ELSEVIER

Pattern Recognition Letters 23 (2002) 1215–1227

Pattern Recognition Letters

www.elsevier.com/locate/patrec

A generic fuzzy rule based image segmentation algorithm

Gour C. Karmakar *, Laurence S. Dooley

Gippsland School of Computing and Information Technology, Monash University, Churchill, Vic. 3842, Australia

Received 7 May 2001; received in revised form 1 November 2001

Abstract

Fuzzy rule based image segmentation techniques tend in general, to be application dependent with the structure of the membership functions being predefined and in certain cases, the corresponding parameters being manually determined. The net result is that the overall performance of the segmentation technique is very sensitive to parameter value selections. This paper addresses these issues by introducing a generic fuzzy rule based image segmentation (GFRIS) algorithm, which is both application independent and exploits inter-pixel spatial relationships. The GFRIS algorithm automatically approximates both the key weighting factor and threshold value in the definitions of the fuzzy rule and neighbourhood system, respectively. A quantitative evaluation is presented between the segmentation results obtained using GFRIS and the popular fuzzy c-means (FCM) and possibilistic c-means (PCM) algorithms. The results demonstrate that GFRIS exhibits a considerable improvement in performance compared to both FCM and PCM, for many different image types. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Generic fuzzy rules; Image segmentation; Spatial information; Fuzzy clustering

1. Introduction

Classical, so-called "crisp" image segmentation techniques, while effective for images containing well-defined structures such as edges, do not perform as well in the presence of ill-defined data. In such circumstances, the processing of images that possess ambiguities is better performed using fuzzy segmentation techniques, which are more adept at dealing with imprecise data. Fuzzy techniques may be broadly classified into five main categories: fuzzy clustering, fuzzy rule based, fuzzy geometry, fuzzy thresholding, and fuzzy integral based segmentation techniques (Tizhoosh, 1998). Of these, the most widely used are fuzzy clustering and fuzzy rule based segmentation.

The two most popular fuzzy clustering techniques are the fuzzy c-means (FCM) (Bezdek, 1981; Chi et al., 1996) and possibilistic c-means (PCM) algorithms (Krishnapuram and Keller, 1993). While both these methods have been applied extensively, neither integrates human expert knowledge nor includes information about pixel spatial relations. Image segmentation which relies upon only feature based information without considering inter-pixel relationships, does not generally

* Corresponding author. Tel.: +51-223-884; fax: +99-026-842.

E-mail addresses: Gour.Karmakar@infotech.monash.edu.au (G.C. Karmakar), Laurence.Dooley@infotech.monash.edu.au (L.S. Dooley).

PII: S0167-8655(02)00069-7

produce good results, because there are usually a large number of overlapping pixel values between different regions.

In contrast, fuzzy rule based image segmentation techniques are able to integrate expert knowledge and are less computationally expensive compared with fuzzy clustering. They are also able to interpret linguistic as well as numeric variables (Chang et al., 1998). The performance of fuzzy rule based segmentation in many applications however, is sensitive to both the structure of the membership functions and associated parameters used in each membership function. For example, the fuzzy rule based segmentation technique proposed by Chi and Yan (1993) for geographic map images, intuitively defined the structure of the membership functions with the related parameters being automatically determined, while Hall and Namasivayam (1998) and Chang et al. (1998) used a different approach for segmenting magnetic resonance images (MRI) of the brain. They predefined the membership functions so the corresponding parameters could be automatically derived. Another approach (Sasaki et al., 1999) was used for segmenting the menisci region from MRI slices, with the structure of the membership functions defined from the anatomical knowledge of the knee and the parameters being taken from actual MRI device data. A different strategy was proposed by Park et al. (1998) who used perceptually selected structures and parameters for the membership functions, in the segmentation of intrathoracic airways trees in computer tomography (CT) images.

Karmakar et al. (2000) presented a contemporary review of fuzzy rule based image segmentation techniques, and confirmed that despite being used in a wide range of applications, both the structure of membership functions and derivation of their relevant parameters were still very much application domain and image dependent.

This paper presents a new generic fuzzy rule based image segmentation (GFRIS) algorithm, which addresses a number of the aforementioned issues, most crucially by incorporating spatial pixel information and automatically data-mining both the key fuzzy rule weighting factor and its threshold (Karmakar and Dooley, 2001). The paper is organised as follows: In Section 2, the three membership functions used in the GFRIS algorithm are defined. The fuzzy rule definition and underlying theory, together with the data-mining algorithm for obtaining both the key weighting factor and threshold are presented in Sections 3 and 4, respectively. Section 5 details the full GFRIS algorithm, while Section 6 discusses the experimental results and performance of this new segmentation technique when applied to a range of different images. All the results are quantitatively evaluated using the empirical objective segmentation assessing method (Zhang, 1996), "*discrepancy based on the number of mis-segmented pixels*". Finally, Section 7 concludes the paper.

