Automatic liver segmentation based on appearance and context information

Yongchang Zheng¹, Danni Ai²*, Jinrong Mu², Weijian Cong², Xuan Wang³, Haitao Zhao¹ and Jian Yang²

Abstract

Background: Automated image segmentation has benefits for reducing clinicians' workload, quicker diagnosis, and a standardization of the diagnosis.

Methods: This study proposes an automatic liver segmentation approach based on appearance and context information. The relationship between neighboring pixels in blocks is utilized to estimate appearance information, which is used for training the first classifier and obtaining the probability distribution map. The map is used for extracting context information, along with appearance features, to train the next classifier. The prior probability distribution map is achieved after iterations and refined through an improved random walk for liver segmentation without user interaction.

Results: The proposed approach is evaluated using CT images with eight contemporary approaches, and it achieves the highest VOE, RVD, ASD, RMSD and MSD. It also achieves a high average score of 76 using the MICCAI-2007 Grand Challenge scoring system.

Conclusions: Experimental results show that the proposed method is superior to eight other state of the art methods.

Background

Liver segmentation plays a key role in various clinical applications, such as liver diagnosis, tumor segmentation, and liver transplantation [1, 2]. This method faces a challenging problem because of the large variability, irregular shape, low contrast, and tissue adhesion of the live as well as intensified overlapping between organs. At present, liver segmentation can be divided into two categories: image- and statistical model-based methods.

Image-based segmentation uses image brightness, gradient, or texture [3, 4]. Peng et al. [5] proposed a convex variational model for liver semi-automatic segmentation in which gradient, local context-based edge detector, and regional appearance constraint are integrated to enhance the liver detection. Afifi et al. [6] proposed a liver segmentation approach in which the related information obtained between neighboring slices is integrated with the graph cuts algorithm to segment the liver in each CT slice. Chen et al. [7] introduced a strategic combination of the active appearance model, live wire, and graph cuts for liver segmentation. Peng et al. [8] proposed a multiregion-appearance-based approach with graph cuts to delineate liver edge. A geodesic distance-based appearance selection is introduced, and the energy functions incorporating both
boundary and region are modeled. Ai et al. [9] proposed a feature-learning-based random walk method for liver segmentation using CT images. Seed points on the original test image were automatically selected.

Statistical model-based segmentation uses a large number of data to establish a general liver location or shape model. Xu et al. [10] developed a 3D-scale invariant feature transform-based registration and designed an objective function to label the target image for liver segmentation. Salman et al. [11] discovered a feature-constrained Mahalanobis distance cost function to determine the active shape model, and liver segmentation is further achieved through a 3D graph cut. He et al. [12] proposed a three-level AdaBoost-guided active shape model for rough segmentation, surface optimization and shape-fitting refinement.

This study proposes an automatic context iteration-based liver segmentation by integrating appearance and statistical information. The gray level co-occurrence matrices (GLCM) are first extracted for appearance representation and used to learn the first classifier. Then, the context features are obtained from a large number of training data with ground truth along with image appearance features to learn the next classifiers. After several iterations, a prior probability liver model is constructed and further segmented through an improved random walk. The main contributions of this work are summarized as follows: (1) the context information rather than only appearance feature is used to construct the prior liver model, and (2) the improved random walk is introduced by integrating both the probability distribution map and original intensity image.

The rest of this paper is organized as follows. After the introduction, the methodology, including prior liver model construction and liver boundary determination, is described in “Methods” section. Evaluation results are presented in “Experimental results” section. “Discussion and conclusion” section concludes the paper.

**Methods**

Motivated in part by context feature extraction and machine learning, we propose an automatic liver segmentation method with two parts and shown in Fig. 1.

1. The prior model construction part contains the training step and test step. The appearance features are first extracted from the test image and training images with label map for classification. Then, the test probability distribution and training probability distribution with label map are obtained to extract the context feature. To achieve the final probability distribution map, both appearance and context features are used for classification on the several iterations.

2. Both intensity and probabilistic information are used by a prior model-based random walk to achieve the final segmentation result.

