Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 2, Issue. 7, July 2013, pg.200 – 206

SURVEY ARTICLE

A Survey of Image Segmentation Algorithms Based On Fuzzy Clustering

R. Ravindraiah[1], K. Tejaswini[2]
[1]Assistant Professor, Electronics and Communication Engineering,
Madanapalle Institute of Technology &science, Andhra Pradesh, India
[2]Student, Electronics and Communication Engineering,
Madanapalle Institute of Technology &science, Andhra Pradesh, India

[1] ravindra.ranga@gmail.com; [2] tejaswini4b7@gmail.com

Abstract— Medical image segmentation plays a vital role in one of the most challenging fields of engineering. Imaging modality provides detailed information about anatomy. It is also helpful in the finding of the disease and its progressive treatment. More research and work on it has enhanced more effectiveness as far as the subject is concerned. Different methods are used for medical image segmentation such as Clustering methods, Thresholding method, Classifier, Region Growing, Deformable Model, Markov Random Model etc. The main purpose of this survey is to provide a comprehensive reference source for the researchers involved in Fuzzy C Means based medical image processing. There are different types of FCM algorithms for medical image. Their advantages and disadvantages are discussed.

Key Terms: - Image segmentation; Medical Image Processing; Fuzzy C Means

I. INTRODUCTION

Segmentation is the process of separating a digital image into different regions which have similar properties such as gray level, colour, texture, brightness etc. So that the image can be more understandable and helpful to analyzing. On the basis of pixel intensity we can differentiate the boundaries of different objects. Segmentation identifies separate object within an image and also find boundary between different regions. Segmentation can be classified into two types: local segmentation and global segmentation. Local segmentation is small windows on a whole image and deal with segmenting sub image. Global segmentation deals with segmenting whole image. Global segmentation mostly deals with relatively large no of pixel. But local segmentation deal with lower no of pixel as compare to global segmentation. Image segmentation is one of the classical problems in image processing and computer vision. Using of Image segmentation we can able to understand the fundamental of digital image processing. Image segmentation is used to enhancement of image and also useful to different medical application. Image segmentation can also use for analysis of the image and further pre-processing of the image. After a segmentation process each phase of image treated differently. Now we are going through about medical images like Ultrasound Images (US) which is widely used today.

Medical images play a vital role in assisting health care providers to access patients for diagnosis and treatment. Studying medical images depends mainly on the visual interpretation of the radiologists. However, this consumes time and usually subjective, depending on the experience of the radiologist. Consequently the use of computer-aided systems becomes very necessary to overcome these limitations. Artificial Intelligence methods such as digital image processing when combined with others like machine learning, fuzzy logic and pattern recognition are so valuable in Image techniques. The computerization of medical image segmentation plays an important role in medical imaging applications. It has found wide application in different areas such as diagnosis, localization of pathology, study of anatomical structure, treatment planning, and computer-integrated surgery. However, the variability and the complexity of the anatomical structures in the human body have resulted in medical image segmentation remaining a hard problem.

Atherosclerosis is a specific form of arteriosclerosis (thickening & hardening of arterial walls) affecting primarily the intima of large and medium-sized muscular arteries and is characterized by the presence of fibrofatty plaques. The term atherosclerosis is derived from “athero” referring to the soft lipid-rich material in the centre of atheroma, and “sclerosis” referring to connective tissue in the plaques. Atherosclerosis is often called as arteriosclerosis.

Arteriosclerosis can occur in several forms, including atherosclerosis. Atherosclerosis is a heart disease, and is a type of arteriosclerosis or hardening of the arteries. Coronary artery mainly consists of three layers. The inner layer is intima, middle layer is media and the outer one is adventitia. An artery is made up of several layers: an inner lining called the endothelium, an elastic membrane that allows the artery to expand and contract, a layer of smooth muscle, and a layer of connective tissue. Arteriosclerosis is a broad term that includes a hardening of the inner and middle layers of the artery. It can be caused by normal aging, high blood pressure, mental stress, lack of inactivity and by diseases such as diabetes.