2. Definition of membership functions

The definition of the membership function lies at the heart of any fuzzy logic system and the capability of fuzzy rule based techniques significantly depend upon it. The eminent psychologist Gestalt, discovered that visual elements may be perceptually grouped together based on the principles of: proximity, similarity, common fate, good continuation, surroundedness, closure, relative size and symmetry (Wertheimer, 1923). In this section, three membership function types are defined to respectively represent the: (i) region pixel distributions, (ii) closeness to a region's centre, and (iii) pixel spatial relations. The second membership function for instance, characterises similarity based on gray level pixel intensity, while the third reflects the characteristics of proximity and good continuation. Each membership function has a corresponding membership value for every region, which indicates the degree of belonging to that region.

2.1. Membership function for region pixel distributions

In gray level images, every region has a distinctive pixel distribution, which characterises to some extent that region's properties. The approach adopted here is to automatically define the membership function including its structure from the pixel distribution of that particular region. This is achieved in three steps:

1. Segment the original image into a desired number of regions by applying a clustering algorithm such as FCM.
2. Generate the gray level pixel intensity histogram for every region and normalise the frequency for each gray level into the range [0 1].
3. Use a polynomial representation to approximate each region. The polynomial value of a region, for every gray level pixel corresponds to the membership value of that pixel in that region, with the actual gray level intensity values being the parameters of the membership function.

As an example, the reference image shown in Fig. 1(a) is classified into two separate regions, namely R_1 (cloud) and R_2 (urban scene) using the standard FCM algorithm. The respective pixel distribution of each region is used to produce the corresponding membership function and a gray level intensity histogram (*gray level histogram*) is generated for both regions, with the frequencies of occurrence being normalised. A polynomial then approximates the histogram of each region. As an example, a 3rd order polynomial is given by

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3, \tag{1}$$

where x is an independent variable, which in this example is the 8-bit gray level pixel intensity.

The coefficients a_0, a_1, a_2, and a_3 are computed by applying a least squares (LS) fit to the histogram for each region. The values of $f(x)$ are constrained between 0 and 1, and represent the membership value of each gray level pixel. The 3rd order polynomials for the segmented regions R_1 and R_2 in the example image, are shown in Fig. 1(b) and (c), respectively.

The degree of belonging to a region of a candidate pixel, that is the pixel to be classified, is determined from the respective membership function. Hence, for a pixel having a gray level value of 150, the membership values for regions R_1 and R_2 can be easily determined from the respective polynomials as 0.425 and 0.125, respectively. Considering the general case of a pixel with a gray level value of $P_{s,t}$ at location (s,t), then the two membership functions $\mu_{DR_1}(P_{s,t})$ and $\mu_{DR_2}(P_{s,t})$ for the pixel distribution of regions R_1 and R_2, respectively, are expressed as:

$$\mu_{DR_1}(P_{s,t}) = f_{R_1}(P_{s,t}) \tag{2}$$

and

$$\mu_{DR_2}(P_{s,t}) = f_{R_2}(P_{s,t}), \tag{3}$$

where $f_{R_1}(P_{s,t})$ and $f_{R_2}(P_{s,t})$ are the respective polynomials of regions R_1 and R_2.

2.2. Membership function to measure the closeness of a region

This membership function represents the similarity between a candidate pixel and the centre of a region based on gray level pixel intensity and is measured using the *city block distance*. A pixel must always be closer to the belonging region than any other region and the degree of *belongingness* of a candidate pixel to a region is determined from

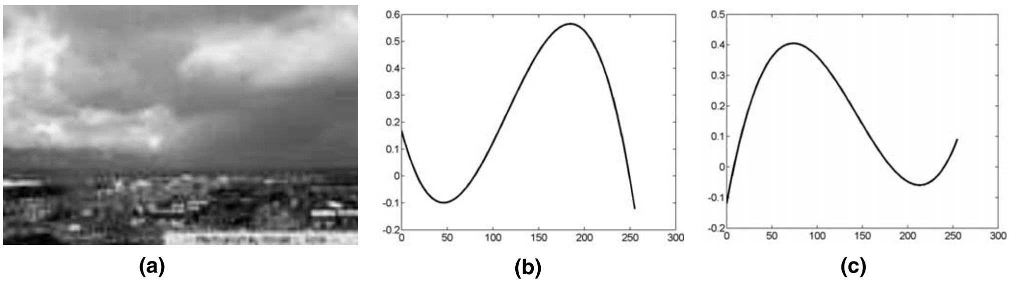

Fig. 1. Reference image and its membership function for each region: (a) original image, (b) membership function for R_1, (c) membership function for R_2.

the k-means clustering algorithm (Gose et al., 1996). When a candidate pixel joins its nearest region, the centre of that particular region is recomputed. The centroid of a region R_j is defined as

$$C(R_j) = \frac{1}{N_j} \sum_{i=1}^{N_j} P_j(i), \tag{4}$$

where N_j is the number of pixels and $P_j(i)$ represents the ith pixel gray level intensity in the jth region.

A membership function should reflect the axiom that "*the closer a pixel is to a region, the larger the membership value that pixel should have*". Hence, the membership function $\mu_{CR_j}(P_{s,t})$, which determines the degree of belongingness of a candidate pixel $P_{s,t}$ at location (s,t), to a region R_j is defined as

$$\mu_{CR_j}(P_{s,t}) = 1 - \frac{|C(R_j) - P_{s,t}|}{D}, \tag{5}$$

where D is a constant equal to the difference between the maximum and minimum gray level intensity values in an image, so using an 8-bit gray scale, $D = 255$.

