MVA2015 IAPR International Conference on Machine Vision Applications, May 18-22, 2015, Tokyo, JAPAN

Unsupervised Figure-ground Segmentation using Edge Detection and Game-theoretical Graph-cut Approach

Yu-Min Hsiao
Computer Science
National Tsing Hua University
a26482289@yahoo.com.tw

Long-Wen Chang
Computer Science
National Tsing Hua University
lchang@cs.nyhu.edu.tw

Abstract

Figure-ground segmentation is to separate the object from background. It can be used in object detection or many other applications. Recently, a lot of methods have been proposed for solving figure-ground segmentation problems. However, most of them are supervised approaches. In other words, those methods need some interactions of users. It makes those methods unfavorable. For example, Graph-Cut needs users to select a part of foreground and background to be foreground seeds and background seeds. A graph and min-cut theory is used to separate the foreground from the image. We proposed an unsupervised figure-ground approach. It uses an edge-based method to grab required information for Graph-Cut. Then, we use game-theoretical Graph-Cut to divide the image into foreground and background. According to our experiment results, our method does not need user interaction and performs very well compared with the previous Graph-Cut method.

1. Introduction

Image segmentation is an important and difficult issue in computer vision and image processing. It usually uses similarity of pixels to divide the original image into segments. The similarity of pixels can be defined by histogram, color deference or texture deference. Image segmentation can be solved by clustering approach which cluster pixels to labels. Figure-ground segmentation is one kind of image segmentation which separates foreground from the image.

The result of figure-ground segmentation can be used in object detection or many applications. For examples, Carreira et al. use constrained parametric min-cuts segmentation and machine learning to achieve object detection [5]. Vijay et al. apply figure-ground segmentation algorithm to detect human faces [6]. Recently, a lot of methods have been proposed for solving the figure-ground segmentation problem. Most of them are supervised approaches. In other words, the procedures of those methods need interactions of users. Some methods let users to mark a stroke of foreground and a stroke of background [2, 10]. The others let users to mark a bounding box [1, 3, 4] and make those methods unfavorable.

We proposed an unsupervised figure-ground approach. It does not need user interactions. Furthermore, it increases the performance of Graph-cut based methods and makes the boundary of segmented foreground closer to the boundary of ground truth.

2. Proposed Method

Our approach divides into four steps shown in Fig. 1. First, we apply canny edge detection [7] for original image. The reason for using it is because canny edge detection would apply non-maximum suppression to let the width of edges fix to one pixel. After doing canny edge detection, we grab the pixels in two sides of the tangent line of each edge as shown in Fig 2.
We let $\|im1 - im2\|$ the color difference of those two side pixels be the boundary value for each edge, where im1 and im2 are the color values of two side pixels. We use a threshold to remove the edges with small boundary values, leave the edges with high boundary values to be true boundary. The color difference in two sides of the tangent line of boundary can be enhanced and it can eliminate the texture edges. The procedure of this step is shown in Fig. 3. When complete boundary detection, we divide the image into small patches, and only remain the patches which contain unbroken boundary. After that, we apply boundary-based segmentation on each patch. It use the boundary to separate pixels of a patch into two labels. For example, if the boundary in a patch is a straight line from the middle of top to the middle of bottom, the pixels in the left side of boundary form a label and the others (the pixels in the right side of boundary) form another label. The boundary-based segmentation is shown in Fig. 4. After boundary-based segmentation, we merge labels to reduce the number of labels. If the boundaries in two patches can match, we merge the labels of those two patches. The merge step is shown in Fig. 5 and the whole edge-based procedures are shown in Fig. 6. After edge-based procedures, we have many labels information in the image. However, the graph-cut approach only need foreground seeds and background seeds. We need to apply a pairwise clustering method to cluster those labels into foreground seeds and background seeds. In this step, we choose dominant set clustering [8] to achieve pairwise clustering.

Dominant set clustering is a pairwise clustering approach based on game theory. Let the labels be clustered represents as an undirected edge-weighted graph $G = \{ V, E, w \}$ which does not have self-loop, and $V = \{ 1, 2, 3 \ldots n \}$ is the nodes of each label. We use the average color to be the feature of a label. The similarity between two labels is the distance of the features. In other words, the similarity between two labels is the average color difference between those two labels. We can represent that $E \in V \times V$ is the edges set and the weight of an edge is defined as the similarity between two labels.

