MULTIREGION GRAPH CUT IMAGE SEGMENTATION

Mohamed Ben Salah[1], Ismail Ben Ayed[2] and Amar Mitiche[3]
[1,3] *Institut National de la Recherche Scientifique, INRS-EMT, 800 De La Gauchetire Ouest , Montral, QC, H5A 1K6, Canada*
[2] *GE healthcare, 268 Grosvenor, E5-137, London, ON, N6A 4V2, Canada*
[1]*bensalah@emt.inrs.ca,* [2]*ismail.benayed@ge.com,* [3]*mitiche@emt.inrs.ca*

Keywords: Combinatorial optimization, graph cuts, parametric modeling, image segmentation.

Abstract: The purpose of our study is two-fold: (1) investigate an image segmentation method which combines parametric modeling of the image data and graph cut combinatorial optimization and, (2) use a prior which allows the number of labels/regions to decrease when the number of regions is not known and the algorithm initialized with a larger number. Experimental verification shows that the method results in good segmentations and runs faster than conventional graph cut methods.

1 INTRODUCTION

Image segmentation occurs in many important applications. It consists of partitioning an image into regions homogeneous with respect to a given description. Energy minimization formulations have led to flexible, transparent, and effective algorithms. These formulations can be divided into two broad categories: continuous and discrete (Zhu and Yuille, 1996) (Ayed et al., 2006) (Chan and Vese, 2001) (Boykov et al., 2006). Continuous formulations, which seek a partition of the image domain by active curves via a level set representation, have segmented accurately a variety of difficult images.

Discrete formulations use objective functions which contain terms similar to those in continuous formulations, generally a data term which measures the conformity of a segmentation to the image data, and a regularization term. A graph cut formulation states segmentation as a labeling problem. The graph cut algorithm assigns each pixel a grey level label in the set of all possible labels. Thus, the optimization of the objective function results in a segmentation described by a subset of the set of possible labels. Since no explicit number of regions guides the algorithm, the solution is generally an implicit oversegmentation. Also, the use of a set of all possible labels lengthens the execution time unnecessarily. When the number of regions is known, one can reduce the number of labels to equal the number of regions. Regions can then be described by positive parametric functions, in which case, the data term of the objective function measures the conformity of the image data to a parametric model in each region. This formulation was used to state the problem of concurrent segmentation and estimation of image motion (Schoenemann and Cremers, 2006). This formulation leads to an optimization which iterates two steps: continuous optimization with respect to the region parameters and segmentation by discrete graph cut optimization. However, the regularization term in (Schoenemann and Cremers, 2006), which estimates the length of the segmentation boundaries assumes the number of regions known. When the number of regions is not known beforehand, one sensible formulation would be to use a larger number initially, to be decreased by the graph cut optimization algorithm (Boykov et al., 2001). However, and contrary to the priors used in (Boykov et al., 2001), the length prior does not allow this, in general, because merging two disjoint regions does not change the value of the prior.

The purpose of our study is two-fold: (1) investigate image segmentation which combines parametric modeling of the image data and graph cut combinatorial optimization and (2) use a prior which allows the number of labels/regions to decrease when the number of regions is not known and the algorithm initialized with a larger number. As in (Schoenemann and

Cremers, 2006), the method iterates two steps: closed form region parameters update and segmentation by graph cut combinatorial optimization but our regularization term allows region merging.

2 OBJECTIVE FUNCTION

As in (Boykov et al., 2001), and other studies as well (Chan and Vese, 2001), the objective function contains two terms: a data term to measure the conformity of the segmentation to a parametric model, a piecewise constant model in this case, and a regularization term to bias the segmentation toward partitions of the image domain with fewer boundary edges. The goal is to segment the image into a fixed but arbitrary number of regions using graph cut optimization of the objective function. To do so, we use a labeling function $\lambda : \Omega \longrightarrow \mathcal{L}$ so that each pixel $p \in \Omega$ is assigned the label $\lambda(p)$ in a finite set of labels $\mathcal{L}$ of cardinality N_{reg} equal to the number of regions. Let R_l be the region of pixels with label l. The solution of the minimization, or the final labeling, can be represented either by the function λ or the partition $\mathbf{P} = \{R_l \mid l \in \mathcal{L}\}$, where $R_l = \{p \in \Omega \mid \lambda(p) = l\}$. The energy functional to minimize, is formulated using the labeling function λ, a data term and the regularization term, $\mathcal{R}$:

