

Quality Measurement in Segmentation of Medical Images

G. Karthick

karthick.sgs@gmail.com

R.Harikumar

harikumarrajaguru@gmail.com

B.Vinoth Kumar

bojanvinoth@gmail.com

Abstract — The main objective of this paper is to calculate the performance of quality measures towards the segmentation of medical images using K-means clustering. Four types of medical images such as MRI, X-rays, CT and Ultrasonic images are studied. The K-means clustering shows that the non-intactness of the clusters. As cluster size increases the edges are brittle and compactness of the clusters get altered. The quality measures like PSNR, average difference, structural content, image fidelity and normalize coefficients are calculated for K-means methods. The K-means (K) increase in PSNR (db) values than the K-means clustering.

Key Words — Segmentation, K-means clustering, EM models, Quality measures.

I. INTRODUCTION

Segmentation is a fundamental process in digital image processing which has found extensive applications in areas such as medical image processing, compression, diagnosis arthritis from joint image [1] [2], automatic text hand writing analysis [3], and remote sensing. Numerous algorithms using different approaches have been proposed for image segmentation. These approaches include local edge detection [4], deformable curves [5], morphological region-based approaches [6] [7], global optimization approaches on energy functions and stochastic model-based methods [8] [9]. Some intensity-based methods such as thresholding and histogram-based finite mixture models are easy to be formulated and fast. However they often fail to segment objects with low contrast or noisy images with varying background. It is noted that these methods don't use the spatial morphological images information [10]. On the other hand, some other methods such as morphological segmentation, region growing and deformable curves, mainly focus on spatial information such as local structures or regions. Unfortunately, the majority of these techniques are not suitable for satellite image segmentation [11].

In this context, it is important to develop unsupervised segmentation methods capable of analyzing and classifying these images. Indeed, recently, there has been considerable interest in stochastic model-based image segmentation techniques. In such techniques, an image is separated into a set of disjoint regions with each region associated with one of finite numbers of classes that are characterized by distinct parameters. In fact, it exist two different statistic image segmentation approaches, the supervised approach, and the unsupervised approach. In supervised approach, it is usually assumed that training data are available for the image classes; therefore, the parameters can be estimated from the training data before segmentation. But, this is rather unrealistic in many

practical situations. For unsupervised techniques, the objective is to estimate the parameters and segment the image simultaneously [15]. Most of the proposed solutions to the unsupervised segmentation problem can be classified into two broad categories; one is a two-step procedure, estimating the parameters for each class and then using a relaxation scheme to do segmentation [16]. The other is an iterative procedure, which starts with initial parameters and alternatively segments the image based on current parameters and estimates parameters based on current segmentation as if it were correct. In the first category, clustering algorithms are usually adopted to estimate the classes parameters. In the second category, parameters are estimated in each iteration using the current segmentation as if it was correct and the estimated parameters are used in the next segmentation as if they were true parameters. Although these techniques have demonstrated substantial success for satellite image, they have some limitations. Indeed, most of the statistical image segmentation techniques need a manual initial input such as the classes number. Meanwhile, these methods are often sensitive to these initial conditions. Moreover, all these methods cannot segment correctly the entire image, if the initialization step is not optimal. In this paper an optimal initialization methods based on K-means approaches, for a better unsupervised medical image segmentation is studied. A two-step EM model is utilized for segmentation of medical images. The quality measures are calculated for both the methods and their efficacy are reported.

II. GENERAL CLUSTERING ALGORITHM

During last decades, growing attention has been put on data clustering as robust technique in data analysis. Clustering or data grouping describes important technique of unsupervised classification that arranges pattern data (most often vectors in multidimensional space) in the clusters (or groups). Patterns or vectors in the same cluster are similar according to predefined criteria, in contrast to distinct patterns from different clusters [12], [13].

Possible areas of application of clustering algorithms include data mining, statistical data analysis, compression, vector quantization and pattern recognition [12], [13]. Image analysis is the area where grouping data into meaningful regions (image segmentation) presents the first step into more detailed routines and procedures in computer vision and image understanding. Clustering problem understood as grouping input data by means of minimizing certain criteria presents NP-hard combinatorial optimization task.

