CLUSTER BASED ON SEGMENTATION MEDICAL IMAGES USING MATLAB

Mandeep
Research Scholar
Department Of Electronic & Communication Engg.
Galaxy Global Group of Institutions, Dinarpur

Vikas Chawla
Faculty,Department of Electronic & Communication Engg.
Galaxy Global Group of Institutions, Dinarpur
Email: chawla.vikas1@gmail.com

Saranjeet Singh
Faculty,Department of Electronic & Communication Engg.
Galaxy Global Group of Institutions, Dinarpur
Email: mail2saranjeet@gmail.com

Abstract
Image processing is one of most growing research area these days and now it is very much included with the medical and biotechnology field. Image Processing can be used to analyze different medical image to get the abnormality in the image. In this paper, a new segmentation scheme for the white blood cells from microscopic images is proposed. The is based on the K-means clustering technique. The RGB test images are converted to the L*a*b color space, and then the two color machinery (a and b) are used as features to the K-means clustering algorithm. The success of image analysis depends on segmentation reliability. The accurate partition of the image into regions is a challenging task. K-Means Clustering algorithm is the popular unsupervised clustering for dividing the images into multiple regions based on image color property.

Key words: **DSP,MSE.PSNR,K-Mean**

1.0 Introduction

Digital images provide important information that on one hand is used in most real-world applications such as remote sensing, defense medical imaging processing for diagnosis of diseases, face recognition for security purposes, satellite images used for weather updates and so on. Hence it is very important that images should be noise free for various processing in image sector. On the other hand, users of these applications suffer from quality degradation issues caused during image transmissions over the networks. Nature of the problem depends on the type of noise added to the image. Extracting useful information from corrupted images without having the knowledge about the noise type from which they are polluted has been an issue in real world applications. The noises that may wipe out image textures include Gaussian, speckle, and salt and pepper noise. Various techniques have been developed to suppress noise and boost image quality such as linear filters, nonlinear filters and wavelet threshold-based techniques. Almost all the techniques developed so far do not attempt to diminish the effects of multiple noises. Hence, numerous hybrid techniques have been proposed to increase visual quality of a received image by removing multiple noises [1] [2].

2.0 Related Work

Pranay Yadav() presents a Modified Adaptive Threshold Median Filter (MATMF) for impulse noise removal for both low and high-density noise levels. They proposed filter for color image random valued impulse noise reduction. For color images, impulse noise removal there are two main stages, firstly, the detection of the impulse noises on the basis of maximum and minimum value of pixels in a small window in the filtering stage, the noise-free pixels remain unchanged in all RGB frames, windows and noisy pixels are restored using a median filter. Experimental results show that their method is superior to the conventional methods in Peak Signal Noise Ratio [8] (PSNR) values and

Mean Square Error (MSE) of the image performs, with and without noise compared. **Isma Irum** proposed Impulse noise reduction or removal was a very active research area of image processing. A nonlinear hybrid filter for removing fixed impulse noise (salt & pepper) noise from color images. Technique was based on mathematical morphology and trimmed standard median filter. This filter was composed of a sequence of morphological standard and well known operations erosion-dilation and trimmed standard median filter. It removes the fixed impulse noise (salt & pepper) very well without distorting the image features, color components and edges. It does not introduce blurring and moving effects even in high noise densities (up to 90%).. A unique sequence of these operators with trimmed median filter has been composed that effectively removes the salt & pepper noise from color images. The results evaluate the performance of hybrid filter. This filter very effectively removes the noise from edge regions and inner regions of noisy images without distorting the features and color components [4]. **Kehan Shi** proposed Utilizing local Hölder seminorm and nonlocal operator, they propose two efficient salt-and-pepper noise removal algorithms. They first minimize a local Hölder seminorm based functional which has a great capacity to restore natural images. Then by the definition of nonlocal operator, a new TV-based functional which inherits the advantage of nonlocal method and not only suppresses the noise but also restores the geometrical and texture features of noisy images. An alternative numerical scheme is also proposed to solve our functional which reduces the computational complexity greatly. These algorithms are efficient even when the noise level is as high as 90 %. Zhichang Guo proposed two functionals for salt-and pepper noise removal based on local Hölder seminorm and nonlocal operator, respectively[8].

3.0 Segmentation Based On Artificial Neural Network

Neural Network based segmentation is totally different from conventional segmentation algorithms. In this, an image is firstly mapped into a Neural Network. Where every Neuron stands for a pixel [16]thus image segmentation problem is converted into energy minimization problem. The neural network was trained with training sample set in order to determine the connection and weights between nodes. Then the new images were segmented with trained neural network, for example, we can extract image edges by using dynamic equations which direct the state of every neuron towards minimum energy defined by neural network.

