

SURVEY ON CLUSTERING BASED COLOR IMAGE SEGMENTATION AND NOVEL APPROACHES TO FCM ALGORITHM

Monika Xess¹, S. Akila Agnes²

¹PG Student, Department of Computer Science and Engineering, Karunya University, Tamil Nadu, India,
monika.xess92@email.com

² Assistant Professor, Department of Computer Science and Engineering, Karunya University, Tamil Nadu, India,
akilaagnes@karunya.edu

Abstract

Segmentation is an important image processing technique that helps to analyze an image automatically. Applications involving detection or recognition of objects in images often include segmentation process. This paper describes two unsupervised clustering based color image segmentation techniques namely K-means clustering and Fuzzy C-means (FCM) clustering. The advantages and disadvantages of both K-means and Fuzzy C-means algorithm are also presented in this paper. K-means algorithm takes less computation time as compared to Fuzzy C-means algorithm which produces result close to that of K-means. On the other hand in FCM algorithm each pixel of an image can have membership to more than one cluster which is not in case of K-means algorithm, an advantage to FCM method. Color images contain wide variety of information and are more complicated than gray scale images. In image processing, though color image segmentation is a challenging task but provides a path for image analysis in practical application fields. Secondly some novel approaches to FCM algorithm for better image segmentation are also discussed such as SFCM (Spatial FCM) and THFCM (Thresholding FCM). Basic FCM algorithm does not take into consideration the spatial information of the image. SFCM specially focus on spatial details and contribute towards image segmentation results for image analysis. It introduces spatial function into FCM algorithm membership function and then operates with available spatial information. THFCM is another approach that focus on thresholding technique for image segmentation. Its main task is to find a discerning cluster that will act as automatic threshold. These two approaches show how better segmentation results can be obtained.

Key Words: K-means clustering, Fuzzy C-means clustering, SFCM, THFCM.

1. INTRODUCTION

Image segmentation is a technique of dividing a digital image into several homogeneous parts or segments. Color images consist of more attributes or components as compared to gray images. Applications of color image segmentation can be seen in the field of video surveillance, recognition tasks (face recognition, fingerprint recognition), object detection, content-based image retrieval, and medical imaging. Color image segmentation is the challenging task in image processing and contains two critical issues, firstly which color model to be used and secondly, which segmentation technique should be applied.

A color space is a geometrical representation of color in a 2D- 3D space. Different color models used are RGB (Red, Green, Blue), YIQ, YUV, HSI (Hue, Saturation, Intensity), HSL (Hue, Saturation, Luminance), CIE, YCbCr. All the color models have their own application and advantages. For example HSL (Hue, Saturation, Luminance) color model is good for color image segmentation while YCbCr color

model finds application in image and video processing. Several algorithms and techniques have been developed for image segmentation. Some of the techniques are described below [1].

In histogram based segmentation technique a histogram of the image is created, then significant peaks and valleys are identified which are used to locate the clusters in the image. In feature space clustering approach a feature vector of attributes such as intensity etc is calculated at each pixel. The feature space is then divided into different groups or clusters and individual pixel is labelled with the cluster that consists its feature vector. Region-based segmentation is a method for determining the regions which includes region-growing, region-splitting and region merging steps. Edge detection is the technique of discovering the outside limit of an image surface and based on that forms the segments. Fuzzy approach is based on (IF-THEN) inference rules and neural networks are used to perform clustering. Clustering based techniques focus on forming clusters or group of similar pixels or segments that is dividing the source image

into multiple similar parts. In view of this clustering based segmentation technique involves methods such as K-means clustering, Fuzzy c-means (FCM), Self-Organizing Maps (SOMs). Also combination of above mentioned different techniques can be used for image segmentation.

The paper is organized as follows: section 2 presents clustering based color image segmentation techniques. Section 3 describes the work contributed towards FCM algorithm for color image segmentation. Finally section 4 draws the conclusion.

2. CLUSTERING BASED COLOR IMAGE SEGMENTATION

Clustering is a process of grouping the similar pixels that is forming clusters. Here the two clustering based algorithm K-means clustering and Fuzzy C-means are discussed in detail.

2.1 K-means clustering algorithm

The K-means algorithm (Mac Queen, 1967) is an unsupervised clustering algorithm that is used to divide an image into k clusters. Steps in the algorithm are as follows [2]:

Step1. Choose the number k of clusters, either randomly or based on some heuristic.

