A review and analysis of fuzzy-c means clustering techniques

Kalema Kiweewa Abdu, Bukenya Faiza,.Rose Alinda Alias
Faculty of Computing, Universiti Tekinologi Malaysia

Abstract-In research, change detection result accuracy does not only depend on the image registration accuracy but also depends on the segmentation and classification accuracy since the errors that arise from image registration are minimised during the segmentation and classification phase. Among the techniques that have been proposed to deal with the misalignment problem that arises from image registration is fuzzy c-means clustering (this technique is a combination of c-means used during grouping pixels in to objects and fuzzy logic used for classification of image objects in to changed and un-changed objects) . Fuzzy c-means clustering technique works by partitioning the image pixels into a set of fuzzy clusters through an iterative optimisation of the objective function adopted from Tariq Rashid, with the update of membership and the cluster centre. Many fuzzy c-means clustering techniques have been proposed by researchers to deal with the image registration errors. This paper reviews some of them and Fuzzy-Statistics-Based Affinity Propagation C-means clustering for segmentation-classification phase is proposed.This is donesince the method is fast, it enables incorporation of spatial information, which leads to prevention of noise as well as the intensity variation in images. It also treats clusters equally which prevents poor solutions caused by hard decisions during clustering, hence avoiding misclassification of clusters especially small clusters, leading to quality segmentation and good classification results.

Index Terms— Image Segmentation, Change Detection, classification, Fuzzy c-means clustering

INTRODUCTION

In change detection, some of the crucial steps that determine the change detection result accuracy are the segmentation-classification phase. In this phase, its where the image registration errors (i.e misalignment problem) are reduced, therefore better techniques are required for better change detection results. According to [11][10][9] true changes can be isolated from false changes if Image object information is manipulated. Researchers have tried to come up with techniques that can minimise the image registration error (ie misalignment problem) but research shows there is still much needed. Among the techniques that have been proposed to be good in dealing with the misalignment problem that arise from image registration is fuzzy c-means clustering [12][13][14]. Fuzzy c-means clustering technique works by partitioning the image pixels into a set of fuzzy clusters through an iterative optimisation of the objective function adopted from Tariq Rashid, with the update of membership and the cluster centre. In this algorithm, centroids are defined for each of the clusters and the quality of results depend on the initial set of clusters, the value of centroid and the number of iteration[8][7]. Many fuzzy c-means clustering techniques have been proposed by researchers to deal with the image registration errors. Section 2.0 of this paper reviews some of the fuzzy c-means clustering and threshold level based fuzzy c means clustering is proposed. Section 3 discusses the findings and also concludes the paper.

2 FUZZY C-MEAN CLUSTERING TECHNIQUES

Among the presented fuzzy clustering techniques, threshold level-based fuzzy clustering is the only fuzzy clustering technique that caters for intensity variation in the images, besides being robust to noise and intra-class dependencies. Threshold level- based fuzzy clustering was applied on the biological image during segmentation of cell, membership connectedness; size-weighted fuzzy clustering and fuzzy statistics-based affinity propagation were applied on multispectral images. Spatially weighted fuzzy c-means (SWFCM) clustering algorithm was applied on synthetic images. The five recently presented fuzzy clustering techniques are described in the subsections that follow.

(I) SPATIALLY WEIGHTED FUZZY C-MEANS (SWFCM) CLUSTERING ALGORITHM

The technique was presented by [3] incorporating the spatial neighbourhood information into the standard FCM clustering algorithm. In this algorithm, k-nearest neighbour (k-NN) algorithm was used to calculate the weight in the SWFCM algorithm in order to enhance the performance of image thresholding. Iterations were carried out with gray level histogram thresholding of the image for a faster segmentation process. According to the author SWFCM is fast and the algorithm is less sensitive to noise since it incorporates spatial information.

But however better segmentation results would be achieved if multi-thresholding was applied instead of single level thresholding.

(ii) THRESHOLD LEVEL-BASED ALGORITHM BASED ON 3 CLASS FUZZY C-MEAN CLUSTERING

The technique was proposed by [5], inorder to segment the cells. In this technique, segments the images (sensed image and reference image) into three classes using FCM clustering thresholding. The threshold is obtained by averaging the maximum in the class with the smallest centre and the minimum in the class with the middle centre. According to the author, the method is robust to noise, prevents weak-edge leakage and has the ability to handle interior intensity variation, also it can used for further analysis but however the method was not tried on othercell types or rather on other applications.

(iii) MEMBERSHIP CONNECTEDNESS METHOD

[4] presented this technique in order to utilize spatial information, in this algorithm, a combination of fuzzy connectedness and fuzzy clustering were applied during the functional form of membership connectedness in order to automatically select the required seeds. According to the author the membership connectedness segmentation method considers the local and global spatial relations which leads to reduced sensitivity to noise and false segmentation results, the method also utilizes the expert knowledge in the form of selected seeds and segments the images as desired by the expert which increases the chance of its success.

