Interactive Image Segmentation using Color and Texture Features

Sajan Kor
Computer Science and Engineering. Krishna Engineering College, Uttar Pradesh Technical University, Ghaziabad, Uttar Pradesh, India

Pramod Kumar Sethy
Computer Science and Engineering. Krishna Engineering College, Uttar Pradesh Technical University, Ghaziabad, Uttar Pradesh, India

ABSTRACT

Image segmentation is the process of selecting objects of interest in background images. Since the fully automatic segmentation is difficult for natural images due to the complex color and texture features. This paper presents the idea of image segmentation using maximum similarity based region merging with minimum user inputs using simple brush strokes called markers(object and background markers).Therefore, the proposed system presents a new region merging method that grow regions from foreground/background seeds based on the color and texture features for interactive image segmentation. An initial segmentation is required to partition the image into homogeneous regions for merging. After the completion of the initial segmentation, initial segmented regions are represented by means of some descriptors such as color and texture to guide the region merging process. Therefore, the proposed method is very effective and it can quickly and accurately segment a wide variety of natural images with ease.

Keywords

Image Segmentation, maximum similarity, Region Merging, Mean Shift

1. INTRODUCTION

Image segmentation is a process of identifying and at the same time separating appropriate objects and structures in an image. This is a basic problem in image analysis – accurate and exact segmentation of particular object is often required before any further processing and analysis can be performed. Regardless of years of active research, fully automatic segmentation of arbitrary images remains an unsolved problem.

This paper deals with the problem of interactive image segmentation – the process of assisting a human user in cutting out a desired object from the image. The aim of interactive segmentation methods is to minimize the user interaction time, while maintaining adequate control to ensure appropriate results.

INTERACTIVE image segmentation involves very minimal user input to include user intention into the segmentation process and is an active research area in recent years because it can achieve satisfactory segmentation results are unachievable by the state-of-the-art automatic image segmentation algorithms.

In this paper, a novel interactive region merging method based on the initial segmentation of mean shift[2] is proposed. In the proposed scheme, markers are used as interactive information provided by the user to roughly indicate the features and position of the object as well as background. The markers can be the simple color strokes

(see Fig. 1b). Then the proposed method will calculate the similarity of different regions and merge them based on the proposed rule for maximum similarity with the help of these markers. And at the end of the merging process, the object will be extracted from the background (Fig. 1c shows an example of segmentation result).
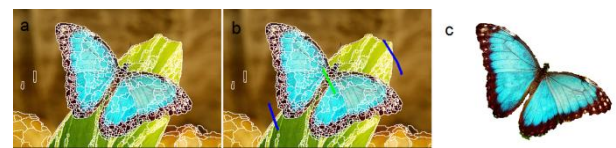
Fig. 1. (a) Initial mean shift segmentation; (b) Marker input by the user. Green line is object marker and blue ones are background marker; (c) proposed method Segmentation result

The key contribution of the proposed method is a novel maximal similarity based region merging (MSRM) mechanism. In this preset threshold is not required and is adaptive to image content. With the proposed algorithm, the unmarked background regions will automatically get merged and labeled, while the unmarked object regions will be identified and circumvented from getting merged with background. Once all the unmarked regions are labeled, then the object contour can readily be extracted from the image. The proposed algorithm is simple and at the same time successfully extracts required objects from even a complex scenes. The paper is further organized as follows. Section 2 presents the related works done till the time. Section 3 describes proposed region merging method. Section 4 performs extensive experiments to verify the proposed method.

2. RELATED WORKS

Boykov and Jolly[4] were the first to propose Interactive graph cuts. The system forms a graph where each node represents a pixel and the edges connect adjacent pixels. The algorithm cut the graph in the best place to separate out the background from the foreground, while using the seeds as restraints in an order to minimize a cost function.

