

Liver Segmentation with Semi-Supervised Learning

Yonghui Gao^{1,a}, Xiaoxiao Li^{2,b} and Jingjing Liu³

¹School of Medical Instrument and Food Engineering,
University of Shanghai for Science and Technology, Shanghai, China

²College of Sciences, Shanghai Institute of Technology, Shanghai, China

^aemail: gaoyonghui1978@163.com, ^bemail: xiaoxiao_li@sina.com

Keywords: Liver Segmentation; Interactive Method; Maximal Similarity; Blocks Merging; Semi-supervised Learning

Abstract. Efficient liver segmentation from volume data provides important assistance for minimal invasive surgery and treatment. However, this task suffers from the special anatomy and topological changes. This paper presents a robust interactive method, which treats it as a semi-supervised learning task. An initial classification is performed to partition the volume data into homogeneous blocks to guide the segmentation. It is easy to implement and a more general linear or nonlinear model can be formed by virtue of semi-supervised learning. Experimental results demonstrate the performance of the proposed scheme in liver contours extracting.

Introduction

Traditional surgical planning of liver uses volumetric information stored in a stack of intensity-based images usually from computerized tomography (CT) and MRI scanners [1]. Based on a number of image slices, surgeons build 3D model of liver, tumor, and vasculature. Quickly and accurately digitized contour model is the basic premise of the surgical simulation and biomechanical evaluation in liver minimal invasive surgery. To move, arrange and remove individual part in a 3D model, each object contour need to be accurately extracted [2]. This task is difficult, even for experienced surgeons. Moreover, tumors can create anatomical variability, which further complicates the situation in volumetric data sets. As a consequence, surgeons can miss important information or draw incorrect conclusions, which can lead to suboptimal treatment strategy decisions. Because complete liver 3D structure is crucial in computer-aided diagnostic [3–4], volume segmentation has been intensively applied with guaranteed results [5].

To deal with the liver segmentation problem, numerous algorithms have been investigated [6]. Commonly, a proper shape or intensity prior is beneficial to separate individual object and create virtual boundaries from surrounding tissues. Level set method maybe a nice solution for this complex problem in theory. However, the scheme is very sensitive to the initialization and suffers from the convergence to the local minima. Gao and Chae [7] proposed an active contour tracking algorithm to separate the touching tissues and adaptively deal with the topological changes, which can obtain individual model with higher accuracy but needed more manual interventions. Gang Chen and Li xu Gu [8] proposed a novel multiple-initialization LSM approach to overcome the leakage and over-segmentation problems in segmenting the liver region from MRIs. Xing Zhang and Jie Tian [9] present an approach for automatic liver segmentation from computed tomography (CT) scans that is based on a statistical shape model (SSM) integrated with an optimal-surface-detection strategy, but requires a multi-dimension parameter space and high computational cost. Because contour extracting in liver image suffers from topological changes, this fact makes liver segmentation is a challenging issue [10].

A novel semi-automatic algorithm is proposed to efficiently achieve the desired contours through a simple strategy. Firstly, the volume data is partitioned into homogeneous blocks by 3D mean shift classification [11]. Then, user interaction is performed in the selected MPR (Multi-Planar Reformation) image planes to indicate the desired object and background, i.e. initialize the labeled

and unlabeled instances as the inputs of subsequent semi-supervised learning (SSL). The object contours can be extracted by labeling all the blocks as either object or background (labeled as 1 and 0).

The rest of the paper is organized as follows. Section 2 lists the previous works. Section 3 presents the proposed semi-supervised learning based algorithm. Section 4 performs the experiment to verify the proposed method. Section 5 concludes the paper.