However the method is faced with the inherent intraclass redundancies in multispectral images caused by difference in the cluster size (the big cluster influence the learning stage which lead to small clusters from being misclassified or even missed during segmentation) which seriously decrease the accuracy and efficiency of the method, it also leads to spatial dependency caused by similarity of the spectral information of neighbouring pixels in the image.

(iv) SIZE-WEIGHTEDFUZZY CLUSTERING

The technique was presented in order to deal with the spatial and the intraclass redundancy that exists in the image faced by the membership connected method, in this unsupervised algorithm, watershed transform was used to deal with the spatial redundancy., Size-weighted fuzzy clustering was also used to deal with intra-class redundancy. During watershed phase, the spatial redundancy is omitted, and in the Membership connectedness construction phase, the spatial relation among image pixels is taken into account. According to the [2], the method can not only detect small regions (that often appear in remote-sensing images) but also it can detect the overlapped regions in the satellite images, although the method does not provide detailed information about the clusters.

(v) FUZZY STATISTICS-BASED AFFINITY PROPAGATION (FS-AP)

The technique was presented in order to extract land-cover information in multispectral images, in this algorithm, affinity propagation (a clustering algorithm) was used to find clusters with small error in large datasets and fuzzy statistical similarity measure was used to determine the clusters to which the pixels belong to through obtaining objective estimates of how closely two pixel vectors resemble each other. According to the [6], the method is fast ,robust to noise, and works well in the case of mixed pixels, since it treats all pixels equally which reduces on dependence and avoids poor solutions caused by hard decisions during clustering. However fuzzy statistics-based affinity propagation works well with application of multispectral remote sensing images.

(vi) ANT COLONY OPTIMIZED FUZZY C-MEANS CLUSTERING

The techniques was presented by [14] to obtain obvious edge and segmentation result, in this algorithm the centres and number of clusters are determined by ant colony optimization algorithm, then later initialization fuzzy C-means algorithm is used for remote sensing image classification. According to[14] the method is fast and the visual interpretation of segmentation results is improved but the method is good on land cover/use detection, i.e. Impervious area, water, etc.

2.1 COMPARISON OF FUZZY C-MEAN CLUSTERING TECHNIQUES

Table 2 shows comparison of each of the fuzzy clustering methods discussed previously emphasising their advantages, disadvantages, the types of their input they can handle. Among the new presented fuzzy clustering techniques, the threshold level based fuzzy clustering is the only fuzzy clustering technique that has the ability to handle the intensity variations in the image.

TABLE 2
COMPARISON OF THE FUZZY CLUSTERING METHODS

Author	Technique	Objective	Advantage	Disadvantage	input

ISSN 2229-5518

Yong, Y., Z. Chongxun.2004	Spartially weighted fuzzy c-means clustering	To incorporate spartial neighbourhood information in to standard FCM	-it is fast -it less sensitive to noise	Single level thresholding	Synthetic image
Guanglei Xiong, 2006	Threshold level based algorithm based of fuzzy c-means algorithm	To segment cell	-to prevent weak edge leakage -it is robust to noise -interior intensity variation, incorporates spatial data	-It was not tried on other different cell types	- Drosophila RNAi Fluorescence Cellular Images
Maryam Hasanzadeh and Shohreh Kasaei,2008	Membership Connectedness	To utilise global spatial relations besides the fuzzy clustering stage in order to improve segmentation accuracy.	-reduced sensitivity to noise. -reduced false segmentation results. - can utilize the expert knowledge in the form of selected seeds and segments	- faced with intraclass redundancies. -it also leads to spatial dependency	- multispectral image
M. Hasanzadeh and S. Kasaei, Senior Member, IEEE,2010	Size-weighted fuzzy clustering method.	to deal with the spatial and the intra-class redundancy that exist in the image faced by the membership connected method	-to detect small regions that often appear in remote sensed images -It can detect the overlapped regions in the satellite images.	-the method does not provide detailed information about the clusters	- multispectral image

Chen Yang, Bruzzone, L,2010	A Fuzzy-Statistics-Based Affinity Propagation Technique for Clustering	- To extract land-cover information in multispectral images	-it is fast -it is roburst to noise - Reduces on dependence among pixels. - Avoids poor solutions caused by hard decisions during clustering	-works well with application of multispectral remote sensing images. - Does not handle variation in the intensity	-multispectral image
Jingfeng Yan,2014	ant colony optimized fuzzy C-means clustering	to obtain obvious edge and segmentation result.	-improved visual interpretation of segmentation results. -Reduces on the computation time.	- good for land cover/use detection, i.e. impervious area, water, etc.	Middle spatial resolution remote sensing images.

3 DISCUSSION AND CONCLUSION

During segmentation, incorporation of spatial information is an important practice in order to avoid noise, and also to treat clusters equally. In order to avoid poor solutions caused by hard decisions (pixels finding it hard to decide cluster to belong to) during clustering hence avoiding misclassification of clusters especially small clusters leading to quality segmentation and classification results. Subsequently intensity variation (which is as result of image being captured by sensor at different intensity) is an important issue as it may lead to misclassification of the pixels, (water pixel may be treated as non-water pixel) which leads to false segmentation results hence leading to misclassification of object which leads to false changes in the changed image.

