ISSN 2319 - 7595

Volume 1, No.1, September - October 2012
International Journal of Information Systems and Computer Sciences
Available Online at http://warse.org/pdfs/ijiscs01112012.pdf

Similarity Based Region Merging Interactive Image Segmentation

[1]Banudevareddy. B, [2]Prasanna Kumari. P
[1]PG Student, LIET, Hyderabad, India, banudevareddy.bandi@gmail.com
[2] Associate Professor, LIET, Hyderabad, India, prasannakumaripotteti@gmail.com

ABSTRACT

Efficient and effective image segmentation is an important task in computer vision and object recognition. Since fully automatic image segmentation is usually very hard for natural images, interactive schemes with a few simple user inputs produce good solutions. This work presents a new region merging based interactive image segmentation method. The users only need to roughly indicate the location and region of the object and background by using strokes, which are called markers. A novel maximal-similarity based region merging mechanism is proposed to guide the merging process with the help of markers. A region R is merged with its adjacent region Q if Q has the highest similarity with Q among all Q's adjacent regions. The proposed method automatically merges the regions that are initially segmented by mean shift segmentation, and then effectively extracts the object contour by labeling all the non-marker regions as either background or object. The region merging process is adaptive to the image content and it does not need to set the similarity threshold in advance. Extensive experiments are performed and the results show that the proposed scheme can reliably extract the object contour from the complex background.

Keywords: Images, Segmentation, Region, Mean shift, Monochrome

I. INTRODUCTION

Most attention to image segmentation has been focused on grey-scale (or monochrome) images. A common problem in the segmentation of grey-scale images occurs when an image has a background of varying grey level, such as, gradually changing shades, or when regions assume some broad range of grey levels. This problem is inherent since intensity is the only available information from monochrome images. It is known that the human eye can detect only in the neighborhood of one or two dozen intensity levels at any point in a complex image due to brightness adaptation, but can differentiate thousands of color shades and intensities. There are currently a large number of color image segmentation techniques available. They can be categorized into four general groups: pixel-based, edge based, region-based, and model-based techniques. These techniques are either based on concepts of similarity (edge-based) or on discontinuity (pixel-based and region-based) of pixel values. Model-based techniques, where segmentation is posed as a statistical optimization problem, have become popular in the past decade.

The low level image segmentation methods, such as mean shift, watershed, level set and super-pixel, usually divide the image into many small regions **[2, 3 and 4]**. Although may have severe over segmentation, these low level segmentation methods provide a good basis for the subsequent high level operations, such as region merging. As a popular segmentation scheme for color image, mean shift can have less over segmentation than watershed while preserving well the edge information of the object. In general, the color and texture features in a natural image are very complex so that the fully automatic segmentation of the object from the background is very hard. Therefore, semi-automatic segmentation methods incorporating user interactions may perform better. Semi-automatic segmentation algorithms are becoming more and more popular.

The Proposed segmentation is a novel interactive region merging method based on the initial segmentation of mean shift **[6]**. In the proposed scheme, the interactive information is introduced as markers, which are in put by the users to roughly indicate the position and main features of the object and background. The markers can be the simple strokes (e.g. the green and blue lines). The proposed method will calculate the similarity of different regions and merge them based on the proposed maximal similarity rule with the help of these markers. The object will then be extracted from the background when the merging process ends.

2.LITERATURE SURVEY

The main objective of image segmentation is to divide an image into regions that can be considered homogeneous with respect to a given criterion such as color or texture. For this reason, a considerable care is taken to improve the probability of a successful segmentation. Image segmentation has taken a central place in numerous applications, including, but not limited to, multimedia databases, color image and video transmission over the Internet, digital broadcasting, interactive TV, video-on-demand, computer-based training, distance education, video-conferencing, tele-medicine, and, with the development of the hardware and communications infrastructure, to support visual applications. The field has become a principal area of research, not only in electrical engineering, but also in other academic disciplines, such as computer science, geography, medical imaging, criminal justice, and remote sensing.

