



Improved technique for object detection using Similarity Based Region Merging

Garima Singh Rawat¹, Joy Bhattacharjee², Roopali Soni³

¹M.Tech Student (IV Sem)

^{2,3} Asst. Prof. Dept. of C.S.E

Oriental College of Technology, Bhopal (M.P.), INDIA

Abstract: Object segmentation is an important task in image processing and computer vision. We address a method to effectively discover a user's concept when object of interest are involved in input image. In this paper our method corporate frameworks for object retrieval using semi-automatic method for object detection because fully automatic segmentation is very hard for natural images. To improve the effectiveness of region merging we propose a new similarity based object retrieval. The users only need to roughly indicate the after which steps desired objects closed contour is obtained during the automatic merging of similar regions. A similarity based region merging mechanism is proposed to guide the merging process with the help of initial segmentation technique. Any two or more regions are merged with its adjacent regions on the basis of similarity. The proposed method automatically merges the regions that are initially segmented through initial segmentation technique, and then effectively extracts the object contour by merging regions.

Keywords: Initial segmentation, similar regions, similarity distance measure, region merging, flood fill.

I. Introduction

Image Segmentation is a process which partitioned image into multiple unique regions, where region is set of similar pixels. If I is set of all image pixels, then by applying segmentation we get different-different unique regions like $\{R_1, R_2, R_3, \dots, R_n\}$ which when combined formed the image I . Many approaches have been proposed earlier which includes probabilistic graphical models[1], normalized cuts[2], graph-cut method [3], region growing[4] etc. The automatic segmentation approaches may also fail even though they use prior information in segmentation process. The main reason behind this is the complexity of object segmentation of image in real applications. The complexity is due to the several reasons like shadow, low contrast areas, occlusion, cluttering and noise in the image. These reasons make the segmentation process quite difficult and challenging. The use of interactive image segmentation process is the solutions of such problems. In the interactive image segmentation the proper user's intervention is required for segmentation process. Because user gives the clue for segmenting the image in the process the results are more improved and satisfactory. Therefore Semi-automatic or interactive segmentation method is proposed, which use human expert knowledge as additional input and makes the segmentation problem more tractable. The goal of semi- interactive segmentation methods is to minimize the required user interaction time, while maintaining tight user control to guarantee the correctness of the results. In image segmentation image similarity measure play an important role. Region merging for object retrieval is an important task in many image processing applications. It has a wide application in area of crime prevention, intellectual properties, medical diagnosis, web searching etc. The object segmentation results are very much influence with how to groups sub regions. So, accurate object segmentation is possible if we combine both high level and low level priors effectively. To achieve this we introduce in our paper new approach object segmentation with region labeling and region growing.

II. Previous Work

Lei Zhang, Zhi Zeng, and Qiang Ji [5] proposed a method to extend the Chain Graph (CG) model to with more general topology and the associated methods for learning and inference. CG is a hybrid Probabilistic Graphical Model (PGM) which Contains both directed and undirected links. Its representation is powerful enough to capture heterogeneous relationships among image entities. For CG they first oversegment the image into superpixels and find out different heterogeneous relationships among image entities (superpixels, vertices or junctions, edges, regions etc.) They construct the CG model with parameterization of links with derived Joint Probability Distribution (JPD). They represent these links by either potential function or conditional probabilities. They first create a Directed Master Graph then create directed sub-graphs for some terms in the