Theorem 1. *The maximum value of the membership function $\mu_{CR_j}(P_{s,t})$ will always be at the centre of the region and the structure of the function will be symmetrical about a vertical line that passes through the centre of the region.*

Proof. For positive values of D,

$$\frac{|C(R_j) - P_{s,t}|}{D} \geqslant 0.$$

The function $\mu_{CR_j}(P_{s,t})$ will therefore be a maximum whenever $|C(R_j) - P_{s,t}| = 0$, i.e. when $C(R_j) = P_{s,t}$, so the maximum always occurs at $C(R_j)$, which is the centre of region R_j.

To prove the membership function is symmetrical about $C(R_j)$, consider the values of $\mu_{CR_j}(P_{s,t})$ for $P_{s,t} = C(R_j) + \delta$ and $P_{s,t} = C(R_j) - \delta$, where δ is an arbitrary constant.

$$\mu_{CR_j}(C(R_j) + \delta) = 1 - \frac{|C(R_j) - C(R_j) - \delta|}{D} = 1 - \frac{|\delta|}{D},$$

$$\mu_{CR_j}(C(R_j) - \delta) = 1 - \frac{|C(R_j) - C(R_j) + \delta|}{D} = 1 - \frac{|\delta|}{D}.$$

Since $\mu_{CR_j}(C(R_j) + \delta) = \mu_{CR_j}(C(R_j) - \delta)$, $\mu_{CR_j}(P_{s,t})$ is also symmetrical about a vertical line passing through the centre of region R_j. □

2.3. Membership function for spatial relations

The principles of proximity and good continuation are used to define this particular membership function. Wherever pixels are close together and exhibit relatively smooth variations, there is an obvious expectation that strong spatial relationships will exist between neighbouring pixels within that region. In the preceding sections, the respective membership functions have been constructed using only feature values, i.e. gray level pixel intensities. Spatial relations between pixels within an identified region have not been considered, yet are vital since they characterise the geometric features of a region as any spatial object contains two descriptors: feature and geometric (Kellogg et al., 1996; Yip and Zhao, 1996).

In many natural images, there are a large number of overlapping pixels between regions, so that effective segmentation cannot be expected unless these overlapping pixels are taken into account. By considering the neighbourhood relationship between the candidate pixel and the pixels of a region that surround it, a large number of overlapping pixels can be reduced. Based on the neighbourhood relations, the candidate pixel can then be assigned to the appropriate region.

Many approaches exist to define neighbourhood relations (Tuceryan, 2000), such as minimum spanning tree, fixed size neighbourhoods, and Voronoi tessellation. This paper concentrates upon only fixed size neighbourhoods around the candidate pixel, since the number of pixels and their distances from a candidate pixel has to be calculated.

The neighbourhood pixel configurations for $r = 1$, $r = 2$, and $r = 4$ are shown in the Fig. 2(a)–(c), respectively, (Geman and Geman, 1984) where $r \geqslant 1$ denotes the neighbourhood radius, while $\circ$ and # represent the candidate and neighbourhood

```
                                 #  #  #  #  #
   #            #  #  #          #  #  #  #  #
#  O  #         #  O  #          #  #  O  #  #
   #            #  #  #          #  #  #  #  #
                                 #  #  #  #  #
```

(a) r=1 **(b)** r=2 **(c)** r=4

Fig. 2. Neighbourhood configurations.

pixels, respectively. The number of neighbours will be $(r+1)^2$ for $r=1$ and $(r+1)^2-1$ otherwise.

As previously mentioned, the principles of proximity and good continuation imply that pixels, which are close together and have smooth variations should be part of the same region, that is, segmented regions are homogeneous and mutually exclusive. It is thus assumed that the variation of neighbouring pixels in a region is limited to some threshold T, and the neighbourhood system of a region based on this premise is defined as

Definition 1 (*Neighbourhood system*). A neighbourhood system $\zeta(P_{s,t},r)$ with radius r, of a candidate pixel $P_{s,t}$ is the set of all pixels $P_{x,y}$ such that $\zeta(P_{s,t},r)=\{P_{x,y}|(d(P_{x,y},P_{s,t})\leqslant r)\Lambda((P_{x,y}\sim P_{s,t})\leqslant T)\}$ where the distance, $d(P_{x,y},P_{s,t})=|x-s|+|y-t|$, $P_{x,y}$ is the gray level value of the pixel at Cartesian coordinates (x,y), $(P_{x,y}\sim P_{s,t})$ is the absolute value of the difference between the gray level values of the pixels $P_{x,y}$ and $P_{s,t}$, and T is the threshold.

To construct a membership function, the number of neighbourhood pixels and their distances from the candidate pixel must be considered. The membership function μ should possess the following properties:

1. $\mu\propto N$ where N is the number of neighbours.
2. $\mu\propto(1/d(P_{x,y},P_{s,t}))$,

where $d(P_{x,y},P_{s,t})$ is the distance between pixels $P_{x,y}$ and $P_{s,t}$.

