

# Segmentation 3D multi-objet par “Graph Cut” et diagramme de Voronoï centroidal

Razmig KÉCHICHIAN<sup>1</sup>, Sébastien VALETTE<sup>1</sup>, Michel DESVIGNES<sup>2</sup>, Rémy PROST<sup>1</sup>

<sup>1</sup>Creatis, INSA de Lyon

7, avenue Jean Capelle, 69621 Villeurbanne Cedex, France

<sup>2</sup>Gipsa-lab, Grenoble INP

961, rue de la Houille Blanche, BP 46, 38402 Grenoble Cedex, France

<sup>1</sup>prénom.nom@creatis.insa-lyon.fr, <sup>2</sup>prénom.nom@gipsa-lab.grenoble-inp.fr

**Résumé** – Nous proposons une application des algorithmes d’optimisation multi-étiquette “Graph Cut” à la segmentation simultanée de plusieurs structures anatomiques présentes dans une image médicale 3D. Les graphes utilisés par les algorithmes “Graph Cut” dans cette application sont consommateurs de mémoire et dégradent les performances de la segmentation. Nous proposons de simplifier l’image de manière adaptative avec une partition calculée par un diagramme de Voronoï centroidal, et d’initialiser l’algorithme de segmentation avec celle-ci. Par comparaison à des segmentations calculées à partir des voxels d’une image, notre approche améliore la vitesse d’exécution et l’empreinte mémoire d’un ordre de grandeur, permettant ainsi le traitement de grands volumes impossibles à traiter en voxels.

**Abstract** – We propose an application of multi-label “Graph Cut” optimization algorithms to the simultaneous segmentation of multiple anatomical structures in 3D medical images, initialized via an over-segmentation of the image computed by a centroidal Voronoi diagram clustering algorithm. With respect to comparable segmentations computed directly on the voxels of an image, we demonstrate performance improvements on both execution speed and memory footprint by an order of magnitude, making it possible to process large volumes on commodity hardware which could not be processed pixel-wise.

## 1 Introduction

La segmentation des images médicales (IRM, CT, ultrasonores) est particulièrement difficile, à cause de la présence d’un niveau de bruit important, de niveaux d’intensité similaires des organes voisins et de l’effet de volume partiel dégradant la netteté de ses frontières.

Dans les dernières années, les approches de segmentation automatiques se sont trouvées en concurrence avec les approches semi-automatiques, où un paradigme simple d’interactivité, reposant sur un faible nombre d’interventions par un expert, permet de résoudre les décisions ambiguës prises par un algorithme de segmentation et de raffiner le résultat permettant ainsi à l’expert de maîtriser le traitement.

De nombreuses méthodes ont été proposées dans le domaine de la segmentation semi-automatique [1]. L’intervention de l’utilisateur se traduit souvent par l’attribution des étiquettes, “germes”, à certaines zones de l’image, chaque objet étant représenté par un type d’étiquette. Ceci permet à l’utilisateur de poser ses propres contraintes et fournit des indices à l’algorithme de segmentation qui peut recueillir des statistiques à propos des objets ciblés.

Les approches “Graph Cut” s’inscrivent dans le cadre des méthodes d’optimisation globale. Classiquement binaires, elles sont couramment appliquées aux problèmes de segmentation

mono-objet [2]. Ces approches s’appuient sur la minimisation exacte [3] d’une fonction de coût caractérisant la segmentation, représentée par un graphe pondéré défini sur les voxels de l’image, sur lequel la coupe minimale, égale au flot maximal, est calculée par un algorithme polynomial [4]. La coupe minimale minimise la fonction parmi toutes les coupes possibles produisant ainsi la segmentation optimale. Les généralisations multi-étiquettes de l’approche [5] ne sont pas optimales, le problème de base étant NP-difficile. Elles sont cependant capables de trouver une solution à distance connue de l’optimum global. Ces approches ont peu été exploitées dans la problématique de segmentation multi-objet.