JPD of Directed Master Graph. They segment the image into two parts foreground and background. Finally applying the probabilistic inference model applies in the foreground to find out Most Probable Explanation (MPE). Li Zhang and Qiang Ji [6] have proposed a Bayesian Network (BN) Model for both Automatic (Unsupervised) and Interactive (Supervised) image segmentation. They Constructed a Multilayer BN from the oversegmentation of an image, which find object boundaries according to the measurements of regions, edges and vertices formed in the oversegmentation of the image and model the relationships among the superpixel regions, edge segment, vertices, angles and their measurements. For Automatic Image Segmentation after the construction of BN model and belief propagation segmented image is produced. For Interactive Image Segmentation if segmentation results are not satisfactory then by the human intervention active input selection are again carried out for segmentation. Kittipat Kampa, Duangmanee Putthividhya and Jose C. Principe [7] design a probabilistic unsupervised framework called Irregular Tree Structure Bayesian Network (ITSBN). The ITSBN is made according to the similarity of image regions in an input image. ITSBN is a Directed acyclic graph (DAG) with two disjoint sets of random variables hidden and observed. The original image is oversegmented in multiscale hierarchical manner then they extracted features from the input image corresponding to each superpixel. According to these superpixels ITSBN is built for each level. After applying the learning and inference algorithms the segmented image is produced. Costas Panagiotakis, Ilias Grinias, and Georgios Tziritas [8] proposed a framework for image segmentation which uses feature extraction and clustering in the feature space followed by flooding and region merging techniques in the spatial domain, based on the computed features of classes. A new block-based unsupervised clustering method is introduced which ensures spatial coherence using an efficient hierarchical tree equipartition algorithm. They divide the image into different-different blocks based on the feature description computation. The image is partitioned using minimum spanning tree relationship and mallows distance. Then they apply K-centroid clustering algorithm and Bhattacharya distance and compute the posteriori distributions and distances and perform initial labelling. Priority multiclass flooding algorithm is applied and in the end regions are merged so that segmented image is produced. Eric N. Mortensen and Jin Jia [9] proposed a two layer BN model for image segmentation, which captures the relationships between edge segments and their vertices. Given a user input seed path, they use minimum path spanning tree graph search to find the most likely object boundaries. They also encode a statistical similarity measure between the adjacent regions of an edge into its a priori probability therefore implicitly integrating region information. In the early study [10] they use the similar BN model for both automatic and interactive segmentation. Their approach can find multiple non-overlapping closed contours before any given user's intervention. The intervention will serve as an evidence to help select a single closed contour that covers the object of interest. Lei Zhang, and Qiang Ji [11] develop a unified graphical model in which they combined directed graphical model and undirected graphical model. The combination allows capturing more complex and heterogeneous relationships among image entities. The unified model is more expressive and more powerful. But it only used for automatic segmentation not for interactive segmentation. They first propose to employ Conditional Random Field (CRF) to model the spatial relationships among image superpixel regions and their measurements. They introduce a multilayer Bayesian Network (BN) to model the causal dependencies that naturally exist among different image entities, including image regions, edges, and vertices. The CRF model and the BN model are then systematically and seamlessly combined through the theories of Factor Graph to form a unified probabilistic graphical model that captures the complex relationships among different image entities. Using the unified graphical model, image segmentation can be performed through a principled probabilistic inference. Lei Ding and Alper Yilmaz [12] develop a probabilistic hypergraph model which is the generalization of graphical model. In this model, Unlike those in standard graphs or hypergraphs are defined probabilistically and are learned from image features In addition ,hyperedge construction method is based on mean-shift segmentation. This method involves merging of superpixels within the hyperedge after initial segmentation based on probability. Jifeng Ning, LeiZhang, DavidZhang, and ChengkeWu[13] develop a image segmentation model based on maximal similar 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 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 R among all R 's adjacent regions. The 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.

III. Proposed Work

In our proposed method, an initial low level segmentation is required to partition the image into homogeneous regions and similarity function are used for region merging. Any existing low level segmentation methods, such as super-pixel [14], mean shift [15], watershed [16] and level set [17], can be used for this step. In the proposed method mean shift method is used also it is best for initial segmentation because it has less over segmentation

and can well preserve the object boundaries. In the proposed method, we only focus on the region merging and object extraction using flood fill.

A. Similarity Measure Using Metric Descriptor

In our method, an initial segmentation is required to partition the image into homogeneous regions, for merging. For initial segmentation we use low level segmentation known as mean shift method [15], we have a number of small regions. To guide the following region merging process, we need to represent these regions using some descriptor and define a rule for merging. Color descriptor is very useful to represent the object color features. In the context of region merging based segmentation, color descriptor is more robust than other feature descriptors because shape and size feature is very different while the colors of different regions from the same object will have high similarity. Therefore we use color histogram that represent each region in this paper. The RGB color space is used to compute the color histogram of each region in this paper. Next problem is how to merge the region based on their color histograms so that the desired object can be extracted. The key issue in region merging is how to determine the similarity between different segmented regions of image so that the similar regions can be merged by some logic control. Therefore we need to define a similarity measure formula [13] between two regions R and Q to accommodate the comparison between various regions and to perform merging process.

