



A Review of an Enhanced Algorithm for Color Image Segmentation

Sumant V. Joshi *

Department of Electronics and Telecommunication,
Dr. Bhausaheb Nandurkar C. o. E. T., Yavatmal

Atul. N. Shire

Department of Electronics and Telecommunication,
Dr. Bhausaheb Nandurkar C. o. E. T., Yavatmal

Abstract— Image segmentation is very essential and critical to image processing and pattern recognition. This paper presents a summary of color image segmentation techniques. Basically, color segmentation algorithms are based on monochrome segmentation approaches operating in different color spaces. The segmentation algorithms are categorized as edge based, region based and graph based segmentation. In this paper, various segmentation techniques have been summarized with emphasis on the Mean Shift (MS) segmentation and the Normalized Cut (Ncut) partitioning methods as an integrated approach.

Keywords— Image segmentation, Image Processing, Mean Shift, Normalized Cut, Clustering

I. INTRODUCTION

Segmentation is the very first step in almost all the image processing application where the properties of objects in image need to be analyzed e.g. in medical imaging problems in automotive vision in vehicle detection; object recognition in content based image retrieval etc. The objective of the image segmentation is to extract the dominant colors. The image segmentation is very important to simplify an information extraction from images, such as color, texture, shape, and structure. The applications of image segmentation are diversely in many fields such as image compression, image retrieval, object detection, image enhancement, and medical image processing. Several approaches have been already introduced for image segmentation. A formal definition of image segmentation is as follows: If $P()$ is a homogeneity predicate defined on groups of connected pixels, then segmentation is a partition of the set F into connected subsets or regions (S_1, S_2, \dots, S_n) such that $\bigcup_{i=1}^n S_i = F$ with $S_i \cap S_j = \emptyset$ ($i \neq j$). The uniformity predicate $P(S_i) = \text{true}$ for all regions, S_i and $P(S_i \cup S_j) = \text{false}$, when $i \neq j$ and S_i and S_j are neighbours[1].

In computer vision, segmentation refers to the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels) such that each segment is homogeneous and the union of no two adjacent segments is homogeneous. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyse. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristics. Image segmentation is a key step in many applications in pattern recognition, computer vision and image understanding to allow further image content exploitation in an efficient way. According to several authors, segmentation terminates when the observer's goal is satisfied but still there needs to develop a unique method for it. There are many algorithms and methods are available for segmentation. The classification of existing algorithms for image segmentation can be categorized as i.e., feature-space-based segmentation, spatial segmentation, and graph-based segmentation. In feature-space-based clustering approaches, [2], [3] the global characteristics of the image captured through the selection and the image features are calculated, normally based on the color or texture. With the use of a particular distance measuring technique not considering the spatial information, the feature samples are handled as vectors, and the effort is to group them into compact, but well-separated clusters. The feature-space-based clustering approaches have some severe shortcomings. There is no provision for preservation of the spatial structure and the detailed edge information of an image and if feature spaces of pixels overlap, then pixels from disconnected regions of the image may be clustered in one group. In order to keep the edge information intact alongwith the spatial relationship between the pixels on the image plane, the images are to be handled in spatial domain [4]. The watershed algorithm [5] is an extensively used technique for this purpose. A very large number of small but quasi-homogenous regions are produced. Hence, some merging algorithm need to be applied to these regions [4]-[6]. Graph-based approaches are based on the fusion of the feature and spatial information and referred as image perceptual grouping and organization methods. In such approaches, several key factors such as similarity, proximity, and continuation are decisive to visual group. Generally, in these approaches, there is the formation of a weighted graph, where each vertex corresponds to n image pixel or a region, and the weight of each edge connecting two pixels or two regions represents the likelihood that they belong to the same segment. The color and

texture features, as well as the spatial characteristic of the corresponding pixels or regions decide the weights. There is a partitioning of the graph into multiple components for the minimization of some cost function of the vertices. Until now, several graph cut-based methods have been developed for image segmentations [7], [8]. For example, Wang and Siskind [8] developed an image-partitioning approach by using a complicated graph reduction.

