ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)
Vol. 2 Issue Special 1 Jan-March 2014
International Journal of Advanced Research in Computer Science & Technology (IJARCST 2014)

Fuzzy C-Means Clustering With Local Information and Kernel Metric For Image Segmentation

[I]P.Sivasangareswari, [II]K.Sathish Kumar
[I]PG Student, [II]Assistant Professor
[I,II]Dept. of CSE, Sethu Institute of Technology, Affiliated to Anna University, Kariapatti, Tamilnadu, India

Abstract

Image segmentation is defined as action of splitting an image into band of pixels. Aim of this process is to group pixels into regions. This is used in locating objects in satellite images, face recognition, iris recognition, agricultural imaging, medical imaging. Clustering is process that make group of similar objects and is widely used technique for brain tumour detection. We proposed a new method Modified FCM algorithm based on the Distance metric for segmentation of images that have been corrupted by intensity inhomogeneities and noise. We define a new trade-off weighted fuzzy factor to adaptively control the local neighbor relationship.

Keywords

Image Segmentation, Fuzzy c-means, Modified FCM, Clustering

I. Introduction

Image segmentation is one of the key techniques in image understanding and computer vision. The task of image segmentation is to divide an image into a number of non overlapping regions, which have same characteristics such as gray level, color, tone, texture, etc. A lot of clustering based methods have been proposed for image segmentation. we define a new trade-off weighted fuzzy factor to adaptively control the local neighbor relationship. This factor depends on space distance of all neighbor pixels and their gray level discrepancy simultaneously, Image processing operations can be roughly divided into three major categories, Image Compression, Image Enhancement and Restoration, and Measurement Extraction. It involves reducing the amount of memory needed to store a digital image. Image defects which could be caused by the digitization process or by faults in the imaging set-up (for example, bad lighting) can be corrected using Image Enhancement techniques. Once the image is in good condition, the Measurement Extraction operations can be used to obtain useful information from the image. Some examples of Image Enhancement and Measurement Extraction are given below. The examples shown all operate on 256 grey-scale images. This means that each pixel in the image is stored as a number between 0 to 255, where 0 represents a black pixel, 255 represents a white pixel and values in-between represent shades of grey. These operations can be extended to operate on colour images. The examples below represent only a few of the many techniques available for operating on images. Details about the inner workings of the operations have not been given, but some references to books containing this information are given at the end for the interested reader.

A. Images and digital images

Suppose we take an image, a photo, say. For the moment, lets make things easy and suppose the photo is black and white (that is, lots of shades of grey), so no colour. We may consider this image as being a two dimensional function, where the function values give the brightness of the image at any given point. We may assume that in such an image brightness values can be any real numbers in the range (black) (white). A digital image from a photo in that the values are all discrete. Usually they take on only integer values. The brightness values also ranging from 0 (black) to 255 (white). A digital image can be considered as a large array of discrete dots, each of which has a brightness associated with it. These dots are called picture elements, or more simply pixels. The pixels surrounding a given pixel constitute its neighborhood. A neighborhood can be characterized by its shape in the same way as a matrix: we can speak of a neighborhood,. Except in very special circumstances, neighborhoods have odd numbers of rows and columns; this ensures that the current pixel is in the centre of the neighborhood.

B. Applications

Image processing has an enormous range of applications; almost every area of science and technology can make use of image processing methods. Here is a short list just to give some indication of the range of image processing applications.1. Medicine Inspection and interpretation of images obtained from X-rays, MRI or CAT scans, analysis of cell images, of chromosome karyotypes.2. Agriculture: Satellite/aerial views of land, for example to determine how much land is being used for different purposes, or to investigate the suitability of different regions for different crops, inspection of fruit and vegetables distinguishing good and fresh produce from old.3. Industry: Automatic inspection of items on a production line, inspection of paper samples.4. Law enforcement: Fingerprint analysis, sharpening or de-blurring of speed-camera images.

