



Efficient multi-object segmentation of 3D medical images using clustering and graph cuts

Razmig Kéchichian, Sébastien Valette, Michel Desvignes, Rémy Prost

► To cite this version:

Razmig Kéchichian, Sébastien Valette, Michel Desvignes, Rémy Prost. Efficient multi-object segmentation of 3D medical images using clustering and graph cuts. ICIP 2011 - 18th IEEE International Conference on Image Processing, Sep 2011, Bruxelles, Belgium. pp.2149-2152, 10.1109/ICIP.2011.6116036 . hal-00658028

HAL Id: hal-00658028

<https://hal.science/hal-00658028>

Submitted on 9 Jan 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EFFICIENT MULTI-OBJECT SEGMENTATION OF 3D MEDICAL IMAGES USING CLUSTERING AND GRAPH CUTS

Razmig Kéchichian¹, Sébastien Valette¹, Michel Desvignes², Rémy Prost¹

¹: CREATIS; CNRS UMR5220; Inserm U630; INSA-Lyon; Université de Lyon, France

²: GIPSA-LAB/Grenoble-INPG, France

ABSTRACT

We propose an application of multi-label “Graph Cut” optimization algorithms to the simultaneous segmentation of multiple anatomical structures, initialized via an over-segmentation of the image computed by a fast centroidal Voronoi diagram (CVD) clustering algorithm. With respect to comparable segmentations computed directly on the voxels of image volumes, we demonstrate performance improvements on both execution speed and memory footprint by, at least, an order of magnitude, making it possible to process large volumes on commodity hardware which could not be processed pixel-wise.

Index Terms— Medical image segmentation, clustering, graph-cuts

1. INTRODUCTION

The analysis of CT medical images presents certain challenges. The presence of noise can create deceptive borders inside anatomical structures, which, in their turn, are often surrounded by other structures having similar appearances. This results in weak object edges which complicate organ extraction tasks. Moreover, the form of targeted structures often varies between individuals.

In recent years, semi-automatic organ extraction methods have become the alternative to fully-automatic methods. A simple paradigm of interactivity relying on reasonable amount of physician input, which can be refined as needed, can bring resolution to ambiguous decisions made by an algorithm, also making it possible to control the result. Such an approach represents an attractive replacement to manual organ segmentation methods which require significant investments of time and effort.

Many methods have been proposed in the application domain of semi-automatic segmentation [1]. Our preference goes to global methods of proven optimality and statistical robustness that describe a segmentation with a cost function, defined in terms of regional and boundary properties of segments. User interaction is carried out by the attribution of pixel labels (seeds) to targeted structures, one for each type of

structure. This provides clues on what the user intends to segment and therefore can be used to collect appearance statistics of targeted objects and to constrain the solution space of the algorithm.

“Graph Cut” approaches are a family of traditionally binary global optimization methods that have been widely applied to mono-object segmentation problems [2], where the segmentation cost function is represented by a weighted regular graph defined on image voxels. The minimum cut on this graph, equal to the maximum flow, is computed exactly in low-order polynomial time [3]. It optimizes the cost function globally hence finding the optimal segmentation. Multi-label generalizations of the approach [4] are not globally optimal, but are nevertheless capable of finding a local solution within a constant factor of the global optimum. These methods have been little exploited in multi-object segmentation.

In an attempt to improve the performance of “Graph Cut” algorithms, methods have been devised to reduce its data space yielding smaller graphs [5, 6]. Exploiting the fact that these algorithms lend themselves equally well to irregular non-grid graphs, we over-segment an image by a fast centroidal Voronoi diagram (CVD) clustering [7], which has ideal cluster size and color compactness properties, and use cluster barycenters and adjacency to define the graph representing our segmentation cost function to be optimized via multi-label “Graph Cut” algorithms. The efficiency of our approach is demonstrated in section 3 in two applications; simultaneous segmentation of principal thoracic cage structures and complete hip bone structure extraction from Visible Human CT datasets [8]. We are able to achieve performance improvements on both execution speed and memory footprint by, at least, an order of magnitude, while obtaining comparable results to those of segmentations computed directly on the voxels of the image.

