

Semi-supervised Method of Multiple Object Segmentation with a Region Labeling and Flood Fill

Uday Pratap Singh¹, Kanak Saxena², Sanjeev Jain³

¹Lakshmi Narain College of Technology, Bhopal,
usinghiitg@gmail.com

²Samrat Ashok Technological Institute, Vidisha,
kanak.saxena@gmail.com

³Madhav Institute of Technology & Science, Gwalior,
dr_sanjeevjain@yahoo.com

ABSTRACT

Efficient and efficient multiple object segmentation is an important task in computer vision and object recognition. In this work; we address a method to effectively discover a user's concept when multiple objects of interest are involved in content based image retrieval. The proposed method incorporate a framework for multiple object retrieval using semi-supervised method of similar region merging and flood fill which models the spatial and appearance relations among image pixels. To improve the effectiveness of similarity based region merging we propose a new similarity based object retrieval. The users only need to roughly indicate the after which steps desired objects contour is obtained during the automatic merging of similar regions. A novel similarity based region merging mechanism is proposed to guide the merging process with the help of mean shift technique and objects detection using region labeling and flood fill. A region R is merged with its adjacent regions Q if Q has highest similarity with Q (using Bhattacharyya descriptor) among all Q 's adjacent regions. The proposed method automatically merges the regions that are initially segmented through mean shift technique, and then effectively extracts the object contour by merging all similar regions. Extensive experiments are performed on 12 object classes (224 images total) show promising results.

KEYWORDS:

oversegmentation, similar regions, Bhattacharyya distance, region merging, mean shift, flood fill.

1. INTRODUCTION

CLASS-SPECIFIC (or category-level) multiple object segmentation is one of the fundamental problems in computer vision and object recognition. Although a human can delineate the object boundaries with much ease, segmenting images is not as ease for a computer. Its goal to segment an image into regions with each region solely containing object(s) of a class. As object segmentation requires that each segmented region to be a semantic object, it is much more

challenging than traditionally object segmentation [1, 2, 3, 4]. There has been a substantial amount of research on image segmentation including clustering based methods, region growing methods [5], histogram based methods [6], and more recent one such as adaptive threshold methods [7], level set methods [8], graph based methods [4, 9] etc. To achieve multiple object segmentation object classes should be represented appropriately. As the variance of object color/texture, shape within an object class can be large, it is difficult to obtain class-specific features that can describe object class accurately. In this regards, multiple object segmentation is a difficult problem. However multiple object segmentation is feasible due to the recent development of recognition and over segmentation (we shall use this in place of image segmentation) technique in computer vision.

The subjectivity of human perception of visual concept play an important role in content based image retrieval (CBIR) system. The fixed image similarity measure cannot meet need to adapt to different focuses of attention of different users. Similarity-based region merging for object retrieval is an important task in many image database applications. It is an object retrieval technique where an input image (which is called query image that is a part of input image) is matched against the images in the database. It has a wide application in area of crime prevention, intellectual properties, medical diagnosis, web searching and education etc. The multiple object segmentation results are very much influence with how to groups the obtained subregions. 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 a new approach known as multiple object segmentation with semi-supervised technique and flood fill.

The existing object segmentation algorithms are broken into three categories: supervised, semi-supervised and unsupervised. Supervised algorithm requires either manually segmented masks in training images [10, 11, 12] object part configurations [13], class fragments [14] or the specification of shape templates [11]. These algorithms are applicable only a particular object class [15] provided that class dependent priors are available. However, as a practical object segmentation systems need to handles a large number of classes of objects and most classes requires many training samples due to significant intra class shape and appearance of variances, it is important that the learning does not involve any human interaction. This makes unsupervised algorithms are more appealing. There has been sparse research in this direction.

Borenstien and Ullman [16] used the overlap between automatically extracted object fragments to determine foreground and background. Winn and Jojic [17] combined all images together to find a consistent segmentation based on assumption and the object shape and color distribution pattern are consistent within class and that the color/texture variability within object of a class is limited. Despite many years of research, unsupervised image segmentation techniques without human interaction still do not produce satisfactory results [18]. Therefore semi-supervised segmentation methods incorporating user interactions have been proposed [19, 20, 21, 22, 23] and are becoming more and more popular. For instance, in the active contour model (ACM), i.e. snake algorithm [19], a proper selection of initial curve by user lead to a good convergence of the true object contour. In order to do semantically meaningful image segmentation, it is essential to take priori (e.g. object part configuration [13], or class fragments [14]) information about the image into account.

The low level image segmentation methods, such as mean shift [24, 25], watershed [26] and super pixels [27], usually divide the image into small regions. These low level segmentation methods

provide a good basis for the subsequent high level operations, such as region merging. As a popular segmentation technique for color images, mean-shift [25] can have less segmented parts in comparison to watershed [23, 28] while preserving well the edge information of the objects. Because of less number of segmentation, the statistical features of each region, which will be exploited by the proposed semi-supervised similar region merging method and object detection can be more robustly calculated and then be used in guiding the region merging process.

In this paper, we proposed semi-supervised similar region merging method based on initial segmentation of mean shift. The proposed method will calculate the similarity of different regions and merge them based on maximum similarity. The object will then extract from the background when merging process ends. Although the idea of region merging is first introduced by [29] this paper uses the region merging for obtaining the contour for object and then extracting desired object from image. The key contribution of the proposed method is a novel maximal similarity based region merging technique, which is adaptive to image content and does not requires a present threshold. With the proposed region merging algorithm, the segmented region will be automatically merged and labeled, when the desired object contour is identified and avoided from background, the object contour can be readily extracted from background. The proposed algorithm is very simple but it can successfully extract the objects from complex scenes.

The rest of the paper is organized as follows; section 2 presents the region labeling and flood fill method with different type of statistical metric and merging rules using these metric. Section 3 performs extensive experiments for multi class segmentation to verify the proposed method and analysis. Section 4 performs experimental analysis for different color space and comparison with various existing algorithms. Section 5 concludes the paper and future works.

2. Region Labeling with Flood Filling

The most important task of object retrieval is to find out which pixels belong to which regions i.e. how many regions in the image and where these regions are located. This part of our proposed method is called region labeling or region coloring. Neighbouring pixels are pieced together in a step wise manner to build regions in which all pixels in that regions are assigned to unique number which we call label. The region label identification we use flood fill method, starting from a seed pixel and using neighbours connectivity e.g. using 4-connectivity or 8-connectivity. In the region marking algorithm we use following conversion first covert an RGB image into gray image say

$I(u,v)$.