Previous Works

The efficiency of user interaction is important to guarantee the precision. A Cellular Automaton (CA) based method [12] has the merits of few interventions, which can balance the contradiction of precision and computing time. Another Maximal Similarity based Region Merging (MSRM) segmentation [13] presents a more robust interactive method. According to MSRM, the scheme is divided into two stages. In the first stage, MSRM only merges the marker background regions. For each region $B \in M_B$ and its adjacent region A_i that $A_i \in \overline{S_B}, \overline{S_B} = \{A_i\}_{i=1,2,\dots,r}$ are merged into one region if

$$\rho(A_i, B) = \max_{j=1,2,\dots,k} \rho(A_i, S_j^{A_i}), A_i \notin M_B \quad (1)$$

and the remained non-marker regions are merged in the second stage. For each non-marker region $P \in N$ and its adjacent region H_i that $H_i \in \overline{S_P}, \overline{S_P} = \{H_i\}_{i=1,2,\dots,p}, \overline{S_{H_i}} = \{S_j^{H_i}\}_{j=1,2,\dots,p}$ are merged into one region if

$$\rho(P, H_i) = \max_{j=1,2,\dots,k} \rho(H_i, S_j^{H_i}), H_i \notin M_B, H_i \notin M_O \quad (2)$$

where ρ is the similarity measurement. M_O and M_B are the marker regions labeled as object and background respectively. N is the non-marker region set. $\overline{S_Q} = \{S_i^Q\}_{i=1,2,\dots,q}$ is the formed set of Q 's adjacent regions. This two-stage method suffers from the complex implementation. Inspiring by the proposed SSL based method for digital matting problem [14], if the initially labeled objects are used as the input of SSL, the rest objects will be labeled automatically through solving a quadratic optimization problem.

Semi-Supervised Based Liver Segmentation

In liver segmentation, a proper initial classification could lead to a good convergence to the true object contour [15], 3D mean shift is firstly used to partition the volume data into homogeneous blocks for subsequent interaction. We mark the liver regions on 2D MPR image planes to indicate desired 3D objects. Blocks merging is performed with a simple CA rule [16] based on maximal similarity to produce initial labeled and unlabeled instances. Then, the rest parts will be extracted automatically by the proposed semi-supervised learning mechanism.

Liver Marking and Similarity Based Blocks Merging

Let M and \overline{M} denote the sets of marker and non-marker blocks respectively. To guide the merging process, we use Bhattacharyya coefficient [17] as the similarity descriptor. Let q_i be an adjacent block of p and $N_p = \{q_i\}_{i=1,2,\dots,k}$ denotes the set of p 's adjacent blocks of 26 Moore neighborhood. The Bhattacharyya coefficient is defined as:

$$\theta(p, q_i) = \sum \sqrt{Hist_p \cdot Hist_{q_i}} \quad (3)$$

where $Hist_p$ and $Hist_{q_i}$ are the normalized grayscale histograms of p and q_i respectively, and the higher Bhattacharyya coefficient indicates the higher similarity between them. The proposed merging scheme can be described as follows:

$$\begin{aligned}
& \text{while}(\phi \subset \overline{M}^t) \\
& \{ \\
& \quad M^{t+1} = M^t, \overline{M}^{t+1} = \overline{M}^t \\
& \quad \forall p \in M^t, \forall q_i \in N_p \\
& \quad \text{if } \theta(p, q_i) = \max_{i=1,2,\dots,k} \theta(p, N_p) \\
& \quad \quad M^{t+1} = M^t \cup q_i, \overline{M}^{t+1} = \overline{M}^t \setminus q_i \\
& \quad \text{end if} \\
& \}
\end{aligned} \tag{4}$$

where M^t and \overline{M}^t are the sets of marker and non-marker blocks at t time respectively. The merging procedure stops when the entire \overline{M} shrinks to empty. This strategy merges similar blocks as many as possible. After initial merging, there is still numbers of blocks wrongly labeled. The next step is to adjust this process by a SSL based mechanism so that the desired objects can be correctly extracted. Fig. 1 shows a view of interaction in 2D MPR image. We use green and blue markers to differentiate object and background. The red lines indicate edges between the homogeneous blocks pre-classified by the 3D mean shift. The proposed merging method can efficiently separate desired object contours.

