International Journal of Software Engineering and Its Applications
Vol.8, No.2 (2014), pp.409-416
http://dx.doi.org/10.14257/ijseia.2014.8.2.39

Object Segmentation using Mean-shift with Grid-mask for Grab Cut Algorithm

Kang Han Oh, Sooh Hyung Kim, In Seop Na* and Gwang Bok Kim

School of Electronics & Computer Engineering, Chonnam National University, 77 Yongbongro bukgu Gwangju, 500-757 Korea
blastps @naver.com, shkim@chonnam.ac.kr, ypencil@hanmail.net, loopaz63@gmail.com

Abstract

In this paper, we propose a novel method for automatic object segmentation using Mean shift with Grid-mask for Grab Cut algorithm. The main idea of proposed method is a rapid and exact technique for extracting initial foreground information using Mean shift with Grid-mask then we make initial rectangle around estimated initial foreground. And then Grab Cut algorithm is applied to segment foreground form background based on initial rectangle provided by previous process. In order to evaluate our proposed method, we compare the proposed method with several competitive automatic methods on the MSRA database which has 1000 images with ground truth. The scheme successfully segments objects without prior knowledge with both higher precision and better recall than competitive methods.

Keywords: *Object detection, Segmentation, Mean-shift, Grab Cut*

1. Introduction

As the smart phones mounted with high resolution cameras are widely used in the world, the smart phones can get information for users from camera image. For example, smart phone applications are capable of providing users with information of unknown object by using captured image. Object segmentation is one of the fundamental steps for a number of advanced systems in computer vision such as object recognition, scene reconstruction, image editing, and image retrieval, *etc*. Due to the strong influence of segmentation results in these advanced systems, many researches have published numerous approaches on object segmentation. However, most of these methods need some users' interactions as a prerequisite to control their progress. [1-2].

In order to segment Object area, we propose an approach for detecting potential object area to segment object using Grab Cut. In this paper, Section 2 introduces related work. Proposed method is explained in Section 3. Section 4 shows performance of our method.

2. Related Work

Several methods have been proposed for segmenting object images; this section finds related studies carried out in this research topic. In [1], suggested unconstrained object segmentation system using Saliency map, Gabor filter [10] and Grab cut. In this method, the goal is to generate an initial rectangle automatically for Grab cut. In order to create initial rectangle, author uses Gabor filter and Saliency map and then they uses

* Corresponding author In Seop Na (ypencil@hanmail.net)

ISSN: 1738-9984 IJSEIA

4 features(amount of area, variance, amount of class with boundary area , amount of class with saliency map) to categorize foreground and background. In [2], proposed an object segmentation method based on saliency map, Mean shift and level set method. First, a histogram based contrast method is used to generate the saliency map of the input image. Second, the input image is segmented into clusters using Mean shift. Based on the saliency map, segmented clusters are classified into background and foreground clusters. After that, an initial contour for level set method is determined by applying morphological erosion on foreground clusters. In [3], this approach proposed flower segmentation technic by a level set based on GMM initialization. First, a GMM [11] is applied to the input image to roughly specify flower area. Then a level set searches for the exact flower area based on the initial contour provided by the GMM.

3. Proposed Method

The proposed object segmentation scheme is organized as follows: 1. Mean shift clustering with grid mask, 2. rectangle initialization, and 3. Object segmentation using Grab cut. Figure 1 shows the flowchart of the proposed method

