Fast Graph-based Medical Image Segmentation with Expert Guided Statistical Information

Yu-Chi Hu, Michael D. Grossberg and Gig S. Mageras

Abstract— **In radiotherapy treatment planning, delineation of normal organs at risk in images is one of the most time-consuming tasks carried out routinely by human experts. Previously we proposed a speedy semi-automatic segmentation method based on a statistical graphical model, Conditional Random Field (CRF,) from which an energy function is defined to obtain Maximum-a-posteriori (MAP) estimation of the segmentation via a fast graph cut algorithm. The probabilistic regional and boundary terms in the energy function are estimated from the training samples collected locally from the the human expert via interactive tool or a training database. In this paper, we present a simple acceleration technique that dramatically improves the speed without sacrificing the accuracy of the segmentation. In the context of slice-by-slice medical image segmentation, we accelerate the process by partially reusing the graph constructed from a previous segmented slice based on the likeness of two consecutive images. Experiment results in 5 liver cases show differences between the manually segmented volumes and our estimated volumes were less than 5%. The differences are within the normal variation of manual segmentation from inter- and intra-observers. Accelerated segmentations show no degradation in terms of accuracy compared to full segmentations. The computation time per slice is within 300 millisecond CPU time for full segmentation and 110 millisecond for accelerated segmentation. The semi-automatic segmentation method proposed achieves similar segmentation done by human expert in significantly lesser time while preserving the human oversight required during the treatment planning process.**

I. INTRODUCTION

In radiation treatment planning and surgery planning, normal organ segmentation in medical images is routinely done by human experts to obtain quantified information of the organs, such as position, orientation and volume, to design a treatment plan. With the increasing amount of image data acquired from modern imaging devices, the task become a major bottleneck in the treatment planning process. While reliable automatic segmentation method is the ultimate goal, the fast and accurate semi-automatic segmentation is still very much desirable due to the human oversight required in health care practice. Semi-automatic segmentation techniques require a certain degree of user interaction such as clicking a starting point in the region glowing method[1], placing an initial contour in active contour/level set methods[2][3] or seed points in graph cut methods[4][5] and fitting a model in deformable model[6]. Examples in medical application can be found in Pan et al.[7] for liver segmentation using level set, Pasquier et al.[8] for bladder and rectum segmentation using seeded region growing and Pekar et al.[9] for segmentation in prostatic region using deformable model.

In our previous work[10], we presented a method that obtains segmentation by minimizing an energy function via graph cut. While similar to other graph cut based methods[11][4][5], it is also fundamentally different in how the energy function is derived. Graph cut based methods take advantage of a fast minimization algorithm: *min s-t cut* that separates a graph into two disconnected sets of nodes by cutting the edges connecting them, such that the cost (cost of edges being cut) is minimized. Graph cut based methods define an energy function in a way such that its minimization is equivalent to finding the min s-t cut. Unlike other graph cut based methods, our method derives the energy function from a statistical graphical model: *Conditional Random Fields* (CRFs.) The energy from a CRF is a summation of local potentials that we defined purposely for image segmentation as some probabilistic regional and boundary terms. The two probabilistic terms are estimated for each node from training sets of regional and boundary samples and are the cost of the corresponding edges in the graph for min s-t cut.

Here we introduce a simple but effective way to accelerate the graph based segmentation method proposed in our previous work while maintaining the same accuracy. While the graph cut minimization is computationally efficient, the probability estimation is not. For medical image segmentation, we take advantage of the finer image slice spacing provided by modern imaging devices. The change from two consecutive image slices is small for a normal organ, thus to accelerate the process, we reduce the number of nodes(voxels) that need estimation. For segmenting two consecutive image slices in a 3D medical image, we reuse the edge cost of nodes in the graph from the previous slice that has been segmented if the image features at a given node from the two images are similar enough. That is, if a node has similar image features to those on the previous slice, it should have same regional term that we estimated for the previous slice. Similarly, if a node and its neighbor have similar image features to those on the previous slice, it should have same boundary term. While the CRF framework and energy function minimization scheme in our previous work remains unaltered, this technique significantly reduces the number of nodes that need to be estimated for regional and boundary terms. Thus the segmentation speed is improved

The work is supported by the CCNY/MSKCC Partnership (U54CA137788/U54CA132378) from the National Cancer Institute.

