www.ijarcsse.com
Volume 4, Issue 8, August 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

# PSO Based Clustering for MR Image Segmentation

**Rajlaxmi Garg**
Mahakal Institute of Technology,
Ujjain, India

**Kshitij Pathak**
Mahakal Institute of Technology,
Ujjain, India

***Abstract— Image segmentation refers to the procedure of partitioning a digital image into various sections. It would be highly useful to employ image segmentation in medical images because it could assist in determining tumors in various body parts.This paper presents an image segmentation method,which is based on Particle Swarm Optimization(PSO). The performance of each particle in PSO is measured using a fitness function.Davies–Bouldin clustering validity index(DBI) is used as the fitness function, which is to be minimized in order to produce optimal clustering.It is compared with other image clustering technique K-means using IQI(Image Quality Index) and DBI(Davies–Bouldin Index).Experimental results show that PSO based image clustering technique produces better results in all measured criterion as compared to K-means.***

***Keywords—Image segmentation, Clustering ,K-means, Particle Swarm Optimization,Davies-Bouldin Index.***

## I. INTRODUCTION

### *A. Image Processing*

An image is an array, or a matrix, of square pixels (picture elements) arranged in columns and rows. In a (8-bit) gray scale image, each picture element has an assigned intensity that ranges from 0 to 255. A gray scale image is what people normally call a black and white image, but the name emphasizes that such an image will also include many shades of gray. Each pixel has a value from 0 (black) to 255 (white) fig 1.1[11]. The possible range of the pixel values depend on the color depth of the image, here 8 bit = 256 tones or gray scales

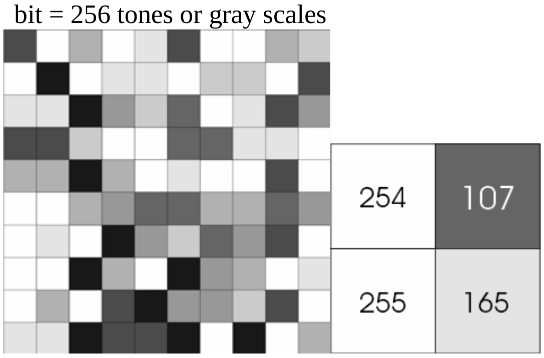



FIGURE 1.1 IMAGE IN THE FORM OF GRAY SCALE

Image processing is a technique in which the data from an image are digitized and various mathematical operations are applied to the data, generally with a digital computer, in order to create an enhanced image that is more useful or pleasing to a human observer, or to perform some of the interpretation and recognition tasks usually performed by humans. Image processing also known as picture processing, deals with images which are two-dimensional entities captured electronically through a scanner or camera system that digitizes the spatially continuous coordinates to a sequence of 0's and 1's. These spatially distinct points holds a number that denotes gray level or color for it, and can be conveniently fed to a digital computer for processing. Image processing algorithm takes as input, an image or a sequence of images and produces an output, which may be a modified image and/or a description of the input image contents or an output based on some analysis of the input image[15]. Some of the image processing operations are Enhancement programs, Math processes programs, Noise filters, Edge detection and Image segmentation etc.

### *B. Image Segmentation*

Tumors are the most common disease among people in this world. This fact has led researchers to continue studying how to treat and detect tumors in human body. Magnetic resonance imaging (MRI) has become a great addition to medical science as an imaging technique dedicated to providing brain tumors, breast tumors screening[3]. But this screening image output must be prominent, so as, it would be easy to detect tumors correctly[16].Segmentation of images can provide this outcome. Segmentation is the process of partitioning a digital image into multiple segments (sets of pixels). It refers to the process of partitioning the image into multiple segments (some non-overlapping meaningful homogeneous regions)[13]. The goal of segmentation is to simplify and/or change the representation of image into something that is more meaningful and easier to analyze.Some of the practical applications of image segmentation are[1]:

1. Medical imaging
   (a) Locate tumors and other pathologies
   (b) Measure tissue volumes
   (c) Computer-guided surgery
   (d) Diagnosis
   (e) Treatment planning
   (f) Study of anatomical structure
2. Locate objects in satellite images (roads, forests, etc.)
3. Face recognition
4. Iris recognition
5. Fingerprint recognition
6. Traffic control systems

Segmentation can be applied to all the above different applications to get a useful and beneficial output result with the help of several methods. Different methods of segmentation as listed below:

1. Thresholding
2. Clustering
3. Histogram based method
4. Edge detection
5. Region growing method.
6. Segmentation through neural networks etc.