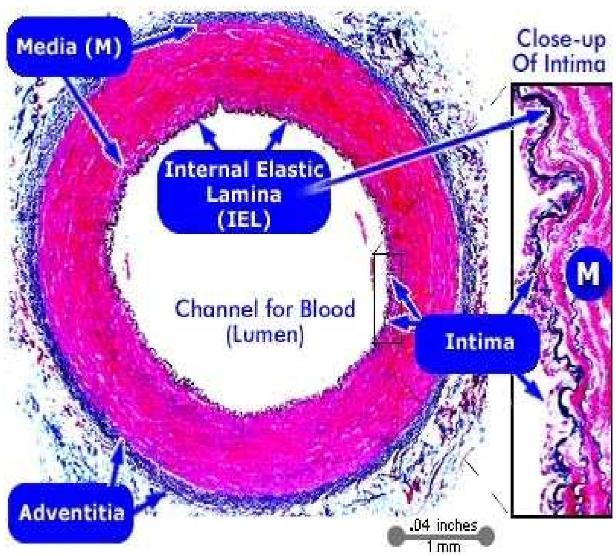


Figure: Cross section of coronary artery

Atherosclerosis is a type of arteriosclerosis that affects only the inner lining of an artery. It is characterized by plaque deposits that block the flow of blood. Plaque is made of fatty substances, cholesterol, waste products from the cells, calcium, and fibrin, a stringy material that helps clot blood. The plaque formation process stimulates the cells of the artery wall to produce substances that accumulate in the inner layer. Fat builds up within these cells and around them, and they form connective tissue and calcium. The inner layer of the artery wall thickens, the artery’s diameter is reduced, and blood flow and oxygen delivery are decreased. Plaques can rupture, causing the sudden formation of a blood clot (thrombosis). Atherosclerosis can cause a heart attack if it completely blocks the blood flow in the heart (coronary) arteries. It can cause a stroke if it completely blocks the brain (carotid) arteries. Atherosclerosis can also occur in the arteries of the neck, kidneys, thighs, and arms, causing kidney failure or gangrene and amputation.

Different types of Fuzzy C Means Algorithms

1. Fuzzy C Means(FCM)
2. Improved Fuzzy C Means(IFCM)
3. Possibilistic C Means(PCM)
4. Fuzzy Possibilistic C Means(FPCM)
5. Modified Fuzzy Possibilistic C Means(MFPCM)
6. Possibilistic Fuzzy C Means(PFCM)
7. Bias Corrected Fuzzy C Means(BCFCM)

8. Kernel Based Fuzzy C Means(KFCM)
9. Gaussian Kernel Based Fuzzy C Means(GKFCM)
10. Robust Kernel Based Fuzzy C Means(RKFCM)

II. FUZZY C MEANS ALGORITHM (FCM)

Fuzzy clustering is a powerful unsupervised method for the analysis of data and construction of models. In many situations, fuzzy clustering is more natural than hard clustering. Objects on the boundaries between several classes are not forced to fully belong to one of the classes, but rather are assigned membership degrees between 0 and 1 indicating their partial membership [1, 2]. Fuzzy c-means algorithm is most widely used. Fuzzy c-means clustering was first reported in the literature for a special case (*m*=2) by Joe Dunn in 1974. The general case (for any *m* greater than 1) was developed by Jim Bezdek in his PhD thesis at Cornell University in 1973. It can be improved by Bezdek in 1981. The FCM employs fuzzy partitioning such that a data point can belong to all groups with different membership grades between 0 and 1.

1. Initialize $U = [u_{ij}]$ matrix, $U^{(0)}$
2. at k-step: calculate the centers vectors $c^{(k)} = [c_j]$ with $U^{(k)}$ $c_{ij} = \frac{\sum_{j=1}^{n} u_{ij}^{m} x_j}{\sum_{j=1}^{n} u_{ij}^{m}}$ (1)
3. Update $U^{(k)}, U^{(k+1)}$
4. $d_{ij} = \sqrt{\sum_{i=1}^{n}(x_i - c_i)}$ (2)

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{kj}}\right)^{2/(m-1)}} \quad (3)$$

5. If $\| U(k+1) - U(k) \| < \epsilon$ then STOP; otherwise return to step 2.

Here m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster *j*, x_i is the *i*th of d-dimensional measured data, c_j is the d-dimension center of the cluster.