Neural network segmentation includes two important steps feature extraction and image segmentation based on neural network. Feature extraction is very crucial as it determines input data of neural network [44], firstly some features are extracted from the images, such that they become suitable for segmentation and then they were the input of the neural network. All of the selected features compose of highly non-linear feature space of cluster boundary.

Neural network based segmentation have three basic characteristics:-

(i). highly parallel ability and fast computing capability which makes it suitable for real time application.

(ii). Improve the segmentation results when the data deviates from the normal situation .

(iii). Robustness making it insensitive to noise.

(iv). Reduced requirement of expert intervention during the image segmentation process.

However there are some drawbacks of neural networks based segmentation either, such as:-

(a). Some kind of segmentation information should be known beforehand.

(b). Initialization may influence the result of image segmentation.

(c). Neural network should be trained using learning process beforehand [16], the period of training may be very long, and we should avoid overtraining at the same time.

3.1 Segmentation Based On Clustering

Clustering is an unsupervised learning task, where one needs to identify a finite set of categories known as clusters to classify pixels [13]. Clustering use no training stages rather train themselves using available data. Clustering is mainly used when classes are known in advance. A similarity criteria is defined between pixels [14], and then similar pixels are grouped together to form clusters. The grouping of pixels into clusters is based on the principle of maximizing the intra class similarity and maximizing the inter class similarity. The quality of a clustering result depends on both the similarity measure used by the method and its implementation. Clustering algorithms are classified as hard clustering, k- means clustering, fuzzy clustering, etc.

3.1.1 Hard Clustering

Hard clustering assumes sharp boundaries between clusters [18]; pixel belongs to one and only one cluster. A popular and well known hard clustering algorithm is K-means clustering algorithm [13].

K-means algorithm is a clustering technique to partition n pixels into k clusters, where $k < n$. K-means algorithm Developed by Mac Queen in 1965 and then refined by Hartigan and Wong in 1979. K-means algorithm is a clustering technique [13], which classify pixels in an image into K number of clusters, where K is a positive integer, according to some similarity feature like grey level intensity of pixels and distance of pixel intensities [14], from centroid pixel intensity. The main advantages of this algorithm are its simplicity and low computational cost, which allow it to run efficiently on large data sets.

The main drawback is that: K the number of clusters must be determined it does not yield the same result each time the algorithm is executed and the resulting clusters depend on the initial assignments of centroids. The process is as follow

1. Randomly choose number of clusters K.
2. Compute the histogram of pixel intensities.
3. Randomly choose K pixels of different intensities as Centroids.
4. Centroids are finding out by calculating mean of pixel values in a region and place Centroids as much far away from each other as possible.
5. Now, compare a pixel to every Centroid and assign pixel to closest Centroid to form a cluster,
$$C(i) := \arg\min ||x^{(i)} - U_j||$$
6. When all pixels have been assigned, initial clustering has been completed.
7. Recalculate position of Centroids in K clusters [13]
$$U_i := \frac{\sum_{i=1}^{m} 1\{c_{(i)} = j\}}{\sum_{i=1}^{m} 1\{c_{(i)} = j\}} x^{(i)}$$
8. Repeat step 5 & 6, until Centroids no longer move.
9. Image separated into K clusters.

3.2 Multi-objective Image Segmentation

Earlier image segmentation problem has been treated as mono-objective. Mono-objective images considers only one objective, because a single segmentation image. Such type of segmented images are of good quality but may not allow a higher level process (as image segmentation considered as low level process & pattern recognition, object tracking & scene analysis as high level process) to extract all information included within the image, so different segmentation results are calculated. Image segmentation is a multi objective optimization problem [13]. The consideration of multiple criteria (objectives) starts from the understanding of image pattern to its selected segmentation process involved (feature selection/extraction, similarity/ dissimilarity measure) and finally the assessment of its output (validity assessment). As there are possibilities of multiple sources of information for a segmentation problem, thus multiple representations have to be considered, for example, feature selection is the process of identifying similarity criteria used in segmentation process, now either only single criteria is used, that is intensity of pixels, or to make it a multi-objective problem consider several similarity criteria to segment same image, which can be intensity, color, texture, shape, spatial information. For instance, in segmenting a medical image based on CT scan, multiple features like intensity, shape and spatial relationship could be considered. Similarly criteria for inter pattern similarity that is grouping can be multiple, spatial coherence vs. feature homogeneity, connectedness vs. compactness, diversity vs. accuracy. For image segmentation multiple methods can be used for getting appropriate output, and there may be a tendency for multiple optimizations and decision making processes where multiple validity assessment should be used. There are two general approaches for Multi-objective optimization problem, the first approach is to combine multiple objective functions into a single composite function, and the second is to determine a set of solutions that are non-dominated with respect to each objective.

