Graph Cut and Image Segmentation using Mean Cut by Means of an Agglomerative Algorithm

Elaine Ayumi Chiba[1], Marco Antonio Garcia de Carvalho[1] and André Luís da Costa[2]
[1]*Computing Visual Lab, School of Technology - FT, University of Campinas - UNICAMP, Limeira - SP, Brazil*
[2]*Department of Computer Engineering and Industrial Automation, School of Electrical and Computer Engineering - FEEC, University of Campinas - UNICAMP, Campinas-SP, Brazil*

Keywords: Image Segmentation, Graph Partitioning, Cut, Mean Cut, Hierarchical Clustering.

Abstract: Graph partitioning, or graph cut, has been studied by several authors as a tool for image segmentation. It refers to partitioning a graph into several subgraphs such that each of them represents a meaningful object of interest in the image. In this work we propose a hierarchical agglomerative clustering algorithm driven by the cut and mean cut criteria. Some preliminary experiments were performed using the benchmark of *Berkeley* BSDS500 with promising results.

1 INTRODUCTION

Image segmentation is an important task in computer vision and image processing domains. It aims at partitioning an image into regions of interest for later analysis (Gonzalez and Woods, 2010).

There are several graph theory based techniques which are used in image segmentation. In particular, the graph cut techniques perform the segmentation by dividing a graph into disjoint subgraphs according to a given measure that takes in account the removed edges (Peng et al., 2013). There are different metrics to evaluate the graph's cut quality. Wu and Leahy (1993) proposed the first graph cut technique for image segmentation, where the graph cut value must be minimized in order to determine the optimal graph partition. However, the cut metric has the bias of finding small components. To address this problem other metrics were introduced, such as the normalized cut (Shi and Malik, 2000) and the mean cut (Wang and Siskind, 2001). The optimization of these metrics are problems with complexity NP-complete for general graphs. Therefore, Shi and Malik (2000) employed spectral graph teory (Cvetković et al., 2010) concepts for finding a graph cut with small normalized cut value, but not optimal. Wang and Siskind (2001) presented an algorithm capable of finding the graph cut with optimal mean cut value, but is restricted to planar graphs.

In a recent work, Costa (2013) proposed a novel algorithm for finding graph partitions with small normalized cut values. This new algorithm uses the normalized cut metric to guide the hierarchical clusterization of the graph nodes, until a given number of clusters are reached. The Costa's algorithm ensures that the subgraphs are connected and achieves a normalized cut value about 40 times smaller than the algorithm proposed by Shi and Malik. Furthermore, the computational performance of the new algorithm has inverse relation and is less dependent on the number of desired region than the former algorithm, which has increasing cost as raises the number of desired regions.

In this paper we utilize the Costa's (Costa, 2013) algorithm structure to create a hierarchical agglomerative clustering algorithm driven by the cut and the mean cut metrics. Although this new algorithm is not able to find the graph partition with optimal mean cut value, it is applicable to general graphs. Indeed, the algorithm's goal is not to optimize the cut measures but, instead, use them for directing the clustering process. Preliminary segmentations of the images from the Berkeley's segmentation benchmark BSDS500 (Arbeláez et al., 2011) are being presented.

The next sections are organized as follows: in Section 2 an overview of the general process of image segmentation by graph cut is given; in Section 3 we introduce the algorithm proposed in this work; the preliminary results obtained with the proposed algorithm are shown in Section 4; finally, in Section 5 are outlined some conclusions and perpectives for future works.

Chiba E., Carvalho M. and Costa A..
Graph Cut and Image Segmentation using Mean Cut by Means of an Agglomerative Algorithm.
DOI: 10.5220/0004858207080712
In *Proceedings of the 9th International Conference on Computer Vision Theory and Applications* (VISAPP-2014), pages 708-712
ISBN: 978-989-758-003-1

2 IMAGE SEGMENTATION BY GRAPH CUT

The general problem of image segmentation using graph cut techniques assumes an image graph representation $G = (V, E, W)$ as an undirected graph, where V is the set of nodes, E is the set of edges and W is the set of weights associated with the edges. Two nodes $u, v \in V$ are adjacent, represented by $u \sim v$, if there exist an edge $\{u, v\} \in E$ linking u and v (Wilson and Watkins, 1990). The graph nodes are associated to the pixels or to groups of pixels, *i. e.*, image regions. In the problem of minimization of the graph cut measures, a weight $w_{\{u,v\}} \in W$ must reflect the similarity between the image elements associated to the nodes u and $v \in V$. Thus we refer to this image representation as a similarity graph.