Here we will use clustering methods to segment images. The methods used will be K-means and PSO based clustering. We will then compare and analyze the results of different clustering methods.

### *C. Clustering*

Clustering is considered as the most important unsupervised learning problem[17]. It deals with finding a structure in a collection of unlabeled data. A loose definition of clustering could be "the process of organizing objects into groups whose members are similar in some way". A cluster is therefore a collection of objects which are "similar" between them and are "dissimilar" to the objects belonging to other clusters[2]. We can show this with a simple graphical example(see fig1.2).

Figure 1.2: Data samples clustered around three different centers (indicated by large black dots). Coloring indicates grouping.

In fig 1.2 we can easily identify the 3 clusters into which the data can be divided; the similarity criterion is distance: two or more objects belong to the same cluster if they are "close" according to a given distance (in this case geometrical distance). This is called distance based clustering.
A mathematical definition of a cluster is follows
Let $X$ be a set of data, that is

$$X=\{x_1,x_2,\ldots\ldots,x_n\}$$

A so called *m-clustering* of $X$ is its partition into $m$ parts (clusters) $C_1,C_2,\ldots.C_m$ so that

1. None of the clusters is empty.
2. Every sample belongs to a cluster.
3. Every sample belongs to a single cluster (crisp clustering).

Various clustering techniques are in use out of which K-means will be used for comparative analysis.
Partitional clustering algorithms(described in 1.3) try to minimize certain criteria(e.g. a squared error function),therefore ,they can be treated as an optimization problem.Particle Swarm Optimization[7] has gained huge popularity as a naturally inspired optimization tool in recent times, and hence we use it in our work.

### *D. Particle Swarm Optimization*

Particle Swarm Optimization was first introduced by Dr. Russell C. Eberhart and Dr.James Kennedy in 1995[7].In this system, a swarm of individuals (called *particles*) fly through the search space. Each particle represents a candidate solution to the optimization problem. Particles have two primary operators: Velocity update and Position update. The position of a particle is influenced by the best position visited by itself (i.e. its own experience) and the position of the best particle in its neighbourhood (i.e. the experience of neighbouring particles). When the neighbourhood of a particle is the entire swarm, the best position in the neighbourhood is referred to as the global best particle, and the resulting algorithm is referred to as a *gbest* PSO.When smaller neighbourhoods are used, the algorithm is generally referred to as a *lbest* PSO [10]. The performance of each particle (i.e. how close the particle is from the global optimum) is measured using a fitness function that varies depending upon the optimization problem.

Each particle i in the swarm maintains the following information:

$x_i$,the current position of the particle;
$v_i$,the current velocity of the particle; and
$y_i$, the personal best position of the particle.

The personal best position of a particle i is the best position that the particle has achieved so far,i.e. a position that produced the highest value of fitness function for that particle[5].

Let f denote the objective function , then the personal best of a particle at a given time t is updated as:

$$y_i(t+1) = \begin{cases} y_i(t) & \text{if } f(x_i(t+1)) \geq f(y_i(t)) \\ & \text{else} \\ x_i(t+1) & \text{if } f(x_i(t+1)) \geq f(y_i(t)) \end{cases} \tag{1}$$

The global best position of a particle in swarm is denoted as ŷ,then

$$\begin{aligned} \hat{y}(t) &\in \{y_0, y_1, \ldots y_s\} \\ &= \min\{f(y_0(t)), f(y_1(t)), \ldots\ldots f(y_s(t))\} \end{aligned} \tag{2}$$

where s is the total number of particles in the swarm.

