



Relaxed Cheeger Cut for Image Segmentation

Ludovic Paulhac, Vinh-Thong Ta, Rémi Mégret

► To cite this version:

Ludovic Paulhac, Vinh-Thong Ta, Rémi Mégret. Relaxed Cheeger Cut for Image Segmentation. International Conference on Pattern Recognition, 2012, Tsukuba, Japan. hal-00708970

HAL Id: hal-00708970

<https://hal.science/hal-00708970>

Submitted on 23 Dec 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.

Relaxed Cheeger Cut for Image Segmentation

Ludovic Paulhac¹, Vinh-Thong Ta¹ and Rémi Mégret²

¹Univ. Bordeaux, LaBRI, UMR 5800, F-33400 Talence, France

²Univ. Bordeaux, IMS, UMR 5218, F-33400 Talence, France

Abstract

Motivated by recent advances in spectral clustering that show the relation between the non linear p -Laplacian graph operator and the Cheeger cut problem, we propose to study and apply this methodology for image segmentation. Based on a ℓ^1 relaxation of the initial clustering problem, we show that these methods can outperform usual well-known graph based approaches, e.g., min-cut/max-flow algorithm or ℓ^2 spectral clustering, for unsupervised and very weakly supervised image segmentation. Experimental results demonstrate the benefits and the relevance of the proposed methodology especially for noisy image or when very few pixels are labeled for interactive image segmentation.

1. Introduction

Graph cut methods consist in dividing a dataset into two or more clusters where the data relations are modeled as a similarity graph. Such approaches have found several applications, for instance in image segmentation, statistics or biology (see [9] and references therein). Solving balanced graph cut problems is well-known to be NP-hard [13]. One of the most popular approximation is to consider spectral graph clustering. These methods are usually based on a ℓ^2 relaxation of the initial problem [8]. They reduce the initial cut problem to the computation and the thresholding of the second smallest eigenvector of the associated graph Laplacian operator. Different definitions of the Laplacian operator correspond to an approximation of different balanced graph cut criteria. We refer the interested reader to [13] for an introduction on spectral clustering. Recently, [3, 7, 11] have proposed a relaxation method based on the general p -Laplacian and show that the relaxed problem provides an exact solution of the Cheeger cut [4] when $p=1$. The authors show the superiority of such approach in comparison with the ℓ^2 relaxation based methods to cluster complex data.

Paper contributions. Inspired by these recent advances in graph clustering, we propose to study and apply such methods for image segmentation. We demonstrate the superiority of these approaches in presence of noise and propose an extension for interactive segmentation, which requires very few labeled pixels to obtain correct segmentation results. To the best of our knowledge, these novel methods have not been used for image processing. Therefore, one of our contributions is to demonstrate the relevance of these methods for image segmentation.

Paper organization. Section 2 recalls the Cheeger cut and its spectral relaxation. Section 3 presents the application of this methodology to image segmentation. Results and comparisons with usual graph cut methods such as min-cut/max-flow algorithm or standard ℓ^2 spectral clustering for noisy and interactive image segmentation are also provided. Finally, Section 4 concludes and presents the perspectives of this work.

2. p -Laplacian and Cheeger Cut

We consider the general situation where a data set is modeled as a graph $G = (V, E)$ where V is a set of vertices that represents the data points, and $E : V \times V$ a set of edges weighted by a function w_{ij} for all $(i, j) \in E$ that captures similarity relationship between the vertices. A method to find the balanced clustering of V is to consider the following Cheeger cut (CC):

$$\text{CC}(A, \bar{A}) = \frac{\text{cut}(A, \bar{A})}{\min(|A|, |\bar{A}|)} = \frac{\sum_{i \in A, j \in \bar{A}} w_{ij}}{\min(|A|, |\bar{A}|)} \quad (1)$$

where $A \subset V$, $\bar{A} = V \setminus A$ and $|\cdot|$ denotes the cardinality of the considered set. One simple interpretation of CC is that the weighted length of the cut boundary is small relatively to the balanced constraint. Because of the balanced condition, finding the global minimum of (1) is NP-hard. The most well-known approach to approximate the solution of (1) is to consider a ℓ^2 relaxation reducing the initial problem to an eigenvector problem associated to the ℓ^2 graph Laplacian [13].