Step2. Generate k clusters and determine the clusters centre.

Step3. Assign each pixel in the image to the clusters that minimize the distance between the pixel and the cluster centre (Distance is the squared or absolute difference between a pixel and a cluster centre).

Step4. Re-compute cluster centre by averaging all of the pixels in the cluster.

Step5. Repeat steps 2 and 3 until convergence is attained (for example cluster centre remains unchanged).

Advantages of K-means algorithm

- When k is small, K-means is computationally faster.
- It may produce clenched clusters than hierarchical clustering if the clusters are globular.

Disadvantages of K-means algorithm

- Difficult to predict k with fixed number of clusters.
- Does not work well with non-globular cluster.

2.2 Fuzzy C-means clustering algorithm

Fuzzy C-means algorithm (Bezdeck 1981) is an important tool for image processing in segmentation of color image. It is an iterative clustering algorithm in which a pixel can belong to more than one cluster and with each pixel a set of membership level is associated. The procedure of FCM algorithm is given as follows [3]:

1. Set the number of fuzzy cluster, A and the cluster centres.
Set iteration time $q=0$;
2. while $\|W^{(q)} - W^{(q-1)}\| \geq \epsilon$ do
3. for $j = 1$ to F^{th} cluster do
4. for $i=1$ to N^{th} image pixel do
5. Calculate u_{ij} , i.e. the membership of pixel x_i to the j^{th} cluster;
6. if $\|x_i - c_j\| = 0$ then
7. $u_{ij} = 1$;
8. reset other membership of pixel i to 0;
9. end if
10. end for
11. Update c_j (cluster center);
12. end for
13. Calculate objective function $W^{(q)}$;
14. $q=q+1$;
15. end while

where, $W^q = \sum_{j=1}^F \sum_{i=1}^N u_{ij}^m \|x_i - c_j\|^2$ is the objective function.

$$u_{ij} = \frac{\|x_i - c_j\|^{-2/m}}{\sum_{k=1}^F \frac{\|x_i - c_k\|^{-2/m}}{\|x_i - c_k\|^{-2/m}}} \quad \text{is defined as}$$

membership of pixel x_i to the cluster i.

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m x_i}{\sum_{i=1}^N u_{ij}^m} \quad \text{is the cluster centre.}$$

m is a constant that defines the fuzziness,

ϵ is a constant and its value is manually determined in the range of [0.01, 0.0001]

Advantages of Fuzzy C-means algorithm

- As compared to K-means algorithm FCM is better.

- As each pixel is assigned membership to each cluster unlike K-means clustering a pixel may belong to more than one cluster.

Demerits of Fuzzy C-means algorithm

- Computational time is more.
- Sensitive to initialization condition of cluster number and cluster center.

3. Novel Approaches to FCM Algorithm

FCM is the most widely used color image segmentation technique. Some of limitations associated with FCM [4]: (a) In terms of performance FCM depends on the initial cluster centre; (b) The cluster number must be fixed before clustering; (c) Computational complexity is more; and (d) spatial information is not considered in conventional FCM algorithm. In view of these limitations many research work has been done. Here we present two of such works that contribute to FCM algorithm for better segmentation results based on required domains.

3.1 Spatial FCM (SFCM)

In [5] Fuzzy C-means clustering with spatial information is proposed for image segmentation. One of the facets of an image is that the nearest pixels are closely related to each other and there is greater chances that they belong to the same group or cluster. In conventional FCM this spatial information is ignored which is an important parameter for clustering. In SFCM spatial information is fused in the membership function to obtain better segmentation results and is defined as follows:

$$u_{ij}^r = \frac{u_{ij}^p h_{ij}^q}{\sum_{k=1}^c u_{kj}^p h_{kj}^q}$$

where p and q are constants to supervise the relative importance of both spatial and membership functions. The spatial function is given as:

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik}$$

where $NB(x_j)$ is a square window centered on pixel x_j in the spatial domain.

The clustering technique is a 2-pass procedure at each iteration. Steps involved in SFCM clustering are as follows:

Step1. In first pass the membership function is computed as similar as in conventional FCM.

Step2. In the second pass, the membership information of each pixel is delinated to the spatial domain, and the spatial function is determined from that.