For each iteration of PSO algorithm, $v_i$ and $x_i$ are updated as follows:

$$\begin{aligned} v_i(t+1) = wv_i(t) + c_1 r_1(t)(y_i(t) - x_i) \\ + c_2 r_2(t)(\hat{y}(t) - x_i(t)) \end{aligned} \tag{3}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{4}$$

where :-

**$c_1$** is the cognition parameter which represents how much the particle trusts its own past experience,
**$c_2$** is the social parameter which represents how much the particle trusts the swarm,
**$r_1$** and **$r_2$** are random numbers
and ***w*** is the inertia weight.
Velocity can be limited to Vmax**.**
$c1+c2\leq 4$ and $c1=c2$ for best results as defined by Kennedy [7]
$c1=c2=1.49$ &$w=0.72$ used by Omran [5]

The Davies–Bouldin clustering validity index is used as the fitness function of the PSO, which is to be minimized to achieve optimal segmentation.A lower value of DB Index is considered to represent good clustering[9].

## II. LITERATURE SURVEY

In [1]Many image segmentation techniques are available in the literature. Some of these techniques use only the gray level histogram, some use spatial details while others use fuzzy set theoretic approaches. Most of these techniques are not suitable for noisy environments. Neural network architectures which help to get the output in real time because of their parallel processing ability, have also been used for segmentation and they work fine even when the noise level is very high. The literature on color image segmentation is not that rich as it is for gray tone images. This paper critically reviews and summarizes some of these techniques. Attempts have been made to cover both fuzzy and non-fuzzy techniques including color image segmentation and neural network based approaches. Adequate attention is paid to segmentation of range images and magnetic resonance images. It also addresses the issue of quantitative evaluation of segmentation results.
In [2] this article, the performance of three clustering algorithms, hard K-Means, single linkage, and a simulated annealing (SA) based technique is evaluated, in conjunction with four cluster validity indices, namely Davies-Bouldin index, Dunn's index, Calinski-Harabasz index, and a recently developed index I. Based on a relation between the index I

and the Dunn's index, a lower bound of the value of the former is theoretically estimated in order to get unique hard K-partition when the data set has distinct substructures. The effectiveness of the different validity indices and clustering methods in automatically evolving the appropriate number of clusters is demonstrated experimentally for both artificial and real-life data sets with the number of clusters varying from two to ten. Once the appropriate number of clusters is determined, the SA-based clustering technique is used for proper partitioning of the data into the said number of clusters.
In [3]Image segmentation is an indispensable process in the visualization of human tissues, particularly during clinical analysis of magnetic resonance (MR) images. Unfortunately, MR images always contain a significant amount of noise caused by operator performance, equipment, and the environment, which can lead to serious inaccuracies with segmentation. A robust segmentation technique based on an extension to the traditional fuzzy c-means (FCM) clustering algorithm is proposed in this paper.. Simulated and real brain MR images with different noise levels are segmented to demonstrate the superiority of the proposed technique compared to other FCM-based methods. This segmentation method is a key component of an MR image-based classification system for brain tumors, currently being developed. compared to other FCM-based methods. This segmentation method is a key component of an MR image-based classification system for brain tumors, currently being developed.
In [4] Medical image databases are a key component in future diagnosis and preventive medicine. Automatic clustering of medical images plays an important role for structuring of a given medical database as well as for searching and retrieval of biomedical images.This paper introduces a new approach for efficient clustering of x-ray images based on various levels of image features. Initially, for each given x-ray image, features have been extracted from three different levels, namely, global, local and pixel. Finally, a new approach, a combination of k-means and hierarchical clustering techniques, has been applied in order to cluster x-ray images.
In [5]An image clustering method that is based on the Particle Swarm Optimization (PSO) is developed.The algorithm finds the centroids of a user specified number of clusters,where each cluster groups together similar image primitives.The proposed image classifies is applied to various synthetic,MRI and satellite images.Experimental results show that the PSO image classifier works better than a conventional image classifier in all measured criteria.The influence of different values of PSO control parameters is also illustrated.
In [6] paper, a new universal objective image quality index is proposed, which is easy to calculate and applicable to various image processing applications. Instead of using traditional error summation methods the proposed index is designed by modeling any image distortion as combination of three factors: loss of correlation, luminance distortion, and contrast distortion. Although the new index is mathematically defined and no human visual system model is explicitly employed, the experiments on various image distortion types indicate that it performs significantly better than the widely used distortion metric mean squared error.
In [7] a concept for the optimization of nonlinear functions using particle swarm methodology is introduced The evolution of several paradigms is outlined, and an implementation of one of the paradigms is discussed. Benchmark testing of the paradigm is described, and applications, including nonlinear function optimization and neural network training, are proposed. The relationships between particle swarm optimization and both artificial life and genetic algorithms are described
This paper [8] presents the use of Computer Aided techniques in agriculture. This paper reviews some of the important segmentation based algorithms and recent trends in image processing techniques which are applicable to crop quality evaluation and defect identification.. A case study of medical image segmentation which adopts similar methodology for the detection of diseases is discussed.. The paper strongly suggests the need to use similar Methodology of Information Technology for the agricultural domain

