Survey on clustering methods : Towards fuzzy clustering for big data

Abdelkarim Ben Ayed, Mohamed Ben Halima, Adel M. Alimi
REGIM-Lab.: REsearch Groups in Intelligent Machines
University of Sfax, National Engineering School of Sfax , BP 1173, Sfax, 3038, Tunisia
Sfax, Tunisia
abdelkarim.benayed.tn@ieee.org, mohamed.benhlima@ieee.org, Adel.alimi@ieee.org

***Abstract*—** In this report, we propose to give a review of the most used clustering methods in the literature. First, we give an introduction about clustering methods, how they work and their main challenges. Second, we present the clustering methods with some comparisons including mainly the classical partitioning clustering methods like well-known k-means algorithms, Gaussian Mixture Modals and their variants, the classical hierarchical clustering methods like the agglomerative algorithm, the fuzzy clustering methods and Big data clustering methods. We present some examples of clustering algorithms comparison. Finally, we present our ideas to build a scalable and noise insensitive clustering system based on fuzzy type-2 clustering methods.

Keywords-clustering; fuzzy; big data.

I. INTRODUCTION

Nowadays, more and more data is created mainly from file hosting websites, social networks, digital cameras, etc. So to organize and classify all those huge amounts of data we use classification and clustering technique which arrange data in homogenous groups called clusters. Data clustering is an essential step in many domains including data mining, pattern recognition, medical fields, spatial databases applications, computer vision, DNA analysis, market analysis, web statistics...

Classification methods are called supervised technique because we already know the classification parameters, characteristics and exact number of cluster groups, although, the clustering methods are called unsupervised technique because we do not know the characteristics of the cluster groups and even their number sometimes. In this case the clustering parameters are computed from learning data. Learning is done on a part of data set to estimate system parameters.

There are two types of learning: off-line (Batch) learning for data available in blocks and on-line learning for data that arrive sequentially. Batch learning converge to better result, but it is not valid for real time application and very large data.

Clustering methods are more appropriate for real, heterogeneous and large data sets with many attributes.

Data clustering is usually anticipated by a step of features (attributes) extraction, and then results are saved in attributes vector, which is used for the calculation of similarity metric between data objects during the classification algorithm.

The most challenges that encounter the clustering methods are initialization parameters, finding the number of clusters, choosing similarity metric (distance), scalability (handling large databases, time complexity, memory requirements), sensibility to noisy data (outliers), incomplete data, overlapping data, validating the clustering results, etc [1].

Saad and Alimi have proposed validity index and number system to validate the clustering results and the number of clusters [2].

All those issues and special characteristics for each type of data and the variety of application domains means that there is not a perfect clustering method for all applications, but, there is more suitable clustering method for each application. So we found many authors that give surveys about clustering algorithms categories, issues and comparing them [1],[3].

In this report, we propose to give a review of the most used clustering methods in the literature including mainly the classical partitioning clustering methods, the classical clustering hierarchical methods, the fuzzy clustering methods and Big data clustering methods.

The remaining sections of this paper are organized as follows. In Sections 2 and 3, we present respectively a brief review of classic clustering methods and fuzzy clustering methods, then, in section 4 we present the new trends for big data clustering methods. In section 5 we present some experimental results and comparisons demonstrating the effectiveness of fuzzy and big data clustering methods. Finally, section 6 contains a conclusion and our ideas to build a scalable and noise insensitive clustering system based on fuzzy type-2 clustering methods.

II. CLASSIC CLUSTERING METHODS

In this section, we present the popular clustering methods mainly the classical partitioning clustering methods, the classical hierarchical clustering methods and density based clustering methods.

A. Partitioning clustering methods

The partitioning methods tries to directly cluster data objects. They divide data into homogenous clusters. First, cluster centers are assigned arbitrary, then, data points are assigned to the nearest cluster center based on distance (like k-means algorithm and its variants) or probability (like Gaussian Mixture Modals and variants) similarity function, then an iterative algorithm optimize clusters until convergence.