Fig.1. Enumerates the block diagram of cluster based segmentation of medical images using K-means clustering and EM models. The motto of this research is to segment the images and measure their quality parameters of the above methods. The quality measures are obtained for the center based K-means clustering and EM models. Comparisons of both the methods are analyzed.

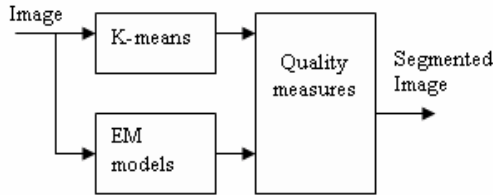


Figure1. Block Diagram for Cluster Based Segmentation of Medical Images

“Center- based clustering” refers to the family of algorithms that use a number of centers to represent and group input data. General iterative model for partition center-based clustering algorithms has the following form [14], [15] and [16]:

1. Data initialization by assigning some values to the cluster centers.
2. For each data point x_i , calculate its membership valued $m(c_j | x_i)$ to all clusters c_j and its weight $w(x_i)$.
3. For each cluster center c_j , recalculate its location taking into account all points x_i assigned to this cluster according to the membership and weight values:

$$C_j = \frac{\sum_{i=1}^n m(c_j | x_i) w(x_i) x_i}{\sum_{i=1}^n m(c_j | x_i) w(x_i)} \quad (1)$$

4. Repeat steps 2 and 3 until some termination criteria are met.

Data clustering presents unsupervised process that finally requires some sort of quality evaluation of generated clusters. This requirement can be satisfied by using cluster validity indices, [17] [18]. In general, three distinctive approaches to cluster validity are possible. The first approach relies on external criteria that investigate the existence of some predefined structure in clustered data set. The second approach makes use of internal criteria and the clustering results are evaluated by quantities describing the data set such as proximity matrix etc. Approaches based on internal and external criteria make use of statistical tests and their disadvantage is high computational cost. The third approach makes use of relative criteria and relies on finding the best clustering scheme that meets certain assumptions and requires predefined input parameters values.

III. K-MEANS CLUSTERING

K-means clustering algorithm proposed by Mac Queen in 1967 belongs to partitioning methods, which is widely

used because of its simpleness and fast convergence. The primary process can be expressed as follows [19] [20].

1. Initialize K cluster centers chosen randomly.
2. Assign each x_i to its nearest cluster center c_k by Euclidean Distance (d).

$$KM(X, C) = \sum_{i=1}^n \min_{j \in \{1, \dots, k\}} \|x_i - c_j\|^2 \quad (2)$$

$$d(p, q) = d(q, p) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (3)$$

3. Update each cluster center c_k as the mean of all x_i that belongs to it.
4. Repeat steps 2-4 until the cluster centers are stable.

K-means clustering is associated with computational expensiveness with respect to the cluster similarity distance measures. Hence a two level EM algorithm is proposed to reduce the effect of excessive computation.

IV. RESULTS AND DISCUSSION

To study the relative performance of cluster based segmentation methods the following quality measures are calculated. I. Avicibas et.al [23], M.Mrak et.al [22] and A.M. Eskicioglu [21] in their papers have described the statistical image quality measures.

A. Peak signal to noise ratio(PSNR):

$$PSNR = 10 \log \frac{255 * 255}{MSE} dB \quad (5)$$

Where MSE is mean square error [24]. Ideally it is infinity. Practically it is in the range of 25 to 40dB.

B. Average Difference (AD):

$$\sum_{j=1}^M \sum_{k=1}^N \left| X(j, k) - \hat{X}(j, k) \right| / MN \quad (6)$$

This measure shows the average difference between the pixel values, ideally it should be zero.

