**Available Online at www.ijcsmc.com**

**International Journal of Computer Science and Mobile Computing**

**A Monthly Journal of Computer Science and Information Technology**

**ISSN 2320–088X**
**IMPACT FACTOR: 6.017**

***IJCSMC, Vol. 6, Issue. 3, March 2017, pg.01 – 08***

# Implementation of Color-based Image Segmentation by Clustering Methods

**Divya[1], Pawan Kumar Mishra[2]**

M.Tech Scholar, CSE Department, Uttarakhand Technical University, Dehradun, India[1]

Assistant Professor, CSE Department, Uttarakhand Technical University, Dehradun, India[2]

[1] divya.utu@gmail.com; [2] pawantechno@rediffmail.com

***Abstract— Digital image processing is the fastest growing computer-based technology. It plays a vital role in computer vision and image processing tasks and is generally used for various purposes like medical diagnosis, robotics, remote sensing, industrial inspection etc... Some major operations performed on an image to extract some useful information that corresponds to the key stages in the digital image processing. Digital image processing techniques help in the manipulation of the digital images. While processing the digital image the key phases include pre-processing, enhancement, and display information extraction. The primary objective is to automate multiple tasks together, and the process of image segmentation is one amongst them. The Clustering method is often used for segmenting large scale images for which sometimes pre-processing is required to reduce the volume of data and then other clustering approaches can be applied for better results.***

***This paper experiments the segmentation technique by using the clustering approaches at the same time and considering the parameters like color intensity, the number of clusters required to segment the image and on the basis of mean shift bandwidth. The purpose of this report aims at the implementation and comparative analysis of the results generated using the normalized cut, k means and the mean shift clustering technique.***

***Keywords— Image Segmentation, Normalized Cut, K Means, Mean Shift, Hierarchical Clustering***

## I. INTRODUCTION

Image segmentation is a process of assigning a label to each pixel in the image such that pixels with same labels share common visual characteristics which make an image easier to analyze in the image processing tasks. There are numerous image segmentation techniques available but the choice depends upon the level of segmentation to be used in the particular type of image and characteristics of the problem being considered. For segmentation of large-scale images, the clustering method is often used to perform pre-treatment of images. The overall objective of these methods is to divide images into small domains with the less complicated clustering algorithm. Now with the help of those data samples, some features are chosen for the second clustering till completion of the final partition of images. These kinds of methods reduce the volume of data for the second clustering through pre-processing and degrade the overall complexity of the algorithm and improve the efficiency. For segmentation of massive image data, those methods were found feasible.

*© 2017, IJCSMC All Rights Reserved* 

## II. IMAGE SEGMENTATION TECHNIQUES

Image segmentation involves a process of partitioning a digital image into N number of parts. The images are segmented on the basis of the set of pixels or pixels in a region. The pixels that are similar forms the basis of some congruity criteria such as color, intensity or texture and that helps to locate and identify the objects or boundaries in an image. Segmentation bridges the gap between the high-level image processing and high-level image processing. Edge detection, clustering approach, and region growing are the common techniques that are used for the segmentation of images.

Edge detection is the fundamental pre-processing step for image segmentation. It is used to change the input image to a binary image, which indicates either the presence or the absence of an edge. Edges are sets of pixels and thus are important contour features where dissimilar intensity changes or discontinuities in the corresponding image occur. In Digital image brightness characteristics of a pixel have an influence of background noise hence imaged pre-processing is important. Most noise in the image is discrete noise. The median filter can eliminate the discrete noise, inhibition of salt and pepper noise and overcome the image blur. The aim of noise reduction is to suppress the noise while preserving the important fine details and edges which are the boundary to different regions [1].

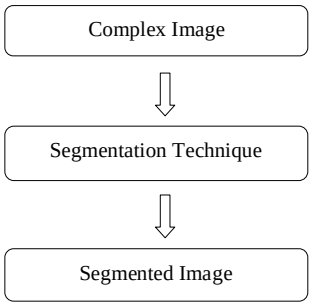



Fig. 1 Segmentation process

Segmentation technique basically converts the complex image into the simple image as shown in Fig. 1... Several segmentation techniques have already been developed like splitting, merging or a combination of both. Some of them are mentioned below.

