Segmentation of Colour Data Base Image by Implementing K-Means Clustering

Madhusmita Sahu
Research scholar, Dept. of ECE, CUTM
Asst.Prof., GGI, Berhampur
ODISHA, INDIA

Dr.K.Parvathi
Professor, School of Electronics Engineering
KIIT University, Bhubaneswar
ODISHA, INDIA

Abstract:- **In image analysis techniques, image segmentation takes a major role for analyzing any type of image. The K-means clustering algorithm is one of the widely used algorithm in image segmentation system. This paper proposes the colour data base image segmentation using the L*a*b* colour space and K-means clustering. This work presents a data base image segmentation based on colour features with K-means clustering unsupervised algorithm developed with MATLAB coding. The entire work is divided into two stages. First enhancement of colour separation of data base colour image using de-correlation stretching is carried out and then the six data base image regions are grouped into a set of three clusters using K-means clustering algorithm. By applying the L*a*b* colour space and k-means clustering algorithm in colour data base image we can only point out the major area of any image. By this process we can isolate the infected area in medical data base colour image and cure the disease easily. For better result we can use some optimization technique like Particle swarm Optimization(PSO).**

Keywords:-image segmentation, De-correlation, K-means, Data base colour image.

I.INTRODUCTION

Generally, image segmentation is defined as: "the search for homogenous regions in an image and later the classification of these regions". It also means the partitioning of an image into meaningful regions based on homogeneity or heterogeneity criteria. Image segmentation techniques can be differentiated into the following basic concepts: pixel oriented, Contour-oriented, region-oriented, model oriented, colour oriented and hybrid. Colour segmentation of image is a crucial operation in image analysis and in many computer vision, image interpretation, and pattern recognition system, with applications in scientific and industrial field(s) such as medicine, Remote Sensing, Microscopy, content based image and video retrieval, document analysis, industrial automation and quality control . The performance of colour segmentation may significantly affect the quality of an image understanding system. The most common features used in image segmentation include texture, shape, grey level intensity, and colour. The constitution of the right data space is a common problem in connection with segmentation/classification. In order to construct realistic classifiers, the features that are sufficiently representative of the physical process must be searched**.** Segmentation evaluation techniques can be generally divided into two categories (supervised and unsupervised). The first category is not applicable to remote sensing because an optimum segmentation (ground truth segmentation) is difficult to obtain. Moreover, available segmentation evaluation techniques have not been thoroughly tested for remotely sensed data. Therefore, for comparison purposes, it is possible to proceed with the classification process and then indirectly assess the segmentation process through the produced classification accuracies. For image segment based classification, the images that need to be classified are segmented into many homogeneous areas with similar spectrum information firstly, and the image segments' features are extracted based on the specific requirements of ground features classification [1].

The colour homogeneity is based on the standard deviation of the spectral colours, while the shape homogeneity is based on the compactness and smoothness of shape. There are two principles in the iteration of parameters:1) In addition to necessary fineness, we should choose a scale value as large as possible to distinguish different regions; 2)

we should use the colour criterion where possible. Because the spectral information is the most important in imagery data, the quality of segmentation would be reduced in high weightiness of shape criterion[2].
This work presents a novel image segmentation based on colour features from the images. In this we did not used any training data and the work is divided into two stages. First enhancing colour separation of data base colour image using decorrelation stretching is carried out and then the regions are grouped into a set of three cluster as per the colour present in the image, after applying the K-means clustering algorithm. But when we take the MR data base and implemented K-means clustering it produce two set of cluster [2,3]. Using this two step process, it is possible to reduce the computational cost avoiding feature calculation for every pixel in the image. Although the colour is not frequently used for image segmentation, it gives a high discriminative power of regions present in the image.
This work is organized as follows: proposed De-correlation stretching, basic k-means clustering algorithm and segmentation of colour data base image using k-means clustering algorithm are explained in section II Experimental results obtained with suggested method are shown in Section III. Finally, section IV concludes with some advantages and remarks.[7]

II. PROPOSED ALGORITHM

A. *De-correlation Stretching*

De-correlation stretching enhances the colour separation of an image with significant band-to-band correlation. The exaggerated colours improve visual interpretation and make feature discrimination easier. We apply de-correlation stretching function. The number of colour bands, in the image is taken three. But we can apply de-correlation stretching regardless of the number of colour bands. The original colour values of the image are mapped to a new set of colour values with a wider range. The colour intensities of each pixel are transformed into the colour eigen space of the correlation matrix, stretched to equalize the band variances, and then transformed back to the original colour bands. To define the band wise statistics, we can use the entire original image, with the subset option, or any selected subset of it.

