Analysis of fuzzy clustering and a generic fuzzy rule-based image segmentation technique

Conference or Workshop Item

How to cite:
Karmakar, Gour C. and Dooley, Laurence S. (2001). Analysis of fuzzy clustering and a generic fuzzy rule-based image segmentation technique. In: International Conference Intelligent Multimedia and Distance Education (ICIMADE ’01), 1-3 Jun 2001, Fargo, North Dakota, USA.

For guidance on citations see FAQs.

© [not recorded]
Version: [not recorded]

Link(s) to article on publisher’s website:
http://www.ece.ndsu.nodak.edu/ece/research/conferences/icimade01/

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.
ANALYSIS OF FUZZY CLUSTERING AND A GENERIC FUZZY RULE-BASED IMAGE SEGMENTATION TECHNIQUE

Gour C Karmakar and Laurence S Dooley
Email:{Gour.Karmakar, Laurence.Dooley}@infotech.monash.edu.au
Gippsland School of Computing and Information Technology
Monash University, Churchill, Victoria- 3842, Australia

ABSTRACT

Many fuzzy clustering based techniques when applied to image segmentation do not incorporate spatial relationships of the pixels, while fuzzy rule-based image segmentation techniques are generally application dependent. Also for most of these techniques, the structure of the membership functions is predefined and parameters have to either automatically or manually derived. This paper addresses some of these issues by introducing a new generic fuzzy rule based image segmentation (GFRIS) technique, which is both application independent and can incorporate the spatial relationships of the pixels as well. A qualitative comparison is presented between the segmentation results obtained using this method and the popular fuzzy c-means (FCM) and possibilistic c-means (PCM) algorithms using an empirical discrepancy method. The results demonstrate this approach exhibits significant improvements over these popular fuzzy clustering algorithms for a wide range of differing image types.

1. INTRODUCTION

Classical so-called “crisp” image segmentation techniques while effective when an image contains well-defined structures, such as edges and regular shapes, do not perform so well in the presence of ill-defined data. In such circumstances, the processing of images that possess ambiguities produce fuzzy regions. Fuzzy image segmentation techniques can cope with the imprecise data well and can be broadly classified into five general classes: fuzzy clustering, fuzzy rule based, fuzzy geometry, fuzzy thresholding, and fuzzy integral based image segmentation techniques [1]. Of these, the foremost are fuzzy clustering and fuzzy rule based segmentation. The most popular and extensively used fuzzy clustering techniques are: fuzzy c-means (FCM) [2-3] and possibilistic c-means (PCM) algorithms [4], although both techniques are unable to integrate either human expert knowledge or spatial relation information.

Fuzzy rule based image segmentation techniques in contrast are able to integrate human expert knowledge. They are also less computationally expensive than fuzzy clustering and able to interpret linguistic as well as numeric variables [5]. They are however very much application dependent and it is difficult to define fuzzy rules that cover all pixels. In most approaches, the structure of the membership functions is predefined and their parameters are manually or automatically determined [5-9]. In addition to these advantages, any fuzzy rule-based image segmentation should be both application and image independent and be capable of incorporating spatial information about the regions. The membership functions and their parameters should also be able to be defined automatically.

This paper explores the development of a novel generic fuzzy rule based image segmentation (GFRIS) technique, which addresses many of the aforementioned issues [11]. Section 2 explores the GFRIS method that defines the membership function, together with the underlying concepts and fuzzy rule definition, while sections 3 and 4 contrast the FCM and PCM fuzzy clustering algorithms respectively. The segmentation results in section 5 are quantitatively evaluated by applying a superior objective segmentation assessment method [12], the “discrepancy based on the number of mis-segmented pixels” on two different image types, namely a gray scale and a medical x-ray image. Finally, some conclusions are presented in section 6.