Y.-C. Hu is with Medical Physics Department, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA. huj@mskcc.org

M. D. Grossberg, is with Department of Computer Science, City College of New York, New York, NY 10031, USA. michaeldg@gmail.com

G. S. Mageras is with Medical Physics Department, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA. magerasg@mskcc.org

978-1-4244-6561-3/10/$26.00 ©2010 IEEE

and better user experience is achieved. Further, comparing with full segmentation in our previous work, this technique offers comparable accuracy. For accuracy comparison of our method with other popular semi-automatic segmentation methods, readers are referred to [10].

II. METHODS

In our previous work[10], we purposed a semi-automatic method based on *Conditional Random Fields* and *Graph Cut*. For a better understanding our acceleration technique described later in this section, we will review these two graph related mathematical concepts in the first two sub-sections.

A. The energy function from Condition Random Fields

2D lattice and 3D grid graphs are a natural fit for representing digital discrete volumetric medical images. Each *voxel* in an image is represented by a *node* in the graph. Throughout the text, the terms voxel and node are exchangeable. There is an *edge* connecting a node and its neighboring node. Mathematically, let $G = (V, E)$ be the graph representing an image with N voxels indexed by i, where $V = \{i | i = 1, 2, 3, \ldots, N\}$, $E = \{\{i, j\} | i \in V, j \in N_i\}$ and N_i are neighbors of voxel i in a neighborhood system such as 4-connected system in 2D and 6-connected system in 3D. We treat the segmentation problem as a classification problem in a stochastic process, that is, we intend to assign a tissue class for each node(voxel) in the graph based on the observed image. In bi-class segmentation, for example, the tissue class could be *target* or *non-target*. Let $X = (X_i)_{i \in V}$ be the multivariate random variable of such assignments. x is an assignment instance and x_i is the class assignment for node i. Let Y be the multivariate random variable of images, y be the observed image instance (a gray-scale or color image) and y_i be the extracted image feature vector at node i. For segmentation of CT images, the 1-dimensional image feature we use is, but not limited to, image intensity. The segmentation problem can be simply described as finding an assignment x such that the conditional probability $P(X = x | Y = y)$ is maximum, that is, obtaining a *Maximum-A-Posteriori* (MAP) estimate of x.

Conditional Random Field (CRF)[12][13] is a type of undirected graphical model (Markov Random Fields, MRF) that models the conditional probability $P(X|Y)$. CRF has Markov property as does MRF, so the conditional probability can be factorized over the cliques in the graph. A clique a complete subgraph in which all nodes are connected to each other by edges. Using the graph G we defined above over the image voxels, a clique c is an actually an edge, that is, $c \in E$. The factorization of the conditional probability is given as product of clique potentials:

$$p(x|y) = \frac{1}{Z} \prod_c \phi_c(x_c, y) \tag{1}$$

Here ϕ_c is the clique potential for clique c and Z is the normalizing constant. ϕ_c is strictly positive and usually expressed as an exponential with Gibbs measure, so we have

$$p(x|y) = \frac{1}{Z} \exp\left(-\sum_c \mathcal{E}_c(x_c, y)\right) = \frac{1}{Z} \exp\left(-\mathcal{E}(x, y)\right) \tag{2}$$

$\mathcal{E}_c$ is interpreted as the energy at the clique c and the total energy $\mathcal{E}$ is obtained by adding the energy of each clique. In the context of medical image segmentation, since a clique is an edge $\{i, j\} \in E$ it is natural to define the clique energy function with both a regional term for associating the voxel i with a possible tissue class and a boundary term for the potential of assigning x_i and x_j to different classes. We intend to avoid any heuristic assumption about the segmentation. An example is that the boundary is more likely to have a stronger image edge feature (gradient), which is not necessarily true for medical images since the connected surrounding tissue near the target tissue class may have stronger contrast. To keep the CRF framework purely statistical, a log-likelihood function is a sensible choice for regional and boundary terms to estimate the assignment given the observed image y. Let r_i and u_{ij} be the regional term and boundary term respectively for a clique $\{i, j\}$. We define the energy as:

$$\mathcal{E}(x, y) = \sum_i [r_i(x_i, y) + \beta \sum_{j \in N_i} u_{ij}(x_i, x_j, y)] \tag{3}$$