$I(u,v) = 0$; for background image

$I(u,v) = 1$; for foreground image (object contour)

Once the regions in a binary image have been found, the next step is often to find the contours (that is, the outlines) of the regions. Like so many other tasks in image processing, at first glance this appears to be an easy one: simply follow along the edge of the region. We will see that, in actuality, describing this apparently simple process algorithmically requires careful thought, which has made contour finding one of the classic problems in image analysis. Before using region merging and flood fill method we will explain different probabilistic similarity measure descriptor. Some similarity measure descriptors are explained in section 2.1 after that in section

2.2 we explain how to use metric descriptor for similarity measure and section 2.4 algorithms for region labeling and flood fill method.

2.1 Different Similarity Measure Metrics

The Mahalanobis distance is simply a modified version of the Euclidean distance. Note, that the Euclidean distance does not take into account the correlations of the dataset, and is sensitive to the scale of the measurements. With the Mahalanobis distance, on the other hand, each dimension is given a weight which is inversely proportional to its variance (in order words, the covariance matrices of the random variables X and Y are two vectors of color histogram of any two different regions. Taking into consideration during the distance computation). The Mahalanobis distance, is expressed as:

$$Q_{\text{MAHANALOBIS}}(X, Y) = (\mu_x - \mu_y)^T \Sigma^{-1} (\mu_x - \mu_y). \quad (1)$$

where Σ is the covariance matrix of the two random variables combined.

The Hotelling's T-Square Statistic is a multivariate generalization of the t-test, which is commonly used in comparing the means of two univariate random variables. Hotelling's T-Square Statistic can be expressed as:

$$Q_{\text{TSQ}}(X, Y) = \frac{n_x n_y}{n_x + n_y} \sum_{i=1}^p \sum_{k=1}^p (\mu_{xi} - \mu_{yi}) C^{ik} (\mu_{xk} - \mu_{yk}) \quad (2)$$

where C^{ik} is the element in the i^{th} row and k^{th} column of the inverse of C, the pooled estimate of the covariance matrix for both populations, expressed as:

$$C = \frac{(n_x - 1)\Sigma_x + (n_y - 1)\Sigma_y}{n_x + n_y - 2} \quad (3)$$

Note that the T-Square statistic is simply the square of the T-test, thereby taking into account the correlation of all features in the set simultaneously.

The Kullback Leibler distance, belongs to a class of distance measures which compute the separation of two pdf's based on the dispersion of the likelihood ratio with respect to one of the densities. The distances in this class are derived from the equation:

$$Q(X, Y) = g \left(E_x \left\{ f \left(\frac{p_y(y)}{p_x(x)} \right) \right\} \right) \quad (4)$$

where g is a function which is continually increasing on the Real Line, E_x is the expectation of the random variable X, and f is a continuous convex function on the positive Real Line. The Kullback Leibler distance is derived by assigning:

$$f(x) = (x-1) \log(x), \text{ and} \quad (5)$$

$$g(x) = x \quad (6)$$

resulting in the expression:

$$Q_{KULLBACK} = \int_S (p_y(\xi) - p_x(\xi)) \log \left(\frac{p_y(\xi)}{p_x(\xi)} \right) d\xi, \quad (7)$$

where S denotes the probability space spanned by the feature vector sets compared. When the random variables are assumed to be Gaussian distribution,

$f_x(X) = \frac{1}{(2\pi)^{n_x/2} |\Sigma_x|^{1/2}} \exp \left\{ -\frac{1}{2} (x - \mu_x)^T \Sigma_x^{-1} (x - \mu_x) \right\}$ is the expression of the Kullback Leibler (KL) distance becomes:

$$Q_{KULLBACK} = \frac{1}{2} (\mu_y - \mu_x)^T (\Sigma_x^{-1} + \Sigma_y^{-1}) (\mu_y - \mu_x) + \frac{1}{2} \ln \left(\frac{|\Sigma_x^{-1} \Sigma_y + \Sigma_y^{-1} \Sigma_x - 2I|}{|\Sigma_x| |\Sigma_y|} \right) \quad (8)$$

where I is the identity matrix. It can be observed that under the assumption of equal covariance matrices for both feature sets, the KL distance is equivalent to the Mahalanobis distance.

The Bhattacharyya distance belongs to the same class as the KL distance. In this case however, the functions f(x) and g(x) in Equation (2.4) are expressed as:

$$f(x) = -\sqrt{x}, \text{ and} \quad (9)$$

$$g(x) = -\text{Log}(x) \quad (10)$$

The general form of the Bhattacharyya distance is given by:

$$Q_{BHATTACHARYA} = \log \left(\rho(p_x(x), p_y(y)) \right), \quad (11)$$

where

$$\rho(p_x(x), p_y(y)) = \sqrt{\int p_x(\xi) p_y(\xi) d\xi} \quad (12)$$

represents the Bhattacharyya coefficient (geometrically interpreted as the cosine of the angle between the pdf's of the two random variables being compared) which measures the amount of similarity between the two arbitrary segmented regions. If the random variables are assumed to be Gaussian, the Bhattacharyya distance may be expressed as:

$$Q_{BHATTACHARYA} = \frac{1}{4} (\mu_y - \mu_x)^T (\Sigma_x + \Sigma_y)^{-1} (\mu_y - \mu_x) + \frac{1}{2} \log \left(\frac{|\Sigma_x + \Sigma_y|}{2\sqrt{|\Sigma_x| |\Sigma_y|}} \right) \quad (13)$$

In proposed method, an initial segmentation is required to partition the image into homogeneous region for merging. For this we use any existing low level image segmentation methods e.g. watershed [26], super-pixel [27], level set [30] and mean-shift [24, 25] can be used for this step. In this paper we use mean-shift method for initial segmentation because it has less over segmentation and well preserve the object boundaries. For the initial segmentation we use the mean shift segmentation software the EDISON system [31] to obtain the initial segmentation map. Fig. 1 shows an example of mean shift initial segmentation. For detailed information about mean shift and EDISON system, please refer to [24, 25, 31, 32]. In this paper we only focus on the region merging.

