




An Improvement on GrabCut with CLAHE for the Segmentation of the Objects with Ambiguous Boundaries

Murat Aykut^(✉)  and Saffet Murat Akturk

Karadeniz Technical University, 61080 Trabzon, Turkey
murat_aykut@ktu.edu.tr

Abstract. Interactive image segmentation scheme provides an opportunity to select/mark initial region(s) of the target object(s) with user interactions. In this way, the foreground objects are segmented easily and successfully from the scenes which have cluttered backgrounds and multiple objects. GrabCut technique that utilize graph theory, Gaussian Mixture Model and iterative energy minimization can be considered in this context. This study concentrate on the weakness that occur on the low-contrast images. Using a contrast enhancement technique as a preprocessing step in GrabCut is proposed to improve the segmentation performance. CLAHE, which is a successful adaptive contrast enhancement method is used with RGB color channels in this work. Experimental results show that the proposed approach gives much better results (4% accuracy improvement) than the original GrabCut method on the images sampled from the Caltech 256 image dataset.

Keywords: GrabCut · Interactive image segmentation · Low-contrast images
CLAHE

1 Introduction

Interactive image segmentation, which uses prior knowledge by interacting with the users, is preferred in many applications due to its superior performance [1]. In this approach the users mark some regions belong to foreground and background and initialize the method with them. The most used interactive segmentation methods are GraphCut [2] and its variants (especially GrabCut).

The GraphCut method considers the image as a graph and utilize the mathematical operations and algorithms developed for graph theory. It equalizes the boundary and region features on all segments [2]. Many studies have been focused on the GraphCut and new methods have been developed. For example: Adding some other information to the energy functions [3, 4]; using region and boundary information together [5]. Among these methods the most interested one is the GrabCut method [1].

GrabCut method has improved the optimization of the GraphCut method, converted it to an iterative procedure and used the “border matting” to enhance the boundary

segmentation performance. It also ask for users to select a rectangular region only, as an initial interaction.

In recent years some works have been implemented to improve the performance of the original GrabCut segmentation method. In [6], the depth information is also used on the energy function. Similarly, in [7] the texture information is used additionally. Khattab et al. [8] enhance the method to give the ability to find more than two objects and reduce the user interaction. In [9], the authors proposed to use the saliency map.

GrabCut is a segmentation method based on iterative energy minimization that use the probability model for color distributions of pixels. Hence, unexpected results may occur when the boundaries between the foreground and background have low contrast. To resolve this problem, the energy function of the method is reformulated in [10].

In this work we propose to use contrast enhancement methods as a preprocessing step for the GrabCut method. Thus, we aim to improve the original GrabCut method's segmentation performance. From the contrast enhancement methods, Contrast Limited Adaptive Histogram Equalization (CLAHE) method [11] has been chosen and used for this study, because it operates on local regions and it is robust to noisy images. In the literature there are two works in which CLAHE and GrabCut methods have been used [12, 13]. On the other hand these works are application-specific and there are some other steps between CLAHE and GrabCut (in [12] CLAHE + brightness preserving dynamic fuzzy histogram + CLAHE + morphological operations + Gabor wavelets + GrabCut; in [13] preprocessing + CLAHE + thresholding and morphological operations + GrabCut have been proposed). Furthermore, they did not applied the CLAHE on all of the RGB channels (e.g. in [13] on the grey-level of the image only).

2 Methodology

2.1 GrabCut Method

GrabCut [1] is one of the best known GraphCut-based methods in the literature. Rother et al. proposed to use color information with GMM; minimize the energy function iteratively; and use incomplete trimaps, additionally. Basic steps of the method are:

Step 1. Initialize the trimap T which consists of known background T_B , unknown T_U and foreground T_F regions, by drawing a rectangle on the image. Outside of the rectangle is determined as T_B and the complement of it as T_U . The initial T_F is empty.

