http://dx.doi.org/10.5392/IJoC.2013.9.1.006

Unconstrained Object Segmentation Using GrabCut Based on Automatic Generation of Initial Boundary

In-Seop Na, Kang-Han Oh, Soo-Hyung Kim*
Chonnam National University, Korea

ABSTRACT

Foreground estimation in object segmentation has been an important issue for last few decades. In this paper we propose a GrabCut based automatic foreground estimation method using block clustering. GrabCut is one of popular algorithms for image segmentation in 2D image. However GrabCut is semi-automatic algorithm. So it requires the user input a rough boundary for foreground and background. Typically, the user draws a rectangle around the object of interest manually. The goal of proposed method is to generate an initial rectangle automatically. In order to create initial rectangle, we use Gabor filter and Saliency map and then we use 4 features (amount of area, variance, amount of class with boundary area, amount of class with saliency map) to categorize foreground and background. From the experimental results, our proposed algorithm can achieve satisfactory accuracy in object segmentation without any prior information by the user.

Keywords: *Gabor filter, Saliency map, Object segmentation, Graph Cut, Foreground, Background, Block Clustering.*

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 an initial and essential stage of image recognition scheme because perfect segmented image has good feature for recognition unknown object.

Over the past years, a number of studies have been made on object segmentation in natural scene image. This paper finds related studies carried out in this research topic. Nilsback and Zisserman [1] designed a flower recognition system using visual vocabularies which represent color, shape and texture feature of flower images. And in order to segment flower area from the natural scene image, the flower is automatically segmented using the concept of interactive graph cut [2] for optimal boundary and the region segmentation. And they addressed the problem of segmenting flower by using geometric model in [3]. This method extracts the petal structure of flower. And segmentation system is tolerant to changes in viewpoint and petal deformation, and the system is applicable in general for any flower class. In [4], the author proposed a method to segment flower using Saliency map(SM) [5] and GMM. This method estimate FG using K-means and Saliency map and then the GMM categorizes all pixels in the input image into 2 class FG and BG. Luca Bertelli [6] proposed a method for supervised learning approach for segmentation, where the only domain specific knowledge required is to select a specific image descriptor. With that and a set of training segmentations using SVM, this algorithm will automatically learn a discriminative model to segment a new image, without any user intervention.

In this paper, we present an approach called "Unconstrained Object Segmentation Using GrabCut Based on Automatic Generation of Initial Boundary". The goal of proposed method is to generate initial rectangle for GrabCut automatically. The proposed method is composed of 3 steps of initial segmentation, rectangle estimation using block clustering and object segmentation using GrabCut.

2. PROPOSED METHOD

This section is organized as follows. In section 2.1, in order to estimate foreground, we use Gabor filter with 5 classes and Saliency map and then section 2.2, system generates rectangle using Grab Cut result. Last section 2.3, the Grab cut algorithm classifies foreground from background. Fig. 1. shows the proposed system process.

* *Corresponding author, Email: shkim@jnu.ac.kr(S.H. Kim)*
Manuscript received Nov. 09, 2012; revised Mar 5, 2013; accepted Mar 15, 2013

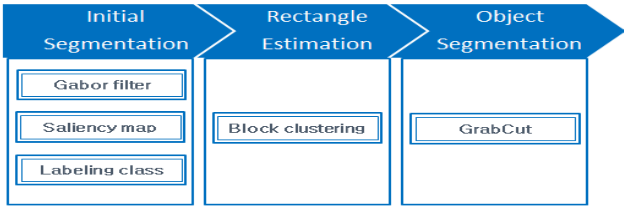


Fig. 1. Proposed system for object segmentation

2.1 Initial Segmentation

In order to estimate feature of FG and BG, a Gabor filter [7] creates 5 classes on the input image. The Gabor filter is a texture filter used in image processing. Frequency and orientation representation of Gabor filter are similar to human visual system, and it has been found to be particularly appropriate for texture representation and discrimination. In this paper, we use a 2D Gabor filter with 5 classes, and it is given by

$$g(x,y;\tau,\theta,\varphi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\tau}+\varphi\right) \quad (1)$$

Where τ represents the wavelength of the cosine, θ is the orientation of the normal to the parallel stripes of a Gabor function, φ is the phase offset, σ is the Gaussian envelope, and γ is the spatial aspect ratio. Fig 2 (c) shows categorized image using Gabor filter with 5. The Saliency map [8] algorithm is a good solution to define salient area in an image using various features (texture, gradient, color). The drawback of the algorithm in the object segmentation is that it is not able to find ever correct boundary pixel of interesting object but it can find correct location of object. Fig 2 (b) shows SM result.