B. *K-means clustering*

k-means clustering is the one of the clustering method of data clustering. Data clustering is the process of identifying natural groupings or clusters within multidimensional data based on some similarity measure (e.g.Euclidean distance) [Jain *et al*. 1999; Jain *et al*. 2000]. It is an important process in image classification. Data clustering is of three types Hierarchical clustering, Partitional clustering, spectral clustering. Partitional clustering is of again classified three types K-means clustering, Fuzzy C-means clustering, and quality threshold clustering. Here only we discuss k-means clustering.

The k-means clustering is belongs to unsupervised image classification. The K-means clustering is also known as C-means clustering has been applied to a variety of areas, including image and speech data compression. The K-means method aims to minimize the sum of squared distances between all points and the cluster center. K-means clustering is to group a number of data vector into a predefined number of cluster[9].
X_j = no. of vectors (j=1,2……..n)
G_i =no. of groups (i=1,2……...c)
And find the cluster center in each group by using the cost function equation

$$J = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} \left(\sum_{k,xk \in Gi} \|X_k - c_i\|^2 \right) \quad \ldots\ldots\ldots\ldots (1)$$

Where X_k is the vector of group j and c_i is the cluster center and Where,
$J_i = \sum_{k,xk \in Gi} \|X_k - c_i\|^2$ is the cost function within group i. Thus, the value of J_i depends on the geometrical properties of G_i and the location of c_i.

Step 1: Initialize the cluster center c_i, i = 1,……,c. This is typically achieves by randomly selecting c points from among all of the data points.

Step 2: Determine the membership matrix U_{ij} by using the Equation

$$U_{ij} = \begin{cases} 1 \text{ if} \|X_j - c_i\|^2 \leq \|X_j - c_k\|^2, \text{for each } k \neq I \end{cases}$$

0 otherwise

Step3: Compute the cost function according to Equation (1) .Stop if either it is below a certain tolerance value of its improvement over previous iteration is below a certain threshold.

Step 4: Update the cluster centers according to Equation

$c_i = \frac{1}{|G_i|} = \sum_{k,xk \in Gi} X_k,$

where $|G_i|$ is the size of G_i , or $|G_i| = \sum^{n}_{i=1} \underline{U}_{ij}$.

Go to step 2.

C. Data base colour image segmentation implementing k-means clustering-

1. Read the data base colour image one by one.
2. For colour separation of an image apply the Decor relation stretching.
3. Now, convert one by one data base colour image from RGB colour to L*a*b colour.
   Here, The L*a*b* space consists of a luminosity layer 'L*', chromaticity-layer 'a*' indicating where colour falls along the red-green axis, and chromaticity-layer 'b*' indicating where the colour falls along the blue-yellow axis. All of the colour information is in the 'a*' and 'b*' layers. We can measure the difference between two colours using the Euclidean distance metric. Convert the image to L*a*b* colour space.
4. Applying the k-means clustering, based upon the Euclidean distance classify the colours of each image in 'a*b*'.
5. Using k-means result labels every pixel in the image.
6. We will get three clusters of Black, White and gray segmented image at the output, which one create segmented data base image by colour.

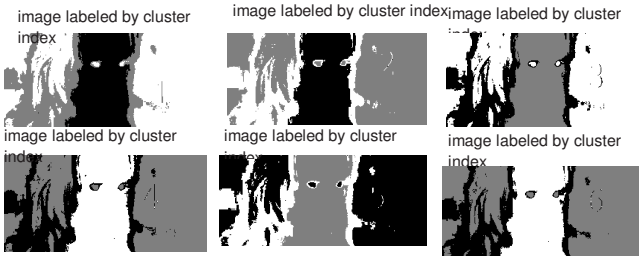
III. EXPERIMENTAL RESULT

In this paper by using MATLAB v7.8 [10] we point out the face region of the colour data base image. The images are different one from another according to their angle of view. The complete process and the standard results are summarized in sequence bellow[4,5].