, where

$$r_i(x_i, y) = -\ln p(y_i | x_i) \tag{4}$$

and

$$u_{ij}(x_i, x_j, y) = -\ln p(y_i, y_j | x_i \neq x_j) \tag{5}$$

Finally, with Eq 3, 4 and 5, the energy function is

$$\mathcal{E}(x, y) = \sum_i [-\ln p(y_i | x_i) - \beta \sum_{j \in N_i} \ln p(y_i, y_j | x_i \neq x_j)] \tag{6}$$

B. Minimization and Graph Cuts

With Eq 2, the Maximum-a-Posteriori (MAP) estimation of x is equivalent to minimization of the the energy function $\mathcal{E}(x, y)$. Let $x_i = 0$ if voxel i is the target tissue class and $x_i = 1$ if voxel i is the non-target tissue class, that is, X is a binary random vector for 2-class segmentation. As Kolmogorov et al [14] suggested, an energy minimization problem over binary variables can be solved by using a graph min s-t cut. The u_{ij} in Eq 5 is 0 when $x_i = x_j$ and non-negative when $x_i \neq x_j$. By Kolmogorov's $\mathcal{F}^2$ class theorem, since

$$u_{ij}(0, 0, y) + u_{ij}(1, 1, y) \leq u_{ij}(1, 0, y) + u_{ij}(0, 1, y) \tag{7}$$

our energy function $\mathcal{E}(x, y)$ is graph representable and can be minimized using a graph cut. We construct a graph $G' = (V', E')$, $V' = \{s, t, x_1, x_2, \ldots, x_n\}$. Here s represents the target tissue class, t represents the non-target tissue class and x_i corresponds to the voxel i. For each x_i, there is a directed edge e_{si} from s to x_i and edge e_{it} from x_i to t. Their edge costs are assigned to the regional term for non-target $-\ln p(y_i | x_i = 1)$ and regional term for target

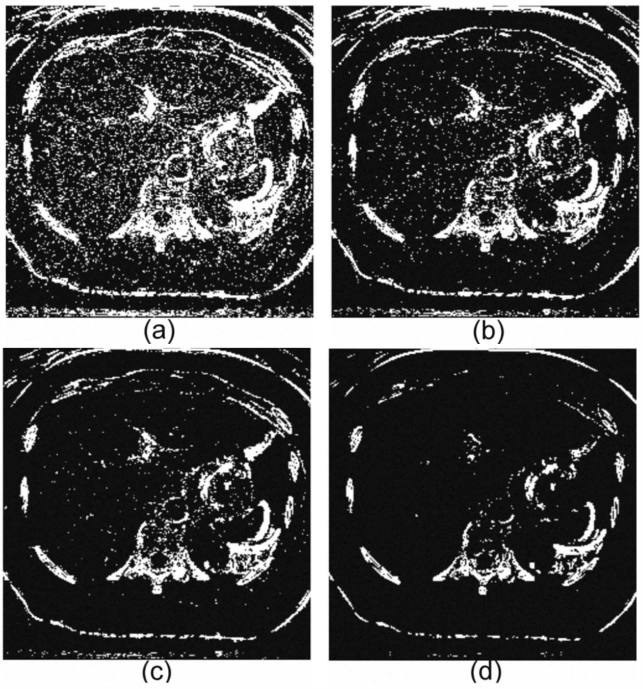

Fig. 1: The difference of two consecutive images. White pixels are where the difference is larger than the threshold and will be re-estimated for regional and boundary terms in the accelerated segmentation. The threshold is determined by regional statistics. In this case, standard deviation of the liver image intensity distribution (σ_{liver}) is used for thresholding. (a) $1\sigma_{liver}$, (b) $1.5\sigma_{liver}$, (c) $2\sigma_{liver}$ and (d) $3\sigma_{liver}$.

