An Experiment with GrabCut Interactive Segmentation Technique on Specific Images

Anuja Deshpande, Dr. Pradeep Dahikar, Dr. Pankaj Agrawal

Abstract—This paper discusses an experiment conducted on GrabCut interactive segmentation technique using Matlab software on select images. The objective is to assess effectiveness of GrabCut interactive segmentation technique on specific *natural*images, which have *complex* image composition in terms of intensity, colour mix, indistinct object boundary, etc. In this experiment, effectiveness of image segmentation has been assessed by computing accuracy measures such as Jaccard Index, Dice Coefficient and Hausdorrf Distance between segmented and ground truth images. Although segmentation appeared to be accurate visually, however, when segmented images were subjected to aforesaid accuracy measures, GrabCut was not found to be effective on selected images, wherein image composition is much more complex or when foreground object boundary is not very distinct from background.

Index Terms—Magic Wand, GrabCut, Graph Cuts, Intelligent Scissors, suitability, interactive image segmentation, effectiveness,Jaccard Index, Dice Coefficient, Hausdorrf Distance.

1 Introduction

Extracting useful image data or segmenting image to solve specific problem has been challenging ever since the need rfdfor the same came into existence. The algorithms which were developed had specific image orientation or segmentation problem for which such algorithms were effective. Over time, while we now have hundreds of algorithms or techniques which have been developed since then, however, it seems image segmentation is still evolving and needs more mature or developed techniques to successfully segment very complex images.

Some of these techniques like Magic Wand [1] used texture (colour) and Intelligent Scissors [2] used edge (contrast) information to segment the images. Technique such as Graph Cuts [3] which is based on optimization, combined both quite successfully. While these works did solve some segmentation problems, those still cannot be generally applied on all images. The challenge is particularly glaring when it comes to segmenting natural images which have complex composition of the object and the background and /or similarity in foreground and background. Experimentation results in [4] have also indicated the same.

While assessing effectiveness or accuracy of the segmentation technique, one must consider the very purpose of segmentation. Segmentation results often need to be extensible to other applications for further processing [5].

This experiment deals with applying GrabCut [6] on select complex images; GrabCut is a great technique and an enhancement of original technique i.e. Graph Cuts. In GrabCut, a more powerful and iterative optimization is employed. The iterative approach allows to have less user interaction and robust border matting technique further enhances the boundary of the extracted foreground. In GrabCut, the two enhancements to Graph Cut are, iterative estimation and incomplete labelling which effectively reduces end user interaction for given quality result. User interaction is limited only to drawing a polygon around the object to be extracted in the background region. Specifically, in this technique, user does not mark foreground. As summarised in [7] below is the summary of GrabCut –

1. User creates an initial trimap by selecting a rectangle. Pixels inside the rectangle are marked as unknown. Pixels outside of rectangle are marked as known background.
2. Computer creates an initial image segmentation, where all unknown pixels are tentatively placed in the foreground class and all known background pixels are placed in the background class.
3. Gaussian Mixture Models (GMMs) are created for initial foreground and background classes using the Orchard-Bouman clustering algorithm.
4. Each pixel in the foreground class is assigned to the most likely Gaussian component in the foreground GMM. Similarly, each pixel in the background is assigned to the most likely background Gaussian component.
5. The GMMs are thrown away and new GMMs are learned from the pixel sets created in the previous set.
6. A graph is built and Graph Cut is run to find a new tentative foreground and background classification of pixels.
7. Steps 4-6 are repeated until the classification converges.

2 Accuracy Measurement Techniques

Similar to and as expressed in [4], in this experiment also we have assessed the accuracy of the segmentation by computing Jaccard Index, Dice Coefficient & Hausdorrf Distance on segmented images by comparing with ground truth.

2.1 Jaccard Index

The Jaccard Index [8], also known as the Jaccard similarity coefficient by Paul Jaccard, is a statistic measure used for comparing the similarity and diversity of sample sets. The Jaccard coefficient measures similarity between finite sample sets A and B, and is defined as the size of the intersection divided by the size of the union of the sample sets:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (1)$$

The Jaccard distance, which measures dissimilarity between sample sets, is complementary to the Jaccard coefficient and is obtained by subtracting the Jaccard coefficient from 1, or, equivalently, by dividing the difference of the sizes of the union and the intersection of two sets by the size of the union:

$$dJ(A,B) = 1 - J(A,B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|} \quad (2)$$