[Fig:1(a) Original Data base image]

[Fig: 1(b) L*a*b Data base image]

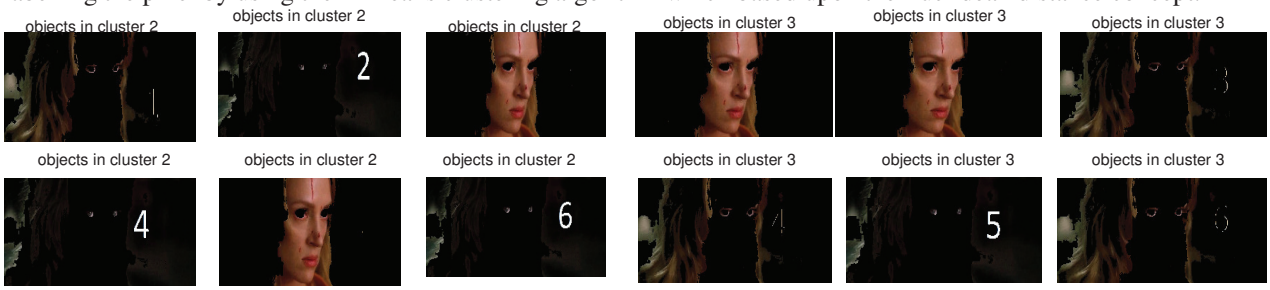
From the above figure we find out that the images are converted into L*a*b colour space image, means if we ignore the brightness of the original image we can easily find out the blue, pink and white colour which are called as L*a*b colour. Here the L*a*b* space consists of a luminosity layer 'L*', chromaticity-layer 'a*' indicating where colour falls along the red-green axis, and chromaticity-layer 'b*' indicating where the colour falls along the blue-yellow axis[2].



[Fig: 1(c) Image Labelled by cluster index]

[. . .Fig: 1(d) . . . . .] Show Black colour region of cluster index image as Cluster 1

From the above figure we find out the images are labelled as per the pixels present in the image and labelling the pixel by using the K-means clustering algorithm which based upon the Euclidean distance concept.



[. . . Fig: 1(e) . . . . .] Show Black colour region of cluster index image as

[. . . . . Fig: 1(f). . . . .] Show White colour region of cluster index image as cluster 2 cluster 3

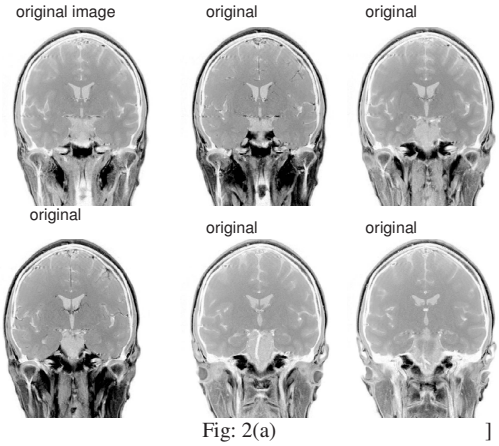
The fig.1(d) shows the face region only of the data base image 4 and 6 because it shows only the Black colour portion of the Labelled cluster index image and it takes as cluster1.

The fig.1(e) shows the face region only of the data base image 3 and 5 because it shows only the Gray colour portion of the Labelled cluster index image and it takes as cluster2.