II. Related Work

Fuzzy c-means algorithm is one of most widely used fuzzy clustering algorithms in image segmentation.FCM algorithm was first introduced by S. Krinidis et.al , suggested that [6] A robust fuzzy local information C-means clustering algorithm," In this paper, we present a c-means algorithm for fuzzy segmentation of magnetic resonance imaging (MRI) data and estimation of intensity inhomogeneities using fuzzy logic. The homomorphic filtering approach to remove the multiplicative effect of the inhomogeneity has been commonly used due to its easy and efficient implementation. To solve the problem of noise sensitivity and computational complexity of Pham and Prince method, we present in this paper a different approach for fuzzy segmentation of MRI data in the presence of intensity inhomogeneities. To overcome the mentioned problems, X. Yin, S. Chen, et.al, Semi-supervised clustering with metric learning: [5] An adaptive kernel method, In this paper presents a new framework for multiple object segmentation in medical images that respects the topological

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

properties and relationships of structures as given by a template. The technique, known as topology-preserving, anatomy-driven segmentation (TOADS), combines advantages of statistical tissue classification, topology-preserving fast marching methods, and image registration to enforce object-level relationships with little constraint over the geometry. When applied to the problem of brain segmentation, it directly provides a cortical surface with spherical topology while segmenting the main cerebral structures. Validation on simulated and real images characteristics the performance of the algorithm with regard to noise, inhomogeneities, and anatomical variations. More recently ,we proposed F. Chung et.al(2009), [3] Generalized fuzzy C-means clustering algorithms with improved fuzzy partitions, To propose a novel contrast image enhancement using intuitionistic fuzzy set theory is suggested and this method is applied on medical images. The method uses window based enhancement and for each window, the image is enhanced accordingly. It considers more uncertainty as compared to the fuzzy set. Therefore, R. Huang, Z. Ding et.al(2011),[8] also proposed A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI it describes an evolutionary approach for unsupervised gray-scale image segmentation that segments an image into its constituent parts automatically. The aim of this algorithm is to produce precise segmentation of images using intensity information along with neighborhood relationships. fuzzy c-means clustering helps in generating the population of Genetic algorithm which there by automatically segments the image. This technique is a powerful method for image segmentation and works for both single and multiple data segmentation.

III. Materials and Methods

The framework of the automated image segmentation system is shown in Fig.1.(a)FCM-based MR brain image segmentation and (b)Modified FCM-based MR brain image segmentation. In case of medical image segmentation, pre-processing techniques must be performed to guarantee superior performance measures. Hence the technique of feature extraction.

In this work ,real time abnormal MR images from four different tumor categories are used for testing the automated systems. These images are collected from M/S Devaki Scan Centre, Madurai, India. These images are gray scale image of size 256x256.

A set of significant features are extracted from these images and used as input for the two fuzzy system. In earlier works, the intensity of the images alone is used as feature for FCM-based segmentation. But in this paper, a set of textures are used to enhance the segmentation efficiency which is not possible with the intensity feature.

The features extraction is to be reduce the original dataset by measuring certain properties. the features used in this work are mean, variance, energy and entropy.

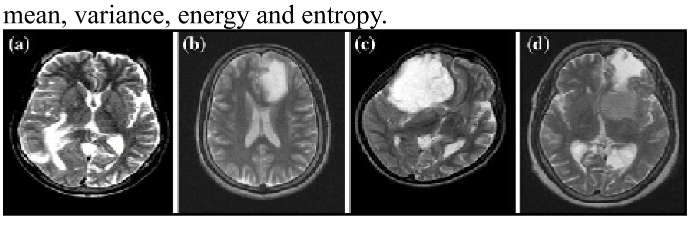

Fig. 1. Sample dataset: (a). metastase, (b). glioma (c). astrocytoma, (d).meningioma

Initially, the entire feature set is given as input to the FCM algorithm and the performance measures are calculated. Then, the modified FCM technique is tested with the reduced input dataset .In this modified system, an approach of dimensionality reduction in the input vector size is adopted to minimize the convergence time period of conventional FCM. This dimensionality reduction is achieved through the concept of distance metrics estimation. The rest of this paper is organized as follows.Sect.4.deals with the Fuzzy Based image segmentation;the conventional FCM algorithm;explains the modified FCM algorithm and sect.5 deals with the Experimental results and Conclusion.

IV. Fuzzy Based Image Segmentation

In this two system are used for image segmentation one is the FCM algorithm and other is the modified FCM. Initially the FCM is dealt briey and further the modified FCM is explained in detail with algorithmic procedures.