2. PROPOSED APPROACH

We define an image I as a finite set of voxels, each voxel v is represented by a vector $v = (x y z I_v)$, where the first three components correspond to voxel coordinates in image space and the last component is the voxel’s intensity, i.e. its gray level.

2.1. Centroidal Voronoi diagrams

Given an image I , and n sites c_i where $0 \leq i \leq n-1$, the Voronoi diagram is defined on the image as n distinct regions C_i such that:

$$C_i = \{v \in I \mid d(v, c_i) < d(v, c_j); 0 \leq j \leq n-1, j \neq i\} \quad (1)$$

where $d(\cdot, \cdot)$ is a distance measure.

A centroidal Voronoi diagram (CVD) has the additional property that each site c_i is also the mass centroid (the barycenter) of the associated region:

$$c_i = \frac{\int_{C_i} v \cdot \rho(v) dv}{\int_{C_i} \rho(v) dv} \quad (2)$$

where $\rho(v)$ is a density function for C_i , which we define to obtain a gradient-adaptive CVD as $\rho(v) = a|\nabla I_v| + b$, where $|\nabla I_v|$ is the magnitude of the image gradient at voxel v .

A CVD can be constructed on an image I by minimizing the following energy function:

$$E_{CVD} = \sum_{i=1}^n \left(\sum_{v_j \in C_i} \int_{v_j} \rho(v) \|v - c_i\|^2 dv \right) \quad (3)$$

which intuitively corresponds to maximizing region compactness in terms of size and color. We use an iterative algorithm [7] that approximates a CVD and minimizes the energy function 3 globally in a computationally-efficient manner, involving only local queries.

2.2. Multi-label ‘‘Graph Cut’’ segmentation

Like many vision problems, segmentation can be formulated in terms of energy minimization. Given an image I and a set of labels L , one label for each targeted object, the task is to find a mapping $f : I \mapsto L; f(v) = l_v$ that minimizes some cost function. We use the following energy function which can be justified on Bayesian grounds in the context of maximum a posteriori estimation of Markov random fields (MAP-MRF) [9]:

$$E_l = \sum_{v \in I} D(l_v) + \sum_{\{u,v\} \in N} K_{\{u,v\}} \cdot T(l_u \neq l_v) \quad (4)$$

where N is a neighborhood system defined on voxels. The unary term $D(\cdot)$ is a function derived from observed data and measures the cost of assigning the label l_v to the voxel v , and the binary term $K_{\{u,v\}} \cdot T(l_u \neq l_v)$ imposes spatial coherence by penalizing the assignment of different labels to neighbor voxels, since $T(\cdot) = 1$ when the condition within the parentheses holds. Defining the energy function 4 on the clusters of a CVD-clustered image $C(I)$ gives the following energy function:

$$E_l = \sum_{C \in C(I)} D(l_C) + \sum_{\{C_1, C_2\} \in N} K_{\{C_1, C_2\}} \cdot T(l_{C_1} \neq l_{C_2}) \quad (5)$$

where the neighborhood system N is defined by cluster adjacency, which is no longer a regular grid graph.

For every identified object in the image, we estimate the probability distribution of object intensity $P_l(\cdot)$ from user-provided seeds, and define the energy function’s unary term $D(\cdot)$ according to the underlying MAP-MRF formulation:

$$D(l_C) = -\log(P_l(I_C)) \quad (6)$$

where I_C is the average intensity of cluster voxels for the cluster C . Note that when a seed is part of a cluster, the unary function for the corresponding label on the cluster is equal to zero, indicating no cost for the assignment of the label to the cluster. Conversely, the unary function for other labels on the cluster is prohibitively large, thus taking advantage of user input to constrain the solution space of the algorithm. We define the constant $K_{\{C_1, C_2\}}$ of the energy function’s binary term so that it preserves sharp discontinuities at object boundaries encouraging constant labeling elsewhere:

$$K_{\{C_1, C_2\}} = |\partial C_1 \cap \partial C_2| e^{-\frac{(I_{C_1} - I_{C_2})^2}{2\sigma^2}} \quad (7)$$

where $|\partial C_1 \cap \partial C_2|$ is the length (in 3D, the surface) of the common boundary of adjacent clusters C_1 and C_2 .

