Semi-automatic Segmentation of 3D Liver Tumors from CT Scans Using Voxel Classification and Propagational Learning

Jiayin Zhou[1], Wei Xiong[2], Qi Tian[2], Yingyi Qi[1], Jiang Liu[2], Wee Keng Leow[3], Thazin Han[1], Sudhakar K Venkatesh[1], and Shih-chang Wang[1]

1. Department of Diganostic Radiology, National University of Singapore, Singapore
2. Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore
3. School of Computing, National University of Singapore, Singapore
dnrzjy@nus.edu.sg

Abstract. A semi-automatic scheme was developed for the segmentation of 3D liver tumors from computed tomography (CT) images. First a support vector machine (SVM) classifier was trained to extract tumor region from one single 2D slice in the intermediate part of a tumor by voxel classification. Then the extracted tumor contour, after some morphological operations, was projected to its neighboring slices for automated sampling, learning and further voxel classification in neighboring slices. This propagation procedure continued till all tumor-containing slices were processed. The method was tested using 3D CT images with 10 liver tumors and a set of quantitative measures were computed, resulted in an averaged overall performance score of 72.

1 Introduction

Tumor size is a primary measure of the severity of cancers and tumor volumetry is used both for cancer management and to assess the treatment response [1]. In addition, accurate lesion localization is a necessary step in some diagnostic and therapeutic procedures. Hence automatic, fast and robust tumor segmentation and quantitation is increasingly receiving attention and research efforts from medical imaging, computer vision and pattern recognition communities.

Computerized segmentation of liver tumor from computed tomography (CT) scans is a challenging task because researchers are facing a variety of difficult situations including tumor shape variations, low contrast between tumor tissue and normal liver tissue, presence of neighboring structures/organs with similar density values, different tumor imaging characteristics at different CT scanning phases (e.g., early phase, arterial phase, portal venous phase, and delayed phase), etc. Mahr et al. evaluated the usability of semiautomatic segmentation algorithms (region growing, isocontour, snake, etc.) for tumor volume determination in an organic human liver phantom [2]. Yim et al. performed the volumetry study for 10 hepatic metastatic lesions including a total of 36 CT slices, by using watershed and active contour algorithms [3]. Snake algorithm with its derivatives has also been studied for semi-automated segmentation

of primary liver tumor from CT images [4]. A multi-stage, automatic hepatic tumor segmentation scheme was proposed by Seo KS [5]. It included liver structure segmentation, hepatic vessels removal, and hepatic tumor extraction at last using composite hypotheses and minimal total probability error [5]. However some limitations exist in these studies: First, each method developed or adopted was only tested by CT images with one particular tumor type (either primary hepatocellular carcinoma or metastatic lesion); Second, most of the testing images are from conventional CT scans with slice thickness of 5-7 mm, thus most slices have high signal-noise-ratio (SNR) which is helpful in segmentation; Third, only a few quantitative measures were used to evaluate these methods and no comprehensive measuring and scoring system was employed to evaluate the overall performance for the quality of lesion segmentation and accuracy of tumor volumetry. Therefore further investigation on tumor segmentation algorithm, proper data for method benchmarking and performance evaluation metric is highly desired.

By exploring the training data and the segmentation references provided by the workshop [6] organizers, some image properties of our data can be summarized as follows: (1) The majority of tumors are isolated objects in 3D volume and their boundaries are closed contours in 2D slices while a few tumors have connections with adjacent lesions; (2) Even within a single tumor, variation of intensity values exists from the intermediate part to the periphery. However due to the thin slice thickness (1 mm or 1.5 mm), the transition between neighboring slices is slight; (3) Tumors in different pathological types (hepatocellular carcinoma or metastatic lesion) show different density values (hypo- or hyper-density); (4) In some cases, tumor region is not homogenous due to necrosis and vascularity inside the lesion; and (5) Generally images show quite low SNR due to the acquisition of thin slices. From the pattern recognition point of view, liver tumor segmentation can be considered as an object extraction mission that is to group pixels/voxels with particular features and discard others, under certain similarity criteria and some knowledge-based constrains. To handle the problems mentioned above, we propose a semi-automatic scheme which uses voxel classification and inter-slice propagation-based learning to segment liver tumors from three-dimensional (3D) CT scans. The details of this scheme as well as algorithms embedded are presented at Section 2. In Section 3, data for algorithm benchmarking with evaluation metric is described and results are given, followed by a discussion and conclusion in Section 4.

