Global Journals LaTeX JournalKaleidoscope™
Artificial Intelligence formulated this projection for compatibility purposes from the original article published at Global Journals. However, this technology is currently in beta. *Therefore, kindly ignore odd layouts, missed formulae, text, tables, or figures.*

Image Segmentation using Rough Set Based Fuzzy K-Means Algorithm

ev reddy[1]

[1] Nagarjuna university

Received: 8 December 2012 Accepted: 5 January 2013 Published: 15 January 2013

Abstract

Image segmentation is critical for many computer vision and information retrieval systems, and has received significant attention from industry and academia over last three decades. Despite notable advances in the area, there is no standard technique for selecting a segmentation algorithm to use in a particular application, nor even is there an agreed upon means of comparing the performance of one method with another. This paper, explores Rough-Fuzzy K-means (RFKM) algorithm, a new intelligent technique used to discover data dependencies, data reduction, approximate set classification, and rule induction from image databases. Rough sets offer an effective approach of managing uncertainties and also used for image segmentation, feature identification, dimensionality reduction, and pattern classification. The proposed algorithm is based on a modified K-means clustering using rough set theory (RFKM) for image segmentation, which is further divided into two parts. Primarily the cluster centers are determined and then in the next phase they are reduced using Rough set theory (RST). K-means clustering algorithm is then applied on the reduced and optimized set of cluster centers with the purpose of segmentation of the images. The existing clustering algorithms require initialization of cluster centers whereas the proposed scheme does not require any such prior information to partition the exact regions. Experimental results show that the proposed method perform well and improve the segmentation results in the vague areas of the image.

Index terms— uncertain images, RGB images, rough set, K-means algorithm.

1 Introduction

mage segmentation is one of the most challenging tasks in image analysis. It is also useful in the field of pattern recognition. Image mining deals with the extraction of implicit knowledge, image data relationship, or other patterns not explicitly stored in the images [5][6] [7]. Image Segmentation is becoming more important for medical diagnosis process. Currently, development an efficient computer aided diagnosis system that assist the radiologist has thus become very interest, the aim being not to replace the radiologist but to over a second opinion [3,4]. Consequently, the need of efficient research on features extracted and their role to the classification results

Author ? : Research Scholar, Acharya Nagarjuna University Guntur. E-mail : evr.eluri@gmail.com Author ? : Principal, University College of Engineering, ANU, Guntur. E-mail : edara_67@yahoo.com makes researchers to select features randomly as input to their systems. In image segmentation an image is divided into different regions with similar features. There are many different types of approaches of image segmentation. Edge-based method, region-based techniques and threshold-based techniques and so on. Images are partitioned according to their global feature distribution by clustering based image segmentation methods. In this paper, a image segmentation method based on K-means using rough set theory is proposed, in which pixels are clustered according to the

intensity and spatial features and then clusters are combined to get the results of final segmentation. The paper is organized as follows. In section 2 rough set theory is described. In section 3 rough set based K-means algorithm is proposed. In section 4 we have shown the experimental results and in section 5 some conclusions have been made.

2 II.

3 Rough Set Concepts

Rough Set Theory was firstly introduced by Pawlak in 1982 [2][3] and is a valuable mathematical tool for dealing with vagueness and uncertainty [4]. Similar or indiscernibility relation is the mathematical basis of the Rough Set theory. The key concept of rough set theory is the approximate equality of sets in a given approximation space [2] [3]. An approximation space A is an ordered pair (,)

U R where U is a certain set called universe, and that equivalence relationR U U ? × is a binary relation called indiscernibility relation; if , x y U ? any (,)

x y R ? , this means that x and y are indistinguishable in A; equivalence classes of the( ) { : [ ] } A X x U x X R = ? ? , ( ) { : [ ] } A X x U x X R ? = ? ? ? , (D D D D D D D D) Year 013 2 F

Where [ ] x R denotes the equivalence class of the relation R containing element x. In addition, the set , relation R are called elementary sets (atoms) in A (an empty set is also elementary), and the set of all atoms in A is denoted by / UR . In the Rough Set approach, any vague concept is characterized by a pair of precise concepts, that is the lower and upper approximation of the vague concept. Let XU ? be a subset ofU , then the lower and upper approximation of X in A are respectively denoted as:( ) ( ) ( ) BN X A X A X A = ? is called a boundary of X in A [2][3]

. If set X is roughly definable in A it means that we can describe the set X with some "approximation" by defining its lower and upper approximation in A [3]. The upper approximation ( ) A X means the least definable set in A containing the objects that possibly belong to the concept, whereas the lower approximation ( ) A X a) Reduction of Attributes Discovering the dependencies between attributes is important for information table analysis in the rough set approach. In order to check whether the set of attributes is independent or not, it is a way to check every attribute whether its removal increases the number of elementary sets in an information system [6]. Let the (, , ,)S U Q V ?