Recently, [3, 7, 11] have introduced a stronger relaxation based on the p -Laplacian: let L_p be the graph p -Laplacian operator defined as

$$(L_p f)_i = \sum_{j \in V} w_{ij} |f_i - f_j|^{p-1} \text{sign}(f_i - f_j) \quad (2)$$

where $i \in V$, $\text{sign}(\cdot) \in [-1, 1]$ and $f : V \rightarrow [0, 1]$. In [3], the authors have shown that the optimal Cheeger cut value $h_{CC} = \inf_{A \subset V} CC(A, \bar{A})$ can be approximated by a relaxation based on (2). They have proved that when $p \rightarrow 1$ the solution of the relaxed problem tends to h_{CC} . Inspired by these results, [11] and later [6], have proved that when $p = 1$ the associated ℓ^1 relaxation problem leads to an exact solution of the CC minimization. Finally, they have proposed to optimize the following problem

$$f^* = \underset{f}{\text{argmin}} \left\{ \frac{\langle f, L_1 f \rangle}{\|f\|_1} = \frac{\sum_{i,j=1}^n w_{ij} |f_i - f_j|}{\|f\|_1} \right\} \quad (3)$$

s.t. $m(f)=0$ where $m(\cdot)$ is the median. To optimize (3), the split Bregman method is used in [11] whereas a primal-dual algorithm is used in [7]. Finally, f^* is thresholded (see for example [6]) to obtain the final graph partitioning. In this paper, we will consider the split Bregman approach.

3. Image Segmentation via Cheeger Cut

In this section, we describe and apply the previous graph cut methodology in the context of image segmentation and we compare the performance of these methods for unsupervised and interactive cases.

3.1. Methodology

Graph construction. An image is represented as a 4-adjacency grid graph where each vertex is associated to an image pixel and edges encode the 4-adjacency spatial relationship. The segmentation is performed only by using the pixel grayscale value as image feature in order to demonstrate the relevance and the efficiency of CC. For the same reason, the weight function uses only the pixel grayscale value. A weight $w_{ij} = \exp(-\|I(i) - I(j)\|^2 / 2\beta_{ij}^2)$ is defined for two adjacent pixels i and j where $I(\cdot)$ is a pixel grayscale value and β is a bandwidth parameter. In the rest of this paper, β will be set to the mean value over all $\|I(i) - I(j)\|$, $\forall i, j \in I$.

Obviously, incorporating higher level information that capture more complex image features in the graph structure and weights could improve the results but are out of the scope of this paper.

Unsupervised segmentation algorithm. Unsupervised image segmentation consists in extracting objects from images without any human intervention. Algorithm 1 describes the optimization of relaxed CC which is applied on the graph constructed on the image.

First, the solution f^* of (3) is computed in steps 2 to 6, where $R(\cdot)$ and $S(\cdot)$ correspond to the numerator and the denominator of (3) respectively. Then, the final partitioning (A, \bar{A}) is obtained with: $\forall i \in V$, $i \in A$ if $f_i^* \leq 0.5$, $i \in \bar{A}$ otherwise.

Algorithm 1 Approximation of (3) and h_{CC} .

- 1: $f^0 = \text{random}$, $\|f^0\| = 1$, $\lambda^0 = R(f^0)/S(f^0)$
 - 2: **repeat**
 - 3: $f^{k+1} = \underset{f}{\text{argmin}} R(f^k) - \lambda^k S(f^k)$
s.t. $m(f^{k+1}) = 0$ and $\|f^{k+1}\|_2 \leq 1$
 - 4: $\lambda^{k+1} = R(f^{k+1})/S(f^{k+1})$
 - 5: **until** $|\lambda^{k+1} - \lambda^k|/\lambda^k < \epsilon$
 - 6: $f^* = f^{k+1}$
 - 7: $(A, \bar{A}) = \text{thresholding of } f^*$
 - 8: **return** (A, \bar{A})
-

The inner problem in step 3 of Algorithm 1 is transformed into the following constrained minimization problem [2, 11]

$$\begin{aligned} \min_{f, g \in [0, 1], d} \|d\|_1 - \lambda \|g\|_1 \\ \text{s.t. } d = Df, g = f \text{ and } \sum_{i \in V} g_i < |V|/2. \end{aligned} \quad (4)$$

D is a linear operator with $\|Df\|_1 = \sum_{i,j=1}^n w_{ij} |f_i - f_j|$. The last constraint on $g \in [0, 1]$ corresponds to a linear approximation of the non linear constraint $m(g) = 0$ (see [2]). Optimization of (4) can be performed with an augmented Lagrangian method where three sub-minimizations w.r.t f , g and d are considered. Due to the lack of space, we refer the reader to [2, 11] for a detailed description of this split Bregman approach and the sub-minimizations resolution.

Adaptation to weakly supervised image segmentation.

Interactive segmentation has received a lot of attention in recent years. In its binary version, it consists in separating the image foreground from the background using initial labels (strokes) defined by the user for both objects to detect. One objective of this approach is to perform the segmentation with the fewest user interaction as possible. Moreover, most of such methods use prior information on images: for example, color likelihood, boundary prior or geometric constraints (see [12] and references therein).