The summation of inverse distances of a region R_j is

$$G_{R_j}=\sum_{i=1}^{N_j}\frac{1}{d_i(P_{x,y},P_{s,t})},\tag{6}$$

where $N_j=|\zeta(P_{s,t},r)|$ is the number of neighbourhood pixels of the candidate pixel $P_{s,t}$ in the region R_j and $d_i(P_{x,y},P_{s,t})$ is the distance between the ith pixel $P_{x,y}$ of region R_j and the candidate pixel $P_{s,t}$.

By considering the number of neighbours N_j and the sum of their inverse distances G_{R_j} from the candidate pixel $P_{s,t}$, the membership function $\mu_{NR_j}(P_{s,t},r)$ of the region R_j becomes

$$\mu_{NR_j}(P_{s,t},r)=\frac{N_j\times G_{R_j}}{\sum_{j=1}^{\Re}(N_j\times G_{R_j})},\tag{7}$$

where $\Re$ is the number of segmented image regions. Eq. (7) shows that the greater the number of neighbours in a region, the larger the membership function value will be for that region. Hence, if all neighbours fall into a single region, the corresponding membership function value will be one for that region, since the sum of the membership function values for all regions always equals unity.

3. Fuzzy rule definition

The definition of the fuzzy rule is the single most important and challenging aspect of fuzzy rule based image segmentation, as its effectiveness is vital to the overall performance. In this paper, the fuzzy rule is heuristically defined using the three membership functions defined in Section 2, in combination with the widely used fuzzy IF–THEN rule structure.

The overall membership value $\mu_{AR_j}(P_{s,t})$ of a pixel $P_{s,t}$ for region R_j represents the overall degree of belonging to that region, and is defined by the weighted average of the three individual membership function values $\mu_{DR_j}(P_{s,t})$, $\mu_{CR_j}(P_{s,t})$ and $\mu_{NR_j}(P_{s,t})$, which are given in Eqs. (2), (5) and (7), respectively.

$$\mu_{AR_j}(P_{s,t})=\frac{W_1\mu_{DR_j}(P_{s,t})+W_2\mu_{CR_j}(P_{s,t})+W_3\mu_{NR_j}(P_{s,t})}{W_1+W_2+W_3}.\tag{8}$$

W_1, W_2, and W_3 are the weightings of the membership values for pixel distribution, closeness to the cluster centres, and neighbourhood relations, respectively. The overall membership value

$\mu_{AR_j}(P_{s,t})$ is used in the antecedent condition of the fuzzy IF–THEN rule.

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

$\mu_{AR_j}(P_{s,t})$ will give support to the region R_j if $\mu_{AR_j}(P_{s,t}) = \max\{\mu_{AR_1}(P_{s,t}), \mu_{AR_2}(P_{s,t}), \ldots, \mu_{AR_{\Re}}(P_{s,t})\}$. This rule is deliberately generic so that it can be applied to any image type thus adhering to one of the key objectives that the GFRIS algorithm should be both image and application independent.

4. Determination of weighting factors and the threshold

The threshold value T introduced in Section 2.3, plays a major role in defining the spatial relationship between pixels in any region, because it regulates the level of variation between the candidate pixel and its neighbours. The greater the variation between a candidate pixel and its neighbours, the larger the standard deviation will be, which pro rata results in poor continuation. Two issues need to be considered in determining the threshold value:

1. The degree to which pixels of one region overlap with those of another region.
2. The pixel standard deviations in each region.

The approximate threshold T_a is computed using 1, by considering the centres of the initially segmented regions, while the status of this approximate threshold as to whether it is actually an overestimation of the final threshold value, is determined using 2. Estimation of both the status and final threshold value is detailed in the algorithm below. If the centre of a particular region is two standard deviations away from the boundary of another region and the pixels in that region are normally distributed, there is at best a 5% probability that the pixels of that region will overlap with the other. The procedure to determine the approximate threshold T_a for two regions may be formalised as follows

Theorem 2. *If two regions with centres c_1 and c_2 have pixels that are normally distributed, then for at least 5% levels of significance, the approximate threshold will be bounded by $T_a \leqslant |c_1 - c_2|/4$.*

Proof. Assuming that the pixels are normally distributed, then in a region having a centre c_1 and standard deviation σ_1, the 5% level of significance means the probability of pixels falling outside $c_1 \pm 2\sigma_1$ will be 0.05 (Zaman et al., 1982). The same is also true for other region, which has a centre c_2 and standard deviation σ_2. Thus, for at least 5% levels of significance,

$$2(\sigma_1 + \sigma_2) \leqslant |c_1 - c_2|.$$

Since the threshold is considered the same for both regions, it may be written as $T_a = (\sigma_1 + \sigma_2/2)$ such that

$$4T_a \leqslant |c_1 - c_2| \Rightarrow T_a \leqslant \frac{|c_1 - c_2|}{4}. \quad \square$$

This theory may be extended to an arbitrary number of regions for determining the weight and the threshold values. If the approximate threshold is overestimated, the minimum value between the standard deviations and the approximate threshold is used as the final threshold. This is conditional on the value not being either zero or very small (less than some arbitrary percentage of T_a), so ensuring that some spatial relationship exists. The weight W_1 in Eq. (8) governs the importance assigned to region pixel distributions, and empirical observations reveal that the resultant segmentation results are not very sensitive to variations in this particular parameter.