Let $G = (V, E, W)$ be a similarity graph and $A = (V_a, E_a, W_a)$ be a subgraph of G. The cut metric is defined as (Wu and Leahy, 1993; Peng et al., 2013)

$$\mathrm{cut}(A) = \sum_{u \in V_a, v \in V \setminus Va} w_{\{u,v\}}. \tag{1}$$

And the meancut metric is defined as (Wang and Siskind, 2001; Peng et al., 2013)

$$\mathrm{meancut}(A) = \frac{\mathrm{cut}(A)}{\sum_{u \in V_a, v \in V \setminus Va} 1}. \tag{2}$$

It is assumed that the best graph partition is the one with the minimal graph cut value (Wu and Leahy, 1993; Shi and Malik, 2000; Wang and Siskind, 2001; Peng et al., 2013). However, such partitions are nearly impossible to find in most cases. There are also several issues related to the graph cut metrics, such as the bias for small regions presented by the cut metric.

3 CUT AND MEAN CUT DRIVEN HIERARQUICAL CLUSTERING

The proposed hierarchical agglomerative algorithm works in a similar fashion to the classical linkage algorithms (Gower and Ross, 1969; Sibson, 1973) in the sense that on each iteration both aims at defining the edge $\{u, v\} \in E$ whose nodes will be merged. However, while the linkage algorithms rely on the edge weights to define such edge, Costa's algorithm rely on the node attributes. The dashed area in the block diagram shown in Figure 1 outline the steps involved on each iteration of the proposed clustering algorithm. Note that one node $u \in V$ is defined first and only after a second node $v \in V | v \sim u$ is determined. The criteria for defining the nodes u and v are distinct. Another difference to the linkage algorithms is that the edges incident to the merged node, and their correspondent nodes, must be updated after each merge in order to keep weights and attributes consistencies.

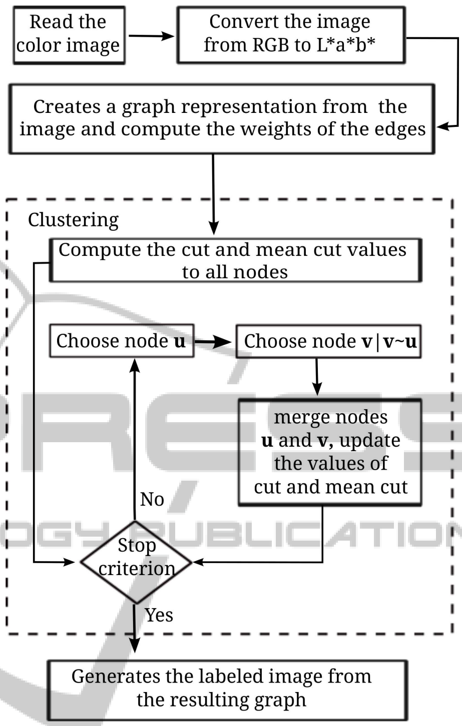


Figure 1: Outline of the image segmentation process using the proposed hierarchical agglomerative algorithm.

Let $G^n = (V^n, E^n, W^n, D^n)$ be the graph used in the iteration $n \in \mathbb{N}^0$, where the sets V^n, E^n, and W^n are defined similarly to the sets V, E, and W from G, and D^n is a set of degrees $d_{\{u,v\}} \in D^n$ associated to the graph edges. The edge degree set is used in the computation of the mean cut values for the nodes $u \in V^n$, as follow

$$\mathrm{meancut}(u) = \frac{\mathrm{cut}(u)}{\sum_{u \in V^n, v \in V^n \setminus \{u\}} d_{\{u,v\}}}, \tag{3}$$

where $\mathrm{cut}(u) = \sum_{u \in V^n, v \in V^n \setminus \{u\}} w_{\{u,v\}}$. The definition of the mean cut for a single node may be confusing at first, but remember that a node $u \in V^n$ corresponds to a whole subgraph $A \in G$.