2 Method

The main idea of this semi-automatic scheme is to extract desired tumor region from one single 2D slice in the intermediate part of a tumor using supervised learning-based voxel classification, then the extracted tumor contour, after some morphological operations, is mapped to its neighboring slices for automated sampling, learning and further voxel classification in the neighboring slices. This scheme is constructed on the important basis that scan is performed on multi-sliced CT and the slice thickness is no more than 1.5 mm such that there are only slight changes in both structural shapes and densities among consecutive slices.

2.1 Overview

The proposed segmentation scheme, illustrated in Fig. 1, has three main stages:

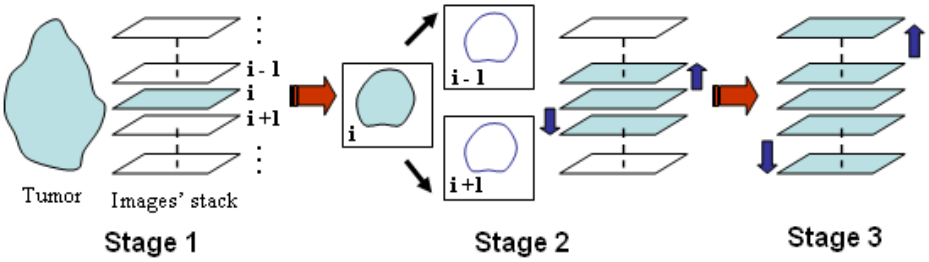

Fig. 1. Three main stages of the proposed segmentation scheme

Stage 1 Desired tumor region in 2D slice i at the intermediate part of the tumor is first segmented out by supervised learning-based voxel classification. In this procedure, tumor samples and non-tumor tissue (healthy liver tissue and other non-liver tissue) samples are manually selected to train a two-class support vector machine (SVM)-based classifier. Then the trained SVM classifier is imposed to a ROI in slice i for voxel classification such that tumor region in the ROI is extracted.

Stage 2 Let C_i be the contour of extracted tumor region from slice i, C_i^D and C_i^E be the contours after morphological dilation and erosion operations are performed on C_i, respectively. As slice thickness of 1 or 1.5 mm was employed for these data, it can be assumed with a high confidence level that the desired tumor contours and image features vary slightly among neighboring slices. Hence both C_i^D and C_i^E are projected to slices i-1 and i+1. In slices i-1 and i+1, voxels enclosed by C_i^E are used for heuristic learning to train the SVM classifier, then the newly trained SVM classifier is imposed to the area enclosed by C_i^D for tumor region extraction in the two propagating slices. This step is the first propagation procedure.

Stage 3 Similarly the whole propagation procedure including contour projection, SVM classifier training and voxel classification is further applied to upper and lower slices for tumor region extraction till all tumor-bearing slices are processed.

2.2 SVM classifier

SVM belongs to the supervised learning methods and is primary used for binary and also one-class and n-class classification problems [7]. It combines linear algorithms with linear or non-linear kernel functions that make it a powerful tool and gain success in the machine learning community with applications in computer vision, automation, data mining and biomedicine. In general for classification, SVM finds the best generalizing hyperplane with maximal margin separating the two classes. Since

an error-free separation is not always possible, the optimization problem can be modified such that misclassified training points are penalized. To be able to apply SVM into non-linear data distributions, the data can be implicitly transformed to a high-dimensional feature space where a separation might become possible.

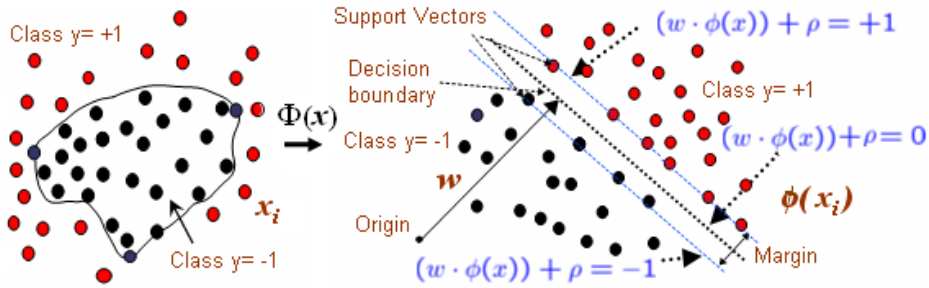

Fig. 2. Illustrator of SVM with kernel mapping for binary classification

For a binary classification, given a set of separable data set with N samples $X = \{x_i\}$, $i = 1, 2, \cdots, N$, labeled as $y_i = \pm 1$. It may be difficult to separate these two classes in the input space directly. Thus, they are mapped into a higher-dimensional feature space by $X^* = \Phi(x)$, as shown in Fig. 2. The decision function can be expressed as