- Edge-Based Segmentation
- Region-Based Segmentation
- Threshold Segmentation
- Fuzzy Theory based Segmentation
- Mean Shift Segmentation

The objective of image segmentation is to make the representation of the image as simple which can easily be understood and for that it is essential to discover the location of objects, boundaries, lines and so on in the digital images. More precisely, image segmentation is the process of assigning a label to every pixel in the image such that pixels with the same label share certain visual characteristics or features [2].

### *A. Edge-Based Segmentation*

In digital image processing and computer vision the edge based segmentation plays a very important role as in an image, the edge represents object boundaries and thus helps in detection and segmentation of objects in an image.
Edge based technique detected edges in an image are assumed to represent object boundaries and used to identify these objects [3]. This method can be preferable for two major reasons are: Algorithms are usually less complex than others and edges are more important features in an image to distinguish their regions as shown in Fig. 2.

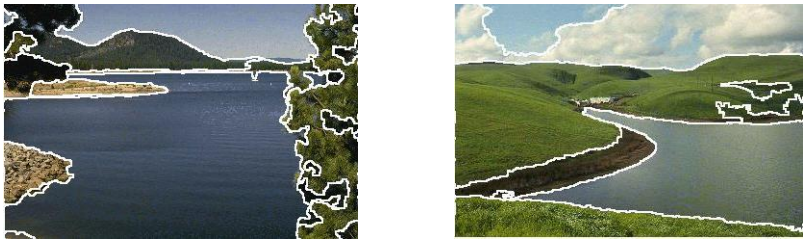


Fig. 2 Edge-based segmentation

*© 2017, IJCSMC All Rights Reserved* 

The edge detection algorithms are suitable for images that are simple and noise free.

*B. Region-Based Segmentation*

Region based segmentation algorithms segment an image into regions that are similar according to a set of predefined criteria [4] and are based on continuity. Region-based segmentation has Region Growing phase, Region Splitting and Merging phase to segment the image.

Region-based methods focus attention on an important aspect of the segmentation process missed with point-based techniques. There a pixel is classified as an object pixel judging solely on its gray value independently of the context. This meant that isolated points or small areas could be classified as object pixels, disregarding the fact that an important characteristic of an object is its connectivity [1]. The ultimate goal of any of the image segmentation technique is to simplify the representation of the images.

Region growing main aim is to partition an image into different regions that are similar and are according to a set of a pre-specified feature as shown in Fig. 3. It requires a set of starting pixels called seeds. Region-based segmentation is a technique that allows us to determine the regions directly.

Fig. 3 Region-Based Segmentation

These techniques are applied in noisy images where it is not easy to detect the edges. These techniques are usually time consuming because it requires lots of computation time.

*C. Threshold Based Segmentation*

Threshold technique is a widely used technique for image segmentation and is the simplest and commonly used technique. For segmentation Histogram thresholding and slicing technique are used. One can apply it directly to an image and can also combine it with pre-processing and post-processing techniques. This technique is based upon a simple concept.

A parameter Θ called the brightness threshold is chosen and applied to the image a[x, y] as follows:
If $a[x, y] \geq \Theta$ then a[x, y] =object=1
Else a[x, y] = background = 0

This technique is appropriate and powerful when the segmentation is to be done in a light object with a dark background [5]. But, when segmentation is to be applied in dark object on a light background then use it as
If $a[x, y] < \Theta$ then a[x, y] = object = 1
Else a[x, y] = background = 0

Image segmentation by using threshold method is quite simple but very powerful approach for segmenting images based on image-space region i.e. characteristics of the image [6]. The threshold method generally used for images that are having light object on darker background or vice versa.

*D. Fuzzy Based Segmentation Technique*

Fuzzy-based segmentation is one of the image segmentation used for the processing of natural images that have uncertainty and ambiguousness. It is a collection of different fuzzy approaches in image processing. The collection of all the approaches that perceive, exemplify and then forms the image, their segments and features them as fuzzy sets.

Fuzzy image processing comprises of three main stages:

- Image fuzzification,
- Modification of membership values and if necessary,
- Image de-fuzzification.

*© 2017, IJCSMC All Rights Reserved* 

This segmentation technique is sensitive to noise, computationally expensive and determination of fuzzy membership is not very easy.

