

# A New Image Binarization Method Using Histogram and Spectral Clustering

Rui Wu<sup>1</sup>, Fang Yin<sup>2</sup>, Jianhua Huang<sup>1</sup>, Xianglong Tang<sup>1</sup>

<sup>1</sup>School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China

<sup>2</sup>School of Computer Science and Technology, Harbin University of Science and Technology, Harbin, China

## Abstract

We present a novel approach of binarization for gray images. The proposed algorithm uses the normalized graph cut(Ncut) as the measure for spectral clustering, and the weighted matrices used in evaluating the graph cuts are based on the gray levels of an image, rather than the commonly used image pixels. Thus, the proposed algorithm requires much smaller spatial costs and much lower computation complexity. Experiments on text images in natural scene show the superior performance of the proposed method compared to the typical thresholding algorithms.

**Keywords:** binarization; graph cut; spectral clustering

## 1. Introduction

In many image processing applications such as document processing, cell motion estimation and automatic target recognition, the object and background of images need to be separated, that is, binarization process, also known as thresholding process. The key of thresholding is how to choose the best value of threshold, and there are many methods of threshold choice such as the histogram technology [1,2], adaptive approach [3,4], the gradient method [5] and all relevant edge detection method

[6,7], and so on. Despite the various ways appearing in recent decades, choosing the optimal threshold automatically is still a pending problem in image processing.

Spectral clustering has gradually gained attention from research on text classification, images segmentation and information retrieval [8,9]. In [9] the proposed algorithm uses the normalized graph cut [8] measure as the thresholding principle to distinguish an object from the background and a large number of examples are presented to show the superior performance of the method. But it is still a thresholding algorithm that has limitations to deal with the cases when backgrounds have the similar color or intensity to that of the objects. In this paper, we propose a new image segmentation method based on gray histogram and spectral clustering. In our approach, the histogram of intensity is used for the object of grouping, we partition the image into two parts using the gray levels of an image rather than the image pixels. For most images, the number of gray levels is much smaller than the number of pixels. Therefore, the proposed algorithm occupies much smaller storage space and requires much lower computational costs and implementation complexity than other similar algorithms.

The rest of the paper is organized as follows. Section 2 introduces the theory of spectral graph partition briefly. Section

3 presents the individual steps of our approach. The experimental results are given in section 4. Finally, section 5 concludes the paper.

## 2. Theory of Spectral Graph

The basic method used by image segmentation is to view an image as a weighted undirected graph  $G = (V, E)$ , where the nodes of the graph are the points in the feature space, and an edge is formed between every pair of nodes. The weight on each edge,  $w(i, j)$ , is a function of the similarity between nodes  $i$  and  $j$ . A graph  $G = (V, E)$  can be partitioned into two disjoint subsets  $A$  and  $B$ , subject to  $A \cap B = \emptyset, A \cup B = V$ , by simply removing edges connecting the two parts. The degree of dissimilarity between these two pieces can be computed as total weight of edges that have been removed. In graph theoretic language, it is called the cut[8]:

$$cut(A, B) = \sum_{u \in A, v \in B} w(u, v) \quad (1)$$

The optimal bipartitioning of a graph is the one that minimizes this cut value. There are many criterions to measure the quality of the final partition results. Then the *Normalized Cut* value of a bipartition result can be defined as follows [8]:

$$Ncut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)} \quad (2)$$

where  $assoc(A, V) = \sum_{u \in A, v \in V} w(u, v)$ ,  $assoc(B, V) = \sum_{u \in B, v \in V} w(u, v)$  respectively, is the

total connection from nodes in  $A$  or  $B$  to all nodes in the graph. And now, the minimal *Ncut* value is just corresponding to the optimal bipartition of the graph. In order to minimize (2), we can transform the optimization problem into solving the eigenvalue system,

$$D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}z = Lz \quad (3)$$

where  $D_{ii} = \sum_{j \in N} w(i, j)$ ,  $W$  is a symmetric

matrix with size of  $N \times N$ ,  $L$  is the eigenvalue and  $z$  is the corresponding eigenvector. Shi and Malik[8] have proved that the second smallest eigenvector of the eigensystem (3) is the real value solution to the normalized cut problem of (2).

When the size of an image is too big, it is difficult to solve the above eigensystem, especially if the affinity matrix  $W$  is constructed by taking each pixel as a node, the size of eigensystem would be  $N \times N$  ( $N$  is the total number of pixels in an image).

