VOL. 10, NO. 8, MAY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

HYBRIDIZATION OF FUZZY C-MEANS AND COMPETITIVE AGGLOMERATION FOR IMAGE SEGMENTATION

I. Nandhin[1] and V. Mohan[2]
[1]M.E Communication Systems, Saranathan College of Engineering, Tiruchirapalli, Tamil Nadu, India
[2]Saranathan College of Engineering, Tiruchirapalli, Tamil Nadu, India
E-Mail: nandhiniiyyan20@gmail.com

ABSTRACT

Image segmentation is the process of partitioning a digital image into multiple segments. Different methodologies have been proposed for the segmentation based on normal techniques such as region growing, threshold technique, watershed transform. The disadvantage of these methods leads to the development of segmentation based on clustering techniques. The main concept of data clustering is to use the centroid to represent each cluster. Also it is based on the similarity between the input vectors to that of the centroid to represent each cluster. Parametric and Non-parametric methods are the broad classes of the clustering methods. Non-parametric method involves finding natural groupings in a dataset using a Euclidean distance between the samples of the dataset. Non-parametric clustering includes k-means, hierarchical, spectral clustering. The disadvantages of these methods are lack of sufficient robustness to image noise. So, a fuzzy segmentation methodology has been widely applied in image clustering and segmentation. The important problem in fuzzy c-means is to specify the number of clusters and selection of objective function. So, a fuzzy clustering-based vector quantization algorithm is used. This algorithm utilizes a specialized objective function, which involves the fuzzy c-means along with a competitive agglomeration term. This algorithm is a fast process and the reconstructed images maintain high quality.

Keywords: competitive agglomeration, fuzzy c-means, fuzzy learning vector quantization (FLVQ).

1. INTRODUCTION

Image segmentation is the process of partitioning a digital image into multiple segments. Image segmentation is used to locate objects and boundaries in images. It is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics. The result of image segmentation is a set of segments that collectively cover the entire image, or set contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristics such as color, intensity, texture. Clustering algorithms groups the samples of a set such that two samples in the same cluster are more similar to one another than two samples from different clusters. Clustering methods can be categorized into two broad classes: non-parametric and parametric methods. Non-parametric clustering involves finding natural groupings (clusters) in a dataset using an assessment of the degree of difference (such as Euclidean distance) between the samples of the dataset. It requires the measure of similarity between samples, defining a criterion function for clustering, and defining an algorithm to minimize (or maximize) the criterion function. Popular non-parametric clustering algorithms include k-means, Hierarchical clustering algorithms and Spectral clustering [1].

Image segmentation process is close to the clustering problem. Clustering methods [3] have been successfully used to segment an image into a number of clusters (segments). Clustering-based segmentation techniques [9] have used several control parameters, e.g., the predefined number of clusters to be found or some tunable thresholds. These parameters are adjusted to obtain the best image segmentation. The parameters value is a nontrivial task. This section provides a basic idea for the segmentation based on clustering technique.

2. PROPOSED WORK

The proposed work is the hybridization of fuzzy c- means and competitive agglomeration for image segmentation. It is fuzzy clustering-based vector quantization algorithm. Fuzzy clustering vector quantizer which linearly combined the fuzzy c- means and CA term as described in section III. This algorithm utilizes a specialized objective function, which involves the fuzzy c-means along with a competitive agglomeration term. A widespread classification of clustering-based VQ methods distinguishes them into crisp and fuzzy. Crisp VQ is mainly based on the c- means method. The c-means is very sensitive on initialization. Fuzzy techniques are mainly based on the fuzzy c-means algorithm. Fuzzy Learning Vector Quantization (FLVQ) manipulates the fuzziness parameter from large to small values. FLVQ also results in high computational cost. The second tries to accomplish the transition from fuzzy to crisp mode by incorporating special mechanisms to reduce the number of distance calculations, while keeping the fuzziness parameter constant. Finally the third constitutes a combination fuzzy c-means and CA term. The implementation of vector quantization (VQ) is based on code words and codebook. The images are taken and converted into blocks. The vectors are formed from the blocks of images. Then the clustering process is carried out. The blocks form the feature vectors and designing a codebook that minimizes the distortion measure as given in Equation (1). Then the code vectors again are reconstructed into blocks. Then the corresponding