A. FCM

In existing system they use a FCM algorithm has been used for segmentation of medical images for kernel metric. Kernel metric for segmentation of images use a FCM algorithm it can be corrupted by inhomogenities and noise. The major drawback for Convergence rate is highly iterative in nature for both processes, which consumes high computational time period, High Computational complexity. The FCM algorithm assigns pixels to each category by using fuzzy memberships. Let X=(x1, x2… xn) denotes an image with N pixels to be partitioned into c clusters, where xi represents multispectral (features) data ([2]-[3]). The algorithm is an iterative optimization that minimizes the cost function defined as follows:

$$J=\sum_{j=1}^{N}\sum_{j=1}^{N} u_{ij}^{m} \parallel x_j - v_i \parallel^{2} \tag{1}$$

Where uij represents the membership of pixel xj in the ith cluster, vi is the ith cluster center, ‖.‖ is a norm metric, and m is a constant. The parameter m controls the fuzziness of the resulting partition and m values is assign to two.One of the important characteristics of an image is that neighboring pixels have similar feature values, and the probability of getting same cluster is comparatively high. The spatial information is important in clustering, but the standard FCM algorithm does not fully utilized it, to exploit the spatial information, a modified membership function is defined as,

$$u_{ij} = \frac{u_{ij}^{m}\, S_{ij}^{n}}{\sum_{k=1}^{c} u_{kj}^{m}\, S_{kj}^{n}} \tag{2}$$

Where

$$S_{ij} = \sum_{k\in N(x_j)} u_{ik}$$

is called spatial function, and $N(x_j)$ represents a square window centered on pixel xj in the spatial domain. The spatial function Sij represents the probability that pixel xj belongs to i th cluster. The spatial function of a pixel for a cluster is large if the majority of its neighborhood belongs to the same square window of size 3×3 is used. In a homogenous region, the spatial functions enhance the original membership, and the clustering result remains unchanged. However, for a noisy pixel, it will reduce the weighting of a noisy cluster by the labels of its neighboring pixel. As a result, misclassified pixels from noisy regions can be easily corrected. There are two steps in FCM algorithm ([2]-[3]). The first step is to calculate the membership function in the spectral domain and the second step is to map the membership information of each pixel to the spatial domain and then compute the spatial function from that. The iteration proceeds with the new membership function that is incorporated with the spatial function. The iteration is stopped when the maximum difference between two cluster centroids is less than a threshold value (=0.01). After the convergence, defuzzification is applied to assign each pixel to a specific cluster for which the membership is maximal.

B. Proposed System

The details of the proposed algorithm for using modified FCM with Distance metric. Distance metric means the concept of dimensionality reduction, it is used to reduce the random number of variables using brain images using modified FCM mainly based on segmentation to detect a tumor. Distance metric for segmentation of images that have been corrupted by intensity in homogeneities and noise. Clustering is process that make group of similar objects and is widely used technique for brain tumor detection. This Proposed system provides Better accuracy, Convergence time period Efficiency, less Computational complexity. Modified FCM yields accurate results within less time period.

Algorithm

Step 1: Set the number of clusters c and the parameter m.

$$J_m(u,v) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} d^2(x_j \quad , v_i) \tag{3}$$

Initialize the fuzzy Cluster centroid vector $V = [v_{1,} \dots \dots \dots, v_c]$ randomly and set ε = 0.01.

Step 2: compute u_{ij}

$$u_{ij} = \left[\sum_{k=1}^{c} \left[\frac{d(x_{j,} v_i)}{d(x_j, v_k)} \right]^{2/(m-1)} \right]^{-1} \tag{4}$$

Step 3: compute v_i .

$$v_i = \frac{\sum_{j=1}^{n} u_{ij}^{m} x_j}{\sum_{j=1}^{n} \quad u_{ij}^{m}} \tag{5}$$

Step 4: update u_{ij}.

Step 5: update v_i

Repeat Steps 4 and 5 until the following termination criterion is satisfied:

$| v_{new} - v_{old} | < \epsilon$

Two types of cluster validity functions, fuzzy partition and feature structure, are often used to evaluate the performance of clustering in different clustering methods.

The representative functions for the fuzzy partition are partition coefficient Vpc and partition entropy Vpe. The reduction in Vpe with the MFCM, as compared with the conventional technique, is 50% for original images and 64.6% for noisy data.