3.1 Heuristic

The main difference from the Costa's algorithm to the algorithm proposed in this paper lies on the heuristic employed to determine the edge whose nodes are merged on each iteration. While the former uses the normalized cut, the second uses a combination of the cut and mean cut metrics.

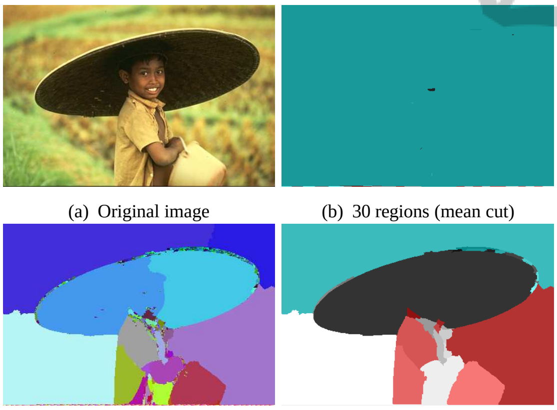
First Rule: this rule define the first node $u \in V^n$. At the begining of the clustering process the node u is defined such that

$$\arg\max_{u}(\text{meancut}(u)). \tag{4}$$

The reason for this criterion is that the node with highest mean cut value will have better chances to have its mean cut value decreased when merged with one of its neighbors. However, we observed that this criterion alone has the tendency to generate regions with unbalanced area, leading to partitions with very small regions along with very big ones. To address this problem the rule is switched when the number of nodes in the graph G^n reaches a given threshold $t \in \mathbb{N}^+$. The new rule selects a node $u \in V^n$ such that

$$\arg\min_{u}(\text{cut}(u)). \tag{5}$$

The new rule causes the smaller regions to be merged with one of its neighbors, yielding to a more balanced segmentation. Figure 2 show the unbalanced regions produced by the mean cut rule, as well as the more balanced result produced by mixing the two rules with $t = 1836$.



(a) Original image (b) 30 regions (mean cut)

(c) 1836 regions (mean cut) (d) 30 regions ($t = 1836$)

Figure 2: Segmentations produced using the mean cut rule and the mixing of the mean cut and cut rules. (b) segmentation by using the mean cut rule until 30 regions. (c) segmentation by using the mean cut rule until 1836 regions. (d) segmentation by mixing the mean cut and cut rules with $t = 1836$.

Second Rule: this rule define the second node $v \in V^n$, which must be adjacent to the node u already chosen. The node $v \in V^n | v \sim u$ is selected such that

$$\arg\min_{v}\left(\frac{\text{cut}(v)}{w_{\{u,v\}}}\right). \tag{6}$$

This criterion selects the neighbor of u with best relation of having the smallest cut value and the greatest similarity to u.

4 EXPERIMENTS

Firstly, we create an image graph representation by using the pixel grid model (Shi and Malik, 2000), where each graph node is associated to a pixel, and two nodes are connected by an edge if their associated pixels are within the radius $r = \sqrt{2}$. The edge weights are given by

$$w_{\{u,v\}} = e^{-\frac{\Delta C}{\sigma(\Delta C)}}, \tag{7}$$

where $\Delta C = \sqrt{(L_u - L_v)^2 + (a_u - a_v)^2 + (b_u - b_v)^2}$ is the color difference, in L*a*b* color space calibrated in D50, between the pixels associated to the nodes u and v. This color space was chosen because of its ability to mimic the nonlinear responses of the human eye (Gonzalez and Woods, 2010).