$$\mathcal{F}(\lambda) = \sum_{p \in \Omega} D_p(\lambda(p)) \; + \; \alpha \mathcal{R}(\lambda) \tag{1}$$

where α is a positive weight. Typically, $D_p(\lambda(p))$ is $(\mu_l - I_p)^2$, where I_p is the intensity at pixel p and μ_l is the value of the label assigned to region R_l for $1 \leq l \leq N_{reg}$. We use the following regularization term:

$$\mathcal{R}(\lambda) = \sum_{\{p,q\} \in \mathcal{N}} r_{\{p,q\}}(\lambda(p), \lambda(q)) \tag{2}$$

where $\mathcal{N}$ is the neighborhood set. Function $r_{\{p,q\}}(\lambda(p), \lambda(q))$ is a smoothness regularization function. We consider a regularization function which is piecewise smooth. More precisely, we consider the truncated squared absolute difference model $r(\lambda(p), \lambda(q)) = \min(const, |\mu_{\lambda(p)} - \mu_{\lambda(q)}|^2)$ where *const* is, in all experimentations, the same used in (Boykov et al., 2001). The objective function can be written, for $1 \leq l \leq N_{reg}$, as follows:

$$\begin{aligned} \mathcal{F}(\{\mu_l\}, \lambda) &= \sum_{l \in \mathcal{L}} \sum_{p \in R_l} (\mu_l - I_p)^2 \\ &+ \alpha \sum_{\{p,q\} \in \mathcal{N}} r_{\{p,q\}}(\lambda(p), \lambda(q)) \end{aligned} \tag{3}$$

3 MINIMIZATION

Minimization of functional (1) is done by alternating graph cut optimization for segmentation and updating of labels via continuous optimisation. Given the segmentation, regions parameters are estimated according to a closed-form solution. Given the parameters, the segmentation is updated by graph cut optimization.

3.1 Labels Updating

Given a segmentation, a labeling λ is an assignment of a label $\lambda(p)$ to each pixel p of the image domain Ω. Equivalently, it is a partition of Ω into disjoint regions R_l where l is the label assigned to pixels of the region R_l. The process of updating labels consists in estimating these parameters in a way to optimize the functional when the segmentation is given.
To optimize the functional (3) for a given segmentation, we proceed by derivation towards the value of a label μ_l:

$$\begin{aligned} \frac{\partial \mathcal{F}(\{\mu_l\}, \lambda)}{\partial \mu_l} &= 2 \sum_{p \in R_l} (\mu_l - I_p) + \\ 2\alpha & \sum_{\substack{\{p,q\} \in \mathcal{N}, \lambda(p) = l \\ |\mu_{\lambda(p)} - \mu_{\lambda(q)}| < const}} (\mu_{\lambda(p)} - \mu_{\lambda(q)}) \end{aligned} \tag{4}$$

We choose a neighborhood system, $\mathcal{N}$, of size 4 and let $\overline{\mathcal{N}}_p$ be the set of pixels q neighbors of p which are assigned a label $\lambda(q)$ different from, $\lambda(p)$, the label assigned to p and verifying $|\mu_{\lambda(p)} - \mu_{\lambda(q)}| < const$. Consequently, equation (4) can be written as

$$\frac{\partial \mathcal{F}}{\partial \mu_l} = 2 \sum_{p \in R_l} (\mu_l - I_p) \; + \; 2\alpha \sum_{p \in \mathcal{C}_l} \sum_{q \in \overline{\mathcal{N}}_p} (\mu_l - \mu_{\lambda(q)}) \tag{5}$$

where $\mathcal{C}_l$ is the boundary of the region R_l. The label value optimizing the functional, given a segmentation, is deduced from the last equation and expressed in a closed-form as follows

$$\mu_l = \frac{\sum_{p \in R_l} I_p \; + \; \alpha \sum_{p \in \mathcal{C}_l} \sum_{q \in \overline{\mathcal{N}}_p} \mu_{\lambda(q)}}{\sharp R_l \; + \; \alpha \sum_{p \in \mathcal{C}_l} \sharp \overline{\mathcal{N}}_p} \tag{6}$$

where $\sharp$ designates set cardinality.