In the field of computer vision, graph-cuts segmentation can be employed to solve a wide variety of low-level computer vision problems (early vision), such as image smoothing. Although many computer vision algorithms involve cutting a graph (e.g., normalized cuts), the term "graph cuts" is applied specifically to those models which employ a max-flow/min-cut optimization **[5, 7]**. "Binary" problems (such as denoising a binary image) can be solved exactly using this approach; problems where pixels can be labeled with more than two different labels cannot be solved exactly, but solutions produced are usually near the global optimum. Graph-cuts segmentation is only able to find a global optimum for binary labeling (i.e., two labels) problems, such as foreground/background image segmentation **[8, 9 and 10]**. Extensions have been proposed that can find approximate solutions for multi-label graph cuts problems.

3. PROPOSED METHOD

In computer vision, segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super-pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s). An initial segmentation is required to partition the image into homogeneous regions for merging. The mean-shift segmentation for initial segmentation has been chosen. The Mean Shift segmentation is a local homogenization technique that is very useful for damping shading or tonality differences in localized objects. It replaces each pixel with the mean of the pixels in a range-r neighborhood and whose value is within a distance d.

The Mean Shift takes usually 3 inputs:

- A distance function for measuring distances between pixels.
- A radius. All pixels within this radius will be accounted for the calculation.
- A value difference. From all pixels inside radius r, we will take only those whose values are within this difference for calculating the mean.

Let us take below example to describe the behavior of man-shift segmentation. The following Figure 1 is an example of the mean shift initial segmentation. After a few iterations, a stable non-isotropic configuration will be arrived as shown in the Figure 1a.

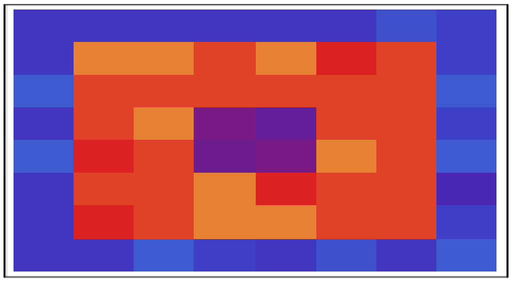

Figure 1: Mean shift initial segmentation

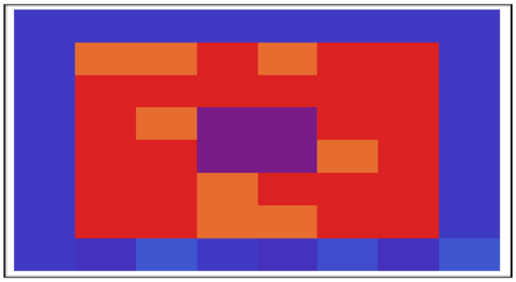

Figure 1a: Mean shift initial segmentation, a stable non-isotropic configuration

3.1 Maximum similarity Region Merging

After mean shift initial segmentation, we have many small regions available. To direct the following region merging process, we need to represent these regions using some descriptor and define a rule for merging. A region can be described in many aspects, such as the color, edge, texture, shape and size of the region. Among them the color histogram is an effective descriptor to represent the object color feature statistics and it is widely used in pattern recognition and object tracking , etc. In the context of region merging based segmentation, color histogram is more robust than the other feature descriptors. This is because the initially segmented small regions of the desired object often vary a lot in size and shape, while the colors of different regions from the same object will have high similarity. Therefore, the color histogram to represent each region is used.

The RGB color space is used to compute the color histogram in this work. Each color channel into 16 levels and then the histogram of each region is calculated in the feature space of 16×16×16 = 4096 bins are uniformly quantized. Denote by HistR the normalized histogram of a region R. The next problem is how to merge the regions based on their color histograms so that the desired object can be extracted.

