ISSN: 0975-766X
CODEN: IJPTFI
*Available Online through* *Research Article*
**www.ijptonline.com**

# SURVEY ON GRID CLUSTERING APPROACH WITH INTUITIONISTIC FUZZY HISTOGRAM

**Payal Nemade, Aarti Sharma, Eklendra Jadam, C.L.Chowdhary**
School of Informational Technology and Engineering, VIT University, Vellore, India.
*Email: setconfpaper@gmail.com*

*Received on 25-10-2016* *Accepted on 02-11-2016*

**Abstract:**

A lot of work was executed on conventional clustering methods like k-mean. Later some fuzzy c-mean clustering was proposed. Fuzzy c-mean algorithm was applied on different areas like pattern recognition, data mining issues, image processing approaches and medical imaging. Clustering is more challenging when applied on medical images. Thus, the medical image segmentation is an important process in image processing used to partition the images into different regions. Many proposed ideas by various researchers have been used for clustering techniques k-mean, fuzzy c-mean etc.

**Keywords:** K-mean, fuzzy c-mean, medical image segmentation, clustering.

## 1. Introduction:

Cluster is the set of object such that the object is that group are more similar to each other than the other group of cluster. It is used to classify the object in the interval area. A cluster is an aggregate of points in the test space such that the distance between any two points in the cluster is less than the distance between any point in the cluster and any point not in it. Its goal is to determine the set of unlabelled data so that the object in the same group of clustering is similar then in the other group of clustering.

There are many fields in which clustering can be used, it is mainly used in Scalability, Dealing with different types of attributes, Discovering clusters with arbitrary shape, Minimal requirements for domain knowledge to determine input parameters, Ability to deal with noise and outliers, Insensitivity to order of input records, its dimensionality is high, Interpretability and usability.There are many clustering techniques, like EM clustering, Lloyd's K-means [1], spectral clustering, K-mean, fuzzy C-mean, were used by different researchers. In our survey we may use the techniques like K-mean, Fuzzy c-mean and the grid based clustering with the intuitionistic fuzzy histogram for getting the better

results in the images. For obtaining better accuracy in our proposed paper we are using advanced grid-based clustering. We may use intuitionistic fuzzy histograms approach for better results.

## 2. Background

In our survey, we may use some clustering techniques which are k-means, fuzzy c-means and grid based clustering. The most useful techniques are k-means and fuzzy c-means which are derived from the clusters. K-means and fuzzy c-means are the prototype-based clusters. As like these clustering techniques there are So much techniques but we are going to use this techniques mainly in our paper. In our paper we are going to deal with this techniques are as given below:

### 2.1 Fuzzy c-means:

The method is developed by Dunn in 1973 and improved by Bedeck in 1981 in frequently used in pattern recognition and Fuzzy C-means needs to do a full inverse-distance weighting. Fluffy c-implies (FCM) is a grouping systems which is utilized as a part of information, in which each point has a level of fitting in with the bunches as in fluffy rationale instead of having a place with one bunch. It is a strategy for bunches which permits one bit of information to have a place with two or more groups.

- Algorithm for fuzzy c-means:
- Firstly, select a number of clusters in a given image.
- Now, Assign randomly to each point coefficients for begins in a cluster.
- Repeat the step until convergence criterion is met.
- The centre of each cluster should be computed.
- At last, compute its coefficients of
- Being in the clusters for each point.

### 2.2 K-means:

K-means is one of the main clustering techniques and is very famous because it is simpler in use. K-means is a method of vector quantization, k-means bunching partition the n perceptions into k-groups where every perception has a place with the bunches. K-means needs to do a distance calculation from the centroid. K-means is a nice method to quickly sort your data into clusters. K-means is a fast, robust and easier to understand. There are many applications of the k-means is Vector Quantization, Feature Learning, Cluster Analysis. The formula which is mainly used in the k-means clustering is:

$J(V) = \sum_{i=1}^{C} \sum_{j=1}^{Ci} (||x_i - v_j||)^2$

Algorithm for the k-means clustering:

a) Firstly, Place K points into the space which is represented by the objects that are being clustered in the group.

b) Secondly, these points are used for represent the initial group of centroid.

c) The last one is it assign each object to the group that has the closest centroid.