1) K-means algorithm

k-means algorithm appears since 1965 [4] and is by the far the most used clustering algorithm for its simplicity of implementation and its effectiveness. It is unsupervised clustering and N*k*d difficult algorithm. K-means algorithm uses Euclidean distance and its objective is to minimize

distance inside the same cluster and maximizes distance between clusters by minimizing the objective function J.

$$J = \sum_{i=1}^{n}\sum_{k=1}^{n} Z_i^k \parallel x_{ik} - v_i \parallel^2 \quad [4] \qquad (1)$$

where

- n: number of data points
- x_i: data points
- Z: is a membership function
- $Z_i^k = \begin{cases} 1, & if\ x_i \in cluster\ k \\ 0, & if\ not \end{cases}$
- v_i: cluster centers

$$v_i = \frac{\sum_{k=1}^{n} x_{ik} x_{kj}}{\sum_{k=1}^{n} x_{ik}} \quad [4] \qquad (2)$$

k-means algorithm steps are:
1) Manually choose the parameter 'k' (number of clusters)
2) Cluster centers v_i are randomly chosen
3) Data points x_{ij} are assigned to the nearest cluster
4) Re-compute cluster centers v_i using equation (2)
5) Repeat steps 3) and 4) until J is invariant (variance $<\varepsilon$)

The most important issues for k-means clustering is that the clustering result depends strongly to the initialization of the cluster centers and their number. K-means algorithms is valid only for numerical data. Besides, k-means do not converge to global minimum but converge to a local minimum. To solve the last issues, we can do many k-means iterations, and then choose the one with least objective function value.

In addition, k-means algorithm do not take data distribution in consideration and do not take in consideration the fact that real objects have no equal importance (unlike Euclidean distance does).

2) K-means variants

To resolve some limitations in k-means algorithms, many improvements (variants) have been proposed like k-medians algorithm that uses cluster medians in spite of cluster means so that other distances can be used other than Euclidian distance.

Using Euclidean distance k-means algorithm finds only spherical clusters. We can also use Mahal distance to find ellipsoidal cluster, but with a higher complexity.

K-medoids (or general k-means) can work with any distance and with categorical data. It can also handle outliers, but has much more complexity and require more iterations.

Kernel k-means algorithm allows clustering irregular shapes and different densities groups based on the use of eigenvectors and minimizing sum of squared errors. Kernel k-means gives better results than k-means but it has much more complexity (N^2*k) and so it has less scalability [5].

Possibilistic c-means produce typical membership that has a good explanation for the degree of belonging to the data and has better clustering performance than c-means mainly with the presence of noise, but still depends on initialization, and can generate coincident clusters because of the independency of some data points to the other data [6].

3) EM algorithm

Dempster has proposed the EM algorithm in 1977 [7]. It is a model based algorithm that cluster data points based on probability. It assumes that data is generated from a mixture of distribution. The algorithm is based on finding maximum likelihood parameters of probabilistic modals. Clustering models are provided by a finite mixture of distributions f.

$$f(x/\theta) = \sum_{i=1}^{K} p_i f_i(x/\alpha_i) \quad [8] \qquad (3)$$

Where:

p_i is the propotion of the class I ($p_i \geq 0\ and\ \sum_i p_i = 1$)

$$\alpha_i = (\mu_i \sum_i),$$

μ_i and $\sum_i$ are respectively the center and the variance matrix of k^{th} normal component $f(./\alpha_i)$.

The log-likelihood of the global parameter θ is then maximized to evaluate the previous parameters.

$$\ln f\ (X/\theta) = \sum_{j=1}^{n} \ln f\ (x_j/\theta) \quad [8] \qquad (4)$$

Where

$X = (x_1, \ldots, x_n)$

A normalized product of a likelihood such the Gaussian Mixture Modals (GMM) is used to express the distribution.

The EM algorithm uses iteratively expectation and maximization until it converges. The expectation step is used to estimate the likelihood then the maximization of likelihood step using estimated parameters from the other step. The result of maximization is then used by the other expectation step, etc. The algorithm converges when variance of the parameter θ is below a fixed threshold ε.

The EM algorithm is appropriate mainly for data with different size and correlated data but is noise sensitive. Other variants of EM algorithm uses adaptive distance ADEM [8] using image histogram and spatial information to handle noisy data.

4) Density based methods

Density based methods search for connected dense regions to identify areas heavily populated with data but they are not suitable for high dimension data, ex: DBCURE, DBSCANN, Optics, DENCLUE, etc. They are used mainly to build clusters with irregular shapes [1].