*E. Neural Network-Based Segmentation*

Neural network-based segmentation technique is the fastest computing technique as compared to other segmentation technique. In this algorithm, an image is firstly mapped into a neural network where every neuron represents a pixel [6], [7].
In order to find the connection and weights between nodes the neural network is trained with training sample set. After that the new image is segmented with trained neural network.

Neural network segmentation includes two important steps:

*a. Feature extraction*

Basically, in this step, the input data of neural network is determined. For image segmentation, some important features are extracted.

*b. Image segmentation*

The image is segmented in this step based on the features extracted from the images.

Neural network-based segmentation improves segmentation results when the data is deviates from a normal situation. Its High robustness makes it immune to noise.

*F. Mean Shift Segmentation*

The mean shift based technique is the most powerful segmentation technique. It has become widely-used in the image processing applications. This is a non-parametric iterative algorithm, an advanced clustering based algorithm. The mean shift segmentation is a popular method for estimating the probability. This technique comprises of two basic steps: a mean shift filtering of the original image data and clustering of the filtered data points that is filtering and clustering.

## III. IMPLEMENTATION USING NORMALIZED CUT, KMEANS AND MEAN SHIFT

*Segmentation using Normalized Cut*

Modification of the minimum-cut problem to resolve this difficulty is called "Normalized Cut." A normalized cut is one that seeks to create a cut, which divides the input graph into "object-like" regions [8]. While maximizing the capacity within the regions created by the cut, it minimizes the capacity across the cut. The Normalized Cut' algorithm is a clustering approach that finds an approximation of the Normalized Cut by minimizing only one piece of the Normalized Cut equation.

Steps that are followed in normalized cut technique:

1. Compute weight matrix W, and D.
2. Solve generalized eigen system.
3. Bipartition the graph that normalized cut is minimized.
4. Decide if the current partition is should be divided. If either of partition is too small, stop recursion. If normalized cut is larger, stop recursion.
5. Recursively repartition.

(a) (b)

Fig. 4 Results using Normalized cut segmentation (Color intensity: 7)
a) Input Image b) Normalized Cut

Fig. 4 shows the results by taking one input image and taking the color intensity as integer value, and the output result is displayed using the normalized cut technique.

*© 2017, IJCSMC All Rights Reserved* 

(a) (b)

Fig. 5 Results using Normalized cut segmentation (Color intensity: 10)
a) Input image b) Output Image

The color intensity value in Fig. 5, was increased to 10 for segmentation using normalized cut and the result were analyzed which shows that minimizing the normalized cut maximizes the normalized association.

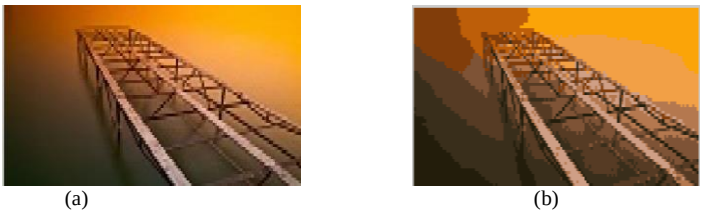
*Segmentation using Kmeans*

In the Kmeans Clustering the letter K is referred to the number of clusters that is to be decided in the starting of the algorithm. In this, we have to define the K centers, one center for each cluster. The center should be far from the others so that the distance could be calculated easily and the data points could predict their clusters very precisely. Then the distance between each data point and the clusters center are calculated. Then data points are assigned to the cluster whose distance from the cluster center is lowest. The distance can be calculated with the help of conventional mathematical perpendicular concept. Now for this iterative process, the same data points are chosen [9], the distance is calculated. Then the data points are assigned which are too nearer to cluster. This process is repeated till we observe the shifting in the center of the cluster. Grouping of clusters is based on maximizing the similarities, if inter class similarities are increased then the quantity of clusters are automatically increased to get the optimum results [10].

(a) (b)

Fig. 6 Results using kmeans clustering (Number of Clusters: 8)
a) Input image b) Output image

In the above Fig. 6 results are obtained using kmeans clustering algorithm, considering the value of k. K-mean clustering is a clustering technique that group n pixels of an image into K number of clusters, where K < n and K is here a positive integer. Initially, the centroids of the predefined clusters are initialized randomly.