Finally we introduce the cluster size $|C|$ in the unary term and rewrite equation 5 so that its dimensions are coherent:

$$E_l = \sum_{C \in C(I)} \sqrt{-\log(P_l(I_C))} \cdot |C| \quad (8)$$

$$+ \sum_{\{C_1, C_2\} \in N} |\partial C_1 \cap \partial C_2| e^{-\frac{(I_{C_1} - I_{C_2})^2}{2\sigma^2}} \cdot T(l_{C_1} \neq l_{C_2})$$

This is the energy function we minimize to compute a segmentation. In mono-object segmentation, where the number of labels $|L| = 2$, the global minimum of the energy function 8 can be computed exactly in low-order polynomial time by finding the minimum cut on a weighted graph representing the energy function, which is equal to the maximum flow in the graph [3]. In multi-object segmentation tasks however, where $|L| > 2$, energy minimization is equivalent to the multi-way graph cut problem, which is NP-hard. Algorithms that compute a local minimum within a known factor of the global minimum have been proposed [4]. α -expansion is one such algorithm that finds a solution within a factor of 2 from the global minimum for our energy function. It can be applied when the binary energy term is sub-modular [10]:

$$\forall a, b, c \in L; E(a, a) + E(b, c) \leq E(a, c) + E(b, a) \quad (9)$$

which holds for our energy function.

3. EXPERIMENTAL RESULTS

We give the results and performance figures of several experimental mono and multi-object segmentations on CT images

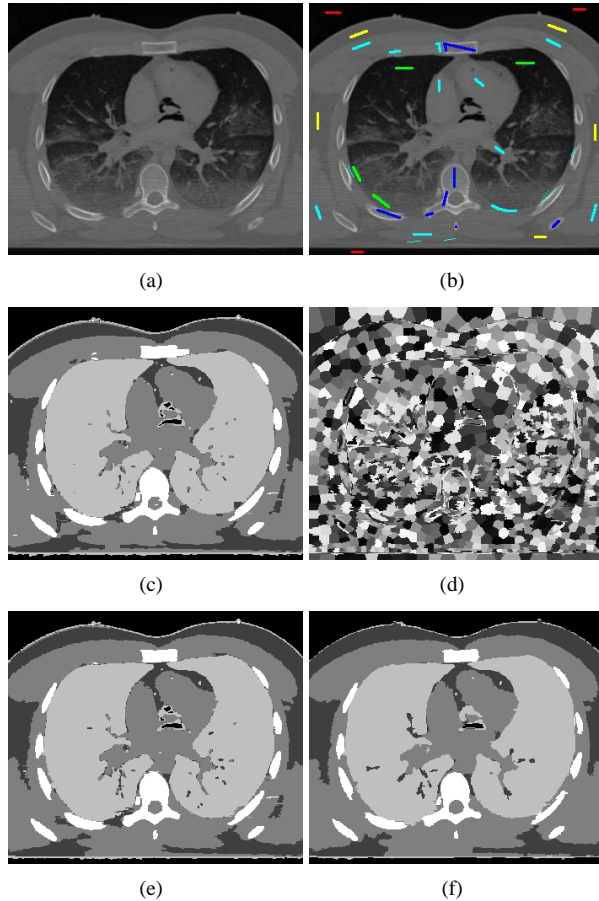


Fig. 1. Multi-object segmentation in 2D: a) CT cross-section at heart level b) seed-marked image, one color for each type of structure c) result of α -expansion multi-label segmentation applied to image pixels d) CVD clustering where number of clusters = 2% of number of pixels e) α -expansion multi-label segmentation applied to a CVD clustering where number of clusters = 10% of number of pixels and f) 2% of number of pixels. Percentage of pixels labelled differently with respect to c) 2.78% for e) and 5.3% for f)

from the Visible Human male and female datasets [8]. Our implementation uses as its back-end max-flow/min-cut and multi-label optimization libraries developed by the Computer Vision Research Group at the University of Western Ontario [3, 4, 10]. In all the experiments, a small number of seeds were placed inside targeted structures in a few selected slices of the image volume according to an anatomical atlas. Figure 1 compares three multi-object segmentation results in 2D, computed on image voxels and two CVD clusterings thereof. The aim was to extract bones, lungs, muscular and other tissues from the image.