In this section, the appearance and context-based liver segmentation algorithm is presented in detail. This algorithm is comprised of prior liver model construction and prior model-based random walk.
Prior liver model construction

Appearance feature is extracted by gray level co-occurrence matrix (GLCM), which reflects the distance and direction between different pixels. In the patch $P$ centered on $x$, GLCM calculates probability $p(a, b|d, \theta)$ wherein the intensity value $a$ occurs with intensity value $b$ in spatial distance $d$ and direction $\theta$. Level $Q$ ($a, b \in Q$) in an image determines the size of GLCM ($Q \times Q$). A number of GLCMs are produced for $x$ according to different values of $d$ and $\theta$. Twelve textural features $U = 12$ are used for each GLCM measuring the characteristics of texture statistics. These features are energy, contrast, correlation, homogeneity, entropy, autocorrelation, dissimilarity, cluster shade, cluster tendency, maximum probability [13], statistics variance, and sum mean [14]. The appearance feature for $x$ is composed of textural features $f_{GLCM}^{d_j, \theta_j}$, $(u = 1, 2, \ldots, U)$ with given distance $d_j$ and direction $\theta_j$ $(j = 1, 2, \ldots, J)$.

$$\mathbf{f}_{GLCM}(x) = \begin{bmatrix} f_{11, \theta_1}^{d_1} & f_{12, \theta_1}^{d_1} & \cdots & f_{11, \theta_1}^{d_J} \\ f_{21, \theta_2}^{d_1} & f_{22, \theta_2}^{d_1} & \cdots & f_{21, \theta_2}^{d_J} \\ \vdots & \vdots & \ddots & \vdots \\ f_{U1, \theta_U}^{d_1} & f_{U2, \theta_U}^{d_1} & \cdots & f_{U1, \theta_U}^{d_J} \end{bmatrix}^T \tag{1}$$

The training data set is denoted as $V_i \in \mathbb{R}^{N \times M}$ and the corresponding label map is denoted as $V_{yi} \in \mathbb{R}^{N \times M}$, $i = 1, 2, \ldots, I$, and $I$ is the number of training data. Training sample points are expressed with appearance feature $f_{GLCM}(x_t)$ and the corresponding label $y_t$.

$$S_0 = \{(y_t, f_{GLCM}(x_t)) \mid t = 1, 2 \ldots T\} \tag{2}$$

The first classifier is trained through the AdaBoost algorithm based on the appearance features calculated on local image patch. Probability distribution maps $P_0(V_t)$ are obtained for the training data set using the learned classifier. We have

$$p_0(y_t = 1|x_t) = \frac{e^{H_0(f_{GLCM}(x_t))}}{e^{H_0(f_{GLCM}(x_t))} + e^{-H_0(f_{GLCM}(x_t))}} \tag{3}$$
As the largest organ located on the right side of human body, the liver features a special shape and is relatively fixed near other organs. The appearance and context information should be used for liver segmentation simultaneously. Figure 2 shows the appearance and context information extracted from the original image and the probability distribution map, respectively.

In the probability distribution maps, eight rays with $45^\circ$ intervals are stretched out from $x_t$, and a radius sequence is used to sample the context locations on each ray. The probabilities on the locations are used as context features $f_{\text{ctx}}(x_t)$. Afterward, a series of new training data set is constructed as follows:

$$S^{(u)} = \left\{ \left( y_t, \left( f_{\text{GLCM}}(x_t), f_{\text{ctx}}(x_t) \right) \right), \quad t = 1, 2, \ldots, T \right\},$$

where $u = 1, 2, \ldots, U$ denotes the iteration number. The second classifier based on the appearance and context features $(f_{\text{GLCM}}(x_t), f_{\text{ctx}}(x_t))$ is identified to construct new probability distribution maps $P_1(V_t)$ in which the classification is expressed as

$$p_1(y_t = 1|x_t) = \frac{e^{H_U\left( (f_{\text{GLCM}}(x_t), f_{\text{ctx}}^{(1)}(x_t)) \right)}}{e^{H_U\left( (f_{\text{GLCM}}(x_t), f_{\text{ctx}}^{(1)}(x_t)) \right)} + e^{H_U\left( (f_{\text{GLCM}}(x_t), f_{\text{ctx}}^{(1)}(x_t)) \right)}}$$