From above step, we can define a similarity matrix A = (a_{ij}) :

$$a_{ij} = e^{-\|im_i - im_j\|_2}$$

where a_{ij} is equal to the weight of the edge between label node i and label node j. im_i is average color value of node i. If the edge (i , j) ∉ E, a_{ij} = 0. In game theory, we correspond every node to a player and correspond similarity to a payoff. The dominant set clustering method defines the average weighted degree for each label node i with regard to a cluster S as:

$$\mathrm{awdeg}_S(i) = \frac{1}{|S|}\sum_{j\in S} a_{ij}$$

We can observe that if S only contain i, awdegs(i) = 0. Also,

$$\varphi_S(i, j) = a_{ij} - \mathrm{awdeg}_S(i)$$

is defined. If the payoff which node j gives to node i greater than the average payoff which node i receives, it represent that node j gives node i payoff very much and φ_S (i, j) is positive. On the contrary, if φ_S (i, j) is negative, it means that node j only gives node i a little payoff. The measurement of how important the node i is in a cluster S can be defined by recursively

$$W_S(i) = \begin{cases} 1 & , \text{if } |S| = 1 \\ \sum_{j\in S\setminus\{i\}} \varphi_{S\setminus\{i\}}(i,j)\cdot W_{S\setminus\{i\}}(j) & , \text{otherwise} \end{cases}$$

The cluster S is a dominant set if:

1. W_S (i) > 0 , for all i ∈ S
2. $W_{S\cup\{i\}}$ (i) < 0 , for all i ∉ S

The elements in the dominant set S are all important with respect to S and the elements outside S are not important with respect to S.

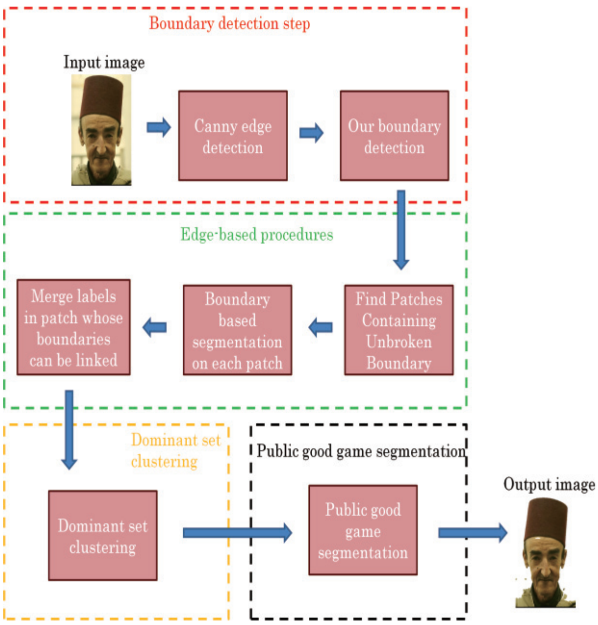


Figure.1. The flow diagram of the proposed figure-ground segmentation method.

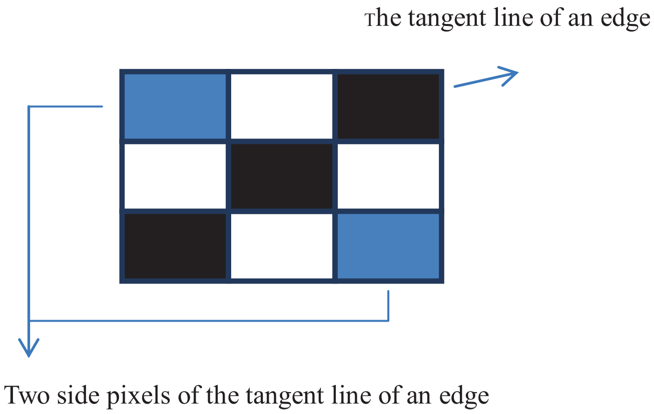


Figure.2.The pixels in two side of the tangent line of each edge.