VOL. 10, NO. 8, MAY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

reconstructed image is obtained. The distortion measure is calculated as follows:

$$D = \frac{1}{n} \sum_{k=1}^{n} \min \left\{ \left\| x_k - v_i \right\|^2 \right\} \qquad (1)$$

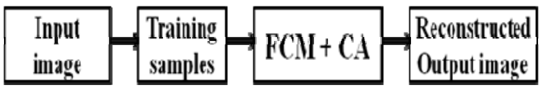


Figure-1. Block diagram for the proposed system.

3. MATHEMATICAL DESCRIPTION

A. Fuzzy c-means algorithm

In fuzzy clustering, the data vectors are associated with membership grades which indicate the degree to which the data vectors belong to the different clusters. The main objective of fuzzy clustering algorithm is to partition the data vectors into clusters so that the similarity of data vectors within each cluster is maximized and the similarity of data vectors in different clusters is minimized. Their fuzzy nature makes the clustering procedure able to retain more original image information than the crisp or hard clustering methodologies. Fuzzy algorithm allows gradual memberships of data vectors to clusters measured as degrees in [0, 1]. This gives the flexibility to express that data vectors belong to more than one cluster.

From the image database, the images are taken and converted into blocks. The vectors are formed from the blocks of images. Then the clustering process is carried out. The cluster centers are input arguments from the vectors. Then the input vector and the cluster centers are compared based on the Euclidian distance function. The minimum distance between the input vector and cluster center is chosen and the corresponding label of the cluster center value is assigned as index to the corresponding input vector. This process is iteratively done until for all the iterations will be over. Then the minimum distance of input vector corresponding to which cluster center is taken for reconstructing the images as given in Eq. (2). Then the code vectors again are reconstructed into blocks. Then the corresponding reconstructed image is obtained.

Iterations: In each iteration of the FCM algorithm, the following objective function J is minimized.

$$J = \sum_{i=1}^{N} \sum_{j=1}^{N} u_{ij}^{m} \left\| x_i - c_j \right\|^2 \qquad (2)$$

where

N = number of vectors,

c = number of clusters,

c_j = centre vector for cluster j,

u_{ij} = degree of membership for the ith data vector x_i in cluster j.

$\|*\|$ The norm, measures the similarity (or closeness) of the data vector x_i to the centre vector c_j of cluster j.

In the iteration, a centre vector for each of the clusters is maintained. These vectors are calculated as the weighted average of the data vectors, where the weights are given by the degrees of membership.

Degree of membership for a given data vector u_{ij}, the degree of its membership to cluster c_j calculated as follows:

$$u_{ij} = \frac{1}{\left(\frac{\left\| x_i - c_j \right\|}{\left\| x_i - c_j \right\|} \right)^2} \qquad (3)$$

where

m denotes the fuzziness coefficient.

The centre vector c_j is calculated as follows.

$$c_j = \left(\frac{\sum_{i=1}^{N} u_{ij}^{m} * x_i}{\sum_{j=1}^{N} u_{ij}^{m}} \right) \qquad (4)$$

Fuzziness coefficient in the above equations the fuzziness coefficient m, where 1 < m < 1, measured the tolerance of the required clustering. The clusters are overlapped with one another is determined by the value of m.The overlapping between clusters is larger for large values of m. The higher the value of fuzziness coefficient a larger number of data vectors will fall inside a `fuzzy' band where the degree of membership is neither 0 nor 1, but somewhere in between.

