Fuzzy C clustering for computation of Cluster Density Distribution in a Mammogram for Breast Cancer Diagnosis

Tripty Singh[1], M.Nagraja[2]
[1]Department of Information Science and Engineering
[2]Department of Medical Electronics
MVJ College of Engineering, Bangalore
India
{triptysmart@gmail.com,Raj5rathna@yahoo.com}

ABSTRACT: *In recent years, the number and variety of applications of fuzzy logic have increased significantly. In this paper A fuzzy C Means technique in conjunction with 'Subtractive Clustering' is used to detect a micro calcification pattern and its density distribution. Fuzzy C Means Clustering FCM provides a method that shows how to group data points that populate some multidimensional space into a specific number of different clusters. The subtractive clustering method builds upon the function to provide a fast, one-pass method to take input-output training data and generate a Sugeno-type fuzzy inference system that models the data behavior. This FCM was designed for detecting Micro-calcifications and suspicious areas. In the process of detecting, it may detect other areas that look like a Micro-calcifications. It is up to the user to decide whether the resulting detection is a Micro-calcifications or some other area. The algorithm is simple and based on a fuzzy technique where the size of the Micro-calcifications n can be "about the size of a Micro-calcifications. The results showed that the genetic algorithm described in the present study was able to produce accurate results in the classification of breast cancer data and the classification rule identified was more acceptable and comprehensible*

Keywords: Fuzzy C Means Clustering (FCM), Subtractive clustering, Rule base clustering, Micro-calcifications

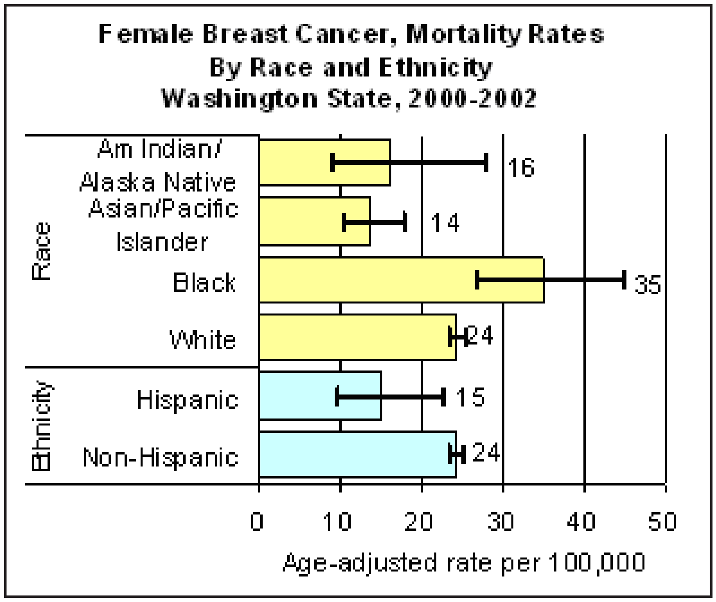
Received: 2 September 2011, Revised 8 October 2010, Accepted 13 October 2010

1. Introduction

An intelligent computer-aided diagnosis system can be very helpful for radiologist in detecting and diagnosing micro-calcifications' patterns earlier and faster than typical screening programs.Fuzzy logic has two different meanings. In a narrow sense, fuzzy logic is a logical system, which is an extension of multivalued logic. However, in a wider sense fuzzy logic (FL) is almost synonymous with the theory of fuzzy sets, a theory which relates to classes of objects with un-sharp boundaries in which membership is a matter of degree. In this perspective, fuzzy logic in its narrow sense is a branch of FL. Even in its more narrow definition, fuzzy logic differs both in concept and substance from traditional multivalued logical systems.

Another basic concept in FL, which plays a central role in most of its applications, is that of a fuzzy if-then rule or, simply, fuzzy rule. Although rule-based systems have a long history of use in Artificial Intelligence (AI), what is missing in such systems is a mechanism for dealing with fuzzy consequents and fuzzy antecedents. In fuzzy logic, this mechanism is provided by the calculus of fuzzy rules. The calculus of fuzzy rules serves as a basis for what might be called the Fuzzy Dependency and Command Language (FDCL). Although FDCL is not used explicitly in the toolbox, it is effectively one of its principal constituents. In most of the applications of fuzzy logic, a fuzzy logic solution is, in reality, a translation of a human solution into FDCL.