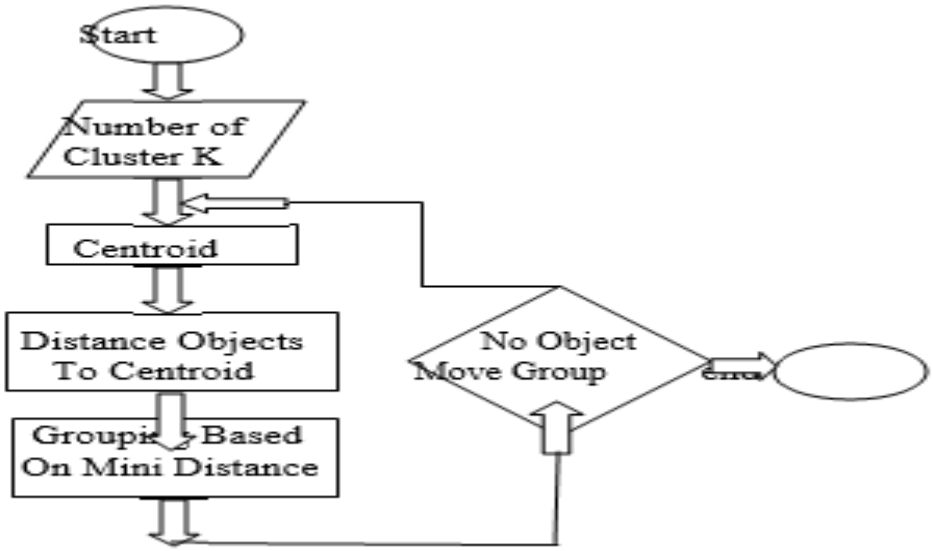


**Fig 1: Architecture Diagram.**

**2.3 Grid Based Clustering:**

The Grid Based Clustering which partitions the data space into a finite number of cells to form a grid structure and performs all clustering operation to group similar spatial object into classes on this obtain grid structure is an efficient clustering. Grid Based clustering method used the multi-resolution grid data structures. It has several interesting methods are STING (a Statistical Information Grid approach), CLIQUE (Clustering in Quest). The main advantages of the grid based clustering algorithm is fast, it has no distance computations and the shapes are limited to union of the grid-cells and this clustering is a significant reduction in the time complexity especially for very data sets. Grid-Based clustering algorithm:

a) Define a set of grids.

b) Then, calculating the cell density for each cell.

c) Eliminate those cells, whose density is below a certain threshold.

d) In the next step, identifies the cluster centres.

e) At last step, traversal of neighbour cells.

**Advantage**

- The major advantage of this method is fast processing time.
- It is dependent only on the number of cells in each dimension in the quantized space.

**2.4 Intuitionistic Fuzzy Histogram:**

Intuitionistic Fuzzy Histogram plays very important role in our proposed paper because we are using intuitionistic fuzzy histogram with the different clustering techniques like grid based, k-mean, fuzzy c-mean to apply the intuitionistic fuzzy sets which seem to be a very better tool for representation. The membership and non-membership functions are able to assign one column, there is a way of assigning the membership and non-membership function for intuitionistic fuzzy sets is relative frequency distributions (histograms).the intuitionistic fuzzy set (IFS) theory is related to definitions of new objects and their properties. The terminology we will use below in the definitions of intuitionistic fuzzy histogram is IFS (intuitionistic fuzzy sets).

**Definition 1.** An intuitionistic fuzzy set (IFS) P in E is defined as an object of the following form

$$P= \{(x, \mu_P(x), \nu_P(x))| \, x \in E\} \quad (1)$$

Where the functions:

$$\mu_P : E \rightarrow [0,1] \quad (2)$$

$$\nu_P : E \rightarrow [0,1] \quad (3)$$

Define the degree of membership and the degree of non-membership of the element $x \in E$, respectively and for every $x \in E$:

$$0 \leq \mu_P(x) + \nu_P(x) \leq 1$$

And we can write the, each fuzzy set may be written as

$$\{(x, \mu_P(x), 1- \mu_P (x)) \, | x \in E\}.$$

**Definition 2.** The value of

$$\pi_P(x) = 1- \mu_P (x) - \nu_P(x) \quad (4)$$

Which is called the degree of the uncertainty of the element $x \in E$ to the intuitionistic fuzzy set P.

Simply, in the case of ordinary fuzzy sets, $\pi_P(x) = 0$ for every $x \in E$.