Termination condition the required accuracy of the degree of membership is determined by the number of iterations completed by the FCM algorithm. This measure of accuracy is calculated using the degree of membership from one iteration to the next, taking the largest of these values across all vectors considering all of the clusters.

B. Modified fuzzy c-means with competitive agglomeration algorithm

The competitive agglomeration (CA) algorithm is a powerful technique that refines good from spurious and badly delineated clusters by minimizing the following objective function.

$$J(U, V, X) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^2 \left\| x_k - v_i \right\|^2 - \gamma \sum_{i=1}^{c} \left(\sum_{k=1}^{n} u_{ik} \right)^2 \qquad (5)$$

subject to the constraint

$$\sum_{j=1}^{c} u_{jk} = 1 \qquad \forall k \qquad (6)$$

where

ARPN Journal of Engineering and Applied Sciences

u_{ik} = membership degree of the kth training vector to the ith cluster.

The first component in Eq. (5) is similar to the fuzzy c-means objective function with the fuzziness parameter. The second component is the CA term whose basic contribution is to control the number of clusters. The parameter is chosen properly, the final partition will minimize the sum of intra-cluster distances, while partitioning the available data set into the smallest number of clusters.

$$\gamma^t = \eta^t \frac{\sum_{k=1}^{n}\sum_{i=1}^{c} (u_{ik})^2 \|x_k - v_i\|^2}{\sum_{i=1}^{c}\left(\sum_{k=1}^{n} u_{ik}\right)^2} \tag{7}$$

$$\eta^t = \eta_0 \ell^{\frac{-t}{\tau}} \tag{8}$$

where

η_0 = initial value of η
τ = time constant of the learning process.

The membership degrees and the cluster centers are calculated as follows

$$v_i = \frac{\sum_{k=1}^{n} (u_{ik})^2 x_k}{\sum_{k=1}^{n} (u_{ik})^2} \tag{9}$$

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\|x_k - v_i\| / x_k - v_j\right)^2} + \frac{\gamma}{\|x_k - v_i\|^2} \mathrm{N}((C_i) - \mathrm{N}(x_k)) \tag{10}$$

$$\mathrm{N}(c_i) = \sum_{k=1}^{n} u_{ik} \tag{11}$$

$$\tilde{\mathrm{N}}(x_k) = \sum_{j=1}^{c} \frac{\mathrm{N}(c_j)}{\|x_k - v_j\|^2} \Big/ \sum_{j=1}^{c} \frac{1}{\|x_k - v_j\|^2} \tag{12}$$

where
$\mathrm{N}(c_i)$ = fuzzy cardinality of the ith cluster.
$\mathrm{N}(x_k)$ = weighted average cardinality calculated with respect to the training vector.

The Competitive Agglomeration term creates large clusters while continuously shrinking the small clusters until their size becomes less than a predefined threshold. In the migration process, a codeword migration strategy is used. The code words of small clusters are detected and they are relocated in specific positions close to large clusters. Due to the competition to achieve an increasing size between the clusters is increased. The competitive agglomeration creates the large clusters, while the low cardinality clusters become smaller as the iteration number increases and are, ultimately discarded.

4. FLOW CHART FOR THE PROPOSED ALGORITHM

1. Initialize the partition matrix U, n, t, c, τ .
2. Initialize v_i .
3. Calculate the distortion measure as in Equation (1).