Shi and Malik [8] proposed a general image segmentation approach based on normalized cut (Ncut) by solving an eigen system. The normalized cut criterion measures both the total dissimilarity between the different groups as well as the total similarity within the groups. The Ncut method can robustly generate balanced clusters and shown that it outperformed the other spectral graph partitioning methods, such as average cut and average association [8]. However, Ncut based image segmentation approach in general, requires high computation complexity and, therefore, cannot be applied for real-time processing [8]. An efficient remedy to this concern is to derive the regions by some region segmentation method and then employ the graph representation strategy on these regions. In [6], the authors have developed an image segmentation method that integrates region based segmentation and graph-partitioning scheme. Firstly, a set of oversegmented regions from an image is produced by using the watershed algorithm, and then a graph structure is then applied to represent the relationship between these regions. To solve these problems, the a novel approach is proposed in this correspondence that provides effective and robust image segmentation with low computational complexity by incorporating the mean shift (MS) and the Ncut methods. In the above stated method, initially image region segmentation is performed by using the MS algorithm [9], and then treated these regions as nodes in the image plane and applies a graph structure to represent them. The final step is to apply the Ncut method to partition these regions.

II. SEGMENTATION TECHNIQUES

Various segmentation techniques has been discussed in the literature, all this techniques has the common aim of partitioning an image into different regions having homogeneous properties some of these along with mean shift clustering and normalized cut method are discussed as follows:

2.1 Histogram Thresholding :

An image histogram is a type of histogram that acts as a graphical representation of the intensity distribution in an image. It plots the number of pixels for image intensity values. By looking at the histogram of image observer will be able to give the intensity distribution of an entire image. Histogram thresholding is most uncomplicated image segmentation process. Since thresholding is fast and economical in computation. For segmenting background and objects, a threshold is selected which is depend on the brightness of an image is used if single threshold is defined for different elements in an image is known as local threshold whereas, if single thresholds is defined for complete image is known as global threshold[10]. For determining the thresholds, threshold recognition approaches were used. Threshold recognition approaches are optimal thresholding, p-tile thresholding and histogram shape analysis. Optimal thresholding is used when there is closest gray level corresponding to maximum two or more intensity distributions present in an image. However this segmentation is less error prone. For color or multi band images multi-spectral thresholding is suitable. Mathematically, a histogram is a function z_i that counts the number of pixels that fall into each of the different intensity levels Thus, let a be the total number of pixels and b be the total number of image intensities, the histogram z_i meets the following conditions,

$$a = \sum_{i=1}^b z_i$$

If a Single threshold is defined for complete image is known as global thresholding. Fig.3.1 shows the histogram corresponds to an image $f(x,y)$, which consist of light object on a dark background and it is considered that object and background pixels intensity values are grouped into two dominant modes. Then by selecting the suitable threshold Thr , the object can is extracted from the background such that point (x, y) by in the image where $f(x, y) > Thr$ determines the object point otherwise indicates background point. Mathematically, segmented image $P(x, y)$ is given by.

$$P(x, y) = \begin{cases} 1, & \text{if } f(x, y) > Thr \\ 0, & \text{if } f(x, y) < Thr \end{cases}$$

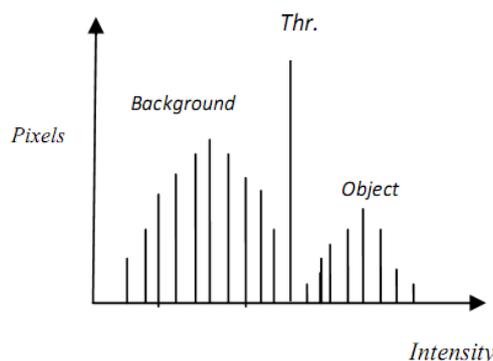


Fig.3.1 Intensity histogram that can be partitioned by single threshold to separate object

2.2 Region Growing :

As discussed in previous section the main objective of segmentation is to partition an image into regions. Segmentation is done with threshold based on distribution of the intensities of the pixels in an image. Region growing is a technique of segmentation in which pixels with similar intensities are grouped in order to find the regions directly. [10] This group of pixels belonging to the region of focus is known as seeds. The similarity criteria of the pixels depend not only on the problem under consideration, but also on the type of image data available.