The important weighting factors are W_2 and W_3, as their values represent a trade-off between the gray level pixel intensity and spatial relationship. Prominence was initially given to the former, because it contributed more to the human visual perception and for this reason, following empirical evaluation; W_2 was set equal to 1.8, with the other two weighting factors being set to one. If the standard deviation in a number of regions is high with respect to the approximate threshold, then the spatial relationship will be ineffective and greater emphasis needs to be given to W_2 by increasing its value. In all other instances, importance should be given to the pixel spatial relationships so that the value of W_2 should be reduced. The following details the various stages of the algorithm to automatically determine this key weighting factor and its threshold.

1. Set the initial values for the three weighting factors as $W_1 = 1$; $W_2 = 1.8$; $W_3 = 1$.
2. Define a set of all regions (R) and a set of centre pairs of all regions (V)
   $$R = \{R_i | (1 \leqslant i \leqslant \Re)\},$$
   $$V = \big\{(C(R_i), C(R_j)) | (\forall i, jR_i, R_j \in R) \wedge (i \neq j)\big\}.$$
3. Compute the absolute sum of differences (*sofd*) between the elements of all pairs
   $$sofd = \sum_{i=1}^{nc2} |V_i(1) - V_i(2)|,$$
   where $nc2$ is the number of combination pairs of all regions.
4. Determine the approximate threshold T_a using Theorem 2
   $$T_a = \frac{sofd}{nc2 \times 4}.$$
5. Calculate the average sum of differences (*arstd*) between the various standard deviations and approximate threshold
   $$arstd = \frac{\sum_{i=1}^{\Re} (rstd_i - T_a)}{\Re},$$
   where $rstd_i$ is the standard deviation of the ith region.
6. If the approximate threshold is overestimated, ($arstd < 0$), then the minimum of the standard deviation and T_a is taken as the final threshold value T, provided this value is neither too small (less than K% of T_a, where K is an arbitrary constant) nor zero. If this condition is invalid, then T_a becomes the final threshold.
7. Normalise the average sum of differences between the standard deviation and approximate threshold
   $$narstd = \frac{arstd}{\max(rstd_i, T_a)}.$$
8. Adjust the weight W_2 accordingly
   $$W_2 = W_2 + narstd.$$

This algorithm has been experimentally tested upon various different image types and as results will prove in Section 6, the automatic data mining of the key weighting factor and threshold value is a significant reason for the superior performance of the GFRIS algorithm.

5. The GFRIS algorithm

The detailed stages involved in the GFRIS algorithm can now be formalised as follows:

1. Classify the pixels of an image into a desired number of regions using any appropriate clustering algorithm.
2. Derive the key weight and threshold value by applying the data-mining algorithm in Section 4, and the membership function for each pixel distribution from the theory given in Section 2.1.
3. Initialise the centre of all regions required to define the membership function in Section 2.2, with the respective centres produced by the clustering algorithm in step 1.
4. Sequentially select an unclassified pixel from the image and calculate each membership function value in each region for that pixel.
5. Classify the pixel into a region applying the fuzzy rule defined in Section 3.
6. Return to step 4 until every pixel is classified.

6. Discussion of experimental results

The GFRIS algorithm, FCM, and PCM were all implemented using MATLAB version 6.0. In order to evaluate the performance of the new GFRIS algorithm, a variety of different image types were applied possessing diverse characteristics, including homogeneous and non-homogeneous regions, low pixel contrast regions and perceptually distinct regions. Three images in particular, Figs. 1(a), 5(a) and 6(a), were used for demonstration and numerical evaluation.

All quantitative evaluations were performed using the powerful empirical discrepancy method (Zhang, 1996) *discrepancy based on the number of* *mis-segmented pixels*. The confusion matrix C is a $\Re \times \Re$ square matrix, where C_{ij} denotes the number of jth region pixels wrongly classified in the ith region by the segmentation algorithm. Two error measures Type I, $errorI_i$ and Type II $errorII_i$, were defined as performance measures:

$$errorI_i = \frac{\left(\sum_{j=1}^{\Re} C_{ji} - C_{ii}\right)}{\sum_{j=1}^{\Re} C_{ji}} \times 100, \tag{9}$$

$$errorII_i = \frac{\left(\sum_{j=1}^{\Re} C_{ij} - C_{ii}\right)}{\left(\sum_{i=1}^{\Re}\sum_{j=1}^{\Re} C_{ij} - \sum_{j=1}^{\Re} C_{ji}\right)} \times 100. \tag{10}$$

Type I, $errorI_i$ represents the percentage error of all ith region pixels that are not classified in the ith region, whereas Type II, $errorII_i$, is the percentage error of all other region pixels wrongly classified in the ith region. The two manually segmented reference regions of the image in Fig. 1(a) used in the evaluation, are shown in Fig. 3.