Early interactive image segmentation algorithms use either regional properties such as Adobe's magic wand or boundary properties e.g. active contour and Intelligent Scissors. Graph cut based methods [3], [5] consider both regional and boundary properties. With the graph cut based methods, an image is displayed as a graph wherein each node of this graph represents a pixel and the adjacent nodes are connected with a weighted edge defined as the distance between the pixel values. Additionally, this algorithm [3] models foreground and background pixel values according to histograms. Mainly, two virtual nodes are added to the graph to define the background and foreground models. The max-flow/min-cut

algorithm [4] is used to classify the nodes. To initialize the models, both foreground and background seeds are required. The LazySnapping work [10] integrates intuitive user interfaces, including foreground/ background strokes and boundary polygon editing, with the graph cut algorithm for easy interactive image segmentation. The Grab-Cut method [6] expands the graph cut framework to segment color images in the developed Gaussian mixture models. The Grab-Cut method supports various kinds of user inputs. Like, a lasso input for complex images, foreground and background markers for local editing, a boundary brush for matting etc. The GrabCut method can attain good performance in segmenting the images whose foreground and background colors are well divisible, but its performance is often unacceptable for the images whose foreground and background share similar color distributions such as in cluttered or invisible images. One intrinsic limit for graph cut based methods [3], [5] lies in its fundamental assumption that an object's shape is best described by the shape with smallest boundary length. This does not hold for very complicated shapes such as bush branches or hair. Due to assumptions, these methods doesn't do well with noise or interlace effects in videos.

The Random Walks algorithm and its extensions [11] have been adopted for various image processing tasks. These also model an image as a graph. Even though For image segmentation, Random Walks algorithm perform better than the Graph Cut algorithm. However, the Random Walks algorithm is very sensitive towards the positions and quantities of foreground and background seeds as it lacks in global color distribution model.

3. PROPOSED METHOD

In this paper, the novel interactive region merging method based on mean shift[9] initial segmentation is proposed. In the proposed scheme, user input is provided through markers to roughly indicate the position and main features of the object and background. The markers with different colors are simple brush strokes. The proposed method computes the similarity of different regions and based on the proposed maximal similarity rule, merge the regions that are maximum similar. Once the merging process ends, the object gets extracted from the background. This process is adaptive to the image content and does not require similarity threshold to be set in advance.

A general schematic flowchart of proposed algorithm is given in Fig.4. It consists of three main steps: 1) initial over-segmentation; 2) user inputs; 3) region merging. Step 3 is our major focus here.
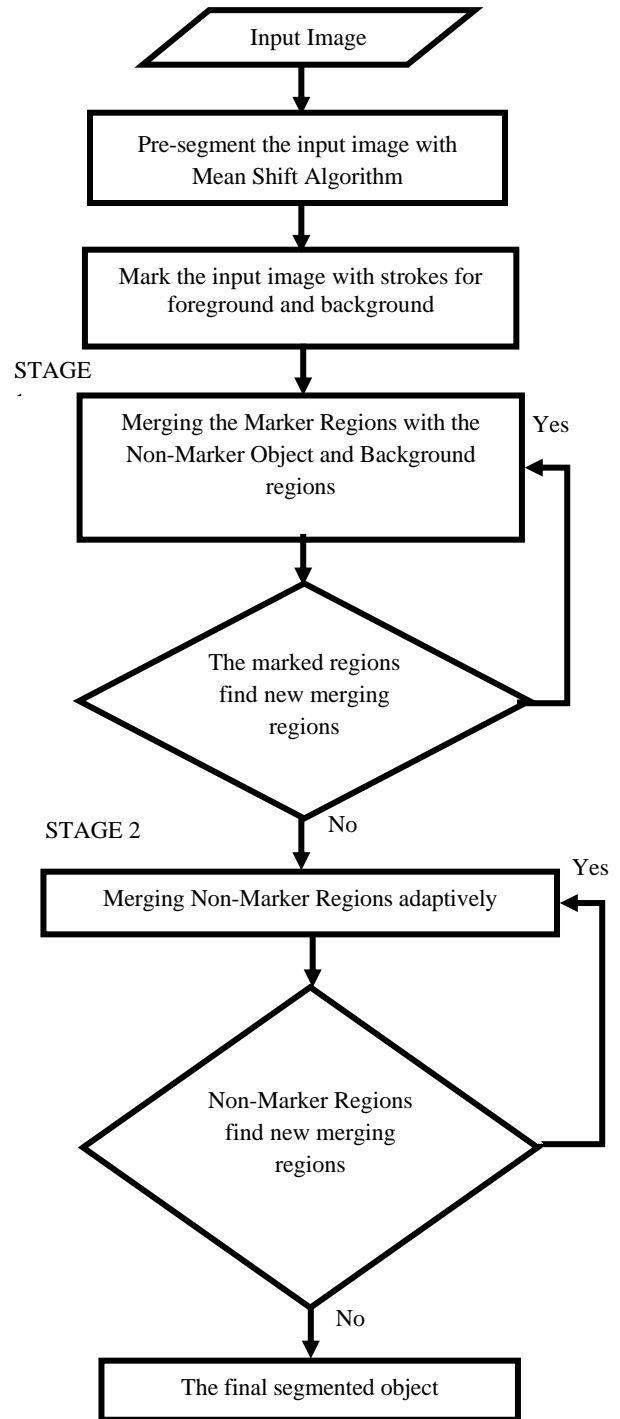