2. GENERIC FUZZY RULE BASED IMAGE SEGMENTATION TECHNIQUE (GFRIS)

The GFRIS technique uses three types of membership function to respectively represent the region pixel distributions, the closeness to their centres and the spatial relations among the pixels in a particular region. Each membership function possesses a membership value for every region, which indicates the degree of belonging to that particular region [10,11]. It also uses a single fuzzy rule. Details of the algorithm applied to automatically define the membership function and fuzzy rule are described in the following sections.

2.1. Membership Function for Region Pixel Distributions

This section outlines the stages used to automatically define the membership function including its structure from the region pixel distributions. The three steps required to define the membership function are:

i. Classify the image into a desired number of regions using manual segmentation or automatically by applying any fuzzy clustering algorithm.

ii. Generate the gray level pixel intensity histogram for each region and normalise the frequency for each gray level into the range [0 1].
iii. Approximate the polynomial for each region. This polynomial represents the membership function for that particular region and the value of the polynomial for each gray level denotes the membership value of that particular gray level value.

The degree of belonging of a candidate pixel (the pixel to be classified) to a region is determined from the respective membership function. The membership function \( \mu_{\text{rel}}(P_{s,t}) \) of the region \( R_j \) for the pixel distribution is defined as

\[
\mu_{\text{DR}_j}(P_{s,t}) = f_{R_j}(P_{s,t})
\]  

(1)

where \( f_{R_j}(P_{s,t}) \) and \( P_{s,t} \) are the polynomial of the region \( R_j \) and the pixel at position \((s,t)\) respectively.

### 2.2 Membership Function to Measure the Closeness of the Region

This type of membership function represents the similarity between the candidate pixel and the centre of a region on the gray level intensity, and is based upon a Euclidean distance measure. The degree of belongingness of a candidate pixel to a region is determined by applying the \textit{k-means} clustering algorithm. Candidate pixels join their nearest region and after joining, the centre of that particular region is recomputed. The centroid of a region \( R_j \) is defined as

\[
C(R_j) = \frac{1}{N_j} \sum_{i=1}^{N_j} P_j(i)
\]  

(2)

where \( N_j \) and \( P_j(i) \) represent the number of pixels and the \( i \)th pixel gray level intensity of the \( j \)th region respectively.

The membership function should reflect the axiom that "the closer to a region, the larger the membership value a pixel should have". So the membership function \( \mu_{\text{cr}_j}(P_{s,t}) \), which determines the degree of belongingness of a candidate pixel \( P_{s,t} \) at a location \((s,t)\) to a region \( R_j \) can be defined as

\[
\mu_{\text{cr}_j}(P_{s,t}) = 1 - \frac{|C(R_j) - P_{s,t}|}{D}
\]  

(3)

where the constant \( D \) is defined as the difference between the maximum and minimum gray level intensity values of an image, so for an 8 bit gray scale image, \( D=255 \). The maximum value of the membership function will always be at the centre of the region and the structure of the membership function will be symmetrical about the vertical line that passes through the centre of the region.

### 2.3. Membership Functions for Spatial Relation

In the previous sections, the membership functions have been developed based only on feature values i.e. gray level intensities of a particular image, and thus did not consider any spatial relationships of the pixels within an identified region. Clearly, there is an expectation that strong spatial relationships will exist between neighbouring pixels within a region, while at the same time there also could be a considerable number of overlapping pixels between the regions. Good segmentation cannot therefore be expected unless these overlapping pixels are taken into account. By considering the neighbourhood relation between a candidate pixel and the classified pixels of the regions, the number of overlapping pixels can be reduced. Based on the neighbourhood relation the candidate pixel can be assigned to the appropriate group. In this paper, we concentrate especially on fixed size neighbourhoods around a candidate pixel. The neighbourhood configurations of the pixels for \( r=1 \), \( r=2 \) and \( r=4 \) are shown in the figures 1(a), (b) and (c) respectively, where \( O \) and \( # \) represent the candidate and neighbourhood pixels respectively.