Let the IFS,

$$P= \{<x, \mu_P (x), \nu_P(x)> | x \in E\} \quad (5)$$

**Definition 3.** We define the characteristic function of an IFS A, $\Omega_P : E^* \rightarrow \{0,1\}$, by

$\Omega_P((x, a, b)) = \begin{cases} 1, & if \ P(x) = a \text{ and } vP(x) = b \\ 0, & otherwise \end{cases}$

Therefore, P is a set in the sense of the set theory in the universe E*. The backward transformation of this result to the ordinary fuzzy sets is trivial.While undertaking of deducing essential associations between concepts as well as courses and workings in surmising an idea organize from a course cross links to exchange learned relations are both new, similar to augmentations to the SVM algorithms exhibited, our work was motivated by strategies in related fields, principally:

In combined filtering via matrix conclusion the works is focused on transfer learning. We look to exchange prerequisite relations between sets of courses inside universities to different combines additionally inside universities, and to link that universities. This is motivated by yet different from the exchange learning writing. Exchange adapting generally looks to exchange enlightening elements, priors, idle structures and all the more as of late regularization punishments. Rather, we move shared ideas in the mappings between course-space and idea space to instigate essential relations. Although the main objective is to detect the prerequisite relation among different courses. One of the direct application of self-induced universal concept graph is task. Other applications includes modularization in designing syllabus by trainers and curriculum scheduling based on different qualifications of students for educational goals of the student. The fact that our Category based illustration scheme (CAT) gave better results as compared to other schemes, gives us more comfort and visual assessment also supports the decision of the inferred links at the concept level, though with infrequent errors.

**3. Proposed System Architecture**

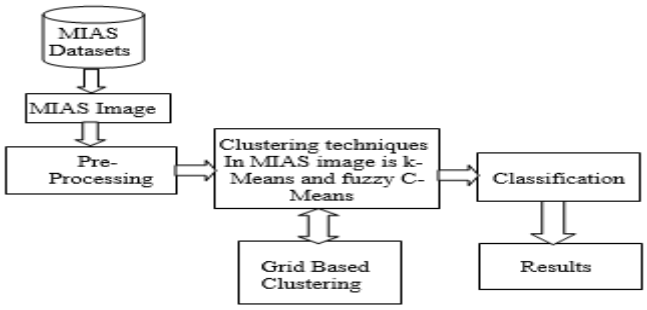


**Fig 2. System Architecture.**

**3.1 MIAS Datasets:**

We will utilize the MIAS Datasets for our execution. The techniques those we are going to utilize, was connected on the MIAS dataset. Mammography Image Analysis Society, which is an association of United Kingdom examination gatherings intrigued by the comprehension of mammograms, has created a computerized mammography database.

The images of the database originate from a film-screen mammographic imaging process. The image those we use, were annotated according to their breast density by the radiologists.

**3.2 MIAS Images:**

MIAS is stands for Mammography Image Analysis Society. Our methods we applied on MIA's images for detecting the breast cancer. The sound waves are used for breast ultra sound to make a computer picture of the inside of the breast (sonogram). For detecting the breast cancer the imaging methods is use is Mammogram. A mammogram is an X-rays types of image of the breast. It is a two dimensional images. We will use the MIAS image in our implementation for applied all the techniques on it.

**3.3 Pre-processing:**

Pre-processing is used to process the data before the data is parsed. It is also used to analyse, transform and filter data. Pre-processing algorithm those is used in our proposed paper to in order to enhance the image and improve the overall image quality like better contrast of region and the less noise. In our paper, we are going to use the pre-processing algorithms are Histogram equalization and the median filtering [6]. This both algorithms, we are used for the better contrast and for remove the noise. Median filtering is applied to the image of m*m sub regions, it preserves the edges of the image and it is done pixel by pixel.

**Mathematical evaluation:**

The following mathematical matrices, we are using in our paper for the evaluation of the quality of the image.

1) Peak Signal to Noise Ratio (PSNR): PSNR is expressed in terms of the logarithm decibel scale. It's mathematical form is

$$PSNR=10.\log_{10}(MAX^2/MSE)=20.\log_{10}(MAX/\sqrt{MSE})$$

MAX is the maximum pixel value of the image. Generally, Maximum possible value of MAX is $2^B$-1, B is stands for bits per sample.

2) Mean Squared Error (MSE): The most commonly use of the PSNR is a measure of the quality of reconstruction in image compression. Mean Squared Error is defined for two m*n monochrome images.

$$MSE = 1/mn \left(\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} ||I(i,j) - k(i,j)||^2\right.$$

**3.4 Clustering techniques:**

Clustering techniques is very important part in our paper of system architecture, because it plays very important role after pre-processing. With the help of clustering techniques those we may use like k-means and fuzzy c-means, will

detect or find the breast cancer in MIAS image. The clustering techniques [2] we have discussed above in the brief those we may use in our survey and the grid based clustering is also we are using for getting the better accuracy in our image. It is also very fast in use, we have given the description of the grid clustering above.