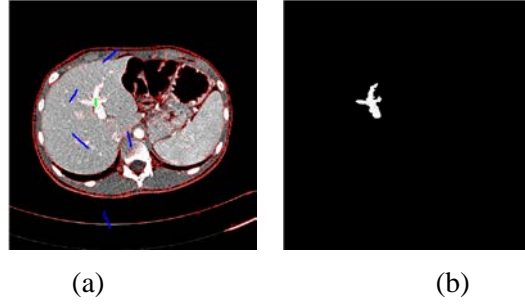


Fig.1. User interaction and merged blocks. (a) Initial classification and user markers in 2D MPR image (Green and blue markers indicate the object and background respectively). (b) 2D view of merged blocks by the proposed method.

Semi-Supervised Segmentation

We treat each block $i \in \Omega$ as a data point denoted by $x_i \in \mathbb{R}^d$. x_i can be set as I_i or other features extracted for block i . In this paper, we set $x_i = I_i$ which is a scale value ($d=1$). The learning problem can be then formulated as follows. Given a data point set $X \subset \mathbb{R}^d$, $X = \{x_i\}_{i \in \Omega}$, the label values $\{l_i\}_{i \in X_l}$ of the labeled data points $X_l = \{x_i\}_{i \in \Omega_l}$, our goal is to compute the accurate label values $\{l_i\}_{i \in X_u}$ of the unlabeled data points $X_u = X - X_l$ through learning methods.

For each block $i \in \Omega$ and its neighboring blocks denoted by $N_i = \{\tau_1, \dots, \tau_m\}$, we assume that its label value l_i can be predicted by a linear combination of the label values $\{l_j\}_{j \in N_i}$ of $N_i \subset \Omega$. If use $\Gamma_i = \{l_{\tau_1}, \dots, l_{\tau_m}\}^T$ where $\tau_i \in N_i$ to denote the vector of label values of N_i , and $f_i = \{f_{i\tau_1}, \dots, f_{i\tau_m}\}^T$ to denote the vector of the linear combination coefficients, the combination for i can be represented by

$$l_i = f_i^T \Gamma_i. \tag{5}$$

We select the blocks in a 26 Moore neighborhood centered at i as the neighbors ($m=26$). The computation of f_i is accomplished by the local learning process, which describes the dependencies between data points $\{x_j\}_{j \in N_i}$ in \mathbb{R}^d and $\{l_j\}_{j \in N_i}$ in \mathbb{R} . f_i can then be used to predict l_i from x_i .

For a data vector $x \in X$, we denote $x' = \begin{bmatrix} x^T & 1 \end{bmatrix}^T$. The local learning can be described as a linear model:

$$l = x^T \beta + \beta_0 = x^T \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix}. \quad (6)$$

where $\beta = [\beta_1, \dots, \beta_d]^T$ and β_0 are the model coefficients. We denote the notation: $X_i = [x_{i1}, \dots, x_{im}]^T$ which is a matrix of size $m \times (d+1)$ and is stacked by the data values of the blocks in N_i . With the ridge regression technique [18], we can estimate β and β_0 by solving a quadratic optimization problem:

$$\arg \min_{\beta, \beta_0} \left\| l_i - X_i \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix} \right\|^2 + \lambda_r \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix}^T \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix}. \quad (7)$$

The optimal solution of (7) can be easily derived as

$$\begin{bmatrix} \hat{\beta} \\ \hat{\beta}_0 \end{bmatrix} = (X_i^T X_i + \lambda_r I_{(d+1)})^{-1} X_i^T a_i = X_i^T (X_i X_i^T + \lambda_r I_{(m)})^{-1} a_i. \quad (8)$$

where $I_{(\cdot)}$ is the corresponding identity matrix.