In the interactive image segmentation, the users will mark some regions as object and background regions. The key issue in region merging is how to determine the similarity between the unmarked regions with the marked regions so that the similar regions can be merged with some logic control. Therefore, we need to define a similarity measure ρ(R, Q) between two regions R and Q to accommodate the comparison between various regions. There are some well-known goodness-of-fit statistical metrics such as the Euclidean distance, Bhattacharyya coefficient and the log-likelihood ratio statistic **[1]**. Here we choose to use the Bhattacharyya coefficient to measure the similarity between R and Q

$$\rho(R,Q)=\sum_{u=1}^{4096}\sqrt{Hist_R^u \cdot Hist_Q^u} \qquad (1)$$

where $\mathbf{Hist_R}$ and $\mathbf{Hist_Q}$ are the normalized histograms of **R** and **Q**, respectively, and the superscript u represents the $\mathbf{u^{th}}$ element of them. Bhattacharyya coefficient **ρ** is a divergence-type measure which has a straightforward geometric interpretation. It is the cosine of the angle between the unit vectors. The higher the Bhattacharyya coefficient between R and Q is, the higher the similarity between them is the geometric explanation of the Bhattacharyya coefficient actually reflects the perceptual similarity between regions. If two regions have similar contents, their histograms will be very similar, and hence their Bhattacharyya coefficient will be very high, i.e. the angle between the two histogram vectors is very small. Certainly, it is possible that two perceptually very different regions may have very similar histograms. Fortunately, such cases are rare because the region histograms are local histograms and they reflect the local features of images. Even in case two perceptually different regions have similar histograms, the similarity between them is rarely the highest one in the neighborhood.

The RGB/Bhattacharyya descriptor is a very simple yet efficient way to represent the regions and measure their similarity. It has been successfully used to measure the similarity between target model and candidate model in the popular kernel based object tracking method. However, it should be stressed that other color spaces, such as the HSI color space, and other distance measures, such as the Euclidean distance between histogram vectors, can also be adopted in the proposed region merging scheme.

3.2 Object and background marking

In the interactive image segmentation, the users need to specify the object and background conceptually. Similar to the users can input interactive information by drawing markers, which could be lines, curves and strokes on the image. The regions that have pixels inside the object markers are thus called object marker regions, while the regions that have pixels inside the background markers are called background marker regions. The green markers to mark the object while using blue markers to represent the background are used. Please note that usually only a small portion of the object regions and background regions will be marked by the user. Actually, the less the required inputs by the users, the more convenient and more robust the interactive algorithm is after object marking, each region will be labeled as one of three kinds of regions: the marker object region, the marker background region and the non-marker region. To completely extract the object contour, we need to automatically assign each non-marker region with a correct label of either object region or background region. For the convenience of the following development, we denote by $\mathbf{M_o}$ and $\mathbf{M_B}$ the sets of marker object regions and marker background regions, respectively, and denote by **N** the set of non-marker regions.

3.3 Maximal similarity based merging rule

After object/background marking, it is still a challenging problem to extract accurately the object contour from the background because only a small portion of the object/background features are indicated by the user. The conventional region merging methods merge two adjacent regions whose similarity is above a preset threshold. These methods have difficulties in adaptive threshold selection. A big threshold will lead to incomplete merging of the regions belonging to the object, while a small threshold can easily cause over-merging, i.e. some object regions are merged into the background. Moreover, it is difficult to judge when the region merging process should stop.

Object and background markers provide some key features of object and background, respectively. Similar to graph cut and marker based watershed, where the marker is the seed and starting point of the algorithm, the proposed region merging method also starts from the initial marker regions and all the non-marker regions will be gradually labeled as either object region or background region. The lazy snapping cutout method proposed, which combines graph cut with watershed based initial segmentation, is actually a region merging method. It is controlled by a max-flow algorithm. In this work, an adaptive maximal similarity based merging mechanism to identify all the non-marker regions under the guidance of object and background markers is presented.

Let **Q** be an adjacent region of R and denote by $\mathbf{S_Q} = \{\mathbf{S_i^Q}\}$ **i = 1, 2, q** the set of Q's adjacent regions. The similarity between Q and all its adjacent regions, i.e. $\rho(Q, \mathbf{S_i^Q})$, i = 1,2, . . . ,q, are calculated. Obviously R is a member of $\mathbf{S_Q}$. If the similarity between R and Q is the maximal one among all the similarities $\rho(Q, \mathbf{S_i^Q})$, we will merge R and Q. The following merging rule is defined: Merge *R* and *Q*

$$\text{if } \rho(R,Q) = \max_{i=1,2,\ldots,q} \rho(Q, S_i^Q) \qquad (2)$$

The merging rule (2) is very simple but it establishes the basis of the proposed region merging process. One important advantage of (2) is that it avoids the presetting of similarity threshold for merging control. Although “max” is an operator that is sensitive to outliers, we empirically found that it works well in our algorithm. This is mainly because that the histogram is a global descriptor of the local region and it is robust to noise and small variations. Meanwhile, the Bhattacharyya coefficient is the inner product of the two histogram vectors and it is also robust to noise and variations.