We used the *Berkeley's* segmentation benckmark BSDS500 (Arbeláez et al., 2011) to evaluate the results. In the BSDS500 dataset, the images are divided into a training set with 200 images, a validation set with 100 images and a test set with 200 images. The metrics F-measure and Region covering, that respectively evalute the segmentation boundary and the overlapping of regions, were used to compute the scores of each segmentation produced. The implementation of these metrics are available in the BSDS500 benchmark. The segmentation were performed in exactly 20 regions. This number was chosen after an analysis of the number of regions in the human segmentations from the training set.

The prelimirary experiments were conducted on the training set, and the goal was to find the threshold t that produced the best overall F-measure and Region Covering scores. Figure 3 show the scores obtained with different values for t.

The sizes of all tested images were 481x321, which give us a total of 154401 nodes in the graph G^0.

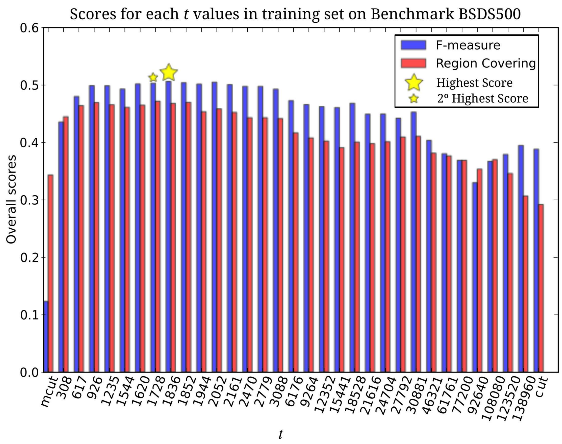


Figure 3: Experiment to find the appropriate parameter for the proposed algorithm.

The first search for the adequated threshold t divided the range $[0, 154401]$ in 10 evenly distributed. The t with the best scores was chosen as a pivot $\hat{t}$, then the same process was recursively performed in the range $[0, \hat{t}]$ until the best parameter was found.

The $t = 1836$ had the highest scores, this represents more than 90% of the clusterization process driven by the mean cut metric. However, using 100% of mean cut or using 100% cut yields to unsatisfactory results. Thus the combination of the two metrics is more suitable.

Figure 4 show a small sample of individual image segmentation into 20 regions using $t = 1836$.

Table 1 shows the overall scores for validation and test set of images, respectively first and second lines, and the scores for the images shown in Figure 4.

Even that the images 35028 and 227092 had low scores, as shown in Table 1, the segmentations shown in Figures 4(p) and 4(z) are visually good.

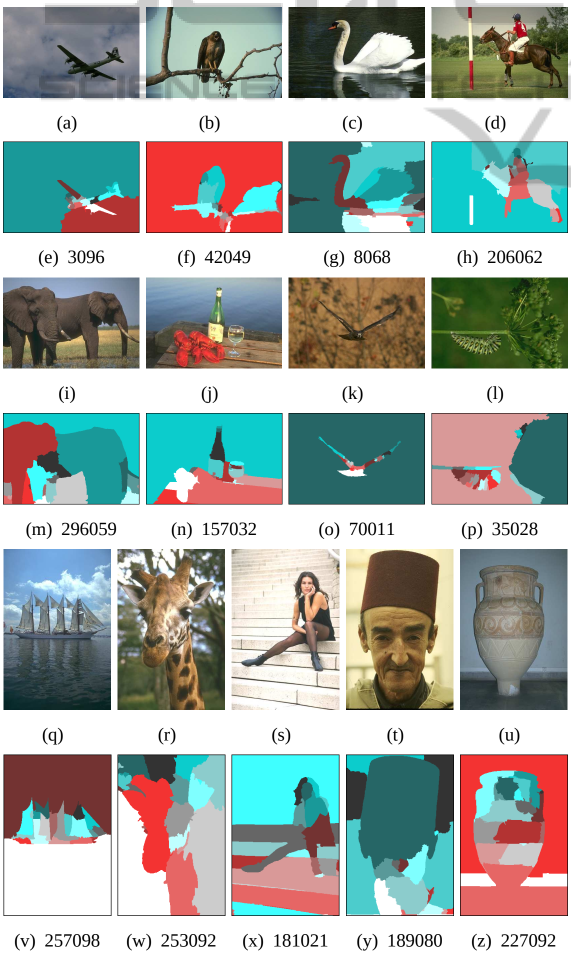


Figure 4: The images (a)-(d), (i)-(l), (q)-(u) are original images from the validation and test sets of the BSDS500 dataset. Images (e)-(h), (m)-(p) and (v)-(z) are the respective segmentation results using the proposed algorithm with $t = 1836$.