B. Merging Rule Using Bhattacharyya Descriptor

It is still a challenging problem to extract accurately the object contour from the background. The conventional region merging methods are merging two adjacent regions whose similarity is above a preset threshold [18, chapter 6.3]. These methods are difficult because of threshold selection. A big threshold will lead to incomplete merging belonging to object, while a small threshold will cause over-merging. Moreover it is difficult to detect when region merging process should stop. Proposed region merging method will start from any random segment part and start automatic region merging process. In our method, an initial segmentation is required to partition the image into homogeneous regions for merging. For initial segmentation we use low level segmentation known as mean shift method, we have a number of small regions. We use color histogram that represent each region and regions having maximum similarity are merged together. The RGB color space is used to compute the color histogram of each region in this paper. Next problem is how to merge the region based on their color histograms so that the desired object can be extracted. The key issue in region merging is how to determine the similarity between different segmented regions of image so that the similar regions can be merged by some logic control. Therefore we need to define a similarity measure formula between two regions R and Q to accommodate the comparison between various regions, for this there are some well known statistical metrics. Here we use Bhattacharyya coefficient [13] to measure the similarity between two regions say R and Q is:

$$P(R, Q) = \sum_{u=1}^{4096} \sqrt{Hist_R^u Hist_Q^u} \quad (1)$$

where $Hist_R$ and $Hist_Q$ are the normalized histogram of R and Q , respectively and superscript u represents the u^{th} element of them.

$$\cos \theta = (\sqrt{Hist_R^1, \dots, Hist_R^{4096}})^T (\sqrt{Hist_Q^1, \dots, Hist_Q^{4096}})^T \quad (2)$$

The higher the Bhattacharyya coefficient between R and Q is the higher the similarity between them i.e. smaller the angle θ . The geometric explanation of Bhattacharyya coefficient actually reflects the perceptual similarity between two regions. If two regions have similar contents then their histogram will be very similar, and their Bhattacharyya coefficient will be very high i.e. angle between histogram vectors is very small. Certainly it is possible that two different regions may have different histogram, such case happens very rarely. Similarity measure between two regions we use Bhattacharyya similarity which works well in proposed region merging method. The Bhattacharyya descriptor is a very simple yet efficient way to represent similarity between regions. However other color spaces e.g. HSV, YCbCr etc. and other distance measure such as the Manhattan, Euclidean and Chernoff can also be adopted for the region merging. Automatic region merging process covers the all part of segmented image, and after every step of merging we will decide whether we want to work on this image or not. Therefore in the automatic region merging process object regions will have high probabilities to be identified as object.

C. The Merging Process and Flood Fill Algorithm

Automatic region merging process covers the all part of segmented image, and after every step of merging we have choice whether we want to work on this image further or want to stop merging process. Therefore in the automatic region merging process object regions will have high probabilities to be identified as object. The whole object retrieval process is working in two stages. In first stage similar region merging process is as follows, our strategy to merge the small segmented image which is start with any randomly selected and merge

this with any of its adjacent regions with high similarity. Proposed method is unsupervised technique of region merging. We will merge segmented image regions with their adjacent regions as: if for each region Q we will set its adjacent regions S_B $i=1, 2, \dots, r$. If the similarity between any R_j for any $i=j$ is maximum. The merging rule [13] is defined below:

$$\rho(R_j, Q) = \max_{i=1,2,\dots,k} \rho(R_j, S_i^{Q_i}) \quad (3)$$

Then Q and R_j are merged into one region and new region is same leveled by

$$Q = QUR_j \quad (4)$$

The above procedure is implemented iteratively. Note that to each and every iterative step we will see whether the desired object is retrieved or not. Specifically the segmented region is shrinking; we will stop iteration when desired object contour is found. After the first stage i.e. when full part of object boundaries or likely to appear which seems in every step we apply second stage of algorithm for this we select a seed pixel on the object and expand this using four connectivity of pixels by using well known Flood Fill method.