According to the table 1, In 2004,[3] incorporated spartial neighbourhood information in to standardize, FCM in order to deal with the reduced sensitivity to noise, the method is fast and less sensitive to noise however, its based on single level thresholding which minimises the chances of getting a better scale for better results, and yet according to [5] a small scale yields better segmentation classification results. The multilevel thresholding technique introduced by [5]enable segmentation of image at many scales which increases the chances of getting a better result. The method also prevents weak edge leakage, interior intensity variation and incorporates spatial data which reduces sensitivity to noise but the method was is good on biological images.

In 2008, [4] reduced sensitivitity to noise by utilising global spatial relations using the membership connectedness in order to utilize the expert knowledge in the form of selected seeds and segments which reduced sensitivity to noise, reduced false segmentation results, however the method has a problem of intraclass redundancies and spatial dependency. [2], dealt with the spatial and the intra-class redundancy that exist in the image faced by the membership connected method with the use of a Size-weighted fuzzy clustering method. This method is capable of detecting small regions that often appear in remote sensed images and can also detect the overlapped regions in the satellite images but the method does not provide detailed information about the clusters. In the same year, [6],introduced A Fuzzy-Statistics-Based Affinity Propagation Technique for Clustering in order to extract land-cover information in multispectral images, the method avoids poor solutions caused by hard decisions during clustering, reduces dependence among pixels, fast and robust to noise but works well with application of multispectral remote sensing images. And does not handle variation in the intensity which increases sensitivity to noise..Recently,[14] optimized fuzzy C-means clustering in order to extract obvious edge and segmentation result from a SAR image, but the method is good on land cover/use detection.

According to the above comparison, a combination of a Fuzzy-Statistics-Based Affinity Propagation Technique and threshold level based fuzzy C-means clustering would yield better segmentation and classification results, since threshold level based fuzzy c-means clustering enable incorporation of spatial information which leads to prevention of noise and deals with intensity variation in images. Moreover, Fuzzy-Statistics-Based Affinity Propagation technique treats clusters equally which prevents poor solutions caused by hard decisions during clustering, hence avoiding misclassification of clusters especially small clusters leading to quality segmentation and classification results. The novelity of this paper is it e proposes a Fuzzy-Statistics-Based Affinity Propagation C-means clustering for segmentation-classification phase, and that makes up the next stage of this research..

REFERENCES

[1] Richards, J. A. and Jia, X, "Remote Sensing Digital Image Analysis: An Introduction.", 2006. Springer-Verlag, Heidelberg, New York, pp. 67-68; pp. 128-130; pp. 342-352; pp. 243.

[2] Hasanzadeh, M. and S. Kasaei, "A multispectral image segmentation method using size-weighted fuzzy clustering and membership connectedness",2010. IEEE Geoscience and Remote Sensing Letters 7(3): 520-524.

[3] Yong, Y., Z. Chongxun. "A Novel Fuzzy C-Means Clustering Algorithm for Image Thresholding",2004. Measurement Science Review 4(1): 11-19.

[4] Maryam Hasanzadeh and Shohreh Kasaei, "Fuzzy image segmentation using membership connectedness",2008.

[5] Guanglei Xiong, X. Zhou,L. ji,. Bradley.S.Wong, "Automated segmentation of Drosophila RNAi fluorescence cellullar Images using deformable models", IEEE Transactions on circuits andsystems.vol 53. No.11,pp.2415-2424, 2006.

[6] Chen Yang, Bruzzone, L, "A Fuzzy-Statistics-Based Affinity Propagation Technique for clustering in multispectral images", 2010.

[7] Andrew Moore,"A Tutorial on Clustering Algorithms",website: http://www.autonlab.org/tutorials/kmeans.html accessed on24th /09/2011.

[8] HongLei Yang,JunHuan Peng,BaiRu Xia,DingXuan Zhang, "An improved algorithm for remote sensing classification",2013.

[9] Bovolo, F, "A multilevel parcel-based approach to change detection in very high

resolution multitemporal images", 2009. IEEE Geoscience and Remote Sensing Letters 6(1): 33-37

[10] F.Bovolo, CMarin,L. Bruzzone, "A Hierarchical Approach to change detection in very high resolution multitemporal SAR mages", 2013

[11] Pacifici F.;Del Frate,F.;Solimini,C.;Emery,W.j, "An innovative neural net method to detect temporal changes in high resolution optical satellite imagery",2007

[12] Usman seljud,Rashid Hussein, "SAR Image clustering by fuzzy c-means clustering technique with thresholding for icebags",2014.

[13] B. Faiza,S. S.Yuhaniz,S. Z. Mohd Hashim,A. K. Kalema, “Detecting Floods Using an Object Based Change Detection Approach”,2012

[14] Jingfeng Yan, “Remote sensing image segmentation based on ant colony optimized fuzzy C-means clustering”, 2014