An Unsupervised Salient Object Extraction Approach using Statistical Method and High Contrast Saliency Map

Tallha Akram[1], Qichang Duan[1], Hongying Xu[1], Atif Amin[2]

[1]College of Automation, Chongqing University, Chongqing 400044, P.R.China
[2]Department of Electrical Engineering, CIIT WahCantt, Pakistan

Abstract
This paper proposes a multilevel unsupervised image segmentation approach aimed at salient object extraction. The fundamental objective is the extrication of pronounced object with confined boundary. Taking full advantage of *HSV* color space, evident of being nearest colorspace to human perception, proposed methodology segment the image using Expectation Maximization (*EM*), to estimate the parameters of Gaussian Mixture Model (*GMM*), using heuristic initialization. The Binary Partitional Tree (*BPT*) then extract the prime focus from reduced color palette on the basis of largest eigenvector. The final amalgamation with high contrast saliency Map diminishes all inutile fragments and excerpt salient object with ensured boundary. The experimental data indicates that hybrid approach leads to improved color segmentation with the apparent assertion of prime object extraction.
Keywords: *EM Segmentation, Binary Partition Tree, Salient Object Extraction, Saliency Map, Histogram.*

1. Introduction

We are dealing with massive amount of data, coming from different type of sources, observations and measurements. The processed data provide different characteristics and properties which may further subjected for records and experiments or analysis and reasoning. The most important activity is to classify the data into groups and categories, so each group must follow same meanings in one sense or another, based on some criteria. Our main concern is to organize visual data to be precise, images into sensible patterns and groups to derive useful conclusions about them. The idea is been applied in many fields, in medicine for medical imaging and *IMRT* segmentation [1,19,20], Computational Biology [26], Social Networks [25], Data Mining [23], Evolutionary Algorithms [9], Robotics [27,30], machine Learning and computer vision, etc. The rapid advancement in machine learning for Computer Vision applications results in the growing need for image analysis and interpretation in a wide range of applications.

The processing of entire image directly to fetch the required information look unpractical and ineffective, so extensive research has been done and numerous image segmentation algorithms were proposed [14,45,28], however, loads of computational issues for precise grouping have still remained unsolved and yet there is no single standard approach to be followed. Image segmentation is the critical preprocessing step to the image recognition [34], compression [7], image visualization, repossession [31] and image retrieval [41]. The process of image segmentation entails the grouping of homogeneous regions, such that resulting pixels in each partitioned regions possesses an identical set of attributes listed, based on spectral values, grey levels , textural properties [3,11,35]. Each pixel in the image get the unique label of the segment where it belongs to, and crowded together to make single group/segment.

No particular image segmentation technique is preeminent for all kind of images, in order to overcome such obstruction, merging strategies were exercised with multiple algorithms support [4,12,29]. In our proposed methodology we were focused on image segmentation with respect to object extraction, an implementation of *CBIR*. We have developed salient object extrication method, from an image, which, behaves like a system, taking advantage of probabilistic techniques in object extraction followed by refinement. The algorithm steps initially with object extraction which exploits probabilistic segmentation of an Image in HSV color space using EM algorithm, which, based on initial set of Gaussian mixture model optimize the partitioning decisions. EM solutions are highly initial conditions dependent for optimal likelihood estimates and fast convergence [6]. Finally BPT will extract the salient object from segmented image on the basis of largest eigenvector. Most salient regions in image considered to

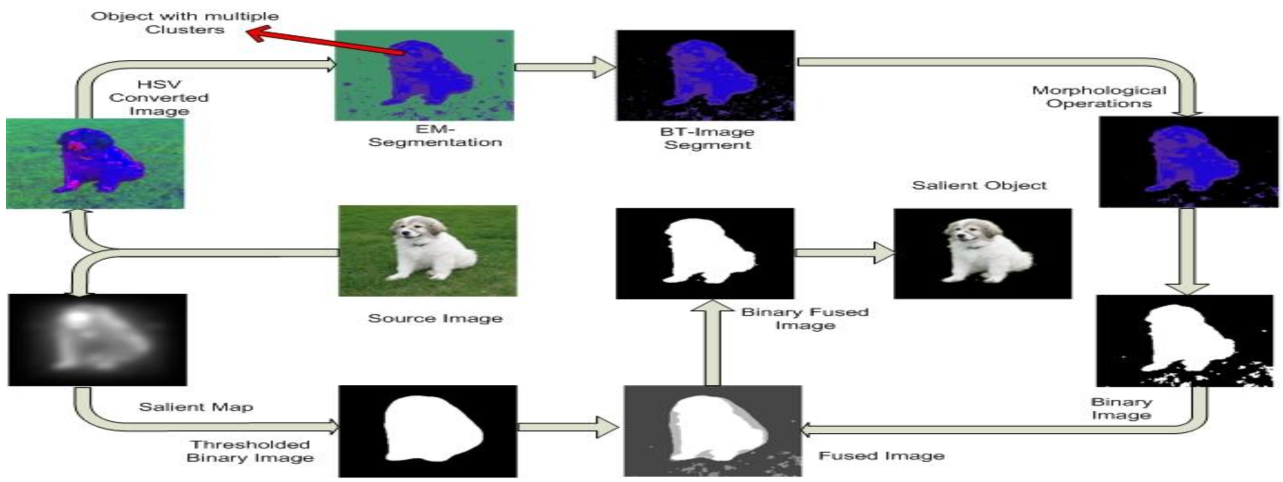


Fig. 1 A Pictorial illustration of Proposed Methodology, Outset with RGB source image (Center Left) to final step of Salient Object Extraction (center right).

have high level contrast, based on this conviction image is processed with low level features at multiple scale and center surround differences were computed. The contrast based saliency map merged with proposed salient clustering algorithm to extract the efficacious salient object, as demonstrated in Fig. 1.

One of the leading confront facing by researchers is to map intelligently between visual features (i.e., shape, color, texture, salient points) and high level semantic or salient objects concepts (e.g., flowers, buildings etc). In this regard selecting visual attention have vital role in image understanding, searching by emphasizing only on salient portions of image, instead of processing image as a whole, for different implementations, e.g., face recognition, object detection and recognition etc. In this context, we can model the relevancy of image with respect to salient objects, only by considering rest of the image details are deemed to be irrelevant and extraneous.

The structure of paper is as follows. In section 2, related work will be presented with the brief introduction of Expectation Maximization, Binary Partitional Tree and Saliency Map Computation followed by section 3, explains our methodology and the algorithm. Section 4 will discuss results and analysis and Conclusion and Future work will be discussed in Section 5.

2. Related Work

Many methodologies were floated, either solitary approaches or merged/hybrid approaches. Despite the virtue of comprehensive literature availability, smattering were selected e.g., [13] proposed approach for region of interest (*ROI*) extraction based on visual attention model and watershed segmentation, using visual attention model for the excerption of winner point among the salient points manifold, which subsequently acts as a seed for watershed segmentation.