Dans l’objectif d’améliorer les performances des algorithmes “Graph Cut”, des méthodes ont été conçues afin de réduire l’espace des données et de construire des graphes ayant des tailles inférieures [6, 7]. Les algorithmes “Graph Cut” étant parfaitement bien définis sur des graphes de topologie quelconque, en particulier non régulière, nous avons choisi de simplifier l’image de manière adaptative avec une partition calculée par un diagramme de Voronoï centroidal [8], créant ainsi des régions compactes en termes de taille et de couleur. Nous utilisons les barycentres et l’adjacence des régions issues de cette partition pour définir le graphe représentant la fonction de coût caractérisant une segmentation multi-objet, qui est ensuite minimisée par les algorithmes d’optimisation multi-étiquette

“Graph Cut”. Notre approche améliore la vitesse d’exécution et l’empreinte mémoire d’un ordre de grandeur tout en assurant une qualité de segmentation comparable à celle d’une segmentation calculée à partir des voxels de l’image.

## 2 Approche proposée

### 2.1 Diagrammes de Voronoï centroidaux

Une image en niveaux de gris  $\mathcal{I}$  consiste en un ensemble fini des voxels  $v = (x, y, z)$ . Soient  $n$  sites dans l’espace image  $c_i$ ;  $0 \leq i \leq n-1$ . Un diagramme de Voronoï est une partition de l’image à  $n$  régions disjointes  $C_i$  tel que :

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

où  $d(\cdot, \cdot)$  est une métrique de distance. Dans un diagramme de Voronoï centroidal (DVC), les sites  $c_i$  correspondent aux barycentres des régions  $C_i$  associées :

$$c_i = \frac{\sum_{v \in C_i} v \cdot \rho(v)}{\sum_{v \in C_i} \rho(v)} \quad (2)$$

$\rho(v)$  étant la fonction de densité associée à la région  $C_i$ . Afin d’obtenir une partition adaptative, nous la définissons à partir de la norme du gradient d’intensité  $\rho(v) = a|\nabla I_v| + b$ .

Soit  $v' = (v, \alpha I_v)$  le vecteur de dimension 4 qui attribue à chaque voxel  $v$  son niveau de gris  $I_v$ ,  $\alpha$  étant un scalaire, et soit  $c'_i = (c_i, \alpha \bar{I}_i)$  le vecteur attribuant au barycentre  $c_i$  l’intensité moyenne  $\bar{I}_i$  de sa région  $C_i$ . Un DVC peut être construit dans l’espace image augmenté en minimisant la fonctionnelle suivante :

$$E_{DVC} = \sum_{i=1}^n \left( \sum_{v \in C_i} \rho(v) \|v' - c'_i\|^2 \right) \quad (3)$$

ce qui équivaut à la maximisation de la compacité des régions en termes de taille et de couleur. Nous minimisons la fonctionnelle (3) avec un algorithme itératif [8] qui construit un DVC approximatif par une optimisation globale efficace à partir des modifications locales aux frontières des régions.

### 2.2 Segmentation multi-objet par “Graph Cut”

Soit  $\mathcal{L}$  un ensemble d’étiquettes représentant les objets ciblés dans l’image  $\mathcal{I}$ . La segmentation équivaut à un problème d’étiquetage consistant à trouver l’application  $f: \mathcal{I} \rightarrow \mathcal{L}$ ;  $f(v) = \ell_v$  minimisant une fonction de coût définie à partir des propriétés des objets à segmenter et des frontières entre ceux-ci. Nous avons choisi une fonctionnelle d’énergie justifiée dans le contexte de l’estimation maximum a posteriori des champs Markoviens (MAP-MRF) [9] :

$$E(\ell) = \sum_{v \in \mathcal{I}} D(\ell_v) + \sum_{\{u, v\} \in N} K_{u, v} \cdot T(\ell_u \neq \ell_v) \quad (4)$$

Le terme unaire  $D(\cdot)$  est dérivé de l’observation et représente le coût d’attribution de l’étiquette  $\ell$  au voxel  $v$ . Le terme binaire  $K_{u, v} \cdot T(\ell_u \neq \ell_v)$ , défini sur un voisinage régulier  $N$ ,

provient de l’a priori Markovien. Il assure une cohérence spatiale en pénalisant l’attribution des étiquettes différentes aux voxels voisins, car  $T(\cdot) = 1$  quand la condition entre les parenthèses est vérifiée.