$$f(x) = \mathbf{w} \cdot \mathbf{x} + \rho, \tag{1}$$

where $\mathbf{w} \cdot \mathbf{x} + \rho = 0$ $(\rho \in R,\ \mathbf{w} \in R^d)$ is a set of hyperplanes to separate the two classes in the new feature space. Therefore for all correctly classified data,

$$y_i f(x) = y_i (\mathbf{w} \cdot \mathbf{x} + \rho) > 0, \quad i = 1, 2, \cdots, N \tag{2}$$

holds. By scaling $\mathbf{w}$ and ρ properly, we can have $f(x) = \mathbf{w} \cdot \mathbf{x} + \rho = 1$ for those data labeled as +1 closest to the optimal hyperplane, and $f(x) = \mathbf{w} \cdot \mathbf{x} + \rho = -1$ for all the data labeled as –1 closest to the optimal hyperplane, as shown in Fig. 2. In order to maximize the margin, the following problem needs to be solved

$$min\left(\|\mathbf{w}\|^2 / 2\right), \tag{3}$$

subject to

$$y_i f(x) = y_i (\mathbf{w} \cdot \mathbf{x} + \rho) \geq 1, \quad i = 1, 2, \cdots, N. \tag{4}$$

It is a quadratic programming problem to maximize the margin, which can be solved using some standard algorithms such as sequential minimization optimization [8].

After optimization, the optimal separating hyperplane can be expressed as:

$$f(x) = \sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + \rho, \tag{5}$$

where $K(\cdot)$ is a kernel function, ρ is a bias, α_is are the solutions of the quadratic programming problem to find the maximum margin. There are only a few training

samples whose α_is are non-zero. They are called the support vectors, which are either on or near the separating hyperplane. The decision boundary, i.e. the separating hyperplane, is along these support vectors, whose decision values $f(x)$ (Eq. 5) approach zero. Compared with the support vectors, the decision values of positive samples have larger positive values and those of negative samples have larger negative values. Therefore, the magnitude of the decision value can also be regarded as the confidence of the classifier. The larger the magnitude of $f(x)$, the more confidence of the classification. By choosing different kernel functions, SVM can model the input space accurately and it shows good generalization by minimizing the errors made on the training set while maximizing the "margin" between different classes.

2.3 Implementation

In this study for voxel classification at each slice using the SVM classifier, voxel density (16 bits) and the median of the densities of the voxel's eight-neighbors in the same slice were used as input features. They represent the information for the voxel and its neighborhood, respectively. In addition, a Gaussian radius basis function (RBF) $K(\mathbf{x},\mathbf{y}) = \exp(\frac{-\|\mathbf{x}-\mathbf{y}\|^2}{\sigma})$ was adopted as the learning kernel in the SVM classifier used, where σ was decided by the variations of tumor samples. For the initial segmentation for slice i in Stage 1, tumor samples were selected by user's manual clicking, non-tumor samples were selected by pulling a rectangular box over the non-tumor region, and the ROI used for tumor region extraction was manually defined by a polygon. An example of initial segmentation is shown in Figs. 3.a and 3.b.