Fig 4: Flowchart of proposed algorithm

3.1 Initial Mean Shift Segmentation

In this paper, for the region merging process, mean shift algorithm is used for the initial segmentation and it preserves the boundary well. A software, "the EDISON System" [8] is used for performing the mean shift initial segmentation. Fig. 2 shows an example of the mean shift initial segmentation using "the EDISON system".

Fig. 2. (a) Original Input Image (b) The result of the mean shift algorithm

3.2 Feature descriptor and a similarity measure

After the mean shift initial segmentation, it is required to represent these regions using some descriptor and define a rule for merging the regions. A region can be described in many aspects, such as shape, size, color, texture and edge of the region. The color histogram is one of the effective descriptor to represent the color feature statistics of object and is evenly used in pattern recognition, object tracking, etc. In context of segmentation based on region merging, color histogram is more robust than the other feature descriptors as different regions from the same object often have low variation in colour whereas they have high variation in other aspects like size and shape.

Therefore, in this work, the color histogram is used to represent each initially segmented region and the RGB color space is used to compute the color histogram.

A normalized histogram for a region X is denoted by HistX. For computing the colour histogram, the RGB colour space is used, wherein each colour is divided into 16 levels with a total feature space of $16\times16\times16 = 4096$ bins.

In the interactive image segmentation, the users will mark part of regions of an image as object and background. The crucial part in region merging is to determine the similarity between the unmarked regions with the marked ones based on some logic control. Therefore, it is necessary to define a similarity measure $\rho(X,Y)$ between two regions X and Y to accommodate the comparison between various regions. Bhattacharyya Coefficient[1] is used to measure the similarity between X and Y.

$$\rho(X,Y) = \sum_{u=1}^{4096} \sqrt{Hist_X^u \cdot Hist_Y^u} \qquad (1)$$

where Hist$_X$ and Hist$_Y$ are the normalized histograms of X and Y, respectively, and the superscript u represents the uth element of them.

The Bhattacharyya coefficient can also be represented as a cosine of angle between two vectors,

$$\left(\sqrt{Hist_X^1}, \dots, \sqrt{Hist_X^{4096}}\right)^T \text{ and } \left(\sqrt{Hist_Y^1}, \dots, \sqrt{Hist_Y^{4096}}\right)^T$$

As per the Bhattacharyya coefficient, higher the similarity, smaller the angle will be and thus higher the Bhattacharyya coefficient.

3.3 User Interaction

In the interactive image segmentation, the users need to specify the inputs on object and background in initially segmented image. Users can input interactive information by drawing markers, which could be simple strokes on the image. The regions having pixels within the object markers are called object marker regions, whereas, the regions having pixels inside the background markers are called background marker regions. The green marker is used to mark the object while blue markers are used to represent the background. A small portion of the object regions and background regions will be marked by the user. Actually, lesser the required input by the user, more robust and convenient will be the interactive algorithm. After object marking, each region will be labeled either as the marker object region, the marker background region or the non-marker region. The proposed method automatically assigns each non-marker region with a correct label of either object region or background region for completely extracting the object contour.

3.4 Merging Rule

Some key features of object and background are provided by object and background markers, respectively. The proposed method starts from the initially marked regions and all the non-marked regions will be gradually labeled as either object region or background region. In this work, an adaptive maximal similarity based merging mechanism is presented to identify all the non-marker regions under the guidance of object and background markers.