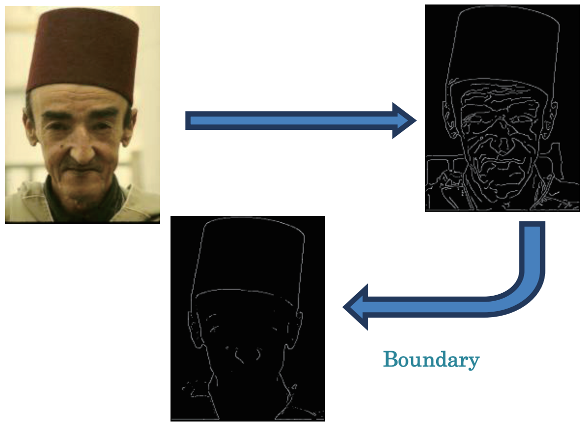


Figure.3 The procedure of our boundary detection.

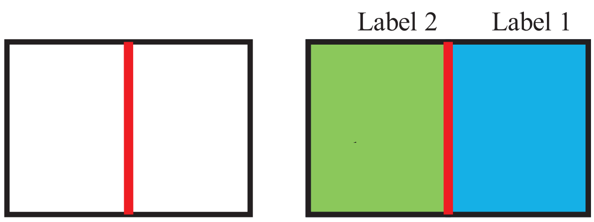


Figure.4 The boundary-based segmentation on a patch. (a) A patch contain unbroken boundary shown by a red line (b) The boundary based segmentation divides the patch into two small patches of label1 and label 2.

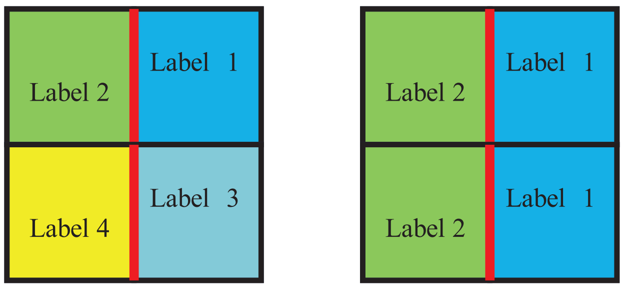


Figure.5 The merge step. (a) The boundaries in two patches that can match together. (b) The merged labels.

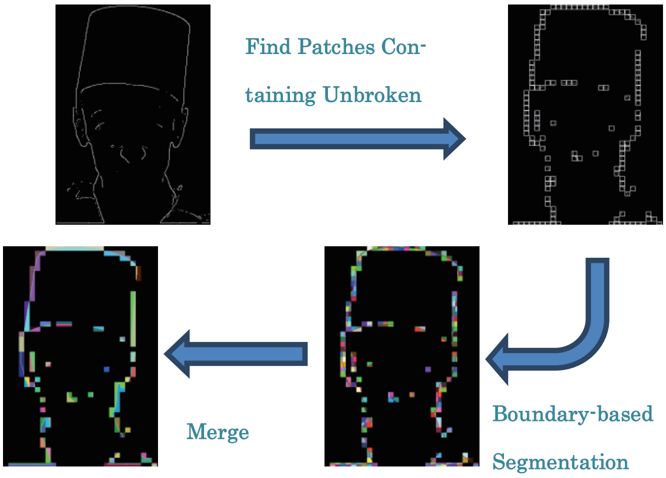
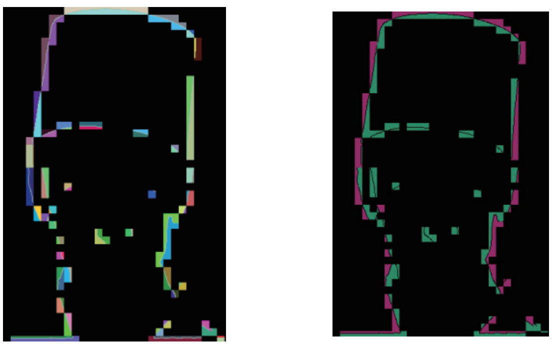


Figure.6 The edge-based procedures.