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

Fig. 2. Proposed system for object segmentation
(a) Input image (b)SM result (c)Gabor filter result (d)Initial segmentation result

To define foreground and background in the Gabor filter result, system generates 4 features using the Gabor filter and SM. It has following stage: (1) the amount of area are considered to estimate FG and BG labels in the each class. Generally, amount of BG area has bigger than FG. (2) A variance of pixels can estimate information about FG and BG. It has same reason with stage 2 (BG variance bigger than FG).The variance of each class can be expressed as follows:

$$variance = \frac{\sum_{((x,y)\in C)}(x-\mu x)(y-\mu y)}{area} \quad (2)$$

here, the normalization equation is given by:

$$\alpha = \frac{variance}{\sum_{i=1}^{i=3} variance_i} \quad (3)$$

Where area is the total area of class , x and y are axes in the class , C is the class. μx and μy are the means of the two axes.

(3) In this stage, we check boundary area of image. Usually, most of image has boundary area with BG. The proposed boundary checking equation is as follows:

$$\beta = \frac{C \cap \tau}{\tau} \quad (4)$$

Where C is each class area and τ is total boundary region.

(4) In the saliency map result; we count number of each class and then we consider high probability with FG class which has biggest number of pixel. It is given by

$$\gamma = \frac{C \cap Saliency}{Saliency} \quad (5)$$

Where C is each class area and $Saliency$ is total area of SM result.

In the last of initial segmentation, we define FG and BG using 4 extracted features. To generate classes into foreground and background, a Threshold value is given by

$$T = \frac{1}{5}\left(\sum_{x=1}^{5}\left(area,\alpha,\beta,\frac{1}{\gamma}\right)\right) \quad (6)$$

And then equation (7) categorize in to 2 class using T value. Formulation is determined as

$$\begin{cases} N_i = \left(area,\alpha,\beta,\frac{1}{\gamma}\right)_i \\ if\ N_i < T,\ foregorund \\ otherwise,\ background \qquad (i = 1..5) \end{cases} \quad (7)$$

Fig 2 (d) shows generating FG and BG result using equation (7).

2.2 Automatic Rectangle Estimation

Second, the initial rectangle is estimated by Block Clustering using first step results. And it has following process: (1) RGB color average of FG and BG is trained by previous result, (2) we divide labeled image into 10*10 (100) blocks, (3) each Block will be categorized to FG and BG.

To decide Block class, we consider amount of each class area and distance of RGB color average from trained color features in the each Block to categorize FG and BG. Equation (8) shows each Gaussian has area in the 1 block.

$$B_f = \frac{G_f \cap block_i}{block_i},\ B_b = \frac{G_b \cap block_i}{block_i},\ i = 1 \dots 100 \quad (8)$$

where G_f and G_b are areas of FG and BG. $block_i$is 1 block from the divided image. And equation (7) measures distance between color average of input block and trained color average (FG, BG)

$$C_f = \sqrt{\left(ic_{rgb}-tcf_{rgb}\right)^2},\ C_b = \sqrt{\left(ic_{rgb}-tcb_{rgb}\right)^2}, \quad (9)$$

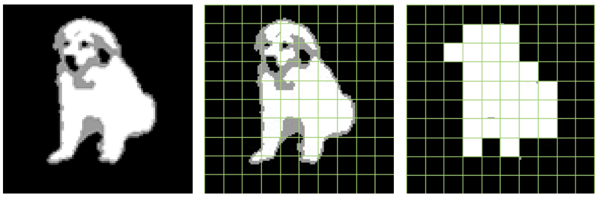
where ic_{rgb} is input color average. tcf_{rgb} and tcb_{rgb} are trained color average(FG,BG).And normalization equation is given by:

$$cd_f = \frac{C_f}{C_f + C_b}, \ cd_b = \frac{C_b}{C_f + C_b} \quad (10)$$

In order to categorize each block into two classes, the clustering formulation is determined as:

$$\begin{cases} if\ B_f + \frac{1}{cd_f} > B_b + \frac{1}{cd_b}, & block_i = foreground \\ otherwise, & block_i = background \end{cases} \quad (11)$$

And the final rectangle for GrabCut is decided by Block Clustering result. Fig 3. shows Rectangle Estimation process.