![Neighbourhood System](image)

**Figure 1: Neighbourhood system**

The neighbourhood system of a region is defined as,

**Definition 1 (Neighbourhood system)** A neighbourhood system with radius \( r \), \( \zeta(P_{s,t}, r) \) of a candidate pixel \( P_{s,t} \) is a set of all pixels \( P_{s,t} \) such that \( \zeta(P_{s,t}, r) = \{P_{s,t}, |d(P_{s,t}, P_{s,t}) \leq r \} \wedge ((P_{s,t} - P_{s,t}) \subseteq T) \) where distance \( d(P_{s,t}, P_{s,t}) = |x-y| + |y-t| \).

\( P_{s,t} \) is a 2D image pixel at Cartesian coordinate \((x,y)\), \( r \) is the radius of the neighbourhood system, and \( T \) is the threshold.

The membership function \( \mu_{\text{sr}_j}(P_{s,t}, r) \) of the region \( R_j \) is defined as

\[
\mu_{\text{sr}_j}(P_{s,t}, r) = \frac{N_j \times G_{R_j}}{\sum_{j=1}^{N_j}(N_j \times G_{R_j})}
\]  

(4)

where \( N_j = \left[ \zeta(P_{s,t}, r) \right] \) is the number of neighbourhood pixels of the candidate pixel \( P_{s,t} \) in the region \( R_j \), \( (G_{R_j}) \) is the sum of inverse pixel distances, and \( \mathcal{R} \) is the number of regions in an image to be segmented.

The membership function of a region defined in (4) considers the number of neighbours and their sum of inverse distances for all regions. The greater the number of neighbours in a region, the larger the value of the membership function will be for that region.

### 2.4 Fuzzy Rule Definition
The overall membership value $\mu_{AR_j}(P_{i,c})$ of a pixel $P_{i,c}$ for the region $R_j$, which represents the overall degree of belonging to the region $R_j$, can be defined by the weighted average of the values of the membership functions $\mu_{ER_j}(P_{i,c})$, $\mu_{VR_j}(P_{i,c})$ and $\mu_{VR_k}(P_{i,c})$.

$$\mu_{AR_j}(P_{i,c}) = \frac{W_{mu_{ER_j}(P_{i,c})} + W_{mu_{VR_j}(P_{i,c})} + W_{mu_{VR_k}(P_{i,c})}}{W_{E} + W_{V} + W_{S}}$$ (5)

where $W_{E}$, $W_{V}$ and $W_{S}$ are the weighting factors of the membership values for the pixel distribution, closeness to the cluster centres and neighbour relations respectively. The overall membership value $\mu_{AR_j}(P_{i,c})$ is used in the antecedent condition of the fuzzy IF-THEN rule, which is defined as,

**Definition 2 (Fuzzy Rule)** IF $\mu_{AR_j}(P_{i,c})$ supports region $R_j$, THEN pixel $P_{i,c}$ belongs to region $R_j$.

$\mu_{AR_j}(P_{i,c})$ will give support to the region $R_j$ if $\mu_{AR_j}(P_{i,c}) = \max\{\mu_{AR_1}(P_{i,c}), \mu_{AR_2}(P_{i,c}), \ldots, \mu_{AR_R}(P_{i,c})\}$

where $R$ indicates the number of regions.

### 3 Fuzzy c-Means (FCM) Algorithm

FCM is the most popular fuzzy based clustering technique. Developed by Bezdek [3], it is still being used today in image segmentation. It performs classification based on the iterative minimization of the following objective function and associated constraints [2].