## 3. Our method

Suppose

$$V = \{(i, j) : i = 0, 1, \dots, n_h - 1; j = 0, 1, \dots, n_w - 1\},$$

$$H = \{H_0, H_1, \dots, H_L\}, \quad LL = \{0, 1, \dots, L\},$$

where  $n_h$  and  $n_w$  is the height and the width of the image, respectively.  $H$  represents the histogram of gray,  $f(x, y)$  is the gray value of position  $(x, y)$ . Then,  $V, H$  and  $f(x, y)$  satisfy the following formulas.

$$(x, y) \in H_l, l \in \{0, 1, \dots, L\}, \quad \forall (x, y) \in V \quad (4)$$

$$H_l = \{(x, y) : f(x, y) = l, (x, y) \in V\}, l \in LL \quad (5)$$

$$\bigcup_{l=0}^L H_l = V, H_i \cap H_j = \emptyset, i \neq j, i, j \in LL \quad (6)$$

Using just the intensity value of the pixels and their spatial location, we can define the graph edge weight connecting the two nodes  $i$  and  $j$  as:

$$w_{ij} = e^{-\frac{\|F(i) - F(j)\|_2^2}{s_i}} * \begin{cases} e^{-\frac{\|X(i) - X(j)\|_2^2}{s_x}} & \text{if } \|X(i) - X(j)\|_2 < r \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

where  $F(i)$  is a feature vector based on intensity of node  $i$ , and  $X(i)$  is the spatial location at that node,  $s_i$  and  $s_x$

are scale factors used to adjust the variation of gray or spatial location between nodes,  $r$  is used to decide the number of nodes from node  $i$  to  $j$ .

And then, we can get a bipartition  $V = \{A, B\}$  corresponding to the graph  $G = (V, E)$ , where

$$A = \bigcup_{k \in L_A} H_k, B = \bigcup_{k \in L_B} H_k, \text{ and}$$

$$L_A \cap L_B = \emptyset, L_A \cup L_B = LL. \text{ Let}$$

$$cut(H_i, H_j) = \sum_{u \in H_i, v \in H_j} w(u, v) \text{ be the total}$$

connection weights from nodes in  $H_i$  with gray level  $i$  to all nodes in  $H_j$  with gray level  $j$ , we can rewrite the above formulas as:

$$cut(A, B) = \sum_{i \in L_A} \sum_{j \in L_B} cut(H_i, H_j) \quad (8)$$

$$asso(A, A) = \sum_{i \in L_A} \sum_{j \in L_A} cut(H_i, H_j) \quad (9)$$

$$asso(B, B) = \sum_{i \in L_B} \sum_{j \in L_B} cut(H_i, H_j) \quad (10)$$

Since  $asso(A, V) = asso(A, A) + cut(A, B)$ ,  $asso(B, V) = asso(B, B) + cut(A, B)$ , we can rewrite (2) as:

$$Ncut(A, B) = \frac{cut(A, B)}{asso(A, A) + cut(A, B)} + \frac{cut(A, B)}{asso(B, B) + cut(A, B)} \quad (11)$$

Given an image, we can construct a histogram-based matrix  $M$  that is similar to [9] by computing the all weights of nodes in the corresponding graph.  $M = [m_{i,j}]$  is an  $L \times L$  symmetrical matrix with  $m_{i,j} = cut(H_i, H_j)$  and  $m_{i,j} = m_{j,i}$ , where  $L$  is the number of gray level of histogram. Now, let  $M$  be the affinity matrix, we can get a complete approach of image segmentation using spectral clustering [8].

Note that the size of the affinity matrix  $M$  depends on the number of gray-level  $L$ , rather than the number of all pixels  $N$  in an image. Meanwhile, the

size of eigensystem to solve is  $L \times L$ , rather than  $N \times N$ , and usually,  $L$  with a fixed size is much smaller than  $N$ . Hence, the complexity of computation and spatial cost reduce greatly.

#### 4. Experimental Results

We perform a series of experiments to test the performance of this method. Our method is compared with two other methods: the *Otsu* thresholding method [10] and the *Ncut-based* thresholding method [9]. We choose them because the *Otsu* method is a simple but classic solution employed by many image segmentation schemes, while the latter is an *Ncut-based* but thresholding solution proposed recently. Firstly, we can see our segmentation result is almost same to the *Otsu*'s from Fig.1, and the parameter settings in formula (7) are  $s_l = 50$ ,  $s_x = 5$ ,  $r = 5$ ,  $L = 100$  which drawn from the experiments.

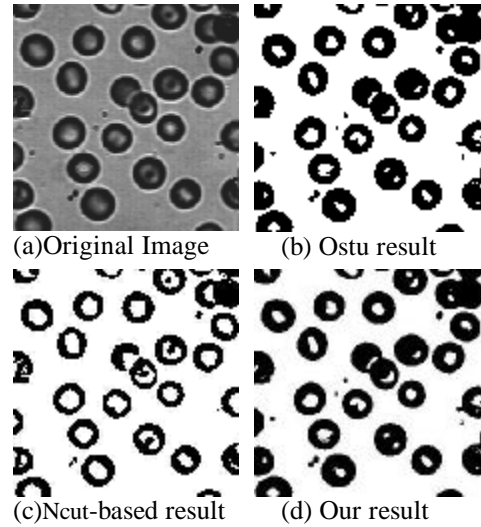


Fig.1. Comparison of three binarization results for the cell image

Secondly, we use the gray images with the text of characters for testing, and which are the real images from the

natural environment. From Fig.2, we can see the latter two methods both based on *Ncut* criterion are superior to the *Otsu* method from the experimental results. In Fig.3, The proposed method can be the right segmentation at a reflective white spots within the part of black spots on the letter 'B', which is difficult to achieve for the conventional thresholding methods.