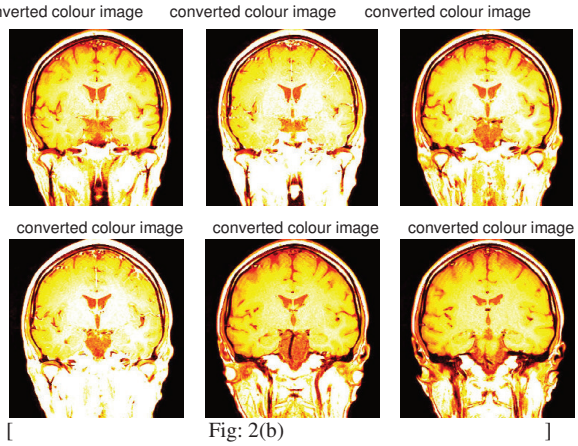
The Fig: 1(f) shows the face region only of the data base image 1 and 2 because it shows only the White colour portion of the Labelled cluster index image and it takes as cluster3.

So, From the above result discussion we have concluded that the data base image has categorized into three clusters that are Black, Gray and White. So as per the original image the clustering images are point out the face region of the data base colour image. The result is different in every run time of the programme because it uses K-means clustering method. So for each iteration the result is different.

Now we implemented the k-means clustering algorithm in another type data base. We take the Magnetic resonance image (MRI) which are in gray colour. We converted that gray colour image into RGB colour and isolated the cancer region from the data base brain image. Here we implemented the k-means clustering algorithm in the MR gray colour data base image and converted into colour image because generally, the colour MR images are not widely used.

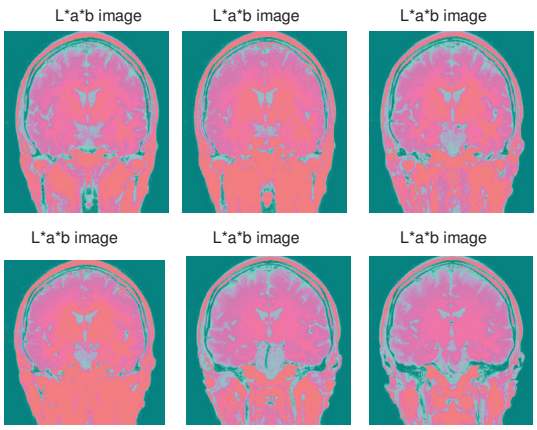


[Fig: 2(a)]
Original Data base MR brain image

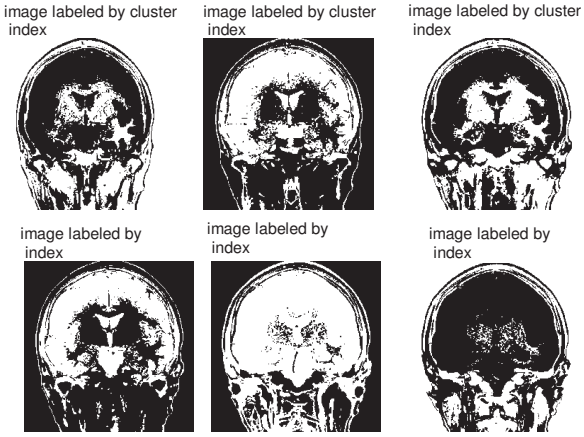


[Fig: 2(b)]
colour converted image

From the fig. 2(b) we find out the converted colour image of MR brain image data base.
colour converted image



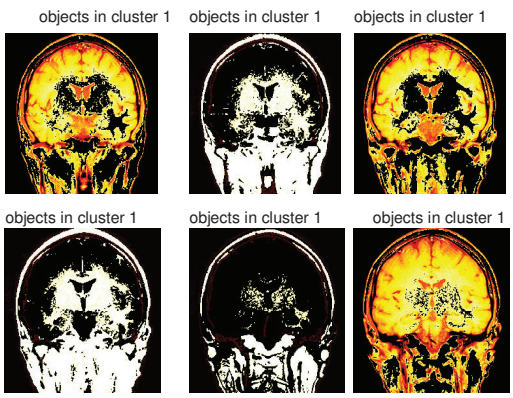
[Fig: 2(c)]
L*a*b Data base MR brain image