$-\ln p(y_i|x_i = 0)$ respectively. There are also bi-directional edges e_{ij} that connect the neighboring voxels i and j. For these edges, the costs are given by the boundary term $-\ln p(y_i, y_j|x_i \neq x_j)$. With the graph constructed, the cost of the cut (sum of the cost of the edges that are cut) is the minimum value of the energy function $\mathcal{E}(x, y)$.

C. Training

Training samples for regional and boundary terms in Eq 4 and 5 can be obtained from pre-segmented training dataset. For the experiments presented in the later section, we collect local samples of the 3D images through the interactive tools provided to expert users. For regional term in Eq 4, we use a similar approach as Boykov et al [4][5]. The expert uses paint brushes to identify locations of target and background tissues. But unlike Boykov et al. for using those voxels as seeds only, we also collect voxels under the brush strokes as samples of intensity distribution for target and non-target classes. The brush strokes are automatically carried over from a previous slice to the next one to save time for interaction and user can re-draw them if the brush strokes are not applicable on the current slice. For the boundary term, boundary samples are collected from a manual drawn contour of the target tissue on a single slice. Histograms are generated from these samples for non-parametric probability density estimation.

D. Acceleration by reusing graph

For 3D image stack, the segmentation process is done in a slice-by-slice fashion. Normally, regional and boundary terms are estimated for each voxel on a slice and we called this a full segmentation. To accelerate the segmentation process in the 3D image, given a image slice k to be segmented, if one of its adjacent slices, say slice $k-1$, has been fully segmented, then we do not need to estimate the regional and boundary terms for all voxels on slice k. Instead, we only estimate those voxels at which the difference between the two image slices is large. In other words, for a voxel indexed at i , if $Diff(y_{k-1,i}, y_{k,i}) < d$, where d is a threshold, the regional term and boundary terms are re-used from the adjacent slice $k-1$:

$$r_{k,i}(x_{k,i}, y_k) = r_{k-1,i}(x_{k-1,i}, y_{k-1}) \quad (8)$$

$$u_{k,ij}(x_{k,i}, x_{k,j}, y_k) = u_{k-1,ij}(x_{k-1,i}, x_{k-1,j}, y_{k-1}) \quad (9)$$

Since we use only the image intensity as the image feature, $Diff$ is simply the residual and we choose d to be σ_{target}, the standard deviation of the image intensity distribution of the target tissue class obtained from the samples. Figure 1 shows the some examples of the difference of two adjacent slices. Pixels in black are the pixels where the difference is smaller than the threshold.

E. Implementation

We implemented our method in an interactive semi-automatic tool in our in house 3D radiation treatment planning system. The code is written in Microsoft Visual Studio .NET C/C++. The pseudo-codes for constructing a graph to minimize the energy function in Eq 3 for full segmentation as well as accelerated segmentation is listed in Algorithm 1. We use the graph cut algorithm implemented by [15] to find the cut.

III. RESULTS

We have studied the segmentation of liver from 5 CT image sets. Experienced physicians manually drew the liver contours which is used as ground truth for comparison. Our method (CRF-GC) were performed twice for each case by an experienced user. One is done with normal full segmentation (CRF-GC) on all the slices, while the other is done with both full segmentation and accelerated segmentation on alternate slices (Fast CRF-GC.) Figure 2 shows some CT image slices with the liver contours delineated by physician, by our method with full segmentation and by our method with accelerated segmentation.

A. Accuracy

We use precision and recall rates with respect to the ground truth segmentation done by physician for accuracy measure and Figure 3 shows the results. The average precision rates for CRF-GC and Fast CRF-GC are 0.950 and 0.948 respectively and the average recall rates are 0.937 and 0.929 respectively. The CRF-GC and Fast CRF-GC are statistically no different in terms of accuracy. We also

Algorithm 1 ConstructCRFGraph(I_k,G_{k-1},C,O,B,σ_o)

I_k - current image slice
G_{k-1} - graph from the previous slice if any
C - boundary samples
O - target samples
B - non-target samples
σ_o - st. dev. of target image feature from samples