2.2 Dice Coefficient

The Sørensen–Dice index [9], also known by other names, is a statistic used for comparing the similarity of two samples. It was independently developed by the botanists Thorvald Sørensen and Lee Raymond Dice. Sorensen's original formula was intended to be applied to presence/absence data, and is –

$$QS = \frac{2|A \cap B|}{|A| + |B|} \quad (3)$$

Where, |A| and |B| are the numbers of species in the two samples. QS is the quotient of similarity and ranges between 0 and 1. It can be viewed as a similarity measure over sets

2.3Hausdorrf Distance

The Hausdorrf distance [10], named after Felix Hausdorrf is also known as Hausdorrf metric, measures how far two subsets of a metric space are from each other. Hausdorrf distance is the greatest of all the distances from a point in one set to the closest point in the other set. Let X and Y be two non-empty subsets of a metric space (M, d). We define their Hausdorrf distance d H(X, Y) as –

$$d_H(X,Y) = inf\ \{\epsilon \geq 0;\ X \subseteq Y\epsilon \text{ and } Y \subseteq X\epsilon\} \quad (4)$$

Where

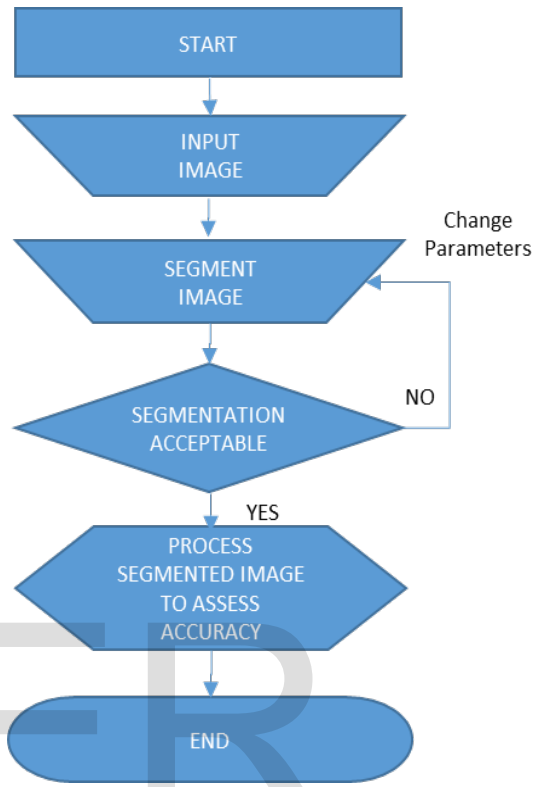
$$X\epsilon := \cup_{x \in X}\{z \in M;\ d(z,x) \leq \epsilon\} \quad (5)$$

3 Experiment

In this experiment, I have studied GrabCut technique as described in [6] and its implementation using [11] on select images using MATLAB, to understand and study effectiveness of GrabCut and accuracy of segmentation by assessing –

1. Visual confirmation
2. Jaccard Index
3. Dice Index
4. Hausdorrf Distance

The original segmentation code in [11] was modified with approval from the author, for three changes so as to facilitate this experiment and the changes were (i) new code was inserted to enable saving segmented image to the disk, (ii) increase iteration limit from 40 to 200 to ensure convergence is successful (as our images were complex) and (iii) increase the max K-cluster limit from 12 to 96 for granular control on segmentation process for better output. Varying values of K-clusters and beta were used during multiple runs of the segmentation process and combination which resulted in best output was chosen for final segmented image.Following steps were performed in this experiment.



(6)

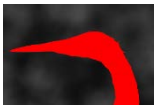
For this experiment, select images from Single Object Image Segmentation Dataset of natural images [12] has been used. This dataset is made freely available for research purposes, by Department of Computer Science and Applied Mathematics, Weizmann Institute of Science. This image dataset provides source image as well as ground truth for comparison. As stated in [12], Ground Truth has been constructed using manual segmentation by human subjects. We have used 5 colour images as an input to the segmentation technique, whose output is also a colour (RGB) image, with extracted foreground and black background. The segmented images were processed using GIMP image editor [13] and using Matlab were converted into binary images for further computations.

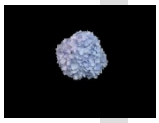
Ground Truth images were also converted to binary images so that a comparison can be done with segmented images. The source images, segmented images and findings are listed below. Following images have been resized using Microsoft Paint to fit this document.