B. Hierarchical clustering methods

The hierarchical methods do not cluster data directly like partitioning methods, but use hierarchical grouping or

division to gradually assemble / disassemble data points in clusters [1].

Hierarchical divisive methods start initially from only one cluster then try to divide it into other clusters. Hierarchical agglomerative methods consider initially each object as the center of a cluster then try to merge them in bigger clusters.

There are many methods to measure distance between groups: single link (minimum distance), average link (average distance), complete link (maximum distance) and ward's method (objective function).

The hierarchical methods have high complexity N^2 which make them not scalable and not appropriate for large datasets.

C. Other clustering methods

There are many other clustering algorithms less popular like those k-nearest neighbors, Bayesian networks, neural networks, graphical clustering, spectral clustering, etc.

Clustering ensembles combines clustering results with co-occurrence matrix using different parameters, different algorithms or using different data representation.

Semi supervised clustering or constraint clustering uses pairwise constraints (must-link or cannot-link) provided by domain experts [3].

Graph clustering algorithm (or spectral clustering) represent elements as nodes in weighted graph, then, extract clusters by minimizing the weight of edge nodes using cut minimizing algorithms. The graph clustering algorithm allows to give more weight for some features, which is needed for some real applications.

D. Limits of classical clustering methods

Those classical clustering methods are famous and are largely used due to their simplicity of implementation and execution time performance mainly for clean, small and synthetic data sets, but they present many limits when used to resolve clustering real life data sets, which are noisy, incomplete, overlapping and large data sets.

III. FUZZY CLUSTERING METHODS

Fuzzy logic is created by Zadeh since 1965 [9] to resolve real life problems. Fuzzy logic tries to act like humans and use simple logic rules to solve real, complex and non-linear problems. Ruspini was the first to implement a fuzzy clustering system in 1969 [10].

Fuzzy clustering methods are based on fuzzy membership, while in classical hard clustering methods, data is assigned to different clusters so each data element belongs to exactly one cluster. In fuzzy clustering methods (also called soft clustering), data elements can be member of more than one cluster, so that objects can belongs to many clusters at the same time.

There are three categories of fuzzy clustering methods: based on fuzzy relation, based on k-nearest neighbor rule and based on objective function. The last category is the most used in fuzzy clustering.

Those fuzzy methods are used to improve clustering results when the boundaries are overlapped, examples Fuzzy C-Means, Fuzzy Gaussian Mixture Modals, etc.

To resolve noisy data (outliers) and other real data problems, zadeh has proposed Fuzzy Logic Type-2.

A. Fuzzy C-means algorithm

Fuzzy C-means algorithm is inspired from the classic c-means, it was developed by Dunn in 1973 [11] then improved by Bezdek in 1981 [12]. It is still very widely used for data clustering. The membership function is not just 0 or 1 but a value between 0 and 1, so that membership vector (for k-means), is replaced by a membership matrix (for c-means).

The membership is represented by c by n matrix, where c is the number of fuzzy subsets and n is the number of objects. Each row represents the membership of all n object to a certain fuzzy subset and each column represents the membership of an object to all c fuzzy subsets.

C-means algorithm is very similar to k-means algorithm, but with new objective function J'.

$$J'(U,V) = \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}^{m}||X_j - V_i||^2 \quad [13] \qquad (5)$$

Where

- U: Fuzzy partition matrix
- V: Set of prototypes
- N: number of prototypes
- C: number of clusters
- X_j: is the j^{th} measured data point
- v_i: is the center of cluster i
  - $v_i = \frac{\sum_{j=1}^{N} u_{ij}^{m} X_j}{\sum_{j=1}^{N} u_{ij}^{m}}$, $1 \le i \le C$ [13] (6)
- u_{ij}: $(0 \le u_{ij} \le 1)$ is the membership value of x_j with respect to cluster i.

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}(\| X_j - V_i \| / \| X_j - V_k \|)^{2/(m-1)}} \qquad (7)$$

- m: $(m \ge 1)$ determines the degree of cluster fuzziness.

The fuzzifier parameter m is fixed according to the application. It's used to reduce the sensibility of the class centers to noise. Bigger value of m leads to a more fuzzy clusters, and small value of $(m \cong 1)$ leads to hard clusters. Bezdek proved that fuzzifier m should be set in [1.5 ; 2.5]. In general, m is set to 2.