Table 1: Overall scores for validation and test set and the individual score for small sample os images

Images	F - measure	Region Covering
Validation set	0.482	0.433
Test set	0.493	0.431
2960591	0.746	0.601
8068	0.727	0.544
3096	0.713	0.701
257098	0.704	0.655
253092	0.699	0.404
42049	0.697	0.378
181021	0.679	0.405
157032	0.634	0.574
206062	0.636	0.445
70011	0.561	0.821
189080	0.590	0.391
35028	0.376	0.407
227092	0.444	0.297

In a First Comparison Study: we have segmented the BSDS500 images from the test and validation sets into 20 regions with both the algorithm proposed in this paper with $t = 1836$, and Costa's normalized cut algorithm. The overall results are shown in Table 2 for the test set and in Table 3 for the validation set.

Table 2: Overall scores for test set.

Algorithm	F-measure	Region Covering
Costa's	0.487	0.350
Proposed	**0.492**	**0.430**

Table 3: Overall scores for validation set.

Algorithm	F-measure	Region Covering
Costa's	0.479	0.344
Proposed	**0.481**	**0.432**

Table 2 and 3 shown that the proposed algorithm scores were higher than the scores from Costa's algorithm for this particular experiment. Remarkable, the score difference was greater for the region covering metric.

5 CONCLUSIONS AND FUTURE WORKS

In this paper we proposed a novel approach of hierarchical agglomerative clustering algorithm guided by mean cut and cut criteria to segment images. We showed the general structure and the heuristic of our algorithm. After many experiments, we found a threshold t that resulted in good segmentations. However, a fixed threshold may not be suitable for all images. Therefore, as future work we plan on defining an adaptative threshold that could be more adequated and lead to better segmentations.

The preliminary results were promising. However, other experiments and studies need to be performed in order to obtain more information about the algorithm behavior. This is an ongoing work and many related issues are not even being mentioned. We hope to give more contributions to the field in future publications.

ACKNOWLEDGEMENTS

This work is supported by CAPES Brazilian Agency.

REFERENCES

Arbeláez, P., Maire, M., Fowlkes, C., and Malik, J. (2011). Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5):898–916.

Costa, A. L. (2013). Normalized cut on graphs : an aglomerative algorithm for bacterial colonies image segmentation (in portuguese). Master's thesis, mentation (in portuguese). Masters thesis, School of Technology - FT, University of Campinas - UNICAMP, Brazil.

Cvetković, D., Rowlinson, P., and Simić, S. (2010). *An Introduction to the Theory of Graph Spectra*. Cambridge University Press, UK.

Gonzalez, R. and Woods, R. (2010). *Digital Image Processing*. Addison–Wesley, 3nd edition.

Gower, J. C. and Ross, G. J. S. (1969). Minimum spanning trees and single linkage cluster analysis. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 18(1):54–64.

Peng, B., Zhang, L., and Zhang, D. (2013). A survey of graph theoretical approaches to image segmentation. *Pattern Recognition*, 46(3):1020–1038.

Shi, J. and Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1:888–905.

Sibson, R. (1973). Slink: An optimally efficient algorithm for the single-link cluster method. *The Computer Journal*, 16(1).

Wang, S. and Siskind, J. (2001). Image segmentation with minimum mean cut. *Proceedings of the Eighth International Conference on Computer Vision*, 1:517 – 524.

Wilson, R. J. and Watkins, J. J. (1990). *Graphs, an introductory approach*. John Wiley & Sons.

Wu, Z. and Leahy, R. (1993). An optimal graph theoretic approach to data clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1101–1113.