International Journal of Computer Applications Technology and Research
Volume 2– Issue 3, 270 - 276, 2013, ISSN: 2319–8656

Image Segmentation Using Two Weighted Variable Fuzzy K Means

S. Suganya
CMS College of Science and Commerce
Coimbatore, Tamil Nadu, India

Rose Margaret
CMS College of Science and Commerce
Coimbatore, Tamil Nadu, India

Abstract: Image segmentation is the first step in image analysis and pattern recognition. Image segmentation is the process of dividing an image into different regions such that each region is homogeneous. The accurate and effective algorithm for segmenting image is very useful in many fields, especially in medical image. This paper presents a new approach for image segmentation by applying k-means algorithm with two level variable weighting. In image segmentation, clustering algorithms are very popular as they are intuitive and are also easy to implement. The K-means and Fuzzy k-means clustering algorithm is one of the most widely used algorithms in the literature, and many authors successfully compare their new proposal with the results achieved by the k-Means and Fuzzy k-Means. This paper proposes a new clustering algorithm called TW-fuzzy k-means, an automated two-level variable weighting clustering algorithm for segmenting object. In this algorithm, a variable weight is also assigned to each variable on the current partition of data. This could be applied on general images and/or specific images (i.e., medical and microscopic images). The proposed TW-Fuzzy k-means algorithm in terms of providing a better segmentation performance for various type of images. Based on the results obtained, the proposed algorithm gives better visual quality as compared to several other clustering methods.

Keyword —Fuzzy-K-means Clustering (FKM), image segmentation, W-k-Means, variable weighting

1. INTRODUCTION

Image segmentation techniques play an important role in image recognition system. It helps in refining our study of images. One part being edge and line detection techniques highlights the boundaries and the outlines of the image by suppressing the background information. They are used to study adjacent regions by separating them from the boundary

Clustering is a process of grouping a set of objects into classes of similar characteristics. It has been extensively used in many areas, including in the statistics [1], [2], machine learning [3], pattern recognition [4], data mining [5], and image processing [6]. In digital image processing, segmentation is essential for image description and classification. The algorithms are normally based on similarity and particularity, which can be divided into different categories; thresholding template matching [7], region growing [8], edge detection [9], and clustering [10]. Clustering algorithm has been applied as a digital image segmentation technique in various fields. Recently, the application of clustering algorithms has been further applied to the medical field, specifically in the biomedical image analysis wherein images are produced by medical imaging devices. The most widely used and studied is the K-means (KM) clustering. KM is an exclusive clustering algorithm, (i.e., data which belongs to a definite cluster could not be included in another cluster).

There are several clustering algorithms proposed to overcome the aforementioned weaknesses. Fuzzy K-means (FKM), an overlapping clustering that employs yet another fuzzy concept, allows each data to belong to two or more clusters at different degrees of memberships. In the FKM, there is no

clear, significant boundary between the elements if they do, or do not belong to a certain class. In 2010, [11] successfully proposed a modified version of K-means clustering, namely, adaptive Fuzzy k -Means (AFKM) clustering. The study proved that AFKM possesses a great ability in overcoming common problems in clustering, such as dead centers and centre redundancy. In this paper, we introduce a new version of clustering algorithm called two weighted variable - Fuzzy-K-means (TWvFKM) clustering algorithm. In this algorithm to build a cluster-based classification model automatically. In the TWv-Fuzzy k-means algorithm, to distinguish the impacts of different views and different variables in clustering, the weights of views and individual variables are introduced to the distance function. The view weights are computed from the entire variables, whereas the weights of variables in a view are computed from the subset of the data that only includes the variables in the view. Therefore, the view data, while the variable weights in a view only reflect the importance of variables in the view. We present an optimization model for the TWv-Fuzzy-k-means algorithm and introduce the formulae, derived from the model, for computing both view weights and variable weights. K-means algorithm as an extension to the standard -means clustering process with two additional steps to compute view weights and variable weights in each iteration. TW-k-means can automatically compute both view weights and individual variable weights. Moreover, it is a fast clustering algorithm which has the same computation complexity as k-means and FKM. We compared TWv-Fuzzy-k-means (TWvFKM) with various clustering algorithms (K-means, FKM, and AFKM) and the results have shown that the TWv-Fuzzy-k-means algorithm significantly out performed the other algorithms