= be an information system and let , P R Q ? Then, the set of attributes P is said to be dependent on set of attributes R in S (denotationR P ?) iff IND IND R P ?

, whereas the set of attributes , P R are called independent in S iff neither R P ? nor P R ? hold [2]. Moreover, finding the reduction of attributes is another important thing. Let the minimal reducts is called the Y ? ??? core of . P especially, the core is a collection of the most relevant attributes in the table [5] and is the common part of all reducts [6] b) Fuzzy-Rough Sets In many real-world applications, data is often both crisp and real-valued, and this is where traditional rough set theory encounters a problem. It is not possible in the original theory to say whether two attribute values are similar and to what extent they are the same; for example, two close values may only differ as a result of noise, but RST considers them as different as two values of a dissimilar magnitude. It is, therefore desirable to develop techniques which provide a method for knowledge modelling of crisp and real-value attribute datasets which utilise the extent to which values are similar. This can be achieved through the use of fuzzy-rough sets. Fuzzy-rough sets encapsulate the related but distinct concepts of vagueness (for fuzzy sets) and indiscernibility (for rough sets), both of which occur as a result of uncertainty in knowledge. A Ttransitive fuzzy similarity relation is used to approximate a fuzzy concept X the lower and upper approximations are: U ( ) inf ((,), ( )subset of attributes R P Q ? ? such that ( ) ( ) Y Y P R ? ? = ??? ??? is called Y ? ??? reduct) P P R X R X y x I x y y µ µ µ ? =**(1)**

U ( ) sup ((,), ( )) P P R X R X y x T x y y µ µ µ ? = (2)

Here, I is a fuzzy implicator and T a t-norm. R P is the fuzzy similarity relation induced by the subset of features P:(,) { (,)} P a R a P R x y T x y µ µ ? = (**3**) (D) U/D ( ) su p ( ) R P P POS R X X x x µ µ ? =**(4)**

An important issue in data analysis is the discovery of dependencies between attributes. The fuzzy-rough dependency degree of D on the attribute subset P can be defined as:(D) ' U ( ) (D) | U | R P POS x P x µ ? ? = ?**(5)**

A fuzzy-rough reduct R can be defined as a minimal subset of features which preserves the dependency degree of the entire dataset' ' (D) (D) R C ? ? =

The fuzzy k-means clustering algorithm partitions data points into k clusters S l (l = 1, 2,? k) and clusters S l are associated with representatives (cluster center) C l . The relationship between a data point and cluster representative is fuzzy. That is, a membership u i,j ? [0, 1] is used to represent the degree of belongingness of data point X i and cluster center C j .

Denote the set of data points as S = {X i }. The FKM algorithm is based on minimizing the following distortion., 1 1 k N m i j ij j i J u d = = = ??

With respect to the cluster representatives C j and memberships u i, j , where N is the number of data points; m is the fuzzifier parameter; k is the number of clusters; and d ij is the squared Euclidean distance between data point X i and cluster representative C j . It is noted that u i, j should satisfy the following constraint:1, for i=1to N , 1 k u i j j = ? =

The major process of FKM is mapping a given set of representative vectors into an improved one

4 ?

is the degree to which objects x and y are similar for feature a, and may be defined in many ways. In a similar way to the original crisp rough set approach, the fuzzy positive region can be defined as:

through partitioning data points. It begins with a set of initial cluster centers and repeats this mapping process until a stopping criterion is satisfied. It is supposed that no two clusters have the same cluster representative. In the case that two cluster centers coincide, a cluster center should be perturbed to avoid coincidence in the iterative process. If d ij < ?, then u i, j = 1 and u i,l = 0 for l ? j, where ? is a very small positive number. The fuzzy kmeans clustering algorithm is now presented as follows. 1. Input a set of initial cluster centers SC 0 = {C j (0)} and the value of ?. Set p = 1. 2. Given the set of cluster centers SC p , compute d ij for i = 1 to N and j = 1 to k. Update memberships u i, j using the following equation:1 1/ 1 1/ 1 , 1 1 ( ) m k m i j ij l d il u d ? ? ? = ? ? ? ? ? ? ? ? ? ? = ? ? ? ? ? ? ? If d ij < ?, set u i, j = 1,

where ? is a very small positive number.