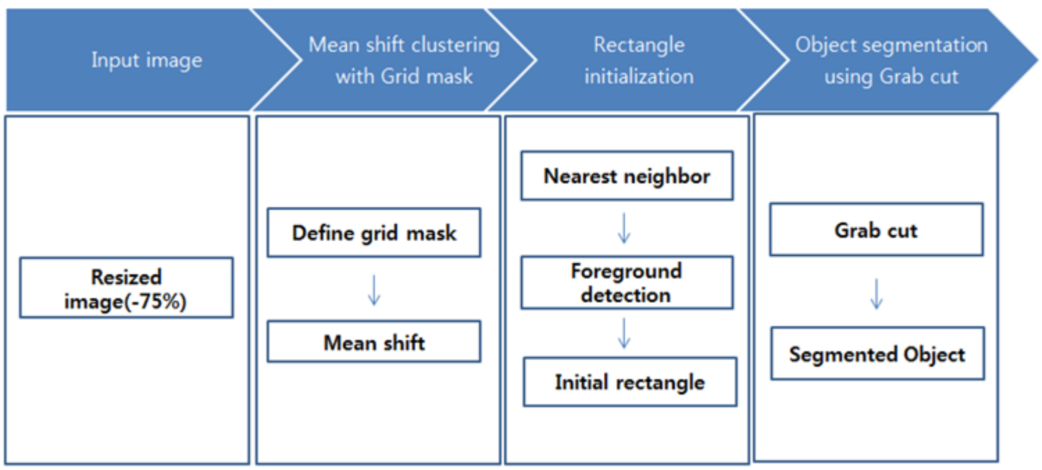


Figure 1. Flowchart of the Proposed Method

3.1. Mean Shift Clustering with Grid Mask

In this section, the goal is clearly estimating number of categorized component and center point of object. The Mean shift algorithm [8] is a very useful nonparametric clustering technique which does not require prior knowledge of the number of clusters, and does not constrain the shape of the clusters. However it is very slow to categorize whole of the pixels from input image. So we made a 20×20 Grid mask (Figure 2) and then mean shit algorithm categorizes whole of pixels on the Grid mask to reduce process time. Because the information from input image on the Grid mask is enough for analyzing image using Mean shift algorithm. This model contains two step, they are:

Step 1: in order to use mean–shift algorithm, a Grid mask is drawn on the original image and then we consider only grid line to extract color features (LAB) to overcome mean-shift drawback. In this paper, average of mean-shift time consuming is 0.2 second on the grid line from reduced image (-75%).

Step 2: extracted features (LAB) from a Grid mask (Figure 2 – white pixel) are categorized by mean-shift. Then we can get number of categorized component and average of (LAB) color value.

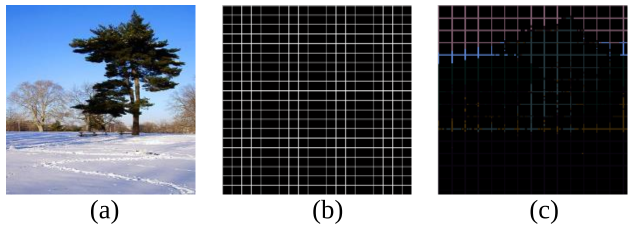


Figure 2. Mean Shift Clustering with Grid Mask, (a) the Input Image, (b) 20*20 Grid Mask (c) Mean Shift Result

3.2. Rectangle Initialization

From the previous step, suppose that we have K-number of classes with average of (LAB) color values and center point. In order to categorize whole of image, this information's are used for train data. In order to estimate foreground area, process has n steps:

Step 1: we use a Nearest Neighbor (NN) classifier to categorize whole of input image. As mentioned previous, the train data and test data are 5-dimensional features vectors (L, A, B, center point x, center point y). Here, the NN classifier equation considering K-number of classes is given by

$$Dist = \sqrt{\sum_{i=1}^{5}(x_i - t_i)^2} \tag{1}$$

$$class = argmin(\sum_{i=1}^{K} Dist(x, t_k)) \tag{2}$$

Where x_i is a input feature vector and t_i means train vector and equation (1) Dist is a Euclidian distance result. The equation (2) computes minimum distance with input features and train features. Figure 3 shows NN results.

Step 2: this step ultimate goal is to estimate classes which has high probability with foreground. Usually the background area has bigger variation in the pixel location than foreground and background classes are located in boundary area on the input image. So in order to detect foreground classes, we make a cost function with considered background features. The cost function given as,

$$E_i = nv_i + nb_i \quad i=1 \ldots K \tag{3}$$

where v means a variation of coordinate of pixels. It is determined as,

$$v = \frac{\sum_{(x,y)\epsilon C}(x-\mu x)(y-\mu y)}{area_{(x,y)\epsilon C}} \tag{4}$$

where area is the total area of class , x and y are coordinate in the class , C is the class. μx and μy are the average of the two axes. Here a normalization equation given by,

$$nv_i = \frac{v_i}{max(\sum_{j=1}^{K} v_j)} \quad i=1 \ldots K \tag{5}$$

where nv has a normalized value from 0 to 1. And then we check boundary area of image. The proposed boundary checking equation is as follows,