$$f_{\mu}(\mu, V, X) = \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^m d_i^2$$ (6)

$$0 \leq \mu_{ij} \leq 1 \quad \forall i \in \{1..c\} \text{ and } j \in \{1..n\}$$ (7)

$$\sum_{i=1}^{c} \mu_{ij} = 1 \quad \forall j \in \{1..n\}$$ (8)

$$0 < \sum_{i=1}^{c} \mu_{ij} < n \quad \forall i \in \{1..c\}$$ (9)

where $c$ and $n$ are the number of cluster and data respectively, $\mu$ is a fuzzy partition matrix containing membership values $[\mu_{ij}]$, $V$ is a prototype vector containing the values of cluster centres $[V_1]$, $m$ is the fuzzifier $(1 \leq m \leq \infty)$, $d$ is the distance between $X_j$ & $V_i$, and $X$ is a data vector. The following two equations are derived after minimization of the function $f_{\mu}(\mu, V, X)$ in (6) with respect to $\mu$ and $V$.

$$\mu_{ij} = \frac{1}{1 + \left(\frac{d_i^2(X_j, V_i)}{\eta_i} \right)^{\frac{1}{m-1}}}$$ (10)

$$v_i = \frac{\sum_{j=1}^{n} \mu_{ij}^m x_j}{\sum_{j=1}^{n} \mu_{ij}^m}$$ (11)

The set of cluster centres is initialised either randomly or by an approximation method. The membership values and cluster centres are updated through an iterative process until the maximum change in $\mu_{ij}$ becomes less than a predefined threshold. The selection of the value of $m$ is important, as if $m=1$, then FCM produces a crisp instead of a fuzzy partitioning. Note, that if any of the distance values $d(X_j, V_i)$ is zero, then equation (11) is undefined.

### 4 Possibilistic c-Means (PCM) Algorithms

FCM arbitrarily divides the data set based on a selected number of clusters. The membership values generated by FCM represent the degrees of sharing. In order to eliminate the constraints in equation (8), Krishnapuram and Keller first introduced PCM whose membership values represent the degrees of typicality, instead of degrees of sharing and clusters are independent with each other [4,13]. They modified the FCM objective function and defined the PCM objective function as,

$$f_{\mu}(\mu, V, X) = \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^m d_i^2 \left(\frac{x_j}{v_i}ight)^2$$ (12)

with the constraints being

$$0 \leq \mu_{ij} \leq 1 \quad \forall i \in \{1..c\} \text{ and } j \in \{1..n\}$$ (13)

$$0 < \sum_{i=1}^{c} \mu_{ij} < n \quad \forall i \in \{1..c\}$$ (14)

$$\max \mu_{ij} > 0 \quad \forall j \in \{1..n\}$$ (15)

where $\eta_i$ is the scale parameter, which determines the zone of influence of a point and other parameters are as defined in section 3. The following are obtained after minimizing the function $f_{\mu}(\mu, V, X)$.

$$\mu_{ij} = \frac{1}{1 + \left(\frac{d_i^2(X_j, V_i)}{\eta_i} \right)^{\frac{1}{m-1}}}$$ (16)
The membership value ($\mu_i$) and prototype centre ($v_i$) are updated using the equations (16) and (17) through an iterative process. As with FCM, when fuzzifier $m=1$, PCM produces a crisp partition. PCM offers more promising results in presence of noise but it is highly dependent on initialisation and estimation of the scale parameters. The output of FCM can be used for initialisation and scale estimation but FCM is very sensitive to noise.

5 EXPERIMENTS

In the proposed system, both FCM and PCM were implemented using MATLAB 5.3.1 and two different example images were used in the experiments, namely a gray scale image showing a cloud and urban scene shown in figure 2(a) and a medical X-ray image of the human vocal tract shown in figure 2(d).

![Figure 2: Original cloud and x-ray of human vocal tract and their reference images](image)

For FCM the initialization of the cluster centre was performed randomly. The maximum number of iterations, the minimum level of improvement and the value of the fuzzifier ($m$) were empirically evaluated as 100, 0.00001 and 2 respectively. For PCM, the initialization of the cluster centres used the output of the FCM. The value of scale parameter $\eta$ was taken as the variance of the cluster $i$ produced by FCM [13].