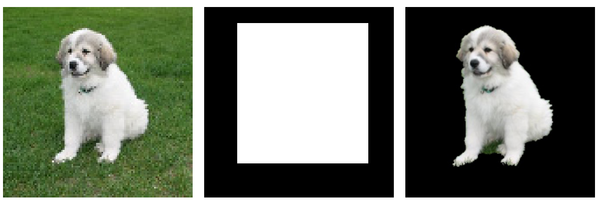


(a) (b) (c)

Fig. 3. Rectangle Estimation (a) input image (b) 10*10 Block in the Gabor filter result (c) Block Clustering result

2.3 Object Segmentation using GrabCut

In this stage, we use a GrabCut algorithm to refine the segmentation of object area using the estimated rectangle. The GrabCut algorithm [9] 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 [9]. 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 [10] (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 [11]. In this paper, the initial rectangle is provided by Box clustering stage for user. Fig 4. shows segmented image using GrabCut.



(a) (b) (c)

Fig. 4. Object segmentation using GrabCut (a) input image (b) estimated rectangle (c) segmented result

3. EXPERIMENTAL RESULTS

We implemented the proposed algorithm using C++ (opencv 2.3) on a Intel(R) Core(TM)2 Quad CPU Q9550 to verify it's performance. The experimental images in this paper consist of 1000 those from Achanta et al. [12] database. In order to evaluate the performance of proposed method, all of the images have ground truth labeled into FG and BG. The performance is measured by computing an overlap score S, between the ground truth and the segmented image (10).

$$S = \frac{(\text{trueforeground} \cap \text{segmented foregorund}) \times 2}{\text{trueforeground} + \text{segmented foreground}} \quad (12)$$

When evaluating this score, only foreground pixels are considered (i.e. the unlabeled pixels have no effect). The score is 1 for perfect segmentation. The segmentation results are compared with original GrabCut that has initial rectangle by user drawing and our proposed approach. The segmentation results are shown in Table 1(segmented result).

Table 1. Segmentation Result

GrabCut (semi-auto)	GMM	K-NN	Proposed Method
0.96	0.87	0.74	0.91

In Table 1, The GMM and K-NN expresses that GMM and K-NN algorithm categorizes input Image into foreground and background using trained data from rectangle estimation part. The proposed method little lower performance compare with original GrabCut. But our method not need initial rectangle for user convenient. Fig 5 shows some qualitative results and process of proposed method. Fig 5(a) is input image. (b) is corresponding saliency map. (c) shows Gabor filter result. (d) is 10*10 block . (e) is categorized block. (f) is initial boundary for Grabcut. (g) shows segmented result using Grabcut.

4. CONCLUSIONS

In this paper, we have demonstrated the object segmentation process using a new method based on GrabCut. The proposed method can make initial rectangle for automatic GrabCut. 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. In the past, we have to make an initial rectangle for GrabCut by hand. However, proposed our method make an initial rectangle automatically without human activity. This is the point out the contribution of paper.