V. Experimental Results

The proposed approach is tested with 240 real time MR images from four abnormal brain tumor categories.The experiments aer performed with the FCM algorithm and the modified FCM algorithm.The performance measures used in this work are convergence time period and segmentation efficiency.Both the qualitative and quantitative analysis is reported in this work. The experiments are carried out on an IBM PC Pentium with processor speed 700 MHZ and 256MB RAM.The software used for the implementation is MATLAB(version7.0).

The segmentation of images to detect a tumour using modified FCM.In First step, preprocessing method median filter is used to remove noise from the input test images. It is often desirable to be able to perform some kind of noise reduction on an image or signal. The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing. Median filtering is very widely used in digital image processing.

In the median filtering operation, the pixel values in the neighborhood window are ranked according to intensity, and the middle value (the median) becomes the output value for the pixel under evaluation. Median filtering does not shift boundaries, as can happen with conventional smoothing filters. Since the median is less sensitive than the mean to extreme values (outliers), those extreme values are more effectively removed. Median filtering preserves the edges.

Clustering is used in segmentation of images that can be used to unionize set of pixels into groups based on similarities among the individual data items in such a way that data points of the same group are more identical to one another than samples belonging to different groups. It is interesting to use fuzzy clustering methods, which holds large information from the image as compared to hard clustering methods. FCM provides flexibility which admits pixels to belong to multiple classes with changing degrees of membership.

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which over fits the training sample and generalizes poorly to new samples. Feature extraction can be used in the area of image processing which involves using algorithms in terms of GLCM to detect and isolate various desired portions or shapes (features) of a digitized image.

A gray level co-occurrence matrix (GLCM) contains information about the positions of pixels having similar gray level values. A co-occurrence matrix is a two-dimensional array, P, in which both the rows and the columns represent a set of possible image values. A GLCM Pd [i,j] is defined by first specifying a displacement vector d=(dx,dy) and counting all pairs of pixels separated by d having gray levels i and j.

The GLCM is defined by:

– Where nij is the number of occurrences of the pixel Values (i,j) lying at distance d in the image.

– The co-occurrence matrix Pd has dimension n× n, Where n is the number of gray levels in the image.

SVM maps input vectors to a higher dimensional vector space where an optimal hyper plane is constructed. Among the many hyper planes available, there is only one hyper plane that maximizes the distance between itself and the nearest data vectors of each category. This hyper plane which maximizes the margin is called the optimal separating hyper plane and the margin is defined as the sum of distances of the hyper plane to the closest training vectors of each category. Classification is used to whether the image is normal or abnormal.

Modified FCM technique includes similar steps as FCM except for the variation in the cluster updation and membership value updation criterions. Computational complexity of the problem is tackled by reducing the size of the input dataset.The Euclidean distance between pairs of objects in m-by-n data matrix X. Rows of X correspond to observations, and columns correspond to variables. D is a row vector of length m(m–1)/2, corresponding to pairs of observations in X. The distances are arranged in the order (2,1), (3,1), ..., (m,1), (3,2), ..., (m,2), ..., (m,m–1)). D is commonly used as a dissimilarity matrix in clustering or multidimensional scaling.

VI. Performance Analysis: Accuracy and Mis-classification

The Accuracy (or Power) is the probability that the test correctly classifies the subjects; the Mis-classification rate is its complement to 1. In statistics, the F1 score (also F-score or F-measure) is a measure of a test's accuracy. Table 1 shows the performance evaluation result.

Table 1 Comparision of Original MRI Images and Noise Added MRI Images

Image	Techniqe	Vpc	Vpe
Original MRI Images	Conventional FCM	0.888	0.234
	Modified FCM	0.924	0.117
Noise added MRI Images	Conventional FCM	0.779	0.414
	Modified FCM	0.909	0.147

It considers both the Precision (positive predictivity) and the Sensitivity of the test to compute the score: P is the number of correct results divided by the number of all returned results. S is the number of correct results divided by the number of results that should have been returned. The F1 score can be interpreted as a weighted average of the Precision and Sensitivity, where an F1 score reaches its best value at 1 and worst score at 0.