Object Retrieval Algorithm

Input: (1) the image (2) the initial mean shift segmentation of input image

Output: desired multiple objects

While there is a merging up to extraction of object contour from input image:

1. Input is image **I** and initial segmentation
2. After step (1) stage of merging of initial segmented image using similar merging rule.
3. After step one number of regions are minimized and again apply similar region merging rule, this is an iterative procedure.
4. After retrieving object contour go to step (5).
5. Apply Region Labeling and after that Flood Fill method on the image after step (4).

Region Labeling (I)

% I: binary Image; $I(u, v)=0$: background, $I(u, v)=1$: foreground %

- 5.1. Let $m \leftarrow 2$
 - 5.2. for all image coordinates (u, v) do
 - 5.3. if $I(u, v)=1$ then
 - 5.4. Flood Fill (I, u, v, m)
 - 5.5. $m \leftarrow m+1$
 - 5.6. return the labeled image **I**.
- % After region labeling we apply Flood Fill method using DFS %
6. FloodFill ($I, u, v, label$)
 - 6.1. Create an empty **Stack S**
 - 6.2. Push ($S, (u, v)$)
 - 6.3. **While** **S** is not empty **do**
 - 6.4. $(x, y) \leftarrow \text{Pop}(Q)$
 - 6.5. If (x, y) is inside image and $I(x, y)=1$ then
 - 6.6. Set $I(x, y)=label$
 - 6.7. Push ($Q, (x+1, y)$)
 - 6.8. Push ($Q, (x, y+1)$)
 - 6.9. Push ($Q, (x-1, y)$)
 - 6.9.1. Push ($Q, (x, y-1)$)
 - 6.10. return

IV. Experimental Results using Proposed Method

Although RGB space and Bhattacharyya distance are used in proposed method, other color spaces and metrics can also be used. In this section, we present some example to verify the performance of unsupervised region merging and flood fill method in RGB color space. Similarity based object segmentation model is very simple as compared to the other existing methods of segmentation. The proposed method is less time consuming and provides better results. Because it is an interactive method so the time taken in the segmentation is totally dependent on the size, number of super pixels of the input image. The segmentation speed mainly depends on the complexity of the region merging and flood fill model. Object extraction time is totally dependent on the size and shape of the object of interest. Results show that our approach is flexible enough to segment different types of images. We achieved encouraging results on these images. As to the images with small changes or similar color of foreground and background, our algorithm will not be able to achieve an ideal segmentation effect. Another kind of error is caused by the clutter. When the background (e.g., the shadow) has a similar appearance as the foreground, the proposed model may not be able to completely separate them but still achieve satisfactory results.

A. Experimental Analysis of proposed Method

Fig.3 shows an example to extract starfish from a picture, After the initial segmentation of mean shift method[15], user need not to input any interactive information like use of object and background markers already seen in previous work[13]over the same image.fig 3a shows the input image from which object is to be extracted ,fig3b shows the initially segmented image using mean-shift method .As fig3c shows the image with object contour after applying region merging rule i-e Bhattacharyya distance[13] ,fig3d shows the gray scale image, for the object to be extracted RGB image is first to be converted into grayscale ,fig3e shows the object mask in binary format before applying floodfill for object retrieval gray scale image is further converted into binary form .Finally,fig3e shows the object (starfish) is well extracted from the complex background after the flood fill is applied to image in binary form.

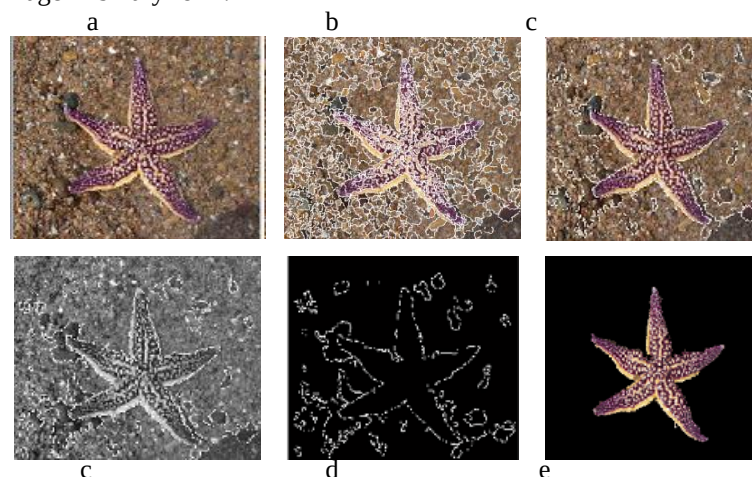


Fig.3. Shows the object retrieval using proposed algorithm (a) The image containing desired objects (b) image after region merging with object contour (c) Gray Scale image with object contour (d) Binary image with object mask(e) desired extracted object.