This paper is organized as follow: Section II give details of the Image Segmentation Section III describes in detail the proposed TWvFKM clustering algorithm. Section III presents the data used and also discusses the type of analyses applied to test the capability of the proposed algorithm. Section IV presents the segmentation results obtained by the proposed algorithm. In addition, a comparison of performance comparison with several selected conventional clustering algorithms is also presented. The comparison is done based on both qualitative and quantitative analyses. Finally, Section V concludes the work focused on of this paper

2. IMAGE SEGMENTATION

Image Segmentation is the process of dividing a digital image into constituent regions or objects [12]. The purpose of segmentation is to simplify the representation of an image into that which is easier to analyze. Image segmentation is typically used to locate objects and boundaries in images. Segmentation algorithms are based on the two basic properties of an image intensity values: discontinuity and similarity. The first step in image analysis is segment the image.

Segmentation subdivides an image into its constituent parts or objects. The level to which this subdivision is carried depends on the problem being viewed. Some time need to segment the object from the background to read the image correctly and identify the content of the image for this reason there are two techniques of segmentation, discontinuity detection technique and Similarity detection technique. In the first technique, one approach is to partition an image based on abrupt changes in gray-level image. The second technique is based on the threshold and region growing.

2.1 Image Segmentation by Clustering

Clustering is a classification technique. Given a vector of N measurements describing each pixel or group of pixels (i.e., region) in an image, a similarity of the measurement vectors and therefore their clustering in the N-dimensional measurement space implies similarity of the corresponding pixels or pixel groups. Therefore, clustering in measurement space may be an indicator of similarity of image regions, and may be used for segmentation purposes.

The vector of measurements describes some useful image feature and thus is also known as a feature vector. Similarity between image regions or pixels implies clustering (small separation distances) in the feature space.

Clustering methods were some of the earliest data segmentation techniques to be developed.

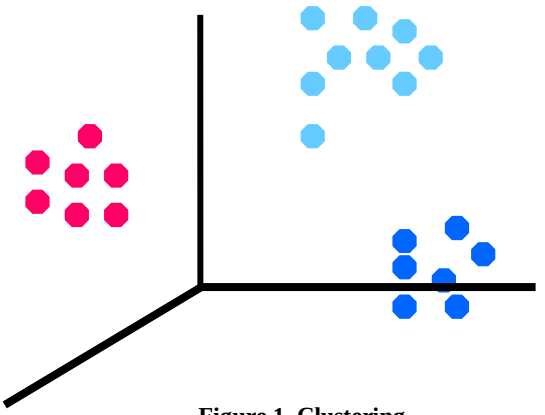


Figure 1. Clustering

Similar data points grouped together into clusters.

Most popular clustering algorithms suffer from two major drawbacks

- **First**, the number of clusters is predefined, which makes them inadequate for batch processing of huge image databases
- **Secondly**, the clusters are represented by their centroid and built using a Euclidean distance therefore inducing generally an hyperspheric cluster shape, which makes them unable to capture the real structure of the data.
- This is especially true in the case of color clustering where clusters are arbitrarily shaped

3. PROPOSED ALGORITHM

3.1 Fuzzy –K- Mean of Clustering Algorithm

TWvFKM is a clustering algorithm, which partitions a data set into clusters according to some defined distance measure. Images are considered as one of the most important medium of conveying information. Understanding images and extracting the information from them such that the information can be used for other tasks is an important aspect of Machine learning. One of the first steps in direction of understanding images is to segment them and find out different objects in them. To do this, we look at the algorithm namely TWvFKM clustering. It has been assumed that the number of segments in the image is known and hence can be passed to the algorithm.

Proposed algorithm namely TWv-Fuzzy-k-means clustering algorithm that can automatically compute variable weights in the k-means clustering process. TWV-Fuzzy-k-means extends the standard k-means algorithm with one additional step to compute variable weights at each iteration of the clustering process. The variable weight is inversely proportional to the sum of the within-cluster variances of the variable. As such, noise variables can be identified and their affection of the cluster result is significantly reduced. This TWvFKM weights both views and individual variables and is an extension to W-*k*-means. Domeniconi et al. [15] have proposed the Locally Adaptive Clustering (LAC) algorithm which assigns a weight to each variable in each cluster.