A similar learning procedure is performed to obtain $U$ learned classifiers. The iteration procedure is shown in Fig. 3. Figure 3a shows the training data $V_t$ (bottom line) and the corresponding segmented liver $V_{tS}$ (top line). Based on the appearance features extracted from the original image, the initial classification, denoted as $p^0(y_t = 1|x_t)$, provides the probability distribution map $P_0(V_t)$ shown in Fig. 3b (top line). The appearance and context features are then integrated and processed through multiple iterative classification (Fig. 3c, d) to obtain the final classification result $p^U(y_t = 1|x_t)$, which is shown in Fig. 3e as the probability distribution map $P_U(V_t)$.

For the test data $V$, the prior liver model $V_s$ is obtained through the sequence of classifiers based on both appearance and context features. The same procedure performed on the training data is implemented, including appearance feature extraction, context

---

**Fig. 2** Appearance and context information extracted from the original image and probability distribution map, respectively. a Appearance information in the original image; b Context information in the probability distribution map.
feature extraction, classifier learning and iteration loop until convergence. The final classification result for each pixel constructs the probability distribution map \( P_U(V) \).

**Prior model-based random walk**

The original text image and probability distribution map are defined as two corresponding graphs. Nodes \( v_i \) are pixels, and edges \( e_{ij} \) are produced between two neighbor pixels. The original random walk [15] has two limitations: (1) the procedure is not automatic such that the seeds have to be manually appointed by the user and (2) only the intensity information is not distinguishable for liver segmentation. To address these problems, prior liver model is used for seed determination [16], that is, pixels \( x_t \) with \( p(y_t = 1|x_t) = 1 \) are denoted as liver seeds, and pixels \( x_t \) with \( p(y_t = 1|x_t) = 0 \) are denoted as background seeds. Both the intensity and probability distribution information are utilized to calculate the most significant parameter of random walk, that is, the weight \( w_{ij} \) of \( e_{ij} \):

\[
\omega_{ij} = \exp \left( -\beta \left( 1 - \alpha \right) \left( G(v_i) - G(v_j) \right)^2 + \alpha \left( P(v_i) - P(v_j) \right)^2 \right)
\]

(6)

where \( G(v_i) \) and \( P(v_i) \) are the pixel intensity value and corresponding probability distribution value of the original image and probability distribution map, respectively, and \( \alpha \) and \( \beta \) are the adjusting parameters. When liver pixel \( v_i \) and background \( v_j \) are near the liver edge with similar intensity values but different probabilities, relatively precise segmentation result can be achieved.

According to the original random walk, the final probability of \( v_j \) is determined by the minimum objective function as follows

\[
P^\text{spatial} = x^T L x = \sum \omega_{ij} (v_i^s - v_j^s)^2, \quad \text{s.t. } L_{v_i v_j} = \begin{cases} 
   d_{v_i} & \text{if } i = j \\
   -\omega_{ij} & v_i, v_j \text{ are neighbors} \\
   0 & \text{others}
\end{cases}
\]

(7)
where $x_i^s$ is the probability of each pixel in the image belonging to the different classes, and $s$ denotes the liver $s = 1$ or background $s = 0$. After finding the value of $x_i^s$, we can determine to which class a certain pixel belongs:

$$
\text{label}(v_i) = \begin{cases} 
1, & x_i^s \geq \frac{1}{2} \\
0, & x_i^s < \frac{1}{2} 
\end{cases}
$$

(8)

**Experimental results**

In this section, the proposed method is evaluated on the MICCAI 2007 grand challenge data [17], which contains 20 3D abdominal CT scans with ground truth. Pixel spacing varies from 0.55 to 0.8 mm in x/y-direction, and slice distance varies from 1 to 3 mm [18]. Tumors, metastasis, and cysts with various sizes are included in most images. Transversal directions were acquired for CT scans with segmented livers. The number of slices in each scan varies between 64 and 394 with 512 × 512 resolution.

Rather than the whole CT scan, only a single slice that contains the largest liver is used as training data. Two scans randomly select in MICCAI 2007 grand challenge database serve as the test data. Eighteen slices contained in the remaining 18 scans are used as training data.