In stage 2 after the initial segmentation at slice i, tumor contour C_i in slice i was dilated/eroded by 2-3 pixels (2 pixels for 1 mm slice thickness and 3 pixels for 1.5 mm slice thickness) to C_i^D / C_i^E. Both C_i^D / C_i^E were projected to slices i-1 and i+1, as shown in Figs. 3.c and 3.d. C_i^D, which encircles almost the whole desired tumor region in slices i-1 and i+1 with some minor margin, was used as the new ROIs for tumor extraction in slices i-1 and i+1. The region enclosed by C_i^E, which is slightly inside the desired tumor region in slices i-1 and i+1 but covers the majority of the tumor, was used as the learning samples for tumor to train the new SVM classifier. In the training procedure, non-tumor samples were the same as those selected in the initial segmentation stage. Moreover, if the amount of tumor samples is more than twice of non-tumor samples, a random re-sampling would be performed on tumor samples to balance sample populations, which is important in the training of SVM classifier. After voxel classification for tumor region extraction, the whole propagation procedure including contour projection, SVM classifier training and voxel classification was further applied to upper and lower slices till all tumor-bearing slices were processed.

In theory, no human supervision was required during the propagation stages. However in this study, re-selecting tumor samples or re-defining the polygonal ROI

or both could be applied during the propagation in case C_i^D or C_i^E gave obviously undesired ROI or tumor samples. Furthermore, an automated morphological processing procedure was applied after voxel classification in each slice to improve the segmentation results.

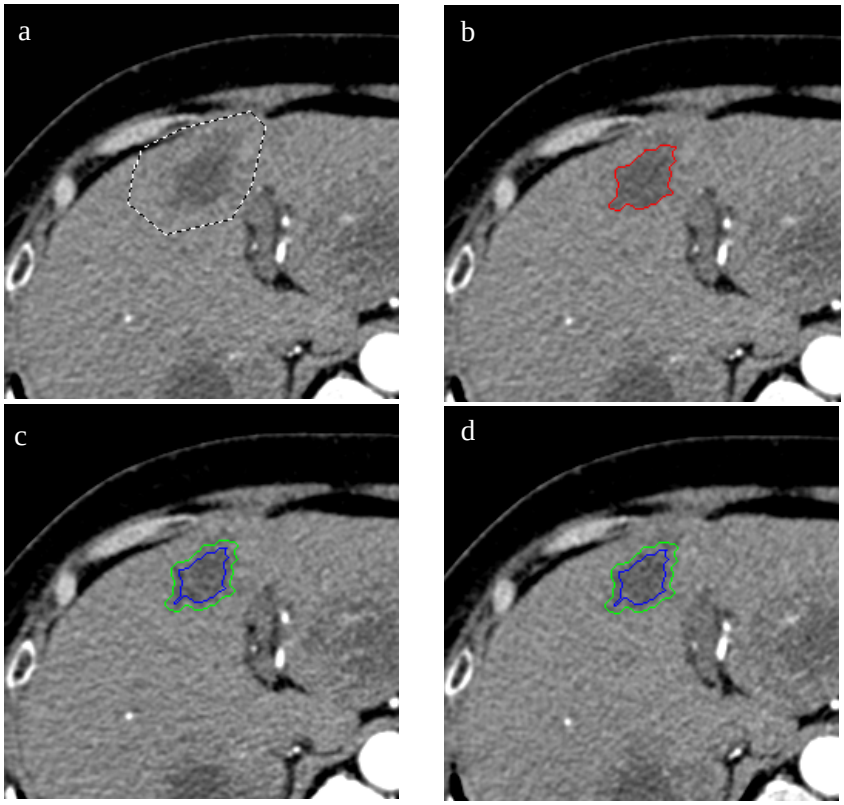

Fig. 3. (a) Slice 149 with ROI in tumor IMG02_L1; (b) segmented tumor contour C_i (red); (c)-(d) C_i^D (green) and C_i^E (blue) projections in slices 148 and 150.

3 Testing Data & Results

Also provided by the workshop organizers, the liver tumor CT image data for this work was acquired on one 64-slice and two 40-slice CT scanners using a standard four-phase contrast enhanced imaging protocol with slice thickness of 1mm or 1.5mm and in-plane resolution of 0.6-0.9 mm. The data are composed of 30 liver tumors, representing a range of patients and pathology. All tumors were manually segmented, by an experienced radiologist and confirmed by another radiologist, as reference for evaluation purposes. Data from 10 liver tumors were used for method training and parameters tuning, another 10 tumors for testing (results presented in this paper) and

the last 10 tumors for the onsite competition. Segmentation program was performed using MATLAB 7.0 R14, on an Intel Core 2 2.66 GHz PC workstation with 3 GB RAM. The computational time for each tumor varied from 7 mins for a tumor with 14 slices involved to around 20 mins for a tumor with 70 slices involved.