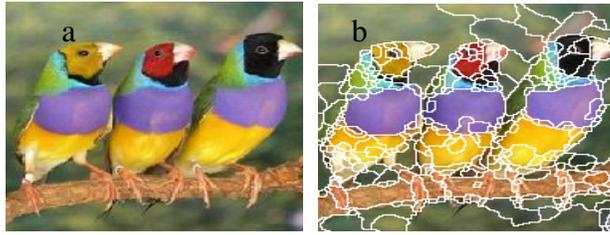


Fig. 1(a) is original image and fig.1 (b) represents Initial segmentation using mean shift.

2.2 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 segmentation method, 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. A region can be described in many aspects, such as texture [33], shape and size and color edge [34] of the regions. Among them color descriptor is very useful for 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 vary lots while the colors of different regions from the same object will have high similarity. Therefore we use color histogram represent each region in this paper. The RGB color space is used to compute the color histogram of each region in this paper. We uniformly quantize each color channels into 16 levels and then the histogram is calculated in the feature space of 4096 bins. 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 (14) between two regions R and Q to accommodate the comparison between various regions, for this there are some well known statistical metrics in section 2.1. Here we use Bhattacharyya coefficient [35, 36, 37, 38] to measure the similarity between two regions say R and Q is:

$$\rho(R, Q) = \sum_{u=1}^{4096} \sqrt{Hist_R^u Hist_Q^u} \quad (14)$$

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.

$$\cos\theta = \left(\sqrt{Hist_R^1 \dots Hist_R^{4096}} \right)^T \cdot \left(\sqrt{Hist_Q^1 \dots Hist_Q^{4096}} \right)^T \quad (15)$$

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 happen very rare. 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 (section 2.1) are also be adopted that for the region merging.

2.3 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 based on threshold [38]. 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. The entire region will be gradually labeled as either object region or background region. The lazy snapping cutout method proposed in [23], which combine graph cut with watershed based initial segmentation, is actually a region merging method. It is controlled by max-flow method [39]. In this paper we present an adaptive similarity based merging technique of regions either in object or in background.

Let Q be the adjacent region of R and denoted by $\overline{S}_Q = \{S_i^{Q_i}\}_{i=1,2,\dots,r}$ its set of Q 's adjacent regions. Using Bhattacharyya coefficient calculate similarity among Q 's adjacent regions $\overline{S}_Q = \{S_i^{Q_i}\}_{i=1,2,\dots,r}$. Obviously R will be one of the adjacent regions of S_Q . If the similarity between R and Q will be maximum then region R will be merged in region Q . We will use merging rule according to the formula defined as:

$$\rho(R_j, Q) = \max_{i=1,2,\dots,r} \rho(R_j, S_i^{Q_i}) \quad (16)$$

Equation (2.2.2) is the merging rule which establish the basis of proposed region merging process. Important advantage of equation (2.2.2) is that it prevents the use threshold for merging control, and the Bhattacharyya coefficient is the inner product of the two histogram vectors and it is robust to small noise and variations. The automatic region merging process cover the all part of segmented image, and after every step of merging we will 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.

2.4 The Merging Process and Flood Fill Algorithm

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. Some two step supervised merging process used in [40, 41] for image pyramid construction. Different from [40, 41] proposed method used image segmentation and it is semi-supervised 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 $S_B \quad i=1, 2, \dots, r$. If the similarity between any R_j for any $i=j$ is maximum i.e.

$$\rho(R_j, Q) = \max_{i=1,2,\dots,r} \rho(R_j, S_i^{Q_i}) \quad (17)$$

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

$$Q = QR_j \quad (18)$$

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

Multiple 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. For the input image (**I**) run EDISON system [31], we obtain initial segmentation
2. After step (1) stage of merging of initial segmented image (by mean shift method) using similar merging rule.
3. After step one number of regions are minimized and again apply similar region merging rule, this is and 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_i, v_i) =0: background, I (u_i, v_i) =1: foreground %

- 5.1. Let m ← -2
- 5.2. for all image coordinates (u_i, v_i) do
- 5.3. if I (u, v) =1 then
- 5.4. Flood Fill (I, u_i, v_i, m)
- 5.5. m ← m+1
- 5.6. return the labeled image I.

% After region labeling we apply Flood Fill method using Depth-First Search %

6. FloodFill (I, u_i, v_i, label)
 - 6.1. Create an empty **stack S**
 - 6.2. PUSH (Q, (u_i, v_i))
 - 6.3. **While** S is not empty **do**
 - 6.4. (x, y) ← POP (S)
 - 6.5. If (x, y) is inside image and I (x, y) =1 then
 - 6.6. Set I (x, y)= label
 - 6.7. PUSH (S, (x+1, y))
 - 6.8. PUSH (S, (x, y+1))
 - 6.9. PUSH (S, (x-1, y))
 - 6.10. PUSH (S, (x, y-1))
 - 6.11. Return
 - 6.12.

The proposed similar region merging method is an iterative method. After doing stage (1) what is the guarantee that the automatic similarity merging method will converge after a certain extent? To answer this question we will prove a proposition stated below.

Proposition: The Similarity region merging method in section 2.3 will converge i.e. every region in the image will be merged after a certain extent.

Proof: If a region Q has the maximal similarity with region R then region R will be merged with region Q i.e. $Q = Q \cup R$, in the first stage of proposed method this procedure is repeatedly and number of segmentation in the image is finite so the desired contour of object is obtained after a certain extent i.e. after k^{th} iteration.

From above analysis we see that the number of regions in image (after mean segmentation) is N (say) it will decrease in the process if iterative region merging. The whole algorithm will stop and all segmented region is in either object or in background. Therefore proposed algorithm converges and it will be label all the region of image.

3. Experiments on Multi-Class Segmentation

The proposed similarity region merging method is an semi-supervised technique, since it will automatically merge the regions and it will labels every regions either object or background. In section 3.1 we will first shows the semi-supervised similarity region merging method qualitatively by several representative examples; in section 3.1 we presents experimental analysis of proposed method considering different objects. In section 3.2 we test our proposed method for different color spaces and different distance metrics.