The algorithm of region growing technique can be states as follows.[11].

1. In the first step pixel or group of pixels which belongs to the region of interest called seeds are formed.
2. In the next step pixels in the region of interest are examined and added to the growing region in accordance with the homogeneity criteria. Until no more pixels can be adjoined to the growing regions, this step continues.
3. And in last step the object illustration is done by all added pixels to the growing regions.

The key advantage of region growing technique is these methods can correctly separate the regions that have the same properties we define. One of the drawbacks of this method is, noise or variation of intensity may result in over segmentation.

2.3 Edge Based Segmentation :

Edge detection approach is most frequently used for segmenting images based on local changes in intensity. In this method boundary or edge on an image is defined by the local pixel intensity gradient. The calculation of first order derivative of the image function is called a gradient. The gradient is denoted by $\nabla f(x,y)$. The magnitude of gradient for a given image $f(x,y)$ can be calculated as,

The three fundamental steps used for edge detection are given as:

1. In the first step image smoothing for reduction of noise is done.
2. Detection of edge points are the second step in which there is extraction of all the points in an image that have intensities to become edge points.
3. The last step of this method is Edge localization in this step there is a selection among the edge points that were selected in the previous step that are true members of the set of points that forms an edge.

Closed region boundaries are required to segment an object from an image. The desired edges are the boundaries between such objects. One of the author developed a method for edge detection of weak edges in image by canny algorithm, by labelling all the 8-connected edge with the different number and then they classify with that edge, for the size of all 8-connected edge circumference being different and then based on this information they plot the histogram according to the size of edge and lastly weak edge of the brain is detected by histogram segmentation.[10][11]

III. GRAPH CUT TO IMAGE SEGMENTATION

This promising technique based on pairwise objects grouping has emerged recently in the machine learning community. The choice of graph cut depends on the ability to segment image through spectral analysis. This grouping procedure can be associated as a pixel graph or characteristic points where the graph nodes are pixels and the graph edges are connected neighbouring pixels to form discrete connections.

3.1 Spectral clustering analysis :

Given a set of points representing image characteristics that we want to cluster in subsets, we can build a weighted graph $G = (V, E, W)$ where V and E are graph nodes and graph edges, and W is the weight which measures the likelihood of neighboring pixels being the same clustering. In image processing, the lattice graph is obtained by computing the score provided by the product of location pixels and their level gray,

$$W_{i,j} = \exp - \frac{\|I_i - I_j\|^2}{2\sigma_I^2} \left\{ \begin{array}{l} \exp - \frac{\|X_i - X_j\|^2}{2\sigma_X^2} \text{ if } d(X_i, X_j) \leq R \\ 0 \text{ elsewhere} \end{array} \right\}$$

where I_i is the grey level at pixel i , X_i is the pixel position of pixel I in image, σ_I is the scale factor controlling the affinity between pixels and σ_X is the coefficient measuring the affinity between position seeds, $d(\cdot)$ is the Euclidian distance between two pixels and an a prior parameter R fixed to threshold the distance from a node i to j . In grouping set of vertices are partitioned into disjoint set V_1, V_2, \dots, V_m where the similarity between the vertices in V_i is high and across different sets is low i.e. between $(V_i$ and $V_j)$. While partitioning the graph two questions arises, first what should be the precise criterion for good partition and how can such partition can be computed. Segmentation criterion used in most of them is based on local properties of the graph. Because perceptual grouping is about extracting the global impressions of the image or scene. As this criterion for partitioning is often falls short of this main goal therefore the Ncut criterion is used[12].