En définissant la fonctionnelle (4) sur une partition d’image par DVC,  $C(\mathcal{I})$ , où le voisinage irrégulier  $Z$  est défini à partir de l’adjacence des régions, nous avons :

$$E(\ell) = \sum_{C \in C(\mathcal{I})} D(\ell_C) + \sum_{\{C_1, C_2\} \in Z} K_{C_1, C_2} \cdot T(\ell_{C_1} \neq \ell_{C_2}) \quad (5)$$

Pour tout objet identifié dans l’image et représenté par une étiquette  $\ell$ , nous estimons la distribution de probabilité conditionnelle pour les niveaux de gris  $Pr(I|L)$  à partir des germes placés par l’utilisateur à l’intérieur des objets ciblés. Nous définissons le terme unaire  $D(\cdot)$  de la fonctionnelle (5) d’après la formulation MAP-MRF :

$$D(\ell_C) = \begin{cases} 0 & \exists s_k \in C, k = \ell \\ \infty & \exists s_k \in C, k \neq \ell \\ -\log(Pr(\bar{I}_C | \ell_C)) & \forall s_k \in C \end{cases} \quad (6)$$

Dans l’équation (6),  $s_k$  représente un germe correspondant à l’objet identifié par l’étiquette  $k$ , et  $\bar{I}_C$  représente l’intensité moyenne de la région  $C$ . Nous définissons la constante du terme binaire de la fonctionnelle (5) afin qu’elle préserve les ruptures aux frontières des objets favorisant un étiquetage constant ailleurs :

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

$|\partial C_1 \cap \partial C_2|$  est l’aire de la frontière commune entre les régions  $C_1$  et  $C_2$ . Enfin nous introduisons la notion de taille de région (le nombre de voxels)  $|C|$  et écrivons la fonctionnelle (5) de la manière suivante afin d’obtenir une équation aux dimensions cohérentes :

$$E(\ell) = \sum_{C \in C(\mathcal{I})} -\log(Pr(\bar{I}_C | \ell_C)) \sqrt{|C|} + \sum_{\{C_1, C_2\} \in Z} |\partial C_1 \cap \partial C_2| e^{-\frac{(\bar{I}_{C_1} - \bar{I}_{C_2})^2}{2\sigma^2}} \cdot T(\ell_{C_1} \neq \ell_{C_2}) \quad (8)$$

En segmentation mono-objet où le nombre d’étiquettes  $|\mathcal{L}| = 2$ , la minimisation exacte de la fonctionnelle (8) est possible. Elle est représentée par un graphe pondéré défini à partir des régions de la partition sur lequel la coupe minimale, égale au flot maximal, est calculée de manière exacte par un algorithme polynomial [4]. En segmentation multi-objet où  $|\mathcal{L}| > 2$ , le calcul de coupe minimale sur le graphe est un problème NP-difficile. Nous utilisons un algorithme d’approximation, appelé  $\alpha$ -expansion [5], capable de trouver un minimum local dans un rapport 2 avec le minimum global, qui peut être appliqué à condition que le terme d’énergie binaire soit sous-modulaire [10] :  $\forall a, b, c \in \mathcal{L}; E(a, a) + E(b, c) \leq E(a, c) + E(b, a)$ . Cette condition est vérifiée par la fonctionnelle d’énergie (8).