3.1 Experimental analysis

Figure 2 shows an example of how semi-supervised similarity region merging method extract object contour from a complex background. After initial segmentation by mean shift, automatic segmentation merging starts and after every step we test our merging results and also after which stage of merging we want to use flood fill method. Figure 2(a) is the initial segmented regions cover only small part but representative features of object and background regions. As shown in figure 2 the semi-supervised similar region merging steps via iterative implementation. Figure 2(a)-2(f), shows that the different steps of well extract object contour from the image and figure 2(h) is object mask. Figure 2(i) shows the extracted object using the two steps object retrieval method.

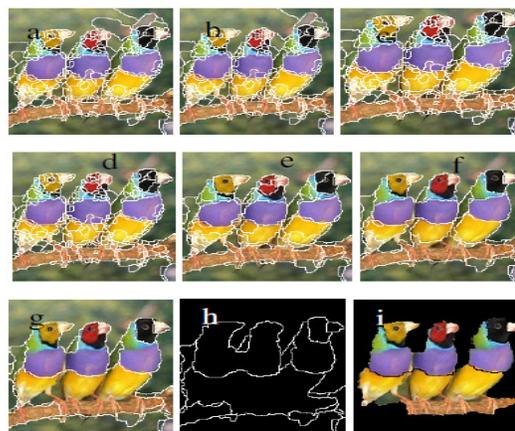


Fig.2. Different region merging Process are shown from 2(a) – 2(g), 2(h) is the object mask and 2(i) is the extracted object from background.

In the second experiment, we want to separate horses from background. Figure 3(b) shows that the initial mean shift segmentation results are serve our segmentation for extraction of object contour from complex background. Figure 3(b)-3(f) shows that different step for fully extracted object contour from input image and fig. 3(g) is the object mask. Fig.3 (h) shows the extracted object using the two steps object retrieval method. The execution time object retrieval using semi-supervised similar region merging and flood fill depends upon a number of factors, including size of image, the initial mean shift segmentation results etc. We implement semi-supervised similar region merging and flood fill algorithm in the MATLAB (R 2008a) 7.6 programming environment and run it on a PC with P4 2.80 GHz CPU and 2.0 GB RAM. Table 1 shows the running time of proposed method on testing image birds air planes, horses, cats and men etc..

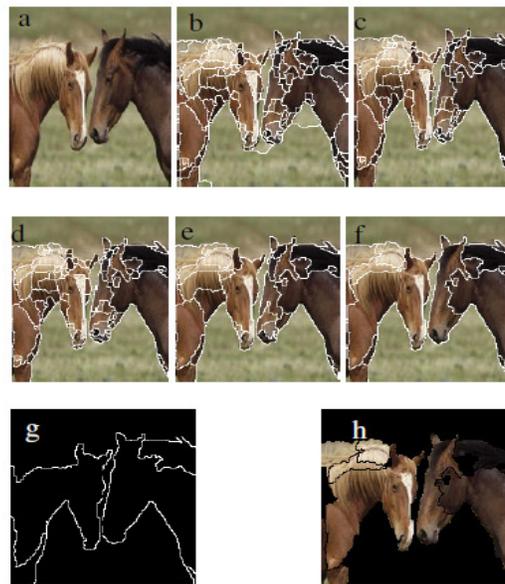


Fig.3 Different region merging steps are shown from 3(b)- 3(f), and fig. 3(g) shows the object mask and fig. 3(h) shows the extraction of both horses from background.

There are no comparison among the multiple object retrieval method using semi-supervised similarity region merging and flood fill method with hybrid graph model (HGM) [42] and well known Normalized cut [43]. But for single object retrieval proposed method is good in comparison to HGM and Normalized cut. Since the original graph cut segmentation is a pixels based method (GC_P) for a fair comparison of proposed method, we extended the original pixel based graph method (GC_P) to a region based graph cut (GC_R) i.e. the nodes in the graph are mean shift segmented region instead of original pixels. Semi-supervised similarity region merging and flood fill method are used for multiple object retrieval where as it is not possible using nHGM and HGM [42].

If we compare these three methods on images, aero planes, birds and horses. We can see that proposed semi-supervised region merging method achieves the best results in comparison to others, while (GC_R) [29] performs better result in comparison to (GC_P) [29]. It can be seen that (GC_R) will miss some object regions and wrongly label background regions as object regions.

3.2. Semi-supervised region merging under different color spaces, distance metrics and initial segmentation

Although RGB space and Bhattacharyya distance are used in proposed method, other color spaces and metrics are also used. In this section, we present some example to verify the performance of semi-supervised region merging and flood fill method. We first test the effect of color space on the region merging result. In this experiment RGB color space is converted into HSV and YCbCr. The Bhattacharyya coefficient is calculated for the histogram of these color spaces. Fig. 4 and figure (5) shows the semi-supervised region merging on the images aero planes and horses, after that we use flood fill method on HSV and YCbCr space for extraction of object.

The figure 6(a), 7(a), 8(a) and 9(a) shows the original images in the HSV and YCbCr color space and figure 6(b), 7(b), 8(b) and 9(b) shows the mask of object and 6(c), 7(c), 8(c) and 9(c) shows the object retrieve by using semi-supervised region merging and after that we use flood fill algorithm for object retrieval. We can see that the results are same as those by using RGB color spaces with Bhattacharyya distance.

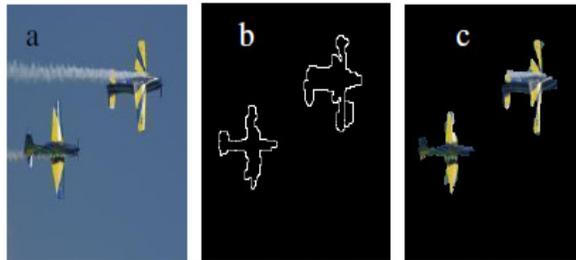


Fig.4. Object segmentation using Euclidean metric (RGB space)

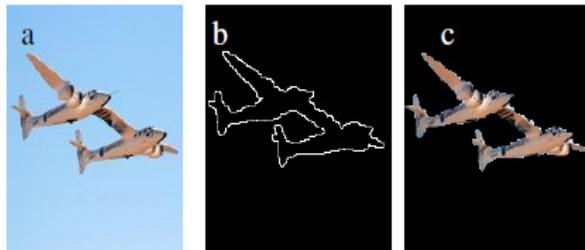


Fig.5. Object segmentation using Manhattan metric (RGB space)

Again we test the effect of distance metric on the segmentation results. In this experiment, RGB color space is used with Euclidean distance, we denote $Hist_R$ and $Hist_Q$ are normalized color histogram of two regions R and Q the Euclidean distance between them is defined as