For the selection of salient regions, low level features were also been utilized and classify the selected regions into different divisions including edges, lines, blobs and texture using correlation with neighboring area, with the region selection preferences [17]. Image assumed to be containing the same objects if region from first image shows high correlation with region of second image.
To adaptively and efficiently locating salient points, [24] uses salient blocks for semantic representation of prime objects. The image was divided into blocks, with the selection of fitting block according to similarity matrix acting on concept specific multi-features spaces.

In the pursuit of primary objects, tree structure approaches were also been scrutinized, e.g., [44] used *BPT* to record merging sequence performed using dissimilarity measures with the exploration of adjacency degree, color difference and area factor for the selection of appropriate subset of nodes. Based on Quadtree data structure, [39] keeps the most salient points in each quadrant according to saliency values from wavelet based methods. Finite mixture models are considered to be more flexible and robust probabilistic modeling tool, certainly it has few deficiencies, particularly, in estimation of number of components, lacking into consideration the spatial information and convergence to local minimum, [37] have tried to overcome few difficulties by addressing spatial information as prior knowledge of total number of components and develop

EM to estimate mixture density by using indirect information.

Saliency based object extraction is one of the leading research area [18,42,8,43] & has many implementations to *CBIR*. [18] Partition image into patches of same size for the computation of saliency region using multiple saliency maps. For the construction of scale invariant saliency map, [42] used multi-resolution feature contrast along with the image partitioned into homogeneous region using Non parametric kernel density Estimate (*NKDE*). Salient objects were extracted with the maximization of region saliency ratio. Similarly [8] has used learning behavior to effectively combine features. [43] has proposed the established sliding window based object detection paradigm using saliency map. Many other researchers have great contributions in the salient object detection using saliency map [32, 16, 10, 22].
[38] proposed contents based image retrieval based on random walk algorithm with relevance feedback. [36] Used a unified approach and used a selective visual attention model for salient contents retrieval.
In the proposed methodology improved saliency map acts as a connecting bridge for edge detection method, used canny edge detection and salient image segmentation map. Stages (1 & 2) combined for object retrieval.

2.1 HSV ColorSpace

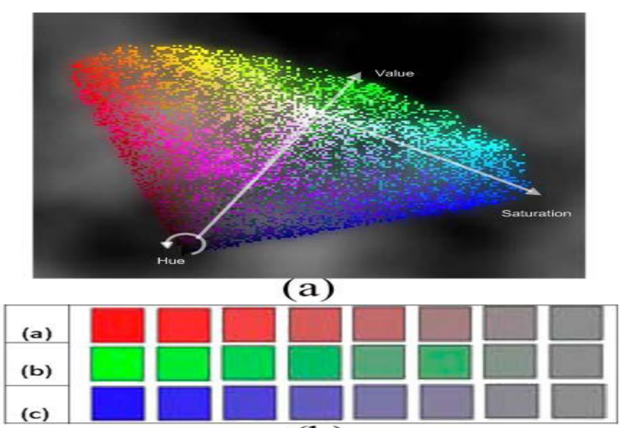


Fig. 2: (a) *HSV* (Hexacone) Colorspace (b) Variation in color perception with saturation, decreasing 1-0 from Left to Right for fixed Intensity.

HSV color space is considered to be the nearest color space for human color perception [9]. Compared with native *RGB* color space for digital images, the *HSV* is more uniform and function for distance computation between two colors in *HSV* is a uniformly distributed function. Transformation from *RGB* to *HSV* color space can be achieved through a non-linear transformation [40]. *HSV* colorspace is composed of hue, saturation and value, a three dimensional Hexacone model, where Hue (H) indicates angle ($0 \rightarrow 2\pi$) with colors such as red, green, blue, etc, within the range of 360^0 measures. Red at angle 0^o, Green at $2\pi/3$, blue at $4\pi/3$ and red again at 2π. The central axis represent the intensity of light, usually measured in percentage from ($0 \rightarrow 100\%$) as white. Saturation indicates the depth or purity of colors, e.g., blue can be divided into dark blue and light blue measured in radial distance from central axis as *0* and outer surface as *1*. For the given intensity and hue, if saturation changes from *0* to *1*, the perceived colors changes form shade gray to most pure form of colors represented by its hue as shown in Figure 2(b). The *HSV* color space is fundamentally differentiated from *RGB* colorspace in respect, it separates out the Intensity (luminance) from the color information (*chromaticity*). Again, of the two chromaticity axes, a difference in Hue of a pixel is found to be visually more prominent compared to that of the Saturation. For each pixel we, therefore, choose either its Hue or the Intensity as the dominant feature based on its saturation [33], and subsection a unique label.

2.2 EM-Frame Work

Suppose we have a dataset of observations, in current scenario, an *HSV* image, $\xi_i = (\xi_i^H, \xi_i^S, \xi_i^V)$ for, $i=1,2,\ldots,N$ is the i^{th} pixel value in *HSV* colorspace. We have represented this data as $(N\times D)$ matrix where dimension *D* represent *Hue*, *Saturation* & Value respectively. ξ_i is suppose to drawn from mixture of multi-variant normal density function. The mixture density estimator can be calculated using following expression with number of densities *K*.

$$p(\xi_i \mid \phi_j) = \sum_{j=1}^{K} \alpha_j p_j(\xi_i ; m_j, \partial_j) \qquad (1)$$

Where α_j is mixing parameters $\sum_{j=1}^{k} \alpha_j = 1$ for each *GMM* & $\phi_j = (\partial_j, m_j)$ where, ∂_j, m_j are standard deviation and mean of mixtures.

Let us introduce *K* dimensional binary random variable *z* having all entries zero except the K^{th} entry, i.e., y_i has actually been generated by the K^{th} normal density of mixtures $z_j = (z_{j1}, z_{j2}, \ldots, z_{jk})$ The value of z_j thus satisfy $z_k \,\varepsilon\, [0,1]$. We shall now define joint distribution $p(\xi, z)$ in terms of marginal distribution $p(z)$ and conditional distribution $p(\xi \mid z)$ is given by $p(z)p(\xi \mid z)$

and the marginal distribution of ξ is obtained by summing joint distribution over all possible state of

$$\sum_{z} p(z)p(\xi \mid z) = \sum_{j=1}^{k} \alpha_j N(\xi \mid m_j, \partial_j) \quad (2)$$

Let $g(\alpha_1, \alpha_2, ..., \alpha_{k-1}; m_1, m_2, ..., m_k; \alpha_1, \alpha_2, ..., \alpha_k)$ be the estimated parameters vector. To get initial value of above parameters, we will follow the algorithm.