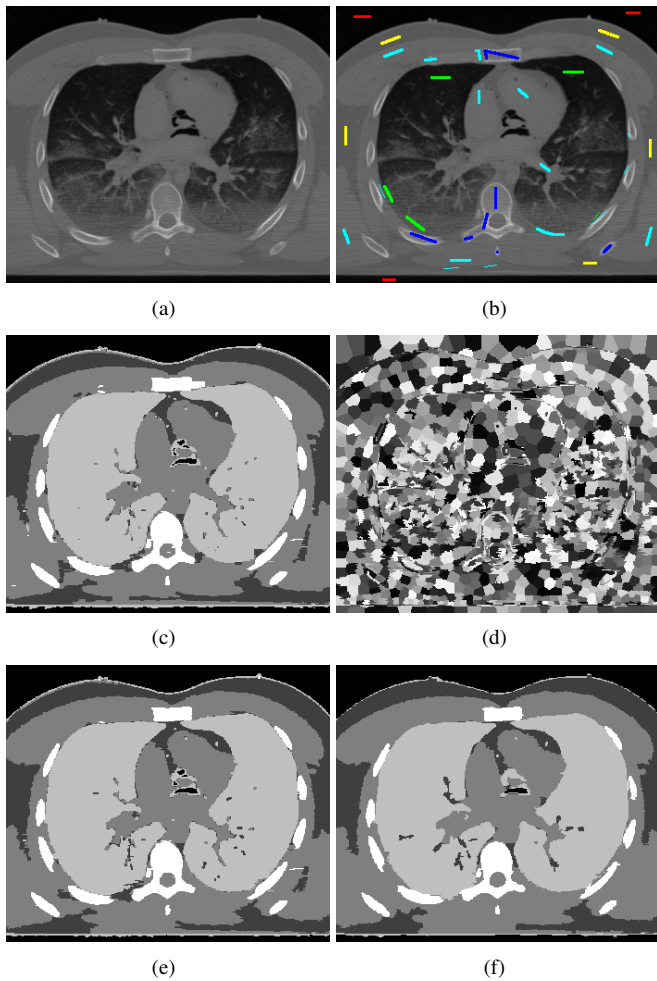


FIG. 1 – Segmentation multi-objet 2D par  $\alpha$ -expansion a) Coupe CT transverse au niveau du cœur b) Image avec germes saisis par l'utilisateur c) Segmentation à partir des pixels d) Partition d'image par DVC où le nombre de régions = 2% du nombre de pixels e) et f) 2 segmentations à partir de 2 partitions d'image où le nombre de régions = 10% et 2% du nombre de pixels respectivement. Pourcentage de pixels étiquetés différemment par rapport à c) 2,78% pour e) et 5,3% pour f)

### 3 Résultats expérimentaux

Nous résumons les résultats et les indices de performances de notre approche appliquée à la segmentation des images CT issues des jeux de données “Visible Human” homme et femme [11]. Notre implémentation s'appuie sur les bibliothèques d'optimisation binaire et multi-étiquette développées par le Computer Vision Research Group à l'Université Western Ontario [4, 5, 10]. Dans toutes les expériences, un faible nombre des germes ont été placés manuellement à l'intérieur des structures anatomiques ciblées dans quelques coupes transversales de l'image 3D selon un atlas anatomique.

La figure 1 compare 3 segmentations multi-objets d'une image 2D, calculées à partir des voxels de l'image ou des régions de 2 partitions différentes de l'image, obtenues par un diagramme de Voronoï centroidal (DVC). L'application cherche

à segmenter les os, les poumons, le tissu musculaire et l'enveloppe corporelle à partir de la cage thoracique.

Les tableaux 1 et 2 résument les indices de performances des segmentations mono et multi-objets de 4 images 3D, calculées à partir des voxels de l'image ou des régions de 3 partitions différentes de chaque image, obtenues par un DVC. En segmentation mono-objet, l'application cherche à segmenter les os de la hanche et des mains, la segmentation multi-objet étant une application 3D de celle démontrée dans la figure 1. La figure 2 illustre les résultats de segmentation pour l'image 2 et l'image 4 dans les tableaux 1 et 2 respectivement.