Step 2. Perform initial segmentation $\alpha = (\alpha_1, \dots, \alpha_p, \dots, \alpha_N)$:

$$\alpha_i = 0, \text{ for } i \in T_B; \quad \alpha_i = 1, \text{ for } i \in T_U \quad (1)$$

Step 3. Initialize two Gaussian Mixture Model (GMM), (one for background and other for foreground) with the previously obtained α segmentation.

Step 4. For each pixel in the unknown trimap T_U , find the most appropriate Gaussian components from the background and foreground GMMs, separately.

$$k_i = \arg \min k_i D_i(\alpha_i, k_i, \theta, z_i) \quad (2)$$

where $k_i \in \{1, \dots, K\}$ is an additional parameter assigned to each pixel to define the most likely GMM component of a pixel, K is the number of GMM components, and D_i is the data term of the Gibbs energy function. The D_i can be calculated by:

$$D(\alpha_i, k_i, \theta, z_i) = -\log \pi(\alpha_i, k_i) + 0.5 \log \det \Sigma(\alpha_i, k_i) + 0.5 [z_i - \mu(\alpha_i, k_i)]^T \Sigma(\alpha_i, k_i) [z_i - \mu(\alpha_i, k_i)] \quad (3)$$

where $\pi(\cdot)$ is the mixture weighting coefficient, μ is the mean vector, Σ is the covariance matrix and z_i is the intensity value of a pixel.

Step 5. Update the GMM parameters from the data previously clustered.

$$\theta = \arg \min_{\theta} \sum_i [D(\alpha_i, k_i, \theta, z_i)] \quad (4)$$

Step 6. To find the new clustering of the pixels perform the min cut algorithm as [1].

$$\min_{\{a_i: i \in Tu\}} \min_k E(\alpha, k, \theta, z) \quad (5)$$

where Gibbs Energy function can be derived by the formula:

$$E(\alpha, k, \theta, z) = \sum_i [D(\alpha_i, k_i, \theta, z_i)] + V(\alpha, z) \quad (6)$$

$$V(\alpha, z) = \gamma \sum_{(i,j) \in C} [\alpha_i \neq \alpha_j] \exp(-\beta \|z_i - z_j\|^2) \quad (7)$$

where V is the smoothness term of the energy function, C is a set of neighboring pixels, γ is the smoothness coefficient, β is a constant.

Step 7. Repeat steps 4–6 until the energy function converges to the predefined value.

In this work, we only use the hard segmentation process described above. The number of the Gaussian components, β constant, and the smoothness coefficient γ are determined empirically as 5, 5 and 50, respectively.

2.2 The Proposed Method

The original GrabCut method gives poor results for some images (especially for the images with low contrast object boundaries). To get rid of this weakness, Khattab et al. [10] proposed to reformulate the energy function of the method. On the contrary, we have not changed the original method, and proposed to use a contrast enhancement method before the original method as a preprocessing step.

In this work the Contrast Limited Adaptive Histogram Equalization (CLAHE) method is preferred for the contrast enhancement task due to these reasons: (1) it enhance the contrast of the local regions, which provides more information; (2) it is not greatly affected by the image noise because of the contrast limitation.

CLAHE Method

The CLAHE method is proposed by Pizer et al. in [11]. Main steps of the CLAHE can be summarized as follows:

Step 1. Divide the image into non-overlapping local regions (grids). Minimum size of the grid should be 32×32 . And calculate the histogram for all grids, separately.

Step 2. Clip the histogram to avoid being affected by the noise. If the number of pixels for any intensity value is greater than a predetermined threshold value, it should be fixed to that threshold. In this case, to equalize the total number of pixels, the clipped number of pixels are distributed to the histogram uniformly.

Step 3. Perform the histogram equalization on the histograms obtained in step 2. Combine the neighboring grids and use the bilinear interpolation to eliminate the boundary artifacts. For the pixels in the center of the grids use interpolation of the four neighboring pixels. Although, the original method is developed for the gray level images, it has been used for color images with different schemes. In this work we use the “CLAHE on RGB model” described in [14].