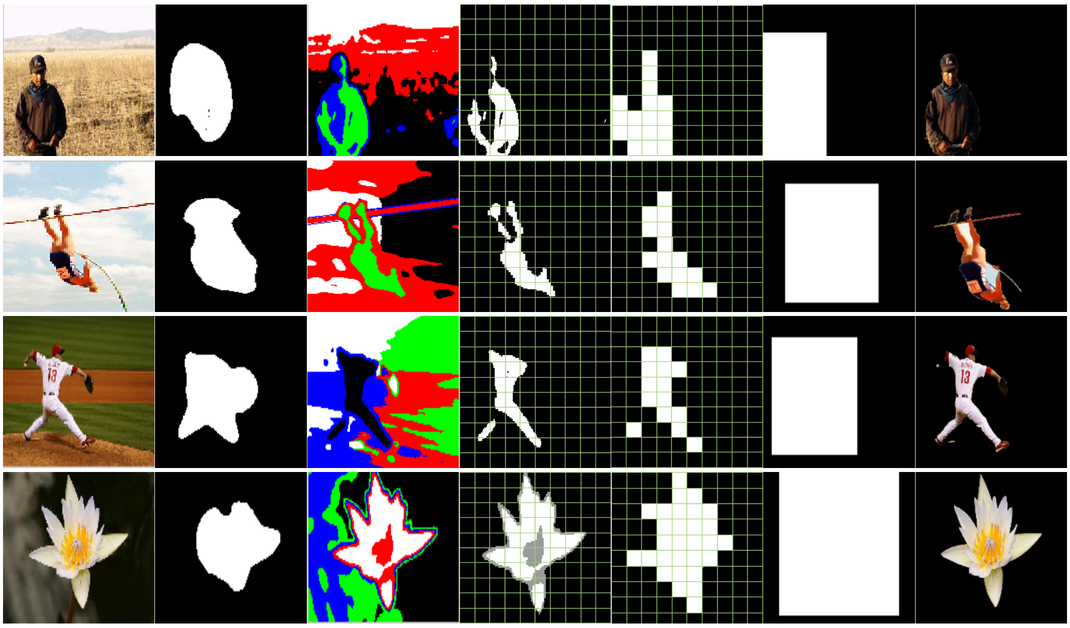


Fig. 5. Results of proposed method: (a) input image (b) SM result (c) Gabor filter with 5classes (d)10*10 block (e) result of block clustering (f) estimated rectangle for GrabCut (g) segmented image using GrabCut

ACKNOWLEDGMENT

"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(2012-047759) and This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency)" (NIPA-2012-H0301-12-3005)"

REFERNCES

[1] Nilsback, M. E. and Zisserman, A. A Visual Vocabulary for flower Classification. In the Proceedings of Computer Vision and Pattern Recognition, vol. 2, 2006, pp. 1447-1454.

[2] Boykov, Y.Y. and Jolly, M.P. 2011.Interactive graph cuts for optimal boundary and region segmentation of objects in N-D imagaes, In Proc. ICCV, volume 2, 2011, pp. 105-112.

[3] Bowman, M., M.-E. Nilsback and A. Zisserman. “Delving into the whorl of flower segmentation.” In BMVC, 2007.

[4] Kang Han Oh, Sang Cheol Park, In Seop Na and Soo Hyung Kim, "Flower Segmentation in Natural Scene Images," The Proc. 1st International Advanced Information Technology and Sensor Application, 2012.

[5] A Shokoufandeh, et al. View-based object recognition using saliency maps. Image and Computing, 17, 1999, pp. 445–460.

[6] L. Bertelli, T. Yu, D. Vu, and B. Gokturk. Kernelized Structural SVM Learning for Supervised Object Segmentation. In CVPR, 2011.

[7] V. Levesque, “Texture segmentation using Gabor filters”, Center for Intelligent Machines Journal, 2000

[8] Moosmann, F., Larlus, D., Jurie, F.: Learning saliency maps for object categorization. European Conference in Computer Vision (ECCV), Inter- national Workshop on The Representation and Use of Prior Knowledge in Vision 2006.

[9] C. Rother, V. Kolmogorov, A. Blake, "GrabCut: Interactive Foreground Extraction using Iterated Graph Guts," SCM Transaction on Graphics (TOG), 2004, pp. 1-6.

[10] Y. Boycov, M. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images," In Proc. IEEE Int. Conf. on Computer Vision, In ICCV, 2001, pp. 105-112.

[11] Tomas Malmer, “Image segmentation using GrabCut” May 12, 2010.

[12] R. Achanta, Freauency-tuned salient region detection. CVPR, 2009, pp.1597-1604.

[13] Kang Han Oh, In Seop Na and Soo Hyung Kim, "Flower Segmentation in Natural Scene Images,"

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 retrieval, image processing, pattern recognition, character recognition and digital library.

Kan Han Oh
He received his B.S. degree in Computer Engineering from Honam University in 2010, and his M.S. degree in Electronic & Computer Engineering from Chonnam National University, Korea in 2013. His research interests are pattern recognition, Artificial Intelligent, Image 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.