Paper Denoising-based Clustering Algorithms for -Corrupted Images

ANNAMALAI K
Lecturer
Instrumentation and Control Engineering,
Seshasayee Institute of Technology, Tiruchirappalli - 10

Abstract —Clustering algorithm is a widely used segmentation method in image processing applications. The algorithm can be easily implemented; however in the occurrence of noise during image acquisition, this might affect the processing results. In order to overcome this drawback, this paper presents a new clustering-based segmentation technique that may be able to find different applications in image segmentation. The proposed algorithm called Denoising-based (DB) clustering algorithm has three variations namely, Denoising-based-K-means (DB-KM), Denoising-based-Fuzzy Cmeans (DB-FCM), and Denoising-based-Moving K-means (DBMKM). The proposed DB-clustering algorithms are able to minimize the effects of the Salt-and-Pepper noise during the segmentation process without degrading the fine details of the images. These methods incorporate a noise detection stage to the clustering algorithm, producing an adaptive segmentation technique specifically for segmenting the noisy images. The results obtained quantitatively and qualitatively have favored the proposed DB-clustering algorithms, which consistently outperform the conventional clustering algorithms in segmenting the noisy images. Thus, these DB-clustering algorithms could be possibly used as pre- or post-processing (i.e., segmenting images into regions of interest) in consumer electronic products such as television and monitor with their capability of reducing noise effect.

Index Terms — clustering, image segmentation, saltand- pepper noise, image processing.

I. INTRODUCTION

Image segmentation is the process of dividing digital images into multiple meaningful regions or sets of pixels. Classically, image segmentation is defined as partitioning an image into separate regions which are homogeneous with respect to some characteristics such as gray level or texture. It is one of the important tasks in image analysis and is widely used in a variety of consumer electronics applications such as object recognition [1]-[3], geographical imaging [4], [5], robot vision [6] and medical imaging.Segmentation algorithms can be classified into different categories based on segmentation techniques used such as the features' thresholding [9], template matching [10], region- based technique [11], [12] and clustering [13], [14]. Those techniques have their own limitations and advantages in terms of suitability, performance and computational cost. For instance, the thresholding technique (providing that the threshold value selected is the suitable one) produces a good quality and rapidity segmentation, but it remains a fact that it is sensitive to noise. Template matching, however becomes time consuming when the image becomes more complex or larger in size while theregion-based technique also suffers from time-consuming and over-segmentation problems. Clustering is an unsupervised classification designed to group a set of data samples with similar characteristics into larger units of analysis (clusters). In image segmentation, clustering algorithm iteratively computesthe characteristics of each cluster and segments the image by classifying each pixel in the closest cluster according to a distance metric. Through clustering technique, a much better results of segmentation can be obtained but over-segmentation is one of the problems that must be faced. Yet, the results of segmentation algorithm (i.e., segmented images) are useful in many ways- it primarily provides easier interpretation of images by highlighting specific objects or features in the image, which serves to be one of the important features in consumer electronics applications. Segmenting images via clustering algorithm has been applied in various fields including in the medical field specifically in the biomedical image analysis. Several previous studies have proven that clustering algorithms is capable in segmenting and determining certain regions of interest on medical images [7], [8], [15]-[18]. It is because in the biomedical image segmentation task, clustering algorithm is often suitable since the number of clusters for the structure of interest is usually known from its anatomical information [14]. In general application, Fuzzy c-means (FCM) is one of the most frequent clustering-based segmentation methods used for image segmentation. However, one disadvantage of FCM is it is very sensitive to noise and other imaging artifacts those of which containing some spatial information [19]. Therefore recently, many researchers have incorporated local spatial information into the original clustering algorithm to improve the performance of image segmentation [19]- [21]. Cai et al. [19] has proposed fast generalized FCM which incorporates both the local special and gray information together to produce a fast and robust FCM framework for image segmentation. Ahmed et al. [20] has modified the object's function of the conventional FCM algorithm to compensate intensity inhomogenities and allow the immediate neighborhood to influence a pixel. The algorithm is called Bias-corrected FCM (BCFCM). They conclude that the BCFCM algorithm tested on MRI data produces good results in terms of its efficiency and effectiveness. Chen and Zhang [21] however, go on to improve BCFCM and propose two variants of algorithm which have simplified the neighborhood term of BCFCM to reduce the computational loads and applied it on the MRI data. Li et al. [22] in their research attempt to improve K-Means (KM) algorithm, an exclusive clustering algorithm that is widely used and applied mainly due to its simplicity and efficiency. However, KM suffers from several major problems such as [22]-[24]; 1. It is dependent on initialization, 2. It is sensitive to outliers and skewed distribution due to standard Euclidean distance employment 3. It treats each pixel as an independent data point. To address the above problems Li et al. [22], have incorporated the kernel-induced distance measurement and neighborhood information into K-harmonic means (KHM) algorithm and replaced the standard Euclidean distance with Mercer kernel-based metric. They also use Markov random fields (MRFs) to model and incorporate the spatial homogeneous information into the kernel-based KHM. They make reference to their new algorithm as spatial kernel-based K-harmonic means (SKKHM) clustering scheme and the algorithm has been found to exhibit a robust performance. Research on image segmentation is not limited to the gray level image. Chaabane et al. [25] propose a new method for the color image segmentation which is based on fuzzy homogeneity and data fusion techniques. Their experimental results demonstrate superiority of their techniques with a significant improved performance in segmentation. In this paper we introduce a version of adaptive clustering based segmentation techniques for low level Salt-and-