Substituting (8) into (6), we can finally get f_i of (5) as

$$f_i = (X_i X_i^T + \lambda_r I_{(m)})^{-1} X_i x_i. \quad (9)$$

which is free from $\{l_j\}_{j \in N_i}$ and only relates to x_i and $\{x_j\}_{j \in N_i}$. λ_r is a parameter for which we set 0.1.

In (5), l_i can also be rewritten in the form of the linear combination of label values of all the blocks. We denote the label values with the vector $\Gamma = [l_1, \dots, l_n]^T$ where the coefficients with $\xi_i = [f_{i1}, \dots, f_{in}]^T$. For ξ_i , the values for the blocks in N_i are equal to the corresponding ones in f_i , and the remaining is zero. We have

$$l_i = \xi_i^T \Gamma. \quad (10)$$

By introducing a new matrix F through stacking $\{\xi_i\}_{i \in \Omega} : F = [\xi_1, \dots, \xi_n]$, we can rewrite (10) as

$$\Gamma = F^T \Gamma. \quad (11)$$

If we know F , a classical way to estimate Γ is to minimize the following quadratic cost

$$\arg \min_{\alpha} \left\| \Gamma - F^T \Gamma \right\|^2 + c \left\| \Gamma_l - \Gamma_l^* \right\|. \quad (12)$$

where Γ_l denotes the vector of the label variables of the labeled blocks in Ω_l and Γ_l^* denotes the vectors of the already known label values of the labeled blocks. Here we set $c = \infty$, which forces the label value to be 1 for the labeled object blocks and 0 for background.

In order to solve (12), we need to reformulate it. We introduce the diagonal matrix C of size $n \times n$, for which the j th diagonal element takes the constant value c if $j \in \Omega_l$, and other diagonal elements are zero. We also bring in the vector Γ_l^* of length n , for which the j th element equals the already known label value of block j if $j \in \Omega_l$. Then, (12) is reformulated as

$$\arg \min_{\alpha \in R^n} \Gamma^T (I_{(n)} - F)(I_{(n)} - F)^T \Gamma + (\Gamma - \Gamma_l^*)^T C (\Gamma - \Gamma_l^*) \quad (13)$$

where $I_{(n)}$ is the $n \times n$ identity matrix.

By taking the first derivative of Γ in (13) and setting it to zero, we get the solution

$$\Gamma = ((I_{(n)} - F)(I_{(n)} - F)^T + C)^{-1} + C \Gamma_l^* \quad (14)$$

Equation (14) can be computed if the linear combination coefficients in (5) are known. Then we can label all blocks by the sign

$$\Gamma(i) = \begin{cases} 1, & \text{if } \Gamma(i) > 0.5 \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

It is worth noting that the initial homogeneous blocks merging is crucial for subsequent learning processing, which can efficiently reduce the dimensions of computing matrix. The proposed segmentation method can be described as a general cluster-then-label algorithm of semi-supervised classification.

Algorithm : Cluster-then-Label

Input: original voxels data $V_{i \in \Omega}$, an unsupervised clustering algorithm A , labeled blocks $(b_1, y_1), \dots, (b_l, y_l)$, unlabeled data b_{l+1}, \dots, b_{l+u} , and a semi-supervised learning algorithm SSL .

Output: labels on unlabeled blocks y_{l+1}, \dots, y_{l+u} .

1. Clustered blocks b_1, \dots, b_{l+u} using A .
2. For each resulting clustered block, let S be the labeled instances. Learn a semi-supervised predictor from $S: f_s = SSL(S)$. Apply f_s to all unlabeled instances.

In step 1, the clustering algorithm A is unsupervised. In step 2, we learn one predictor using the labeled instances to apply for the unlabeled instances. This cluster-then-label algorithm does not necessarily involve a probabilistic mixture model.