Let, X & Y are two adjacent regions of the initially segmented image. If all the adjacent regions of X are denoted by $\bar{S}X = \{S_i^X\}_{i=1,2,\dots,x}$ than Y is a subset of $\bar{S}X$. Now, after calculating the similarity between X and all its adjacent regions, i.e. $\rho(X, \bar{S}X)$, i= 1,2, ... ,x, if maximum similarity is found between X and Y, Y will merge with X. Therefore, the merging rule can be defined as below:

Merge X and Y if,

$$\rho(X, Y) = \max_{i=1,2\dots x} \rho(X, S_i^X) \qquad (2)$$

Since the markers cover only a small portion of the image, according to our merging process, unmarked object regions should be identified but not merged with the background. As all object regions are part of the same object, the unmarked object regions will usually have higher similarity with the marked object regions than the background regions. Therefore, the unmarked object regions are more likely to be identified as object.

3.5 Region Merging Process

The whole region merging process in our work is divided into two stages, which are repeatedly executed until no new merging befalls. Merging as many background regions as possible while keeping out object regions from being merged is the main concept of this region merging process. After merging all the background regions, it is correspond to extracting the desired object.

In the first stage, the marked background regions are merged with their adjacent regions. . For each background region b i.e. $b \in R_B$, the set of its neighbouring regions $\bar{S}b = \{a_i\}_{i=1,2,\dots,r}$ is formed. Then for each ai and $a_i \notin R_B$, a set of its neighbouring regions $\bar{S}_{a_i} = \{S_j^{a_i}\}_{j=1,2,\dots,k}$ is formed.

From the above statement it can be said that $b \in \bar{S}_{a_i}$. The similarity between ai and each element in $\bar{S}_{a_i}$, i.e. $\rho(a_i, S_j^{a_i})$, is calculated. If b and ai satisfy the rule (2), i.e.

$$\rho(a_i, b) = \max_{j=1,2,\dots,k} \rho(a_i, S_j^{a_i}) \qquad (3)$$

then b and a_i are merged into one region and the new region will have the same label as region b:

$$b = b \cup a_i \qquad (4)$$

Otherwise, b and ai will not be merged.

The above procedure is iteratively implemented. The sets R_B and R_N will be updated. Specifically, R_B grows and R_N shrinks. The iteration stops when the entire marker background regions R_B will not find new merging regions.

After this stage of region merging, some unmarked background regions will be merged with the corresponding marked background regions. However, there are still unmarked background regions which cannot be merged because they have higher similarity scores with each other than with the marked background regions.

In the second stage, the focus is on the non-marker regions R_N remained from the first stage. Part of R_N belongs to the background, while part of R_N belongs to the target object. In this stage, the non-marker object regions will be fused each other under the guidance of the maximal similarity rule and so do the non-marker background regions.

For each unmarked (background or object) region $r \in R_N$, a set of its adjacent regions $\bar{S}r = \{Hi\}i=1,2,\dots,r$ is formed. Then for each Hi $H_i \notin \mathbf{R_B}$ and $H_i \notin \mathbf{R_O}$, its set of adjacent regions $\bar{S}_{H_i} = \{S_j^{H_i}\}j=1,2,\dots,k$ is formed. There is is $r \notin \bar{S}_{H_i}$. The similarity between Hi and each element in $\bar{S}_{H_i}$, i.e. $\rho(Hi, S_j^{H_i})$, is calculated. If Z and Hi satisfy the rule (2), i.e.

$$\rho(r, H_i) = \max_{j=1,2,\dots,k} \rho(H_i, S_j^{H_i}) \qquad (5)$$

then r and H_i will be merged into one region

$$r = r \cup H_i \qquad (6)$$

Otherwise, r and Hi will not be combined.

This again is an iterative process and stops only when the entire set of unmarked regions R_N will not find any new region for merging. Fig. 3b shows the merging result after the second stage. It can be seen that some unmarked background regions, as well as some unmarked object regions, are merged, respectively, in this stage.
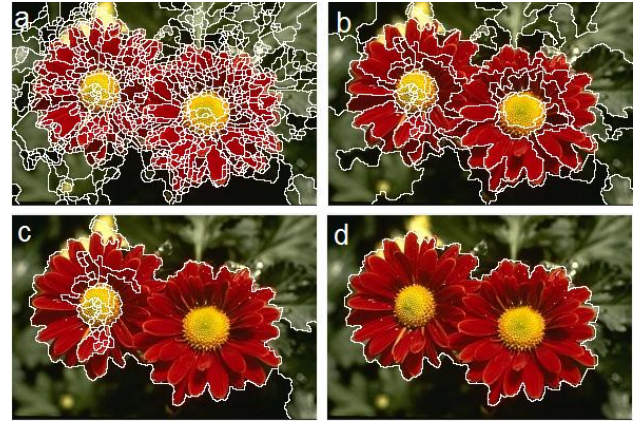

Fig. 3. Region merging process: (a) the first stage (1ˢᵗ round); (b) the second stage (1ˢᵗ round); (c) the first stage(2ⁿᵈ round); and (d) the merging results.