Pepper noise corrupted images. The concept behind this segmentation strategy lies on the fundamental needs on removing impulsive noises i.e., Salt-and-Pepper noise. Salt-and-Pepper noise contaminations are normally detectable at the minimum and maximum intensities and are largely caused by errors in the image acquisition, recording and/or transmission. It is important to eliminate Salt-and-Pepper noise contained in the image before subsequent image processing tasks like image segmentation is carried out. This is because the occurrence of Salt-and-Pepper noise can severely damage the information or data embedded in the original image. Conventionally, in the occurrence of Salt-and-Pepper noise in images, segmentation can only be applied after performing pre-processing tasks like filtering algorithm. To tackle this problem, we propose new clustering algorithms for segmenting noisy images by incorporating the noise detection stage with the clustering algorithm to introduce an adaptive clustering-based segmentation technique. The adaptive behavior enables the clustering algorithm to segment the noisy image properly even in the occurrence of low density Salt-and-Pepper noise without going through any filtering stage beforehand. The inherited noise detection behavior will improve the segmentation results by only selecting noise-free pixels for the process of segmentation. This paper is organized as follows. In Section II the implementation of the conventional clustering algorithms namely KM, FCM and Moving K-means (MKM) as image segmentation is presented. The proposed adaptive clustering algorithms are explained in Section III. Section IV analyses the results obtained from the proposed algorithm and its comparison with several selected conventional clustering algorithms by using both qualitative and quantitative methods. Finally, Section V concludes the work of this paper.

II. CONVENTIONAL CLUSTERING ALGORITHMS Generally, clustering methods can be categorized intohistorical, graph theoretic, decomposing a density function and minimizing an objective function [21]. In this paper, the focus is on clustering methods by minimizing an objective function and their application for segmenting images.

A K-means (KM) Clustering Basically, the conventional algorithm is used to group datato the nearest centre. Consider N as the number of data to beclustered into ncregions or clusters. Let vtthet-th data where t=1,2,…,N and is the k-thcentre. Mathematically, theconventional KM objective function of partitioning adataset into ckis given by [26]:

$J = vt - ck\ n,N\ 2\ K,t$=1 (1)

where . stands for a distance measure that is normally taken to be the Euclidean norm. All data will be assigned to the nearest centre based on the Euclidean distance. The new position for each centre is calculated using:

ck = 1 $nck\ t\in ck\ vt$ (2)

The process is repeated until the values of all centers no longer change.