**Objective evaluation**

Five evaluations are implemented for assessing how the results of the proposed method $B$ will generalize to the ground truth $A$. The five evaluation measurements are generally defined as follows [19].

Volume overlap (VOE):

$$
\text{VOE} = 1 - \frac{\text{vol}(A \cap B)}{\text{vol}(A \cup B)}
$$

(9)

Relative volume difference (RVD):

$$
\text{RVD} = \frac{\text{vol}(A \setminus B)}{\text{vol}(B)}
$$

(10)

Average surface distance (ASD):

$$
\text{ASD}(A,B) = \left( \sum_{s_A \in S(A)} d(s_A, S(B)) + \sum_{s_B \in S(B)} d(s_B, S(A)) \right) / (|S(A)| + |S(B)|)
$$

(11)

Root mean squared error (RMSE):

$$
\text{RMSE}(A,B) = \sqrt{\left( \sum_{s_A \in S(A)} d^2(s_A, S(B)) + \sum_{s_B \in S(B)} d^2(s_B, S(A)) \right) / \sqrt{|S(A)| + |S(B)|}}
$$

(12)

Maximum surface distance (MSD):

$$
\text{MSD}(A,B) = \max \left\{ \max_{s_A \in S(A)} d^2(s_A, S(B)), \max_{s_B \in S(B)} d^2(s_B, S(A)) \right\}
$$

(13)
where $\text{vol}(\ast)$ denotes the volume of the region $\ast$, $S(\ast)$ is the surface voxel of the region $\ast$, $s_\ast$ is one of the voxels on the surface of the region $\ast$, and $d(s_{\ast 1}, S(s_{\ast 2})) = \min_{s_{\ast 2} \in S(s_{\ast 2})} \|s_{\ast 1} - s_{\ast 2}\|$ is the minimum Euclidean distance between corresponding voxels of two data surfaces. According to the above definition, VOE and RVD give the whole comparison for two segmented livers match, while ASD, RMSE and MSD indicate the difference of the pixel surface distance.

**Context iteration-based classification results**

To detect the validity of context iteration, two kinds of classifications are illustrated. One classification is obtained with GLCM features, and the other is achieved with both GLCM and context iteration. In Fig. 4d, the liver cannot be distinguished from muscles and other organs when only GLCM features are utilized. Context iteration considers both texture and context information, as shown in Fig. 4e–h. The significance of the liver is substantially improved, whereas the non-liver areas are weakened by adding the context feature to the texture feature.

The context iteration-based classifications of five slices are shown in Fig. 5. Figure 5a shows the original image. GLCM-based classification results are shown in Fig. 5b. The context iteration-based classification results for the first and fifth iteration are shown in Fig. 5c and d, respectively. Compared with the GLCM-based classification results shown in Fig. 5b1–b3, the probability difference between the liver and the background is large and the significance of the liver is substantially improved based on the five context iteration, as shown in Fig. 5d1–d3. The red and yellow boxes in Fig. 5a4, a5 indicate similar intensity between the liver and the kidney or vessels (red box), and between the liver and muscle (yellow box). The liver cannot be segmented using only GLCM features, as shown in Fig. 5b4, b5. After considering context information, the contrasts between the liver and the other organs are apparent as shown in the red and yellow boxes in Fig. 5d4, d5. The prior probability model effectively provides the shape information of the liver.
Experimental error analysis

The segmentation results in six slices are displayed in the first two lines of Fig. 6. Three segmentation results are compared using different colors, particularly, the ground truth (green curves), GLCM-based segmentation (red curves), and context iteration-based classification (blue curves). The contents in the yellow boxes are enlarged to illustrate the details, as shown in the last two lines of Fig. 6. The GLCM-based classification evidently resulted in over-segmented or under-segmented results because of the similar intensities between the livers and background. Context iteration-based classification can provide more precise segmentation than the GLCM-based classification. However, context iteration-based classification is inadequate in the corner, as shown in Fig. 6d, e. This
finding is caused by the decrease in prior probability. The proposed method improves the segmentation results for the conjunction of the liver and kidney/muscle/vessels.