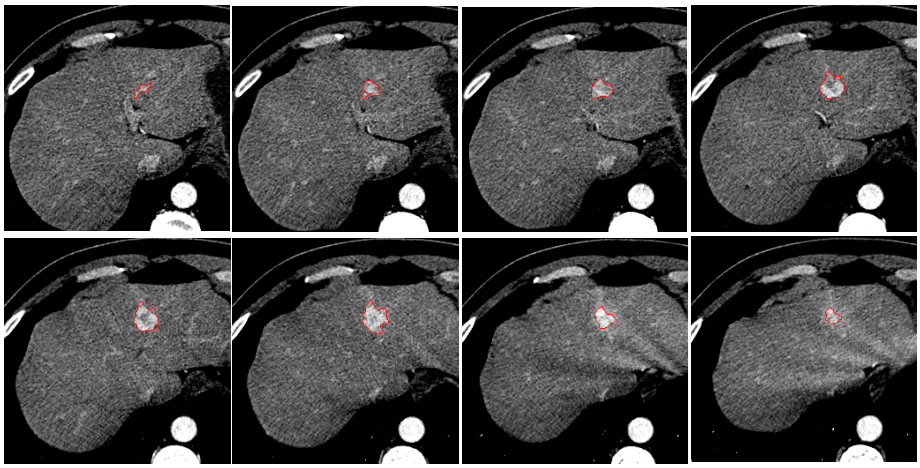

Fig. 4. Segmentation results (tumor contours in red) of selected slices from tumor IMG06_L1

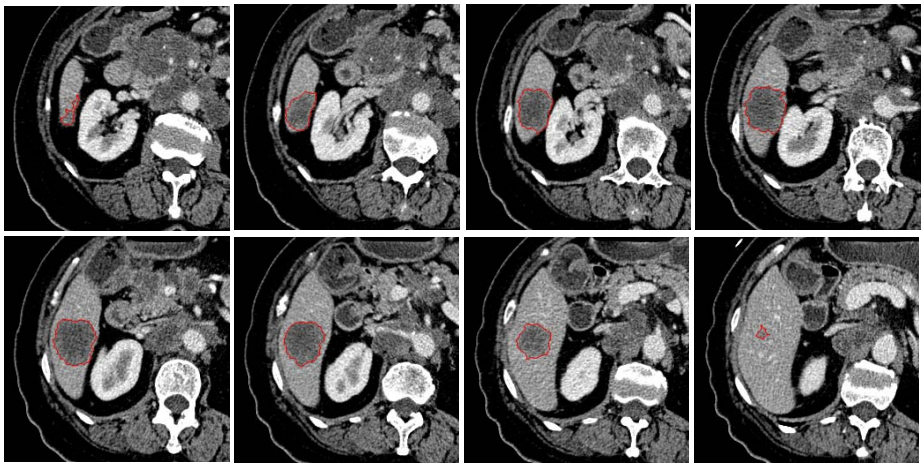

Fig. 5. Segmentation results (tumor contours in red) of selected slices from tumor IMG10_L1

Figures 4 and 5 show the visual examples of segmentation results. By visual inspection, we find that in general boundaries of tumors are detected accurately for each slice. Results were also evaluated quantitatively with respect to the following five measures, (1) relatively absolute volume differences, (2) average symmetric absolute surface distance, (3) symmetric RMS surface distance, (4) maximum symmetric absolute surface distance, and (5) volumetric overlap error. An overall score from 0 to 100 was assigned by an automatic scoring system using results of the

above mentioned five measures [9]. A score of 100 represents the perfect segmentation while 0 is the minimum score one segmentation will get. The evaluation results are summarized in Table 1.

Table 1. Results of the evaluation metrics and scores for all ten testing tumors.