$V = \{s, t\}$
$E = \emptyset$
Create $G_k = (V, E)$
for each voxel $i \in I_k$ **do**
  $G_k.addNode(i)$
  if $G_{k-1} \neq null$ and $Diff(I_k(i), I_{k-1}(i)) \leq \sigma_o$ **then**
    $FastMode = true$
  end if
  if $FastMode$ **then**
    $G_k.addEdge(s, i, G_{k-1}.getEdgeCost(s, i))$
    $G_k.addEdge(i, t, G_{k-1}.getEdgeCost(i, t))$
  else {full segmatation}
    estimate regional term $r_i(0, I_k(i))$ from O
    estimate regional term $r_i(1, I_k(i))$ from B
    $G_k.addEdge(s, i, r_i(1, I_k(i)))$
    $G_k.addEdge(i, t, r_i(0, I_k(i)))$
  end if
  for each voxel $j \in N_i$ **do**
    if $FastMode$ **then**
      $G_k.addEdge(i, j, G_{k-1}.getEdgeCost(i, j))$
    else {full segmatation}
      estimate boundary term $u_{ij}(I_k(i), I_k(j))$ from C
      $G_k.addEdge(i, j, u_{ij}(I_k(i), I_k(j)))$
    end if
  end for
end for
return G_k

calculated the volumes of the segmented liver for comparison. The result is shown in Table I. The <5% difference between physician's segmentation and our methods is within the common variance of manual segmentation from intra- and inter-observers[16].

B. Performance

Using our GUI implementation, the user selects the box region of interest (ROI) on the image to be constructed for graph cut in our methods. The ROI size is usually 250×250 for liver segmentation. For the slices where Fast CRF-GC is performed, using the threshold selection method(Sec. II), the number of nodes that do need re-estimation of regional and boundary terms (edge costs in graph) in our energy function for graph cut minimization is reduced to 40% to 50% of original number of nodes in ROI. We measure the CPU time in one liver case on a dual-Xeon workstation. The average CPU time per slice for constructing the graph (estimation of edge costs) and graph cut are listed in Table II. Note that we implement our method in multi-threading.

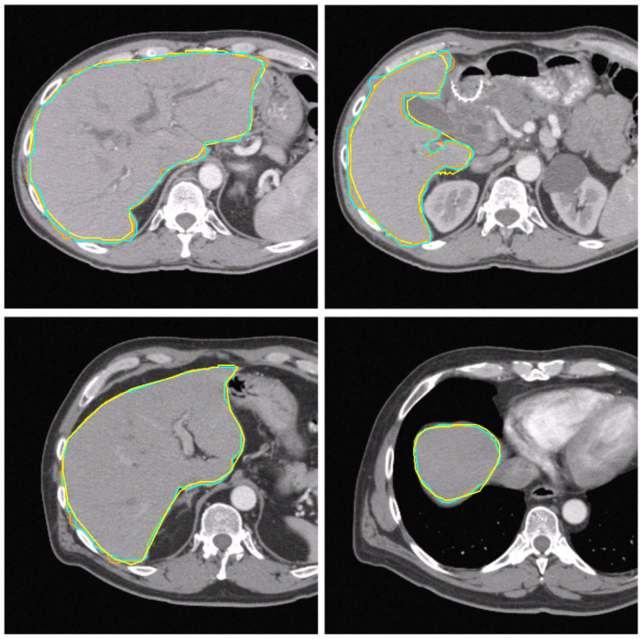

Fig. 2: Examples of liver segmentation. Yellow: physician drawn contour, Orange: CRF-GC, Cyan: CRF-GC Fast.

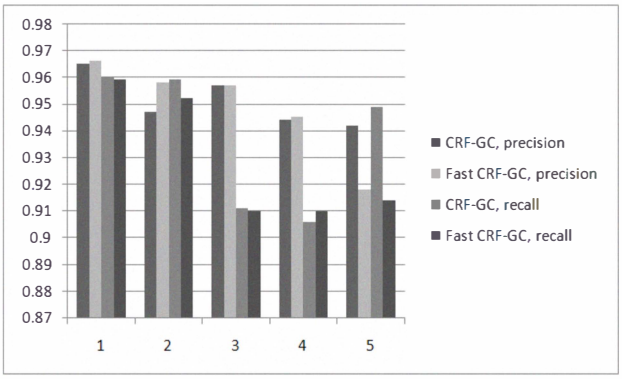

Fig. 3: Pecision/Recall rates of segmentations of 5 livers done by our methods, using physician manual segmentation as ground truth.