3.2 Normalized Cut Algorithm :

Normalized cut is a global criterion for partitioning the graph used for segmentation of image. Normalized cut criterion measures both the total dissimilarity and similarity between different groups. A graph $G = (V, E)$ can be partitioned into two disjoint sets, $A, B, A \cup B = V, A \cap B = \emptyset$ by simply removing edges connecting the two parts. The degree of dissimilarity been removed. In graphical language, it is called the cut.

$$cut(A, B) = \sum_{\mu \in A, \nu \in B} W(\mu, \nu)$$

The successful bi-partitioning of a graph is done when it minimizes this cut value. Although there are a various number of such partitions, in the past lot of work was done for finding the minimum cut of a graph. Wu and Leahy [11] proposed a clustering method based on this minimum cut criterion. They partition a graph into k-sub graphs such that the maximum cut across the subgroups is minimized. By finding the minimum cut this problem can be efficiently solved by them. However the minimum cut criteria favors cutting small sets of individual nodes in the graph. Fig. 3.3 illustrates one such case

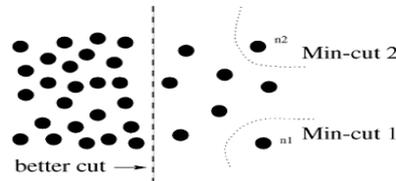


Fig. 3.3 An Example Illustrating that Minimum Cut gives a Bad Partition.

From figure 3.3 any cut those partitions out individual nodes on the right half will definitely have smaller cut value than the cut that partitions the nodes into the left and right halves. To avoid this problem of partitioning small sets of points, [8] propose a new measure of finding cut between two groups. Instead of looking at the value of total edge weight connecting the two partitions, they compute the cut cost as a fraction of the total edge connections to all the nodes in the graph and call this disassociation measure the normalized cut (Ncut).[8]

$$Ncut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)}$$

Where $assoc(A, V) = \sum_{\mu \in A, t \in V} W(\mu, t)$ is a total connection from nodes A to all nodes in the graph and $assoc(B, V)$ is similarly defined. However if there are more pixels in the image, more graph of nodes will be generated and this will cause difficulties to solve this algorithm.[8] .

IV. CLUSTERING

A process of organizing the objects into groups based on its attributes is known as clustering. A cluster is therefore a collection of objects which are “similar” between them and are “dissimilar” to the objects belonging to other clusters. An image can be grouped based on keyword (metadata) or its content (description). In keyword based clustering, a keyword is a form of font which describes about the image keyword of an image refers to its different features. The similar featured images are grouped to form a cluster by assigning value to each feature. In content based clustering [13] content refers to shapes, textures or any other information that can be inherited from the image itself. The tools, techniques and algorithms that are used originate from fields such as statistics, pattern recognition, signal processing etc. Clustering based on the optimization of an overall measure is a fundamental approach explored since the early days of pattern recognition. The most popular method for pattern recognition is K-means clustering. In K-means clustering a centroid vector is computed for every cluster. The centroid must be chosen such that it should minimize the total distance within the clusters. Both supervised and unsupervised clustering techniques are used in image segmentation. In supervised clustering method, grouping is done according to user feedback. In unsupervised clustering, the images with high features similarities to the query may be very different in terms of semantics [13]. This is known as semantic gap. A variety of clustering techniques have been introduced to make the segmentation more effective.

4.1 Kernel density estimation :

Kernel density estimation is a non-parametric way to estimate the probability density function of a random variable. Kernel density estimation is a fundamental data smoothing problem where inferences about the population are made, based on a finite data sample. Let (x_1, x_2, \dots, x_n) be an any sample drawn from some distribution with an unknown density f . We are interested in estimating the shape of this function f . Its kernel density estimator is

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i)$$

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right)$$

Where $K(\cdot)$ is the kernel — a symmetric but not necessarily positive function that integrates to one and $h > 0$ is a smoothing parameter called the bandwidth. A Kernel with subscript h is called the scaled kernel and defined as $K_h(x) = 1/h K(x/h)$, the normal kernel is often used $K(x) = \phi(x)$, where Φ is the standard normal density function. [13]

4.2 Mean shift clustering:

Mean shift clustering is non-parametric clustering technique which does not require prior knowledge of the clusters. Mean shift algorithm clusters an n-dimensional data set. For each point mean shift computes its associated peak by first defining a spherical window at the data point of radius r and computing the mean of points that lie within the window.

Algorithm then shifts the window to the mean and repeats until convergence. At each iteration the window will shift to a more densely populated portion of data set until peak is reached where data is equally distributed.