3.2 Graph Cuts Segmentation Updating

Graph cut optimization was introduced by Greig et al. (Greig et al., 1989) for image restoration. We use the algorithm of (Boykov et al., 2001) which has been used successfully in several vision applications. The algorithm functions as follows. Let $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$ be a weighted graph where $\mathcal{V}$ is the set of vertices (nodes) and $\mathcal{E}$ the set of edges. $\mathcal{V}$ contains a node for each pixel and two additional nodes called terminals. There is an edge $\{p,q\}$ between any two distinct nodes p and q. A *cut* $\mathcal{C} \subset \mathcal{E}$ is a set of edges verifying:

- Terminals are separated in the graph $\mathcal{G}(\mathcal{C}) = \langle \mathcal{V}, \mathcal{E} \setminus \mathcal{C} \rangle$
- No subset of $\mathcal{C}$ separates terminals in $\mathcal{G}(\mathcal{C})$

The minimum cut problem consists in finding the cut $\mathcal{C}$ in a given graph with the lowest cost. The cost of a cut, $|\mathcal{C}|$, is the sum of its edges weights.
In (Boykov et al., 2001), Boykov et al. present two algorithms based on graph cuts which find a local minimum of the objective function (3) with respect to two kinds of large moves: expansion moves and swap moves. Large moves proceed by changing labels of a large set of pixels simultaneously to decrease the objective function. In this work, we use the algorithm based on swap moves because it handles more general energy functions and deals with the regularization function that we have chosen.

4 EXPERIMENTATION

To validate the advantages of the method, we present first experiments with synthetic data for noisy piecewise constant images. Then we experiment with real images. The initialization starts with an arbitrary partition of the desired number of regions to which arbitrary labels are assigned. the unique free parameter is the regularization term penalty α.

4.1 Synthetic Data

Figure 1 shows synthetic images representing geometric shapes in the first row, the segmentations obtained using the method of (Boykov et al., 2001) in the second row and the results of the segmentations with our method in the third row. Contrary to the method of (Boykov et al., 2001) which uses a set of labels corresponding to all the possible grey levels, we have the possibility to initialize the process of segmentation with the right number of regions, which is known here, and with an arbitrary set of regions parameters.
The method of (Boykov et al., 2001) allows the number of labels to decrease due to the regularization term but it initializes with the set of labels corresponding to all possible grey levels, i.e., 256 grey levels. The final segmentation is described by a subset of the initial set of labels and is an implicit oversegmentation since no explicit number of regions guides the segmentation. As shown in figures 1(e)-(h), the segmentation is described, in all cases, by a number of regions which is much larger than the right number of regions: figure 1(e) is segmented into 98 regions, 1(f) is segmented into 51 regions, 1(g) is segmented into 58 regions and 1(h) is segmented into 53 regions. On the other hand, our method allows regions parameters to vary jointly with the segmentation; thus, the initial set of labels is arbitrarily chosen and the number of regions is fixed equal to the right number.
In addition to that, our method realizes a big improvement in term of running time since we deal with a small number of regions compared to the method of (Boykov et al., 2001). We show in Table 1 the running times of the method of (Boykov et al., 2001) and our method in the case of figures 1(a)-(d) and the corresponding energies at convergence. In all these examples, the region merging process allows segmentation of the images into the right number of regions and we obtain exactly the same energy and regions parameters at convergence in much less running time.

4.2 Real Data

Though the number of regions is not known, in general, for real data, and consequently cannot be fixed in advance, in most applications a maximum number of regions is known, i.e., we know that the actual number of regions is less than some maximum number. This maximum number is in all cases much smaller than the number of all possible grey levels as used in (Boykov et al., 2001). In this section, we deal with real images and, depending on the application, we initialize with a number of regions that we consider large enough to segment the image into an acceptable number of regions.
The method of (Boykov et al., 2001) is unable to segment these images into the desired number of regions and it gives, in all cases, an oversegmentation, in addition to the much larger time that it takes to converge. An image such as Figure 2(a) presents smooth variation of intensity inside the region of interest. Methods without control on the number of regions fragment that region into multiple segments. The image in figure 2(c) presents the biggest running time and takes 2.90 seconds which is very small compared to the method of (Boykov et al., 2001) that takes, for the same image, more than 220 seconds because it deals with all the possible grey levels and segments it into 200 segments.