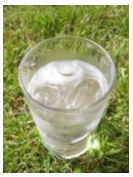
3.1Experiment Results

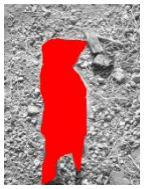
Let us review the segmentation results which seem to be largely successful visually, however, when the images were subjected to the accuracy measures, accuracy was found to be low as is evident from the following results. These segmentations were the best possible as were observed during multiple segmentation runs for varied "k" clusters and Beta values.

Segmentation Set 1		
Original Image	Segmented Image	Ground Truth
		
Jaccard Index = 0.9159		
Dice Coefficient = 0.9561		
Hausdorrf Distance = 41.5692		

Segmentation Set 2		
Original Image	Segmented Image	Ground Truth
		
Jaccard Index = 0.8976		
Dice Coefficient = 0.9460		
Hausdorrf Distance = 42.1189		

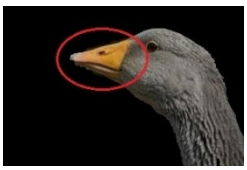
Segmentation Set 3		
Original Image	Segmented Image	Ground Truth
		
Jaccard Index = 0.9039		
Dice Coefficient = 0.9495		
Hausdorrf Distance = 32.7719		

Segmentation Set 4		
Original Image	Segmented Image	Ground Truth
		
Jaccard Index = 0.9452		
Dice Coefficient = 0.9718		
Hausdorrf Distance = 23.3880		

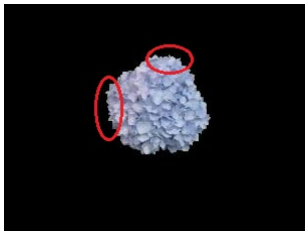
Segmentation Set 5		
Original Image	Segmented Image	Ground Truth
		
Jaccard Index = 0.8271		
Dice Coefficient = 0.9054		
Hausdorrf Distance = 46.7440		

4 Observations

Overall in this experiment, we found Grab cut technique to be quite fast in execution and could successfully identify/segment most of the foreground with excellent accuracy, barring few areas which have complexities.

Segmentation Set 1	Segmentation Set 2
	

In the above observation related to segmentation set 1 and 2, as can be visually confirmed in the original image, in some areas the foreground object boundary is not quite distinct and has similar texture, colour, etc. as that of background (highlighted). Such areas or a portion thereof, seems to have gotten marked as background region during the iterative segmentation process and eventually has led to loss of information in the foreground or segmented image.

Segmentation Set 3


As in the above observation related to segmentation set 3, in the original image, object boundary in few areas (highlighted in above image) seems to have layers partly overlapping each other and the lower layers are not quite crisp, thereby leading to a blur boundary. During segmentation, for such areas or edges this seems to have led to loss of information during the iterative runs.

In the below observation related to segmentation set 4, the high intensity spot and the background which is visible in the transparent foreground seems to be interfering with the segmentation process. Due to the complexity of the image, the finesse of the boundary is also missing at quite a few locations.

Segmentation Set 4

Let us observe segmentation set 5 as shown below, which by far seems to be the most complex of the sets under this experiment. In the original image, the background and foreground have similar colours and also, the object boundary, due to existence of fur, is not quite clear and crisp. During the segmentation process body the fur is lost and also in lot of areas even the image boundary has been missed.

Segmentation Set 5

In this experiment, during the segmentation runs, various combinations of "k" cluster limit and Beta values were tried for best results. It was however observed that increasing these limits beyond a threshold did not necessarily yield successful segmentations. Although in certain cases, it did facilitate the convergence, still the segmentation output was far from desired results. Also, beyond a certain K-cluster limit which was different for every image, the output either did not have any improvements or in fact deteriorated in some cases.

5 Conclusion

Grab Cut technique, which is based on iterative Graph Cut, is quite faster and was largely successful from general purpose utility perspective. The Grab Cut technique could come in handy especially in such images which incorporate lots of edges and also are too close to each other (like in Segmentation Set 4), where other techniques like Intelligent Scissors is likely to face challenge due to limitations on account of boundary tracing.

However, when the segmented images were compared with the ground truth, the accuracy of segmentation was found to be bit low. For specific applications, which require great deal of accuracy, this technique alone may not be suitable.

The said technique might be suitable for moderately complex images only. Images which employ great deal of complexities, the accuracy of this technique might be limiting.