Experimental Results

The proposed method is essentially an adaptive segmentation method. With the simple markers, all the blocks can be automatically labeled as object or background. We first evaluate the proposed method with a series of CT sequences. Then, we compare it quantitatively with the classical MSRM algorithm. The software is designed based on VC++ 9.0 and implemented on an Intel Xeon E5620 2.4 GHz CPU with 24GB RAM.

Experimental Analysis of the Proposed Method

Fig. 2 demonstrates a CT sequence (512 by 512 by 320) with an associate resolution of [0.6, 0.6, 0.6] mm. Fig. 2 (a)-(c) show the interaction in MPR planes; the rest images show the corresponding 2D view of segmentation results with SSL. After the initial classification, we input some interactive information: the green marker represents the liver artery object while the blue markers represent the background in the selected axial, sagittal and coronal image planes. The marker blocks cover only part but representative features of desired artery objects.

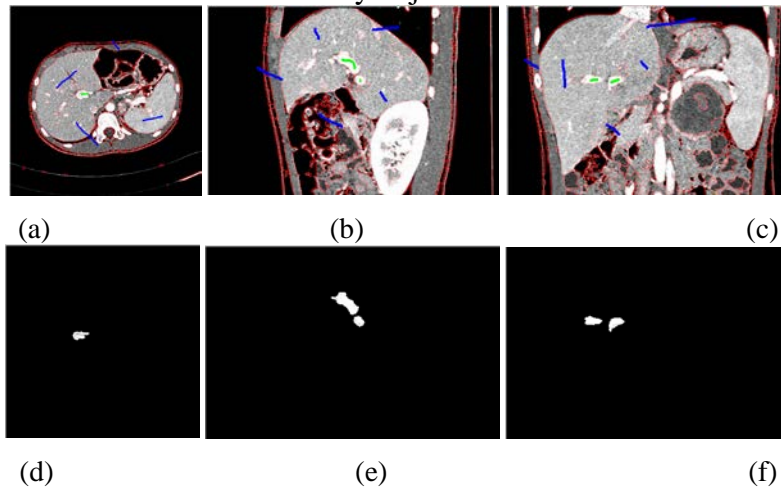


Fig.2. Examples of initial classification and interaction in MPR planes. (a) - (c) Selected axial, sagittal and coronal image planes. (d) - (f) Segmented results in MPR planes.

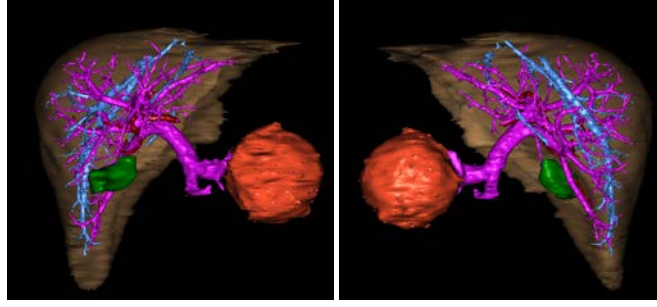


Fig.3. Liver segmentation results. The object contours are rendered as colored mesh models.

The total computing time of the proposed SSL method takes about 4.25 seconds on the whole CT data. We use multiple label values on the homogeneous blocks to extract multiple liver objects. The resulting mesh models are listed in Fig. 3. Another important issue is that the local resulting errors can be re-edited by additional interaction. In realistic systems, user can set data boundary or bounding box to reduce computing time.

Comparison with MSRM

In this section, we compare the SSL algorithm with the MSRM segmentation method [13] using the same input markers. For a fair comparison, we extend the original region based MSRM to a block based version (denote by BMSRM), i.e. the data points in MSRM are mean shift classified blocks instead of the original regions. The total evolution time of BMSRM takes about 27.28 seconds on the same data. The extracted models with SSL and BMSRM are listed in Fig. 4. Fig. 4 (b) demonstrates the difference between BMSRM and the proposed SSL algorithm. We can see that BMSRM is easy to fail in some important objects and wrongly label some background blocks as object (highlighted by green profile). A block is wrongly labeled as liver object, that dues to BMSRM only considers local neighborhood similarity while SSL can achieve a global minimal deviation optimization.

