Brain Tumor Segmentation on MRI Brain image Using PSO and K-means Algorithms

J Thimmiaraja and Dr. N Senthilkumaran

Department of Computer Science and Applications
Gandhigram Rural Institute – Deemed to be University
Gandhigram-624302, Tamil Nadu, India

Abstract : ***Medical image processing is the most emerging field now a day. The Processing of MRI images is one of the parts of this field. Image enhancement technique is used to improve the contrast and removal of noise if present. We improve contrast by using histogram equalization, after the image passed through a median filter the impulse noises present in the images are also enhanced, Such that the resulting image is better than the original image. A K-means Clustering Algorithm Based on the Particle Swarm Optimization Algorithm (PSO) is refined in this paper. A number of clusters specified by a user to find the centroids in the algorithm, where all clusters associate together with related image primitives. In the experimentation, the performance are analyzed and the results compared by using various measures***

IndexTerms - Histogram equalization (HE), median filter, Particle Swarm Optimization (PSO), K-means.

I. INTRODUCTION

Image processing is a vast and demanding area and applications used in various fields like medical images, satellite images and also in industrial applications. Brain tumor detection plays an important and essential role in many medical imaging applications. There are different image modalities are used by the medical experts for proper treatment. The Magnetic Resonance Imaging (MRI) is one of the widely used imaging modality for brain tumor identification [1].

The main purpose of image enhancement is to bring out the hidden part in an image or to enhance the low contrast image. The quality of the image gets improved by contrast manipulation. A very popular performance for contrast enhancement is Histogram Equalization (HE) [1]. Image enhancement can be done by improving contrast and removing noise from the image or if image is blurred then we remove blurring also [2].

It is necessary to smooth the noisy signals to preserve the edge information. The most commonly used smoothing technique is median filtering. The median of a group, containing an odd number of elements, is defined as the middle element, when the elements of the group are sorted. The median computed at this operation is called the running or the moving median [3]. Since the size of the window is constant, the number of incoming elements is equal to the number of outgoing elements.

Clustering is the process of grouping together similar multi-dimensional data vectors into **a** number of clusters. The PSO algorithm has several parameters to be adjusted by empirical approach, if these parameters are not appropriately set, the search will become very slow near the global optimum. The K-means algorithm, on the contrary, has a strong ability to find local optimistic result, but its ability to find the global optimum is weak [4]. The K-means algorithm clusters a group of data vectors into a predefined number of clusters. It starts with a random initial cluster center and keeps reassigning the data objects in the dataset to cluster centers based on the similarity between the data object and the cluster center [5].

This paper is organized as follows. Section II discuss about the Histogram Equalization, Median filtering, PSO, and K-Means. Section III discusses the Analysis of experiments and results for the diagrams display the methods. Finally, Section IV presents the conclusion.

II. TERMS ADOPTED

A. Histogram equalization

Histogram equalization is a technique for adjusting image intensities to enhance contrast. The histogram of an image mostly represents the comparative frequency of occurrence of the various gray levels in the image [1][7].

$$\mathrm{p_n} = \frac{\text{no. of pixels with intensity } \mathrm{n_k}}{\text{total no. of pixels n}} \quad n = 0,1..L-1 \tag{1}$$

Consider an image with gray levels in the range [0, L −1], Probability Distribution Function of the image can be computed as:

$$p_{(r_k)} = \frac{n_k}{n} \quad n = 0,1,...L-1 \tag{2}$$

Where, r_k is the k^{th} gray level and n_k is the number of pixels in the image having gray level rk.

Histogram Equalization (HE) method has two main disadvantages which affect efficiency of this method. Histogram Equalization (HE) assigns one gray level into two different neighbor gray levels with different intensities. If most of an image includes a gray level, Histogram Equalization (HE) assign a gray level with higher intensity to that gray level and gives washed out appearance to the resultant image [1]. HE can initiate a considerable change in brightness of an image obtain a maximum value of the uniformly distributed image.

B. Meadian filtering

Median filtering is a nonlinear filtering technique; it is used to separate noise from medical images. It is widely used technique; it is very efficient in removing noise while preserving edges. The median filter works pixel by pixel by moving through the image, the neighboring pixels changing each value with the median value. The pattern of neighbours is called the "window", which slides, pixel by pixel over the complete image. The median is calculated by first sorting all the pixel values from the window into numerical order, and then replacing the pixel being considered with the median value. The window size is constant; the number of incoming elements is equal to the number of outgoing elements. The dimensions of the filter mask must be odd. Mask sizes are 3x3, 5x5 or 7x7. Minimum mask size is preferred in many cases. In this paper, the mask size is 3x3 [3].

C. Particle swarm optimization

PSO was originally developed by Eberhart and Kennedy in 1995, and was inspired by the social behavior of a flock of birds. In the PSO algorithm, the birds in a flock are symbolically represented as particles. These particles can be considered as simple agents "flying" through a problem space. A particle's location in the multi-dimensional problem space represents one solution for the problem. When a particle moves to a new location, a different problem solution is generated. This solution is evaluated by a fitness function that provides a quantitative value of the solution's utility [7] [8].