Table 1: Running time and energies of the two methods for synthetic images.

Number of regions	CPU(s): Boykov et al.	CPU(s): our method	Energy: Boykov et al.	Energy: our method
5	136.12	2.36	15.738	17.640
4	110.01	0.31	14.210	15.701
3	132.40	0.28	27.295	32.784
2	124.90	0.12	32.396	40.650

5 CONCLUSIONS

We presented an image segmentation method which combines the continuous optimization methods efficiency in describing the segmentation with the computational efficiency of graph cuts methods. It allows initializing with just the needed number of regions when it is known or with a larger number otherwise. Parameters vary jointly with the optimization and convergence is relatively rapid.

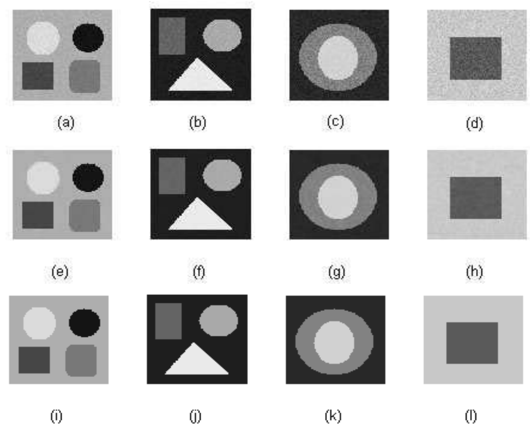

Figure 1: Synthetic images with gaussian noise: (a) 5 regions; (b) 4 regions; (c) 3 regions; (d) 2 regions; (e)-(h): segmentations with Boykov et al. method; (i)-(l): segmentations with our method. Images size: 128×128. $\alpha = 1.2$.

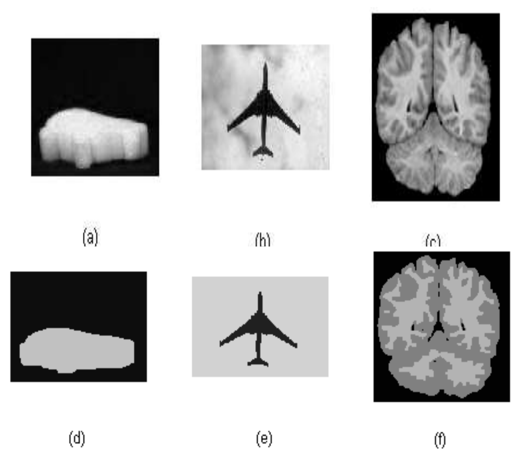

Figure 2: Real data: (a) 2 regions; (b) 2 regions; (c) 3 regions; (d)-(f): segmentations with our method. Images sizes: (a)154×115, (b)288×193, (c)142×144. $\alpha = 2.2$.

REFERENCES

Ayed, I. B., Hennane, I., and Mitiche, A. (2006). Unsupervised variational image segmentation/classification using a weibull observation model. *IEEE Trans. on Image Proc.*, 15(11):3431–3439.

Boykov, Y., Kolmogorov, V., Cremers, D., and Delong, A. (2006). An integral solution to surface evolution pdes via geo-cuts. In *ECCV (3)*, volume 3953, pages 409–422. Springer.

Boykov, Y., Veksler, O., and Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *IEEE Trans. on PAMI*, 23(11):1222–1239.

Chan, T. F. and Vese, L. A. (2001). Active contours without edges. *IEEE Trans. on Image Proc.*, 10(2):266–277.

Greig, D. M., Porteous, B. T., and Seheult, A. (1989). Exact maximum a posteriori estimation for binary images. *Jour. of the Roy. Stat. Soc.*, 51(2):271–279.

Schoenemann, T. and Cremers, D. (2006). Near real-time motion segmentation using graph cuts. In *Pattern Recognition (Proc. DAGM)*, volume 4174 of *LNCS*, pages 455–464, Berlin, Germany. Springer.

Zhu, S. C. and Yuille, A. (1996). Region competetion: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE Trans. on PAMI*, 18(9):884–900.