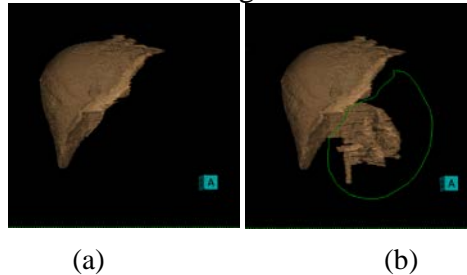


Fig.4. Comparisons of SSL and BMSRM. (a) Segmentation result with SSL. (b) Segmentation result with BMSRM. Segmentation difference is marked out with green profile.

To quantitatively compare these two methods, we manually label the desired objects in the CT slices and take them as ground truth. Then, we compute the true positive rate (TPR) and false positive rate (FPR) for these segmentation results [13]. The TPR is defined as the ratio of the number of correctly classified object pixels to the number of total object pixels in the ground truth, and the FPR is defined as the ratio of the number of background pixels but classified as object pixels to the number of background pixels in the ground truth. Obviously, the higher TPR and lower FPR indicate better efficiency. Table I lists comparison of these two methods on CT data in Fig. 4. The numeral results imply that SSL achieves better segmentation performance.

Table 1. Comparison of SSL and BMSRM

Method	Time(s)	TPR (%)	FPR (%)
SSL	4.25	96.31	2.18
BMSRM	27.28	91.20	7.64

We also perform the validation in numbers of MRI sequences and get the similar results. This validation demonstrates that by grouping the similar pixels into homogenous blocks, mean shift initial classification improve the robustness of SSL to image noises and small pixel variations.

Commonly, the proposed SSL method is not sensitive to the initial classification, this fact makes it is more robust and efficient in image segmentation tasks. Although the proposed algorithm can greatly reduce the requirements of user interaction, but more manually fine markers are still crucial in region with complex structures, which dues to this algorithm only considers the intensity property in liver characteristics. Subsequent work will combine other machine learning method such as active learning to automatically select the confusing instances for manually marking. The liver artery shape and arrangement priors should be considered to improve the precision of segmentation further.

Conclusion

A novel interactive method based on semi-supervised learning is proposed to resolve liver segmentation. Volume data is firstly clustered into homogeneous blocks, and then an iterative merging scheme is performed to produce the initial labeled and unlabeled instances for subsequent semi-supervised learning. The proposed maximal similarity based merging mechanism can efficiently reduce the dimensions of computing matrixes of semi-supervised learning. With simple matrix operations, problem-solving is easy to implement. The proposed method provides a general framework for interactive segmentation, which does not depend essentially on mean shift, and other unsupervised methods can also be used in initial classification. Furthermore, we will explore how to improve the efficiency with parallel computing.

Acknowledgement

Our research is supported by Qingdao Hisense Medical Equipment Co., Ltd. in China. The proposed algorithm and experimental CT dataset are verified with Hisense CAS® system.