(a) (b)

Figure.7 Use dominant set clustering to cluster many labels into foreground label and background label. (a) Many labels where each color represent a label. (b) Foreground label and background label.

Fig. 7 shows that the dominant set clustering method clusters many labels into foreground label and background label. After dominant set clustering, we have foreground seeds and background seeds. We can apply those seeds to execute Graph-Cut based segmentation. In this step, we choose public-good-game segmentation [9] which is a Graph-Cut approach based on game theory. In this game, every player can choose two strategies, one is cooperate and another is defect. When a player chooses to cooperate, he will invest a cost to public good. Otherwise, if a player chooses to defect, he will not invest any cost to public good. Finally, the total cost which all cooperators contribute will multiply by an enhancement factor r and become total contribution. Then the total contribution will be allotted to all players. The payoff of each player is the contribution which he obtains. The above procedures of public-good-game continue until no player change his strategy. After public good game segmentation, we can obtain the label mask shown in Figure 8(a). With this label mask, we can extract the foreground from image as shown in Fig. 8(b).

(a) (b)

Figure.8: The example result of our method. (a) Label mask. (b) the segmented foreground.

3. Experimental Results

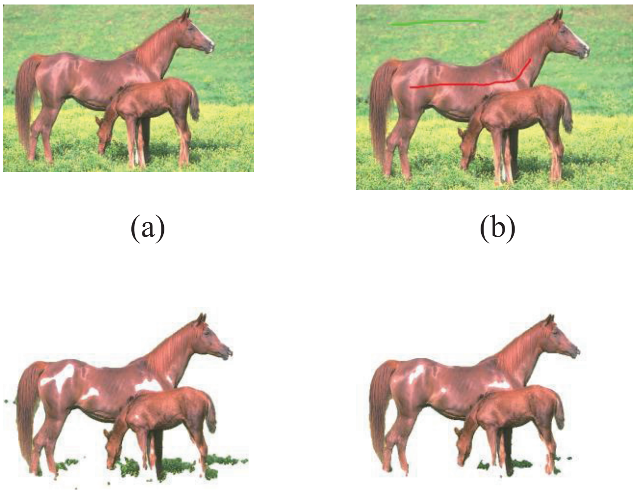
We show some results of our method and compare our method to original public good game segmentation with user interactions on the Berkeley segmentation dataset and icoseg data set. Berkeley segmentation dataset has several diverse images. We only choose the images which are appropriate to do figure-ground segmentation. In other words, the images which we choose from Berkeley segmentation dataset are only contain a consistent background and an object. The number of images we choose from Berkeley segmentation is 10 images. We also choose 10 images from icoseg data set to do our experiment. The final step of our method is public good game segmentation. We use the same parameters of the original public good game segmentation in our method. Fig. 9 shows our experiment results. We can observe that our method can segment the image smoother than the original public good game segmentation..

For measuring the accuracy of our method and the original public good game segmentation, we calculate cover and rand index on each result. Cover and rand index are a common measure method in image segmentation. The higher the value of cover or rand index is, the more accurate the result is. In Table 1 and Table 2 we can see that the accuracy of our method is mostly higher than the accuracy of the original public good game segmentation.

4. Conclusion

We proposed an unsupervised figure-ground approach. It does not need user interaction and makes Graph-Cut based approach convenient. It uses an edge-based method to grab required information for Graph-Cut based approach. However, it has too much information. We apply dominant set clustering, which is a pairwise clustering method, to cluster those information into foreground seeds and background seeds. Finally, we use a game-theoretical graph-cut approach, which is called public-good game segmentation, to divide the image into the foreground and background.

In our experiment results, our method does not need user interaction and performs well on the Berkeley Segmentation Dataset and iCoseg dataset. Also, our results are smoother than the original public-good game segmentation with user interactions.



(a) (b)

(c) (d)

Figure.9: The segmentation results of our method and the original public good game method with Berkeley segmentation data set. (a) Original images. (b) User interaction for original public good game method need. (c) Original public good game method. (d) Our method.