“Breast cancer is increasing in India with such a pace that we may face a serious burden of this disease in coming years,” says Diptendra Sarkar, assistant professor of Surgery at the Kolkata-based Institute of Post Graduate Medical Education and Research. “Thanks to the lifestyle changes in common people and lack of a system to properly facilitate mass-awareness and an early diagnosis and treatment facility in various regions, the incidence of breast cancer is getting increased in India



2. Enhancement Techniques

Step 1: Read Image
Step 2: Use Morphological Opening to Estimate the Background
Step 3: View the Background Approximation as a Surface
Step 4: Subtract the Background Image from the Original Image
Step 5: Increase the Image Contrast
Step 6: Threshold the Image
Step 7: Identify Objects in the Image
Step 8: Examine One Object
Step 9: View All Objects
Step 10: Compute Area of Each Object
Step 11: Compute Area-based Statistics
Step 12: Create Histogram of the Area
Step 13: Overview of ROI Processing
Step 14: Creating a Binary Mask
Step 15 Creating an Region Of Intrest without an Associated Image
Step 16: Creating an ROI Based on Color Values
Step 17: Analysis

3. De-noising The Image

One of the most important problems in image processing is de-noising. Usually the procedure used for de-noising, is dependent on the features of the image, aim of processing and also post-processing algorithms [5]. De-noising by low-pass filtering not only reduces the noise but also blurs the edges. Spatial and frequency domain filters are widely used as tools for image enhancement. Low pass filters smooth the image by blocking detail information. Mass detection aims to extract the edge of the tumor from surrounding normal tissues and background of mammogram images clearly shows that Partial low and high pass filter when applied to mammogram image leads to best Image Quality.

4. Cluster Identification Techniques

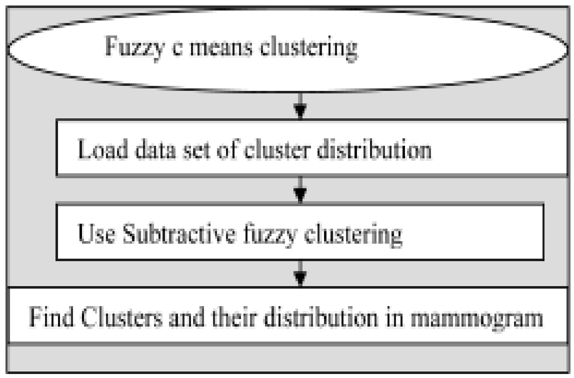


Figure 1. Showing the steps for finding clusters

5. Fuzzy C-Means Clustering

Fuzzy c-means (FCM) is a data clustering technique wherein each data point belongs to a cluster to some degree that is specified by a membership grade. This technique was originally introduced by Jim Bezdek in 1981[1] as an improvement on earlier clustering methods. It provides a method that shows how to group data points that populate some multidimensional space into a specific number of different clusters.

Fuzzy Logic Toolbox command line function fcm starts with an initial guess for the cluster centers, which are intended to mark the mean location of each cluster. The initial guess for these cluster centers is most likely incorrect. Additionally, fcm assigns every data point a membership grade for each cluster. By iteratively updating the cluster centers and the membership grades for each data point, fcm iteratively moves the cluster centers to the right location within a data set. This iteration is based on minimizing an objective function that represents the distance from any given data point to a cluster center weighted by that data point's membership grade.

The command line function fcm outputs a list of cluster centers and several membership grades for each data point. We have used the information returned by fcm to help us building a fuzzy inference system by creating membership functions to represent the fuzzy qualities of each cluster.

5.1 Implimentation of Fuzzy C Means Clustering Algorithm

This algorithm aims at detecting microcalcifications and suspicious areas. In the process of detecting, it may detect other areas that look like a microcalcification. It is up to the user to decide whether the resulting detection is a microcalcification or some other area. The algorithm is simple and based on a fuzzy technique where the size of the microcalcification can be "about the size of a microcalcification." It uses a 8*8 window to scan over the entire digital mammogram and locate microcalcifications or other abnormalities [2]:

*WHILE entire 8*8 image has not been examined by 4 window*
*MOVE 8*8 window to next position*
RECORD x,y position and grey level value of pixel with largest grey level in window
IF pixels ***surrounding*** *the largest pixel are* ***as bright as*** *the largest pixel grey level value AND*
outer *pixels are* ***darker*** *than the largest pixel grey level value*
THEN largest pixel position is the center pixel of a microcalcification
area
END IF
END WHILE

5.2 Obtaining 2-D Clusters for sample Mammogram

We have quasi-random two-dimensional data to illustrate how FCM clustering algorithm.Every time you run this example, the fcm function initializes with different initial conditions. This behavior swaps the order in which the cluster centers are computed and plotted

5.3 Subtractive Clustering

If we do not have a clear idea how many clusters there should be for a given set of data, *Subtractive clustering*,, is a fast, one-pass algorithm for estimating the number of clusters and the cluster centers in a set of data. The cluster estimates, which are obtained from the subclust function, can be used to initialize iterative optimization-based clustering methods (fcm) and model identification methods (like anfis). The subclust function finds the clusters by using the subtractive clustering method.The genfis2 function builds upon the subclust function to provide a fast, one-pass method to take input-output training data and generate a Sugeno-type fuzzy inference system that models the data behavior.

