

Segmentation of Similar Color Objects using Modified GrowCut Method

Sayan Bhowal¹, Bibhas Chandra Dhara²

Abstract

One of the most complex and prominent problems in the field of image processing is the image segmentation. Image segmentations are mainly of two types, namely fully automatic and interactive segmentation. Using fully automatic technique it is not always possible to produce a prominent and visually distinguishable output. As a result interactive segmentation gained popularity in recent years. GrowCut is a well known interactive image segmentation technique, which can localize the marked objects. In this present work, we have modified the GrowCut method so that all similar colored objects are segmented with respect to the marked objects. The proposed method is also faster than the GrowCut method.

Keywords

Image segmentation, interactive image segmentation, GrowCut, similar colored object segmentation, multiple object segmentation

1. Introduction

Segmentation of image into different regions is one of the most studied problems in image processing. Exact segmentation of the interested objects is a hard problem in the field of computer graphics and computer vision. There are two types of segmentation namely fully automatic and semi-automatic (interactive segmentation). Many works had been carried out on designing fully automatic algorithms [1], [2], [3], [4], [5], [6]. Fully automation is not possible unless knowledge of the world is modeled in the algorithm. Analyzing pixel values alone is not enough to accurately determine those segmentation curves (boundaries). Semi-automatic (or interactive image segmentation) [7], [8], [9], [10], [11], [12] involves minimal user interaction (marked some pixels of the foreground and some pixels of the background) to incorporate user intention into the segmentation process.

Sayan Bhowal, Department of Information Technology, Jadavpur University.

Bibhas Chandra Dhara, Department of Information Technology, Jadavpur University.

For a good interactive image segmentation algorithm, there are two basic requirements: (1) given a certain user input, the algorithm should produce intuitive segmentation that reflects the user intent; (2) the algorithm must be efficient so that it can provide instant visual feedback.

The present work is based on GrowCut [13] technique. In GrowCut, pixels are marked for both object and background. GrowCut method only segments the marked objects and disconnected similar colored unmarked objects are considered as background. In this work, GrowCut method is modified so that all the similar colored objects can be segmented, where some objects are marked as object. The proposed method is also faster than the original GrowCut method. The organization of the rest of the paper is as follows. The survey of the related works is given in section 2. The GrowCut method is described in section 3. The proposed method is presented in section 4. Experimental result is reported in section 5. Finally, conclusion is drawn in section 6.

2. Related Works

Interactive segmentation allows user to select some pixels as seeds and then some computations are involved to segment out the region of interest. Normally, interactive image segmentation techniques can be grouped into two classes: hard segmentation and soft segmentation. In hard segmentation [14], [15], [16], a pixel belongs to either foreground or background. While in soft segmentation [10], [17] a fuzzy matte is extracted. Interactive segmentation techniques can be region based, boundary based or hybrid concept of the both.

A typical example of the region based technique is magic wand [18], which is a common tool for many image editors. Another region based technique is intelligent paint [19], which is based on connect-and-collect strategy. This strategy handles the human-computer interaction to extract the region of interest. Recently, a region based interactive image segmentation method, MSRM, has been proposed in [16], [20]. The MSRM method is based on the result of the initial segmentation given by mean shift [21], [22]. Here, the user marks the foreground region and background region. Then, similarities between

different regions (given by mean shift method) are computed and adjacent regions are merged using maximal similarity rule.

Active contour (Snakes) [7], [23] is the most popular and well studied method among the boundary based techniques. Active contours are initialized manually with a rough approximation of the boundary of interest and then the energy minimization function locks the contour to the boundary. The energy function is a combination of internal force, such as boundary curvature, and external force, like image gradient magnitude. This minimization function computes within a loop and it is hard to control the shape of the boundary.