## III. PROBLEM STATEMENT

Given a set of gray scale MR images of brain tumor of size 256*256 and 512*512 and the number of clusters, proposed method aims to detect tumor by segmentation .The image is segmented using PSO based clustering with DB index as fitness function results are compared with K-means on parameters IQI and DBI.

## IV. K-MEANS ALGORITHM

C means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. This algorithm takes an input data(image) and the number of clusters to be constructed.Output produced is image segmented into number of clusters specified(see fig 3)[14].

It aims at minimizing an objective function which is defined below

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \qquad (5)$$

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 centres

The algorithm is composed of the following steps:

1. Place *C* points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the C centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

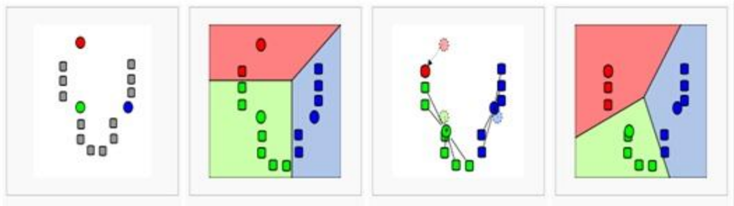


Fig 3.Steps of K-means Clustering

Although it can be proved that the procedure will always terminate, the C means algorithm does not necessarily find the most optimal configuration, corresponding to the global objective function minimum. The algorithm is also significantly sensitive to the initial randomly selected cluster centres. The C Means algorithm can be run multiple times to reduce this effect.

## V. DAVIES- BOULDIN INDEX

Davies Bouldin Index was introduced by David L. Davies and Donald W. Bouldin in 1979 [9].It is a metric for evaluating clustering algorithms.This index (Davies and Bouldin) is a function of the ratio of the sum of within cluster scatter to between-cluster separation.

$$DB = \frac{1}{n}\sum_{i=1}^{n} \max_{i \neq j} \left\{ \frac{S_n(Q_i) + S_n(Q_j)}{S(Q_i, Q_j)} \right\} \qquad (6)$$

Where, n- number of clusters

$S_n$ - average distance of all objects from the cluster to their cluster centre,

*S(Qi,Qj)* - distance between clusters centres.

Davies-Bouldin index will have a small value for a good clustering.Therefore it is required to minimize using PSO.

## VI. PROPOSED SOLUTION

In the context of image clustering individual particle in swarm represents N cluster centroids.That is each particle $x_i$ is constructed as $x_i=(m_{i,1},....m_{i,j},....m_{i,N})$ where $m_{ij}$ refers to the $j^{th}$ cluster centroid vector of the $i^{th}$ particle. The method used in this paper proposed in [5] and modified as follows.