E-Step: Evaluate the responsibilities using current parameters.

$$\beta_{ij}^{(u)} = \frac{\alpha_i p(\xi_i ; m_j^{(u)}, \partial_j^{(u)})}{\sum_{j=1}^{k} \alpha_j p(\xi_i ; m_j^{(u)}, \partial_j^{(u)})} \quad (3)$$

M-Step: Re-estimate the parameters using current responsibilities.

$$m_j^{(u+1)} = \frac{\sum_{i=1}^{N} \beta_{ij}^{(u)} \xi_i}{\sum_{i=1}^{N} \beta_{ij}^{(u)}}$$

$$\partial_j^{(u+1)} = \frac{\sum_{i=1}^{N} \beta_{ij}^{(u)} (\xi_i - m_j^{(u)})T(\xi_i - m_j^{(u)})}{\sum_{i=1}^{N} \beta_{ij}^{(u)}}$$

$$\alpha_j^{(u+1)} = \frac{1}{N} \sum_{i=1}^{N} \beta_{ij}^{(u)}$$

To check the convergence, finding the optimized mean and standard deviation.

$$L_j = max_j \frac{e^{(-1/2)(\xi_i - m_j^{EM})(\partial_j^{EM})^{-1}(\xi_i - m_j^{EM})^T}}{|\partial_j^{EM}|^{-1/2}} \quad (4)$$

Where y_i is the data and m_j^{EM} , ∂_j^{EM} are the means and standard deviation.

2.3 Heuristic Initialization

$g^{(init)}$ were calculated by following multi-dimensional histogram and minimum distance clustering methodology. The process can be summarized by following pseudo-code sequence.

1. Division of *HSV* color space into intervals by counting number of observed color elements into bins within *HSV* colorspace.
2. Selection of *k*-highest density bins to find centroid of *x*-observed elements belongs to bin.
3. If number of centroids less than *K*, rebuild the histogram with wider interval and perform recalculation of centroids.
4. Finalizing, by labeling all observed elements using minimum Euclidean distance classifier.

$$d_{min}(D_i, D_j) = \min_{c \varepsilon D_i, y \varepsilon D_j} d(C, Y)$$

2.4 Binary Tree

Consider image to be $(n \times n)$ size, with all grid points denoted by S for $s \,\varepsilon\, S$ where $s = (i, j)$ having row index and column index. Starting binary tree palette design by partitioning S into K^1 disjoint sets with the constrains of being partitioned into binary tree. Where K^1 is the predetermined palette size. Each node in binary tree can maximum have two nodes usually distinguished as "*Left*" and "*Right*" node. Each node of tree represent subset of its parent node which flows up the hierarchy to head node (ancestor of all nodes), *X*. The set of image pixels corresponding to node *n* is denoted by P_n and for the algorithm illustration, let head or root node is demonstrated by number *'1'* and children of node *n* are *2n* & *2n+1*. The partition of *X* is grouped by leaves with the demonstration of single color in the palette. Each leaf node is stored as linked list. The splitting of node into two sub nodes will not result in relocation of image color data.

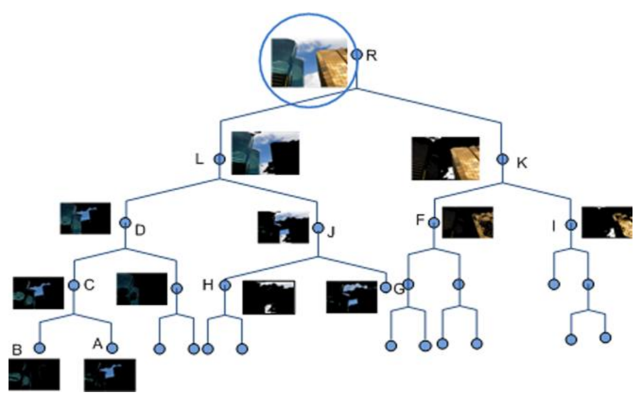


Fig. 3: A pictorial illustration for the algorithm of binary tree Image segmentation

The order of the node is very important in the process of splitting; so as to counter the procedure in an uncomplicated way is to define the second-order statistical properties.

$$R_n = \sum_{s\varepsilon P_n} x_s x_s^t$$
$$O_n = \sum_{s\varepsilon P_n} x_s \tag{5}$$

The mean of the cluster is a point with minimum squared deviation, so with assumption that mean is equal to q, a quantization value of cluster.

$$\mu_n = O_n / | P_n |$$

Cluster covariance is symmetric, defined as

$$\sum_{n+1} = \sum_n - \frac{1}{| P_n |} O_n O_n^{\ t}$$

New quantization level is achieved by splitting a node and selection of a suitable plane for best cluster's colors. This strategy is even favorable for large clusters, which incorporate the findings of greatest cluster variation and then splitting the cluster perpendicular to that direction by passing through clusters mean. The unit vector $\hat{u}$ is defined as

$$\sum_{s\varepsilon P_n} ((x_s - \mu_s)^t \hat{u})^2 = \hat{u}^t \sum_{n+1} \hat{u}$$

The eigenvector $\hat{u}$ corresponds to the largest eigenvector λ_n of $\sum_{n+1}$. Total squared variation in direction of $\hat{u}$ is.

$$\sum_{s\varepsilon P_n} ((x_s - \mu_s)^t \hat{u})^2 = \lambda_n$$

The points in the P_n is sorted as P_{2n}

$$P_{2n} = \{ s\,\varepsilon\, P_n : x_s \hat{u}_n^t \le \mu_n \hat{u}_n^t$$

& P_{2n+1} is defined as

$$P_{(2n+1)} = \{ s\,\varepsilon\, P_n : x_s \hat{u}_n^t > \mu_n \hat{u}_n^t \tag{6}$$

Finally the new statistics are created as

$$R_{2n+1} = R_n - R_{2n}$$

$$O_{2n+1} = O_n - O_{2n}$$

$$P_{2n+1} = | P_n | - | P_{2n} | \tag{7}$$

The splitting order is chosen with objective of maximum reduction in total square error at each stage ,which is defined as,

$$TSE = \sum_{All\,leaves\ n} \ \sum_{s\,\varepsilon\, p_n} \| x_s - \mu_n \| \tag{8}$$

The order in which nodes are split is chosen with the objective of producing the greatest reduction in *TSE* at each stage of the algorithm which is proportional to the total squared error variation along the principal eigenvector u of covariance with the approximation that best allocation of quantization level is to split the leaf, with largest principle eigenvector [21].