Tumor	Overlap Error		Volume Diff.		Avg. Dist.		RMS Dist.		Max. Dist.		Total Score
	[%]	Score	[%]	Score	[mm]	Score	[mm]	Score	[mm]	Score	
IMG05_L1	30.65	76	10.78	89	2.19	45	2.95	59	11.62	71	68
IMG05_L2	32.60	75	12.10	87	1.07	73	1.51	79	6.36	84	80
IMG05_L3	35.39	73	15.41	84	1.44	64	2.08	71	9.69	76	73
IMG06_L1	25.78	80	7.98	92	0.59	85	0.85	88	4.24	89	87
IMG06_L2	64.56	50	96.93	0	2.42	39	3.13	56	12.78	68	43
IMG07_L1	18.67	86	0.35	100	2.21	44	3.25	55	18.45	54	68
IMG07_L2	24.70	81	1.01	99	1.12	72	1.66	77	8.46	79	81
IMG08_L1	17.30	87	9.61	90	1.91	52	2.71	62	14.33	64	71
IMG09_L1	30.30	77	23.20	76	1.20	70	2.11	71	12.50	69	72
IMG10_L1	20.22	84	15.73	84	1.03	74	1.53	79	6.73	83	81
Average	30.02	77	19.31	80	1.52	62	2.18	70	10.52	74	72

4 Discussion & Conclusions

For years research efforts have been put into the development of an accurate, robust and reliable liver tumor volumetry system. In this paper a method which uses machine learning-based voxel classification and shape-based inter-slice propagation is presented for the semi-automatic segmentation of liver tumor from CT data. This method was tested using 3D CT images with 10 liver tumors and a set of quantitative measures were computed, resulted in an averaged overall performance score of 72.

We can observe that among the 10 total scores achieved by testing ten tumor data, the one from tumor IMG06_L2 is the lowest (43 only) and it is also significantly lower than other nine scores (mean, 76; range, 68-87). In this case, the high segmentation overlap error (64.56%) and relative volume difference (96.93%) suggest that our segmented tumor volume is almost the double of the reference segmentation volume. In addition, it only matches one third part of the reference segmentation such that two thirds of the segmented volume belongs to false positives. If we explore data IMG06_L2 we can find that in this arterial phase image, the tumor is a hypo-density focal lesion enclosed by an irregular hyper-density ring, as two slices shown in Fig. 6 a-b. In pathology, it represents a lesion of primary hepatocellular carcinoma (HCC, hypo-density) enclosed by small arteries (hyper-density) for blood supply. Our segmentation results for the corresponding slices, shown in Fig. 6 c-d, include both the hypo-density focal lesion and the hyper-density arteries. The inclusion of the surrounding arteries into the HCC lesion as a part of the total tumor volume was suggested by our consultant radiologists, however for the reference segmentation

given by other experts, surrounding arteries might be probably excluded from the total tumor volume. Hence the poor performance of our method on tumor IMG06_L2 may probably due to the different understandings and definitions from different clinical experts, for the total tumor volume in such cases. On the other hand, the results show the non-linear separation ability of SVM classifier that hypo-density focal lesion and hyper-density arteries are successfully separated from normal liver tissue which is with intermediate density.