The velocity and direction of each particle moving along each dimension of the problem space will be altered with each generation of movement. In combination, the particle's personal experience, P_{id} and its neighbors' experience, P_{gd} influence the movement of each particle through a problem space. The random values $rand_1$ and $rand_2$ are used for the sake of completeness, that is, to make sure that particles explore a wide search space before converging around the optimal solution. The values of c1 and c2 control the weight balance of P_{id} and P_{gd} in deciding the particle's next movement velocity. Mathematically, given a multi-dimensional problem space, the i^{th} particle changes its velocity and location according to the following equations [7] [8].

$$V_{id} = W * V_{id} + C_1 * rand_1 * (P_{id} - X_{id}) + C_2 * rand_2 * (P_{gd} - X_{id}) \quad (3)$$

$$X_{id} = X_{id} + V_{id} \quad (4)$$

At every generation, the new location of the particles computed by adding the velocity, v_{id}, current velocity to its location, $x_{id.}$ where w denotes the inertia weight factor; c_1 and c_2 are constants and are known as acceleration coefficients; p_{id} is the location of the particle that experiences the best fitness value; d denotes the dimension of the problem space; p_{gd} is the location of the particles that experience a global best fitness value; $rand_1$, $rand_2$ are random values in the range of (0, 1).

Equation 3 requires each particle to record its current coordinate X_{id}, its velocity V_{id} that indicates the speed of its movement along the dimensions in a problem space and the coordinates P_{id} and P_{gd} where the best fitness values were computed. The best fitness values are updated at each generation, based on Eq. 5,

$$P_i\,(t+1) = \begin{cases} P_i(t) & f\big(X_i(t+1)\big) \le f\big(X_i(t)\big) \\ X_i(t+1) & f\big(X_i(t+1)\big) > f\big(X_i(t)\big) \end{cases} \quad (5)$$

where the symbol f denotes the fitness function; $P_i(t)$ stands for the best fitness values and the coordination where the value was calculated; and t denotes the generation step.

It is possible to view the clustering problem as an optimization problem that locates the optimal centroids of the clusters rather than finding an optimal partition. This view offers us a chance to apply PSO optimal algorithm on the clustering solution [8].

D. K-means algorithm

The K-means clustering data vectors group's into a predefined number of clusters, similarity measure based on Euclidean distance. Inside a cluster, Data vectors have small Euclidean distances from one another, and combined with one centroid vector, which perform the "midpoint" of that cluster. The centroid vector is the mean of the data vectors that belong to the corresponding cluster. The k-means algorithm attempts to discover the best points in space as the cluster centroids[10].

The standard K-means algorithm is summarized as follows:

(a) For each data vector, assign the vector to the class with the closest centroid vector, where the distance to the centroid is determined using

$$d\big(z_p, m_j\big) = \sqrt{\sum_{k=1}^{N_d} (z_{pk} - m_{jk})^2} \quad (6)$$

where k subscripts the dimension.

(b) Recalculate the cluster centroid vectors, using until **a** stopping criterion is satisfied

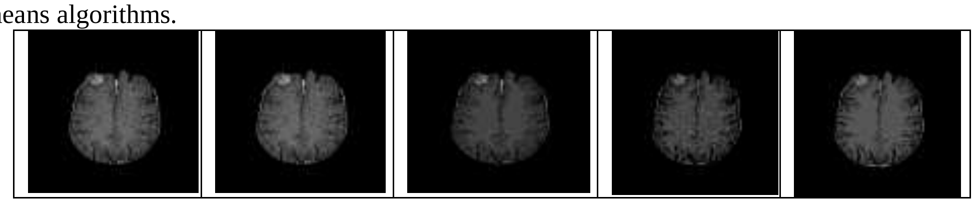
$$m_j = \frac{1}{n_j} \sum_{\forall z_p \in c_j} z_p \quad (7)$$