3. Compute the center for each cluster using next equation below to obtain a new set of cluster representatives SC p+1 .1 1 X ( ) N m ij i i j N m ij i u C p u = = = ? ? If ||C j (p) ? C j (p ? 1)

|| < ? for j = 1 to k, then stop, where ? > 0 is a very small positive number. Otherwise set p + 1 ? p and go to step 2.

The major computational complexity of FKM is from steps 2 and 3. However, the computational complexity of step 3 is much less than that of step 2. Therefore the computational complexity, in terms of the number of distance calculations, of FKM is O(Nkt), where t is the number of iterations.

5 IV.

6 Proposed Method

Fuzzy k-means is one of the traditional algorithms available for the clustering. However this algorithm is crisp as it allows an object to be placed exactly in only one cluster. To overcome the disadvantages of crisp clustering fuzzy based clustering was introduced. The distribution of member is fuzzy based methods can be improved by rough clustering. Based on the lower and upper approximations of rough set, the rough fuzzy k-means clustering algorithm makes the distribution of membership function become more reasonable. a) Rough Set Based Fuzzy K-Means Algorithm Specific steps of the RFKM clustering algorithm are given as follows:

Step 1 : Determine the class number k (2<=k<=n), parameter m, initial matrix of member function, the upper approximate limit A i of class, an appropriate number ? > 0 and s = 0.

Step 2 : We can calculate centroids with the formula given below:n n m m ij j ij j 1 j 1 U X / U i C = = = ? ?**(6)**

Step 3 : If j X ? the upper approximation, then U ij = 0.

Otherwise, update Uij as shown below:ij 2 k ij j 2 l=1 ij 1 U d 1 ,x Rwi( ) d m-1 = ? ?**(7)**

Step 4 : If( ) s+1 | U U || s ? ? < i.

Obtain each Feature's Membership Value First, initial cluster centers {P1, P2? Pc} were generated by randomly choosing c points from an image point set. Where c ? [cmin, cmax], cmin =2, cmax = n (n is the image pixels number). Each cluster centers Pi is represented by n numeric image features {F i , i=1, 2,...n}. Then each feature F i is described in terms of its fuzzy membership values corresponding to three linguistic fuzzy sets, namely, low (L), medium (M), and high (H), which characterized respectively by a ?membership function. i i i i F c for F c i F c for F c F ? ? ? ? ? µ ? = ? ? ? ?

Where ? is the radius of the ? -membership function with c as the central point. To select the center c and radius ?. Thus, we obtain an initial clustering centers set where each cluster center is represented by a collection of fuzzy set.

7 ii. Constitute a Decision Table for the Initial cluster Centers Set

Definition 1 Degree of similarity between two different cluster centers is defined as:

8 Aâ??”B, Bâ??”C? Aâ??”Bâ??”C

Based on what mentioned above, taking initial cluster centers as objects, taking cluster centers features F i , the central point c and the radius ? as conditional attributes, taking degree of similarity between two different cluster centers as decision attribute by computing the -membership function value, then a decision table for the initial cluster centers set can be constituted as follows:

9 T= <U?P?R?C?D>

Where U= {x i , i=1, 2?m}, it denotes a initial cluster centers set; P?R is a finite set of the initial cluster center category attributes (where P is a set of condition attributes, R is a set of decision attributes); C= {p i , i=1, 2?n} (where pi is a domain of the initial cluster center category attribute); D:U?P?R?C is the redundant information mapping function, which defines an indiscernibility relation on U.

iii. Eliminating redundant cluster centers from the initial cluster centers set Assuming D(x) denotes a decision rule, D(x)|P (condition) and D(x)|R (decision) denote the restriction that D(x) to P and R respectively, I and j denotes two different cluster centers respectively, and other assumptions are as the same as what mentioned above. Based on what described above, the initial cluster centers set can be optimized by reduction theory according to the following steps: Based on what descript above, now the procedure for Rough Sets based FKM image segmentation method can be summered as follows:

Step 1 : Randomly initialize the number of clusters to c, where 2 ? c ??n and n is number of image points.

Step 2 : Randomly chooses c pixels from the image data set to be cluster centers.

Step 3 : Optimize the initial cluster centers set by Rough Sets.

Step 4 : Set step variable t=0, and a small positive number ?. Step 6 : Update the cluster centers by equation (7).

Step 7 : Compute the corresponding Xie-Beni index using equation (12).

Step 8 : Repeat step 5-8 until

10 Experiment Results

In this section, experimental results on real images are described in detail. In these experiments, the number of different types of object elements in each image from manual analysis was considered as the number of clusters to be referenced. They were also used as the parameter for FKM. The Xie-Beni index value has been utilized throughout to evaluate the quality of the classification for all algorithms. All experiments were implemented on PC with 1.8 GHz Pentium IV processor using MATLAB (version 9.0). ? If an initial cluster center set decision table are compatible, then when p?P and Ind(P)=Ind (P-p), p is a redundant attribute and it can be leaved out, otherwise p can't be leaved out.