3.2.1 The Conventional Weighted Formula Approach (WFA)

In this approach a Multi-objective problem is Transformed into a problem with single objective [13], that is typically done by assigning a numerical weight to each objective and then combining the values of weighted criteria into a single value by either adding or multiplying weighted criteria. The quality of a given candidate model is given by one of the two kinds of formula:

$$Q = w_1c_1 + w_2c_2 \ldots + w_nc_n$$
$$Q = w_1c_1 \times w_2c_2 \ldots \times w_nc_n$$

$w_i, i = 1, 2 \ldots n$ denotes the weight assigned to criteria c_i and n is the number of evaluation

4.0 Algorithm Steps

There are four steps for detecting a face in a single image using the color model as described below:

4.1 Segmentation: An important task in several vision applications is the image pixel classification in a discrete number of classes. The objective of segmentation is providing an effective and real time classification. The first step for segmentation is obtaining a set of pixel values, which corresponds to the color skin in the images. The set is obtained manually, selecting a small region directly from a frame where the face is located. Then this set to define the maximum and minimum skin pixel values for each channel of the color space. This vector will be employed in subsequent classifications due the skin color constitutes a regular cluster in the color space. Next, we are classifying the pixels by applying the thresholding operation to the frame. Thresholding involves the use of the vector values in the corresponding three-dimensional color space. Every vector values can be considered as a class.

4.2. Regions: After segmentation is performed, then connect all pixels to produce regular regions or blobs. This process is exhaustive and can affect the real time performance. The threshold values have a large influence on the segmentation results. A small threshold value leads to a large number of small regions while with a large threshold value few large regions are calculated.

4.3. Localization: In this stage firstly detect and locate the face in the frame. Then search in all labeled regions for the one that satisfies a specific area (an approximately number of pixels to form a face) and size (high and width considering the region is rectangular). After the region of interest is located, we obtain the position of the region center. Dimension and center position of the face region is computed for each color model. In this way, now we get the many regions as color spaces where the detection was successful for the polling process.

4.4. Polling: Finally, validate the face detection and position in the frame by polling. Simply, then examine all the regions, and if there is a region common to at least three colors spaces, then considered such region as the detected face, in the other way, if the mentioned condition is not satisfied, the region is discarded and the process is performed to the next region or the next frame.

5.0 Proposed methodology

The proposed EKM algorithm employs hard membership function as its clustering concept, where the data are clustered into non-overlapping regions. Consider an image, *X* with *N* data to be clustered into *K* regions. Initially, all centers values are randomly assigned. The *i*-th data, *i v* is assigned to the nearest data cluster, *j c* based on the minimum Euclidean distance, where *i* □1,2,3.....*N* and *j* □1,2,3.....*K* . After completing the assigning process for all data, the new position of centers is
measured according to:

$$C_{J} = 1/N_{J}\ \Sigma v_{i}$$

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2$$

where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j ,is an indicator of the distance of the *n* data points from their respective cluster centers.

The most common algorithm, described below, uses an iterative refinement approach, following these steps:

- Define the initial groups' centroids. This step can be done using different strategies. A very common one is to assign random values for the centroids of all groups. Another approach is to use the values of *K* different entities as being the centroids.
- Assign each entity to the cluster that has the closest centroid. In order to find the cluster with the most similar centroid, the algorithm must calculate the distance between all the entities and each centroid.
- Recalculate the values of the centroids. The values of the centroid's fields are updated, taken as the average of the values of the entities' attributes that are part of the cluster.
- Repeat steps 2 and 3 iteratively until entities can no longer change groups

6.0. Result Anlaysis

This work has been implemented on Intel(R) Core (TM) i3 at 2.20GHz using MATLAB version 7.8.0.347 and tested with many DICOM files obtained from different orientation of CT Scan machine to provide different Coronary Arteries of Heart .The developed program is able to segment the coronary arteries of Heart using a 64 Slice DICOM file of Heart CT Scan Data. The different results obtained from the output of the program are written in the folder specified to assist the physician via directly indicating the region of calcification in the respective coronary arteries of an abnormal patient i.e. patient with calcifications in one or many coronary arteries. This algorithm when applied to a normal patient i.e. a person with no calcification in the coronary arteries shows no region of calcification.