This algorithm works by assigning membership to each data point corresponding to each cluster center on the basis of distance between the cluster center and the data point. More the data is near to the cluster center more is its membership towards the particular cluster center. Clearly, summation of membership of each data point should be equal to one. After each iteration membership and cluster centers are updated according to the formula. The advantage of FCM is Unsupervised and Converges and limitations is Long computational time, Sensitivity to the initial guess (speed, local minima) and Sensitivity to noise and One expects low (or even no) membership degree for outliers (noisy points).

III. IMPROVED FUZZY C MEANS ALGORITHM (IFCM)

The improved FCM algorithm is based on the concept of data compression where the dimensionality of the input is highly reduced. The data compression includes two steps: quantization and aggregation [3]. The quantization of the feature space is performed by masking the lower 'm' bits of the feature value. The quantized output will result in the common intensity values for more than one feature vector. In the process of aggregation, feature vectors which share common intensity values are grouped together. A representative feature vector is chosen from each group and they are given as input for the conventional FCM algorithm. Once the clustering is complete, the representative feature vector membership values are distributed identically to all members of the quantization level. Since the modified FCM algorithm uses a reduced dataset, the convergence rate is highly improved when compared with the conventional FCM.

The improved FCM algorithm uses the same steps of conventional FCM except for the change in the cluster updation and membership value updation criterions. The modified criterions are showed below

$$c_{ij} = \frac{\sum_{j=1}^{n} u_{ij}^{m} x_j}{\sum_{j=1}^{n} u_{ij}^{m}}, \quad u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{kj}}\right)^{2/(m-1)}}$$

Where $d_{ij} = x_j - c_{ij}$ (4)

x = Reduced Dataset

IV. POSSIBILISTIC C-MEANS (PCM)

To overcome difficulties of the FCM, Krishnapuram and Keller [4] proposed a new clustering model named Possibilistic c- Means (PCM).

η_i –determines distance at which the membership value of a point in a cluster becomes 0.5.

$$\eta_i = K\frac{\sum_{j=1}^{N} u_{ij}^{m} d_{ij}^{2}}{\sum_{j=1}^{N} u_{ij}^{m}} \quad (5)$$

The advantage of PCM is clustering noisy data samples and disadvantage is very sensitive to good initialization.

V. FUZZY POSSIBILISTIC C MEANS ALGORITHM (FPCM)

To overcome difficulties of the PCM, Pal defines a clustering technique that integrates the features of both Fuzzy a Possibilistic c-means called Fuzzy Possibilistic c-Means (FPCM). Membership and Typicality's are very significant for the accurate characteristic of data substructure in clustering difficulty [5]. An objective function in the FPCM depending on both membership and typicality's are represented as: Memberships and topicalities is represented as:

$$J_{FPCM}(U,T,V) = \sum_{i=1}^{c}\sum_{j=1}^{n}\left(u_{ij}^{m} + t_{ij}^{\eta}\right)d^{2}\left(x_j, v_i\right) \quad (6)$$

FPCM generates Memberships and possibilities at the same time, together with the usual point prototypes or cluster center for each cluster. The advantage of the FPCM is, it ignores the noise sensitivity deficiency of FCM and overcomes the coincident clusters problem of PCM. The disadvantage is the row sum constraints must be equal to one.