The procedure of situating maximum value of density function among the given isolated data sampled from that function. This is an iterative method, starting with an initial estimate d . Let $K(d_i - d)$ be a given Kernel function. This function determines the weight of nearby points for re-calculation of the mean. Typically Gaussian kernel is, used on the distance to the current estimate, $K(d_i - d) = e^{-\|d_i - d\|}$. The weighted mean of the density in the window determined by K is

$$m(d) = \frac{\sum_{d_i \in N(d)} K(d_i - d) d_i}{\sum_{d_i \in N(d)} K(d_i - d)}$$

Where $N(d)$ is the neighbourhood of x , a set of points for which $K(d) \neq 0$. The mean-shift algorithm now sets, $d_n \leftarrow m(d)$ and repeats the estimation until $m(d)$ converges.

An example illustrating mean shift procedure is shown in Fig.3.4 The shaded and black dots denote the data points of an image and successive window centres respectively. Mean shift procedure starts at Point Y_1 , by defining spherical window of radius r around it, algorithm then calculates the mean of data points that lie within the window and shifts the window to the mean and iterates the same procedure until peak is reached. At each iteration window is shifted to the more densely populated region.[8][14]

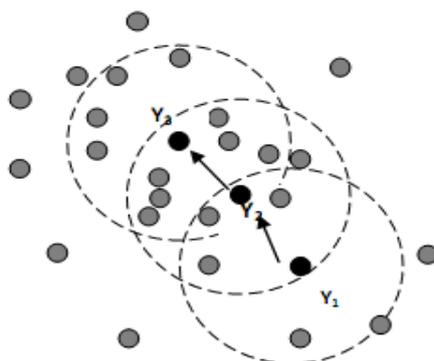


Fig.3.4 Mean shift procedure

TABLE I
COMPARISON OF IVARIOUS SEGMENTATION TECHNIQUES

| Segmentation technique | Method description | Advantages | Disadvantages |
|---------------------------|--|---|--|
| Histogram thresholding | Requires that the histogram of an image has a number of peaks, each corresponds to a region | It does not need prior information of the image. For a wide class of images satisfying the requirement, this method works very well with low computation complexity | (1)Does not work well for an image without any obvious peaks or with broad and flat valleys (2)Does not consider the spatial details, so cannot guarantee that the segmented regions are contiguous |
| Region-based approaches | Group pixels into homogeneous regions. Including region growing, region splitting, region their combination | Work best when the region homogeneity criterion is easy to define. They are also more noise immune than edge detection approach | (1)Are by nature sequential and quite expensive both in computational time and memory (2)Region growing has inherent dependence on the selection of seed region and the order in which pixels and regions are examined (3)The resulting segments by region splitting appear too square due to the splitting scheme |
| Edge detection approaches | Based on the detection of discontinuity, normally tries to locate points with more or less abrupt changes in gray level. Usually classified into two categories: sequential and parallel | Edge detecting technique is the way in which human perceives objects and works well for images having good contrast between regions | (1)Does not work well with images in which the edges are ill-defined or there are too many edges (2)It is not a trivial job to produce a closed curve or boundary (3)Less immune to noise than other techniques, e.g., thresholding and clustering |
| Clustering | Assumes that each region in the image forms a separate clustering the feature space. Can be generally broken into two steps:(1) categorize the points in the feature space into clusters; (2)map the clusters back to the spatial domain | Straightforward for classification and easy for implementation | (1)How to determine the number of clusters (known as cluster validity) (2)Features are often image dependent and how to select features so as to obtain satisfactory segmentation results remains unclear (3)Does not utilize spatial information |

| | | | |
|--|--------------------------|--|--|
| | to form separate regions | | |
|--|--------------------------|--|--|

V. Proposed Approach

The proposed algorithm integrates the two techniques for efficient and effective image segmentation with lower computationally. From a data-flow point of view, the outline of the proposed algorithm can be characterized as the following. First, an image is segmented into multiple separated regions using the MS algorithm. Second, the graph representation of these regions is constructed, and the dissimilarity measure between the regions is defined. Finally, a graph-partitioning algorithm based on the Ncut is employed to form the final segmentation map.