These both stages of the algorithm are executed repeatedly until no new merging happens. Fig. 3c shows the merging output of the first stage in the 2nd round. Since there is no more merging action, the algorithm stops here. At the end, each region is labelled as one of the two classes: object or background. Now the object contour can easily be extracted by extracting only the object regions, as shown in Fig. 3d.

4. RESULTS AND CONCLUSION

The objective of the proposed system is to achieve robust and accurate segmentation in different natural images. Table 1 shows the size, number of regions after the initial segmentation by mean shift algorithm and the running time of different images used in this paper work. The initial segmentation is very important for segmenting the object from its background in the interactive scheme.

Table 1. Initial Segmentation result and running time of Different images

Image	Size of Image	Number of regions after initial segmentation	Running Time (in seconds)
Butterfly	273x410	816	25
Two Flowers	289x193	985	37
Dog	254x372	1042	34
Bear	288x433	646	16
Starfish	448x368	1088	183
Monkey	360x414	843	58

Fig. 5 (a) Original Input Image (b) The source images with user input(c) Result obtained by the proposed method

Fig. 6(a) Original Input Image (b) The source images with user input (c) Result obtained by the proposed method

Fig 5 shows the some examples for extracting the target object from the background with the user input markers. Since the proposed method is mainly based on color and texture features for region merging to extract the desired object, it is very convenient for these examples to get the accurate result. However, for some natural scene images such as complex background, it may require more user input markers to achieve the satisfying result. If the foreground color of the object overlaps with its background color in some images, the color feature is not sufficient information to extract the object. Therefore, the integration of color and texture features is considered in the proposed method for region merging. Fig 6 shows the result of segmenting the foreground from the complex background with less user inputs because of the advantages of the proposed method.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] Kailath, Thomas. "The divergence and Bhattacharyya distance measures in signal selection." Communication Technology, IEEE Transactions on 15.1 (1967): 52-60.

[2] Ning, Jifeng, et al. "Interactive image segmentation by maximal similarity based region merging." Pattern Recognition 43.2 (2010): 445-456.

[3] Felzenszwalb, Pedro F., and Daniel P. Huttenlocher. "Efficient graph-based image segmentation." *International Journal of Computer Vision* 59.2 (2004): 167-181.

[4] Boykov, Yuri, and Vladimir Kolmogorov. "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision." *Pattern Analysis and Machine Intelligence, IEEE Transactions* on 26.9 (2004): 1124-1137.

[5] Çiğla, Cevahir. "Efficient graph-based image segmentation via speeded-up turbo pixels." *Image Processing (ICIP), 2010* 17th *IEEE International Conference* on. IEEE, 2010.

[6] Vezhnevets, Vladimir, and Vadim Konouchine. "GrowCut: Interactive multi-label ND image segmentation by cellular automata." *proc. of Graphicon*. 2005.

[7] Yuan, Ye, and Chuanjiang He. "Variational level set methods for image segmentation based on both L2 and Sobolev gradients." *Nonlinear Analysis: Real World Applications 13.2* (2012): 959-966.

[8] EDISON software. http://www.caip.rutgers.edu/riul/research/code.html.

[9] Bhargava, Neeraj, et al. "Iterative Region Merging and Object Retrieval Method Using Mean Shift Segmentation and Flood Fill Algorithm." Advances in Computing and Communications (ICACC), 2013 Third International Conference on. IEEE, 2013.

[10] Y. Li, J. Sun, C.-K. Tang, and H.-Y. Shum, "Lazy snapping," in *ACM Siggraph*, 2004, pp. 303–308.

[11] Jha, Sonu Kumar, Purnendu Bannerjee, and Subhadeep Banik. "Random Walks based Image Segmentation Using Color Space Graphs." *Procedia Technology* 10 (2013): 271-278.