C. Structural content (SC), Image Fidelity (IF) and Normalized Correlation Coefficient (NK):

These are the co relational based quality measure which normally looks at correlation features between the pixels of original and reconstructed image, they are given as

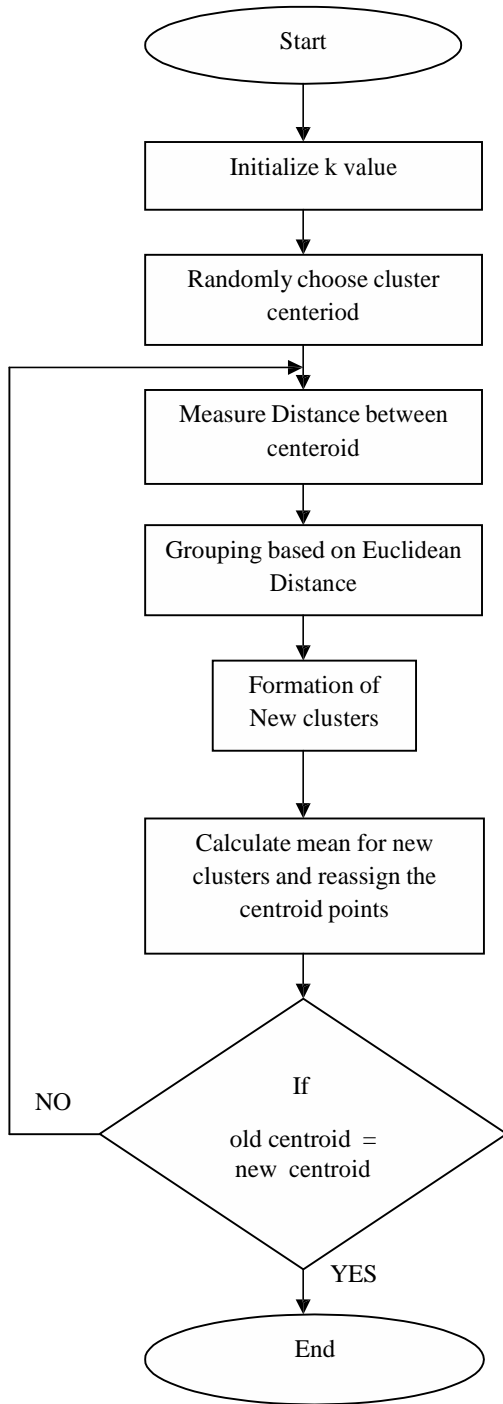


Figure 2: Flow chart for K means Clustering algorithm

$$SC = \frac{\sum_{j=1}^M \sum_{k=1}^N X(j,k)^2}{\sum_{j=1}^M \sum_{k=1}^N \hat{X}(j,k)^2} \quad (7)$$

$$IF = 1 - \frac{\sum_{j=1}^M \sum_{k=1}^N [X(j,k) - \hat{X}(j,k)]^2}{\sum_{j=1}^M \sum_{k=1}^N X(j,k)^2} \quad (8)$$

$$NK = \frac{\sum_{j=1}^M \sum_{k=1}^N [X(j,k) \hat{X}(j,k)]}{\sum_{j=1}^M \sum_{k=1}^N X(j,k)^2} \quad (9)$$

Normally SC, IF and NK are in the range of 0 to 1, very near to or one is the best.

D. Mean structural similarity index (MSSIM) and Universal Quality Index (UQI):

Zhou Wang et.al [24] in their paper proposed a new quality measures, via mean structural similarity index and universal quality index. This compares local patterns of pixel intensities that have been normalized for luminance and contrast. It is given by

$$SSIM = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (10)$$

Where μ and σ are mean and variance respectively, x and y are for original and segmented images. The MSSIM is calculated by taking mean of SSIM and UQI is calculated by substituting the values of C1 and C2 as zero.

The quality measures for MRI, X-Ray and Ultrasonic images are calculated and are shown in Table1, Table2, Table3 and Table 4. It is observed that one dB increment of PSNR values in EM models when compare to K-means clustering techniques. Higher Average difference in EM models indicates intra cluster fragileness. Figure 3 and Figure 4 depicts the segmentation of MRI images using K means clustering and EM models respectively. Through the visual identification, we can observe that the edges of segmented MRI images are well preserved in the EM techniques.