In the second experiment, we want to extract the flower from the background. In fig 4a shows the original input image and fig 4b shows the retrieved object from the complex background. Similarly fig 5b shows another example to extract bird from the background using the proposed scheme. In the proposed scheme of object segmentation, there is no need to explicitly use any object and background marker though it can extract object from the background accurately. We implement the proposed algorithm in the MATLAB 7.10 programming environment.



Fig.4. Shows the object retrieval using proposed algorithm (a) original input image (b) segmentation result by proposed method



Fig.5. Shows the object retrieval using proposed algorithm (a) original input image (b) segmentation result by proposed method

B. Comparison with MSRM Method

In this section, we compare our method of segmentation with the MSRM segmentation method. In the MSRM method of segmentation first we have to select the region of interest (foreground and background) by capturing the object in a contour by dragging the mouse on the object and giving some objects boundaries. Since the MSRM segmentation is a pixel based method, selection of region of interest makes the MSRM method a region based method. Figure 6 shows the segmentation results of the two methods on five test images. The first column shows the input image; the second column shows the results by MSRM method; the third column shows the results produced by our method of segmentation. We can see that with the same images, the proposed method achieves better results. It can be seen that MSRM will miss some object regions and wrongly label some background regions as object regions.

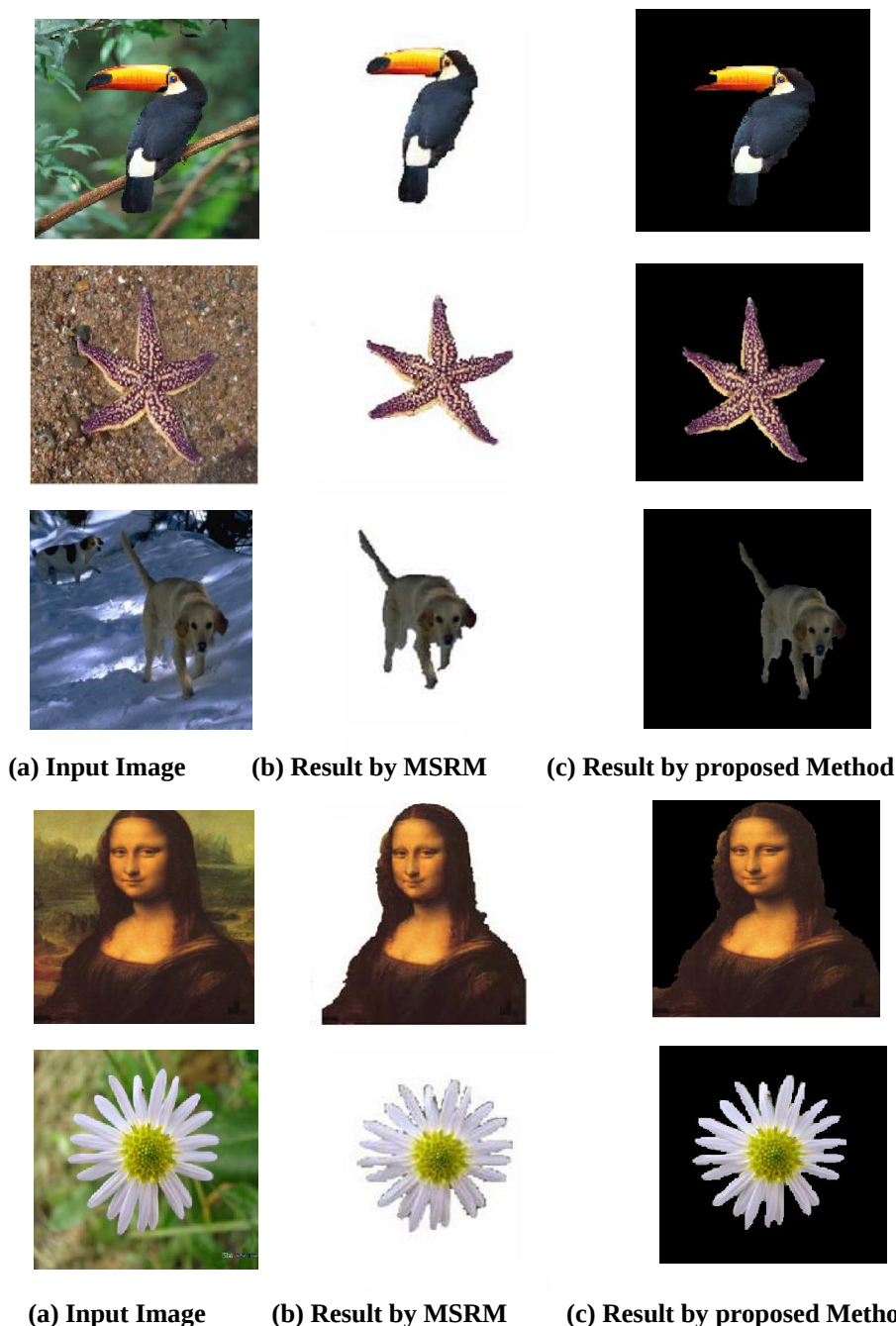


Figure 6: Comparisons between the MSRM Method and our proposed method. First column: Input image; Second column: result produced by MSRM method; third column: Result produced by proposed method.