### A. Algorithm

1.Initialize LB=0,UB=255,No of clusters N,swarm size=20.
2. Initialize each particle to contain N randomly selected cluster means.
3. For t=1 to $t_{max}$ or untill convergence achieved
  (a) for each particle $x_i$
    (i) for each pixel $Z_p$
- calculate $d(Z_p, m_{ij})$ for all cluster $C_{ij}$ using eq5
- assign $Z_p$ to $C_{ij}$ where
  $d(Z_p,m_{ij})=\min_{C_{i=1..N}} d(Z_p,m_{ij})$

    (ii) calculate the fitness using eq 6

  (b) Update personal best using eq 1
  (c) Update global best using eq 2
  (d) Update the cluster centre using eq 3&4.

## VII. IMPLEMENTATION

### A. Evaluation parameter

Image Quality Index is[6] easy to calculate and applicable to various image processing applications. Instead of using traditional error summation methods the proposed index is designed by modeling any image distortion as combination of three factors: loss of correlation, luminance distortion, and contrast distortion.

If $X=\{x_1,x_2,....,x_n\}$ and $Y=\{y_1,y_2,....,y_n\}$ are the original and test image signals, then the new quality index is defined as

$$Q = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x^2 + \sigma_y^2)[(\bar{x})^2 + (\bar{y})^2]},$$

Where $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ , $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ ,

$$\sigma_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2,$$

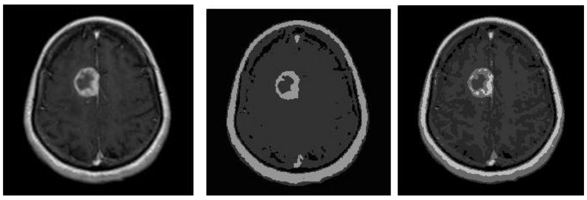
$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right).$$

The dynamic range of Q is [-1,1]. The best value 1 is achieved if and only if $y_i = x_i$ for all $i = 1,2,3,\ldots N$ . The lowest value value (-1) is achieved if $y_i = 2\bar{x} - x_i$ , for all $i = 1,2,3,\ldots N$.

This quality index models any distortion as a combination of three factors : loss of correlation, luminance distortion, and contrast distortion. In order to understand this , we rewrite the definition of *Q* as a product of three components

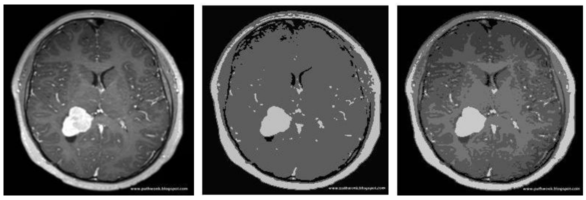
$$Q = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}. \qquad (7)$$

In this section experimental results on medical MR images and satellite images are shown .PSO is implemented in Matlab version 7.12.Where as the inbuilt function of K-means of Matlab 7.12 is used .Number of cluster was choosen 4 .MRI of brain tumor is collected from [12]



4.a 4.b 4.c

Fig 4. Segmentation of MR Image 1

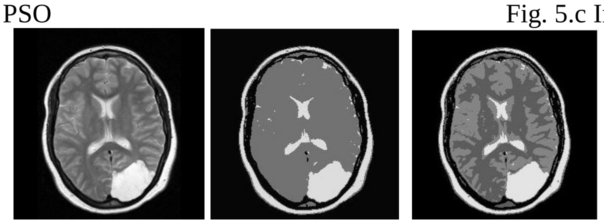


5.a 5.b 5.c

Fig 5. Segmentation of MR Image 2

Fig. 4.a Original gray scale image of human brain
Fig. 4.b Image segmented using k –means
Fig. 4.c Image segmented using PSO

Fig. 5.a Original gray scale image of human brain
Fig. 5.b Image segmented using k –means
Fig. 5.c Image segmented using PSO



6.a 6.b 6.c

Fig 6.Segmentation of MR Image 3

Fig. 6.a Original gray scale image of human brain
Fig. 6.b Image segmented using k –means
Fig. 6.c Image segmented using PSO

## VIII. RESULTS

The IQI(eq 7) and DB (Davies Bouldin) Index(eq 6) values obtained for K-means and PSO clustering algorithms is shown in the Table 1 & 2 below. Greater the value of IQI better is the quality of image and lesser the value of DB Index better is the clustering algorithm(see. Table 1&2).