6 Acknowledgment

We are deeply grateful to Itay Blumenthal, the author of the Matlab code for Grab Cut technique for granting the permission to modify the original source code to facilitate this experiment. It was required that the original Matlab code implementing Grab Cut be modified to ensure that convergence of the graph cut algorithm is successful. The modified Matlab code was used to conduct this experiment, upon approval by the Author.

7 About Authors

1. Anuja Deshpande is currently pursuing Doctoral Research at Department of Electronics, Kamla Nehru Mahavidyalaya, Sakkardara, Nagpur 440024, Maharashtra,India; affiliated to RSTM Nagpur University. E-mail: anuja_1978@yahoo.com

2. Dr. Pradeep Dahikaris currently Head of the Department, Department of Electronics, Kamla Nehru Mahavidyala, Sakkardara, Nagpur 440024, Maharashtra, India. E-mail: pbdahikarns@rediffmail.com

3. Dr. Pankaj Agrawal is currently Professor and Dean (Academics), Department of Electronics and Communication Engineering, G. H. Raisoni Academy of Engineering and Technology, Nagpur 440016, Maharashtra, India. E-mail: Pankaj_rknec@rediffmail.com

References

[1] Magic Wand - ADOBE SYSTEMS INCORP. Adobe Photoshop User Guide.
[2] Mortensen, E, and Barrett W. 1995. "Intelligent scissors for image composition". *Proc. ACM Siggraph*, 191–198.
[3] Y. Boykov and M.P. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images". In Proceedings of International Conference on Computer Vision, Vol. 1, pp. 105-112, July 2001.
[4] Anuja Deshpande(1), Pradeep Dahikar(1) and Pankaj Agrawal(2), "An experiment with Intelligent Scissors interactive segmentation technique on specific images", International Journal of Scientific and Research Publications, Volume 6, Issue 7, July 2016, 60-63.
  [1] Department of Electronics, Kamla Nehru Mahavidyalaya, Sakkardara, Nagpur, Maharashtra, India 440024
  [2] Dept. of Electronics and Communication Engg, G. H. Raisoni Academy of Engineering and Technology, Nagpur. Maharashtra, India 440016
[5] Feng Ge (1), Song Wang (2) and Tiecheng Liu (2): "New benchmark for image segmentation evaluation". Journal of Electronic Imaging 16(3), 033011 July – September 2007.
  [1] Virginia Tech Department of Electrical and Computer Engineering, Blacksburg, Virginia 24061
  [2] University of South Carolina, Department of Computer Science and Engineering, Columbia, South Carolina 29208
[6] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut: inter-active foreground extraction using iterated graph cuts", Proc. ACM SIGGRAPH, 2004.
[7] Justin F. Talbot(1) and Xiaoqian Xu(1), "Implementing GrabCut".
  [1] Brigham Young University
[8] https://en.wikipedia.org/wiki/Jaccard_index
[9] https://en.wikipedia.org/wiki/S%C3%B8rensen%E2%80%93Dice_coefficient
[10] https://en.wikipedia.org/wiki/Hausdorff_distance
[11] Itay Blumenthal authored GrabCut implementation code available at , http://grabcut.weebly.com/code.html
  [1] Yuri Boykov, Olga Veksler, RaminZabih: "Efficient Approximate Energy Minimization via Graph Cuts", IEEE transactions on PAMI, vol. 20, no. 12, p. 1222-1239, November 2001.
  [2] Vladimir Kolmogorov and RaminZabih: "What Energy Functions can be Minimized via Graph Cuts?" To appear in IEEE Transactions on

Pattern Analysis and Machine Intelligence (PAMI). Earlier version appeared in European Conference on Computer Vision (ECCV), May 2002.

[3] Yuri Boykov and Vladimir Kolmogorov: "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision", In IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), September 2004

[12] Single Object Segmentation Evaluation Database(1) maintained by Sharon Alpert, Department of Computer Science and Applied Mathematics, Weizman Institute of Science, http://www.wisdom.weizmann.ac.il/~vision/Seg_Evaluation_DB/index.html

[1] @inproceedings{AlpertGBB07,
author = {"Sharon Alpert and MeiravGalun and Ronen Basri and Achi Brandt"},
title = {"Image Segmentation by Probabilistic Bottom-Up Aggregation andCue Integration."},
booktitle = {"Proceedings of the IEEE Conference on Computer Vision andPattern Recognition"},
month = {"June"},
year = {"2007"}
}

[13] Gimp Image Editor Software, https://www.gimp.org