Intelligent scissors [9], [12], [24] is also well studied boundary based interactive segmentation method. Here, user specifies the starting point and end point of the boundary curve. In this method, each pixel is considered as a graph node and connected with its eight neighboring nodes. The boundary curve between the points is obtained by searching the shortest path. Another well-known boundary based technique is live-wire [25], [26] and its extension live-lane [25]. In these approaches, an image is considered as a directed graph where pixels are assumed as the graph nodes and oriented pixel edges are considered as arcs. A set of features are assigned to the arc and these values characterize the boundaries of the oriented edge. In live-wire, user first selects an initial point on the boundary and for subsequent point is indicated by the cursor; an optimal path from the initial point to the current point is found and displayed in real time. If the live-wire segment describes boundary appropriately, user deposits the cursor which becomes new starting point and the process is continued. In live-lane approach user involvement is more than live-wire. First, user selects the initial point and subsequent points are selected automatically as the cursor is moved in the vicinity of the boundary within a lane of certain width. Then, live-wire segments between successive selected points are computed and displayed in real time.

Algorithms like graph cut based methods [14], [27], [28] have used both regional and boundary properties. In graph cut, an image is modeled as a graph where each pixel is considered as the graph node and two neighboring nodes are connected with a weighted edge. A combinatorial optimization technique has been used in graph cut method. For two labels problem (foreground/background identification) seeds are specified by the user and

then optimal pixel labeling is done by using max-flow/min-cut algorithm. Grabcut [28] is an extended version of the graph-cut method. This is an iterative approach where graph-cut is used as the intermediate step. Grabcut method supports various types of user inputs including a bounding box to enclose the object of interest, foreground and background strokes for local editing, etc. In Grabcut, color statistics is computed and then image graph is re-weighted and graph cut is applied for refined segmentation.

Recently, random walks and its extensions [15], [29] have been adopted in various image processing tasks. In this approach, an image is modeled as a graph. Each edge is assigned a weight corresponding to the likelihood that a random walker will cross that edge. Random algorithm requires the input of foreground and background seeds. Random walks algorithm can achieve better segmentation performance than graph cut [15]. Random walks algorithm is sensitive to the position and quantities of the foreground and background seeds.

Another popular interactive segmentation approach is SIOX method [30], [31]. This method uses the color signature of the selected foreground and background area.

3. GrowCut method

GrowCut [13] is a cellular automaton based technique where the user specifies certain pixels as seeds for the object of interest and background. Here, the task is to assign label to all other pixels. The labeling process is an iterative technique. During the iteration, this method gives the feedback to the user and allows correcting the segmentation with user input where the segmentation is difficult to compute. The GrowCut method has some interesting properties like capable to solve moderately hard segmentation tasks, works with images of any dimensions $N \geq I$, performs multi-label image segmentation, is extensible (i.e. allowing construction of new families of segmentation algorithms with specific properties), interactively (i.e. segmentation is refined with each iteration user can observe the evolution and refine the segmentation “on the fly”), algorithm is simple in both understanding and implementation, allows fast parallel implementation.

3.1 Basics of GrowCut

The A cellular automata is an algorithm discrete in both space and time, that operates on a lattice of sites $p \in P \subseteq Z^n$. A cellular automaton is a

triplet $A = (S, N_g, \delta)$, where S is a non-empty set of states, N_g is neighborhood system, and $\delta: S^{N_g} \rightarrow S$ is the local transition function. The transition function δ defines the state at $t+1$ time step with help of the states of the neighbor cells at time step t . For the image, a cell (p) state $S = (l_p, \theta_p, \vec{C}_p)$ where l_p is the label of the current cell, $\theta_p \in [0,1]$ is the strength of the cell and \vec{C}_p is the cell feature vector.

For an image of size $k \times m$, the cellular space (P) is defined by an array of size $k \times m$. So, there is one-to-one correspondence between a pixel of the image and a cell of the automata and hence p is used to represent a pixel or a cell. The initial states for $\forall p \in P$ is $l_p=0, \theta_p=0, \vec{C}_p = (R, G, B)_p$ is the intensity profile of the pixel p . For the image, different neighboring system (N_g) can be used like 4-neighbors, 8-neighbors etc.

Before starting the segmentation, user specified the seeds and accordingly the label and strength of the seed pixels are determined. The seed strength can be controlled by the initial strength which is set by the user's stroke. For hard segmentation, strength is set to 1 and for soft segmentation initial strength is < 1 . The label and strength of p at step t is denoted as l_p^t and θ_p^t , respectively. The labels and strength of the cells are updated, at step $t+1$, using **Algorithm: Automata evolution rule** given below.