The algorithms steps are similar to the hard k-means algorithm:

1) Chose C centers randomly (vector V)
2) Compute the partition matrix U
3) update the centers v_i

4) Compute objective function J

5) Repeat steps 2 to 4 until convergence ($\| var(J) \leq \varepsilon \|$)

The most important issues for FCM algorithm are the computational time, it's much slower than the hard c-means due to iterative nature, choice of metric distance (Euclidian distance isn't always the best), choice of initial clusters, choice of the number of clusters, choice of fuzzifier parameter m, sensibility to edge noise, etc [8].

B. *Fuzzy C-means variants*

Many authors have proposed variants of FCM to solve some of those issues, like using other distance functions, using other membership function or using hierarchical agglomerative algorithm or khonen networks to find the number of clusters.

Gustafson and Kessel [14] have proposed to use adaptive distance norm inducing an n by n matrix to optimize the distance function.

Gath and Geva [15] have proposed FMLE (Fuzzy Maximum Likelihood Estimation) with distance norm distance. This method gives better results, especially for hyper ellipsoidal classes.

Ahmed has proposed in 2002 [16] a modified FCM algorithm based on a modified objective function based on neighbor's information. This algorithm is more robust than c-means clustering and give good classification, mainly for noisy MRI medical images.

Other authors have proposed SFCM (Supressed FCM) [17]. The algorithm consists in magnifying and suppressing largest membership degrees. This algorithm converges faster than FCM but its performance depends strongly on the random parameter α. A modified SFCM determine the parameter α with prototype driven learning approach. This method was tested mainly for MRI medical images.

Nikhil, et al. have proposed possibilistic FCM in 2005 [18], which is a fuzzy version of possibilistic c-means (2003). PFCM has both advantages of fuzzy membership and PCM typicality. PFCM uses spatial information in the objective function to give better quality clustering results. Unlike FCM algorithm that forces outliers to belong to one or more clusters, which distort the result of clustering, PFCM is more robust for outliers. PFCM objective function has two randomly fixed parameters a (degree of fuzzy) and b (degree of typicality).

In 2011 Saad, et al. has proposed Enhanced PFCM [19], in which parameters a and b have no random values but computed according to data points to have better clustering results. In 2014 Zhao [20] has proposed another modified SFCM.

C. *Type-2 Fuzzy logic*

Zadeh has proposed type-2 fuzzy logic in 1975 [21]. It is a generalization of type-1 fuzzy set and consist in using uncertain (or fuzzy) membership function. Unlike fuzzy logic, which uses 2-dimensional memberships, type-2 fuzzy logic is 3-dimensional memberships, so it can handle more uncertainties and it is better for identifying outliers. For type-2 fuzzy logic the membership function is called FOU (Footprint Of Uncertainty) and it is bounded by two membership function (upper membership function and lower membership function). There are two categories: generalized type-2 fuzzy using a 3-dimensional FOU and interval type-2 fuzzy using a 2-dimensional FOU. The latter is the most used due to its simplicity and reduced complexity.

Type-2 based clustering algorithms such as T2FCM, T2FGMM, T2FHMM, … give better results than similar and fuzzy type-1 algorithms, but, have higher complexity [22].

IV. BIG DATA CLUSTERING METHODS

For some applications like medicals, aerospace … we can have very large databases called also "Big Data". For such databases, a huge amount of memory and time complexity (computation) are required for the corresponding clustering which makes the classic clustering methods unable to handle those huge amounts of data, so new methods are required [1].

To reduce execution time and memory, some methods replace real data by integer values, other methods use samples of the data or summary of the data, etc.

Due to its low computational cost and easily parallelized process, k-means algorithm is still used for big data clustering, but improvements and variants are needed to support multiple core parallelism, to handle outliers, to reduce complexity for very big data and for more convergence efficiency.

To solve scalability issue for kernel k-means, approximate kernel k-means is proposed which use sampling to reduce execution time.

A. *Incremental mining*

Those techniques as DIGNET [1] and siibFCM [23] based on k-means algorithm, but use only one pass, thus allows reducing execution time and handling noise but also affects cluster quality besides the clustering strongly depends on data order.