V. CONCLUSION

In this paper, we consider generic segmentation of the medical images which is carried out for different types of medical images and compared using quality measures such as structural content(SC), Image fidelity(IF), Normalized correlation coefficient(NK),mean structural similarity index(MSSIM) and universal Quality index(UQI). The quality measures like SC, NK, IF and MSSIM are enhanced for the segmented images through the K-means clustering algorithms. The future research is in the direction of a comparison of EM models between heuristic MLP and Elman neural network optimization models.

REFERENCES

- [1] S. Li, T. Fevens, A. Krzyzak and S. Li Automatic, "clinical image segmentation using pathological modeling, PCA and SVM", *Engineering Applications of Artificial Intelligence*, Volume 19, Issue 4, June 2006, pp. 403-410.
- [2] A. Schwaighofer, V. Tresp, P. Mayer, A.K. Scheel and G. Muller, The RA scanner, "prediction of rheumatoid joint inflammation based on laser imaging", *IEEE Trans. Biomed. Imaging* (2003), pp. 375-382.
- [3] Y. Zheng, H. Li and D. Doermann, "Machine printed text and handwriting identification in noisy document images", *IEEE Trans. Pattern Anal. Mach. Intell.* 26, 2004, pp. 337-345.
- [4] W.Y. Manjunath, B.S., Edge flow; A framework of boundary detection and image segmentation. In: *IEEE Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, 1997 pp. 744-749.
- [5] C. Y. and Jerry, L.P., Snakes, shapes and gradient vector flow, *IEEE trans. Image Process.* 73, 1998, pp. 359-369.
- [6] S.C. and Alan, Y., Region competition: Unifying snakes, region growing and Bayes/MDL for multiband image segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* 189, 1996, pp. 884-900.
- [7] G.P and Chuang, Extensive partition operators, gray-level connected operators, and region merging/classification segmentation algorithms; Theoretical links. *IEEE Trans. Image Process.* 109, 2001, pp. 1332-1345.
- [8] K. S. and jayaram, K.U., 2001, "Optimum image thresholding via class uncertainty and region homogeneity", *IEEE Trans. Pattern Recog. Mach. Intell.* 237, pp. 689-706.
- [9] J.P., "Stochastic relaxation on partions with connect components and its application to image segmentation", *IEEE Trans. Pattern Anal. Mach. Intell.* 206, 1998, pp. 619-636.
- [10] J. Xie, and H. T. Tsui, "Image segmentation based on maximum-likelihood estimation and optimum entropy-distribution (MLE-OED)", *Pattern Recognition Letters*, Volume 25, Issue 10, 16 July 2004, pp. 1133-1141.
- [11] Y.N. and Manjunath, B.S., 2001, "unsupervised segmentation of color-texture regions in images and video", *IEEE Trans. Pattern Anal. Mach. Intell.* 238, pp. 800-810.
- [12] R. Xu, D. Wunsch II, "Survey of clustering algorithms", *IEEE Transactions on Neural Networks*, 16, 2005, 645-678.
- [13] A. Jain, M. Murty and P. Flynn, "Data Clustering: A review", *ACM Computing Surveys*, 31, 1999, 264-323.
- [14] O. Hall, I. Barak, J.C. Bezdek, "Clustering with a genetically optimized approach", *IEEE Trans. Evo. Computation*, 3, 1999, 103-112.
- [15] G. Hamerly, C. Elkan, "Alternatives to the k-means algorithm that find better clusterings", *Proc. Of the ACM Conference on Information and Knowledge Management, CIKM-2002*, 2002, 600-607.
- [16] G. H. Omran, A. Salman, A.P. Engelbrecht, "Dynamic clustering using particle swarm optimization with application in image segmentation", *Pattern Anal Applic.*, 8, 2006, 332-344.
- [17] M. Halkidi, M. Vazirgiannis, and I. Batistakis, "Quality scheme assessment in the clustering process". In *Proc. Of the 4th European Conf. on Principles of Data Mining and Knowledge Discovery*, LNCS 1910, 2000, 265-267.
- [18] M. Halkidi et al., "Clustering validity checking methods: PartII", *SIGMOD Rec.*, 31, No. 3, 2002, 19-27.
- [19] Stephen J Redmond and Conor Heneghan, "A method for initializing the k-means clustering algorithm using kd-trees", *Pattern Recognition* 28(2007), 965-973.
- [20] Ujjwal Maulik, Sanghamitra Bandyopadhyay, "Genetic algorithm based clustering technique", *Pattern Recognition* 33(2000), 1455-1465.
- [21] Ahmet M. Eskicioglu and Paul S. Fisher, "Image Quality Measures and their performance" *IEEE Transactions on Communications*, Vol.43, No.12, December 1995.
- [22] Marta Mrak, Sonja Grgic and Mislav Grgic "Picture Quality Measures in Image Compression Systems" *EUROCON 2003*, Ljuijana, Slovenia.
- [23] Ismail Avicibas, Bulent Sankur, Khalid Sayood "Statistical Evaluation of Image Quality Measures" *Journal of Electronic Imaging* 11(2), 206-223 (April 2002).
- [24] Zhou Wang, Alan Conrad Bovik., et.al. "Image Quality Assessment: From Error Visibility to Structural Similarity" *IEEE Transactions on Image processing*, vol. 13, no.4, April 2004.