VI. MODIFIED FUZZY POSSIBILISTIC C MEANS ALGORITHM (MFPCM)

The objective function in [6] is called the modified fuzzy possibilistic c-means (MFPCM) function composed of two expressions: the first is the fuzzy function and it uses a fuzziness weighting exponent, the second is possibilistic function and it uses a typical weighting exponent; but the two coefficients in the objective function are only used as exhibitor of membership and typicality [7]. The objective function of the modified fuzzy possibilistic c-means (MFPCM) can be formulated as follows

$$J_{MFP|CM} = \sum_{i=1}^{c}\sum_{j=1}^{n}(u_{ij}^{m} w_{ji}^{m} d^{2m}\left(x_j, v_i\right) + t_{ij}^{\eta} w_{ji}^{\eta} d^{2\eta}(x_j, v_i)) \quad (7)$$

Where U = $\{u_{ij}\}$ represents a fuzzy partition matrix and is defined by:

$$u_{ij} = \left(\sum_{k=1}^{c}\left(\frac{d^2(x_j, v_i)}{d^2(x_j, v_k}\right)^{2m/(m-1)}\right)^{-1} \quad (8)$$

$T = \{t_{ij}\}$ represents a typical partition matrix, and is defined by:

$$t_{ij} = \left(\sum_{k=1}^{c}\left(\frac{d^2(x_j, v_i)}{d^2(x_j, v_k}\right)^{2\eta/(\eta-1)}\right)^{-1} \quad (9)$$

$V = v_k$ represents c centers of the clusters, and is defined by:

$$v_i = \frac{\sum_{j=1}^{n}\left(u_{ij}^{m} w_{ji}^{m} + t_{ij}^{\eta} w_{ji}^{\eta}\right) * x_j}{\sum_{j=1}^{n}\left(u_{ij}^{m} w_{ji}^{m} + t_{ij}^{\eta} w_{ji}^{\eta}\right)} \quad (10)$$

Where $w_{ij} = exp\left(-\frac{d^2(v_i, v_j)}{\left(\sum_{j=1}^{n} d^2(x_j, \bar{x}_j)\right) * v/n}\right)$, $\bar{x} = \frac{\sum_{j=1}^{n} x_j}{n}$ (11)

The above equations indicate that membership u_{ij} is influenced by all C cluster centers, while possibility t_{ij} is influenced just by the i-th cluster center c_i. The possibilistic term distributes the t_{ij} with respect to every n data points, but not by means of every C clusters. Thus, membership can be described as relative typicality, it determines the degree to which a data fit in to cluster in accordance with other clusters and is helpful in correctly labeling a data point. Possibility can be observed as absolute typicality, it determines the degree to which a data point belongs to a cluster correctly, and it can decrease the consequence of noise. Joining both membership and possibility can yield to good clustering result.

VII. POSSIBILISTIC FUZZY C MEANS ALGORITHM (PFCM)

In FPCM, the constraint corresponding to the sum of all typicality values of all data to a cluster must be equal to one cause problems particularly for a big data set. In order to avoid this problem pal[8] et al propose a new algorithm called Possibilistic Fuzzy c means algorithm (PFCM) The objective function is defined by

$$J_{PFCM}(U,T,V,Z) = \sum_{i=1}^{c}\sum_{k=1}^{n}\left(au_{ik}^{m} + bt_{ik}^{n}\right) \times \| z_k - v_i \|^2 + \sum_{i=1}^{c}\delta_i \sum_{k=1}^{n}(1 - t_{ik})^{n} \qquad (12)$$

The advantages of the PFCM is, it ignores the noise sensitivity deficiency of FCM, overcomes the coincident clusters problem of PCM and eliminates the row sum constraints of FPCM.

VIII. BIAS CORRECTED FUZZY C MEANS (BCFCM)

We now modify the FCM to allow the labeling of a pixel to be influenced by the labels in its immediate neighborhood [9]. We call this modification BCFCM (Bias-Corrected FCM).

$$J_m^{BCFCM}(\mu, a) = \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}\left\|x_j - a_i\right\|^2 + \frac{\alpha}{N_R}\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}\left(\sum_{x_r \in N_j}\|x_r - a_i\|^2\right) \qquad (13)$$

Where N_j stands for the set of neighbors that exist in a window around x_j and N_R is the number of elements in N_k. The effect of the neighbors term is controlled by the parameter α.

IX. KERNEL BASED FUZZY C MEANS (KFCM)

Chen and Zhang [10, 11] pointed out a shortcoming of the BCFCM and then replaced the Euclidean distance $\|x_j - a_i\|$ with a Gaussian kernel-induced distance $1 - K(x_j, a_i) = 1 - \exp\left(-\|x_j - a_i\|^2/\sigma^2\right)$.