B. Fuzzy C-means (FCM) Clustering For the conventional FCM, the objective function ofsegmenting an image as described above into ckclusters is given by [21]:

$J = (Mkt\ m\ )\ vt - ck\ n,N\ 2\ K,t$=1 (3)

wherem, the fuzziness exponent, is an integer, m > 1.

The new position for each centre is calculated using:

$ck = vtMkt\ N\ m\ t$=1 $Mkt\ N\ m\ t$=1 (4)

with each fuzzy membership function,$Mkt\ m$ satisfying

$Mkt\ m$ = 1 $dkt\ dlt$ 2 $n\ (m$ −1) $c\ l$=1 ; dlt > 0(5)

$if\ dlt$ = 0; $Mkt\ m$ = 1 $Mkt\ m$ = 0; $for\ t \neq l$ (6)

Where $dkt = vt - ck$ 2 All processes are repeated until the memberships or cluster centers for successive iteration differ by more than some prescribed value ε(where $\varepsilon$is a termination criterion value between 0 and 1). C. Moving K-Moving (MKM) Clustering

The MKM algorithm is the modified version of KMproposed by [27] where the new position for each cluster is calculated using (2). [27] has introduced the concept of fitness to ensure that each cluster should have a significant number of members and final fitness values before the new position of cluster is calculated. The fitness for each cluster is calculated using:

$f\ ck = vt - ck$ 2 $t\in ck$ (7) [31]

has stated a condition that all centers must satisfy the following condition: f(Cs) ≥ αa f(Cl) (8) whereαa is a small constant value, initially set to be equal toαo , αo is a designed constant with value in range 0<αo

III THE PROPOSED ADAPTIVE CLUSTERING-BASED SEGMENTATION TECHNIQUE We propose a new version of adaptive clusteringbasedsegmentation technique, particularly for images corrupted with a low level of Salt-and-Pepper noise. The technique, known as Denoising Based (DB) clustering is introduced to overcome the problem of noise sensitivity in the segmentation process and to increase the robustness of the clustering algorithms with respect to noise. The proposed technique is divided into two stages. The first stage involves the detection of Salt-and-Pepper noise's intensity and locations. The second stage will perform the clustering process. The 'noise-free' pixels will be totally considered as the input data and they will give full contribution on the clustering process. Otherwise, for the 'noise' pixels, the fuzzy concept is applied to determine the degree of contributions of these 'noise' pixels on the clustering process. The combination of these features (i.e., the noise detection and the clustering), allows more versatile and powerful methods to achieve a better segmentation especially on noisy images.

a) Noise Detection In this stage, the histogram of noisy image will be utilizedto estimate the two Salt-and-Pepper noise intensities. The detection stage begins by searching for two peaks of intensities, where according to [29], an image corrupted with the Salt-and-Pepper noise will produce two peaks at the noisy image histogram (i.e., either a positive impulse or a negative impulse). Let us consider an 8-bit gray scale digital image with 256 gray levels in the interval [0, 255]. Generally, a Salt-andPepper noise takes on the high-end and low-end intensities. It can either be positive or negative where the intensity value for the positive impulse is near 255 (i.e., appears white known as the salt), and the negative impulse with the intensity value of near 0 (i.e., appears black known as the pepper). These two Salt-and-Pepper noise intensities will be used to identify possible 'noise-pixels' in the image. As in [30], [31], a binary noise mask N(i,j) will be created to mark the location of 'noisepixels' by using;

$N(i,j)$ = 0, $X\ i,j = Lsalt\ orLpepper$ 1, $otherwise$ (13)

where X(i,j) is the pixel at the location (i,j) with intensity X, N(i,j)=1 represents the 'noise-free' pixel to be retained in the next clustering stage while N(i,j)=0 represents 'noise'pixels.

b) Clustering Process In order to allow more versatile and powerful methods ofclustering-based segmentation in noisy images, after the binary noise mask N(i,j) is created, a linearly-fuzzy weighted correction value of 'noise' pixel is obtained using:

$X' i,j = 1 - F\ i.j\ .X\ i,j + F\ i.j\ .\ M(i,j)$(14)

where X'(i,j) denotes the corrected 'noise' pixel value, M(i,j) is the median value of the considered pixel and its neighboringpixel in n×nwindow (i.e., n is an odd number and typicallyset to 3), and F(i,j) is the fuzzy membership used to weigh thelinear relationship between the processing pixel, X(i,j), and themedian pixel, M(i,j). Prior to that, the median of the 'noise' pixels is extracted ina 3×3 window as given by:

$m = median\ X(i + k, j + l)$ as k,l∈ (−1,0,1)(15)

After the median pixel is found, the absolute luminance difference, d (i, j) , is computed by using;

$d\ i + k, j + l = X\ i + k, j + l - X(i,j)$ (16)

Then the local information of the 'noise' pixels in 3×3window is calculated by taking the maximum value of the absolute luminance difference given by;

$D\ i,j = max\ d(i + k, j + l)$ (17)

The choice of the maximum operator rather than minimum operator is justified in [30]. Next the fuzzy concept is applied to the extracted local information, D(i, j).The fuzzy membership function F (i, j) is defined by; $F\ i,j = 0$;

$D\ i,j < T1\ D\ i,j - T1\ T2 - T1$;

$T1 \leq D\ i,j < T2\ 0$;

whereby for optimal performance, the threshold value T1 and T2 are set to 10 and 30 respectively as described in [30]. Then the corrected value of noise pixel is calculated using (14). To increase the robustness of KM clustering towards noise, these corrected values (i.e., for the noise pixels) are used to replace original pixels values during the process of assigning the data to their nearest centre. Then the new position for each cluster is calculated using (2). The term vtin (2) is substituted by:

$vt = X\ i,j\ if\ N\ i,j = 1\ X'\ i,j\ if\ N\ i,j = 0$ (19)

By employing this concept in the conventional KM clustering algorithm, the new proposed algorithm is called Denoisingbased-K-means (DB-KM). The same concept can also be applied to the conventional FCM and MKM. The modified versions of those techniques are called Denoising basedFuzzy C-means (DB-FCM) and Denoising-based- Moving Kmeans (DB-MKM) respectively.

IV. EXPERIMENTAL RESUL

In this section, we present the experimental results onseveral standard real images. There is a total of six algorithms used in this study namely the conventional KM, the conventional FCM, the conventional MKM, the proposed DBKM, DB-FCM, and DB-MKM. The results of the proposed algorithms (i.e., the DB-clustering algorithms) are compared with the conventional clustering algorithms (i.e., KM, FCM, MKM).

a) Qualitative Analysis We execute the three proposed clustering algorithms as wellas three conventional clustering algorithms on several standard real test images contaminated by different levels of salt-and-pepper noise to investigate the

b) robustness of the algorithms. Five out of those tested images are chosen to enable the visualization of the performance of the proposed algorithms. The images are called House, Lighthouse, Barn, Barn2, and Butterfly as shown in Figs. 1(a), 2(a), 3(a), 4(a), and 5(a) respectively. Figs. 1(b), 2(b), 3(b), 4(b), and 5(b) are the same aforementioned images corrupted with 10%, 20%, 30%, 40%, and 50% density of salt-and-pepper noise respectively. The segmentation results are shown in Figs. (c) to (h). Fromthe results we can visualise that the results produced by the conventional clustering algorithms are influenced by the noise, which indicates that these algorithms are less robust to the noise mixture, as shown in Figs.(c) to (e). The proposed DBclustering algorithms (i.e., DB-KM, DB-FCM and DBMKM) can significantly remove the effect of noise added to the images (refer to Figs. (f) to(h)). The proposed DBclusteringalgorithms therefore, have achieved satisfactory results. For example, we may be able to notice that for the Houseimage shown in Fig. 1, the proposed DB-clustering algorithmsgive better and clearer segmentation results compared to theconventional clustering algorithms. The black and whiteparticles (i.e., pepper and salt noise respectively) aresignificantly able to be reduced. These findings prove that theproposed algorithms are robust with respect to the noiseeffect. As for the Lighthouse image which is contaminatedwith 20% of salt-and-pepper noise as illustrated in Fig. 2, thenoise contamination has highly affected the segmentationprocess in the conventional clustering algorithms whichcontribute to poor segmentation results.The proposed DB-clustering algorithm produces better segmented images as shown in the figure2(f)to(h). As bran image as shown in fig.3,the proposed DBclustering algorithm out perform the other conventional algorithms by segmenting the door,window and wheel clearly.