For FCM, initialisation of the centre of the regions was performed randomly. The maximum number of iterations, the minimum level of improvement and the value of the fuzzifier (m) were empirically selected as 100, 0.00001 and 2, respectively.

For PCM, initialisation of the centre of the regions utilised the output of FCM. The value of the scale parameter η_i (Krishnapuram and Keller, 1993), was taken as the variance of the region i produced by FCM. The maximum number of iterations, minimum level of improvement and value of fuzzifier (m) were empirically chosen as 200, 0.00001 and 1.5, respectively.

For the GFRIS algorithm, the membership function defined in Section 2.1 was constructed using the regions produced by FCM, with their centre values used to initialise the centre of the regions required to define the membership function (Section 2.2). The respective weighting and threshold values were automatically data mined using the algorithm described in Section 4, with the constant $K = 0.25$. The segmented results of

Fig. 3. Manually segmented reference regions of Fig. 1(a): (a) cloud, (b) urban scene.

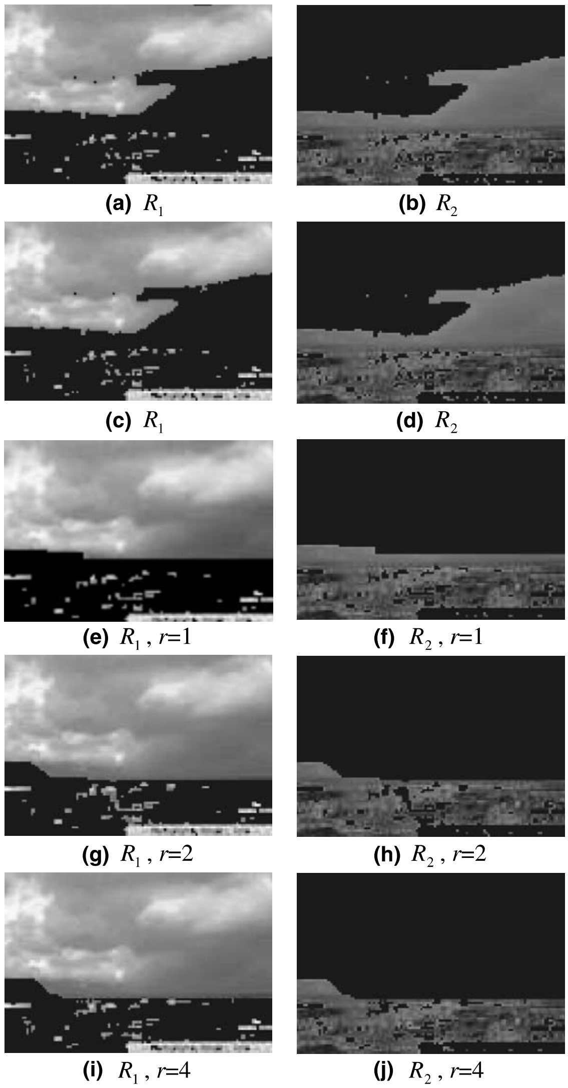

Fig. 4. Automatic segmentation of Fig. 1(a) into two regions using FCM (a)–(b), PCM (c)–(d), and GFRIS (e)–(j).

Table 1
Error percentages for the cloud region (R_1) segmentation in Fig. 1(a)

Algorithm	Error		
	Type I	Type II	Mean
FCM	28.0000	15.7372	21.8686
PCM	26.8939	16.3141	21.6040
GFRIS $r = 1$	7.3333	17.0513	*12.1923*
GFRIS $r = 2$	1.7273	21.2500	*11.4887*
GFRIS $r = 4$	1.8030	23.6218	*12.7124*

the image Fig. 1(a) for the two regions, cloud (R_1) and urban scene (R_2) produced by FCM, PCM and GFRIS, respectively are shown in Fig. 4.

The results confirmed that GFRIS separated almost the entire cloud region from the urban scene and produced significantly better results than both FCM and PCM. FCM and PCM gave approximately equal performance since as alluded earlier, both algorithms do not consider the spatial

Fig. 5. Original iceberg image (a), and the segmented results for three regions produced by FCM (b)–(d), PCM (e)–(g), and GFRIS (h)–(p).

relationships between the pixels in each region. GFRIS also exhibited better results for larger values of neighbourhood radius r, since the pixels of region R_1 (cloud) are homogeneous and possess strong spatial correlation. Evaluation of the segmentation results for the cloud image, compared with the manually segmented reference images in Figs. 3(a) and (b), are shown in Table 1, where the final column is the average of the Type I and Type II errors. Note that only the error rates for the segmented cloud region are displayed in Table 1, because only two regions were identified, and the error rate of one region would be the reverse of that of the other region. The values given in italics correspond to the best GFRIS results.

The average GFRIS error rates for Fig. 1(a) were significantly better than those of both FCM and PCM for each value of the neighbourhood radius r. While GFRIS provided particularly good performance in segmenting the cloud region (R_1), it is worth noting that the error rates of GFRIS for the type II error were higher than those for both PCM and FCM. This was because not all the pixels in this region possessed good continuation due to the abrupt changes in the urban scene, which did not constitute a single object and so opposed the necessary condition for good inter-pixel relationships.