2.5 Saliency Map

The detailed flow can be visually depicted in Figure. 5 where it is broadly partitioned into three main steps, extraction, activation and normalization/combination [15]. In step *1* two spatial scales $(1/2, 1/4)$ were used and Gabor filter with *4* orientations (0^0, 45^0, 90^0, 135^0) was applied to calculate the orientation. The contrast map was computed using luminance variance in a local neighborhood. Center-surround feature maps were yield, with difference between "*center*" fine scale c & "*surround*" courser scale s. where c={2,3} for $(1/2, 1/4)$ & $\delta = 4$.

$$I(c,s) = | I(c)\ \Theta\ I(s) |$$

where Θ is across scale difference between two maps. Initiate with an objective of computing an activation map $A_M : [n]^2 \rightarrow \Re$ with the fact, an image location $I(i,j)\,\varepsilon$ $[n]^2$. Let us consider dissimilarity measure of $M_{DS}(i,j)$ & $M_{DS}(l,m)$.

$$d((i,j);(l,m) = | \log \frac{M_{DS}(i,j)}{M_{DS}(l,m)} | \overset{\Delta}{=} | M(i,j) - M(l,m) |$$

Let G_{AM} be fully connected directed graph of lattice $M(i,j)\varepsilon[n]^2$ with all (n-1) nodes. The weights of edge from node $[(i,j) \rightarrow (l,m)]$ is proportional to dissimilarity \& to their proximity in the domain of M_{DS}.

$$w((i,j),(l,m)) = d((i,j);(l.m)).F(i-l; j-m))$$

Where

$$F(x,y) = exp^{-(\frac{x^2+y^2}{2\partial^2})}$$

Where σ is set to approximately one fifth of map width endorsed as insensitive parameter of perturbation.

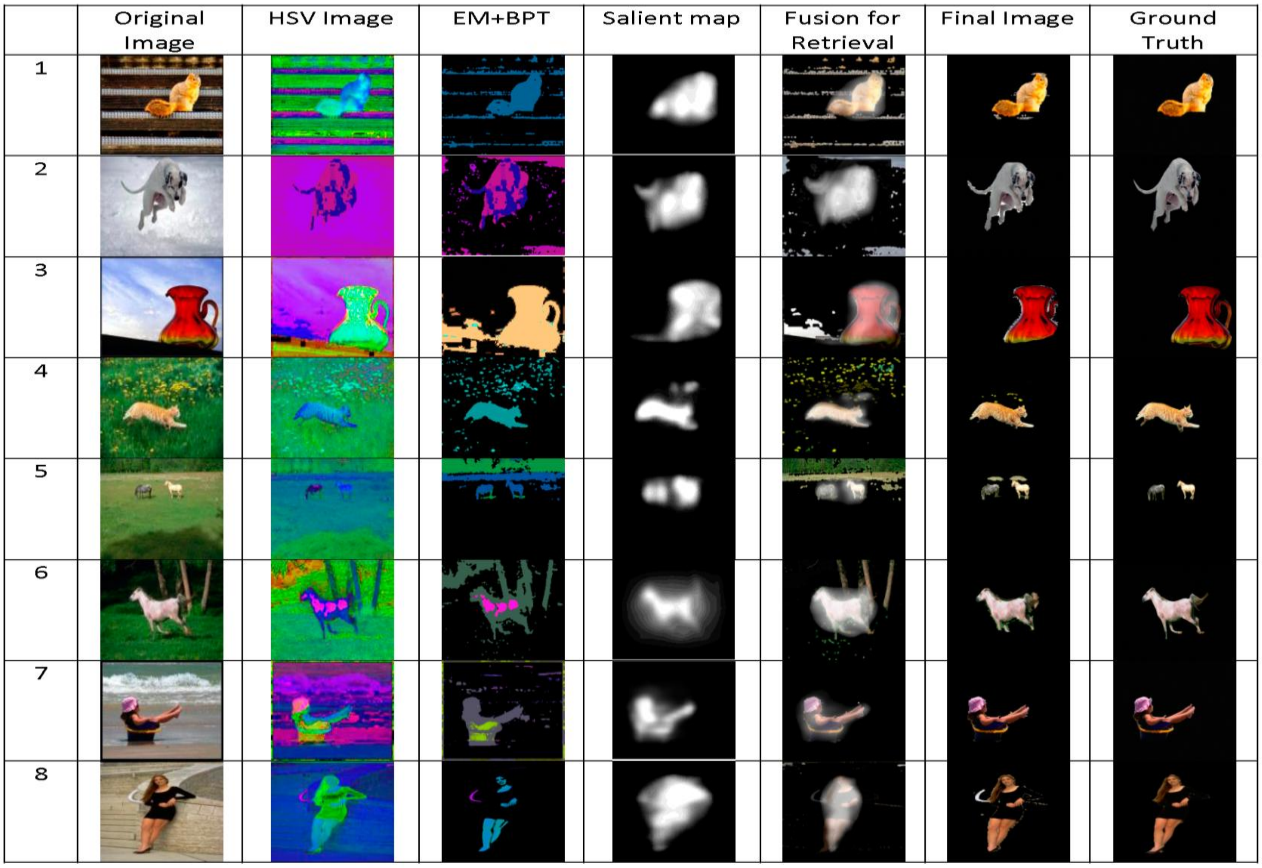


Fig. 4 Detailed Visual Illustration of Proposed Salient Object Extraction Methodology.

that transition into sub graph is likely and unlikely in case of similarity in *M* values [5,2].

The equilibrium distribution is computed by markov matrix repeated multiplication while initiated with uniform vector to yield Principal Eigen vector of matrix. The equilibrium state met with small number of iterations with computational complexity $O(n^4 K)$ when $K << n^2$.

The third \& final step refers to as normalization was mentioned "*Concentration mass on activation maps*", [5]. The resulting master map will be informative if mass is concentrated on individual activation maps preceding additive combination based on above notion, markovian algorithm were followed by following an activation eigenmap $A_{NORM} : [n]^2 \rightarrow R$. The graph G_{Norm} with total number of nodes n^2. For every node (i',j') connected with nodes (l',m') we found weights, Markov Chain was

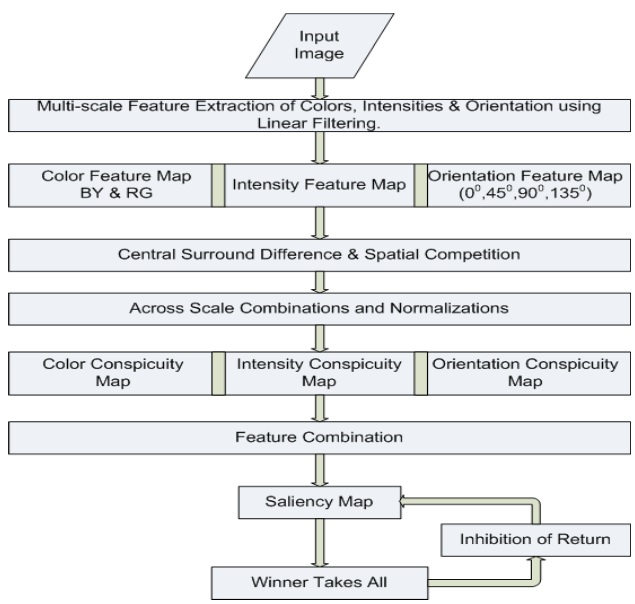


Fig. 5: General Architecture of Salient Flow Model

applied to directed graph G_{AM} by normalizing outward bound edges weights of each node
to unity & depicting correlation between set of parameters like nodes/states & weights & transition probabilities. The equilibrium distribution of this sequence is time to traverse each node, with maximum repetition will consequently shows maximum dissimilarity towards its surrounding nodes with the facts

$$w_2((i', j');(l', m')) = A_{Norm}(l', m').F(i', l'; j'-m')$$

Following the procedure of step *2* by normalizing outward bound edges weight of each node to unity \& treating graph as markov chain. The number of iterations *k ε [2,3,4]* to improve performance. Mass will have low consideration for the nodes with high activation.