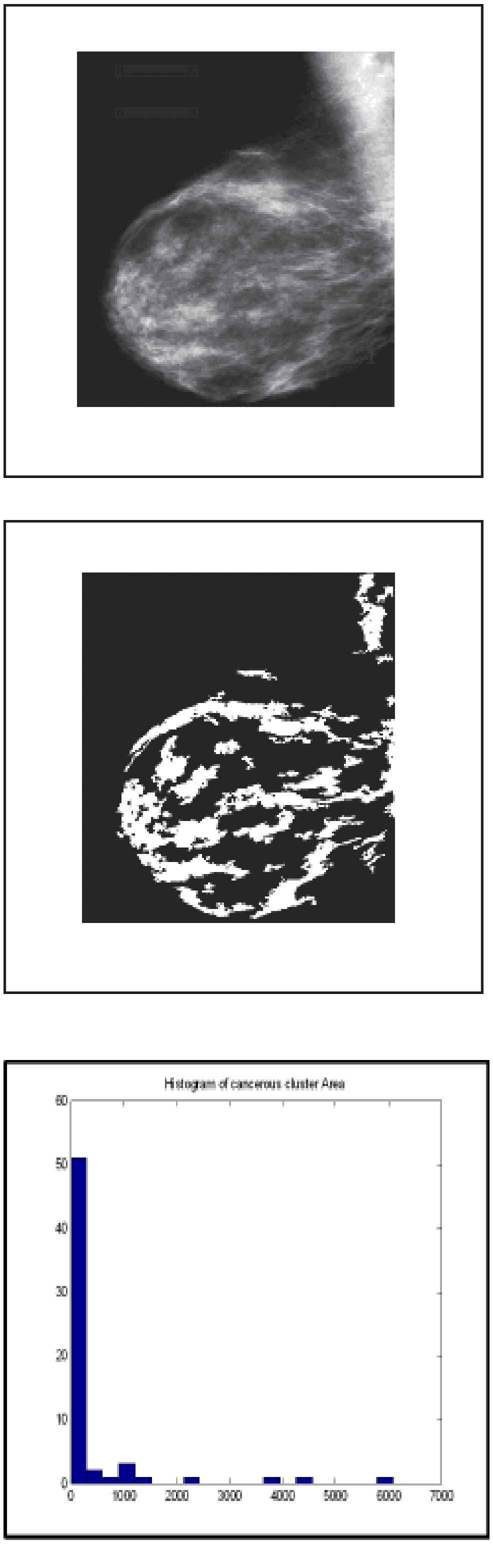
5.4 Find cluster in mammogram

Find cluster opens a GUI to implement either the fuzzy c-means (**fcm**) by using findcluster()function., the fuzzy subtractive clustering (**subtractiv**) using the pull-down tab under **Method** on the GUI, or both. Data is entered using the **Load Data** button. The options for each of these methods are set to default values. These default values can be changed. See fcm reference page for a description of the options for fuzzy c-means. The subclust reference page provides a description of the options for fuzzy subclustering.This tool works on multidimensional data sets, but only displays two of those dimensions. Use the pull-down tabs under **X-axis** and **Y-axis** to select which data dimension you want to view. For example, if you have data that is five-dimensional, this tool labels the data as data_1, data_2, data_3, data_4, data_5, in the order in which the data appears in the data set. **Start** to perform the clustering, and **Save Center** to save the cluster center.When operating on a data set called file.dat, findcluster (file.dat) loads the data set automatically, plotting up to the first two dimensions of the data only. You can still choose which two dimensions of the data you want to cluster after the GUI appears.