The marker regions cover only a small part of the object and background. Those object regions that are not marked by the user, i.e. the non-marker object regions, should be identified and not be merged with the background. Since they are from the same object, the non-marker object regions will usually have higher similarity with the marker object regions than the background regions. Therefore, in the automatic region merging process, the non-marker object regions will have high probabilities to be identified as object.

3.4 Merging Process

The whole maximal similarity based region merging (MSRM) process can be divided into two stages, which are repeatedly executed until no new merging occurs. Our strategy is to merge background regions as many as possible while keep object regions from being merged. Once we merge all the background regions, it is equivalent to extracting the desired object. Some two-step strategies have been used for image pyramid construction. The proposed strategy aims for image segmentation and it is guided by the markers input by users.

In the first stage, we try to merge marker background regions with their adjacent regions. For each region **B** □ $\mathbf{M_B}$, we form the set of its adjacent regions $\mathbf{S_B = \{A_i\}_{i=1,2,\ldots,r}}$. Then for each $\mathbf{A_i}$ and $\mathbf{A_i}$ □□$\mathbf{M_B}$, we form its set of adjacent regions $\mathbf{S_{Ai}}$ = $\mathbf{\{S_{Ai}{}^{j}\}\ j_{=1,2,\ldots,k}}$. It is obvious that **B** □ $\mathbf{S_{Ai}}$. The similarity between $\mathbf{A_i}$ and each element in $\mathbf{S_{Ai}}$, i.e. $\boldsymbol{\rho}\mathbf{(A_i, S_{Ai}{}^{j})}$, is calculated. If **B** and $\mathbf{A_i}$ satisfy the rule (2), i.e.

$$\rho(A_i, B) = \max_{j=1,2,\ldots,k} \rho(A_i, S_j^{A_i}) \qquad (3)$$

Then **B** and $\mathbf{A_i}$ are merged into one region and the new region will have the same label as region **B**:

$$\mathbf{B = B\ U\ Ai} \qquad (4)$$

Otherwise, **B** and $\mathbf{A_i}$ will not merge.

The above procedure is iteratively implemented. Note that in each iteration, the sets $\mathbf{M_B}$ and **N** will be updated. Specifically, $\mathbf{M_B}$ expands and **N** shrinks. The iteration stops when the entire marker background regions $\mathbf{M_B}$ will not find new merging regions. After the region merging of this stage, some non-marker background regions will be merged with the corresponding background markers. However, there are still non-marker background regions which cannot be merged because they have higher similarity scores with each other than with the marker background regions. Figure 2a shows that after the first stage merging, many regions belonging to the background (leaves, branches, etc.) are merged but there is still some non-marker background regions left. To complete the task of target object extraction, in the second stage we will focus on the non-marker regions in **N** remained from the first stage. Part of **N** belongs to the background, while part of **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.

After the first stage, for each non-marker (background or object) region **P** ∈ **N**, we form the set of its adjacent regions $\mathbf{S_P}$ = $\mathbf{\{H_i\}_{i=1,2,\ldots,p}}$. Then for each $\mathbf{H_i}$ that $\mathbf{H_i}$ □□$\mathbf{M_B}$ and $\mathbf{H_i}$ □□$\mathbf{M_o}$, we form its set of adjacent regions $\mathbf{S_{Hi}}$ = $\mathbf{\{S_j{}^{Hi}\}_{j=1,2,\ldots,k}}$. There is **P** □ $\mathbf{S_{Hi}}$. The similarity between $\mathbf{H_i}$ and each element in $\mathbf{S_H}$, i.e. $\boldsymbol{\rho}\mathbf{(H_i, S_j{}^{Hi})}$, is calculated. If **P** and $\mathbf{H_i}$ satisfy the rule(2),i.e.