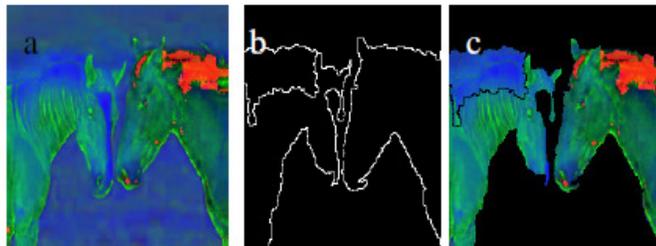


Fig.6. Object segmentation via proposed method (HSV space)

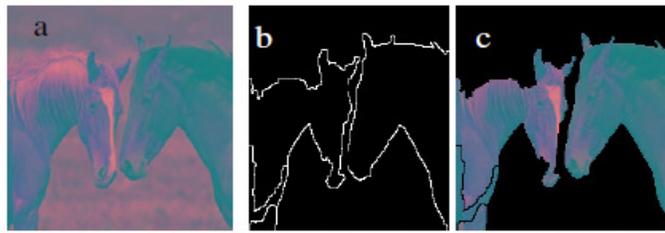


Fig.7. Object segmentation via proposed method (YCbCr space)

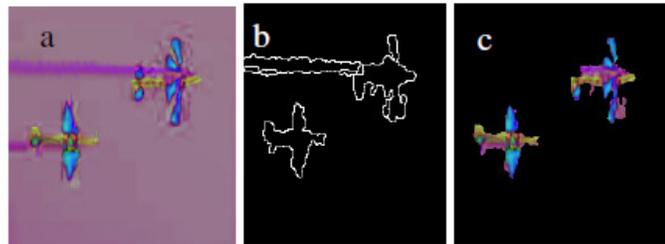


Fig.8. Object segmentation via proposed method (HSV space)

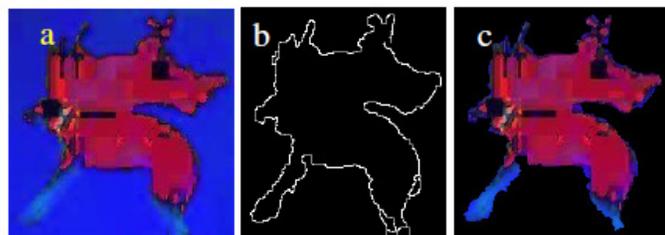
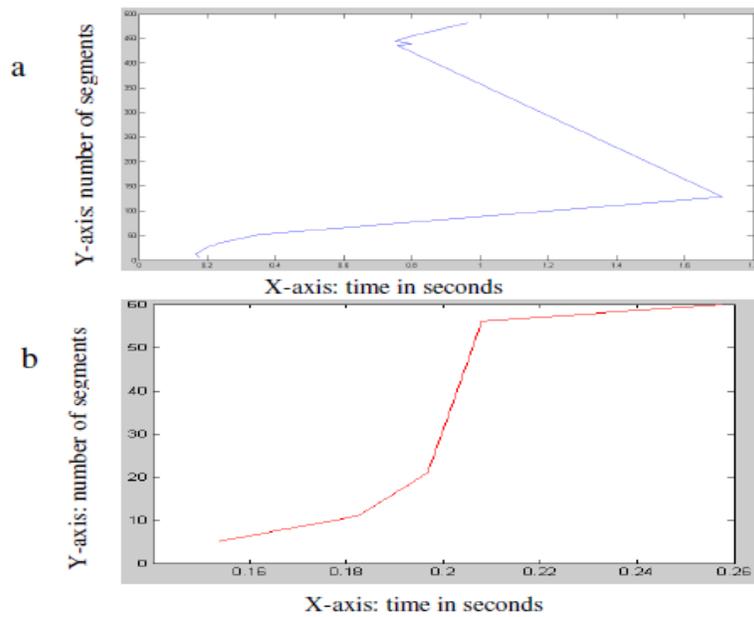


Fig.9. Object segmentation via proposed method (YCbCr space)



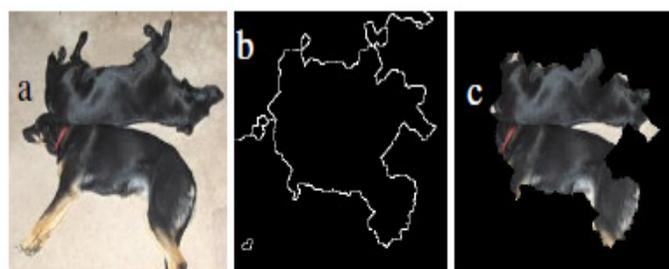
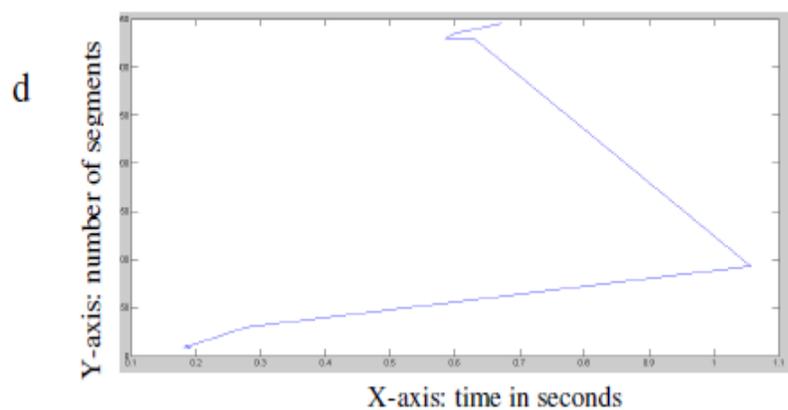
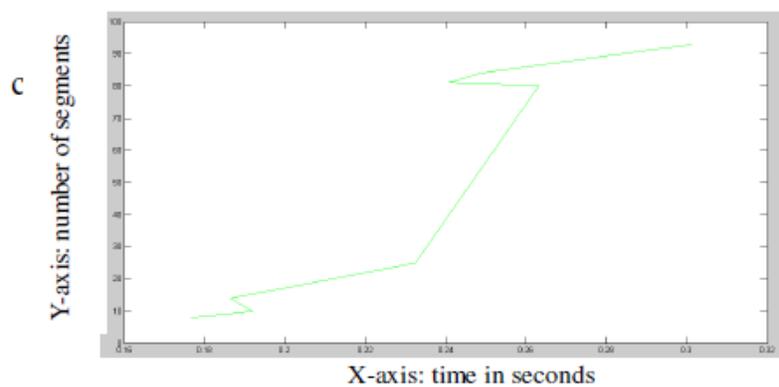