IV. CONCLUSIONS

Previously we have proposed a purely statistical framework for medical image segmentation. In this framework, two graph based components are used: the stochastic Conditional Random Field for defining an energy function with regional and boundary log-likelihoods and graph min s-t cut for minimizing the energy function that is equivalent to finding a Maximum-A-Posteriori estimate of the segmentation. In this paper, we present a simple but effective way to accelerated this segmentation process by reusing the graphs from pre-segmented slices in the same image volume. The results with the accelerated method show comparable accuracy and consistency with our previous method but the performance is improved by factor of two on slices where the accelerated method is used.

TABLE I: Volumes (in c.c.) of liver segmented by physician, our method and our accelerated method.

cases	MD	CRF-GC	Fast CRF-GC	Diff. MD & CRF-GC	Diff. CRF-GC & Fast CRF-GC
1	1270.1	1264.5	1264.9	0.44%	0.03%
2	1581.8	1611.4	1575.8	1.87%	2.21%
3	1755.4	1678.3	1671.7	4.39%	0.39%
4	1086.1	1049.8	1047.3	3.34%	0.24%
5	1967.7	1990.1	2024.1	1.14%	1.71%

TABLE II: Volumes (in c.c.) of liver segmented by physician, our method and our accelerated method.

	CPU time per slice	
	Graph construction	Graph cut
CRF-GC	211ms	15ms
CRF-GC Fast	97ms	15ms

REFERENCES

[1] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641–647, 1994.

[2] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, vol. 1, pp. 321–331, 1988.

[3] R. Malladi, J. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158–175, 1995.

[4] Y. Boykov and M. P. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images," in *Proc. International Conference of Computer Vision*, 2001, pp. 105–112.

[5] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient n-d image segmentation," *International Journal of Computer Vision*, vol. 70, pp. 109–131, 2006.

[6] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: A survey," *Medical Image Analysis*, vol. 1, pp. 91–108, 1996.

[7] S.Pan and B. Dawant, "Automatic 3d segmentation of the liver from abdominal ct images: a level set approach," *Proc SPIE Med Imaging*, vol. 4322, pp. 128–138, 2001.

[8] D. Pasquier, T. Lacornerie, M. Vermandel, J. Rousseau, E. Lartigau, and N. Betrouni, "Automatic segmentation of pelvic structures from magnetic resonance images for prostate cancer radiotherpy," *Int. J. Radiation Oncology Biol. Phys.*, vol. 68, pp. 592–600, 2007.

[9] V. Pekar, T. McNutt, and M. Kaus, "Automated model-based organ delineation for radiotherpy planning in prostatic region," *Int. J. Radiation Oncology Biol. Phys.*, vol. 60, pp. 973–980, 2004.

[10] Y.-C. Hu, M. Grossberg, and G. Mageras, "Semi-automatic medical image segmentation with adaptive local statistics in conditional random fields framework," in *EMBS 2008. 30th Annual International Conference of the IEEE*, 2008, pp. 3099–3102.

[11] Z. Wu and R. Leahy, "An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation," *IEEE Transactions on Pattern Analysis and Machince Intelligence*, vol. 15, no. 13, pp. 1101–1113, 1993.

[12] J. Lafferty, A. McCallum, and A. F. Pereira, "Conditional random fields:probabilistic models for segmenting and labeling sequence data," in *Proc. 18th International Conference on Machine Learning*, 2001, pp. 282–289.

[13] H. M. Wallach, "Conditional random fields: An introduction. technical report ms-cis-04-21." Department of Computer and Information Science, University of Pennsylvania, Tech. Rep., 2004.

[14] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 2, pp. 147–159, 2004.

[15] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machince Intellgence*, vol. 26, no. 9, pp. 1124–1137, 2004.

[16] L. Hermoye, I. L.-A. I, Z. Cao, L. Annet, J. Lerut, B. Dawant, and B. V. Beers, "Liver segmentation in living liver transplant donors: comparison of semiautomatic and manual methods," *Radiology*, vol. 234, pp. 171–8, 2005.