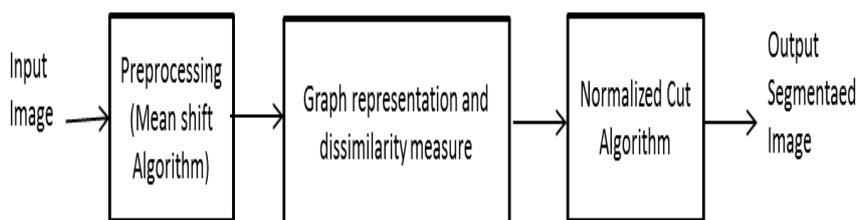


Fig. 5.1 Block diagram of proposed approach.

V. Conclusion

In this paper, various segmentation techniques categorized as The classification of existing algorithms for image segmentation can be categorized as i.e., feature-space-based segmentation, spatial segmentation, and graph-based segmentation alongwith edge detection, region growing and clustering have been reviewed and summarized alongwith their algorithms, merits and demerits. Among which it is observed that mean shift algorithm is an efficient method of clustering, which segments the colour image into multiple separated regions having homogeneous properties. However on applying Ncut method directly on color images takes more time for segmentation because of more graph nodes are generated which causes the difficulties to solve this algorithm. In this paper, a novel image segmentation algorithm has been proposed and it is based on the conventional mean shift algorithm and Ncut algorithm. The effectiveness of the proposed algorithm needs to be verified by extensive experiments.

REFERENCES

- [1]. H. D. Cheng, X. H. Jiang, Y. Sun, and J. Wang, "Color image segmentation: Advances and prospects," *Pattern Recognition.*, vol. 34, no. 12, pp. 2259–2281, Dec. 2001.
- [2]. D. W. Jacobs, D. Weinshall, and Y. Gdalyahu, "Classification with nonmetric distances: Image retrieval and class representation," *IEEETrans. PatternAnal.Mach.Intell.*, vol. 22, no. 6, pp. 583–600, Jun. 2000.
- [3]. A. K. Jain and D. Zongker, "Representation and recognition of handwritten digits using deformable templates," *IEEETrans.PatternAnal.Mach. Intell.*, vol. 19, no. 12, pp. 1386–1391, Dec. 1997.
- [4]. K. Haris, S. N. Efstratiadis, N. Maglaveras, and A. K. Katsaggelos, "Hybrid image segmentation using watersheds and fast region merging," *IEEE Trans.ImageProcess.*, vol. 7, no. 12, pp. 1684–1699, Dec. 1998.
- [5]. L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulation," *IEEETrans.PatternAnal.Mach. Intell.*, vol. 13, no. 6, pp. 583–597, Jun. 1991.
- [6]. S. Makrogiannis, G. Economou, and S. Fotopoulos, "A region dissimilarity relation that combines feature-space and spatial information for color image segmentation," *IEEETrans.Syst.,Man,Cybern.B,Cybern.*, vol. 35, no. 1, pp. 44–53, Feb. 2005.
- [7]. S. Wang and J. M. Siskind, "Image segmentation with ratio cut," *IEEE Trans.PatternAnal.Mach.Intell.*, vol. 25, no. 6, pp. 675–690, Jun. 2003.
- [8]. J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans.PatternAnal.Mach.Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.
- [9]. D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEETrans.PatternAnal.Mach.Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.
- [10]. R.C Gonzalez., R.E Woods and . S.L Eddins, '*Digital image processing using MATLAB*', Second edition, Gatesmark publishing, USA, 2009. (histogram)
- [11]. M. S. Khadem, '*MRI image segmentation using graph cut*' Master of Science Thesis in Communication Engineering chalmers university of technology, Sweden 2010
- [12]. S. X. Yu and J. Shi, "Multiclass spectral clustering," in *Proc.Int.Conf. Comput.Vis.*, 2003, pp. 313–319.
- [13]. S. Thilagamani and N. Shanthi, "A Survey on image segmentation through clustering" International journal of research and Information sciences Vol 1 No. 1, March 2011

- [14]. Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 8, pp. 790–799, Aug. 1995.