(a) (b)

Fig. 7 Results using kmeans clustering (Number of clusters: 10)
a) Input Image b) Output image

The steps that are followed in this process are as follows:

i. Choose the number of clusters K at random.
ii. Choose K pixels of different intensities as Centroids at random.
iii. Calculate the mean of pixel values in a region as centroids. Centroid must be placed as far away with each other.
iv. Comparing a pixel with every Centroid. Assign the pixel to the closest Centroid which thus forms a cluster. After assigning all the pixels the initial clustering is achieved.
v. Recalculate the mean of each cluster and recalculate the position of Centroids in K clusters.
vi. Repeat steps (iv) & (v) until the Centroids no longer move.

*© 2017, IJCSMC All Rights Reserved* 

This process is repeated till we observe the shifting in the center of the cluster. The main drawback having this method is to determine the number of clusters in an image [11]. Thus, Clusters are formed on the basis of some similar features like gray level the distance and intensity of pixels.

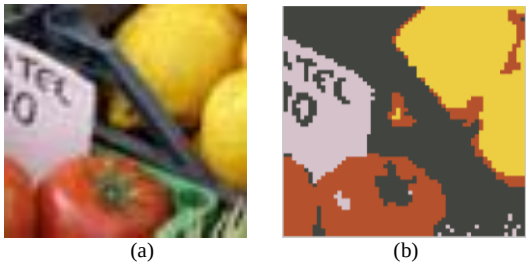
*Segmentation using Mean Shift*

In the field of digital image processing, the mean shift segmentation approach is amongst the most powerful used techniques. The Mean shift algorithm comprises two basic steps for the processing of image segmentation.

1. In the first step, the mean shift filtering of the original image data is performed to provide smoothness to an image while preserving discontinuity i.e. to apply the smoothing process using convergence points.
2. After that clustering is performed on a filtered data points through which the probability density of underlying feature space is analyzed and which effectively converts the filtered points to segmentation.

By using the mean shift segmentation the original image is divided into many regions which form the basis for over segmentation. To remove the over segmentation hierarchical clustering is applied by merging those segmented regions and to achieve the segmentation results.

For image segmentation, mean shift algorithm uses similar principle with smoothing; that is, cutting image on the basis of image smoothing; to put it simply, clustering all pixels converged by one maximum value point of the same density in the input image and assigning cluster's label to all points in the cluster; any one cluster with point number smaller than M after segmentation should be removed. Merge the region where it's pixels fewer than M to its neighboring small region. After pre-segmentation, the image is parted to lots of small regions [12]. In Fig. 8, we see target objects are cut finely after pre-treatment and edge information is well conserved, by defined bandwidth.



(a) (b)

Fig. 8 Results using Mean shift segmentation (Mean shift bandwidth: 0.2002)
a) Input image b) Mean shift segmentation

*Hierarchical Clustering*

Hierarchical clustering forms a cluster tree or *dendrogram* by grouping the data over a variety of scales. The tree forms a multilevel hierarchy where cluster of one level is merged with the cluster of another level. It helps in deciding the level or scale which is appropriate. Hierarchical clustering algorithm follows two approaches. Top-down approach and Bottom-up approach. In Bottom-up Hierarchical clustering algorithms every document is processed as a singleton cluster at the outset and then pairs of clusters are merged one after another until all clusters have been merged into a single cluster that contains all documents. The Bottom-up approach is therefore called *hierarchical agglomerative clustering* or *HAC*. The result of image segmentation by mean shift is shown in Fig. 9.

(a) (b)

Fig. 9 Results using Mean shift segmentation (Mean shift bandwidth: 0.2001)
a) Input image b) Output image

*© 2017, IJCSMC All Rights Reserved* 

Steps that were followed in the proposed work include:

1. Select the input image that is to be segmented.
2. Select the method for segmentation process (Normalized cut, kmeans or mean shift)
3. Select the value from the different parameters that are defined (color intensity, number of clusters, mean shift bandwidth)
4. Process the input image one at a time.
5. Result generated using the selected segmentation technique and the defined parameter.