6. Results

- ✓ Iteration count = 1, obj. fcn = 8.970479
- ✓ Iteration count = 2, obj. fcn = 7.197402
- ✓ Iteration count = 3, obj. fcn = 6.325579
- ✓ Iteration count = 4, obj. fcn = 4.586142
- ✓ Iteration count = 5, obj. fcn = 3.893114
- ✓ Iteration count = 6, obj. fcn = 3.810804
- ✓ Iteration count = 7, obj. fcn = 3.799801
- ✓ Iteration count = 8, obj. fcn = 3.797862
- ✓ Iteration count = 9, obj. fcn = 3.797508
- ✓ Iteration count = 10, obj. fcn = 3.79744
- ✓ Iteration count = 11, obj. fcn = 3.79743
- ✓ Iteration count = 12, obj. fcn = 3.79743
- ✓ Iteration count = 1, obj. fcn = 8.794048
- ✓ Iteration count = 2, obj. fcn = 6.986628
- ✓ Iteration count = 12, obj. fcn = 3.79743

7. Figures and Graphs

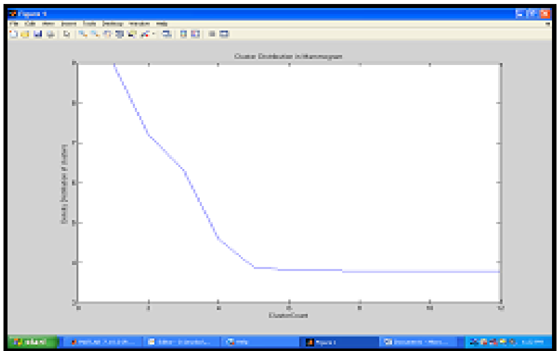


Pre Enhancement Image

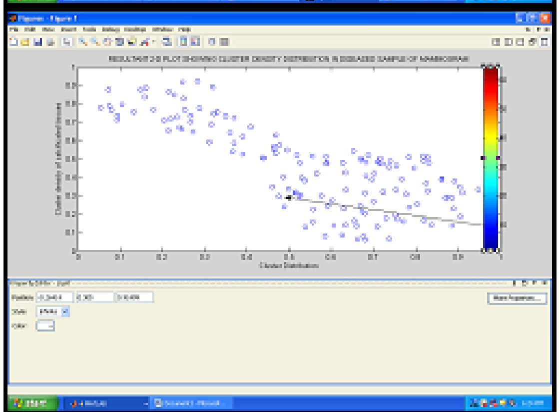
Pre Enhancement Image

Histogram of Post Enhancement Image

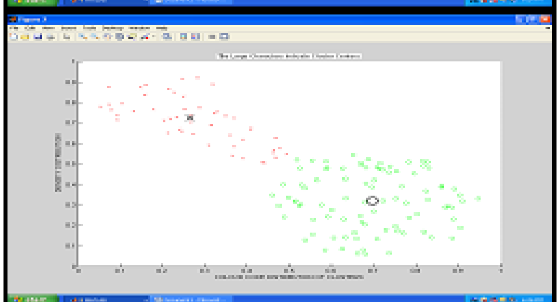
Figure 2. Pre and Post images in Image enhancement using histogram equalization and Partial Filters



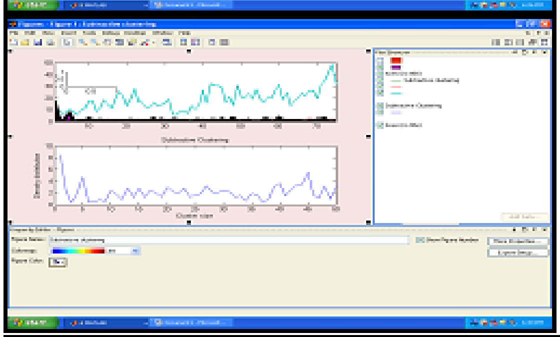
Graph 1. 2-D plot of density distribution of clusters in mammogram



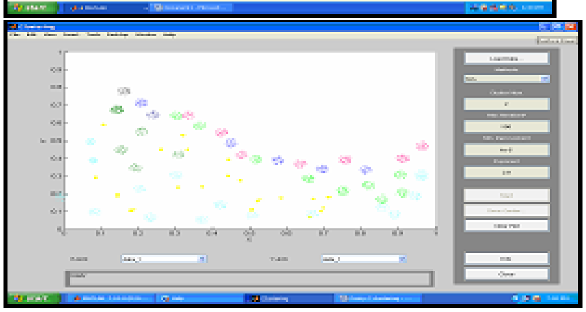
Graph 2. 3-D plot of density distribution of clusters in mammogram with colour bar



Graph 3: 3-D plot of Large characters indicating cluster center



Graph 4. Two plots Showing subtractive clustering



Graph 5. Plot showing clusters distribution of different intensities in mammogram