References

- [1] Dirk Selle, Bernhard Preim, Andrea Schenk, and Heinz-Otto Peitgen, "Analysis of Vasculature for Liver Surgical Planning," *IEEE Trans. Med. Imaging [J]*, vol. 21, no. 11, pp. 355-369, 2002.
- [2] Marco Feuerstein, Thomas Mussack, Sandro M. Heining, and Nassir Navab, "Intraoperative Laparoscope Augmentation for Port Placement and Resection Planning in Minimally Invasive Liver Resection," *IEEE Trans. Med. Imaging [J]*, vol. 27, no. 3, pp. 355-369, 2008.
- [3] Miltiades Gletsos, Stavroula G. Mougiakakou and etc, "A Computer-Aided Diagnostic System to Characterize CT Focal Liver Lesions: Design and Optimization of a Neural Network Classifier," *IEEE Transactions on Information Technology in Biomedicine [J]*, vol. 7, no. 3, pp. 153-162, 2003.
- [4] Hongjian Shi, Aly A. Farag, Rachid Fahmi, and Dongqing Chen, "Validation of Finite Element Models of Liver Tissue Using Micro-CT," *IEEE Transactions on Biomedical Engineering [J]*, vol. 55, no. 3, pp. 978-984, 2008.
- [5] Muriel Mescam, Marek Kretowski, and Johanne Bézy-Wendling, "Multiscale Model of Liver DCE-MRI Towards a Better Understanding of Tumor Complexity," *IEEE Trans. Med. Imaging [J]*, vol. 29, no. 3, pp. 699-707, 2010.
- [6] Tobias Heimann, Bram van Ginneken and etc, "Comparison and Evaluation of Methods for Liver Segmentation From CT Datasets," *IEEE Trans. Med. Imaging [J]*, vol. 28, no. 8, pp. 1251-1265, 2009.
- [7] Hui Gao, Oksam Chae, "Individual tooth segmentation from CT images using level set method with shape and intensity prior," *Pattern Recognition [J]*, vol.43, no. 7, pp. 2406-2417, 2010.
- [8] Gang Chen, Lixu Gu, Lijun Qian, and Jianrong Xu, "An Improved Level Set for Liver Segmentation and Perfusion Analysis in MRIs," *IEEE Transactions on Information Technology in Biomedicine [J]*, vol. 13, no. 1, pp. 94-103, 2009

- [9] Xing Zhang, Jie Tian and etc, "Automatic Liver Segmentation Using a Statistical Shape Model With Optimal Surface Detection," *IEEE Transactions on Biomedical Engineering* [J], vol. 57, no. 10, pp. 2622-2626, 2010.
- [10] Sergio Casciaro, Roberto Franchini, Laurent Massotier and etc, "Fully Automatic Segmentations of Liver and Hepatic Tumors From 3-D Computed Tomography Abdominal Images: Comparative Evaluation of Two Automatic Methods," *IEEE Sensors Journal* [J], vol.12, no. 3, pp. 464-473, 2012.
- [11] D.Comaniciu and P.Meer, "Meanshift: a robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.* [J], vol. 24, no. 5, pp. 603-619, 2002.
- [12] Yonghui G., Jie Y., Xian X. and Feng S., "Efficient cellular automaton segmentation supervised by pyramid on medical volumetric data and real time implementation with graphics processing unit," *Expert Systems with Applications* [J], vol.38, no. 6, pp. 6866 - 6871, 2011.
- [13] Jifeng N., Lei Z., David Z. and Chengke W., "Interactive image segmentation by maximal similarity based region merging," *Pattern Recognition* [J], vol. 43, pp. 445-456, 2010.
- [14] Y. Zheng and C. Kambhamettu, "Learning based digital matting," In *ICCV* [C], pp. 889-896, 2009.
- [15] Vladimir Vezhnevets and Vadim Konouchine, "'GrowCut'-interactive multi-label N-D image segmentation by cellular automata," in *Proceedings of Graphicon* [C], pp. 150-156, 2005.
- [16] Claude K. and N. Piché, "Seeded ND medical image segmentation by cellular automaton on GPU," *International Journal of Computer Assisted Radiology and Surgery* [J], vol. 24, no. 5, pp. 1861-6410, 2009.
- [17] D. Comaniciu, V. Ramesh and P. Meer, "Kernel-based object tracking," *IEEE Trans. Pattern Anal. Mach. Intell.* [J], vol. 25, no. 5, pp. 564-577. 2003.
- [18] N. R. Draper and H. Smith. *Applied Regression Analysis* . JohnWiley, New York, 2nd edition, 1981.