One of the major drawbacks of MSRM method is time complexity. It takes more time in processing the object extraction from image and producing the results. Proposed method is very fast as compared to the MSRM method. We took same images for MSRM and proposed approach, our method produced results more quickly than the MSRM method. Time to retrieve object contours is depends upon the major factor which is number of initial segmented regions .Table1 lists the running time of two comparision methods over five test images,time for the proposed method is assumed to be worst case time required for object extraction. To quantitatively compare the two methods, we manually labelled 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 superpixels (i.e. superpixels to form object) to the number of total object pixels in the ground truth, and the FPR is defined as the ratio of the number of background superpixels (i.e. superpixels to not form object) 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 two comparison methods on the five test images in Figure 6. We can see that our method has the highest TPR and the lowest FPR simultaneously, which implies that it achieves the best segmentation performance. This shows that by grouping the similar pixels into small homogenous regions, mean shift initial segmentation improve the robustness of our method to noise and small pixel variations. Time to retrieve object contours is depends upon the major factor which is number of initial segmented regions shown in figure 6.

Table 1: The running time values of different methods on the test images

Images	Size	No. of superpixels	Method	RunningTime (Sec)
Bird	163*192	170	MSRM	6sec
			Proposed Method	0.5Sec
Dogs	335*295	196	MSRM	12 sec
			Proposed Method	1 Sec
Starfish	448*368	1088	MSRM	80 sec
			Proposed Method	6Sec
Monalisa	376*425	522	MSRM	32 sec
			Proposed Method	4 sec
Flower	229*216	250	MSRM	-
			Proposed Method	2Sec

Table2: The TPR and FPR values of different methods on the test images.

Images	Size	No. of superpixels	Method	TPR	FPR
Bird	163*192	170	MSRM	94.64	0.29
			Proposed Method	98.23	0.16
Dogs	335*295	196	MSRM	92.85	0.11
			Proposed Method	94.89	0.06
Starfish	448*368	1088	MSRM	90.25	0.26
			Proposed Method	97.97	0.09
Monalisa	376*425	522	MSRM	98.85	0.71
			Proposed Method	99.04	0.19
Flower	229*216	250	MSRM	97.59	1.08
			Proposed Method	98.00	0.40