Fig.2. Comparison of three segmentation methods for the text image

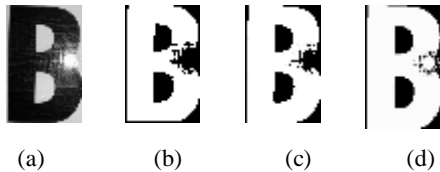


Fig.3. Comparison of three text segmentation methods for an illuminated image ((a)Original image (b) Otsu result (c) Ncut-based result (d) Our result)

Finally, we choose 150 images with clear differences and 50 images with blur text respectively to test the three methods. The aim of the three algorithms is to separate the “foreground (texts)” from the “background (non-texts)”. The precision  $p$  and recall  $r$  are used to evaluate the methods, which are defined as follows.

$$p = \frac{com\_Image_o \cap base\_Image_o}{com\_Image_o} \quad (12)$$

$$r = \frac{com\_Image_o \cap base\_Image_o}{base\_Image_o} \quad (13)$$

where  $com\_Image_o$  is a points set of objects including texts;  $base\_Image_o$  is the ground-truth text region. And then, the comprehensive assessment indicator  $f$  is defined as follows [11]:

$$f = \frac{1}{a/p + (1-a)/r} \quad (14)$$

where  $a$  is a relative rate between precision and recall, letting  $a = 0.5$ .

The result of 150 images shows in table I, and the result of 50 images shows in table II. From the results, we can see the proposed method has a good performance for both the two kinds of images. And our method is superior to other two methods from the point of comprehensive indicator  $f$ .

Table I The segmentation result of three methods for normal images

Method	$p$	$r$	$f$
Otsu method	0.8089	0.8361	0.8223
Ncut-based thresholding method	0.8364	0.8011	0.8184
Our method	0.8058	0.8802	0.8413

Table II The segmentation result of three methods for abnormal images

Method	$p$	$r$	$f$
Otsu method	0.5307	0.7734	0.6295
Ncut-based thresholding method	0.5815	0.8316	0.6844
Our method	0.6842	0.8015	0.7382

## 5. Conclusion

Image segmentation is a long-standing problem in image processing. Due to its simplicity and efficiency, thresholding is a widely used method for solving this problem. But, it could not handle the cases when backgrounds have the similar

color or intensity to that of the objects. In this paper, the proposed spectral clustering can do with the issue by using spectral graph theory. And this method controls the complexity of algorithm effectively by changing the clustering objects from pixels to gray levels. The experiment results from text images have proved its superiority to the traditional thresholding method.

## 6. References

- [1] Kapurj N., Sahoop K. ; Wong A. K. C., "A New Method for Gray-level Picture Thresholding Using the Entropy of the Histogram," *Computer vision, graphics, and image processing*, vol. 29, no.3, pp.273-285, 1985.
- [2] C.A.Glasbey, "An analysis of Histogram-based Thresholding Algorithms," *CVGIP: Graphical Models and ImageProcessing*, 55(6), pp.532-537, 1993.
- [3] Felix Abramovich, Yoav Benjamini, "Adaptive Thresholding of Wavelet Coefficients," *Computational Statistics&Data Analysis*, 22:351-361, 1996.
- [4] B. Gatos, I. Pratikakis, S. J. Perantonis, "Adaptive Degraded Document Image Binarization," *Pattern Recognition*, Vol.39, No.3, pp.317-327, 2006.
- [5] D Wang, "A Multiscale Gradient Algorithm for Image Segmentation Using Watersheds," *Pattern Recognition* 30(12), pp.2043-2052, 1997.
- [6] Mo Dai, Pierre Baylou, Louis Humbert, Mohamed Najim, "Image Segmentation by a Dynamic Thresholding Using Edge Detection Based on Cascaded Uniform Filters," *Signal Processing*, 52(1), pp.49-63, 1996.
- [7] Qiang Chen, Quan-sen Sun, Pheng Ann Heng, De-shen Xia, "A Double-threshold Image Binarization Method Based on Edge Detector," *Pattern Recognition*, 41(4), pp.1254-1267, 2008.
- [8] Jianbo Shi, Jitendra Malik, "Normalized Cuts and Image Segmentation," *IEEE Transactions Pattern Analysis and Machine Intelligence*, 22(8):888-905, 2000.
- [9] Tao Wen-Bing, Jin Hai, "A New Image Thresholding Method Based on Graph Spectral Theory," *Chinese Journal of Computers(in Chinese)*, Vol.30, No.1, pp.110-118, 2007.
- [10] Otsu N, "A threshold selection method from grey - level histograms," *IEEE Trans System. Man Cybernet*, SMC-9: 62 - 66, 1979.
- [11] Lucas S M, Panaretos A, Sosa L, et al, "ICDAR 2003 Robust Reading Competition," *Proc. of 7<sup>th</sup> Int'l Conference on Document Analysis and Recognition*, Scotland, pp.682-687, 2003.