They gave the kernel version $JS_m^{KFCM\_S}$ of JS_m^{BCFCM} as

$$JS_m^{k}(\mu, a) = \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}(1 - K(x_j, a_i)) + \alpha\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}(1 - K(\bar{x}_j, a_i)) \qquad (14)$$

X. GAUSSIAN KERNEL BASED FCM (GKFCM)

The parameter α in BCFCM and KFCM_ S_1 heavily affects the final clustering results. For estimating the parameter σ and learning the parameter α , Yang[12] and Tsai [13] proposed a generalized type of BCFCM and KFCM_ S_1 where the parameter α can be automatically estimated and learned from the data, called GKFCM. The modified objective function JS_m^{GKFCM} with

$$JS_m^{GKFCM} = \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}(1 - K(x_j, a_i)) + \sum_{i=1}^{c}\sum_{j=1}^{n}\eta_i\mu_{ij}^{m}(1 - K(\bar{x}_j, a_i)) \qquad (15)$$

Where K(x, y) = $\exp\left(-\|x_j - a_i\|^2/\sigma^2\right)$ and $\sigma^2 = \sum_{j=1}^{n}\|x_j - \bar{x}\|^2/n$.

The parameter η_i is estimated as follows:

$$\eta_i = \frac{\min_{i' \neq i}(1 - K(a_{i'}, a_i))}{\max_k(1 - K(a_k, \bar{x}))} \qquad (16)$$

XI. ROBUST KERNEL BASED FUZZY C MEANS (RKFCM)

Zhang and Chen considered a kernel version of FCM by replacing the Euclidean distance $\|x_j - a_i\|$ with the kernel substitution as

$$\left\|\phi(x_j) - \phi(a_i)\right\|^2 = K(x_j, x_j) + K(a_i, a_i) - 2K(x_j, a_i) \qquad (17)$$

Where ϕ is a nonlinear map from the data space into the feature space with its corresponding kernel K, $K(x, y) = \phi(x)^T\phi(y)$ is an inner product kernel function.

They specially assumed $K(x,x)=1$, and then proposed the kernel-type objective function JS_m^{KFCM} and

$$JS_m^{KFCM}(\mu,a)=\sum_{j=1}^{n}\sum_{i=1}^{c}(\mu_{ij})^m\left\|\phi(x_j)-\phi(a_i)\right\|^2$$
$$=\sum_{j=1}^{n}\sum_{i=1}^{c}(\mu_{ij})^m\left(K(x_j,x_j)+K(a_i,a_i)-2K(x_j,a_i)\right)$$
$$=2\sum_{j=1}^{n}\sum_{i=1}^{c}(\mu_{ij})^m\left(1-K(x_j,a_i)\right) \quad (18)$$

If we take Gaussian kernel with $K(x_j,a_i)=\exp(-\left\|x_j-a_i\right\|^2/\sigma^2))$ then $K(x,x)=1$, the update equation for the necessary conditions of minimizing $J_m^{\phi}(\mu,a)$ are as follows:

$$a_i=\sum_{j=1}^{n}\mu_{ij}^m K(x_j,a_i)x_j/\sum_{j=1}^{n}\mu_{ij}^m K(x_j,a_i), i=1,2,\ldots,c; j=1,\ldots,n \quad (19)$$

$$\mu_{ij}=(1-K(x_j,a_i))^{\frac{-1}{m-1}}/\sum_{k=1}^{c}(1-K(x_j,a_k))^{\frac{-1}{m-1}}, i=1,\ldots,c; j=1,\ldots,n. \quad (20)$$

To speed the GKFCM, we adopt the suppression idea of Fan et al. [14] and Hung et al. [15] to KFCM algorithm.

REFERENCES

[1] Fan, J.L., Zhen, W.Z., Xie, W.X.: Suppressed fuzzy c-means clustering algorithm Pattern Recognition Lett. 24, 1607–1612 (2003)

[2] M.S.Yang, "A Survey of fuzzy clustering" Mathl. Comput. Modelling Vol. 18, No. 11, pp. 1-16, 1993.