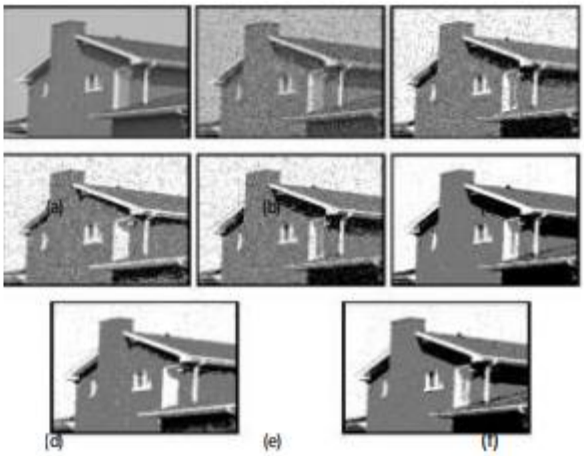
H

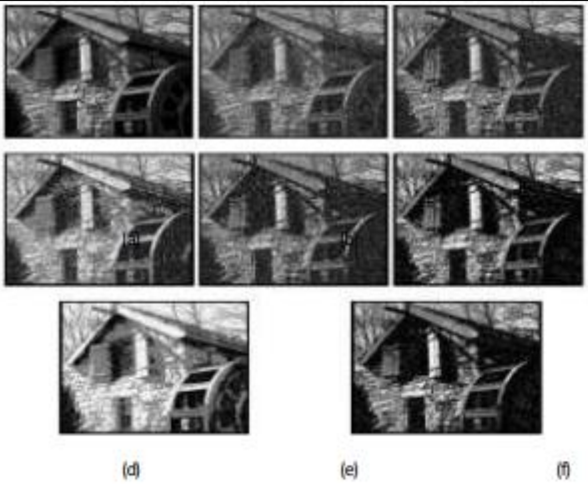
(a) (b) (c)

c) (b) (c) (d) (e) (f}

(g) (h)

Fig. 1. Segmentation results on House with 10% density of salt-andpepper noise ; (a) original image, (b) noisy image, (c) conventional KM (d)conventional FCM, (e) conventional MKM, (f) DB-KM, (g) DB- FC





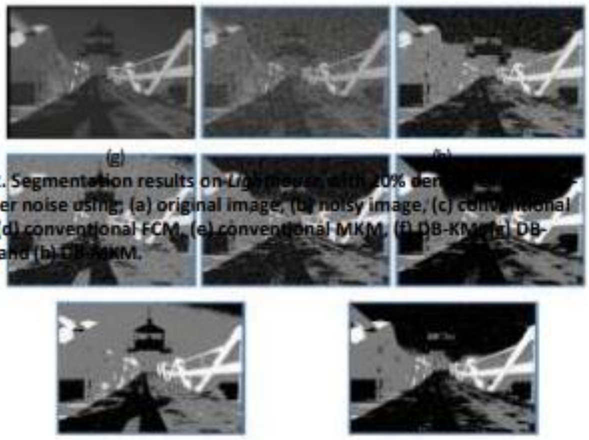


Fig. 2. Segmentation results on Lighthouse with 20% density of salt-and-pepper noise using; (a) original image, (b) noisy image, (c) conventional KM, (d) conventional FCM, (e) conventional MKM, (f) DB-KM, (g) DB-FCM and (b) DB-MKM.

Fig. 4. Segmentation results on *Barn2* with 40% density of salt-and-pepper noise; (a) original image, (b) noisy image, (c) using conventional KM, (d), using conventional FCM, (e) using conventional MKM, (f) using DB-KM, (g) using DB-FCM and (h) using DB-MKM.