3. Methodology

Our proposed salient object extraction technique is considered to be an amalgamation of two major image segmentation techniques, in addition to some preprocessing step, conversion of *RGB* space to *HSV* color space. Figure.1 shows complete methodology, with the instigation of image selection with at least one salient object to the extraction of that prime image object. As a preprocessing step, *RGB* image was converted to *HSV* colorspace in an attempt to be more intuitive and perceptually relevant.

As mentioned in section. 2.1, expectation maximization is immensely dependent on initial conditions for convergence, to coup with stated problem, multi-dimensional histogram and minimum distance technique were applied with the selection of high density bin. The resultant *EM* segmented image was visually precise with the reduction of color labels of the image. The texture details of an object may lead to the fusion of multiple clusters within the salient object, mentioned in Figure. 1.

To overcome such predicament, hierarchical structure were utilized which has the advantage of substantially reducing the computation requirement and partitioning of an image on the basis of principal eigenvector with an optimal covariance selection criteria on each iterative step. The variance is normally reduced when the node splits in the direction of principal eigenvector, so it is reasonable to make an assumption, total square error should be proportional to total squared variation along the principle eigenvector direction and salient object is selected with optimal covariance. The node splitting

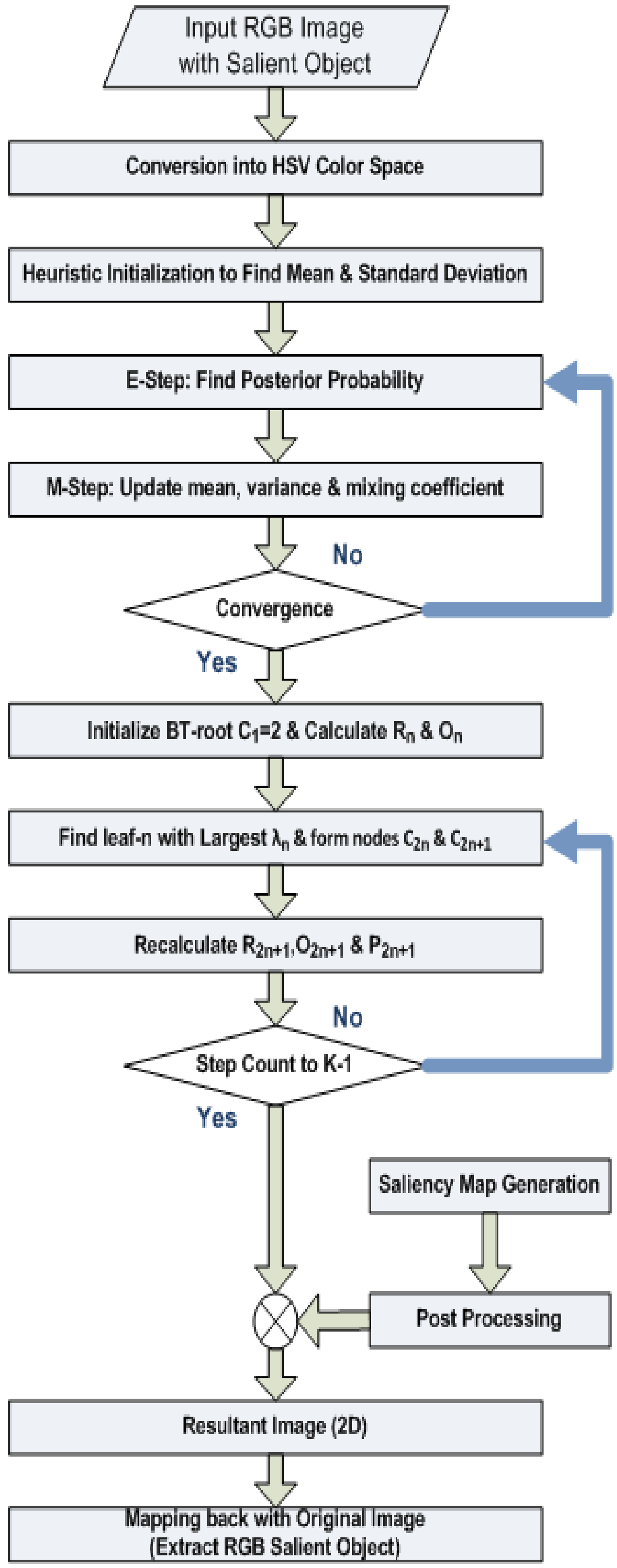


Fig. 6: Flow Diagram of Proposed Object Extraction Methodology

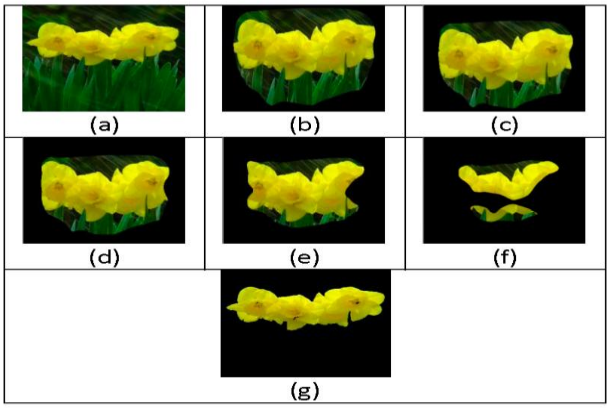


Fig. 7: Saliency Computation (a) Original Image (b) Threshold selected [*0.1*] (c) Threshold [*0.20*] (d)Threshold [*0.25*] (e) Threshold [*0.30*] (f) Threshold [*0.40*] (g) Proposed Methodology

orders were selected with the objective of reduction in total square error, *TSE*.

The high contrast saliency map was computed to get the impression of salient object which further fused to get the Pronounced object. Multiple fragments were detected in image belong to salient object cluster which were diminished in the fusion step with object extracted with hierarchical algorithm, visually depicted in Figure. (1,4), On the basis of threshold, opted in the range of ($0.01 \rightarrow .1$), tested for optimal results, image were converted to binary for further subjection to image fusion process as shown in Figure. 1.

All irrelevant information was expunged from the image using fusion procedure as shown in Figure. (8, 9). Detailed stepping of our proposed methodology is demonstrated in Figure. 6.