In order to incorporate user provided initial labels, we slightly modify the optimization of (4). Let V_f and

V_b be the initial foreground and background labeled pixel sets. Function f^0 in Algorithm 1 is initialized as $f_i^0 = 0$ if $i \in V_b$, $f_i^0 = 1$ if $i \in V_f$ and $f_i^0 = 0.5$ otherwise. Throughout the optimization process, the values of g at the labeled pixels are maintained to their initial value at the end of each sub-minimization w.r.t. g , i.e., $g_i = 0$ and $g_i = 1$ if $i \in V_b$ and $i \in V_f$, respectively.

3.2. Results

In this section, we apply the previously describe methodology for unsupervised and interactive segmentation. In the following results, boundaries of obtained segmentations (presented in Figures 1, 2 and 3) are superimposed on initial test images and are artificially dilated to increase the visibility.

Unsupervised image segmentation. We first compare CC with the popular min-cut/max-flow algorithm (MC) [1] and ℓ^2 spectral methods: the Ratio Cut (RC) [5] and the Normalized Cut (NC) [10]. Results for MC are obtained by using publicly available C code from authors where the sink and source vertices are set to a pixel within the partitions to segment. Since we use [1] to optimize the min-cut problem, the data fidelity term involving in [1] energy is not used. For RC and NC, different 2-Laplacian operators are used, and the cut is obtained by thresholding the associated second smallest eigenvector [13].

Figure 1 shows unsupervised segmentation results with two synthetic examples where different additive Gaussian noise level are applied on binary images to illustrate the effect of noise. The σ parameters presented in these results are chosen in order to illustrate the noise level at which these methods begin to fail. The green, yellow, blue and red colors in Figure 1 correspond to MC, RC, NC and CC, respectively. To give a quantitative evaluation of the segmentation results, an error rate is computed. It is defined as a percentage of misclassified pixels with respect to the ground truth. Table 1 presents the corresponding measures computed for each segmentation in Figure 1 where best results are presented in bold.

Regarding the measures obtained in the first part of Table 1 that correspond to segmentation results of Figure 1(a)-(d), one can see that CC outperforms the other methods whatever the noise level. We can observe that MC fails up to $\sigma = 0.07$ where segmentation results correspond to the initialization involving in the algorithm (Figures 1(b) to 1(d)). As shown in Figures 1(c) and 1(d), RC and NC are not able to provide correct segmentation when the noise level increases. The second part of Table 1 and Figures 1(e)-(h) show approximately

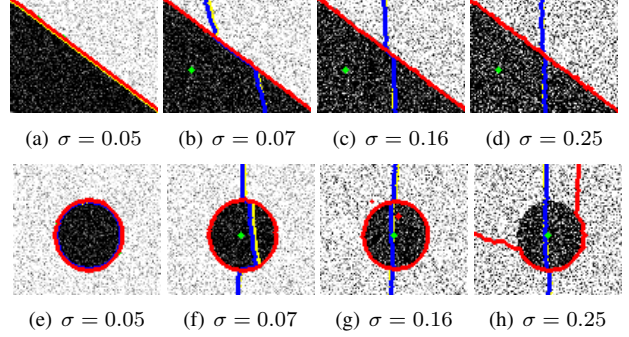


Figure 1. Qualitative comparison. MC: green, RC: yellow, NC: blue, CC: red (see text for more details)

	σ	MC	RC	NC	CC
(a)	0.05	0.02	0.02	0.02	0.02
(b)	0.07	48.87	11.88	11.48	0.06
(c)	0.16	48.87	22.88	22.58	0.22
(d)	0.25	48.87	24.89	24.68	0.46
(e)	0.05	0.02	0.02	0.02	0.02
(f)	0.07	21.5	38.54	38.54	0.02
(g)	0.16	21.5	48.35	47.89	0.07
(h)	0.25	21.5	48.72	48.00	18.69

Table 1. Error rates for segmentations of Figure 1 (see text for more details)

the same behaviors. Nevertheless, we can notice that for $\sigma = 0.25$, CC results are better on the first image than on the second one where objects are not of the same size.

The previous results clearly demonstrate the efficiency of CC for unsupervised noisy image segmentation. They also show the benefit of the ℓ^1 relaxation as compared to ℓ^2 one for balanced cuts approximations.

Very weakly supervised image segmentation. We show how CC can be applied to interactive image segmentation where very few labels (some pixels) are used as initialization and without any data fidelity or prior term while obtaining satisfying qualitative results.