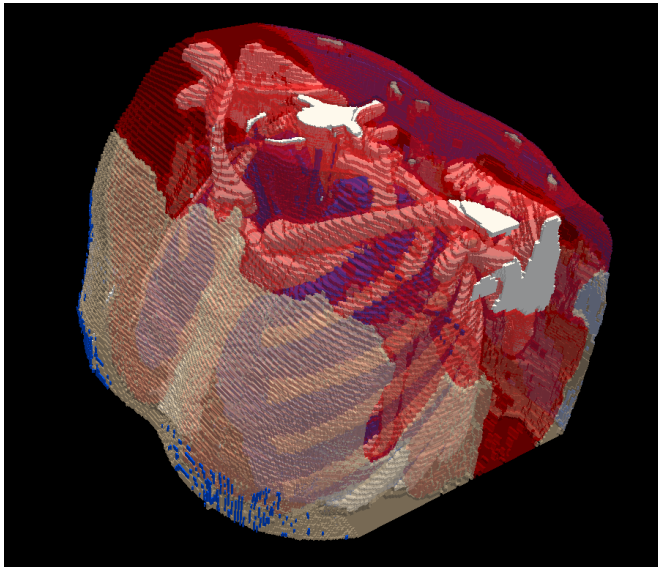
Toutes les expériences ont été effectuées sur une machine à 4 processeurs cadencés de 2,84GHz avec 4Go de mémoire vive. Dans les tableaux, “N/A” correspond aux cas où la librairie d'optimisation était incapable d'allouer l'espace mémoire nécessaire rendant le calcul impossible. D'après les tableaux, on observe que le temps d'exécution est linéaire en fonction du nombre de nœuds de graphe et que la partition par DVC permet de segmenter de grands volumes qui ne peuvent pas être traités en voxels. Bien que le temps de partition soit relativement long, nous pensons que cela n'est pas gênant, car la partition peut s'effectuer “hors-ligne” avant que l'image soit présentée à l'utilisateur pour une segmentation interactive.

TAB. 1 – Tableau de performances de segmentations multi-objets 3D par  $\alpha$ -expansion. Colonnes I, II et III indiquent le temps d'exécution en secondes de la partition par DVC, la construction de graphe et l'optimisation respectivement. “N/A” indique un échec d'allocation de mémoire.

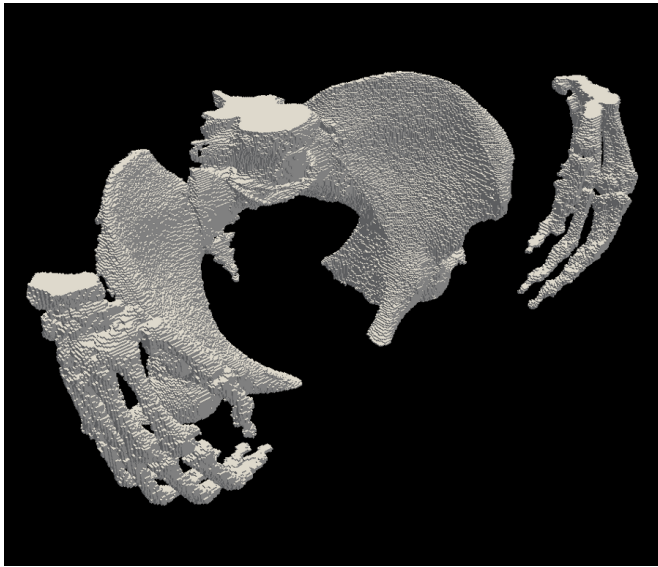
Image / # voxels	Nœuds	I (s)	II (s)	III (s)	Mém (Mo)
Image 1 $5,25 \times 10^6$	voxels		36,3	27,2	3027
	DVC 1%	151	1,7	1,1	238
	DVC 5%	90	5,0	9,2	615
	DVC 10%	69	8,3	19,2	1030
Image 2 $10,48 \times 10^6$	voxels		N/A	N/A	N/A
	DVC 1%	264	3,2	2,6	476
	DVC 5%	152	9,9	20,6	1141
	DVC 10%	125	16,6	43,8	1573

TAB. 2 – Tableau de performances de segmentations mono-objets 3D par flot-max. Colonnes I, II et III indiquent le temps d'exécution en secondes de la partition par DVC, la construction de graphe et l'optimisation respectivement. “N/A” indique un échec d'allocation de mémoire.