4. Results And Analysis

In order to subjectively evaluate the performance of our proposed methodology, set of images were selected from Berkeley segmentation dataset, can be seen at (*http://www.elib.cs.berkeley.edu/segmentation*), Coral database, (*http://www.wang.ist.psu.edu/docs/related/*) and our own images with the perspective of finding at least one obvious salient object in test images. Analysis was made based on image segmentation from the perspective of salient object extraction. From several images, fewer were selected for visual analysis in this article as shown in the Figure. (4, 8, 9).

The current approach mainly relies on expectation maximization algorithm for grouping and quantization, which is typical mixture density estimate method to classify interclass difference. The performance of our proposed approach is little discriminatory in regions with diminutive color difference, in respect of, adequate color variations is exceptional. In contrary to fact that salient object with high color variations were grouped into different clusters, which was overcome by using hierarchical approach of binary tree partitioning on the basis of Covariance and principle eigenvector. Exhaustive steps of our technique can be observed visually in Figure (1, 4), where each step of an algorithm was mentioned. Column 1 & 2 shows original picture and its *HSV* conversion where as column *3* shows *EM* segment of proposed *HSV* image and extraction of salient object in addition with fragments as a noise using BPT.

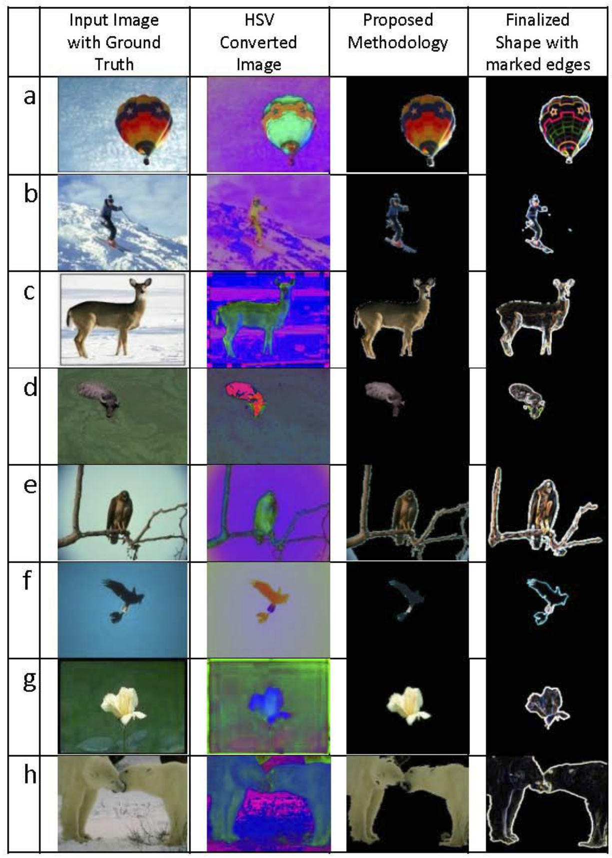


Fig. 8(a): Example of Images with Pronounced Salient Object Extraction. Column 4 shows filtered image using 'Sobel' with edge detection

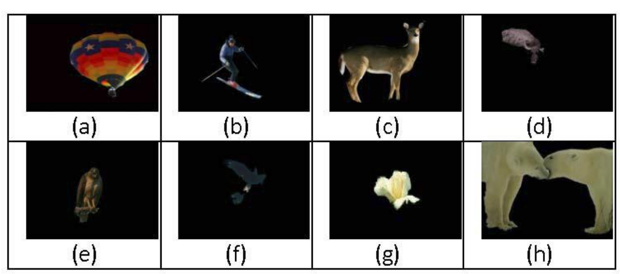


Fig8(b): Idealized image sample (Ground Truth)

High Contrast saliency Map is demonstrated in column *4*. The fusion state is visually shown in column *5* with the merging of saliency map and (*EM+BPT*) strategy. Final refined image object in column 6 where as manually segmented ground truth images are shown in column *7*. Analyzing Figure. 4 in depth revealed us few prominent misclassification, e.g., image (*3*) shows diminishing of right bottom of jar, similarly image 4 and 5 shows background fragment attached, the high resemblance of foreground salient object color with background color results in same group. Despite of mentioned diminutive errors, our proposed method worked well in comparison to the other techniques.

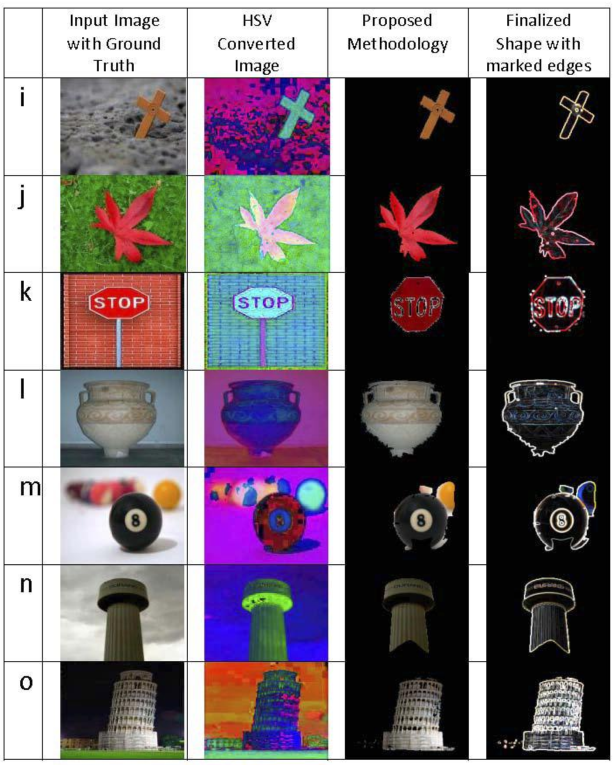


Fig. 9(a): Few More examples of Images with Pronounced salient object extracted.

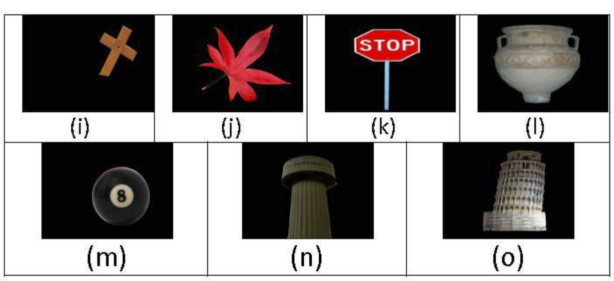


Fig. 9(b): Ground Truth Images

Although exact segmentation of color images is still a complicated problem but object boundaries can be identified with our strategy in a ways more similar to human perception. Segmentation results generated using our approach for test images from different categories such as flowers, signs, buildings, natural views, demonstrated in Figure 8*(a)* & 9*(a)* with ground truth or manually segmented image in Figure 8*(b)* and 9*(b)*. Column (3-4*)* shows the proposed procedure results and finalized edges. To highlight small fragments present with the object, we have applied sobel filter for edge detection. The proposed image contrast algorithm is presented in as pseudo code example and results can be visually analyzed in Figure. 10, where selected threshold was selected to be *0.45* for optimal results.