It is interesting to point out that, contrarily to the widely-held opinion that “Graph Cut” segmentation methods are inherently incapable of extracting thin elongated objects [2], figure 1 shows a reasonable extraction of pulmonary veins by means of what is essentially an unrefined segmentation.

Tables 1 and 2 chart the performance of mono and multi-object segmentations performed on four 3D images, applied to image voxels and CVD-clustered images using a number of clusters equal to 1, 5 and 10 percent of the number of voxels. Results were obtained on a 4-core machine running at 2.84GHz with 4 GBs of main memory. Times are given in seconds and memory footprints in megabytes. “N/A” is indicated where a back-end was unable to allocate memory, making it impossible to proceed with segmentation. From the tables, it is obvious that execution time is linear in the number of graph nodes, and that the clustering allows us to segment large volumes which could not be processed pixel-wise. We believe that the relatively-long clustering time is insignificant since it can be performed off-line before the image is presented to a physician for interactive segmentation. Figure 2 shows segmentation results for the second image in each table. It illustrates segmentations of thoracic cage structures and complete hip and hand bones on CT images of Visible Human male and female datasets respectively.

4. CONCLUSIONS AND FUTURE WORK

We have presented an efficient coupling of centroidal Voronoi diagram (CVD) clustering and multi-label “Graph Cut” optimization algorithms applied to the simultaneous segmentation of multiple anatomical structures in medical images, improving significantly on execution speed and memory footprint of segmentations computed directly on the voxels of the image, without compromising segmentation quality.

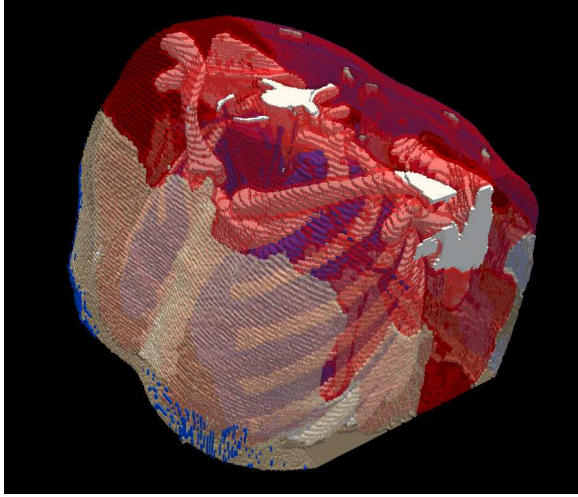
While the approach allows for demonstrably good segmentations as it is, there is much room for further improvement and experimentation. The approach can be readily applied to the analysis of medical images originating from different modalities, importantly MR imaging. Higher-order cluster statistics can be used to develop more expressive energy functions. Ultimately, statistical priors on organ appearance derived from the imaging modality and interorgan spatial relationships could be introduced in the framework of multi-label “Graph Cut” optimization, which would result in a principled multi-object segmentation approach based on medical imaging and anatomical concepts.

Acknowledgements

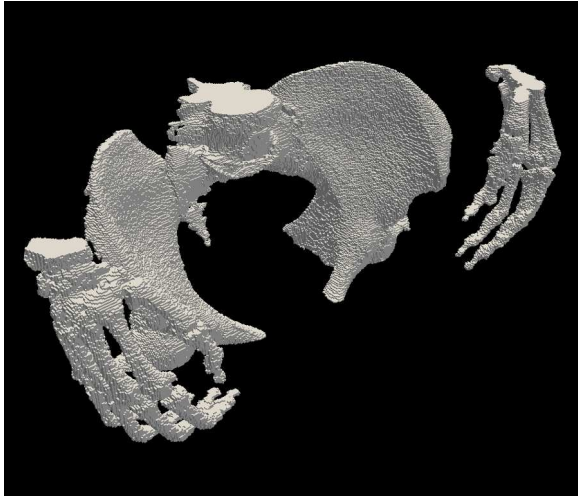
This work is supported by the Région Rhône-Alpes via the SIMED project of the research cluster ISLE.