Fig.10. Some object segmentation using superpixel segmentation



Fig.11. Some object segmentation using superpixel segmentation

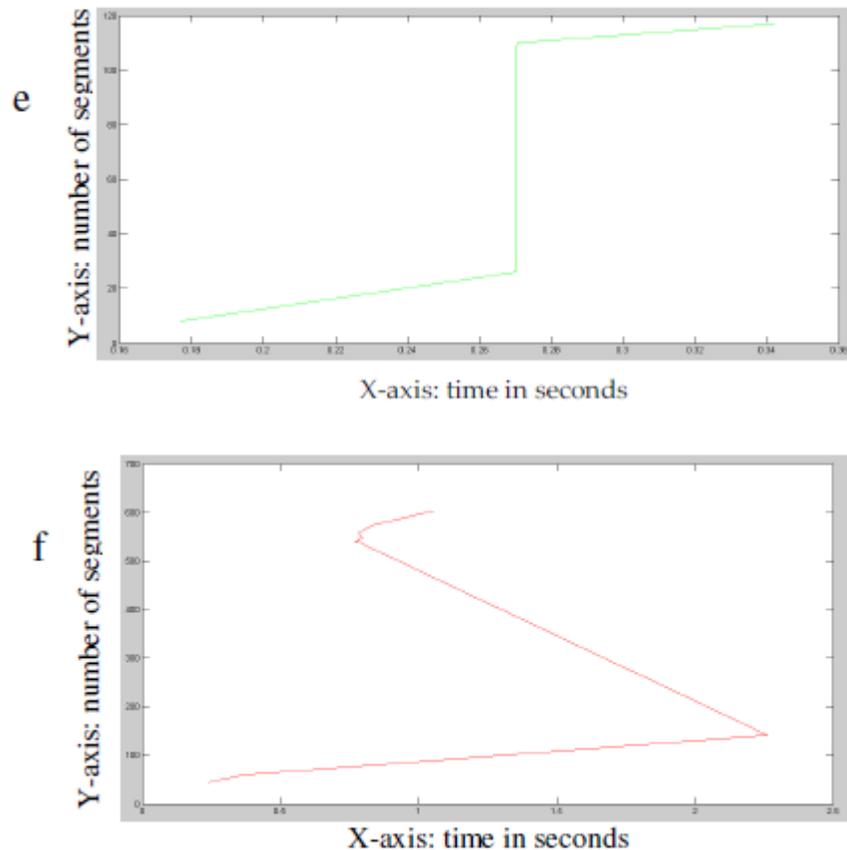


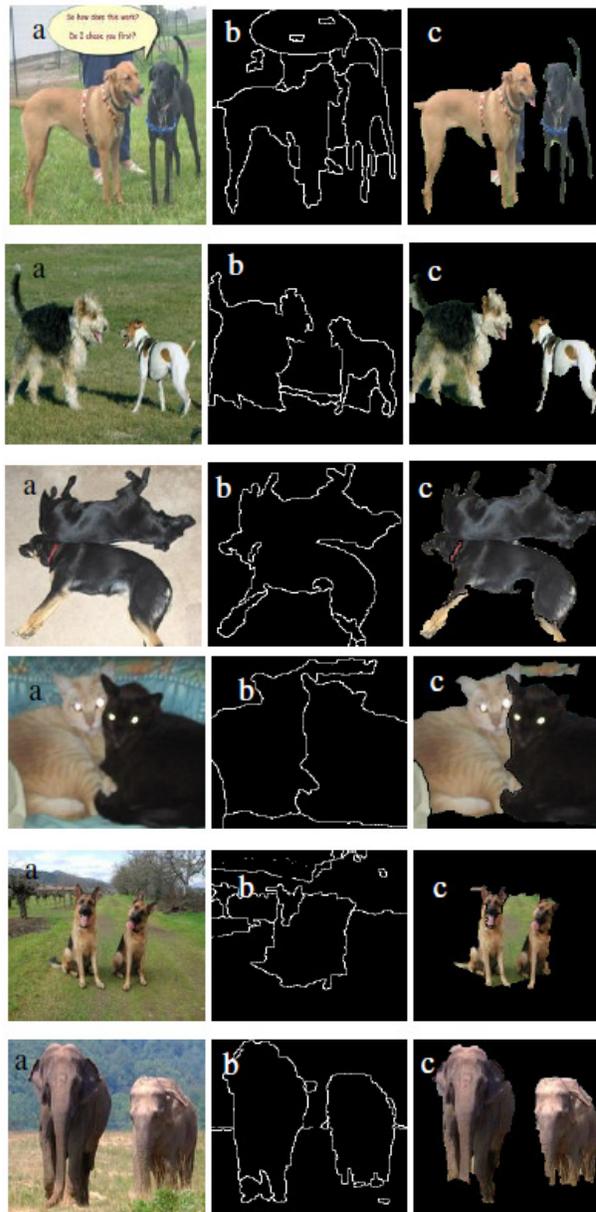
Fig. 12. shows the curve between time (sec.) and no. of pixels merged at each time segments (a) curve for fig. (6), (b) curve for fig. (7), (c) curve for fig. (8), (d) curve for fig. (9), (e) curve for fig. (10), (f) curve for fig. (11).

Fig. 10 and figure (11) shows the segmentation results on the images of dogs using initial segmentation superpixels method. We can see that the results are not good as those by initial mean shift segmentation.

4. Experimental Results using Proposed Method (for RGB and other color space)

To see how Similar Region Merging Flood Fill produces promising segmentation results in the case that there is a large variation in shape (including position, size, profile and pose) within an object class. We refer to the Fig. 6, 7, 8, 9. The task is to segment an object from the background scene. To segment a new image that may contain object of several classes, we use initial mean shift segmentation method to segment the image into K regions in which all containing instance (s) of object class. We assume that K is known a priori for each test image. In this section, we present the example to verify the performance of proposed method under the different color spaces like HSV, YCbCr etc. and for different initial segmentation like watershed [20], super pixels [21].