5. Reference

[1] V. Lempitsky, P. Kohli, C. Rother, and T. Sharp. "Image segmentation with a bounding box prior. " Page. 277 - 284 in ICCV, 2009.

[2] Viet-Quoc Pham , Univ. of Tokyo, Tokyo, Japan , Takahashi, K. and Naemura, T. "Bounding-Box Based Segmentation with Single Min-cut Using Distant Pixel Similarity." Pattern Recognition (ICPR), 2010 20th International Conference on, Page 4420 – 4423, 2010

[3] C. Rother, V. Kolmogorov and A. Blake. "GrabCut: Interactive Foreground Extraction using Iterated Graph Cuts. " ACM Transactions on Graphics (SIGGRAPH'04), Page 309 – 314, 2004

[4] Meng Tang, Lena Gorelick, Olga Veksler and Yuri Boykov. "GrabCut in One Cut." ICCV, Sydney, Australia, Page 1769 – 1776 , Dec. 2013

[5] Joao Carreira and Cristian Sminchisescu. "CPMC: Automatic Object Segmentation Using Constrained Parametric Min-Cuts." Computer Vision and Pattern Recognition (CVPR) , Page 1312 – 1328 , 2010

[6] H C Vijay Lakshmi and S. PatilKulakarni. "Segmentation Algorithm for Multiple Face Detection in Color Images with Skin Tone Regions using Color Spaces and Edge Detection Techniques." International Journal of Computer Theory and Engineering, Vol. 2, No. 4, Page 162 – 166 , August, 2010

[7] Canny, J., A. "A Computational Approach To Edge Detection" IEEE Transactions on pattern analysis and machine intelligence, VOL. PAMI-8, NO. 6, Page 679 – 698, November 1986

[8] Massimiliano Pavan and Marcello Pelillo. "Dominant Sets and Pairwise Clustering" IEEE Transactions on pattern analysis and machine intelligence, VOL. 29, NO. 1, Page 167 – 172, January 2007

[9] Jing Li, Gang Zeng, Rui Gan, Hongbin Zha and Long Wang. "A Game-Theoretical Approach to Image Segmentation." Computational Visual Media Lecture Notes in Computer Science Volume 7633, Page 33 – 42, 2012

[10] Wenxian Yang , Sch. of Comput. Eng., Nanyang Technol. Univ., Singapore, Singapore , Jianfei Cai , Jianmin Zheng and Jiebo Luo. "User-Friendly Interactive Image Segmentation Through Unified Combinatorial User Inputs." Image Processing, IEEE Transactions on, Page 2470 – 2479, 16 August 2010 [11]

ID	(2)	(3)	(4)	(5)
3063	0.94849	0.95252	0.96057	0.96033
3096	0.91765	0.90915	0.95788	0.94957
8068	0.91299	0.91062	0.90315	0.89747
42049	0.82927	0.84199	0.8619	0.86664
113044	0.6854	0.68287	0.89185	0.89267
147091	0.8662	0.86889	0.84824	0.85035
181021	0.85951	0.86841	0.87129	0.86782
189080	0.91303	0.91663	0.91178	0.91539
196027	0.67479	0.61487	0.71633	0.69404
253036	0.92904	0.93036	0.95869	0.958

Table.1 The results of 10 images in Berkeley segmentation dataset. The first column shows image ID. The second column and third column show cover score and rand index of the original public good game results. The fourth column and fifth column show cover score and rand index of our results.

ID	(2)	(3)	(3)	(5)
006	0.95378	0.94801	0.95139	0.9451
014	0.93968	0.92889	0.92827	0.91862
020	0.78775	0.80328	0.94985	0.95426
025	0.9754	0.97109	0.98305	0.98467
025	0.94898	0.93908	0.96054	0.96246
026	0.95254	0.94897	0.9723	0.97484
032	0.93947	0.93948	0.96657	0.96832
037	0.88366	0.87153	0.91507	0.92592
038	0.96321	0.96341	0.97397	0.97892
050	0.92586	0.92681	0.95648	0.9611

Table.2 The results of 10 images in icoseg dataset. The first column shows image IDs. The second column and third column show cover score and rand index of the original public good game results. The fourth column and fifth column show cover score and rand index f our results.