5. REFERENCES

- [1] Vladimir Vezhnevets and Vadim Konouchine, ““Grow-Cut” - interactive Multi-Label N-D image segmentation,” in *Graphicon*, 2005, pp. 150–156.



(a)



(b)

Fig. 2. CVD-clustered 3D segmentation: a) bones, lungs, muscular and other tissues extracted from a 10.48×10^6 voxel volume by multi-object segmentation with number of clusters = 10% of number of voxels b) complete hip bone structure and hand bones extracted from a 20.6×10^6 voxel volume by mono-object segmentation with number of clusters = 10% of number of voxels

# Voxels	Approach	I (s)	II (s)	III (s)	Mem (Mb)
10.3×10^6	voxels		53.2	2.95	3994
	clusters 1%	319	3.3	0.1	433
	clusters 5%	176	8.8	0.9	902
	clusters 10%	147	14.1	2.0	1367
20.6×10^6	voxels		N/A	N/A	N/A
	clusters 1%	739	7.0	0.5	826
	clusters 5%	365	18.3	4.5	2000
	clusters 10%	327	29.5	8.2	2766

Table 1. Performance chart of 3D mono-object segmentation by max-flow. Columns I, II and III indicate the time spent for clustering, graph construction and optimization, respectively.

# Voxels	Approach	I (s)	II (s)	III (s)	Mem (Mb)
5.25×10^6	voxels		36.3	27.2	3027
	clusters 1%	151	1.7	1.1	238
	clusters 5%	90	5.0	9.2	615
	clusters 10%	69	8.3	19.2	1030
10.48×10^6	voxels		N/A	N/A	N/A
	clusters 1%	264	3.2	2.6	476
	clusters 5%	152	9.9	20.6	1141
	clusters 10%	125	16.6	43.8	1573

Table 2. Performance chart of 3D multi-object segmentation by α -expansion. Columns I, II and III give the time spent for clustering, graph construction and optimization, respectively.

- [2] Yuri Boykov and Gareth Funka-Lea, “Graph cuts and efficient N-D image segmentation,” *International Journal of Computer Vision*, vol. 70, no. 2, pp. 109–131, 2006.
- [3] Yuri Boykov and Vladimir Kolmogorov, “An experimental comparison of Min-Cut/Max-Flow algorithms for energy minimization in vision,” *IEEE PAMI*, vol. 26, no. 9, pp. 1124–1137, 2004.
- [4] Yuri Boykov, Olga Veksler, and Ramin Zabih, “Fast approximate energy minimization via graph cuts,” *IEEE PAMI*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [5] Alper Ayvaci and Daniel Freedman, “Joint segmentation-registration of organs using geometric models,” Lyon, France, 2007, Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings, pp. 5251–5254.
- [6] Chen Liu, Fengxia Li, Yan Zhang, and Haiyang Gu, “Interactive image segmentation based on hierarchical graph-cut optimization with generic shape prior,” in *ICIAR 2009*, Halifax, NS, Canada, 2009, vol. 5627 LNCS, pp. 201–210, Springer Verlag.
- [7] Sebastien Valette and Jean-Marc Chassery, “Approximated centroidal voronoi diagrams for uniform polygonal mesh coarsening,” *Computer Graphics Forum*, vol. 23, no. 3, pp. 381–389, 2004.
- [8] Victor Spitzer, Michael J Ackerman, Ann L Scherzinger, and David Whitlock, “The visible human male: a technical report,” *Journal of the American Medical Informatics Association*, vol. 3, no. 2, pp. 118–130, 1996.
- [9] D. M. Greig, B. T. Porteous, and A. H. Seheult, “Exact maximum a posteriori estimation for binary images,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 51, no. 2, pp. 271–279, 1989.
- [10] Vladimir Kolmogorov and Ramin 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.