Algorithm: Automata evolution rule

```

1   for all  $p \in P$  // for each cell...
2    $l_p^{t+1} = l_p^t; \theta_p^{t+1} = \theta_p^t; //$  copy previous state
3       for all  $q \in N_g(p)$ 
4           if  $g(\|\vec{C}_p - \vec{C}_q\|_2) \cdot \theta_q^t > \theta_p^t$ 
5                $l_p^{t+1} = l_p^t;$ 
6                $\theta_p^{t+1} = g(\|\vec{C}_p - \vec{C}_q\|_2) \cdot \theta_q^t$ 
7           end if
8       end for
9   end for
    
```

Where g is a function bounded to $[0, 1]$ and considered as $g(x) = 1 - \frac{x}{\max \|\vec{C}\|_2}$.

This method starts from the seed pixels and try to label the all other pixels of the image. When the strength of the current cell (p) is less than the attacking force of the neighbor cell (q), then the attacking force and label of q are set as the strength and label of p . Attacking force of q depends on the strength θ_q and the distance between feature vectors \vec{C}_p and \vec{C}_q . In this local operation, the strongest attacker occupies the cell p and its label and corresponding attack force is set as the label and strength of p . This computation continues until automata converge to stable configuration (i.e. no change in state of any cell).

4. Proposed method

The GrowCut method is an iterative method, where in each of the iterations the strength and label of the cells are updated considering the label, strength and attacking force of the neighboring cells. It may be noted that unlabeled cell (where strength =0) has no impact in the labeling process (see line 4 of the **Algorithm: Automata evolution rule**). The strongest attacker occupies the current cell p and corresponding attacking force and label is set as the strength and label of p . This is not reflected in the given code. To reflect this line 4 of the **Algorithm: Automata evolution rule** needs to be modified. The correct statement would

be $g(\|\vec{C}_p - \vec{C}_q\|_2) \cdot \theta_q^t > \theta_p^{t+1}$. In GrowCut, the labeling starts from the seed pixels and executed in a loop, for all pixels. This labeling step is iterated if there is change in the cells state of the automata.

We have some observations on GrowCut method: i) the pixel labeling method is time consuming. Within loop of the iteration, pixels are labeled layer-by-layer (as unlabeled pixels have no influence to change the state of any pixel); ii) the state of an unknown pixel is influenced by the closest seed pixel. So, it is not always possible to get the desired output, i.e., some foreground part of the image may be covered by the background and conversely (see Fig.1 (b)); iii) if there are multiple, disconnected and similar colored objects and some pixels of the one object are marked as the foreground seeds, then it may not be possible to segment out all the objects from the background.

The major problem of the GrowCut method is to determine the correct label of the pixels. To resolve the above stated problems, we do the following in this experiment: i) from the seeded pixels we compute the color statistics; ii) using the color statistics some of the unlabeled pixels are temporarily labeled as foreground or background. Still, some pixels would remain unlabeled. Then automata evolution algorithm is executed to fix-up the label of the pixels. Before starting the evolution algorithm, using the above pre-processing steps, number of unlabeled pixels is reduced heavily. So, the evolution algorithm converges very fast. This pre-processing also helps to assign same label (temporarily) to the pixels having similar color profile and hence the distinct similar colored objects are labeled by same label, which helps to find all the similar color objects. The modified segmentation algorithm is as given below.

Modified GrowCut method	
1	Input foreground and background seeds
2	Compute the color statistics for the foreground and background
2.1	mean and std. dev of the foreground is m_f and σ_f
2.2	mean and std. dev of the background is m_b and σ_b
3	Compute temporary label of the unlabeled pixels p
3.1	if $(R, G, B)_p \in (m_f - 2.\sigma_f, m_f + 2.\sigma_f)$
3.1.a	consider p as the part of foreground
3.1.b	end if
3.2	if $(R, G, B)_p \in (m_b - 2.\sigma_b, m_b + 2.\sigma_b)$
3.2.a	consider p as the part of background
3.2.b	end if
4	Execute automata evolution rule