3 Experiments and Results

3.1 Data Set

To evaluate the performance of the proposed approach we began with constituting a data set. The images are selected from the Caltech 256 images dataset [15] according to the following criteria: (1) do not select the images that have uniform backgrounds; (2) maximally select one image per object category; (3) select images which have cluttered backgrounds and complex shaped objects. Although most of the studies performed the experiments on 25 images, 40 images are chosen for this work.

3.2 Experimental Results

The performance of the proposed method has been evaluated visually and quantitatively with the segmentation results of the images on the constituted data set. Six well known performance metrics were used in the experiments for quantitative evaluation: Accuracy, Dice Coefficient, Jaccard Index, Precision, Sensitivity, and Specificity. In the experiments the region that is marked as known background is not taken into account, because there is no possibility to segment erroneously. Table 1. shows the average values of the quality metrics, for the proposed method compared to original method.

It is clear from Table 1. that the proposed method achieved better results than the original GrabCut method for all the metrics. There is approximately 4% improvement with the proposed method.

To interpret the results much better, all the segmentation results of the images for both methods were analyzed visually. Figure 1 shows some images from the data set and their segmentation results in a comparative manner.

Table 1. Quantitative results of the proposed and original GrabCut methods on the image set.

Performance metric	Segmentation method	
	GrabCut	Proposed
Accuracy	0.9208	0.9580
Dice coefficient	0.9012	0.9557
Jaccard Index	0.8422	0.9117
Precision	0.8849	0.9564
Sensitivity	0.9149	0.9511
Specificity	0.9292	0.9698

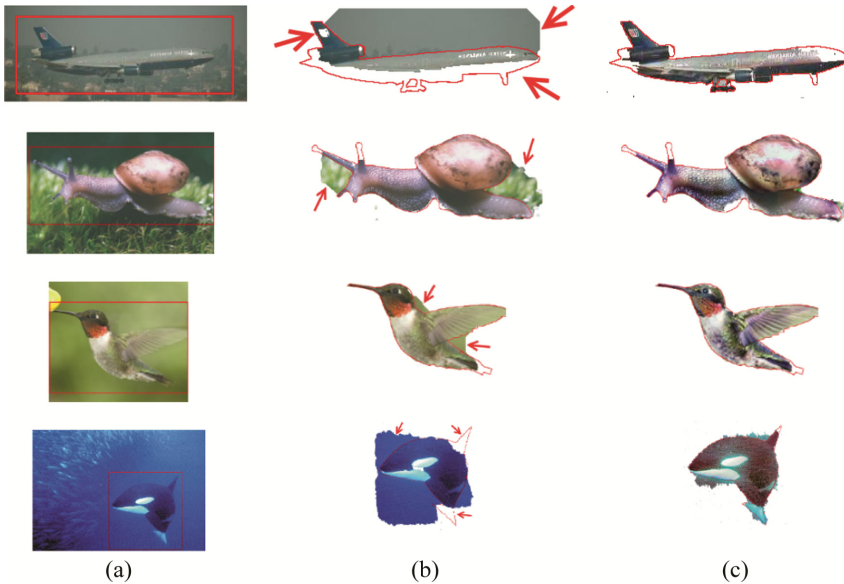


Fig. 1. Sample images and segmentation results for various scenarios. (a) Original image and initial bounding box; b, (c) Segmentation results of the GrabCut and the proposed method.

In the first two lines sample images which have low contrast are given. For these images, the proposed method gives better results noticeably. In the first image, although a part of the sky is included in the foreground; the bottom casing of the airplane and the logo of the airline company are not included in the foreground with the GrabCut method. In the second image, the snail is not successfully separated from the background with the original method. Third image contains a humming bird which flaps its wings at high frequency that results with some uncertainties on the boundaries. The proposed method also overcome this problem by enhancing the contrast in a local manner. Another scene is underexposed images taken under the sea. These images have low contrast and high noise as shown in the last line. Although the original GrabCut method gives weak performance, the proposed method provides sufficiently good results.

For the other schemes which are not visually shown above, both methods give very similar segmentation results.