TWvF-K- Mean algorithm is one of the most important clustering algorithms, the first samples are divided into two or more clusters. In this fuzzy algorithm the number of clusters has been already specified. In FWv-Fuzzy- K Mean of clustering algorithm the main function is:

$$J = \sum_{I=0}^{C} \sum_{k=0}^{n} U_{ik}^{m} d_{ik}^{2} = \sum_{I=0}^{C} \sum_{k=0}^{n} U_{ik}^{m} |x_k - v_i|_k^2$$

In formula 1: m is a real number which is bigger than 1. In most of the cases, m=2. If m=1, the non-fuzzy c-mean of main clustering function is obtained. In above formula X^k is the k^{th} sample, and V^i is the center of it he cluster and n is the number of samples. U^{ik} shows the dependency of I^{th} sample in k^{th} cluster. $|X|$ is determined the similarity of sample(distance) from the center of cluster and can use every function that shows the similarity of sample or the center of cluster.

Steps of k-fuzzy mean algorithm [13]:

- For the first clusters initial value for k, m, and U should be estimated.
- The center of clusters should be calculated by second formula.
- The dependence matrix should be calculated by in second step.

If $||U^{l+1}-U^{l}|| \leq \varepsilon$ the algorithm is finished, visa versa go to second step.

3.2 The TW-Fuzzy-k-means Clustering Algorithm

Input: The number of clusters k and two positive real parametersλ,η

Output: Optimal values **of** U, Z, V and W

randomly choose K cluster centres Z^0**:**

For t=1 to T **do**

$w_t^0 \leftarrow 1/T$

For all j ∈ G_t **do**

$v_j^0 \leftarrow 1/|G_t|$

End for

End for

r ← 0

Repeat

Update U^{r+1}

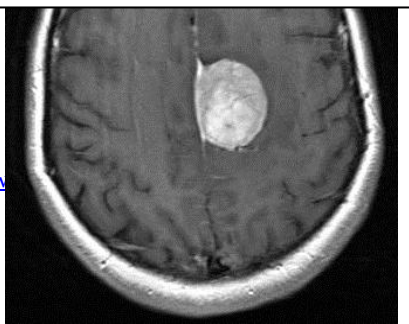
Update Z^{r+1}

Update V^{r+1}

Update W^{r+1}

r ←r+1

until: the objective function obtained its local minimum value;



4. EXEPERIMENT RESULTS

The experiments on the medical images have been carried out in MATLAB v7.10 TWv-Fuzzy - K-means segmentation is a clustering based segmentation algorithm. In clustering based segmentation changing in the distance metric will change the output. Euclidean distance is the default distance used in the algorithm, replacing it with the cosine distance gives better segmented areas in the medical images. In the Figure 2 we can see the original medical images. Figure 3 shows the cluster index images by the applying variable weight is 7 in Figure 2. Now compare it with Figure 4, which are cluster index images by applying variable weighting is 10. We can see that segmentation of areas is good in Figure 5 than in other figures. The Figure 5 has variable weighting in 15. The Figure 6 is another resulted image an applying weight is 20. Comparing those images the Figure 5 is better than another. It has variable weighted is 15. Now we analyze various images to apple TWv-Fuzzy-k-means with weight 15, the table1 has resulted images. Table 1 shows various image analysis results.