B. *Probability based for clustering big data*

CLIQUE is an EM algorithm variant with much higher scalability. It can handle high dimensional data and big data requirements [1].

C. *Squashing techniques*

Those techniques (ex: BIRCH, Bubble, Bubble-FM) scan data to compute certain data summaries then cluster the summaries using height balanced tree of nodes [1].

D. *Distributed methods*

Distributed or parallel methods distribute computation on many computers to improve scalability, like DBCURE-MR [24], P-PIC [25] Google Map Reduce algorithm [26], HDFS, Hadoop, etc.

V. EXAMPLES OF COMPARATIVE STUDIES ON CLUSTERING ALGORITHMS

In this section, we present some results of comparison between clustering methods [16], [23]. Then, we present our comparisons.

In [23] the authors compare the traditional FCM with their size insensitive integrity based FCM (siibFCM) on synthetic and medical images, in most cases they reduce the error rate from more than 30% to 0% mainly for elliptical clusters. Results are illustrated in figure 1.

In [16] the authors compared FCM with EM and their Bias corrected FCM (BCFCM) on medical images. The results show that EM (85%) is better than FCM (79%) algorithm and that the proposed algorithm BCFCM gives the best results (93%) mainly for noisy images. Results are illustrated in figure 2.

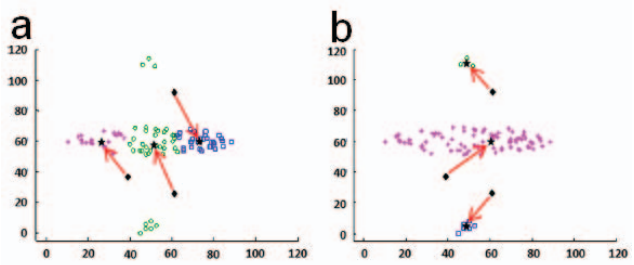

Figure 1. Clustering results on a synthetic image; (a) using FCM; (b) using siibFCM. [23]

In [22] Zeng and Lui have presented a survey of comparisons between the most popular fuzzy type-1 clustering methods (FCM, GMM, HMM) with their fuzzy type-2 correspondents (T2FCM, T2GMM, T2HMM) giving better results mainly with uncertain datasets.

In [27] Wu and Mendel observed a significant improvement when comparing type-2 fuzzy logic classifiers with type-1 fuzzy logic classifiers on Battlefield ground vehicle dataset. The used dataset has variable acoustic features of multiple terrains that contains lot of uncertainties. Their results as shown in the table I.

TABLE I. Mean of the classification error rates over more than 800 experiments [27]

Dataset	T2 fuzzy classifiers	T1 fuzzy classifiers
Battlefield ground vehicle	9.13%	12.8%

We compare k-means, hierarchical ward's, EM and FCM algorithms using Matlab on a synthetic image that contain 4000 random created data points. We get similar results for k-means, ward's and FCM. Results are shown in figure 3.

VI. CONCLUSION

The choice of clustering algorithm depends on the application domain. Most clustering variants are dedicated to certain application domain. Some clustering algorithms give similar results like k-means, EM, spectral clustering and ward's linkage algorithms.

Fuzzy and big data clustering methods give in general better results than classical clustering methods and even better if merged. There have to be used when classical methods are useless because Fuzzy and big data clustering methods have more complex implementation and use require more computational resources.

We propose in the near future to develop a system that has the advantages of noise insensibility of Type-2 fuzzy clustering methods and scalability of big data clustering methods.