Using the proposed Generic Fuzzy Rule Based Image Segmentation (GFRIS) technique, the membership function defined in section 2.1 was developed using the clusters produced by FCM and their centre values were used to initialise the centres of the clusters required to define the membership function, as described in section 2.2. The respective values of the weights and threshold were determined empirically as $W_1 = 1$, $W_2 = 2$, $W_3 = 1$, $T=25$, for the image in figure 2(a) and $W_1 = 1$, $W_2 = 1.5$, $W_3 = 1$, $T=30$ for the X-ray image in figure 2(d). The segmented results of the gray scale image for the two regions (cloud and urban scene) produced by FCM, PCM and GFRIS respectively are displayed in the figure 3.

![Figure 3: The segmented results of the cloud image with two regions by FCM, PCM and GFRIS](image)

The results clearly show that GFRIS separated almost all the cloud from the image and produced significantly better results than both FCM and PCM. FCM and PCM gave approximately equal performance since as alluded earlier, both techniques do not consider the spatial relationships between the pixels comprising each region. GFRIS also exhibited better results for larger values of neighbourhood radius $r$, because the pixels of region 1 (cloud) are homogeneous and possess very strong spatial correlation.

The quantitative evaluations were performed using one of the most powerful empirical discrepancy methods [12] based upon the number of wrongly segmented pixels. The confusion matrix $C_{ij}$ is a $R \times R$ square matrix where $R$ represents the number of segmented region and $C_{ij}$ denotes the number of $i^{th}$ region pixels classified as region $j$ by segmentation. Type I error, $errorI$ is defined as,

$$errorI_i = \frac{\sum_{j=1}^{R} C_{ji} - C_{ii}}{\sum_{j=1}^{R} C_{ji}} \times 100$$

(18)

while a Type II error, $errorII$ is defined as,

$$errorII_i = \frac{\sum_{j=1}^{R} C_{ij} - C_{ii}}{\sum_{i=1}^{R} \sum_{j=1}^{R} C_{ij}} \times 100$$

(19)

The reference images in figure 2 were again used for evaluation purposes. The results of the cloud image segmentation with respect to reference images (figures 2(b) and 2(c)) are shown in Table 1.

![Image of the human vocal tract shown in figure 2(d).](image)
In the above table, the image is segmented into two regions, so the error rates refer to incorrect segmentation for region $R_1$ (clouds). Since the error rate of one region will be the inverse of the error rate of other region, the results reveal that GFRIS provides superior performance for region $R_1$, which indicates that GFRIS successfully separated the cloud from the image and represents the underlying structure of data far better than FCM and PCM. The error rates of GFRIS for type II error are higher than for both PCM and FCM because the pixels in this region do not have good continuation i.e. they are abruptly changing, which oppose a strong spatial relation. In fact, the urban scene is not a single object. Good continuation is one of the seven properties of grouping of the visual elements [14]. The average error rates of the three techniques are shown in the figure 4.

This graph shows that the average error rates of GFRIS are much less than those of PCM and FCM. Average error rate of GFRIS for $r=4$ is higher than that of for $r=2$ because there is no sharp boundary between cloud and urban scene. For this case, GFRIS interpreted some sections of the urban scene as cloud for $r=4$. PCM again showed slightly better performance than FCM.

A second series of experiments were performed using a medical x-ray image of the human vocal tract (figure 2(d)). The segmentation was again for two regions, namely the human vocal tract (region $R_1$, figure 2(e)) and general background (region $R_2$). The corresponding results produced by FCM, PCM and GFRIS are presented in figure 5.