Two different segmentation results of the test data are shown in Fig. 7. Figure 7a–c illustrate the 2D segmentation results on the coronal, vertical, and transverse sections. The red contour profile indicates the ground truth, and the green contour profile indicates the context-based iteration segmentation. The segmentation results accurately fit the liver edge, even in the concave region of the liver. The 3D segmentations of the ground truth and proposed methods are shown in Fig. 7d. Precise segmentation is obtained in the area indicated by the yellow arrow. Under-segmentation is produced in the area of the yellow box and should be improved in our future work.

Five evaluations, namely, VOE, RVD, ASD, RMSD and MSD, are implemented for assessing nine automatic segmentation methods (Table 1). The proposed method
achieves 7.83% VOE, 5.06% RVD, 1.06 mm ASD, 1.39 mm RMSD, 1.39 mm RMSD, and 11.12 mm MSD. The corresponding score for test data is obtained by [18]. The comparative results show that the proposed method is superior to eight other methods with regard to the total score and exhibits a good applicability for liver segmentation. The total score of the proposed method is 76, which is higher than the standard score provided by [18]. This result indicates that the proposed method satisfies the requirement of clinical liver segmentation.

**Discussion and conclusion**

In this paper, we developed machine learning techniques for the automatic liver segmentation in CT images. Appearance and context features are extracted for classifying image pixels based on AdaBoost algorithm. This classification provides the prior probability distribution map, which is combined with the original image to segment the liver.
through the improved random walk. Five evaluations, namely, VOE, RVD, ASD, RMSD and MSD, are implemented for assessing the efficiency of the proposed segmentation method in MICCAI 2007 grand challenge database with 20 abdominal CT images. Experimental results show that the proposed method achieves significantly more accurate results than the other eight state-of-art segmentation methods by comparing VOE, RVD, ASD, RMSD and MSD measurements. In addition, the total score of the proposed method is also higher than that of the ground truth which is manually labeled by medical experts.

The context features take full advantage of the context and high-level information. The intra-object configuration, and inter-object relationship are both considered to distinguish the organs with similar intensity. The saliency of liver is significantly improved compared to other organs.

However, several limitations of the proposed method still exist and should be improved in future work: (1) only an appearance is connected with context feature, which is somewhat limited and requires more explicit shape information; (2) after more features concatenated for expressing, the feature fusion method should be utilized to reduce the redundancy of features (3) the running time takes a long for the image analysis, because all pixels containing in the image have to be analyzed on the serial calculation. GPU accelerated algorithm should be developed in our future work.

**Abbreviations**

GLCM: gray level co-occurrence matrix; VOE: volume overlap; RVD: relative volume difference; ASD: average surface distance; RMSD: root mean squared error; MSD: maximum surface distance.

**Declarations**

**Author details**

1 Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing 100730, China. 2 Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Electronics, Beijing Institute of Technology, Beijing 100081, China. 3 Department of Radiology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing 100730, China.

**Authors’ contributions**

ZYC participated in the design of the study and drafted the manuscript. ADN suggested the algorithm for liver segmentation, implemented it and drafted the manuscript. MJR drew a part of figures and performed the image analysis. CWJ expressed opinions on the evaluation metric of the segmentation results. WX performed the acquisition of the CT images. ZHT expressed opinions on the overall framework of the manuscript. YJ conceived of the study, and participated in its design and coordination. All authors read and approved the final manuscript.

**Acknowledgements**

None.

**Competing interests**

The authors declare that they have no competing interests.

**Availability of data and materials**

All data analyzed during this study are download from MICCAI grand challenge on 3D segmentation (2007) http://mbi.dkfz-heidelberg.de/grand-challenge2007/.

**Funding**

National Basic Research Program of China (2013CB328806), PI: Jian Yang; the Key Projects in the National Science & Technology Pillar Program (2013BA01801), PI: Jian Yang; the National Hi-Tech Research and Development Program (2015AA042035), PI: Jian Yang, and the National Science Foundation Program of China (61433039, 61501030), PI: Danni Ai.

Received: 26 October 2016   Accepted: 8 December 2016

Published online: 14 January 2017
References

1. Chen Y, Shi L, Feng Q, Yang J, Shu H, Luo L, Coatrieux J-L, Chen W. Artifact suppressed dictionary learning for low-dose CT image processing. IEEE Trans Med Imaging. 2014;33(12):2271–92.