Figure 3.a MRI Knee original image



Figure 3.b Segmentation of MRI for K means cluster K=4



Figure 3.c Segmentation of MRI for K means cluster K=6



Figure 3.d Segmentation of MRI for K means cluster K=8



B. Vinoth Kumar received his B.E (EEE) degree from Bannari Amman Institute of Technology, Sathyamangalam, Bharathiar University, Coimbatore in 2002. He obtained his M.E (Applied Electronics) degree from Government College of Technology, Coimbatore in 2005. He has six year five months of teaching experience

at college level. Currently he is working as Assistant Professor (Senior Grade) in Department of EEE at Bannari Amman Institute of Technology, Sathyamangalam. His area of interest is Medical Image Processing, Soft computing and Embedded systems.



G. Karthick received his B.E (ECE) degree from KSR College of Engineering Tiruchengode, Anna University Chennai in 2009. He is pursuing his PG programme in Applied Electronics at Department of ECE, Bannari Amman Institute of Technology Sathyamangalam. His area of interest is Medical Image Processing, Soft computing, Embedded systems and VLSI.

AUTHOR'S PROFILE



Dr. R. Harikumar received his B.E (ECE) degree from REC Trichy 1988. He obtained his M.E (Applied Electronics) degree from College of Engineering, Guindy, Anna university Chennai in 1990. Engineering He was awarded Ph.D. in I&C Engg from Anna university Chennai in 2009. He has 21 years of teaching experience at college level. He worked as faculty at

Department of ECE, PSNA College of Engineering & Technology, Dindigul. He was Assistant Professor in IT at PSG College of Technology, Coimbatore. He also worked as Assistant Professor in ECE at Amrita Institute of Technology, Coimbatore. Currently he is working as Professor ECE at Bannari Amman Institute of Technology, Sathyamangalam. He is guiding eight PhD theses in these areas. He has published Thirty papers in International and National Journals and also published around seventy one papers in International and National Conferences conducted both in India and abroad. His area of interest is Bio signal Processing, Soft computing, VLSI Design and Communication Engineering. He is life member of IETE, IEEE and ISTE.

Table 1. Quality Measures of MRI Image

A.KNEE

K Means							
CLUSTER SIZE	PSNR	AD	SC	NK	IF	SSIM	UQI
K=4	29.8519	0.3303	1.0286	0.9763	0.9664	0.9802	0.9801
K=5	30.3278	0.5222	1.0219	0.9829	0.9727	0.9872	0.9871
K=6	31.134	0.714	1.0201	0.9858	0.9744	0.9914	0.9914
K=7	32.0844	0.383	1.0167	0.9887	0.9777	0.9939	0.9939
K=8	32.89	0.3954	1.0129	0.9914	0.9814	0.9956	0.9955