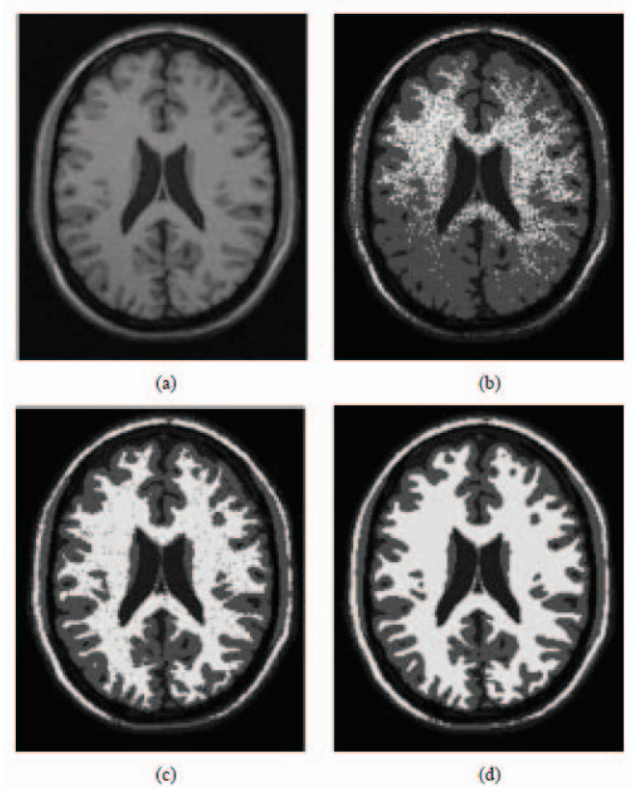

Figure 2. Clustering results on a medical MRI image; (a) original image; (b) using FCM; (c) using EM; (d) using BCFCM. [16]

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support of this work by grants from General Direction of Scientific Research (DGRST), Tunisia, under the ARUB program.

REFERENCES

[1] J. Kogan, C. Nicholas, M. Teboulle, "A Survey of Clustering Data Mining Techniques", Springer Berlin Heidelberg, 2006, pp. 25-71, Berkhin, P.

[2] MF. Saad, Adel M. Alimi, “Validity Index and number of clusters”, International Journal of Computer Science Issues (IJCSI), 2012.

[3] Anil K. Jain, “Data clustering: 50 years beyond K-means”, Pattern Recognition Letters, vol. 31, Issue 8, 1 June 2010, pp 651-666.

[4] E. W. Forgy, “Cluster analysis of multivariate data: efficiency vs interpretability of classifications.”, 1965, Biometrics, 21, pp 768–769.

[5] B. Scholkopf, A. Smola, and K.-R. Muller, “Nonlinear component analysis as a kernel eigenvalue problem.”, Neural Computation, 10, 1998, pp 1299–1319.

[6] J. L. Fan, W. Z. Zhen and W. X. Xie, “Suppressed fuzzy c-means clustering algorithm”, Pattern Recognition Letters, 24(9), pp. 1607-1612, June 2003.

[7] A. P. Dempster, N. M. LAIRD, D. B. RUBIN, “Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal”, 1977, Statistical Society, 39, pp. 1–38.

[8] K. Kalti, M. A. Mahjoub, "Image Segmentation by Gaussian Mixture Models and Modified FCM Algorithm", The International Arab Journal of Information Technology, vol. 11, No. 1, January 2014.

[9] L.A. Zadeh, “Fuzzy sets”, Inf.control, 8, pp.338-352, 1965.

[10] E. R. Ruspini, “Numerical methods for fuzzy clustering”, Information Science, 2, pp. 319-350, 1970.

[11] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters", J. Cybernetics, 3(3), pp. 32-57, 1973.

[12] J. C. Bezdek, "Pattern recognition with fuzzy objective function algorithms", Plenum New York,1981.

[13] S. Ara Begum, M. O. Devi, “A Rough Type-2 Fuzzy Clustering Algorithm for MR Image Segmentation”, International Journal of Computer Applications 54(4), pp. 4-11, September 2012. Published by Foundation of Computer Science, New York, USA.

[14] D. Gustafson and W. Kessel, “Fuzzy clustering with a fuzzy covariance matrix”, Proceedings of IEEE Conference on Decision Control, San Diego, CA, pp. 761-766, January 10-12, 1979.

[15] I. Gath and A. B. Geva, “Unsupervised optimal fuzzy clustering”, IEEE Transactions Pattern Anal. Machine Intell., PAMI-11(7), pp. 773-781, July 1989.

[16] M. Ahmed, S.M. Yamany, N. Mohamed, A. A. Farag, T. Moriarty, "A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data", IEEE Transactions on Medical Imaging, vol.. 21, NO. 3, March 2002.

[17] R. Krishnapuram and J. Keller, “A possibilistic approch to clustering”, IEEE Transactions on Fuzzy Systems, 1(2), May 1993.

[18] R. Nikhil, K. Pal, James M. Keller and J. C. Bezdek, "A possibilistic fuzzy c-Means clustering algorithm", IEEE Transactions on Fuzzy Systems, 13(4), August 2005.