7. Conclusion

Fuzzy logic is not a cure-all. When should you not use fuzzy logic? The safest statement is the first one made in this introduction: fuzzy logic is a convenient way to map an input space to an output space. If you find it's not convenient, try something else. If a simpler solution already exists, use it. Fuzzy logic is the codification of common sense — use common sense when you implement it and you will probably make the right decision. Many radiologist, for example, do a fine job without using fuzzy logic. However, if you take the time to become familiar with fuzzy logic, you'll see it can be a very powerful tool for dealing quickly and efficiently with imprecision and nonlinearity

8. Acknowledgement

We would like to express our sincere appreciation to **Dr. Sarita Bhadauria and Dr Gopal Krishnan Nair** for guiding me through the project, for his thoughtful advice, insight and detailed suggestions. Many thanks and appreciation go to **Dr, Swarnalata Rao**, **Dr. A.K.Wadhwani** and **Dr. Sulochana Wadhwani** for serving as my graduate committee members and previewing the paper. I thank **Dr. B.M.Girish** for giving me the chance working at Desert Research Institute, we not only learned knowledge from working on the project but also learned a lot from his great personality. I thank **Kavita R** for helping me edit the paper.

References

[1] Yang, W.T., Lam, W.W.M., Cheung, H., Suen, M., King W.W.K., Metreweli, C. (1997). Sonographic, magnetic resonance imaging, and mammographic assessments of preoperative size of breast cancer, *J. Ultrasound Med.* 16, 791–797.

[2] Boetes, C., Mus, R.D., Holland, R,, Barentsz, J.O., Strijk, S.P., Wobbes, T., Hendriks, J.H., Ruys, S.H. (1990). Breast tumors: comparative accuracy of MR imaging relative to mammography and US for demonstrating extent. *Radiology* 197, 743-747 (1995).

[3] Davis, P.L., Staiger, M.J., Harris, K.B., Ganott, M.A., Klementaviciene, J., Mccarty, K.S. Jr, Tobon, H. (1996). Breast cancer measurements with magnetic resonance imaging, ultrasonography, and mammography. *Breast Cancer Res. Treat.* 37, 1-9.

[4] Fischer, U., Kopka, L., Grabbe, E.(1999). Breast carcinoma: effect of preoperative contrast-enhanced MR imaging on the therapeutic approach. *Radiology* 213, 881-888.

[5] Schelfout, K., Van Goethem, M., Kersschot, E., Colpaert, C., Schelfhout, A.M., Leyman, P., Verslegers, I., Biltjes, I., Van Den Haute, J., Gillardin, J.P., Tjalma, W., Van Der Auwera, J.C., Buytaert, P., De Schepper, A. (2004). Contrast-enhanced MR imaging of breast lesions and effect on treatment. *Eur. J Surg. Oncol.* 30, 501-507.

[6] Mumtaz, H., Hall-Craggs, M.A., Davidson, T., Walmsley, K., Thurell, W., Kissin, M.W., Taylor, I. (1997). Staging of symptomatic primary breast cancer with MR imaging. *AJR Am. J. Roentgenol.* 169, 417-424.

[7] Tripty Singh , Sarita Singh, Bhadauoria, Tomar , Ranjeet Singh., Wadwani, AK (2009). Analysis of Breast Cancer Using Image Processing Tecniques, Published IEEE, IEEE Computer Society Washington, DC, USA. Bibliometrics. This paper appears in: Computer Modeling and Simulation, 2009. EMS'09. Third uksim European Symposium on, Issue Date: 25-27 Nov. 2009. p. 251 – 256.

[8] Tripty Singh, Sarita Bhadauoria, Sulochana Wadhwani, Wadwani , AK(2010). Image Processing of a Mammogram for Micro Calcification, *In*: ICACCT 2010.*IEEE 4th International conference.* 30th Oct 2010,420-424

[9] Tripty Singh, Sarita Bhadauoria, Sulochana Wadhwani, Wadwani ,AK (2010).Expert System Design and Analysis for Breast Cancer, Deaprtment of Electronics and Electrical Engineering. MITS Gwalior , 2 (12) 31st December 2010.IJEST International Journal of Engineering Science and Technology.

[10] Tripty Singh, Bhadauoria, Sarita Singh., Wadhwani, S., Wadwani, A.K (2010). Design Issues of Expert System for Breast Cancer Detection, *International Journal of Information Studies,* 3 (1) 5.

[11] Tripty Singh, Sarita Singh Bhadauoria, Wadhwani, S., Wadwani, AK (2001). Image Analysis and Interpretation of a Mammogram for Microcalcification, *International journal of CSI*, 'IJCSI-2011-8-1-62' IJCSI. 8 (1).