V. CONCLUSION This paper presents new clustering-based segmentation algorithms named Denoising-based clustering algorithm for adaptive segmentation, i.e., for segmenting noisecorrupted images. The qualitative and quantitative analyses favor the proposed algorithms as good segmentation algorithms where the segmentation is concerned. The proposed algorithms also produce better results as compared to the conventional algorithms through its inclusion of the noise detection stage in its clustering process. This stage could reduce the effect of noise during the segmentation process. Simulation results show that the proposed algorithms are able to remove low density of salt-and-pepper noise (i.e.,up to 50%) during the segmentation process. In addition, the recommended algorithms have also successfully preserved important features on digital images. This finding is proven by smaller values ofF(I), F'(I), Q(I), and E(I) produced by the proposed DB-clustering algorithm. Furthermore, this finding suggests the DB-clustering as a good technique for the segmentation of noisy images, which could be used as preor post-processing technique in the consumers' electronics fields.

REFERENCES

[1] S. X. Yu, R. Gross, and J. Shi, “Concurrent Object Recognition and Segmentation by Graph Partitioning,” Advances in Neural InformationProcessing Systems, vol. 15, pp.1407- 1414, 2003.

[2] J. Shotton, J.Winn, C.Rother, and A. Criminisi, “Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling appearance, shape and context,” International Journalof Computer Vision, 2007.

[3] J. Shotton, A. Blake, and R. Cipolla, “Efficiently combining contour and texture cues for object recognition,” British Machine Vision Conf. 2008.

[4] N. Lee, R. T. Smith, and A. F. Laine, “Interactive Segmentation for Geographic Atrophy in Retinal Fundus Images,” Proceedings ofAsilomar Conference on Signals, Systems and Computers, IEEE Signal Processing Society, 2008.

[5] S. K. Antani, T.M. Deserno, L.R. Long, and G.R. Thoma, “Geographically distributed complementary content-based image retrieval systems for biomedical image informatics,” Proceedings of12th World Congress on Medical Informatics (MEDINFO), 2007.

[6] T. Kubik, and M. Sugisaka, “Image segmentation techniques and their use in artificial life robot implementation,” Artif Life Robotics, vol. 7, pp 12-15, 2003.

[7] N. A. Mat-Isa, M.Y. Mashor, and N. Othman, “Pap Smear Image Segmentation Using Modified Moving K-Mean Clustering,” International Conference on Biomedical Engineering. Kuala Lumpur Malaysia. 2002.

[8] N. A. Mat-Isa, M. Y. Mashor, N. H. Othman and S. N. Sulaiman, “Application of Moving K-Means Clustering for Pap Smear Image Processing,” Proceeding of International Conference on Robotics,Vision, Information and Signal Processing, Penang Malaysia, 2002.

[9] M. Cheriet, J.N. Said, and C.Y. Suen, “A Recursive Thresholding Technique For Image Segmentation,” IEEE Transactions on ImageProcessing, vol.7, no. 6, pp 918-921, 1998.

[10] 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.

[11] Y. Seunghwan, and P. Rae-Hong, "Red-Eye Detection And Correction Using Inpainting In Digital Photographs," IEEE Transactions onConsumer Electronics, vol.55, no.3, pp 1006-1014, 2009.

[12] N. A. Mat-Isa, M.Y. Mashor, and N.H. Othman, "Automatic Seed Based Region Growing for Pap Smear Image Segmentation," InternationalConference on Biomedical Engineering, Kuala Lumpur, Malaysia, 2002.

[13] 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.

[14] N. A. Mat-Isa, A. S. Samy, U. K. Ngah, "Adaptive fuzzy moving Kmeans algorithm for image segmentation," IEEE Transactions onConsumer Electronics, vol. 55, no. 4, pp. 2145-2153, 2009

[15] A. E. O. Boudraa, J. J. Mallet, J. E. Besson, S. E. Bouyoucef, and J. Champier, "Left Ventricle Automated Detection Method In Gated Isotopic Ventriculography Using Fuzzy Clustering," IEEE Transactionson Medical Imaging, vol. 12, no.3, pp 451-465, 1993