A second series of experiments were performed using the image in Fig. 5(a), which comprised three distinct regions, namely water (R_1), iceberg (R_2), and sky (R_3). The segmentation performance for the three regions using FCM, PCM and GFRIS is presented in Figs. 5(b)–(p).

It was visually apparent again that the GFRIS algorithm produced more distinctive regions for all values of neighbourhood radius r and hence considerably outperformed both FCM and PCM. PCM divided the iceberg image into only two regions (Figs. 5(e) and (f)) instead of three, because it was unable to distinguish between regions having a poor gray level contrast. The error rates for the segmentation of the iceberg image compared with the manually segmented reference images are given in Table 2.

The mean error rates of GFRIS for the iceberg and sky regions were considerably lower than for both FCM and PCM, while the error was slightly higher for the water region. This was due to floating ice on the water, which was classified as water in the manually segmented reference region but was misclassified as sky using GFRIS.

Table 2
Error percentages for the iceberg image segmentation in Fig. 5(a)

Algorithm	Region	Error		
		Type I	Type II	Mean
FCM	Water	7.2228	20.7483	13.9856
	Iceberg	62.5797	0.8486	31.7141
	Sky	1.0421	24.3015	12.6718
PCM	Water	8.9581	19.1153	14.0367
	Iceberg	28.3612	59.5832	43.9722
	Sky	100.0000	0.0000	50.0000
GFRIS $r = 1$	Water	7.4898	21.3213	14.4055
	Iceberg	51.5495	0.9331	*26.2413*
	Sky	1.1869	15.8559	*8.5214*
GFRIS $r = 2$	Water	7.0449	22.2586	14.6517
	Iceberg	51.8344	0.9299	*26.3822*
	Sky	1.3027	14.9659	*8.1343*
GFRIS $r = 4$	Water	9.1435	21.4849	15.3142
	Iceberg	51.7933	0.9006	*26.3470*
	Sky	1.1406	16.3272	*8.7339*

In the above experiments, the number of segmented regions was constrained to two and three, respectively. In order to examine the discriminating potential of the GFRIS algorithm for a larger number of regions, a comparison was made with FCM and PCM algorithms on the image in Fig. 6(a) that possessed five regions. These were: egg (R_1), glass of milk (R_2), curtain (R_3), cheese (R_4) and table (R_5). Fig. 6 shows the segmentation performance of all three algorithms.

From Fig. 6(b)–(k), it is clear that both FCM and PCM arbitrarily divided the image into five regions without considering any semantic meaning of the data. The results produced by GFRIS for $r = 1$ and $r = 2$, in Figs. 6(l)–(u) showed more typical information of the regions. There are some regions such as egg and milk, curtain and cheese, which overlap with each other because their gray level pixel intensities are very similar. The most

Fig. 6. Original food image (a), and its segmented results for five regions produced by FCM (b)–(f), PCM (g)–(k), and GFRIS (l)–(z).

promising results in Fig. 6(v)–(z) were obtained for GFRIS using $r = 4$, with the exception of region R_4 (cheese) in Fig. 6(y), which partially merged with region R_2 (milk) as shown in Fig. 6(w). Again the GFRIS algorithm considered the underlying meaning of data far better than both the FCM and PCM techniques when compared with the manually segmented results.

Table 3
Error percentages for the food image segmentation in Fig. 6(a)

Algorithm	Region	Error		
		Type I	Type II	Mean
FCM	Egg	53.8987	27.7937	40.8462
	Milk	78.1723	17.5717	47.8720
	Curtain	57.7310	19.3766	38.5538
	Cheese	73.6814	18.3165	45.9990
	Table	64.1724	1.6680	32.9202
PCM	Egg	24.5806	59.3575	41.9690
	Milk	97.2167	3.8489	50.5328
	Curtain	98.2103	1.0998	49.6551
	Cheese	61.2456	30.5258	45.8857
	Table	100.0000	2.3314	51.1657
GFRIS $r = 1$	Egg	27.5875	19.8809	*23.7342*
	Milk	82.0478	18.3831	50.2155
	Curtain	34.9451	15.1914	*25.0683*
	Cheese	72.7393	18.4654	*45.6024*
	Table	69.8608	2.7701	36.3155
GFRIS $r = 2$	Egg	21.2948	25.2192	*23.2570*
	Milk	91.3547	9.4606	50.4077
	Curtain	16.2273	19.9142	*18.0708*
	Cheese	81.0402	12.0240	46.5321
	Table	51.6803	2.1541	*26.9172*
GFRIS $r = 4$	Egg	5.8837	0.2062	*3.0450*
	Milk	14.8141	33.2056	*24.0099*
	Curtain	49.5865	6.2929	*27.9397*
	Cheese	81.4295	11.2236	46.3266
	Table	46.0001	3.0249	*24.5125*

The numerical evaluations of the image segmentation given in Table 3, revealed that the mean error rates for the egg, curtain and cheese, egg, curtain and table, and egg, milk, curtain and table regions were appreciably lower using GFRIS with $r = 1$, $r = 2$, and $r = 4$, respectively than for either FCM or PCM. Overall the results confirmed that a significant improvement was achieved for all regions using GFRIS with neighbourhood radius $r = 4$, except for the cheese (R_4) region, for the reason alluded to above.