**3.5 Classification:**

Firstly, we will take MIAS image which is a two-dimensional image after that we will the pre-processing techniques for the contrast and for removing noise from the image after that the Clustering techniques we will use on this image and also use grid based clustering after that we will get classify it then we will be able to the accurate image according to our expected Result.

**4. Conclusion:**

Paper based on fuzzy C-means and K- means clustering methods. The fuzzy c-means algorithm is sensitive to initialization & is easily find in local optima. An intensity-based intuitionistic fuzzy image representation was considered. The experimental evaluation of the proposed approach focused on object clustering especially in the presence of noise and image segmentation. In the presented example they are obtained by clustering analysis. The histograms can be used for the results visualization and the values analysis. In our proposed paper we are using intuitionistic fuzzy histogram for getting better accuracy in the images. The computational results show that the performance of the proposed algorithm is better than the other existing algorithms.

**References:**

1. E. Szmidt, K. Atanassov, J. Kacprzyk (2013). On intuitionistic fuzzy pairs, Notes on IntuitionisticFuzzy Sets, Vol. 19(3), pp. 1–13.
2. C.L. Chowdharyand D.P. Acharjya, 2016. A Hybrid Scheme for Breast Cancer Detection using IntuitionisticFuzzy Rough Set Technique, International Journal of Healthcare Information Systems and Informatics. Vol. 11(2), pp. 38-61..
3. S. Chatterjee, G. Sheikholeslami, A. Zhang (2000). WaveCluster: a wavelet-based clustering approach for spatial data in very large databases, The VLDB Journal,Vol. 8, pp. 289-304.
4. C. L. Chowdhary, P. G.Shynu, 2011. Applications of Extendable Embedded Web Servers in Medical Diagnosing,International Journal of Computer Applications, Vol. 38 (6), pp. 34-38.
5. K. Atanassov (1993). A universal operator over intuitionistic fuzzy sets. Comptes Rendus de l'Academie bulgare des Sciences, Vol. 46 (1), pp. 13–15.

6. C. L. Chowdhary, 2011. Linear feature extraction techniques for object recognition: study of PCA and ICA,Journal of the Serbian Society for Computational Mechanics,Vol. 5 (1), pp. 19-26.
7. G. Gluhchev, L. Atanassova, K. Atanassov (2010). On intuitionistic fuzzy histograms, Notes on Intuitionistic Fuzzy Sets, Vol. 16(4), pp. 31–36.
8. R. Manavalan, B.Sathya (2011). Image Segmentation by Clustering Methods: Performance Analysis", IJCA, Vol. 29(11).
9. U. C. Sekhar, L. D. Dwivedi and C. L. Chowdhary, 2013. Classification of ECG- Beats using Features from TwoStageTwo-Band Wavelet Decomposition, Journal of Theoretical and Applied Information Technology, Vol. 49(3), pp. 922-928.
10. C. L. Chowdhary, G. V. K. Sai and D. P. Acharjya, 2016. Decrease in False Assumption for Detection usingDigital Mammography, In Springer Proceedings under AISC series, International Conference on ComputationalIntelligence in Data Mining (ICCIDM-2015), Vol. 2, pp. 325-333.
11. C. Pantofaru, R. Unnikrishnan, and M. Hebert (2007). Toward objective evaluation of image segmentation a algorithms, IEEE Trans. Pattern Anal. Mach. Intell, Vol. 29(6), pp. 929-944.
12. C. L. Chowdhary and D. P. Acharjya, 2016. Breast Cancer Detection using Intuitionistic Fuzzy Histogram Hyperbolization and Possibilitic Fuzzy c-mean Clustering algorithms with texture feature based Classification on Mammography Images, AICTC '16 Proceedings of the International Conference on Advances in Information Communication Technology & Computing, August 13-14, 2016.
13. C. L. Chowdhary, 2016.A review of feature extraction application areas in medical imaging, International Journal of Pharmacy and Technology, Vol. 8 (3), pp. 4501-4509.
13. C. L. Chowdhary, K. Muatjitjeja, D. S. Jat, Three-dimensional object recognition based intelligence system for identification, Proceedings of 2015 International Conference on Emerging Trends in Networks and Computer Communications, ETNCC 2015, pp. 162-166.
14. C. L. Chowdhary, A. Ranjan and D. S. Jat, 2016. Categorical database information-theoretic approach of outlier detection model, Annals. Computer Science Series. 14th Tome 2nd Fasc. – 2016, pp. 29-36.