5. Experimental Results

In this work, we have modified the GrowCut method to locate multiple similar colored objects. The proposed method is faster than original GrowCut method. The performance of the present method is evaluated with some images, each of which contains multiple similar objects (see Fig. 1(a)). For GrowCut method, we have used the source code available at [32] and the proposed extension is also implemented on the same platform. In the proposed method, the **Algorithm: Automata evolution rule** is executed with some modifications. The modifications are: i) the line 4 of the **Algorithm: Automata evolution**

rule is modified as $g(\|\vec{C}_p - \vec{C}_q\|_2) \cdot \theta_q^t > \theta_q^{t+1}$; ii) the function 'g' (<1) (used to determine the attacking force) is modified as

$$g(\|\vec{C}_p - \vec{C}_q\|_2) = \frac{1}{diff * dist}, \text{ where } diff \text{ is the}$$

weighted sum of the RGB difference between the current pixel p and neighboring pixel q and $dist$ is the Euclidean distance between p and q ; iii) the neighboring system (N_g) is defined within 5×5 window.

The stability of the cell label is highly dependent on the neighbor of the cell. If neighbor size is small, then it takes more iteration to reach to stable configuration. At the same time, if the size of the neighbor is larger then the execution time of a particular iteration will be more. The total execution time of the evolution algorithm depends on both number of iteration and execution time per iteration. Experimentally we have decided that 5×5 neighboring system is the best option compared to 3×3 and 7×7 neighboring system. These modifications and the proposed pre-processing steps help to converge the evolution the cellular array very fast. The proposed method not only extracts multiple similar colored objects but also it is much faster than the GrowCut method. The comparison of the execution time and number of iterations between the GrowCut and proposed method is given in Table 1.

The comparative output of the GrowCut method and the proposed method is given in Fig 1. Here, green line is used to mark the foreground and background is marked by blue line. Fig. 1(b) shows the output of the GrowCut method and that of proposed method is shown in Fig. 1(c). The GrowCut method is not always good enough to extract all the desired objects. Most of the cases, the GrowCut method gives part of both foreground and background objects in the output. Most of the times, the output of the proposed method is satisfactory. However, the performance of the proposed method is not good enough for 'Image2' and 'Image9' where some unwanted objects are also included in the output.

Table 1: Comparative study in terms of execution time (in second) and number of iterations

Image	GrowCut		Proposed method	
	no. of iteration	execution time	no. of iteration	execution time
Image1	740	43.42	05	2.68
Image2	736	62.59	02	1.47
Image3	1749	221.80	12	14.42
Image4	984	70.34	11	7.35
Image5	989	61.94	09	4.90
Image6	858	32.52	03	1.02
Image7	897	50.40	06	3.03
Image8	1064	129.86	04	4.38
Image9	856	20.23	06	1.22

6. Conclusion

In this work, we have modified the GrowCut method to segment out all similar colored objects from the image. The performance of the proposed method is quite good and the method is faster than GrowCut. The proposed method is not giving exact result for all the images; there are some unwanted objects in the output. Our next target is to include the noisy images and also to improve the result.