## Table 1: Percentage errors for cloud (region $R_1$) segmentation in figure 2(a).

<table>
<thead>
<tr>
<th>Method</th>
<th>Error Type I</th>
<th>Error Type II</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCM</td>
<td>28.7335</td>
<td>17.4194</td>
</tr>
<tr>
<td>PCM</td>
<td>27.1375</td>
<td>18.3409</td>
</tr>
<tr>
<td>GFRIS r=1</td>
<td>8.8332</td>
<td>20.4783</td>
</tr>
<tr>
<td>GFRIS r=2</td>
<td>1.9749</td>
<td>21.4497</td>
</tr>
<tr>
<td>GFRIS r=4</td>
<td>2.0388</td>
<td>23.9742</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: Error percentage for human vocal tract (region $R_1$) of x-ray of human vocal tract segmentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>FCM</td>
</tr>
<tr>
<td>PCM</td>
</tr>
<tr>
<td>GFRIS r=1</td>
</tr>
<tr>
<td>GFRIS r=2</td>
</tr>
<tr>
<td>GFRIS r=4</td>
</tr>
</tbody>
</table>

It is visually evident that the proposed technique GFRIS considerably outperforms both the FCM and PCM techniques for this image type as well. The image (figure 2(d)) contains two regions, the vocal tract (comprising the lips, tongues, teeth, aural cavity) and general background. The soft part of the human vocal tract is not clearly visible and has low local contrast pixels [15]. Almost the entire vocal tract had been successfully separated by GFRIS using $r=4$, which confirms that the larger values of $r$, provide a better representation of the spatial relation. Here PCM showed slightly better performance than FCM. The error rates of human vocal tract segmentation with respect to the reference images (figures 2(e) and 2(f)) are shown in the table 2. Both types of errors for human vocal tract segmentation are less than FCM and PCM except the error rate of error type II of GFRIS using $r=4$. This is caused by the fact there is good continuation of low contrast pixels of human vocal tract with the background and it takes some portion of the background as a part of human vocal tract for higher order of spatial relation i.e. $r=4$. 

This graph shows that the average error rates of GFRIS are much less than those of PCM and FCM. Average error rate of GFRIS for $r=4$ is higher than that of for $r=2$ because there is no sharp boundary between cloud and urban scene. For this case, GFRIS interpreted some sections of the urban scene as cloud for $r=4$. PCM again showed slightly better performance than FCM.
Figure 6: Average error rates of PCM, FCM and GFRIS for human vocal tract segmentation

All the average error rates for GFRIS are less than those of FCM and PCM. The error rate is decreasing rapidly for higher orders of spatial relation, because the pixels of both regions are almost homogeneous. The error rate of FCM is higher than PCM.

For all of the above experiments the number of regions to be segmented was two. It is important however to use a larger number of regions in order to check the underlying meaning of data. To achieve this, another experiment was performed using the above techniques identifying three regions to be segmented. From the experimental results, it was shown that GFRIS considered the underlying meaning of data better than FCM and PCM and out performed both of them for both types of images for three regions. PCM again showed better underlying structure of the data than FCM for both types of images when three regions were to be segmented.

6 CONCLUSIONS

In this paper both quantitative and qualitative performance analysis of FCM, PCM and GFRIS have been performed based on the standard segmentation evaluation method, empirical discrepancy method using two different types of images. Both results proved that the new generic fuzzy-clustering algorithm (GFRIS) provided significantly better results than either FCM or PCM. The reasons for this are that GFRIS has considered spatial relationships very well and hence represented the underlying meaning of data better than both FCM and PCM. PCM has considered the underlying structure of data in some extent but PCM has arbitrarily divided the data into region without considering any underlying meaning of data.

The values of the weighting factors $W_1$, $W_2$, and $W_3$ of GFRIS were determined empirically. More research needs to be undertaken in order to determine the suitable values of both the three weighting factors as well as the threshold.

Finally, as the proposed technique is fuzzy rule-based, it is capable of incorporating any type of attribute of any special application. It is thus possible to add membership functions from the high level semantics of an object for object-based image segmentation, such as in MPEG-4 applications. Like FCM and PCM, the GFRIS technique needs to be provided with the desired number of regions to be segmented. It also needs further investigation for automatically determining the optimal number of regions.

REFERENCES