VII. Conclusion and Future Work

In this paper, MRI brain image segmentation is performed using Modified fuzzy C-Mean algorithm. And in order to get a better classification rate, different statistical feature were extracted by using GLCM. SVM is used to classify the input, which is MRI Brain image into normal and abnormal classification. It provides better accuracy and time period, proposed a modified FCM approach for computational complexity. Impact of image segmentation which leads to the help for future treatment. This kernel technique will help to get more accurate result. In the proposed work about 90% are classified accurately and 10% are misclassified. In future, 3D image Analysis can be done for classification of MRI brain images. .

VIII. Acknowledgment

The authors wish to thank Dr.S.Raja of M/s.Devaki Scan Center, Madurai, Tamilnadu, India, for his help regarding MRI data and validation

References

[1] S. Krinidis and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm," IEEE Trans. Image Process., vol. 19, no. 5, pp. 1328–1337, May 2010.

[2] X. Yin, S. Chen, E. Hu, and D. Zhang, "Semi-supervised clustering with metric learning: An adaptive kernel method," Pattern Recognit., vol. 43, no. 4, pp. 1320–1333, Apr. 2010.

[3] L. Zhu, F. Chung, and S. Wang, "Generalized fuzzy C-means clustering algorithms with improved fuzzy partitions," IEEE Trans. Syst., Man, Cybern., B, Cybern., vol. 39, no. 3, pp. 578–591, Jun. 2009.

[4] S. Tan and N. A. M. Isa, "Color image segmentation using histogram thresholding fuzzy C-means hybrid approach," Pattern Recognit., vol. 44, no. 1, pp. 1–15, 2011.

[5] C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, and J. C. Gore, "A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI," IEEE Trans. Image Process., vol. 20, no. 7, pp. 2007–2016, Jul. 2011.

[6] J. Dunn, "A fuzzy relative of the ISODATA process and its use indetecting compact well-Separated clusters," J. Cybern., vol. 3, no. 3,pp. 32–57, 1974.

[7] J. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms. New York: Plenum, 1981.

[8] M. Ahmed, S. Yamany, N. Mohamed, A. Farag, and T. Moriarty,"A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data," IEEE Trans. Med. Imag., vol. 21, no. 3,pp. 193–199, Mar. 2002.

[9] S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," IEEE Trans. Syst., Man, Cybern., B, Cybern., vol. 34, no. 4, pp. 1907–1916, Aug. 2004.

[10] L. Szilagyi, Z. Benyo, S. Szilagyii, and H. Adam, "MR brain image segmentation using an enhanced fuzzy C-means algorithm," in Proc. 25th Annu. Int. Conf. IEEE EMBS, Nov. 2003, pp. 17–21.

[11] W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy C-means clustering algorithms incorporating local information for image segmentation," Pattern Recognit., vol. 40, no. 3, pp. 825–838, Mar.2007.

[12] M. Gong, Z. Zhou, and J. Ma, "Change detection in syntheticaperture radar images based on image fusion and fuzzy clustering,"IEEE Trans. Image Process., vol. 21, no. 4, pp. 2141–2151, Apr.2012.

[13] N. Cristianini and J. S. Taylor, An Introduction to SVM's and Other Kernel-Based Learning Methods. Cambridge, U.K.: Cambridge Univ.Press, 2000.

[14] X. Yang and G. Zhang, "A kernel fuzzy C-means clustering-based fuzzy support vector machine algorithm for classification problems with outliers or noises," IEEE Trans. Fuzzy Syst., vol. 19, no. 1, pp. 105–115,Feb. 2011.

[15] V. Roth and V. Steinhage, "Nonlinear discriminant analysis using kernel functions," in Advances in Neural Information Processing Systems 12, S. A Solla, T. K. Leen, and K.-R. Muller, Eds. Cambridge, MA: MIT Press, 2000, pp. 568–574.

[16]Maoguo Gong, Member, IEEE, 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.

Siva Sangareswari.P received her B.Tech degree in Information Technology from Sethu Institute of Technology, Virudhunagar, in 2012. She is currently pursuing her M.E degree in Computer Science and Engineering from Sethu Institute of Technology, Virudhunagar. Her research interests include Image Processing and Data Mining.

Sathish kumar received his B.E degree in Computer Science and Engineering from Madurai Kamaraj University, Tamilnadu, India, in 2002. He has received his M.E degree in Computer Science and Engineering from Anna University Trichy, Tamilnadu, India, in 2010. He is currently pursuing his Ph.D degree in Anna University, Chennai. His research interests include cryptography and Network Security.