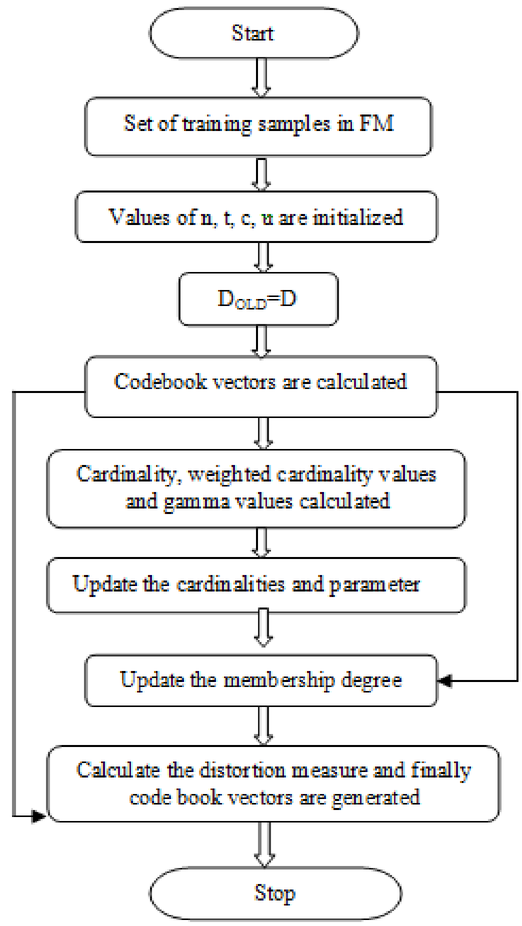
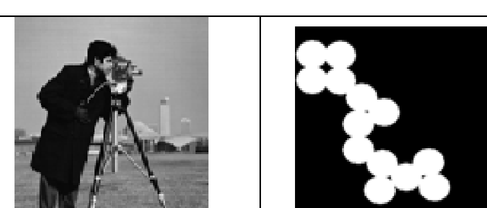


Figure-2. The flow chart illustrating the procedure for the competitive agglomeration algorithm.

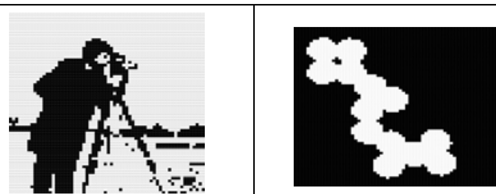
5. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, we experimentally evaluate our proposed method for a set of images. The experiments have been developed in Mat lab R2009b, and are executed on an Intel Pentium Dual-Core 2.2 GHZ CPU, 2G RAM.



(a) (b)

Figure-3. (a) and (b) Original images.



(c) (d)

Figure-4. (c) and (d) Image segmentation results.

6. CONCLUSIONS

In this paper, we propose a new simple and effective Fuzzy clustering-based vector quantization using fuzzy c-means along with CA term. The joint effect is a learning process where the number of code words affected by a specific training sample is gradually reducing and therefore, the number of distance calculations is also reducing. Thus, the computational cost becomes smaller. The proposed algorithm is a fast process and is insensitive with respect to its design parameters. The algorithm is based on distortion measure function. The reconstructed images maintain high quality, which is quantified in terms of the distortion measure.

REFERENCES

[1] J. Yu, “General C-Means Clustering Model,” IEEE 773-779, 1999.

[2] Jianbo Shi and Jitendra Malik “Normalized Cuts and Image Segmentation” IEEE transactions on pattern analysis and machine intelligence, vol. 22, no. 8, august 2000.

[3] Yongyue Zhang, Michael Brady, and Stephen Smith “segmentation of brain MRI images through a hidden markov random field model and the expectation-maximization algorithm” IEEE transactions on medical imaging, vol. 20, no. 1, January 2001 45.

[4] Mohamed N. Ahmed, Sameh M. Yamany “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.

[5] Balkis Abidi, Amel Bouzeghoub “a new algorithm for fuzzy clustering able to find the optimal number of clusters” 1082-3409/12 $26.00 © 2012 IEEE.

[6] Gerardo Di Martino, Antonio Iodice, Daniele Riccio” a novel approach for disaster monitoring fractal models and tools” IEEE transactions on geosciences and remote sensing, vol. 45, no. 6, June 2007.

