlation method based on FFT is actually used in this scheme. The details of this method are described in the following paragraphs.

- (1) Given a query image  $a$  and select a image  $b$  from a CT image database
- (2) Apply a Hamming window function to both images  $a$ ,  $b$  and calculate the discrete 3D Fourier transform to obtain  $A = F(a)$ ,  $B = F(b)$ .
- (3) Calculate the cross-power spectrum  $R$  by taking the complex conjugate of the  $A$  and  $B$ , multiplying the Fourier transforms together element-wise, and normalizing this product element-wise to get  $R = (A B^*) / |A B^*|$
- (4) Calculate the normalized cross-correlation  $r = F^{-1}(R)$  by applying the inverse Fourier transform.
- (5) Determine the location and value of the peak in  $r$  and obtain  $s = \max(r)$  and translative offset in  $x$ ,  $y$ ,  $z$  directions between two images  $t_{x,y,z} = \operatorname{argmax}(r)$ .
- (6) Use  $s$  as the similarity measure and  $t_{x,y,z}$  as the parameters for deforming the organ region (binary image) in  $b$  to  $c$  that under the same spatial coordinates of  $a$ .
- (7) Repeats step (1)-(5) to compare the query image  $a$  with all the CT cases  $b$  stored in the database and select  $n$  CT cases  $c_i$  ( $i=[0-n]$ ) that have the higher similarities as the output.
- (8) Generate an atlas  $p$  for  $a$  by a weighted averaging of the  $c_i$  with its similarity measure  $s_i$ .

The atlas  $p$  shows the prior probability of the organ existence on each voxel within the bounding box of the input image.

### 2.4 Atlas-based organ segmentation

The organ segmentation is based on a Bayesian frame work that estimates the posterior probability of an organ existence on each voxel of the inputted image. This posterior probability can be formulated by a multiplication of a prior with the likelihood. The prior probability is obtained in Section 2.3 and the likelihood is calculated based on the CT number on each voxel. Here, we assume that the CT number distribution of a target organ is a Gaussian distribution. The parameters (mean and variance) of the Gaussian distribution are estimated from the CT number histogram within the bounding box of the target organ in the inputted CT images. Actually, the CT number which was the maximum peak in the histogram was selected as the mean value and a Gaussian distribution which has the same FWHM (full width half maximum) with the observed density histogram was decided [5]. The likelihood is generated from a non-linear density translation of the CT number histogram by using the Gaussian distribution as the translation curve as shown in Figure 4.

The target organ region can be extracted by simply selecting the voxel  $i$  with the condition that the posterior probability of  $i$  is larger than a threshold  $th$ , and then, the selected regions are refined by a binary morphological processing [using a ball kernel with a radius =  $r$ ]. At last, the biggest connected component in 3D was decided as the target region.

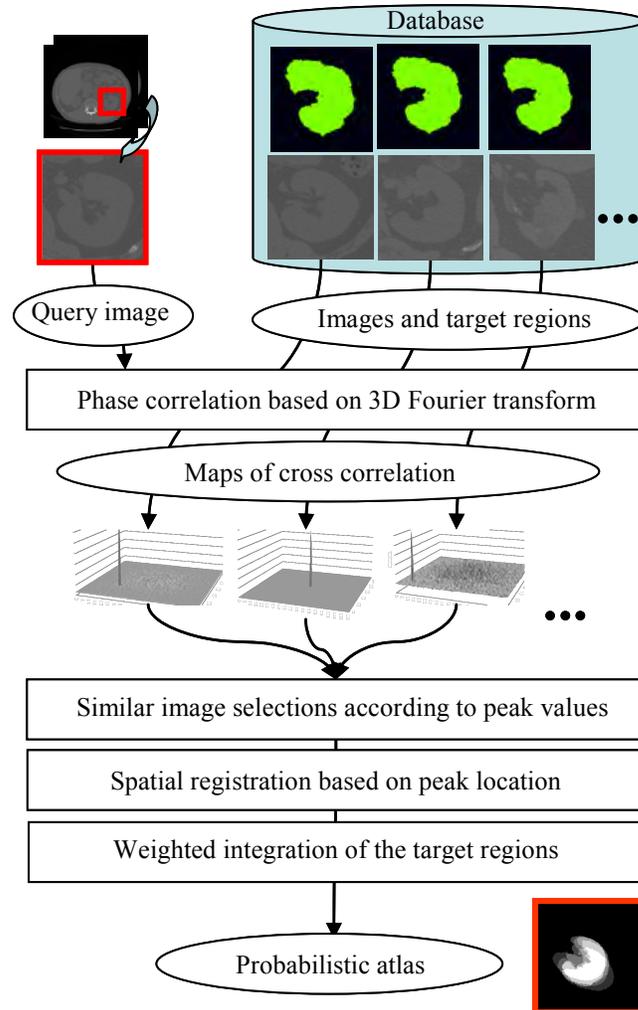


Fig.3 Processing flow of the atlas construction based on content based image retrieval.