Table 1. IQI value for different clustering algorithms

| Clustering Algorithm | Image1 | Image2 | Image3 |
|---|---|---|---|
| K-means | .291 | .537 | .407 |
| PSO | .471 | .575 | .588 |

Table 2. DB index value for different clustering algorithms

| Clustering Algorithm | Image1 | Image2 | Image3 |
|---|---|---|---|
| K-means | .250 | .291 | .269 |
| PSO | .235 | .252 | .213 |

## IX. CONCLUSION

The implementation of Particle Swarm Optimization for image segmentation is done using Davies-Bouldin Index as fitness function.A comparision is made with K-means and experimental results show that PSO based image segmentation technique fairs better in both criteria viz, IQI and DB Index.

## REFRENCES


[1] Pal, N.R., Pal, S.K "A Review on Image Segmentation Techniques ,Pattern Recognition 26, 1277–1294 (1993) .

[2] U. Maulik, S. Bandyopadhyay "Performance evaluation of some clustering algorithms and validity indices",in, IEEE Transactions on Pattern analysis and machine intelligence, Volume**:**24, pp 1650 – 1654, December 2002.

[3] Van Lung, Huynh Kim, and Jong-Myon, "A generalized spatial fuzzy c-means algorithm for medical image segmentation", In: Proc. of IEEE International Conference on Fuzzy Systems, 409-414, 2009

[4] Chhanda Ray, Krishnendu Sasmal "A New Approach for Clustering of X-ray Images" ,in proc, IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 4, No 8, July 2010

[5] M. Omran A. P. Engelbrecht Salman"Particle Swarm Optimization Method For Image Clustering", In Int. J. Patt. Recogn. Artif. Intell. 19, 297 (2005)

[6] Z. Wang, Alan C. Bovik "A Universal Image Quality Index",in proc. IEEE Signal Processing Letters, Vol. 9, No. 3, March 2002

[7] James Kennedy' and Russell Eberhart" Particle Swarm Optimization", In Proceedings of IEEE International Conference on Neural Networks , volume IV, pages 1942{1948, Perth, Australia, 1995.

[8] Dr.Hari Ramakrishna,T.S.Praveena,Prof. N.Rama Devi, "Segmentation Based Image Mining For Agriculture",In Proc. International Journal of Engineering Research & Technology (IJERT) Vol. 1 Issue 6, August – 2012.

[9] David L. Davies And Donald W.Bouldin,"A Cluster Separation Measure" IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. Pami-1, No. 2, April 1979.

[10] Y. Shi and R.C. Eberhart. "Parameter selection in particle swarm optimization". In Proceedings of Evolutionary Programming VII (EP98) , pages 591{ 600, 1998.}

[11] R.C.Gonzalez, Digital Image Processing,2$^{nd}$ ed Prentice hall,New Jersey.

[12] Sri Aurobindo Institute of Medical Sciences,Indore.

[13] H..P.Narkhede Review of Image Segmentation Techniques,In Proc of International Journal of Science and Modern Engineering (IJISME ) ISSN: 2319 - 6386, Volume 1, Issue 8, July 2013.

[14] A.K.Jain,M.N.Murty,P.J.Flynn,"Data Clustering: A Review",In Proc of ACM Computing Surveys, Vol. 31, No. 3, September 1999

[15] Richard O.Duda, Peter E. Hart, David G. Stork Pattern Classification 2$^{nd}$ ed Wiley Publication,oct2001

[16] M.C. Jobin Christ and R.M.S. Parvathi," Segmentation of Medical Image using Clustering and Watershed Algorithms",In proc of American Journal of Applied Sciences 8 (12): 1349-1352, 2011ISSN 1546-9239.

[17] Anil K. Jain ,Richard C. Dubes , Algorithms for Clustering Data, Prentice Hall Advanced Reference Series : March 1, 1988