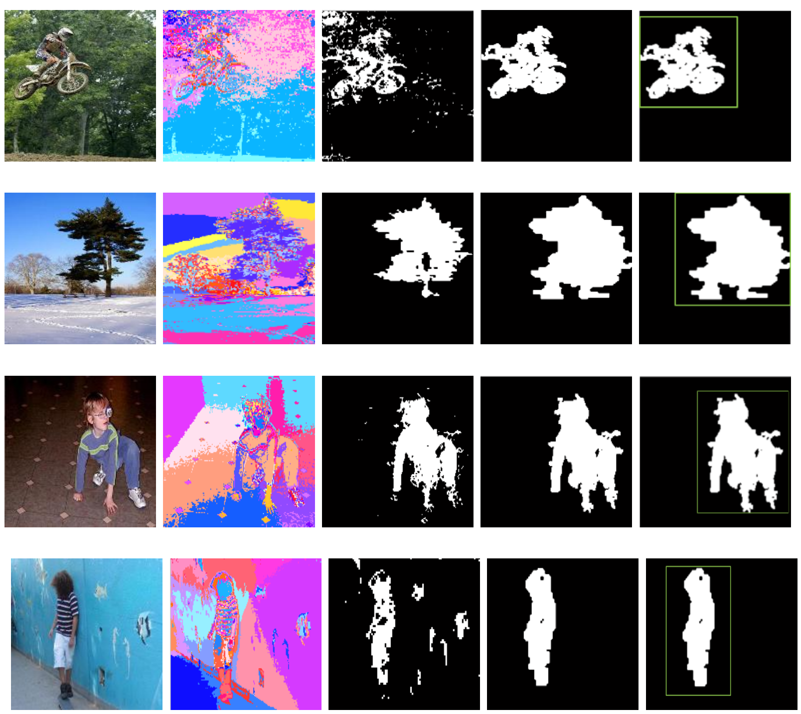
$$nb_i = \frac{C_i \cap \tau}{\tau} \quad i=1 \ldots K \tag{6}$$

Where C denotes each class area and τ is total boundary region. nb has a normalized value from 0 to 1.

Finally, from the E_i which has K-number results, we define threshold value using average of E_i and then equation (7) selects foreground classes. It is defined as,

$$\begin{cases} T = average(E_i) \\ \\ if \ E_i < \mathrm{T}, \ foregorund \\ otherwise, \ background \qquad (i = 1..K) \end{cases} \tag{7}$$

Figure 3 (c-d) shows a result sample of estimated foreground. Connected component algorithm remains biggest component and then we apply the result image with morphological dilatation to create rectangle around estimated foreground easy.



(a) (b) (c) (d) (e)

Figure 3. Rectangle Initialization (a) the Input Image, (b) NN Result (c) Defined Foreground (d) Morphological Dilatation and Remained Biggest Component Result (e) Rectangle Result

3.3. Object Segmentation using GrabCut

We use a GrabCut algorithm to refine the segmentation of object area using the estimated rectangle. The GrabCut algorithm [4] is semi-automatic segmentation algorithm that has recently gained popularity for use in object segmentation in computer vision. And this method is based on graph cut which is proposed by Boykov and Jolly [5]. The basic steps for the GrabCut algorithm are as follows: (1) users input initial rectangle around foreground. This rectangle provides feature that rectangle the inside area is unknown and the outside is BG. (2) the FG and BG area modeled as GMM using Orchard-Bouman clustering algorithm [5] (3) all pixel in the FG assigned most probable Gaussian component. The same process is done with the pixels in the BG. (4) new GMM parameters are learned from previous process. (5) A graph is built and Graph cut is used to find a new classification of FG and BG. (6) Repeat process 3-5 until the classification converges. In this section, the rectangle around object made form saliency detection part is applied in to GrabCut initial parameter. Figure 4 shows proposed ethod results easy.

4. Experimental Result

We implemented the proposed algorithm using MATLAB 2012 on a Intel(R) Core(TM)2 Quad CPU Q9550 to verify it's performance. The result of proposed method is evaluated on public MSRA datasets used by Liu *et al.*, [6]. MSRA datasets has 5000 color images and human labeled ROI around object. But it is not appropriate for evaluating performance. So we used the 1000 MSRA benchmark images in [7] which has ground truth. We used ground-truth mask categorized into background and foreground. To evaluate the performance of our proposed method, Average precision, recall and F-Measure are compared over the entire ground truth database, with F-Measure defined as

$$F_\beta = \frac{(1+\beta^2) Precision \times Recall}{\beta^2\ Precision + Recall} \tag{8}$$

We use β^2=o.3 to weigh precision more than recall.

Table 1. Segmentation Accuracies

Method	Dataset	Precision	Recall	F-Measure
[1]	MSRA (1000)	0.75	0.78	0.75
[2]	MSRA (1000)	0.71	0.67	0.67
[3]	MSRA (1000)	0.79	0.84	0.79
Proposed Method (without GrabCut)	MSRA (1000)	0.79	0.75	0.78
Proposed Method	MSRA (1000)	0.85	0.79	0.85

Table 1. shows segmentation results. We compare the proposed method with 3completive methods based on automatic object segmentation models. In the Table 1, proposed method (without GrabCut) means a evaluated performance with ground truth and estimated foreground (Figure 3 (d)) using equation(8). Our proposed method is the most competent among three methods. Especially, we can get high F-Measure score without Grab cut using

only estimated foreground result. Table 2 Shows the Mean shift process time and F-Measure form Number of grid mask.