### 3. EXPERIMENTS

A database (DB) that includes 100 cases of 3D volumetric CT cases was used in the experiment. These CT cases were collected at Gifu University Hospital by two kinds of multi-slice CT scanners (LightSpeed Ultra16 of GE Healthcare and Brilliance 64 of Philips Medical Systems). Each CT case used a common protocol (120 kV/Auto mA) and covered the entire human torso region. Each 3D CT case has approximately 800-1200 axial CT slices with an isotropic spatial resolution of approximately 0.625 mm and a density (CT number) resolution of 12 bits. All of these CT images were taken from patients with certain real or suspicious abnormalities.

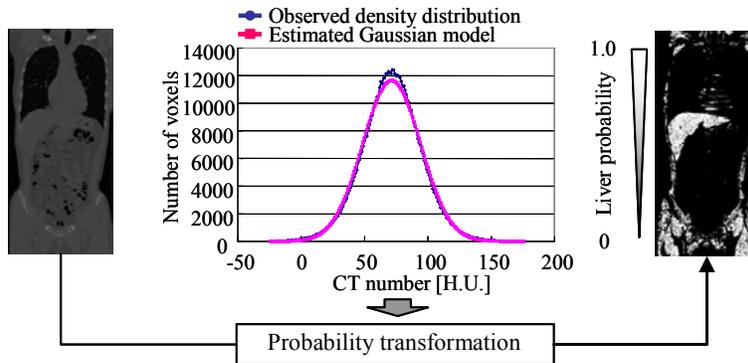


Fig.4 Processing flow of the likelihood estimation of liver region on CT images [5].

The heart, liver, spleen, left kidney, and right kidney are selected as the segmentation targets for evaluating the performance of the proposed approach. One of the authors (A.W.) manually extracted liver region in 38 CT cases, spleen region in 60 CT cases, left kidney in 93 CT cases, and right kidney in 35 CT cases. These regions were used as the ground truth for accuracy evaluation. A leave-one-out cross validation is used in the experiment. We pick one CT case from the DB and segment each target region respectively by querying similar image patterns from the other CT images left in the DB. The Jaccard similarity coefficient (JSC) between the segmentation result and ground truth is used as the accuracy measure. The parameters  $n$ ,  $th$ ,  $r$  were given as 10, 0.1, 1 empirically. Some examples of the segmented results were shown in 2D and 3D with the original CT images in Figure 5.