We first test effect of color space on region merging result. In this experiment RGB color space is converted into HSV, YCbCr color spaces. Figure (5) shows the object retrieval from RGB space, where as Fig. 6, 7, 8, and 9 shows the object retrieval from HSV and YCbCr color spaces. Also figure (10) and figure (11) shows that object retrieval using different initial segmentation besides mean shift, watersheds [26] and super pixels is another important method of initial segmentation. Different from mean shift it partitions the image into more number of regions. Due to large number of regions in the initial segmentation of images using super pixels its running time is more in comparison to mean shift initial segmentation and also results shows that the retrieve object after processing in similar fashion is not good as comparison with mean shift.



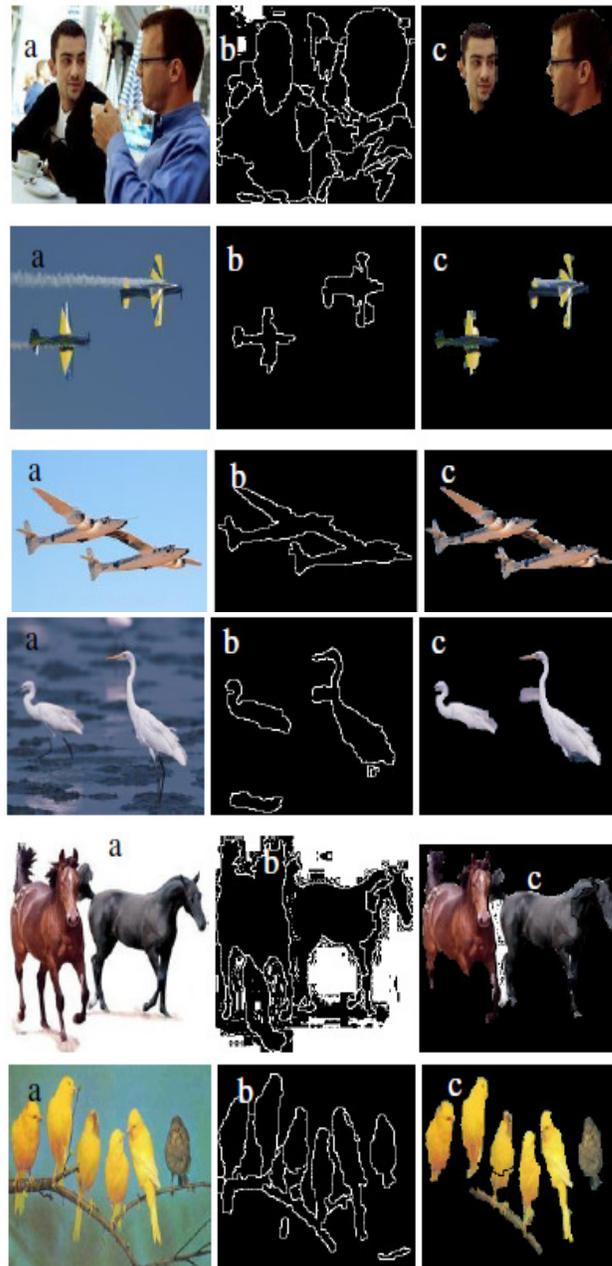


Fig. 13. Shows the object retrieval using proposed algorithm (a) The image containing multiple objects (b) objects mask (c) the final segmented results for multiple objects.

5. Conclusion

In this paper proposed a class specific multiple object segmentation method using semi-supervised similar region merging and flood fill algorithm. The image is initially segmented using mean-shift segmentation and automatic start of merging with any random segmented region and after each merging we check whether the object contours is obtained or not, if at any particular stage of merging object contours is obtained then use region labeling and flood fill algorithm and

choose seed pixels which objects that we want to extract from complex scene of background. The proposed scheme is simple yet powerful and it is image content adaptive.

In future we can extract multiple objects from input image by using unsupervised and supervised methods by merging similar regions using some metric and also merging is to be done on the basis of correlation coefficient (i.e. to find correlation coefficient between any two regions). Extensive experiments were conducted to validate the proposed method of extracting multiple objects from complex scenes. The proposed method is efficiently exploits the color similarity of the target. The proposed method provides a general region merging framework, it depends upon initially mean shift segmentation method or other color image segmentation methods [26, 30, 31, 44] can also be used for segmentation.

References:

- [1]. E. Sharon, A. Brandt, and R. Basri (2001) "Segmentation and Boundary Detection using Multiscale Intensity measurements" *Proc. IEEE conf. Computer Vision and Pattern Recognition*, pp.469-476.
- [2]. M. Galun, E. Sharon, R. Basri and A. Brandt (2003) "Texture Segmentation by Multiscale Aggregation of Filter Responses and Shape Elements" *Proc. IEEE Int'l conf. Computer Vision and Pattern Recognition*, pp. 716-723.
- [3]. E. Sharon, M. Galun, D. Sharon, R. Basri and A. Brandt, (June 2006) "Hierarchy and Adaptivity in Segmenting Visual Scenes," *Nature* vol. 442, no. 7104, pp. 810-813.
- [4]. J. Shi and J. Malik, (August 2000) "Normalized Cuts and Image Segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888-905.
- [5]. D.A. Forsyth, J. Ponce, (2002) *Computer Vision: A Modern Approach*, Prentice-Hall Englewood Cliffs, NJ.
- [6]. N. Bonnet, J. Cutrona, M. Herbin, (2002) A no Threshold Histogram Based Image Segmentation Method, *Pattern Recognition* vol. 35, no.10, pp. 2319-2322.
- [7]. E. Navon, O. Miller, A. Averbuch, (2005) "Color Image Segmentation Based on Adaptive Local Thresholds," *Image and Vision Computing* vol. 23, no.1, pp. 69-85.
- [8]. S. Osher, N. Paragios, (2003) "Geometric Level Set Methods in Imaging," *Vision and Graphics, Springer*, New York.
- [9]. Y. Boykov, G. Funka-Lei, (2006) "Graph cuts and efficient n-d image segmentation" *International Journal of Computer vision*, vol 70. no.2, pp.109-131.
- [10]. B. Leibe, A. Leonardis, and B. Schiele, (2004) "Combined Object Categorization and Segmentation with an Implicit Shape model," *Proc. Workshop Sixth European Conf. Computer Vision*, pp. 17-32.
- [11]. Z. Tu, X. Chen, A.L. Yuille, and S.C. Zhu, (2006) "Image parsing: Unifying Segmentation, Detection, and Recognition," *Proc. Towards Category-Level Object Recognition*, pp. 545-576.
- [12]. S.X. Yu and J. Shi, (2003) "Object-Specific Figure Ground Segmentation," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 39-45.
- [13]. S.X. Yu, R. Gross and J. Shi, (2002) "Concurrent Object Recognition and Segmentation by Graph Partitioning," *Proc. Neural Information Processing System*, pp. 1383-1390.
- [14]. E. Borenstein and S. Ullman, (2002) "Class-Specific, Top-Down Segmentation," *Proc. Workshop Seventh European Conf. Computer Vision*, pp. 109-124.
- [15]. A.L. Yuille, P.W. Hallinan, and D.S. Cohen, (1992) "Features Extraction from Faces Using Deformable Templates," *Int'l J. Computer vision*, vol. 8, no.2, pp. 99-111.