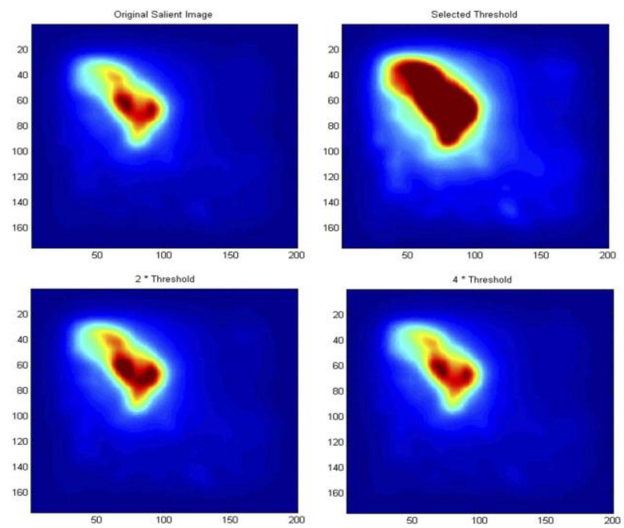


Fig. 10 Contrast Enhancement (a) Original Saliency Map (b) Proposed threshold selected, Image enhancement (c) Bifolded amplification in threshold (d) Fourfolded amplification in threshold.

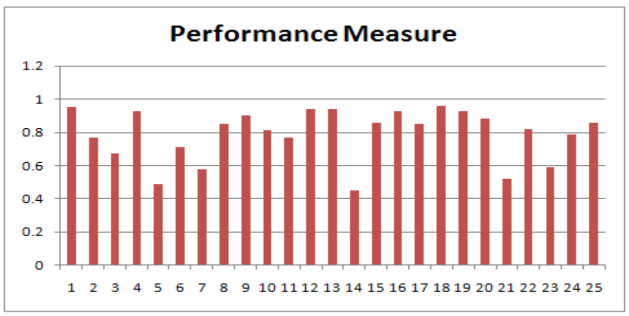
4.2 Pseudocode Example

```
Load salient map image and rescale in the range
of (0 → 1) and assign thresh = 0.45;
rscaled = rscaled − min(min(rscaled));
rscaled = rscaled./max(max(rscaled));
maxval = max(max(rscaled));
minval = min(min(rscaled));
con thresh = thresh ∗ (maxval − minval) + minval;
con threshtwo = 2 ∗ con thresh;
con threshthr = 4 ∗ con thresh;
fimage = min(rscaled, con thresh);
fimagetwo = min(rscaled, con threshtwo);
fimagethree = min(rscaled, con threshthr);
```

In order to objectively evaluate the segmentation performance from the perspective of salient object extraction, we categorize salient object under two groups [45], one under the assumption, manual segmentation produced is the ideal scenario, acknowledged as a ground truth and other group with our proposed object extraction procedure. The image segmentation method only consists of two segments, one for background and one for foreground. To measure the segmentation performance, it is consider to be concurrence among ground truth and the segmentation results. Let us consider S_{GT} is the foreground structure of Ground truth and S_{AS} is the foreground structure of algorithm segmented image. The region based segmentation accuracy will be defined as

$$P(S_{AS}, S_{GT}) = \frac{|S_{AS} \cap S_{GT}|}{|S_{AS} \cup S_{GT}|} = \frac{|S_{AS} \cap S_{GT}|}{|S_{AS}| + |S_{GS}| - |S_{AS} \cap S_{GT}|}$$

Where | . | is the operation of computing the segment area. The graph order follows the image sequences as in Figure. (1, 4, 7, 8, 9) respectively, having total number of *25* test images.



5. Conclusion and Future Work

In this paper, we have presented an efficient image segmentation approach aimed at salient object extraction, based on self initialized expectation maximization and BPT analysis in fusion with saliency map calculation. Subsequent to preprocessing step of *RGB* to *HSV* conversion, image was segmented with *EM* model of Gaussian mixture with self initialized mean and variance to quantize the image into segments of different clusters. Binary partition tree then extricate the salient object on the basis of principle eigenvector, covariance. The amalgamation of described approach with saliency image results in accurate salient object shape. Although the hybrid segmentation technique worked well when tested with several images, but manual selection of salient object from multiple partitions after *BPT* acts as a snarl, which will be undertaken as a future work. The accurate image extractions leads to precise object shapes which will be used as a starting step for Contents based image retrieval.

6. Acknowledgement

This work is supported by Chongqing Science Foundation, grant number (CSTC2011AB6054). We are also thankful to authors of Curve Evolution approach for segmentation using Low depth field to make code available online.

References

[1]. A.Meye-Base,A.Saalbach,O.Lange,A.Wismuller. "Unsupervised Clustering of FMRI and MRI time series." *Biomedical Signal Processing And Control* 2, no. 4,2007: 295-310.

[2]. A.Borji. "Modeling the influence of action on spatial attention in visual interactive environments". *IEEE International Conference on Robotics and Automation (ICRA),* (2012), 444-450

[3]. A.Rosenfeld, A.C.Kak. *Digital Image Processing.* 2nd . Vol. 1. Academic Press, 1982.

[4]. A.S.Frisch. "Unsupervised construction of fuzzy measures through self-organizing feature maps and its application in color image segmentation." *International Journal of Approximate Reasoning* 41, no. 1 (2006): 23-42.

[5]. C.Koch,E.Niebur. "A model of saliency-based visual attention for rapid scene analysis".*IEEE Transaction on Pattern Analysis and Machine learning*, 20, no 11 (1998), 1254-1259.

[6]. C.Biernachi, G.G.Celeux. "Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian Mixture Models." *Computational Statistics & Data Analysis* 41,2003: 561-575.

[7]. C.Zhang, T.S. Chen. "A survey on Image-based rendering-representation, sampling and Compression." *Signal Processing : Image Communication* 19, no. 1 ,2004: 1-28.

[8]. D.M.Ramik, C.Sabourine, and K.Madani. "Hybrid Salient Object Extraction Approach with Automatic Estimation of Visual Attention Scale ." *Signal Image Technology and Internet Based Systems.* France: IEEE, 2011. 438-445.

[9]. D.B.Burdescu, M.Brezovan, E.Ganea, L.Stanescu," A new method for segmentation of images represented in a HSV color Space", *Lecture Notes on Computer Science, Advanced Concept for Intelligent Vision Systems*, 2009, no 5807, 606-617.

[10]. F.Perazzi, P.Krahenbuhl, Y.Pritch, and A.Hornung. "Saliency filters: Contrast based filtering for salient region detection ." *Computer vision and Pattern Recognition.* IEEE, 2012. 733-740.