[3] Jude hemanth.D, D.Selvathi and J.Anitha,"Effective Fuzzy Clustering Algorithm for Abnormal MR Brain Image Segmentation", Page Number 609-614, International/Advance Computing Conference (IACC 2009), IEEE,2009.

[4] R. Krishnapuram and J.M. Keller, "A possibilistic approach to clustering", IEEE Trans. Fuzzy Systems, Vol. 1, Pp. 98-110, 1993.

[5] Vuda Sreenivasarao and Dr.S. Vidyavathi, "Comparative Analysis of Fuzzy C- Mean and Modified Fuzzy Possibilistic C -Mean Algorithms in Data Mining", IJCST Vol. 1, No. 1, Pp. 104-106, 2010.

[6] A.Rajendran, R.Dhanasekaran, "MRI brain image tissue segmentation analysis using possibilistic fuzzy c-means method", International Journal on Computer Science and Engineering (IJCSE), vol. 3 no. 12, pp. 3832-3836, 2011

[7] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty (2002), —A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data, IEEE Trans. Med. Imag., vol. 21, no. 3, pp. 123–122, Mar. 2002.

[8] Pal N.R, Pal K, Keller J.M. and Bezdek J.C, "A Possibilistic Fuzzy c-Means Clustering Algorithm", IEEE Transactions on Fuzzy Systems, Vol. 13, No. 4, Pp. 517–530, 2005.

[9] Liao, L., Lin, T., Li, B.: MRI brain image segmentation and bias field correction based on fast spatially constrained kernel clustering approach. Pattern Recognition Lett. 29, 1580– 1588 (2008)

[10] Chen, S.C., Zhang, D.Q.: Robust image segmentation using FCM with spatial constrains based on new kernel-induced distance measure. IEEE Trans. Systems Man Cybernet.-Part B 34, 1907–1916 (2004)

[11] Zhang, D.Q., Chen, S.C.: Clustering incomplete data using kernel-based fuzzy c-means algorithm. Neural Processing Letters 18, 155–162 (2003)

[12] Hung, W.L., Yang, M.S., Chen, D.H.: Parameter selection for suppressed fuzzy c-means with an application to MRI segmentation. Pattern Recognition Lett. 27, 424–438 (2006)

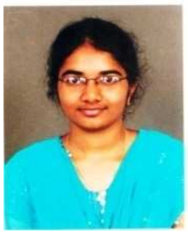
[13] Yang, M.S., Tsai, H.S.: A Gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction. Pattern Recognition Lett. 29, 1713–1725 (2008)

[14] Fan, J.L., Zhen, W.Z., Xie, W.X.: Suppressed fuzzy c-means clustering algorithm. Pattern Recognition Lett. 24, 1607–1612 (2003)

[15] Hung, W.L., Yang, M.S., Chen, D.H.: Parameter selection for suppressed fuzzy c-means with an application to MRI segmentation. Pattern Recognition Lett. 27, 424–438 (2006)

Authors Bibliography

Mr. R. Ravindraiah received his B.Tech degree from G.Pullareddy Engineering College, Kurnool, India in 2008 and M.Tech from AITS, Rajampet, India in 2011. He is doing research work in Bio-Medical Image Processing and presented some papers in reputed journals, IEEE & ElSIEVER Conferences. His research work includes analysis and processing of images of pathologies like Atherosclerosis, Retinopathy and Leukemia. Currently he is the Assistant Professor in department of Electronics and Communication Engineering, Madanapalle institute of technology and science, Madanapalle. Under his guidance many UG and PG students have completed their projects.



Ms. K. Tejaswini received her B.Tech degree in Electronics and Communication Engineering from Mekapati Rajamohan Reddy Institute of Technolgy and Science, Udayagiri, India in 2011.Currently she is pursuing M.Tech degree Digital Electronics and Communication Engineering in Madanapalle institute of technology and science, Madanapalle, India.