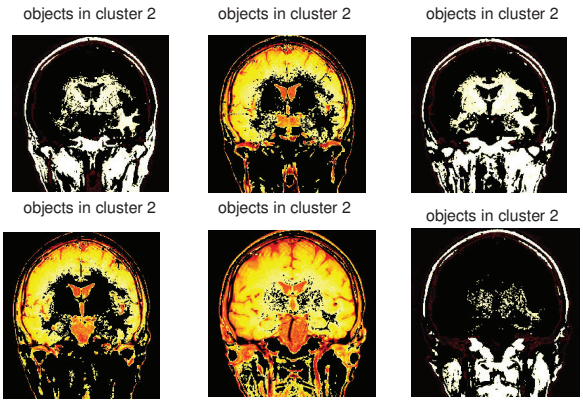
[Fig: 2(d)]
Image Labelled by cluster index

From the Fig: 2(c) we find out that the images are converted into L*a*b colour space image, the L*a*b colour space was already discussed in the Fig1 (b).

From the Fig: 2(d) we find out the images are labelled as per the pixels present in the image and labelling the pixel by using the K-means clustering algorithm.



[Fig: 2(e)]
Show Black colour region of cluster index image as cluster1



[Fig: 2(f)]
Show White colour region of cluster index image as cluster2

The Fig: 2(e) dominatingly shows the infected region of cancer cell of the data base image 1,3 and 6 because it shows only the Black colour portion of the Labelled cluster index image and it takes as cluster1.

The Fig: 2(f) dominatingly shows the infected region of cancer cell of the data base image 2,4 and 5 because it shows only the white colour portion of the Labelled cluster index image and it takes as cluster2.

From the above discussion we find out that by converting the gray colour MR image into colour image we can isolated the brain cancer cell and some important nerves region, which are connected to the brain. If we take the total data base image of brain which have taken in different point of view then it will easy for us to find out the cancer cell spread region in brain.

IV. CONCLUSION

Using colour based image segmentation; it is possible to reduce the computational cost avoiding feature calculation for every pixel in the image. Although the colours is not frequently used for image segmentation, it gives a high discriminative power of regions present in the image. As we seen above, by implementing the algorithm at first figure we point out the face region of the data base image and when we use the same algorithm in MRI data base cancer cell image, this result isolate the infected area of cancer cell from the converted colour data base cancer cell image that images were different according to their point of view of a persons cancer cell. Though this experiment uses the MATLAB software to solve its problem so effectively it has work done in a very few of cost. But still it is not clearly view of the cancer cell region from the MR images, so for this overcomes for a more accurate result we can use the optimization techniques like PSO (Particle swarm optimization)with the k-means clustering algorithm which is called as Hybrid K-means +PSO clustering method [3].

REFERENCES

[1] Image Segmentation for the Purpose Of Object-Based Classification, by Ahmed Darwish 2003 IEEE,pp 2039-2041
[2] Colour based image segmentation using k-means clustering by Anil Z chitade published in IJEST,vol.2,2010.
[3] Brain tumor detection of MR images based on colour converted hybrid PSO+Kmeans clustering segmentation published in European jornal of scientific research vol.70(2012).
[4] Data base image classification using k-means clustering & PSO by Madhusmita Sahu published in IJSAA journal, vol.2, may 2012.
[5] A colour image segmentation algorithm based on region growing by **Jun Tang** IEEE, vol 6, 2010, pp 634-637.
[6] Evaluation of colour image segmentation hierarchies by **Darren MacDonald** Proceeding of the 3rd Canadian conference on computer and Robot Vision, IEEE, 2006.
[7] Digital image processing by R.C.Gonzalez and R. E. Woods,chapter 10, paper 567-630 2nd edition.
[8] Image Segmentation via Adaptive –Mean Clustering and Knowledge-Based Morphological Operations with Biomedical Applications by Chang Wen Chen, Jiebo Luo, and Kevin J. Parker ,published in IEEE transaction on IMAGE PROCESSING, VOL. 7, NO. 12, DECEMBER 1998.
[9] Neuro-Fuzzy and soft computing Text book by jyh-shubg Roger jang , page no-450.
[10] R. C. Gonzalez and R. E. Woods, "Digital Image Processing using MATLAB",2nd edition.
[11] www.imageprocessingplace.com
[12] www.mathworks.com
[13] www.wikipedia.org
[14] http://ieeexplore.ieee.org/Xplore/guesthome.js
[15] http://www.sciencedirect.com