- [16]. E. Borenstein and S Ullman, (2004) "Lerning to Segments," Proc. Workshop Eighth European Conf. Computer Vision, pp. 315-328.
- [17]. J. M. Winn and N. Jovic, (2005) "Locus: Learning Object Classes with Unsupervised Segmentation," Proc. 10th IEEE Int'l Conf. Computer Vision, pp. 756-763.
- [18]. D. Martin, C.Folwkes, D. Tal, J.Malik, (2001) "A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics," *Internatinal Conference on Computer Vision*, pp. 416-423.
- [19]. M. Kass, A. Witkin, D. Terzopoulos, (1987) (1990) "Snake: Active Contour Models," *International Journal of Computer Vision*, vol.1, no.4, pp. 321-331.
- [20]. F. Meyer, S. Beucher, "Morphological Segmentation," *Journal of Visual Communication and Image representation*, vol.1, no. 1, pp. 21-46.
- [21]. P. Felzenszwalb, D. Huttenlocher, (2004) "Efficient Graph Based Image Segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167-181.
- [22]. Q. Yang, C.Wang, X. Tang, M. Chang and Z. Ye, (2007) "Progressive cut and Image Cutout Algorithm That Models User Intention," *IEEE Multimedia* vol.14, no.3, pp. 56-66.
- [23]. Y. Li, J. Sun, C. Tang, H. Shum, (2004) "Lazy Snapping" *SIGGRAPH* vol. 23, pp.303-308.
- [24]. Y. Cheng, (1995) "Mean Shift, Mode Seeking and Clustering," *IEEE Transaction on Pattern and Machine Intelligence*, vol. 17, no.8, pp. 790-799.
- [25]. D. Comaniciu, P. Meer, (2002) "Mean Shift: A Robust Approach Towards Feature Space Analysis," *IEEE Trnsaction on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603-619.
- [26]. L. Vincent, P. Soille, (1991) "Watersheds in Digital Space: An Efficient Algorithm Based on Immersion Simulations," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583-598.
- [27]. X. Ren, J. Malik, (2003) "Learning a Classification Model for Segmentation," *ICCV03*, vol.1, pp. 10-17, Nice.
- [28]. Y. Li, J.Sun, H. Shum, (2005) "Vedio Object cut and Paste," *SIGGRAPH* vol. 24, pp. 595-600.
- [29]. J.Ning, et al., (2009) "Interactive Image Segmentation by Maximal Similarity Based Region Merging," *Pattern Recognition*, vol. 44, no.9 .
- [30]. B. Sumengen, "Variational Image Segmentation and Curve Evolution on Natural Images," *Ph.D. Thesis, University of California*.
- [31]. EDISON software. (<http://www.caip.rutgers.edu/riul/research/code.html>).
- [32]. C. Christoudias, B. Georgescu, P. Meer, (2002) "Synergism in Low Level Vision," *Proceedings of Internatinal Conference on Pattern Recognition*, vol. 4, pp. 150-155.
- [33]. T. Ozala, M. Pietikainen, T. Maenpaa, (2002) "Multi-resolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987.
- [34]. S. Birchfield, (1998) "Elliptical Head Tracking Using Intensity Gradient and Color Histograms," *Proceeding of IEEE Conference Computer Vision and Pattern Recognition*, pp. 232-237.
- [35]. K. Fukunaga, "Introduction to Statistical Pattern Recognition," *Second ed., Academic Press*, 1990.
- [36]. T. Kailath, (1967) "The Divergence and Bhattacharyya Distance Measures in Signal Selection," *IEEE Transaction on Communication Technology*, vol. 15, no. 1, pp. 52-60.

- [37]. D. Comaniciu, V. Ramesh, P. Meer, (2003) "Kernel-Based Object Tracking," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 25, no. 5, pp. 564-577.
- [38]. M. Sonka, V. Hlavac, R. Boyle, (2007) "Image Processing, Analysis and Computer Vision," *Thomson*.
- [39]. Y. Boykov, V. Kolmogorov, (2004) "An Experimental Comparison of min-cut/max-flow Algorithm for Energy Minimization in Vision," *IEEE Trasaction on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124-1137,.
- [40]. P. Meer, (1989) Stochastic Image Pyramids, *Computer Vision, Graphics and Image processing (CVGIP)*, vol. 43, no. 3, pp. 269-294.
- [41]. J.M. Jolion, (1992) "The Adaptive Pyramid a framework for 2D Image Analysis," *Computer Vision, Graphics and Image Processing (CVGIP), Image Understanding* vol. 55, no.3, pp. 339-348,.
- [42]. G. Liu, Z. Lin, X.Tang, and Y. Yu, (2007) "A Hybrid Graph Model for Unsupervised Object Segmentation," *Proc. 11th IEEE Int'l Conf. Computer Vision*, pp. 1-8. .
- [43]. J. Shi and J. Malik, (2000) "Normalized Cuts and Image Segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 1546-1562.
- [44]. J.Wang, V. Thiesson, Y.Xu, M. F. Cohen, (2004) "Image and Vedio Segmentation by Anisotropic Kerne Mean Shift," *Proceeding of the European Conference on Computer Vision, prague, Czech Republic*, vol. 3022, pp. 238-249.