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

then **P** and $\mathbf{H_i}$ are merged into one region

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

Otherwise, **P** and $\mathbf{H_i}$ will not merge.

Figure 2:. Region merging process

The above procedure is iteratively implemented and the iteration stops when the entire non-marker region set **N** will not find new merging regions. Figure 2b shows the merging result after the second stage. We see that some non-marker background regions, as well as some non-marker object regions, are merged, respectively, in this stage. The first and second stages of the algorithm are executed repeatedly until no new merging occurs. Figure 2c shows the merging output of the first stage in the 2nd round. Since there is no more merging action, the algorithm stops here. In the end, each region is labeled as one of the two classes: object or background. Then we can easily extract the object contour by extracting only the object regions, as shown in Figure. 2d. In most of our experiments, the algorithm will end within 2–3 rounds.

4. EXPERIMENTAL RESULTS

The proposed MSRM method is essentially an adaptive region merging method. With the markers input by the user, it will automatically merge regions and label the non-marker regions as object or background. The MSRM methods qualitatively by several representative examples are evaluated. The MSRM under different color spaces, distance metrics and initial segmentation are tested; the robustness of MSRM to user input markers as well as the failure cases of it.

In the second experiment, we want to separate a starfish from the complex background. Figure 3a shows that the mean shift initial segmentation results in severe over segmentation for both the target object and background. In this image, since the starfish lies relatively in the center of the image, we implicitly specify the regions which locate in the border of the image as background markers. Therefore we only need to draw the object markers (green strokes) in the image. As shown in Figure 3b, although there is no explicit user input background marker, the proposed MSRM method can still extract the desired object accurately.

Figure 3 (a): Initial mean shift segmentation and object markers and (b) the extracted object using the proposed method.

The proposed MSRM scheme can be naturally extended to extract multiple objects. Figure 4 shows an example to extract the two dogs in the snow background. Although the skin is the smaller dog in the left part of the scene is somewhat similar to the snow background, the proposed method still successfully separates it from the background

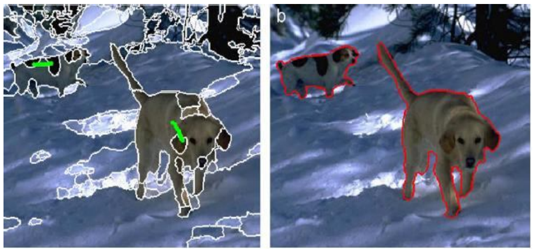

Figure 4: Multiple object extraction initial mean shift segmentation and interactive information. The two green markers mark two objects. (b) The two extracted objects

Meanwhile, although the contour of the bigger dog is complex, a simple marker was used to extract it out. The execution time of the MSRM depends on a couple of factors, including the size of the image, the initial segmentation result, the user input markers and the content of the image. We implement the MSRM algorithm in the MATLAB 7.0 programming environment and run it on a PC with P4 2.6 GHz CPU and 1024 MB RAM.

Although the RGB color space and Bhattacharyya distance are used in the proposed MSRM method, other color spaces and distance metrics can also be used in MSRM. In this section, we present examples to verify the performance of MSRM under different color spaces and distance metrics, as well as different initial segmentation.

The effect of color space on the region merging result. In this experiment, the RGB color images are converted into the HSI color space, and the HSI color histograms are then built and tested. The Bhattacharyya coefficient is calculated by using the HSI color histograms as in (1) for similarity measurement. Figure 5 shows the MSRM segmentation results on images bird and dogs. The left column shows the initially segmented images in the HSI color space, and the right column shows the finally segmented images by using the MSRM algorithm. We can see that the results are the same as those by using RGB color space with the Bhattacharyya distance.