7. Conclusions

This paper has presented a new generic fuzzy rule based image segmentation (GFRIS) algorithm, which crucially has incorporated spatial relationships between pixels. It has been experimentally shown that in comparison with both FCM and PCM, GFRIS provided significantly superior results for a variety of different image types, including image examples having multiple regions. Its performance in considering the underlying meaning of data was also better when the results were compared with the manually segmented reference regions.

A single fuzzy rule was defined in order to classify the pixels, and three weighting factors W_1, W_2, and W_3 applied to stress the importance attached to feature based and spatial information in the image. Another important advantage of the GFRIS algorithm was that the structure of the membership functions and associated parameters were automatically derived and a new data-mining algorithm presented to determine both the key weighting factor and threshold value. The vital role of the threshold to the performance of GFRIS in controlling the maximum permitted pixel intensity variation between neighbouring and candidate pixels was highlighted.

From a computational perspective, since the three membership functions are independent of each other, the GFRIS algorithm possesses a high degree of inherent concurrency, which could be exploited by a parallel implementation, with a dedicated processor being used for each function.

Finally, as GFRIS is fuzzy rule based, the algorithm has the capability of incorporating any type of image attribute in any special application, by simply defining new membership functions, so making this solution both image and application independent.

Acknowledgements

The authors wish to thank Dr. Manzur Murshed of the Gippsland School of Computing and IT, Monash University for his suggestions and also the two formal reviewers for their thorough reading, numerous valuable comments and corrections.

References

Bezdek, J.C., 1981. Pattern Recognition With Fuzzy Objective Function Algorithms. Plenum, New York.

Chang, C.W., Ying, H., Hillman, G.R., Kent, T.A., Yen, J., 1998. A rule-based fuzzy segmentation system with automatic generation of membership functions for pathological brain MR images. Computers and Biomedical Research. Available from http://gopher.cs.tamu.edu/yen/ publications/index.html.

Chi, Z., Yan, H., 1993. Segmentation of geographic map images using fuzzy rules. In: Conf. Proc. DICTA-93, Digital Image Computing, Techniques and Applications, Vol. 1. Australian Pattern Recognition Society, Broadway, NSW, Australia, pp. 95–101.

Chi, Z., Yan, H., Pham, T., 1996. Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition. World Scientific, Singapore.

Geman, S., Geman, D., 1984. Stochas. Prentice Hall River, NJ.

Gose, E., Johnsonbaugh, R., Jost, S., 1996. Pattern Recognition and Image Analysis. Prentice Hall River, NJ.

Hall, L.O., Namasivayam, A., 1998. Using adaptive fuzzy rules for image segmentation. FUZZ-IEEE'98. Available from http://modern.csee.usf.edu/~hall/adrules/segment.html.

Karmakar, G.C., Dooley, L., 2001. Generic fuzzy rule based technique for image segmentation. In: IEEE Internat. Conf. on Acoustics, Speech, and Signal Process., (ICASSP'2001), Salt Lake City, Utah, ISBN 0-7803-7043-0.

Karmakar, G.C., Dooley, L., Rahman, S.M., 2000. A survey of fuzzy rule based image segmentation techniques. In: 1st IEEE Pacific-Rim Conf. on Multimedia, Sydney, Australia, pp. 350–353.

Kellogg, C.B., Zhao, F., Yip, K., 1996. Spatial aggregation: language and applications. In: Proc. AAAI-96. Available from http://www.cis.ohio-state.edu/insight/pubs.html.

Krishnapuram, R., Keller, J., 1993. A possibilistic approach to clustering. IEEE Trans. Fuzzy Syst. 1, 98–110.

Park, W., Hoffman, E.A., Sonka, M., 1998. Segmentation of intrathoracic airway trees: a fuzzy logic approach. IEEE Trans. Med. Imaging 17 (4), 489–497.

Sasaki, T., Hata, Y., Ando, Y., Ishikawa, M., Ishikawa, H., 1999. Fuzzy rule based approach to segment the menisci region from MR images. In: Proc. SPIE Medical Imaging, San Diego, CA, USA, Vol. 3661, p. 258.

Tizhoosh, H.R., 1998. Fuzzy image processing. Available from http://pmt05.et.uni-magdeburg.de/~hamid/-segment.html.

Tuceryan, M., 2000. Computational geometry. Available from http://klingon.cs.iupui.edu/~tuceryan.

Wertheimer, M., 1923. Laws of organization in perceptual forms. Pshychol. Forsch. 6.

Yip, K., Zhao, F., 1996. Spatial aggregation: theory and applications. J. Artif. Intell. Res. 5, 1–26.

Zaman, S.M.H., Rahim, K., Howlader, M., 1982. Simple Lesson From Biometry. Bangladesh Rice Research Institute, Joydebpur, Dhaka, Bangladesh.

Zhang, Y.J., 1996. A survey on evaluation algorithms for image segmentation. Pattern Recognition 29 (8), 1335–1346.