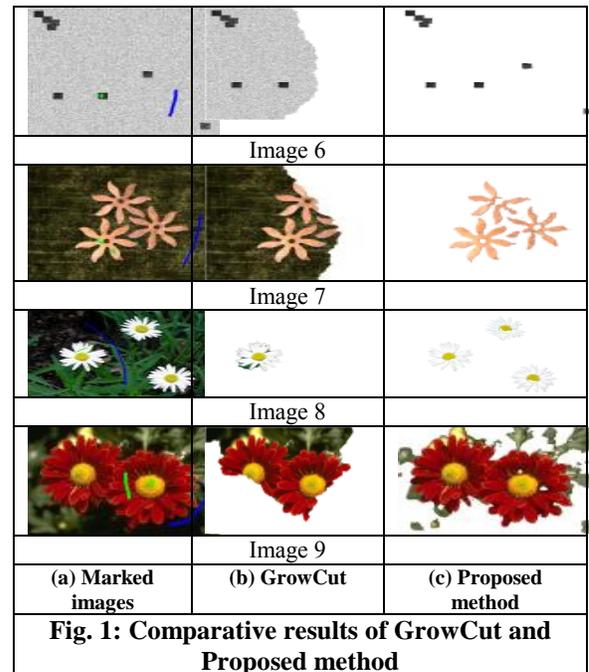
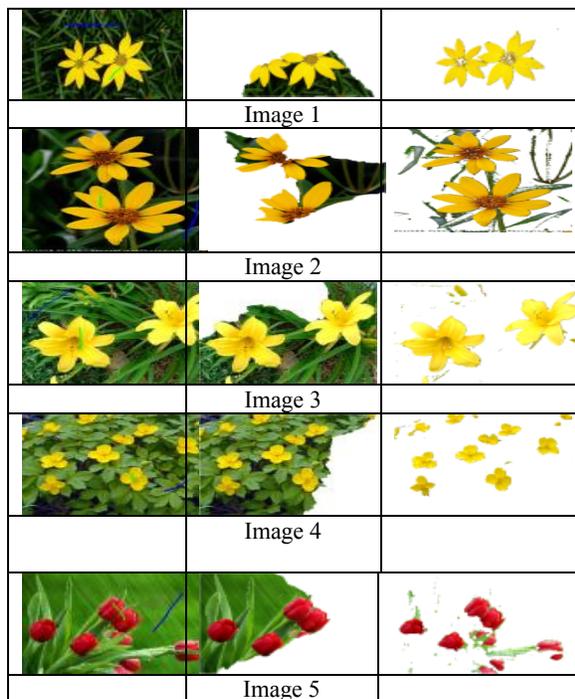


Fig. 1: Comparative results of GrowCut and Proposed method

References

- [1] Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1990) 629 – 639.
- [2] Roerdink, J.B.T.M., Meijster, A.: The watershed transform: Definitions, algorithms and parallelization strategies. *Fundamenta Informaticae* 41 (2001) 187 – 228.
- [3] Dong, L., Yu, G.: Fast search for thresholds from 1d and 2d histogram by an iterative algorithm for image segmentation. In: *Proceedings of the international conference on Systems, Man, and Cybernetics*. (2004) 3057 – 3062.
- [4] Busin, L., Vandehbroucke, N., Macaire, L., Postaire, J.G.: Color space selection for unsupervised color image segmentation by histogram multi-thresholding. In: *Proceedings of the international conference on image processing*. (2004) 203 – 206.
- [5] Hojjatoleslami, S.A., Kittler, J.: Region growing: A new approach. *IEEE Transactions on Image Processing* 7 (1998) 1079 – 1084.
- [6] Chen, S., Lin, W., Chen, C.: Split-and-merge image segmentation based on localized feature analysis and statistical tests. *CVGIP: Graph Models Image Processing* 53 (1991) 457 – 475.
- [7] Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. *International Journal of Computer Vision* 1 (1988) 321 – 331.
- [8] Blake, A., Rother, C., Brown, M., Perez, P., Torr, P.: Interactive image segmentation using an adaptive gmmrf model. In: *Proc. of the European*