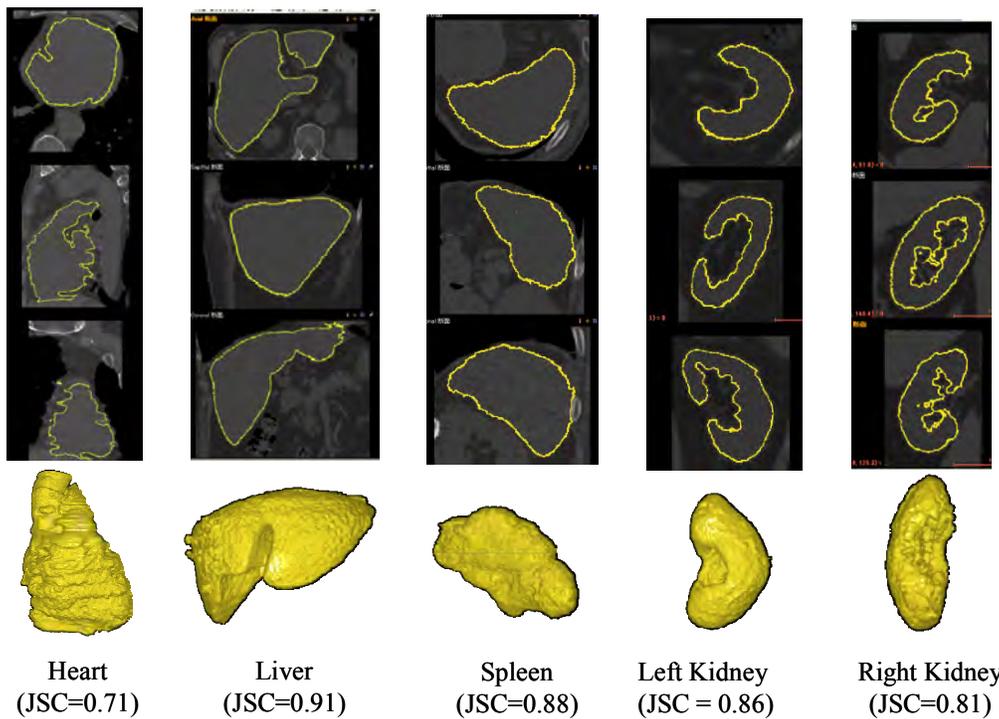


Fig.5. Examples of the segmentation results of heart, liver, spleen, left kidney, and right kidney. Each of the segmented organ regions is shown in 3 typical 2D image slices and a 3D surface rendering, where the yellow lines indicate the contour of the segmented regions. (JSC: Jaccard similarity coefficient between the segmentation result and ground truth).

## 4. DISCUSSION

The goal of this research is to build a frame work that can solve the automatic segmentation problem of the different organ regions by using the same algorithm. The experimental results showed that proposed scheme can solve the segmentations of the five different inner organs by using one algorithm. We confirmed that those five target organs were segmented automatically and correctly in all CT cases. The coincidence ratios showed by JSC values were centered on 0.67 for heart, 0.78 for spleen, 0.83 for liver, 0.77 for right kidney, and 0.73 for left kidney. Our method was very robust for different organ segmentations in both normal and abnormal CT cases. The segmentation results were approximate to the manual inputs of the human operators.

Content-based image retrieval is used as the key solution for different organ segmentation by providing the references of the anatomical structures for the inputted CT case. We found that the phase correlation method can output the similarity measure efficiently and correctly for two CT cases with a very similar organ appearance. However, the output of phase correlation function decreased quickly with the increasing of the dissimilarity between two images. In order to ensure the usefulness of the query results, a database including a large number of CT cases is necessary. We varied the numbers of the query results for atlas construction from 10-30 and found that the accuracy of the atlas did not change significantly.

The deterioration on the JSC values of the final segmentation results was caused by the difference of the definitions of the target regions between the automatic segmentation and manual inputs. The segmentation process assumed that the target organ region should be homogenous in CT number and the human operators tended to sketch a smoothing contour surrounding the organ region and ignore some other tissues (for example vessel region) inside the organ regions. This problem can be solved by adding the shape information to refine the segmentation result from this method.

## 5. CONCLUSION

We proposed a universal scheme for segmentation of different massive-organ regions automatically in 3D CT cases. Machine-learning-based organ localization, content-based image retrieval, and atlas-based organ segmentation techniques have been combined in this scheme to accomplish segmentation for the different organ regions by using the same algorithm. This scheme was applied to the segmentations of heart, liver, spleen, left and right kidneys in 100 cases of CT cases, and its efficiency and accuracy were showed in preliminary results. This scheme will be used for the model construction for understanding and commuting the human anatomy based on medical images [6].

## ACKNOWLEDGEMENTS

Authors thank to members of the Fujita Laboratory. 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. 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*, in press, 2012.
2. 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, 403-418, IGI Global, USA, 2012.
3. P. Viola and M. J. Jones: Rapid object detection using a boosted cascade of simple features. *Proc. Intl. Conf. on Computer Vision and Pattern Recognition (CVPR)*, 511-518, 2001.
4. R. Lienhart and J. Maydt: An extended set of haar-like features for rapid object detection, *Proc. Intl. Conf. on Image Processing (ICIP)*, 1, 900-903, 2002.

5. 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, 4191, 856-863, Springer Berlin/Heidelberg, 2006.
6. H.Fujita, T.Hara, X.Zhou, X.Zhang, T.Hayashi, N.Kamiya, H.Chen, and H.Hoshi: A01-3 Model construction for computational anatomy: Progress overview FY2010, Proc. of the Second International Symposium on the Project "Computational Anatomy", 13-17, Nagoya Univ., March 6th and 7th, 2011.