Figure 2. The Original Medical Image

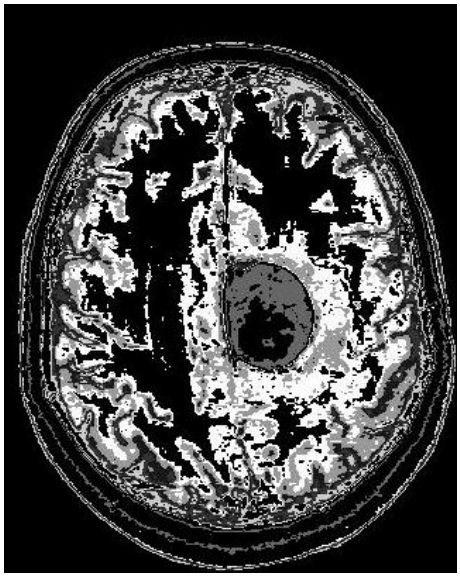


Figure 3: medical image has weighed is 7

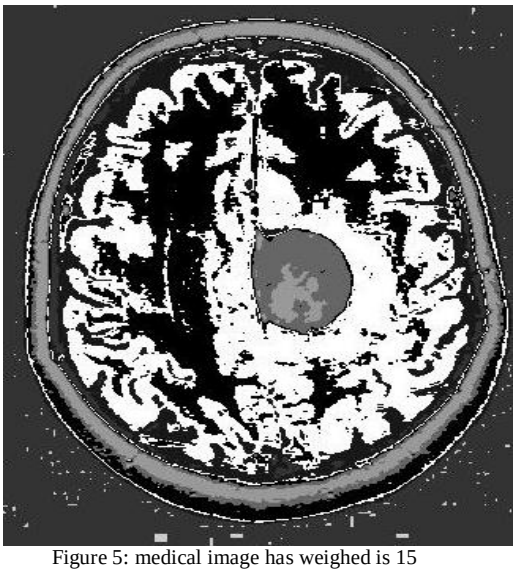


Figure 5: medical image has weighed is 15

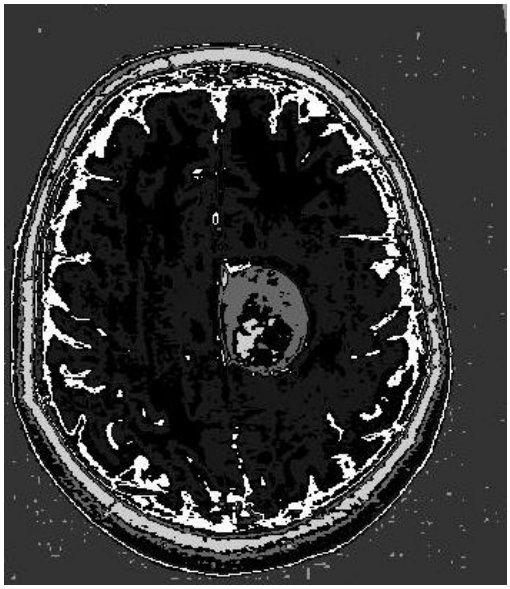


Figure 4: medical image has weighed is 10

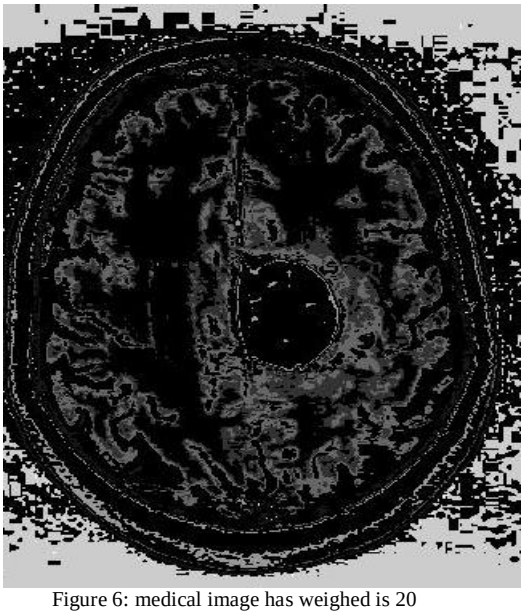
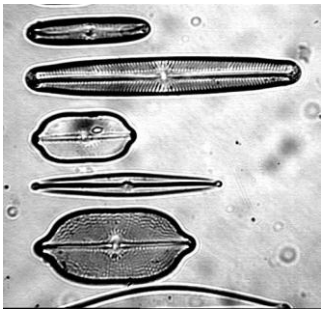
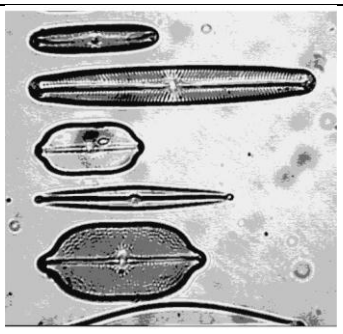
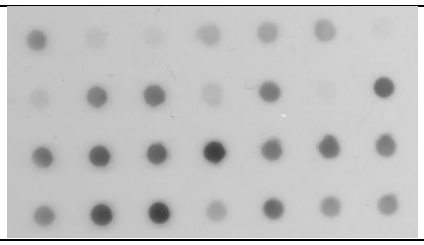
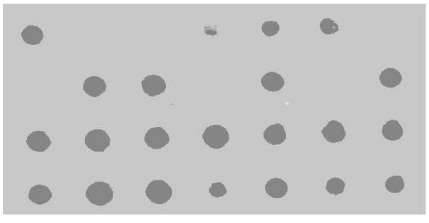


Figure 6: medical image has weighed is 20

Table 1: Analyzing multiple images with algorithm TWv-Fuzzy-k-means with weight 15

Image name	Original	Resulted image Weighted at 15
LENA		
Boat		
Bridge		
Diatoms		
Dot blot		