2. Chen Y, Yang Z, Yining H, Yang G, Luo L, Chen W, Toumoulin C. Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means. Phys Med Biol. 2012;57(9):2667–88.

3. Zhao Y, Liu Y, Wu X, Harding SP, Zheng Y. Retinal vessel segmentation: an efficient graph cut approach with retinaex and local phase. PLoS ONE. 2015. doi:10.1371/journal.pone.0122332.

4. Zhao Y, Rada L, Chen K, Harding S, Zheng Y. Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retina images. IEEE Trans Med Imaging. 2015;34(9):1797–807.

5. Peng J, Wang Y, Kong DX. Liver segmentation with constrained convex variational model. Pattern Recogn Lett. 2014;43(1):81–8. doi:10.1016/j.patrec.2013.07.010.

6. Affif A, Nakaguchi T. Liver segmentation approach using graph cuts and iteratively estimated shape and intensity constrains. In: Medical image computing & computer-assisted intervention. Miccai international conference on medical image computing & computer-assisted intervention. 2012. p. 395–403.

7. Chen X, Udupa JK, Bagci U, et al. Medical image segmentation by combining graph cuts and oriented active appearance models. IEEE Trans Image Process. 2012;21(4):2035–46. doi:10.1109/TIP.2012.2186306.

8. Peng J, et al. 3D liver segmentation using multiple region appearances and graph cuts. Med Phys. 2015;42(12):6840–52.

9. Zheng Y, Ai D, Zhang P, Gao Y, Xia L, Du S, Sang X, Yang J. Feature learning based random walk for liver segmentation. PLoS ONE. 2016;11(11):1–17.

10. Xu Y, Xu C, Kuang X, et al. 3D-SIFT-Flow for atlas-based CT liver image segmentation. Med Phys. 2016;43(5):2229–41.

11. Salman AlShaikhli SD, Yang MY, Rosenhahn B. 3D automatic liver segmentation using feature-constrained Mahalanobis distance in CT images. Biomedizinische Technik Biomed Eng. 2015;61(4):401–12.

12. He B, et al. Fast automatic 3D liver segmentation based on a three-level AdaBoost-guided active shape model. Med Phys. 2016;43(5):2421.

13. Soh LK, Tsatsoulis C. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE Trans Geosci Remote Sens. 1999;37(2):780–95.

14. Haralick RM, Shanmugam K, Dinstein IH. Textural features for image classification[J]. IEEE Trans Syst Man Cybern. 1973;6:610–21.

15. Grady L. Random walks for image segmentation. IEEE Trans Pattern Anal Mach Intell. 2006;28(11):1768–83.

16. Grady L. Multilabel random walker image segmentation using prior models. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR’05), 20–25 June 2005.

17. Heimann T, van Ginneken B, Styner M. MICCAI grand challenge on 3D segmentation. 2007. http://mbi.dkfz-heidelberg.de/grand-challenge2007/.

18. Heimann T, van Ginneken B, Styner MA, et al. Comparison and evaluation of methods for liver segmentation from CT datasets. IEEE Trans Med Imaging. 2009;28(8):1251–65.

19. Ségonne F, Dale AM, Busa E, Glessner M, Salat D, Hahn HK, Fischl B. A hybrid approach to the skull stripping problem in MRI. Neuroimage. 2004;22(3):1060–75.

20. Heimann T, Münzing S, Meinzer HP, Wolf I. A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation. Information processing in medical imaging. Berlin: Springer; 2007. p. 1–12.

21. Schmidt G, Athelogou MA, Schönmann M, Korn R, Binning G. Cognition network technology for a fully automated 3-D segmentation of liver. In: Proceedings of MICCAI workshop 3-D segmentation. Clinic: a grand challenge. 2007. p. 207−14.

22. van Rikxoort E, Arzhaeva Y, van Ginneken B. Automatic segmentation of the liver in computed tomography scans with voxel classification and atlas matching. In: Proceedings of MICCAI workshop 3-D segmentation. Clinic: a grand challenge. 2007. p. 101−8.