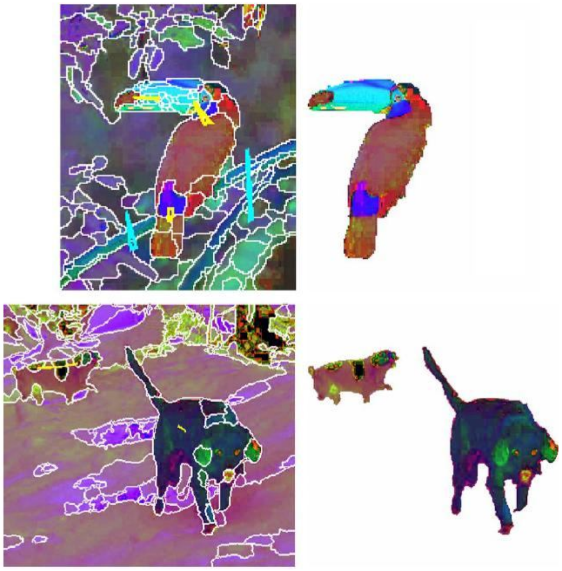

Figure 5: Left column: initial segmentation by mean-shift and the user input interactive information; right column: segmentation result by MSRM.

The effect of distance metric on the segmentation result is tested. In this experiment, the RGB color space is used but we replace the Bhattacharyya distance by the Euclidean distance. Denote by HistR and HistQ the normalized RGB color histograms of two regions R and Q, the Euclidean distance between them is defined as

$$\rho(R,Q) = -\sqrt{\sum_{u=1}^{4096}(\mathrm{Hist}_R^u - \mathrm{Hist}_Q^u)^2} \qquad (7)$$

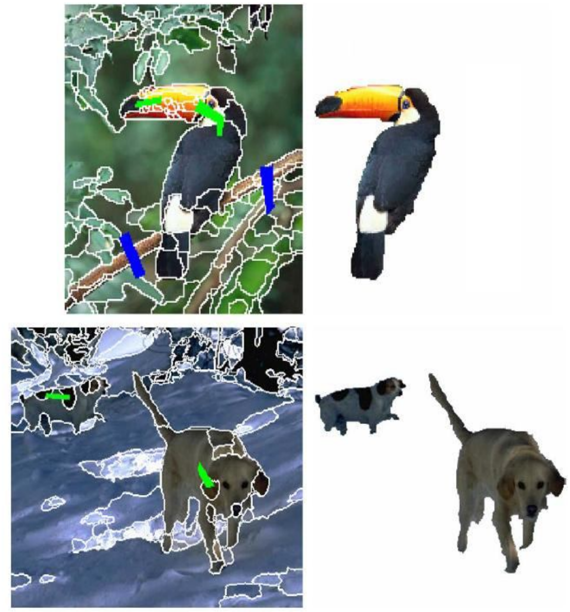

Figure 6: Left column: initial segmentation by mean shift and the user input interactive information; right column: region merging result by using the Euclidean distance for similarity measurement.

The Figure 6 shows the segmentation results on images bird and dogs. We see that the results are the same as those by Bhattacharyya distance. At last we test the MSRM algorithm with other initial segmentation. Besides mean shift, the super-pixel is another popular initial segmentation method. Different from mean shift, it partitions evenly the image but into many small regions.

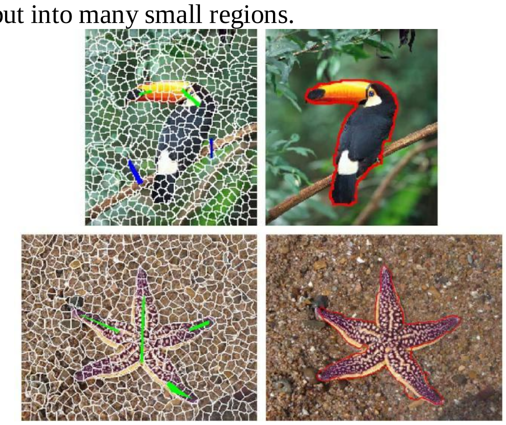

Figure 7: Left column: initial segmentation by super-pixel method and the user input interactive information; right column: region merging result by the proposed MSRM method

In this experiment, the super-pixel method is used for initial segmentation. Figure 7 shows the results on images bird and starfish-1. It can be seen that super-pixel leads to similar region merging results to those by mean shift.

However, for some images, e.g. the starfish-1, it may require more user input markers. This is mainly because super-pixel has more over segmentation than mean shift, and hence the statistics of some regions segmented by super-pixel is not as robust as that by mean shift initial segmentation. For compensation, more markers may be required for the same result.