5. CONCLUSION

This paper presents a new clustering algorithm named the Two Weighted variable Fuzzy-K-Means algorithm for segmentation purposes. TWv-Fuzzy-*k*-means can compute weights for views and individual variables simultaneously in the clustering process. With the weights effect of low-quality views and noise variables can be reduced. Therefore, TWv-Fuzzy-*k*-means can obtain better clustering results than individual variable weighting clustering algorithms from multi-view data. We discussed the difference of the weights between TWv-Fuzzy-*k*-means. For medium values weighted of Fuzzy k-means algorithms give good results. For larger and smaller values of Weight, the segmentation is very coarse; many clusters appear in the images at discrete places. The conclusion of this paper sees the proposed algorithm outperforming the conventional FCM, AFKM and MKM algorithms by successfully producing better segmented images. The proposed TWvFKM also successfully preserves important features on digital images. Thus, it is recommendable for this algorithm to be applied in the post image processing in consumer electronic products such as the digital camera for general applications and the CCD camera which is extensively used with the microscope in capturing microscopic images, especially in segmenting medical images.

6. REFERENCES

[1]. D. Auber, and M. Delest, "A clustering algorithm for huge trees," *Advances in Applied Mathematics*, vol. 31, no. 1 pp. 46-60 2003.

[2]. S. Mahani, A. E. Carlsson, and R. Wessel, "Motion repulsion arises from stimulus statistics when analyzed with a clustering algorithm," *Biological Cybernetics*, vol. 92, no. 4, pp. 288-291, 2005.

[3]. T. Abeel, Y. V. d. Peer, and Y. Saeys, "Java-ML: A machine learning library," *Journal of Machine Learning Research*, vol. 10, pp. 931-934, 2009.

[4]. M. J. Rattigan, M. Maier, and D. Jensen. "Graph clustering with network structure indices," in *Proceedings of the 24th International Conference on Machine Learning*. 2007. Corvallis, OR.

[5]. H. Wang, W. Wang, J. Yang, and P. S. Yu, "Clustering by pattern similarity in large data sets," in *Proceedings of the 2002 ACM SIGMOD International Conference on Management of data.* 2002, Madison, Wisconsin

[6]. S. K. Singh, K. Shishir, G. S. Tomar, K. Ravi, and G. K. A. Santhalia,"Modified framework of a clustering algorithm for image processingapplications," *in First Asia International Conference on Modelling & Simulation, AMS '07,* 2007.

[7]. S. K. Warfield, K. Michael, F. A. Jolesz, and K. Ron, "Adaptive, template moderated, spatially varying statistical classification," *Medical Image Analysis*, vol. 4, no. 1, pp. 43-55, 2000.

[8]. N. A. Mat-Isa, M. Y. Mashor, and N. H. Othman, "Automatic seed based region growing for pap smear image segmentation," in *Kuala Lumpur International Conference on Biomedical Engineering.* 2002. Kuala Lumpur, Malaysia.

[9]. J. K. Paik, Y. C. Park, and S. W. Park, "An edge detection approach to digital image stabilization based on tri-state adaptive linear neurons," *IEEE Transactions on Consumer Electronics*, vol. 37, no. 3, pp. 521-530, 1991

[10]. X. Yang, Z. Weidong, C. Yufei, and F. Xin., "Image segmentation with a fuzzy clustering algorithm based on Ant-Tree," *Signal Processing*, vol. 88, no. 10, pp. 2453-2462, 2008.

[11]. S. N. Sulaiman and N. A. M. Isa:, "**Adaptive Fuzzy-K-means Clustering Algorithm for Image Segmentation**," IEEE Transactions on Consumer Electronics, Vol. 56, No. 4, November 2010

[12]. N.Senthilkumaram, R.Rajesh "Edge detection techniques for image segmentation-A survey of soft computing approach", International journal of recent trends in engineering , vol.1, No.2, May 2009,pp. 250-254

[13]. Farhad Soleimanian Gharehchopogh, Neda Jabbari, Zeinab Ghaffari Azar ," Evaluation of Fuzzy K-Means And K-Means Clustering Algorithms In Intrusion Detection Systems", International Journal Of Scientific & Technology Research Volume 1, Issue 11, December 2012

[14]. Vance Faber,‖Clustring and the Continuous K-means Algorithm‖, Los Almas since Number22, pp: 138-144, 1994.

[15]. C. Domeniconi, D. Gunopulos, S. Ma, B. Yan, M. Al-Razgan, and D. Papadopoulos. Locally adaptive metrics for clustering high dimensional data. Data Mining and Knowledge Discovery, 14(1):63–97, 2007.