4 Discussion

In this paper, an improvement to the GrabCut method is proposed to deal with the segmentation difficulties occurred when the images have low-contrast regions. The improvement has been done by using a contrast enhancement method (CLAHE on RGB colour bands) before the GrabCut method as a preprocessing step. The performance tests have been achieved on a data set consists of 40 images. According to the obtained results it is evident that the proposed method is superior to the original GrabCut method, especially for the low-contrast images. It has no drawbacks for the other images.

For the future work, the effect of the other contrast enhancement and image preprocessing techniques to the original GrabCut method might be studied.

References

1. Rother, C., Kolmogorov, V., Blake, A.: “GrabCut”: interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.* **23**, 309–314 (2004)
2. Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In: *International Conference on Computer Vision*, pp. 105–112. IEEE Press, Vancouver (2001)
3. Shan, J., Tu, J., Lu, X., Yao, J., Li, L.: Optimal seamline detection for multiple image mosaicking via graph cuts. *ISPRS J. Photogramm. Remote Sens.* **113**, 1–16 (2016)
4. Najjar, A., Gamra, S.B., Zagrouba, E.: Model-based graph-cut method for automatic flower segmentation with spatial constraints. *Image Vis. Comput.* **32**, 1007–1020 (2014)
5. Han, S., Chen, Q., Sun, Q., Ji, Z., Wang, T.: Image segmentation based on weighting boundary information via graph cut. *J. Vis. Commun. Image Represent.* **33**, 10–19 (2015)
6. Vaiapury, K., Aksay, A., Izquierdo, E.: GrabcutD: improved GrabCut using depth information. In: *ACM International Conference on Multimedia*, pp. 57–62. ACM Press, Firenze (2010)
7. Han, S., Tao, W., Wang, D., Tai, X.C., Wu, X.: Image segmentation based on GrabCut framework integrating multiscale nonlinear structure tensor. *IEEE Trans. Image Process.* **18**, 2289–2302 (2009)
8. Khattab, D., Ebied, H.M., Hussein, A.S., Tolba, M.F.: Multi-label automatic GrabCut for image segmentation. In: *14th International Conference on Hybrid Intelligent Systems*, pp. 152–157. IEEE Press, Kuwait (2014)
9. Kim, K.S., Yoon, Y.J., Kang, M.C., Sun, J.Y., Ko, S.J.: An improved GrabCut using a saliency map. In: *3rd Global Conference on Consumer Electronics*, pp. 317–318. IEEE, Tokyo (2014)
10. Khattab, D., Theobalt, C., Hussein, S.A., Tolba, F.M.: Modified GrabCut for human face segmentation. *Ain Shams Eng. J.* **5**, 1083–1091 (2014)
11. Pizer, S.M., Amburn, E.P., Austin, J.D., Cromartie, R., Geselowitz, A., Greer, T., Romeny, B.H., Zimmerman, J.B., Zuiderveld, K.: Adaptive histogram equalization and its variations. *Comput. Vis. Graph. Image Process.* **39**, 335–368 (1987)

12. Gutierrez, J.E., Barrena, J.T., Aroca, P.R., Valls, A., Puig, D.: Interactive optic disk segmentation via discrete convexity shape knowledge using high-order functionals. In: International Conference of the Catalan Association for Artificial Intelligence, pp. 39–44. UPC, Barcelona (2016)
13. Okuboyejo, DA., Olugbara, OO., Odunaike, SA.: CLAHE inspired segmentation of dermoscopic images using mixture of methods. In: World Congress on Engineering and Computer Science (WCECS), pp. 355–365. IAENG Press, San Francisco (2013)
14. Hitam, M.S., Awalludin, E.A., Yussof, W.N., Bachok, Z.: Mixture contrast limited adaptive histogram equalization for underwater image enhancement. In: International Conference on Computer Applications Technology (ICCAT), pp. 1–5. IEEE Press, Sousse (2013)
15. Griffin, G., Holub, A., Perona, P.: The Caltech-256 Object Category Dataset. Technical report, Caltech (2017)