Image / # voxels	Nœuds	I (s)	II (s)	III (s)	Mém (Mo)
Image 3 $10,3 \times 10^6$	voxels		53,2	2,95	3994
	DVC 1%	319	3,3	0,1	433
	DVC 5%	176	8,8	0,9	902
	DVC 10%	147	14,1	2,0	1367
Image 4 $20,6 \times 10^6$	voxels		N/A	N/A	N/A
	DVC 1%	739	7,0	0,5	826
	DVC 5%	365	18,3	4,5	2000
	DVC 10%	327	29,5	8,2	2766



(a)



(b)

FIG. 2 – Segmentations sur une partition d’image 3D par DVC où le nombre de régions = 10% du nombre de voxels a) Segmentation multi-objet des os, des poumons, du tissu musculaire et de l’enveloppe corporelle à partir d’un volume à  $10,48 \times 10^6$  voxels par  $\alpha$ -expansion b) Segmentation mono-objet des os de la hanche et des mains à partir d’un volume à  $20,6 \times 10^6$  voxels par flot-max

## 4 Conclusions et perspectives

Nous avons présenté un couplage efficace de la partition d’image par un diagramme de Voronoï centroidal (DVC) et des algorithmes d’optimisation multi-étiquette “Graph Cut”. Son application à la segmentation simultanée de plusieurs structures anatomiques présentes dans une image médicale 3D améliore la vitesse d’exécution et l’empreinte mémoire, par comparaison à des segmentations calculées à partir des voxels de l’image, sans compromettre la qualité de segmentation.

La suite de ce travail portera sur l’amélioration des performances de l’algorithme de partition par DVC et l’exploitation des statistiques d’ordre supérieur afin de définir des fonctionnelles d’énergie plus expressives. Nous envisageons l’introduction d’a priori statistique sur l’apparence des organes provenant de la modalité d’imagerie visée ainsi que sur les relations spatiales inter-organes dans le formalisme de segmentation multi-objet par “Graph Cut”. Ces relations permettent de lever facilement certaines ambiguïtés, mais il faudrait pouvoir les exprimer dans un référentiel image exploitable par les fonctionnelles d’énergie minimisables par les algorithmes “Graph Cut”.

## Remerciements

Nous remercions la Région Rhône-Alpes du soutien apporté à ce travail par une allocation doctorale de recherche dans le cadre du projet SIMED du cluster de recherche ISLE.

## Références

- [1] V. Vezhnevets et V. Konouchine. “Grow-Cut” - *Interactive Multi-Label N-D Image Segmentation*. Graphicon, 2005.
- [2] Y. Boykov et G. Funka-Lea. *Graph cuts and efficient N-D image segmentation*. IJCV, vol. 70, no. 2, 2006.
- [3] D. M. Greig, B. T. Porteous et A. H. Seheult. *Exact maximum a posteriori estimation for binary images*. Journal de RSS, série B, vol. 51, no. 2, 1989.
- [4] Y. Boykov et V. Kolmogorov. *An experimental comparison of Min-Cut/Max-Flow algorithms for energy minimization in vision*. IEEE PAMI, vol. 26, no. 9, 2004.
- [5] Y. Boykov, O. Veksler et R. Zabih. *Fast approximate energy minimization via graph cuts*. IEEE PAMI, vol. 23, no. 11, 2001.
- [6] A. Ayvaci et D. Freedman. *Joint segmentation-registration of organs using geometric models*. IEEE EMB, 2007.
- [7] C. Liu et al. *Interactive image segmentation based on hierarchical graph-cut optimization with generic shape prior*. ICIAR, 2009.
- [8] S. Valette et J-M. Chassery. *Approximated centroidal voronoi diagrams for uniform polygonal meshing*. Computer Graphics Forum, vol. 23, no. 3, 2004.
- [9] S. Z. Li. *Markov random field models in computer vision*. ECCV’94, vol. 801, 1994.
- [10] V. Kolmogorov et R. Zabih. *What energy functions can be minimized via graph cuts ?*. IEEE PAMI, vol. 26, no. 2, 2004.
- [11] V. Spitzer, M. J. Ackerman, A. Scherzinger et D. Whitlock. *The visible human male*. Journal of the American Medical Informatics Association, vol. 3, no. 2, 1996.