[7] Moumen El-Melegy, E.A.Zanaty, Walaa M.Abd-Elhafiez, and Aly Farag “on cluster validity indexes in fuzzy and hard clustering algorithms for image segmentation” 1-4244-1437-7/07/$20.00 ©2007 IEEE.S

[8] Turgay Celik “unsupervised change detection in satellite images using principal component analysis and *k*-means clustering” IEEE geosciences and remote sensing letters, vol. 6, no. 4, October 2009.

[9] Mahnhoon Lee “On fuzzy cluster validity indices for the objects of mixed features” fuzz-IEEE 2009, Korea, August 20-24, 2009.

[10] Tai-Ning Yang, Chih-Jen Lee and Shi-Jim Yen “Fuzzy objective functions for robust pattern recognition” IEEE 2009, Korea, August 20-24, 2009.

[11] Mingrui Zhang, Wei Zhang, Hugues Sicotte and Ping yang “a new validity measure for a correlation-based fuzzy c-means” 978-1-4244-3296-7/09/$25.00 ©2009 IEEE clustering algorithm.

[12] Xiangrong Zhang, Zemin Li, Biao Hou, Licheng Jiao “spectral clustering based unsupervised change detection in SAR images” 978-1-4577-1005-6/11/$26.00 ©2011 IEEE.

[13] Maoguo Gong, Yan Liang, Jiao Shi, Wenping Ma, and Jingjing Ma “fuzzy c-means clustering with local information and kernel metric for image segmentation” IEEE transactions on image processing, vol. 22, no. 2, February 2013.

[14] Xiao fang Liu, Xiaowen Li, Ying Zhang, Cunjian Yang, Wenbo Xu, Min Li, Huanmin Luo “ Remote Sensing Image Classification Based on Dot Density Function Weighted FCM Clustering Algorithm” 1-4244-1212-9/07/$25.00 ©2007 IEEE.

[15] Jocelyn Chanussot, Jon Atli Benediktsson, and Mathieu Fauvel “classification of remote sensing images from urban areas using a fuzzy possibilistic model” IEEE geosciences and remote sensing letters, vol. 3, no. 1, January 2006.

[16] Turgay Celik and Hwee Kuan Lee “a robust fuzzy local information c-means clustering algorithm” IEEE transactions on image processing, vol. 22, no. 3, March 2010.

[17] Dongju Liu, JianYu “objective function of Otsu method is equivalent to that of K means method in multilevel thresholding” vol. 45, no. 4, pp. 1907-1916, 2004.

[18] Tse-Wei Chen, Yi-Ling Chen, Shao-Yi Chien. 2010. proposed "Fast Image Segmentation Based on K-Means Clustering with Histograms in HSV Color Space” IEEE Trans vol. 32, no. 3, pp. 359–369, 1996.

[19] Nan Li, Hong Huo, Yu-ming Zhao, Xi Chen, and Tao Fang “a spatial clustering method with edge weighting for image segmentation” IEEE geosciences and remote sensing letters, vol. 10, no. 5, September 2013.

VOL. 10, NO. 8, MAY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

[20] Xiangrong Zhang, Yaoguo Zheng, Jie Feng, Shuiping Gou "sar image change detection based on low rank matrix decomposition" 978-1-4673-1159-5/12/$31.00 ©2012 IEEE.

[21] Timothy C. Havens, James C. Bezdek "Cluster validity for Kernel Fuzzy Clustering" WCCI 2012 IEEE World Congress on Computational Intelligence June, 10-15, 2012.

[22] Mark W. Liu, Mutlu Ozdogan, and Xiaojin Zhu "crop type classification by simultaneous use of satellite images of different resolutions" IEEE transactions on geosciences and remote sensing, vol. 52, no. 6, June 2014.

[23] S. Chen and D. Zhang, "Robust Image Segmentation using FCM with Spatial Constraints based on new Kernel-induced Distance Measure," IEEE Transactions on Systems, Man and Cybernetics, vol. 34, no. 4, pp. 1907-1916, 2004.