Figure 2 shows results and comparisons of CC and MC. the boundaries of the segmentation results and the initial labels are superimposed on the initial images where white color are used for boundaries, red and blue colors for initial foreground and background labels, respectively (initial labels are artificially dilated to increase the visibility). Figures 2(a) and 2(c) show interactive CC segmentation results on natural images where

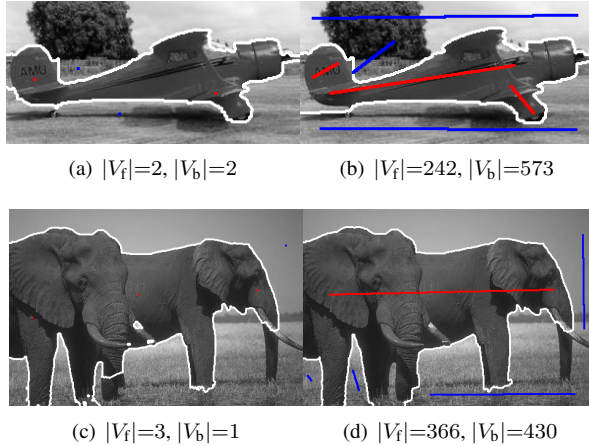


Figure 2. Results of interactive segmentation. At left CC with very few labeled pixels. At right MC with large number of labeled pixels (see text for more details)

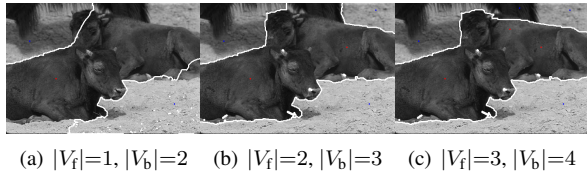


Figure 3. Iterative segmentation in three steps (see text for more details)

only 4 pixels are needed to obtain correct final segmentation. In our experiments, using the same few labeled pixels with MC produce totally incorrect segmentations similarly to Figure 1. Figures 2(b) and 2(d) illustrate the correctness of the MC segmentation but only after increasing the number of initial labels. Figure 2 clearly demonstrates the potential of CC in interactive segmentation allowing less human interaction while not requiring a global color model or prior as compared to usual approaches.

Finally, Figure 3 illustrates the potential of this methodology in iterative interactive segmentation process. Starting from 3 pixels to obtain a first segmentation, this method allows the user to add few pixels (Figures 3(b) and 3(c)) to have the final segmentation.

4. Conclusion and Future Work

In this paper, we have presented a novel methodology based on ℓ^1 relaxation of Cheeger cut problem for

image segmentation. Our study shows the potential and the relevance of this novel approach in a noisy context and for interactive segmentation where very few initialization is needed to obtain comparable result with the state of the art.

Due to the iterative optimization process, methods presented in this work are less efficient in term of computation time compared to, for instance, the min-cut/max-flow algorithm. Nevertheless, due to the locality of computations, it could be implemented on CPU/GPU platforms.

Finally, the ability of the proposed method to handle a very sparse initial labeling could be leveraged by allowing automated initialization by non dense interest point detectors. This opens new possibilities in the initial phase of the segmentation, when only little information is known about the objects to extract.

References

- [1] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE TPAMI*, 26:1124–1137, 2004.
- [2] X. Bresson, X.-C. Tai, T. F. Chan, and A. Szlam. Multi-class transductive learning based on ℓ^1 relaxations of cheeger cut and mumford-shah-potts model. Technical report, UCLA CAM-Reports, 2012.
- [3] T. Bühler and M. Hein. Spectral clustering based on the graph p -Laplacian. In *Proc. ICML*, pages 81–88, 2009.
- [4] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in Analysis*, pages 195–199, 1970.
- [5] L. Hagen and A. Kahng. Fast spectral methods for ratio cut partitioning and clustering. In *Proc. ICCAD*, pages 10–13, 1991.
- [6] M. Hein and T. Bühler. An inverse power method for nonlinear eigenproblems with applications in l-spectral clustering and sparse PCA. In *Proc. NIPS*, pages 323–343, 2010.
- [7] M. Hein and S. Setzer. Beyond spectral clustering - tight relaxations of balanced graph cuts. In *Proc. NIPS*, pages 2366–2374, 2011.
- [8] M. Nascimento and A. de Carvalho. Spectral methods for graph clustering - a survey. *European J. of Operational Research*, 211:221–231, 2011.
- [9] S. E. Schaeffer. Graph clustering. *Computer Science Review*, 1:27–64, 2007.
- [10] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE TPAMI*, 22:888 – 905, 2001.
- [11] A. Szlam and X. Bresson. Total variation and cheeger cuts. In *Proc. ICML*, pages 1039–1046, 2010.
- [12] S. Vicente, V. Kolmogorov, and C. Rother. Graph cut based image segmentation with connectivity priors. In *Proc. CVPR*, pages 1–8, 2008.
- [13] U. von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17:395–416, 2007.