V. Conclusion

This paper proposed a novel semi- interactive object segmentation framework using similar region merging and flood fill method. The similar region merging and flood fill method can systematically capture the relationships among different image regions to perform effective object segmentation. The image is first initially segmented into superpixels and use need not to roughly indicate any object and background marker, Since the object regions having similarity are merged after applying region merging based on Bhattacharyya rule. The model performs region-merging based on the Bhattacharyya rule on the regions. After region merging the number of super pixels is minimized and we are closure to object contour. Maximum similarity of a region shows that the region will be merged to other smaller regions present in the image and results in a bigger region. It is an iterative procedure and number of iterations depends on the user satisfaction. It can find that the object segmentation applications in many different areas like computer vision problems including object tracking, object recognition etc. Our experimental results demonstrate the promising capability for effective semi-interactive image segmentation. The proposed scheme is simple yet powerful and it is image content adaptive. Extensive experiments were conducted to validate our method in extracting object from complex scenes. We apply proposed method on database Caltech-256. Proposed method is very fast as compared to the MSRM method. With the same input images our method produced results quickly. In the future we can extend this method for extracting multiple objects from input image by using interactive image segmentation method as well as automatic segmentation method and merging similar regions using probability or by some other object metric descriptor

References

- [1] J. Lafferty, A. Mc Callum, and F. Pereira, "Conditional Random Fields: Probabilistic Models for segmenting and labeling Sequence Data," *Proc. Int'l Conf. Machine Learning*, pp. 282-289, 2001.
- [2] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888-905, August 2000.
- [3] Y. Boykov, G. Funka-Lei, "Graph cuts and efficient n-d image segmentation" *International Journal of Computer vision*, vol 70, no.2, pp.109-131, 2006
- [4] Felipe Calderero and Ferran Marques, "Region Merging Techniques using Information Theory Statistical Measures", *IEEE Transactions on Image Processing*, Vol. 19, No. 6, pp. 1567-1586, Jun. 2010.
- [5] Lei Zhang, Zhi Zeng, and Qiang Ji "Probabilistic Image Modeling With an Extended Chain Graph for Human Activity Recognition and Image Segmentation", *IEEE Transaction on Image Processing*, vol. 20, no. 9, September 2011
- [6] Lei Zhang and Qiang Ji, "A Bayesian Network Model for Automatic and Interactive Image Segmentation", *IEEE Transaction on Image Processing*, vol. 20, no. 9, September 2011.
- [7] Kittipat, Kampa, Duangmanee Putthividhya and Jose C. Principe, "Irregular Tree-Structured Bayesian Network for Image Segmentation", *IEEE International Workshop on Machine Learning for Signal Processing* September 18-21, 2011, Beijing, China.
- [8] Costas Panagiotakis, Ilias Grinias, and Georgeios Tziritas "Natural Image Segmentation Based on Tree Equipartition, Bayesian Flooding and Region Merging", *IEEE Transactions on Image Processing*, vol. 20, no. 8, August 2011.
- [9] E. N. Mortensen and J. Jia, "Real-time semi-automatic segmentation using a Bayesian network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2006, pp. 1007-1014.
- [10] E. Mortensen and J. Jia, "A Bayesian network framework for real-time object selection," in *Workshop Percept. Org. Comput. Vis.*, 2004, pp. 44-49.
- [11] L. Zhang and Qiang Ji, "Image Segmentation with a Unified Graphical Model," *IEEE Transactions Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1406-1425, August 2010.
- [12] Lei Ding, Alper Yilmaz "Interactive image segmentation using probabilistic hypergraphs," vol.43, pp.1863-1873, 2010.
- [13] Jifeng Ning, Lei Zhang, David Zhang, Chengke Wu, "Interactive image segmentation by maximal similarity based region merging," *Pattern Recognition*, Vol.43, pp. 445-456, 2010.
- [14] X. Ren, J. Malik, Learning a classification model for segmentation, *ICCV03*, vol. 1, pp. 10-17, Nice, 2003
- [15] Y. Cheng, Mean shift, mode seeking, and clustering, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (8) (1995) 790-799.
- [16] L. Vincent, P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (6) (1991) 583-598.
- [17] B. Sumengen, Variational image segmentation and curve evolution on natural images, Ph.D. Thesis, University of California
- [18] M. Sonka, V. Hlavac, R. Boyle, *Image Processing, Analysis and Computer Vision*, Thomson, 2007.