Step3. Next the FCM iterative procedure continues until some convergence is reached that is when the maximum difference between cluster centroids at two successive iterations is less than a threshold = 0.02.

Step4. Finally defuzzification process is used to allocate each pixel to a distinct cluster for which the membership is maximal.

Advantages of spatial FCM are:

- Eliminates noisy spots
- Reduces false blobs
- Less sensitive to noise
- More homogeneous regions are obtained

3.2 THFCM (Thresholding by fuzzy c-means) algorithm

In [6], hybrid approach based on thresholding by FCM algorithm is proposed. Thresholding is a widely used methodology in image segmentation which partitions an image into segments using particular threshold value. It is used in differentiating the foreground from the background. In the hybrid thresholding and FCM (THFCM) approach the aim is to find a discerner cluster that is treated as automatic threshold. Steps in THFCM algorithm are as follows:

Step1. Create a histogram for the image, where the x-axis value lies between [0-255] and y-axis is for the frequencies for each pixel.

Step2. Smooth the histogram by removing insignificant values. Use filter to eliminated insignificant peaks and valleys from the image histogram created.

Step3. FCM algorithm is applied on y-axis values (In conventional FCM, it is applied to the x-axis values). y-axis values define the number of occurrence for each pixel. The iterative FCM methodology is then applied.

Step4. Determine the discerner cluster among the clusters which has maximum frequency in the y-axis value. This discerner cluster acts as automatic threshold.

If number of clusters are three, then discerner cluster represent the foreground with pixels inside it as (1) while other two clusters represent background (0). This iterative approach overcomes the drawback of the existing method to determine a threshold.

4. CONCLUSION

In this paper, we have provided an extensive overview of clustering based color image segmentation techniques and highlight two approaches to FCM algorithm for better segmentation results. Clustering based segmentation methodologies involve K-means and Fuzzy C-means clustering algorithm. K-means is faster when k is small and hence form tighter cluster, but it is quite difficult to predict k with fixed number of clusters. On the other hand Fuzzy C-means algorithm is comparatively better than K-means and a pixel can belong to more than one cluster, but it takes more computation time. Furthermore, SFCM overcomes the limitation of conventional FCM towards noisy image and provides better segmentation results. Also, THFCM approach solves the problem of existing method to determine a threshold for excellent segmentation. Thus by analyzing new problem areas related to image segmentation further works can be done on basic K-means and Fuzzy C-means algorithm to obtain more faster and accurate segmentation result.

REFERENCES

- [1] H.D. Cheng, X.H. Jiang, Y. Sun, J. Wang, Color image segmentation: advances and prospects, *Pattern Recognition* 34 (2001) 2259–2281.
- [2] B. R. Jipkate, V.V Gohokar, “A Comparative Analysis of Fuzzy C-Means Clustering and K Means Clustering Algorithms” *IJCER* May-June 2012 Vol. 2 Issue No.3 737-739.
- [3] K. S. Tan, H.W. Lim, N. A. M. Isa, Novel initialization scheme for Fuzzy C-Means algorithm on color image segmentation, *Applied Soft Computing* 13 (2013) 1832–1852.
- [4] K. Chen, Y. Ma, L. Jun, Segmentation by fusion of self-adaptive SFCM cluster in multi-color space components, *International Journal of Image Processing (IJIP)* 6 (2012) 157–166.
- [5] K.S. Chuang, H.L. Tzeng, S. Chen, J. Wu, T.J. Chen, Fuzzy C-means clustering with spatial information for image segmentation, *Computerized Medical Imaging and Graphics* 30 (2006) 9–15.
- [6] F. A. Jassim, “Hybrid Image Segmentation using Discerner Cluster in FCM and Histogram Thresholding”, *International Journal of Graphics & Image Processing* Vol.2 Issue 4 November 2012.

BIOGRAPHIES



Monika Xess obtained her bachelor's degree in Computer Science Engineering from CSVTU university and currently pursuing M.Tech degree in Computer Science Engineering from Karunya University Coimbatore, Tamil Nadu, India. Research area of interest is Image Processing.

S. Akila Agnes, M.E in Computer Science Engineering. Currently working as Assistant Professor in Department of Computer Science and Engineering, Karunya University Coimbatore, Tamil Nadu, India. Area of interest is Image Processing.