[11]. Gonzales, Woods. *Digital Image Processing.* 3. Prentice Hall, 2008.

[12]. I.Dagher, K.E.Tom. "WaterBalloons: A hybrid watershed Balloon Snake segmentation." *Image And Vision Computing* 26, no. 7 (2008): 905-912.

[13]. J.Zhang, L.Zhou, L.S.Lan. "Regions of Interest extraction based on visual attention model and watershed

segmentation ." *Neural Networks and Signal Processing.* IEEE, 2008. 375-378.

[14]. K.S.Fu, J.K.Mui. "A survey on image segmentation." *Pattern Recognition* 13, no. 1 (1981): 3-16.

[15]. J.Harel, C.Koch, P.Perona. "Graph-based visual saliency.", *Advances in Neural Information Processing Systems*,(2007):545-552, MIT Press

[16]. L.Itti, C.Koch, and E.Niebur. "A model of saliency-based visual attention for rapid scene analysis." *IEEE Transaction on Pattern Analysis and Machine Learning* 20, no. 11 (1998): 1254-1259.

[17]. L.Shao, and M.Brady. "Specific object retrieval based on salient regions." *Pattern Recognition* 39, no. 10 (2006): 1932-1948.

[18]. L.Tang, H.L.Li,T.Chen. "Extract salient objects from natural images." *Intelligent Signal Processing and Communication Systems.* IEEE, 2010. 1-4.

[19]. M.A.Sutton, J.C.Bezedek,T.C.Cahoon. "6-Image Segmentaton by Fuzzy Clustering: Methods And Issues." *Handbook of Medical Imaging*, 2000: 87-106.

[20]. M.S.Yang, U.J.Hu,K.C.Ren,C.L.Lin. "Segmentation techniques for tissue differentiation in MRI of Ophthalmology using Fuzzy clustering Algorithm." *Magnetic Resonance Imaging* 20, no.2, 2002): 173-179.

[21]. M.T.Orchard. "Color Quantization of Images." *IEEE Transaction on Signal Processing* 39, no. 12 (1991): 2677-2690.

[22]. N.Bruce, and J.Tsotsos. "Saliency based on Information maximization." *Advances in Neural Information Processing Systems* 18 (2006): 155-162.

[23]. O.H.Odukoya, G.A.Aderounmu, E.R.Adagenodo. "An Improved Data Clustering Algorithm for Mining Web Documents." *International Conference on Computational Intelligence And Software Engineering.* IEEE, 2010. 1-8.

[24]. Q.Zhang, Izquierdo, and E.Ebroul. "Adaptive Salient Block Based Image Retrieval in Multi-Feature Space ." *Content based Multimedia Indexing.* IEEE, 2007. 106-113.

[25]. R.L.Breiger. "An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling." *Journal of MAthematical Psychology* 12, no. 3 (1975): 328-383.

[26]. R.L.Prentice. "A design for studying the clustering of plant or animal species using quadrat sizes in geometric progression." *Journal of Theoratical Biology* 39, no. 3 (June 1973): 601-608.

[27]. R.Palm, B.lliev,B.Kadmiry. "Recognition of human grasps by time-clustering and fuzzy modeling." *Robotis And Autonomous Systems* 57, no. 5 (2009): 484-495.

[28]. S.D.Olabarriaga, A.W.M Smeulders. "Interaction in the Segmentation of Medical Images: A Survey." *Medical Image Analysis* 5, no. 2 (2001): 127-142.

[29]. S.Feng, X.De, and X.Yang. "Attention-driven salient edge(s) and region(s) extraction with application to CBIR." *Signal Processing* 90, no. 1 (2010): 1-15.

[30]. S.Kazadi, A.A.khaliq, R.Goodman. "On the convergence of puck clustering systems." *Robotics And Autonomous Systems* 38, no. 2 (2002): 93-117.

[31]. S.L.Hartmann, Galloway.R.L. "Depth-buffer targeting for spatially accurate 3-D visualization of medical images." (IEEE Transaction on Medical Imaging) 19, no. 10 (2000): 1024-1031 .

[32]. S.R.Bulo, M.Rabbi, and M.Pelillo. "Content-based image retrieval with relevance feedback using random walks." *Pattern Recognition* 44, no. 9 (2011): 2109-2122.

[33]. S.Sural, G.Qian, S.Pramanik. "Segmentation and Histogram Generation Using the HSV Color Space for Image Retrieval." *International Conference on Image Processing.* 2002. 589-592.

[34]. T.Lei, W.Sewchand. "Object detection and Recognition via stochastic model-based image segmentation." *Multidimensional signal processing workshop.* IEEE , 1989. 17-18.

[35]. W.K.Pratt. *Digital Image Processing.* 2nd. New York: Wiley, 1991.

[36]. X.Shen, and Y.Wu. "A unified approach to salient object detection via low rank matrix recovery ." *Computer Vision and Pattern Recognition.* IEEE, 2012. 853-860.

[37]. X.Yang, and S.M.Krishnan. "Image segmentation using finite mixtures and spatial information." *Image and Vision Computing* 22, no. 9 (2004): 735-745.

[38]. Y.H.Lu, X.H.Zhang, J.Kong, X.F.Wang, and J.B.Zhang. "A Novel Objects of Interest Extraction Approach Using Attention-Driven Model for Content-Based Image Retrieval ." *Image and Signal Processing.* IEEE, 2008. 339-343.

[39]. Y.H.Tsai. "Hierarchical Salient Point Selection for image retrieval." *Pattern Recognition Letters* 33, no. 12 (2012): 1587-1593.

[40]. M.Miyahara, Y.Yoshida,"Mathematical Transform of (r,g,b) Colour Data to Munsell (h,s,v) Colour Data", *SPIE Visual Communications and Image Processing*, 1988, 650-657.

[41]. Yixin Chen, J.X.Wang. "A region based fuzzy feature matching approach to content-based image retrieval." *IEEE Trans* 24, no. 9 (2002): 1252-1267.

[42]. Z.Han, Z.Q.Liu, Z.Y.Zhang, Y.Lu, and W.Li. "Salient Object Extraction Based on Region Saliency Ratio ." *Computer and INformation Science.* IEEE, 2009. 611-615.

[43]. Z.J.Yuan,J.Sun, J.Wang, N.N.Zheng, X.Tang, H.Y.Shum. "Learning to Detect a Salient Object." *Pattern Analysis and Machine Intelligence* 33, no. 2 (2011): 353-367.

[44]. Z.Liu, L.Shen,Z.Zhang. "Unsupervised image segmentation based on analysis of binary partition tree for salient object extraction." *Signal Processing* 91, no. 2 (2011): 290-299.

[45]. F.Ge, S.Wang, T.Liu. "Image-Segmentation Evaluation From the Perspective of Salient Object Extraction". *IEEE, Conference on Computer Vision and Pattern Recognition*}.2006. 1146-1153.