For the given input RGB image, the first process is to determine the number of clusters K which is equivalent to number of regions. The number of clusters K is then passed to fast K-Means algorithm to segment the input RGB color image into K regions. A GUI Interface Screen is implemented for the proposed system to implement automatic color image segmentation and the segmented regions of Heart image.

Time Complexity

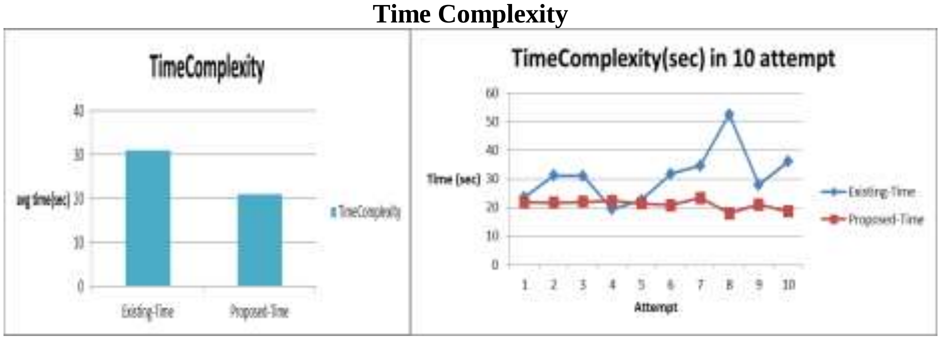

Figure 5.1

Fig5.1 shows the Time Complexity values comparison of different algorithms as a function of attempt for heart image. Here as the density varied from 10-90 % i.e. even at high density the proposed algorithm shows minimum Time Complexity values in comparison with the existing algorithms, showing the effectiveness of proposed algorithm. A plot of PSNR values has been presented in Fig.4.6.

Peak Signal-to-Noise Ratio PSNR

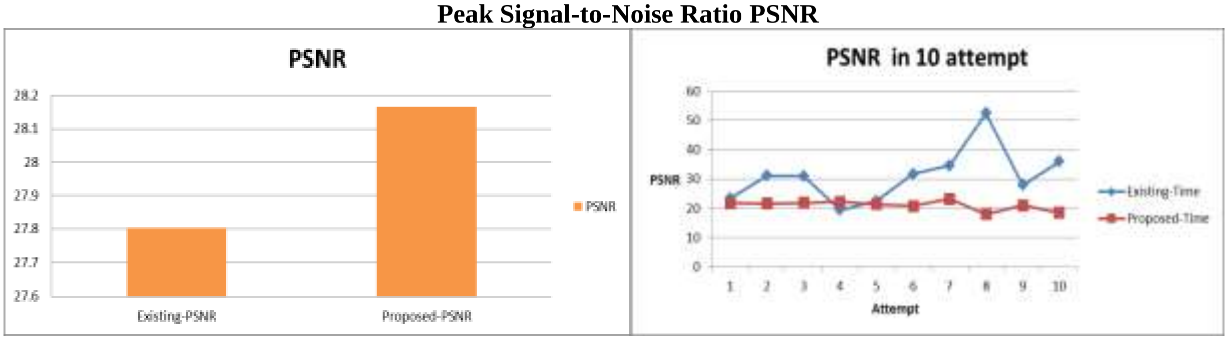

Figure 5.2

Fig 5.2 shows the PSNR values comparison of different algorithms as a function of attempt for heart image. Here as the density varied from 10-90 % i.e. even at high density the proposed algorithm shows minimum PSNR values in comparison with the existing algorithms, showing the effectiveness of proposed algorithm.

Mean Square Error (MSE)

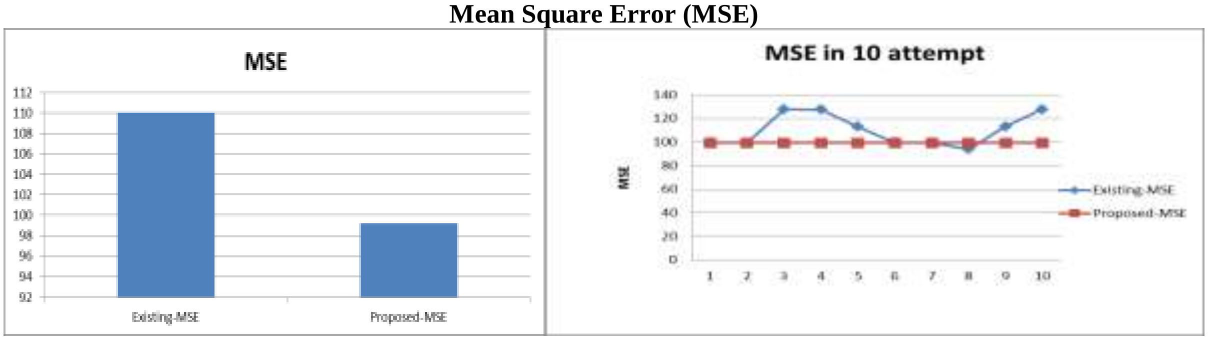

Fig 5.3

Fig 5.3 shows the mean square error comparison of different algorithms as a function of noise density for Heart image. Here as the noise density varied from 10-90 % i.e. even at density the proposed algorithm shows minimum MSE values in comparison with the existing algorithms, showing the effectiveness of proposed algorithm