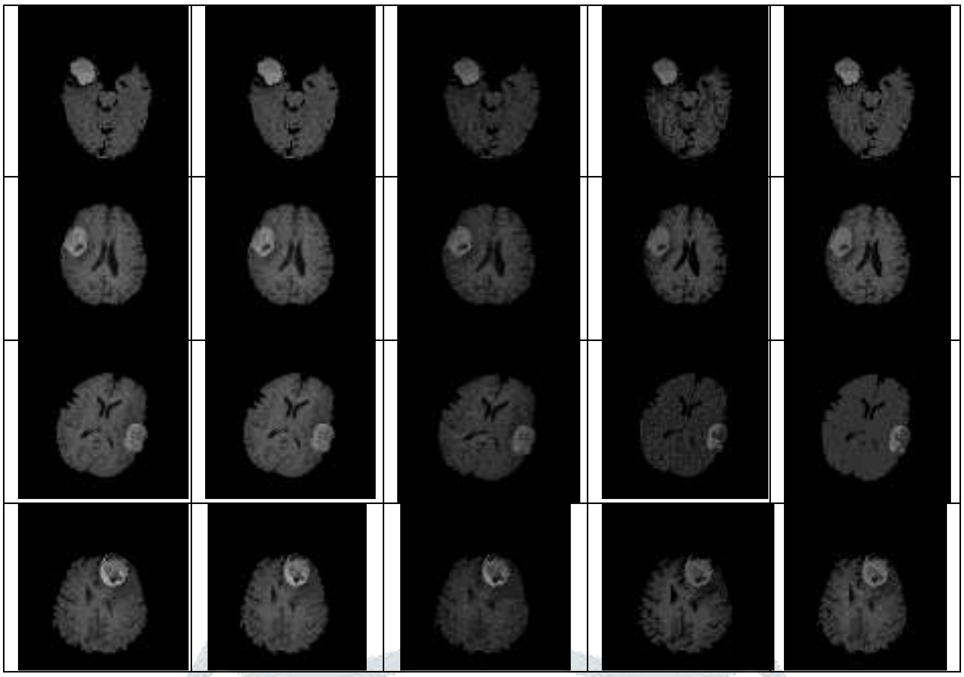
N_d denotes the input dimension, i.e. the number of parameters of each data vector; N_o denotes the number of data vectors to be clustered; N_c denotes the number of cluster centroids (as provided by the user), i.e. the number of clusters to be formed; z_p denotes the p-th data vector; m_3 denotes the centroid vector of cluster j; n_j is the number of data vectors in cluster j; C_j is the subset of data vectors that form cluster [9].

The K-means clustering process can he stopped when any one of the following criteria are satisfied: when the maximum number of iterations has been exceeded, when there is little change in the centroid vectors over **a** number of iterations or when there are no cluster membership changes. For the purposes of this study, the algorithm is stopped when a user-specified number of iterations have been exceeded [10].

III. RESULT AND DISCUSSION

In this paper, images are enhanced by histogram equalization and median respectively. Then the enhanced and filtered image compared between the pso and k-means algorithms.





(a) (b) (c) (d) (e)

Fig. 1 Results of MRI brain images. (a) Original, (b) Histogram Equalization, (c) Median Filter, (d) PSO, (e) K-means

In order to evaluate clustering algorithm performance, using two experiments using synthetic and real data. In these experiments we compare the method with stand alone PSO clustering and k-means clustering. The performance evaluated of each method using both error rate and mean square error criterion. The clusters are well separated, the result of clustering for PSO is better than kmeans [9].

Table 1. Information for sysnthetic and real datasets

		PSO	KMeans
Image1	Error rate	4.7%	5%
	MSE	27.21	28.13
Image2	Error rate	3.6%	3.5%
	MSE	28.24	29.12
Image3	Error rate	4.1%	4.8%
	MSE	24.29	26.24
Image4	Error rate	4.3%	4.9%
	MSE	28.29	29.27
Image5	Error rate	3.7%	4%
	MSE	24.28	28.11

IV. CONCLUSION

In this paper, images are enhanced by histogram equalization respectively. Then a comparison is made between the two clustering algorithms based on the PSO and Kmeans clustering were compared, which showed that the PSO approaches have better convergence to lower quantization errors, and in general, larger inter-cluster distances and smaller intra-cluster distances. The K-means algorithm is applied for refining and generating the final result. The results illustrate that the PSO algorithm can generate more compact clustering results than the Kmeans algorithm to found the better results from the MRI brain Image.

REFERENCES

[1] Senthilkumaran N, Thimmiaraja J, "Histogram Equalization for Image Enhancement Using MRI brain images", World Congress on Computing and Communication Technologies, IEEE, 2014.

[2] Deepika Bansal, Dr. Renu Bagoria, Vinod Sharma, "Digital Image Enhancement by Improving Contrast, Removal of Noise and Motion Blurring", International Journal of Innovative Research in Science, Engineering and Technology, 2015.

[3] P. Jagatheeswari , S.Suresh Kumar, M. Rajaram, "Contrast Enhancement for Medical Images Based on Histogram Equalization Followed by Median Filter", International Conference on Man-Machine System 2009.

[4] Alireza Ahmadyfard, Hamidreza Modares, "Combining PSO and k-means to Enhance Data Clustering", IEEE, 2008.

[5] Xiaohui Cui and Thomas E. Potok, "Document Clustering Analysis Based on Hybrid PSO+K-means Algorithm", Journal of Computer Sciences, 2005.

[6] Annadurai s, Shanmugalakshmi R, "Fundamentals of digital image processing", Pearson Education 2007.

[7] Kennedy J., Eberhart R. C., Shi Y., "Swarm Intelligence", New York, 2001.

[8] Xiaohui Cui, Thomas E. Potok, Paul Palathingal, "Document Clustering using Particle Swarm Optimization", IEEE, 2005.

[9] Alireza Ahmadyfard, Hamidreza Modares, "Combining PSO and k-means to Enhance Data Clustering", IEEE, 2008.

[10] DW van der Merwe, AP Engelhrecht, "Data Clustering using Particle Swarm Optimization", IEEE, 2003.