The Mean shift algorithm is based on probabilistic intuitions and for that, the kernel density estimation needs to be understood, for which to define the distribution the data is used. In measuring kernel density, with relatively small number of samples, the approximate value of density is not estimated. It improves significantly with more number of data samples. While the normalized cut clustering and is based on approximation & minimizing the capacity of cuts.

### IV. RESULTS AND ANALYSIS

Implementation of the above techniques was done on multiple images to partition the image on the basis of various parameters like color intensity, the number of clusters required to segment the image and the mean shift bandwidth; we have limited ourselves to fewer parameter. The parameters were applied accordingly in normalized cut, kmeans and mean shift algorithm and then the results of the output image of these techniques were analysed. An optimal solution cannot be achieved through kmeans clustering as the quality depends upon the number of clusters and to improve the performance for large-sclae images the clusters is to be computed and then reassign the values to the larger image if required. Considering more complicated images there is more isolation and noise points and the time complexity is associated with the number of cluster. While in normalized cut by minimizing the normalized cut leads to maximum association. Mean shift clustering process is highly efficient than kmeans clustering.

### V. CONCLUSION & FUTURE WORK

In this paper we basically focused on the image segmentation technique based on clustering like use of mean shift, normalized cut and kmeans. From the evaluation we have found that the result of the image segmentation depends on selection of various factors like color intensity, pixel color, image content, number of clusters etc. In this paper we have done an experiment on an input image and used parameters like color intensity, number of clusters, mean shift bandwidth and found that there is no particular method for the image segmentation which can give appropriate results and evaluate the effectiveness. We have analyzed the results using normalized cut technique, kmeans and mean shift. The computation can be reduces to half in the future, as because of symmetric matrix.

Thus, Image processing depends on the processing speed and time, we have considered both quality and speed. But different clock speed gives different results. Thus it is very difficult to find any particular method for image segmentation and clustering approach which can give an appropriate result and can be found better than other techniques.

## REFERENCES

[1] Varshali Jaiswal, Aruna Tiwari," A Survey of Image Segmentation based on Artificial Intelligence and Evolutionary Approach", IOSR Journal of Computer Engineering (IOSR-JCE), e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 15, Issue 3, Nov. - Dec. 2013.

[2] Nandhini. P and Dr. J. Jaya, Image Segmentation for Food Quality Evaluation Using Computer Vision System. Int. Journal of Engineering Research and Applications www.ijera.com ISSN : 2248-9622, Vol. 4, Issue 2( Version 5), February 2014.

[3] Bock J D, Smet P D Philips W, Image segmentation using watersheds and normalized cuts, Proceedings of SPIE 2005.

[4] K. Dehariya, S. K. Shrivastava, R. C. Jain, "Clustering of Image Data Set Using K-Means and Fuzzy K-Means Algorithms", International conference on CICN, pp. 386-391, 2010.

[5] Linda G. Shapiro and George C. Stockman: "Computer Vision", pp. 279-325, New Jersey, Prentice-Hall, 2001.

[6] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", 2nd ed., Beijing: Publishing House of Electronics Industry, 2007.

[7] Tao W B, Jin H, Zhang Y M. Color image sSegmentation based on mean shift and normalized cuts. IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, 2007.

[8] W. X. Kang, Q. Q. Yang, R. R. Liang,"The Comparative Research on Image Segmentation Algorithms", IEEE Conference on ETCS, pp. 703-707, 2009.

*© 2017, IJCSMC All Rights Reserved* 

[9] Rozy Kumari and Narinder Sharma, A Study on the Different Image Segmentation Technique. International Journal of Engineering and Innovative Technology (IJEIT) Volume 4, Issue 1, July 2014.
[10] S. Karkra ,J.B. Patel " ATLAS BASED MEDICAL SEGMENTATION TECHNIQUES –A REVIEW "GE-INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH VOL-3,ISSUE-5, May 2015.
[11] Naveen Tokas et al," Comparison of Digital Image Segmentation Techniques- A Research Review", International Journal of Computer Science and Mobile Computing, Vol.5 Issue.5, , pg. 215-220, May- 2016.
[12] Lihua Tian, Liguo Han and Junhua Yue," Research on Image Segmentation based on Clustering Algorithm", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.9, No.2, 2016.

*© 2017, IJCSMC All Rights Reserved* 