- Conference on Communication Vision. (2004) 428 – 441.
- [9] Wong, K.C.H., Heng, P.A., Wong, T.T.: Accelerating ‘intelligent scissors’ using slimmed graphs. *Journal of Graphics Tools* 5 (2000) 1 – 13.
- [10] Levin, A., Rav-Acha, A., Lischinski, D.: Spectral matting. In: Proc. of the international conference on Computer Vision and Pattern Recognition. (2007) 1 – 8.
- [11] Felzenszwalb, P., Huttenlocher, D.: Efficient graph-based image segmentation. *International Journal of Computer Vision* 59 (2004) 167 – 181.
- [12] Mortensen, E.N., Barrett, W.A.: Intelligent scissors for image composition. In: Proc. of the conference on Computer graphics and interactive techniques (SIGGRAPH). (1995) 191 – 198.
- [13] Vezhnevets, V., Konouchine, V.: “Growcut” - interactive multi-label n-d image segmentation by cellular.
- [14] Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. In: Proc. of the international conference on Computer Vision. (2001) 105 – 112.
- [15] Grady, L.: Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (2006) 1768 – 1783.
- [16] Ning, J., Zhang, L., Zhang, D., Wu, C.: Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition* 43 (2010) 445 – 456.
- [17] Wang, J., Cohen, M.F.: An iterative optimization approach for unified image segmentation and matting. In: Proc. of the international conference on Computer Vision. (2005) 936 – 943.
- [18] Dayton, L., Davis, J.: *The Photoshop 6 Wow!* Peachpit press (2001).
- [19] Reese, L.: *Intelligent paint: Region-Based Interactive Image Segmentation*. Master Thesis, Department of Computer Science, Brigham Young University (1999).
- [20] Dhara, B.C., Chanda, B.: A fast interactive image segmentation to locate multiple similar-colored objects. In: Proc. of the national conference on Computer Vision, Pattern Recognition, Image Processing and Graphics. (2011) 25 – 28.
- [21] Cheng, Y.: Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (1995) 790 – 799.
- [22] Comaniciu, D., Meer, P.: Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (2002) 603 – 619.
- [23] Paragios, N., Deriche, R.: Geodesic active regions and level set methods for supervised texture segmentation. *International Journal of Computer Vision* 46 (2002) 223 – 247.
- [24] Stalling, D., Hege, H.C.: Intelligent scissors for medical image segmentation. In Arnolds, B., Muller, H., Saupe, D., Tolxdorff, T., eds.: *Tagungsband zum 4. Freiburger Workshop Digitale Bildverarbeitung in der Medizin*. (1996) 32 – 36.
- [25] Falcao, A.X., Udupa, J.K., Samarasekera, S., Sharma, S., Hirsch, B.E., Lotufo, R.A.: User-steered image segmentation paradigms: Live wire and live lane. *Graphical Models and Image Processing* 60 (1998) 233 – 260.
- [26] Falcao, A.X., Udupa, J.K.: A 3d generalization of user-steered live-wire segmentation. *Medical Image Analysis* 4 (2000) 389 – 402.
- [27] Li, Y., Sun, J., Tang, C., Shum, H.Y.: Lazy snapping. In: Proc. of the conference on Computer graphics and interactive techniques (SIGGRAPH). (2004) 303 – 308.
- [28] Rother, C., Kolmogorov, V., Blake, A.: Grabcut: Interactive foreground extraction using iterated graph cuts. In: Proc. of the conference on Computer graphics and interactive Techniques (SIGGRAPH). (2004) 309 – 314.
- [29] Grady, L.: Multilabel random walker image segmentation using prior models. In: Proc. of the international conference on Computer Vision and Pattern Recognition. (2005) 763 – 770.
- [30] Friedland, G., Jantz, K., Rojas, R.: Siox: Simple interactive object extraction in still images. In: *Proceedings of the Seventh IEEE International Symposium on Multimedia*. (2005) 253 – 260.
- [31] Friedland, G., Jantz, K., Lenz, T., Wiesel, F., Rojas, R.: A practical approach to boundary accurate multi-object extraction from still images and videos. In: *Eighth IEEE International Symposium on Multimedia*. (2006) 307 – 316.
- [32] GrowCut segmentation code in matlab. (<http://www.shawnlankton.com/2008/03/growcut-segmentation-in-matlab/>).

Sayan Bhowal: He received B.Tech degree in Computer Science and Engineering from West Bengal University of Technology, India in 2010. He earned M.E. in Software Engineering from Jadavpur University in 2012. His research area and interest include Image Processing, Video Processing and Pattern Recognition.

Bibhas Chandra Dhara: He received B.Sc. (Hons) degree in Mathematics and B.Tech degree in Computer Science and Engineering from University of Calcutta, India in 1997 and 2000, respectively. He earned M.Tech and PhD both in Computer Science from Indian Statistical Institute in 2002 and 2008, respectively. Currently, he is working as Assistant Professor in the Department of Information Technology, Jadavpur University, India. His research area and interest include Image Processing, Video Processing, Pattern Recognition and Audio Processing.