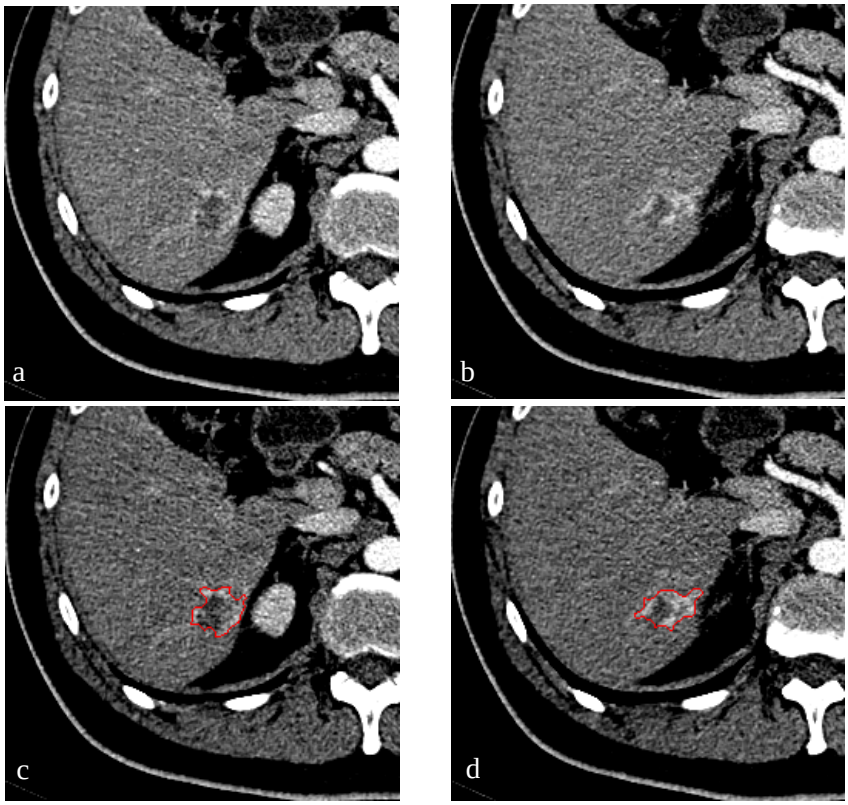

Fig. 6. (a) Slice 109 of tumor IMG06_L2; (b) Slice 112 of the same tumor; (c)-(d) Corresponding segmentation contours (red) for the two slices.

A limitation of the proposed method is that although 3D information was used in this study to determine the sampling and classification areas, the actual voxel classification was still performed on the basis of 2D slice. Therefore the segmented tumor object may not keep good smoothness and continuity in 3D in this direction, though the extracted tumor contour in 2D can reflect tiny and subtle changes along the tumorous/healthy tissue interface. Abrupt sticks and concavo-convex curve can be seen on the contours even after the usage of morphological operations. Manual tracing by human experts, on the other hand, trends to give smooth contours but without many subtle details, due to the insufficient spatial resolution of the manual tracking devices we use, i.e., mouse and touch-screen pen. This may also partially explain why

among the five measures, the “average symmetric absolute surface distance” got the lowest average score (62) among our segmentation results.

In general, the promising benchmarking results achieved demonstrate that our learning-classification-propagation strategy, with the help from knowledge-based constraints on tumor shape and location, succeeds generally. Further improvement on liver tumor segmentation techniques is of course still possible. Besides that 3D shape prior constraints can be better utilized to keep the smoothness and continuity of the segmentation result in 3D, normalized cut (an extension version of graph cut [10] with spectral clustering approach) also has high potential to be explored in this area.

Acknowledgements

This work is supported by a research grant (SBIC RP C-008/2006) from the Singapore BioImaging Consortium, Agency for Science, Technology and Research.

References

1. Quivey, J.M., Castro, J.R., Chen, G.T., Moss, A., Marks, W.M.: Computerized tomography in the quantitative assessment of tumour response. Br. J. Cancer Suppl., 4, 30--34 (1980)
2. Mahr, A., Levegrün, S, Bahner, M.L., Kress, J., Zuna, J., Schlegel W.: Usability of semiautomatic segmentation algorithm for tumor volume determination. Invest. Radiol. 34, 143--150 (1999)
3. Yim, P.J., Foran, D.J.: Volumetry of hepatic metastases in computed tomography using the watershed and active contour algorithms. In: Proceedings of the 16th IEEE Symposium on Computer-based Medical Systems, pp. 329--335. (2003)
4. Lu, R., Marziliano, P., Thng, C.H.: Liver tumor volume estimation by semi-automatic segmentation method. In: Proceedings of the 27th Annual Conference of IEEE Engineering in Medicine and Biology, pp. 3296--3299. (2005)
5. Seo, K.S. Automatic hepatic tumor segmentation using composite hypotheses. In: Kamel, M., Campilho A. (Eds.) ICIAR 2005. LNCS, vol. 3656, pp. 922--929. Springer, Berlin (2005)
6. 3D Segmentation in the Clinic: A Grand Challenge II, MICCAI 2008 Workshop. URL: http://grand-challenge2008.bigr.nl/
7. Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. University of Cambridge, Cambridge (2000)
8. Schölkopf, B., Smola, A.J. Learning with Kernels Support Vector Machines: Regularization, Optimization and Beyond. MIT, Cambridge (2002)
9. Gerig, M., Chakos, M.V.: A new validation tool for assessing and improving 3D object segmentation. In: Niessen, W.J., Viergever, M.A. (Eds.) MICCAI 2001. LNCS, vol. 2208, pp. 516--523. Springer, Berlin (2001)
10. Esneault, S., Hraiech, N., Delabrousse, É., Dillenseger, J.L.: Graph cut liver segmentation for interstitial ultrasound therapy. In: Proceedings of the 29th Annual Conference of IEEE Engineering in Medicine and Biology, pp. 5247--5250. (2007)