If an initial cluster center set decision table are not compatible, then computing its positive region POS(P, R). If p?P, when POS(P, R)=POS (P-p, D), then p can be leaved out, otherwise p can't be leaved out. Proposed algorithm applied on all the images shown above. This RFKM image segmentation method partitions into different regions exactly. Visually as well as theoretically our method gives better results other than state of the art methods like, FCM, RFCM. We present a clustering time of experiment for 2 experiments and shows that RFKM performs better than FCM and RFCM.

11 Conclusions

We employed Rough Sets to FKM image segmentation. By reduction theory (the core of Rough Sets), the vagueness and uncertainty information inherent in a given initial cluster center set is analyzed, and those redundant initial cluster centers in the initial cluster set is then eliminated, the reduced initial cluster center set as input to FKM for the soft evaluation of the segments, this is very useful for overcoming the drawbacks of conventional FKM segmentation overdependence on initial value. To evaluate the quality of clusters, the Xie-Beni index was used as the cluster validity index. Experimental results indicate the superiority of the proposed method in image segmentation. [1] [2]

[1] FImage Segmentation using Rough set Based Fuzzy K-Means Algorithm

[2] F Image Segmentation using Rough set Based Fuzzy K-Means Algorithm

Figure 1: ?

Figure 2:

5

Figure 3: Step 5 :

Figure 4: F

1

Figure 5: Figure 1 :

1. Deducing the compatibility of each rule of an initial cluster center set decision table z If D(I)|P (condition)= D(j)|P (condition) and D(i)|R (decision)= D(j)|R (decision), then rules of an initial cluster center set decision table are compatible;
If D(i)|P (condition)= D(j)|P (condition), but D(i)|R (decision) ?D(j)|R (decision), then rules of an initial cluster center set decision table are not compatible.
2. Ascertaining redundant conditional attributes of an initial cluster center set decision table.
3.

Figure 6:

1

	Average of the XB index values	Clustering time (in sec)
FCM	0.034024	13.64
RFCM	0.031578	6.48
RFKM	0.029197	5.92

Figure 7: Table 1 :

[Yong et al. ()] 'A Novel Fuzzy C-Means Clustering Algorithm for Image Thresholding'. Y Yong , Z Chongxun , L Pan . *Measurement Science Review* 2004. 4 (1) .

[Ht Farrah Wong and Ramachandran (2001)] 'An image segmentation method using fuzzy-based threshold'. Nagarajan Ht Farrah Wong , Ramachandran . *International symposium on signal processing and its applications (ISSPA)*, August (2001. p. .

[Rao and Vidyavathi ()] 'Comparative Investigations and Performance Analysis of FCM and MFPCM Algorithms on Iris data'. V S Rao , Dr S Vidyavathi . *Indian Journal of Computer Science and Engineering* 2010. 1 (2) p. .

[Rui and Sousa ()] 'Comparison of fuzzy clustering algorithms for Classification'. A Rui , J M C Sousa . *International Symposium on Evolving Fuzzy Systems*, 2006. p. .

[Russo ()] 'Edge detection in noisy images using fuzzy reasoning'. F Russo . *IEEE transactions on instrumenttation and measurement*, 1998. 47 p. .

[Borji and Hamidi (2007)] 'Evolving a fuzzy rule base for image segmentation'. A Borji , M Hamidi . *Proceedings of world academy of science, engineering and technology*, (world academy of science, engineering and technology) July 2007. 22 p. .

[Hui et al. ()] 'K-Means clustering versus validation measures: A data distribution perspective'. X Hui , J Wu , C Jian . *IEEE Transactions on Systems, Man, and cybernetics* 2009. 39 (2) p. .

[Pawlak ()] 'Rough sets'. Z Pawlak . *Internat. J. Comput. Inform. Sci* 1982. 11 p. .

[Pawlak ()] *Rough sets: Theoretical aspects of reasoning about data*, Z Pawlak . 1991. 9. Knowledge Engineering and Problem Solving. Dordrecht: Kluwer

[Nirulata and Meher] 'Skin Tumor Segmentation using Fuzzy c-means Clustering with Neighbourhood Attraction'. K Nirulata , S Meher . *Communicated to International Journal of Computers and Electrical Engineering*

[Rakhlin and Caponnetto ()] 'Stability of K-Means clustering'. A Rakhlin , A Caponnetto . *Advances in Neural Information Processing Systems*, (Cambridge, MA) 2007. MIT Press. p. .