4.1 Robust analysis and failure cases

The proposed MSRM method is an interactive scheme, i.e. the users need to input markers. Therefore, the marker input by the user is important to segmentation.

Figure 8: (a) The initial mean shift segmentation and different markers (green, blue and cyan markers) input by the user; (b) the extracted object is the same under different user inputs

By our many experiments, we find that the object can be correctly extracted as long as the markers can cover the main feature regions. To better illustrate this, we use an image with relatively simple features as an example. Refer to Figure 8a, by using the three different (green, blue and cyan) object markers as user inputs, in Figure 8b we can obtain the same object extraction result. This is because all the three object markers cover the main features (i.e. white and yellow colors) of the flower, i.e. the white petals and the yellow core.

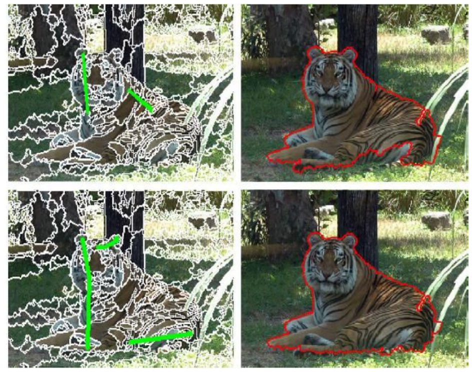

Figure 9: Segment a tiger by the proposed MSRM method with two groups of markers.

In the Figure 9, a tiger from the complex background with two groups of markers are separated. Obviously, the MSRM with more markers performs better than with few markers. Nonetheless, it still extracts the rough contour of tiger with even fewer markers. In general, the proposed MSRM algorithm could reliably extract the object contour from different backgrounds if the user input markers cover the main features of object and background.

However, it may fail when shadow, low-contrast edges and ambiguous areas occur. For example, in Figure 10a parts of the object regions are very similar to background. Although many markers were used to cover the object and background features, in some regions the proposed method does not achieve satisfying result. In Figure 10b parts of the object are present in the background, so the final segmentation is not very good.

Figure 10: Two failure examples of the proposed method: (a) parts of the object and background have very similar color features and (b) parts of the object are present in the background.

In addition, the proposed method is based on some initial segmentation such as mean shift or super-pixel. Therefore, if the initial segmentation does not provide a good basis to region merging, the proposed method may fail. Fortunately, many works have been proposed or are under development to improve the mean shift segmentation, which will make the proposed method more robust and efficient in image segmentation tasks.

4.2 Comparison of Results

In this section, the MSRM algorithm with the well known graph cut segmentation method under the same user input markers are tested. Since the original graph cut segmentation is a pixel based method, for a fair comparison with the proposed region based method, we extended the original pixel based graph cut (denoted by GCP) to a region based graph cut (denote by GCR), i.e. the nodes in the graph are mean shift segmented regions instead of the original pixels.

Figure 11 shows the segmentation results of the three methods on eight test images. The first column shows the mean shift initial segmentation result and the input markers (for the last four images, the image boundary is set as the background marker); the second column shows the results by GCP ; the third column shows the results by GCR ; and the fourth column gives the results by MSRM. We can see that with the same user input markers, the proposed MSRM method achieves the best results, while GCR performs better than GCP. It can be seen that GCR will miss some object regions and wrongly label some background regions as object regions.

To quantitatively compare the three methods, we manually labeled the desired objects in the test images and took them as ground truth. Then we computed the true positive rate (TPR) and false positive rate (FPR) for these segmentation results. The TPR is defined as the ratio of the number of correctly classified object pixels to the number of total object pixels in the ground truth, and the FPR is defined as the ratio of the number of background pixels but classified as object pixels to the number of background pixels in the ground truth. Obviously, the higher the TPR is and the lower the FPR is, the better the method is. Table 2 lists the TPR and FPR results by the three comparison methods on the eight test images in Figure 11. We can see that MSRM has the highest TPR and the lowest FPR simultaneously, which implies that it achieves the best segmentation performance.

Table 2
The TPR and FPR values of different methods on the test images.