Table 2. Accuracies and Time form Number of Grid Mask

Number of grid mask	F-Measure	Mean shift time(s)
5	0.76	0.1
10	0.81	0.1
20	0.85	0.2
50	0.84	0.4
70	0.78	0.5

5. Conclusions

In this paper, an automatic object segmentation method based on Mean shift with grid mask and Grab cut. We efficiently reduce Mean shift process time using the grid mask. The experiments have shown that our proposed algorithm can produce satisfactory accuracy in object segmentation without any prior information (initial rectangle) by user. Although the results are encouraging, future works are required to segment object area in more complex scenes and the initialization rectangle process should be accurate.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology(2013-022495)

References

[1] I.-S. Na, K.-H. Oh and S.-H. Kim, “Unconstrained Object Segmentation Using GrabCut Based on Automatic Generation of Initial Boundary”, International Journal of Contents, vol. 9, no. 1, **(2013)**, pp. 6-10.

[2] H. Le, S.-H. Kim and I.-S. Na, “Object Segmentation Based on Adaptive Contour Initialization for Level Set Method”, Proc. IEEE Int. Symp. Signal Processing and Information Technology 2012 (CD-Pub), Hochiminh City, Vietnam, **(2012)** December 8-13.

[3] S. C. Park, S.-H. Kim, G. S. Lee, C. W. Lee, I. S. Na, A. V. Ha Le and D. Yen, “Flower Area Segmentation by a Level Set based on GMM Initialization”, Proc. KSII 3rd Int. Conf. Internet (ICONI), Sepang, Malaysia, **(2011)** December 17-21.

[4] C. Rother, V. Kolmogorov and A. Blake, “Image Segmentation Based on GrabCut Framework Integrating Multiscale Nonlinear Structure Tensor”, Image Processing, IEEE Transactions, vol. 18, **(2009)**, pp. 2289-2302.

[5] Y. Boycov and M. Jolly, “Graph Cuts and Efficient N-D Image Segmentation”, International Journal of Computer Vision, vol. 70, **(2006)**, pp. 109-131.

[6] J. Harel, C. Koch and P. Perona, “Graph-Based Visual Saliency”, Advances in Neural Information Processing Systems, vol. 19, **(2007)**, pp. 545-552.

[7] T. Liu, J. Sun, N.-N. Zheng, X. Tang and H.-Y. Shum, “Learning to detect a salient object”, Pattern Analysis and Machine Intelligence, IEEE Transactions, vol. 33, **(2007)**, pp. 353-367.

[8] D. Comaniciu and P. Meer, “Mean shift: A robust approach toward feature space analysis”, IEEE Trans. Pattern Anal. Machine Intell., vol. 24, **(2002)**, pp. 603-619.

[9] A. Teuner, O. Pichler, B. Hosticka and J. Bedrich, “Unsupervised texture segmentation of images using tuned matched Gabor filters”, Image Processing, IEEE Transactions, vol. 4, **(1995)**, pp. 863-870.

[10] S.-H. Lee and J. Keun Seo, “Level set-based bimodal segmentation with stationary global minimum”, Image Processing, IEEE Transactions, vol. 15, **(2005)**, pp. 869-884.

Authors

Kangh Han Oh, he received his B.E. degree in Computer Engineering from Honnam University in 2010 and his M.E. degree in Electronic & Computer Engineering from Chonnam National University, Korea in 2013. He has been taking Ph.D course in Electronics & Computer Engineering at Chonnam National University, Korea. His research interests are pattern recognition, Signal processing and Object recognition.

Soo Hyung Kim, he received his B.S. degree in Computer Engineering from Seoul National University in 1986, and his M.S. and Ph.D degrees in Computer Science from Korea Advanced Institute of Science and Technology in 1988 and 1993, respectively. From 1990 to 1996, he was a senior member of research staff in Multimedia Research Center of Samsung Electronics Co., Korea. Since 1997, he has been a professor in the Department of Computer Science, Chonnam National University, Korea. His research interests are pattern recognition, document image processing, medical image processing, and ubiquitous computing.

In Seop Na, he received his B.S., M.S. and Ph.D. degree in Computer Science from Chonnam National University, Korea in 1997, 1999 and 2008, respectively. Since 2012, he has been a research professor in Department of Computer Science, Chonnam National University, Korea. His research interests are image processing, pattern recognition, character recognition and digital library.

Gwang Bok Kim, he received his B.E. degree in in Electronic & Computer Engineering from Chonnam National University, Korea in 2013. He has been taking the M.S. course in Electronics & Computer Engineering at Chonnam National University, Korea. His research interests are pattern recognition, machine learning and Image processing.