B.SIDE HEAD

K Means							
CLUSTER SIZE	PSNR	AD	SC	NK	IF	SSIM	UQI
K=4	30.3616	0.2806	1.0195	0.9803	0.9744	0.9782	0.9781
K=5	31.2356	0.5394	1.019	0.985	0.975	0.9878	0.9877
K=6	32.1353	0.3628	1.0183	0.9872	0.9756	0.9917	0.9917
K=7	32.9971	0.5858	1.016	0.9894	0.9778	0.9941	0.9941
K=8	33.5428	0.6378	1.0173	0.9894	0.9766	0.9954	0.9954

Table 2. Quality Measures of X-ray image

A.HAND

K Means							
CLUSTER SIZE	PSNR	AD	SC	NK	IF	SSIM	UQI
K=4	33.2785	0.1651	1.0209	0.9814	0.9741	0.9892	0.9891
K=5	34.3837	0.5142	1.0157	0.9874	0.9791	0.9937	0.9936

K=6	35.1403	0.4962	1.0152	0.9893	0.9796	0.9959	0.9959
K=7	35.917	0.3943	1.0125	0.9915	0.9822	0.997	0.997
K=8	36.7625	0.4169	1.0143	0.9912	0.9804	0.9978	0.9978

B.CHEST

K Means							
CLUSTER SIZE	PSNR	AD	SC	NK	IF	SSIM	UQI
K=4	29.296	0.6269	1.0285	0.9784	0.9655	0.9713	0.9712
K=5	30.1746	0.7301	1.0185	0.9864	0.9751	0.9837	0.9836
K=6	30.9338	0.6603	1.0181	0.988	0.9755	0.9885	0.9884
K=7	31.9018	0.7378	1.0136	0.9909	0.9798	0.9916	0.9916
K=8	32.8045	0.6315	1.0134	0.9915	0.9801	0.9933	0.9933

Table 3. Quality Measures of CT image

A.CT BRAIN

K Means							
CLUSTER SIZE	PSNR	AD	SC	NK	IF	SSIM	UQI
K=4	32.2861	0.2498	1.0078	0.9908	0.9873	0.9877	0.9877
K=5	33.0137	0.115	1.0145	0.9898	0.9808	0.9924	0.9924
K=6	32.9529	0.5025	1.0137	0.9909	0.9816	0.9939	0.9938
K=7	33.6582	0.4793	1.0145	0.9911	0.9808	0.9952	0.9952
K=8	35.783	0.3807	1.0125	0.9932	0.9828	0.9973	0.9973

B.LUNG

K Means							
CLUSTER SIZE	PSNR	AD	SC	NK	IF	SSIM	UQI
K=4	33.2458	0.287	1.0197	0.986	0.9735	0.9913	0.9913
K=5	34.7117	0.1892	1.0174	0.9888	0.9757	0.9947	0.9946
K=6	35.4737	0.6858	1.0178	0.9894	0.9753	0.9964	0.9964
K=7	37.1373	0.5631	1.0162	0.9909	0.9768	0.9977	0.9977
K=8	38.1009	0.6456	1.0172	0.9907	0.9759	0.9983	0.9983

Table 4. Quality Measures of Ultrasonic Image

A.BABY

K Means							
CLUSTER SIZE	PSNR	AD	SC	NK	IF	SSIM	UQI
K=4	30.951	0.9019	0.9672	0.989	0.9936	0.8928	0.8882
K=5	31.8368	0.6636	0.966	0.9921	0.9949	0.8923	0.8877
K=6	32.8323	0.6236	0.9647	0.994	0.9962	0.8917	0.887
K=7	33.6023	0.8675	0.9635	0.9937	0.9975	0.891	0.8862
K=8	34.5931	0.6355	0.9622	0.9961	0.9989	0.8901	0.8852

B.SPINE

K Means							
CLUSTER SIZE	PSNR	AD	SC	NK	IF	SSIM	UQI
K=4	30.3786	0.6163	1.0597	0.9498	0.9338	0.9581	0.9576
K=5	31.3336	0.4186	1.0397	0.9682	0.952	0.9762	0.9759
K=6	32.5776	0.3935	1.0335	0.9759	0.9577	0.9853	0.9851
K=7	33.7688	0.2837	1.0317	0.9792	0.9594	0.9898	0.9897
K=8	34.8629	0.5774	1.0282	0.9825	0.9628	0.993	0.9929