Image	Method	TPR (%)	FPR (%)
Fruit	GC_P	93.14	2.37
	GC_R	96.56	3.37
	MSRM	98.97	0.37
Woman	GC_P	97.58	2.99
	GC_R	96.82	0.73
	MSRM	98.53	0.44
Bird	GC_P	87.49	3.64
	GC_R	90.62	3.55
	MSRM	94.64	0.29
Dogs	GC_P	66.79	0.68
	GC_R	78.99	0.32
	MSRM	92.85	0.11
Mona Lisa	GC_P	54.08	2.02
	GC_R	90.71	2.34
	MSRM	98.85	0.71
Flower	GC_P	95.20	2.09
	GC_R	96.67	2.46
	MSRM	97.59	1.08
Tiger	GC_P	68.50	12.53
	GC_R	79.20	2.42
	MSRM	91.70	0.75
Starfish-1	GC_P	77.50	2.35
	GC_R	87.42	2.66
	MSRM	90.25	0.26

It can also be seen that GCR has better performance than GCP. This shows that by grouping the similar pixels into small homogenous regions, mean shift initial segmentation improve the robustness of graph cut to noise and small pixel variations.

Figure 11: Comparisons between the graph-cut and proposed method.

5. CONCLUSION AND FUTURE SCOPE OF WORK

This proposed scheme is a novel region merging based interactive image segmentation method. The image is initially segmented by mean shift segmentation and the users only need to roughly indicate the main features of the object and background by using some strokes, which are called markers. Since the object regions will have high similarity to the marked object regions and so do the background regions, a novel maximal similarity based region merging mechanism was proposed to extract the object. The proposed scheme is simple yet powerful and it is image content adaptive. With the similarity based merging rule, a two stage iterative merging algorithm was presented to gradually label each non-marker region as either object or background. Extensive experiments were conducted to validate the proposed method in extracting single and multiple objects in complex scenes. The proposed scheme efficiently exploits the color similarity of the target object so that it is robust to the variations of input markers.

The proposed method provides a general region merging frame work. It does not depend essentially on mean shift segmentation and other color image segmentation methods can also be used for initial segmentation. Although some marker based interactive image segmentation methods have been proposed, the proposed algorithm firstly exploits a novel adaptive maximal similarity based region merging mechanism. In the future, we will explore how to introduce pixel classification into the merging process to make the algorithm more intelligent.

REFERENCES

1. T. Kailath. **The divergence and Bhattacharyya distance measures in signal selection**, IEEE Transactions on Communications Technology 15 (1), pp. 52–60, 1967

2. F. Meyer and S. Beucher. **Morphological segmentation**, Journal of Visual Communication and Image Representation 1 (1), pp. 21–46, 1990.

3. Y. Cheng. **Mean shift, mode seeking, and clustering**, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (8), pp. 790–799, 1995.

4. D. Comaniciu and P. Meer. **Mean shift: a robust approach toward feature space analysis**, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (5), pp.603–619, 2002.

5. C. Christoudias, B. Georgescu and P. Meer. **Synergism in low level vision**, Proceedings of the International Conference on Pattern Recognition, Vol. 4, pp. 150–155, 2002.

6. Q. Luo and T.M. Khoshgoftaar. **Efficient image segmentation by mean shift clustering and MDL-guided region merging**, IEEE Proceedings of the International Conference on Tools with Artificial Intelligence, pp. 337–343, November 2004.

7. Y. Boykov and V. Kolmogorov. **An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision**, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (9), pp. 1124–1137, 2004.

8. Asha D, Dr. Shivaprakash Koliwad and Jayanth.J, **A Comparative Study of Segmentation in Mixed-Mode Images**", International Journal of Computer Applications (0975 – 8887) Volume 31, No.3, October 2011.

9. D.Maheswari and Dr.V.Radha. **Improved Block Based Segmentation and Compression Techniques for Compound Images**, International Journal of Computer Science & Engineering Technology (IJCSET)-2011.

10. D.Sasirekha et al. **Enhanced Techniques for PDF Image Segmentation and Text Extraction**", International Journal of Computer Science and Information Security, Vol. 10, No. 9, September 2012.