[19] M.F. Saad, J. Lee, O. Kwon and A.M. Alimi, “Context data clustering on modified fuzzy possibilistic c-means algorithm for efficient context-aware computing servives”, International Interdisciplinary Journal Information, 14 (9), pp. 3101-3111, September 2011.

[20] F. Zhao, J. Fan, H. Liu, "Optimal-selection-based suppressed fuzzy c-means clustering algorithm with self tuning non local spatial information for image segmentation", Expert Systems with Applications, vol. 41, Issue 9, July 2014, pp 4083-4093.

[21] L. A. Zadeh, "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning–1," Information Sciences, vol. 8, pp. 199–249, 1975.

[22] J. Zeng, Z.Q. Liu, "Type-2 Fuzzy Sets for Pattern Recognition: the State-of-the- Art"", Journal of Uncertain Systems 1, pp. 163–177 (2007).

[23] P. Lin, P. Huang, C.H. Kuo, Y.H. Lai, "A size-insensitive integrity-based fuzzy c-means method for data clustering", Pattern Recognition, vol. 47, Issue 5, May 2014, pp. 2042-2056.

[24] Y. Kim, K. Shim, M. Kim, J. S. Lee, "DBCURE-MR: An efficient density-based clustering algorithm for large data using MapReduce", Information Systems, vol. 42, June 2014, pp. 15-35.

[25] W. Yan, U. Brahmakshatriya, Y. Xue, M. Gilder, B. Wise, "P-PIC: Parallel power iteration clustering for big data", Journal of Parallel and Distributed Computing, vol. 73, Issue 3, March 2013, pp. 352-359.

[26] J. Dean, S. Ghemawat, "MapReduce: simplified data processing on large clusters", Communications of the ACM - 50th anniversary issue, vol. 51 Issue 1, January 2008, pp. 107-113.

[27] H. Wu and J. M. Mendel, “Classification of battlefield ground vehicles using acoustic features and fuzzy logic rule-based classifiers”, IEEE Trans. Fuzzy Syst., vol. 15, no. 1, pp. 56–72, Feb. 2007.

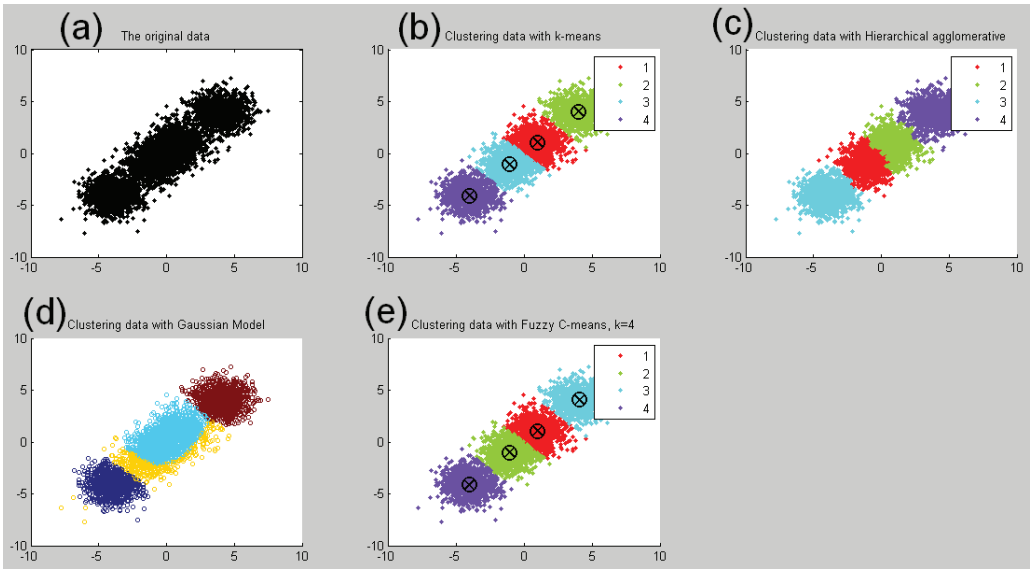

Figure 3. Clustering results on a synthetic image; (a) original image; (b) using k-means algorithm; (c) using ward’s hierarchical algorithm; (d) using EM algorithm; (e) using FCM algorithm