7.0 Conclusion And Future Scope

In this paper we combine segmentation and K-means clustering. The Dissertation is proposing a better way of segmenting the coronary arteries of the Heart and assists the physician to diagnose the calcification regions in the coronary arteries by directly indicating the region of calcification by Blue color and the region of normal artery is indicated by the Red color. The proposed automatic color image segmentation algorithm includes automatic method for determining region count - K .

8.0 References

[1]. Memoona Malik, Faraz Ahsan and Sajjad Mohsin, 2014. Adaptive image denoising using cuckoo algorithm, pp. 926s IEEE.

[2]. Pranay Yada."Color Image Noise Removal by Modified Adaptive Threshold Median Filter for RVIN, pp.175.IEEE.

[3]. Mayank Tiwari and Bhupendra Gupta, 2015. Image Denoising using Spatial Gradient Based Bilateral Filter and Minimum Mean Square Error Filtering, pp.638 IEEE.

[4]. Sma Irum, Muhammad Sharif, Mudassar Raza and Sajjad Mohsin, 2015. A Nonlinear Hybrid Filter for Salt & Pepper Noise Removal from Color Images, pp.79 IEEE.

[5]. Atluri Srikrishna, B. Eswara Reddy and Manasani Pompapathi, 2015. Pixon Based Image Denoising Scheme by Preserving Exact Edge Locations IEEE.

[6]. Sinem Özdemir and Bekir Dizdaroğlu, 2015. PDEs-Based Gaussian Noise Removal from Color Images, pp.246 IEEE.

[7]. Bogdan Smolka, Krystyna Malik and Dariusz Malik, 2015. Adaptive rank weighted switching filter for impulsive noise removal in color images pp.289 IEEE.

[8]. Kehan Shi, Zhichang Guo, Gang Dong, Jiebao Sun, Dazhi Zhang and Boying Wu, 2015. Salt-and-Pepper Noise Removal via Local Hölder Seminorm and Nonlocal Operator for Natural and Texture Image, pp.400 IEEE.

[9].Bernardino Roig and Vicente D. Estruch, 2016. Localised rank-ordered differences vector filter for suppression of high-density impulse noise in colour images, pp.246 IEEE

[10]. Santosh M .Tondare and Niteen S. Tekale, 2015. An Efficient Algorithm for Removal of Impulse Noise in Color Image Through Efficient Modified Decision Based Unsymmetric Trimmed Median Filter, pp.476 IEEE.

[11]. Bora Jin, Su Jeong You and Nam Ik Cho, 2015. Bilateral image denoising in the Laplacian subbands, pp.122 IEEE.

[12]. S.Muthukumar, P.Pasupathi, S.Deepa and Dr. N Krishnan, 2015. An efficient Color Image Denoising method for Gaussian and Impulsive Noises with blur removal IEEE.

[13]. Sara Behjat-Jamal, Recep Demirci and Taymaz Rahkar-Farshi, 2015. Hybrid Bilateral Filter. IEEE.

[14]. Xue Han, Xiaobo Lu, Xuehui Wu and Chunxue Liu, 2015. An Edge Detection Based Anisotropic Denoising Method for Mobile Phone Images, pp.876 IEEE.

[15]Mia Rizkinia and Keiichiro Shirai, 2015. Local Spectral Component Decomposition for Multi-Channel Image Denoising, pp.3208 IEEE.

[16]. Mehdi Teimouri, Ehsan Vahedi, Alireza Nasiri Avanaki and Zabihollah Hasan Shahi, 2015. AN EFFICIENT DENOISING METHOD FOR COLOR IMAGES pp.208 IEEE.

[17]. Stefan Schulte, Samuel Morillas, Valentín Gregori, and Etienne E. Kerre, 2015. A New Fuzzy Color Correlated Impulse Noise Reduction Method, pp.2565 IEEE.

[18]. M. Szczepanskia, B. Smolkaa, K.N. Plataniotisb and A.N. Venetsanopoulosb, 2004. On the distance function approach to color image enhancement, pp. 283 IEEE.